# *Causality versus Serial Correlation: an Asymmetric Portmanteau Test**

Amedeo Andriollo †

February 25, 2026

**Abstract**

This paper studies specification testing in dynamic linear models in the presence of omitted variables. The null hypothesis of interest is weak exogeneity: structural shocks have zero conditional expectation given their own past and the past of omitted variables. Existing tests based on quadratic forms of serial cross-correlations suffer from size distortions because their variance incorporates symmetric dependence in both directions, including causality from past shocks to present omitted variables (inverse causality). This paper proposes a correction that offsets the contribution of inverse causality, yielding an asymmetric Portmanteau statistic that is asymptotically normal under the null, without requiring parametric modeling of the joint dynamics. An empirical application revisits Diercks et al. (2024) and rejects weak exogeneity of Baker et al. (2016)'s EPU shocks. Addressing this failure by augmenting the information set with additional controls leads to a positive inflation response, pointing to a supply-side shock interpretation.

**Keywords:** omitted variables; cross-correlation; weak exogeneity.

*I am sincerely indebted to Eric Renault for his supervision throughout this work. For their valuable suggestions, I am especially thankful to Bertille Antoine, Ao Wang, Giovanni Ricco, Luis Candelaria, Kenichi Nagasawa and Leonardo Melosi. I am grateful for comments from Anna Mikusheva, Carlos Velasco, Mikkel Plagborg-Møller, Majid Al Sadoon, Michele Piffer, Giulio Rossetti, Emanuele Savini, and seminar and conference participants at several venues, including the University of Luxembourg, CREST, the 2025 World Congress of the Econometric Society (ESWC 2025), the Eleventh Italian Congress of Econometrics and Empirical Economics (ICEEE 2025), the Virtual Workshop for Junior Researchers in Time Series, and the Virtual Time Series Seminar (VTTS). I would also like to thank SIdE, the Bank of Italy, and the committee of the Carlo Giannini Prize. Any and all errors are my own.

†Université Paris Dauphine-PSL, DRM, 75775 Paris, France. *E-mail*: amedeo.andriollo@dauphine.psl.eu

# 1 Introduction

Estimating dynamic causal effects through structural models, such as Structural Vector Autoregressions (SVARs), is a central tool in applied macroeconometrics. Following the framework of Sims (1980), identification assumptions allow the innovations of a multivariate time series model to be interpreted as linear functions of the underlying structural shocks. Economists routinely exploit this link by estimating the residuals of a structural model, thus investigating the shocks' propagation through impulse response analysis (Kilian and Lütkepohl, 2017). Related arguments extend to univariate approaches, such as local projections (Jordà, 2023; Plagborg-Møller and Wolf, 2021), when combined with external instrumental variables.

The validity of such analyses hinges critically on the correct specification of the structural dynamics, both for variables included in and external to the model. Structural shocks must be "internally" exogenous to the other current and lagged endogenous variables in the model (Ramey, 2016), and this property cannot be undermined by variables omitted from the model. If omitted variables do interfere, the shocks fail to be "externally" exogenous and no longer represent truly *unanticipated* movements in the macroeconomic system. As a consequence, the history of the observed internal variables is insufficient to recover the shocks, violating invertibility or fundamentalness (Lippi and Reichlin, 1994; Nakamura and Steinsson, 2018).

Existing approaches to assessing this issue have cast it either as a problem of Granger causality testing (Giannone and Reichlin, 2006; Forni and Gambetti, 2014; Plagborg-Møller and Wolf, 2022; Miranda-Agrippino and Ricco, 2023) or as tests of conditional mean independence (Chen et al., 2017). From an econometric standpoint, both approaches can be viewed as addressing dynamic specification testing in the presence of omitted (external) variables. This paper brings these perspectives together by formulating the problem in terms of weak exogeneity, understood here as the property of the structural shocks having zero conditional expectation given the past of both internal and external variables (Mikusheva and Sølvsten, 2025).

The literature on specification testing in the presence of omitted variables can be organized according to whether the practitioner explicitly models the joint dynamics of internal and external variables. When the joint system is estimated, classical specification tests are available (e.g., Wald-type procedures, see Hong and Lee (2005) and related approaches Escanciano and Velasco (2006)). Inference in this case depends on how accurately the interaction between internal and external variables is specified and estimated. Parametric joint modeling is therefore sensitive to misspecification, while semi-parametric or nonparametric alternatives face well-known finite-sample limitations and the curse of dimensionality. When the joint dynamics are left unspecified, practitioners typically rely on nonparametric tests (e.g., Ljung–Box or Box–Pierce Portmanteau tests, see Hong (1996b), Lobato et al. (2002)). Despite not requiring augmentation, these tests are designed to detect symmetric forms of dependence, whereas weak exogeneity imposes a directional restriction. In the context considered here, this implies that both classes of approaches may produce rejections when there is dependence from past structural shocks to current omitted variables. This paper refers to this dependence as *inverse causality*, since it reflects the causal direction opposite to weak exogeneity, broadly speaking from past structural shocks to present omitted variables rather than from past omitted variables to present shocks. This channel becomes particularly salient in applied macroeconometrics because structural shocks, which capture primitive fluctuations of the macroeconomic system, are naturally expected to influence external macro variables over time. As a result, inverse causality can trigger rejections that researchers may mistakenly interpret as evidence against shocks' exogeneity.

This paper addresses these limitations by proposing an asymmetric Portmanteau test that isolates violations of weak exogeneity from inverse causality. The benchmark Portmanteau statistic tests the null of zero cross-correlation between current estimated shocks and lagged omitted variables by aggregating squared sample cross-correlations.[1] A key difficulty is that the quadratic norm introduces a *symmetry*: the variance of the resulting statistic depends not only on dependence from lagged omitted variables to current shocks, which is relevant under the null

[1]For univariate processes, Hong (1996b) defines the Portmanteau statistic as the weighted sum of squared cross-correlation between univariate time series at positive and negative lags, with weights determined by a kernel function. Following Hong (2001) and Bouhaddioui and Roy (2006), this paper regards as representative of the class of tests based on the serial cross-correlation function its one-sided multivariate formulation. Specifically, the benchmark is defined as the weighted sum of the $\ell_2$ norm of the cross-correlation between the two multivariate processes at positive lags.

of weak exogeneity, but also on dependence from lagged shocks to current omitted variables, that is inverse causality. The proposed correction term subtracts the contribution of this latter channel, leading to a modified statistic that targets only the dependence implied by violations of weak exogeneity. By construction, the procedure avoids parametric modeling of the joint dynamics and remains robust to misspecification of how past shocks affect the present of omitted variables, which is particularly useful when prior knowledge of the interaction between omitted variables and the dynamic system is limited.

The asymptotic distribution of the proposed test statistic is studied under the null hypothesis that estimated shocks have zero mean conditionally on their own past and the past of omitted variables. Establishing asymptotic normality requires additional assumptions on the shocks' second and fourth conditional moments. The restriction on second moments serves to isolate the "effects" rather than the "causes" or, in other words, to distinguish between the contribution of squares to the mean versus the variance of the statistic when testing for the conditional mean. In technical terms, this condition ensures correct centering of the test statistic under the null while avoiding stricter assumptions on the joint process and on the inverse causality channel. The fourth moment condition is mainly required for deriving normality.[2] The main asymptotic result is first derived for observed processes, then extended to settings where both shocks and omitted variables are estimated. This generalization permits flexible specifications of conditional mean and variance dynamics for variables external and internal to the structural model. Under the fixed alternatives of nonzero cross-correlation, this paper proves that the asymmetric Portmanteau achieves equivalent asymptotic power to benchmark. Since the statistic has hereditarily limited power against alternatives involving nonlinear non-pairwise dependence, a discussion on possible generalizations of the correction term is provided. A Monte Carlo study then examines the finite sample properties of both tests, corroborating the concerns on the size of the benchmark and demonstrating the desirable properties of the proposed statistic.

As an empirical application, the paper examines the exogeneity of widely used measures of macroeconomic structural shocks, with particular focus on uncer-

[2]When relaxing the assumption of independence, Hong (2001)'s footnote 8 briefly discussed the condition of conditional homokurtosis for establishing the asymptotic normality of his testing procedure.

tainty shocks (Baker et al., 2016; Jurado et al., 2015; Berger et al., 2020) and monetary policy shocks (Jarociński and Karadi, 2020; Miranda-Agrippino and Ricco, 2021; Bu et al., 2021; Bauer and Swanson, 2023; Aruoba and Drechsel, 2024). The analysis controls for economic conditions using four sets of principal components extracted from two large datasets: macroeconomic factors (McCracken and Ng, 2016; Rapach and Zhou, 2021) and financial factors (Giglio and Xiu, 2021; Lettau and Pelger, 2020b). The findings challenge the exogeneity of a substantial proportion of these series. Notably, Baker et al. (2016)'s Economic Policy Uncertainty (EPU) shocks fail the exogeneity test with respect to lagged macro factors. Building on this, the paper revisits Diercks et al. (2024) and strengthens their conclusions: when these additional controls are included, the response of inflation to EPU shocks shifts from modestly negative to drastically positive. Combined with contractionary responses in other variables, this evidence suggests that the EPU structural shock operates as a supply-side negative shock, similar to the 'expectational' shocks discussed in Ascari et al. (2023).

LITERATURE. This paper contributes to three strands of literature. First, it relates to specification testing in dynamic linear models. Early contributions include Hosking (1980) and Li and McLeod (1981), who proposed Portmanteau statistics based on sums of squared auto-correlations of residuals over a fixed number of lags. Rather than modeling the joint process, Haugh (1976) developed a two-step procedure to test for independence between time series: first fitting univariate models, then examining cross-correlations at different lags. Hong (1996a,b) generalized these tests to all lags using kernel-weighted schemes, with Bouhaddioui and Roy (2006) extending the framework to multivariate processes. This paper contributes to the Haugh (1976) approach by introducing a correction term that isolates a specific direction of causality, making the test more suitable for assessing weak exogeneity. Second, the paper relates to tests of the martingale difference hypothesis, which typically require modeling the conditional mean of the joint process (Durlauf, 1991; Hong, 1999; Hong and Lee, 2005; Escanciano and Velasco, 2006). This class of tests can be viewed as extending Ljung and Box (1978)'s approach rather than Haugh (1976)'s. The proposed method improves upon these by avoiding the need to model joint conditional means and variances. Third, this paper contributes to the literature on testing invertibility or fundamentalness of structural shocks. It provides a new testing strategy and offers some empirical insights about the exogeneity of widely used macroeconomic shock measures, with particular focus on Baker et al. (2016)'s Economic Policy Uncertainty (EPU) shock series.

OUTLINE. Section 2 introduces the benchmark and the modified test statistics. In particular, the definition of the asymmetric Portmanteau statistic is in Section 2.2. Section 3 presents the asymptotic theory for the corrected test statistic. Section 4 presents the empirical application. Section 5 concludes.

NOTATION. Throughout, the following standard notation is used. Given two vectors $a$ and $b$, the inner and Kronecker products are: $\langle a, b\rangle = a'b$ and $a \otimes b$. The $\ell_2$ norm is: $||a|| = \sqrt{\langle a, a\rangle}$. For a real positive semidefinite matrix $A$, its square-root is: $B = (A)^{1/2}$ such that $A = BB = BB'$, and its Frobenius norm is: $||A||_F = \sqrt{\text{tr}(A'A)}$. tr$(\cdot)$, vec$(\cdot)$ and diag$(\cdot)$ stand for the trace, vectorization and main diagonal operators. I denote $\xrightarrow{d}$ and $\xrightarrow{p}$ as convergence in distribution and in probability. $\perp$ stands for orthogonality, and $\perp\!\!\!\perp$ for mutual independence.

# 2 The Testing Strategies Based on $\ell_2$-norm

Section 2.1 establishes the primary framework and discusses the class of Portmanteau statistics, clarifying the connection between their variance and inverse causality. Section 2.2 introduces the correction term and defines the asymmetric Portmanteau statistic. Proofs of the first proposition and lemmas appear in Appendix A.

## 2.1 Preliminaries and a Discussion on Portmanteau Statistics

Let $\{X_t, Z_t; t = 1, .., T\}$ denote two zero-mean multivariate square-integrable jointly stationary processes of respective finite dimensions $d_1, d_2 \in \mathbb{N}_+$. Let $\mathcal{I}(t-1)$ be the information set available at period $t-1$ comprising the joint past, $\{X_s, Z_s; s < t\}$. Unless stated otherwise, these processes are standardized.[3]

Throughout, $X$ represents the structural shock series and $Z$ represents the omitted variables. The null hypothesis of interest is weak exogeneity, defined as structural shocks having zero conditional expectation given their own past and the past

[3] That is $\text{Var}[X_t] = \mathbb{E}[X_t X_t'] = I_{d_1}$ and $\text{Var}[Z_t] = \mathbb{E}[Z_t Z_t'] = I_{d_2}$. This assumption is relaxed in Section 3.2.

of omitted variables:

$$\mathcal{H}_0 : \mathbb{E}[X_t|\mathcal{I}(t-1)] = 0 \tag{1}$$

Before introducing the modified test statistic in Eq.(5) (Section 2.2), we discuss testing strategies based on squared serial cross-correlations following Hong (1996b). The benchmark is the one-sided statistic based on weighted quadratic forms:

$$\mathcal{T}_\omega = \sum_{j=1}^{T-1} \omega(j) Q(j) \tag{2}$$

$$Q(j) = ||\widehat{\Gamma}_{XZ}(j)||_F^2 = \text{tr}\left[\widehat{\Gamma}_{XZ}(j)'\widehat{\Gamma}_{XZ}(j)\right] = \left|\left|\text{vec}\left[\widehat{\Gamma}_{XZ}(j)\right]\right|\right|^2 \tag{3}$$

for some nonrandom non-negative weights $\{\omega(j)\}$, where $\widehat{\Gamma}_{XZ}(j)$ is the sample cross-correlation between the (standardized) processes:

$$\widehat{\Gamma}_{XZ}(j) = \frac{1}{T}\sum_{t=j+1}^{T} X_t Z'_{t-j}, \quad \Gamma_{XZ}(j) = \mathbb{E}[X_t Z'_{t-j}], \qquad j = 1, ..., T-1$$

The statistic is one-sided ($j > 0$) because $\mathcal{H}_0$ concerns a particular direction of causality: the influence of past of $Z$, $\{Z_s; s < t\}$, on the present $X$, $\{X_t\}$. The quadratic forms, $\{Q(j)\}$, correspond to the squared $\ell_2$-norms of vectorized sample cross-correlation matrices, equivalently their squared Frobenius norms.[4] Lemma A.1 in Appendix A.1 provides an equivalent spectral domain formulation, drawing the parallel between weights $\{\omega(j)\}$ and kernel estimation of the cross-spectrum.

By construction, quadratic forms treat causality directions symmetrically. This becomes apparent when decomposing the test statistic: the inner product of cross-correlation matrices generates cross-product terms where $X$ and $Z$ enter symmetrically across different time lags. Formally, by means of some algebra (Lemma

[4]This generalization from univariate to multivariate analysis dates to Li and McLeod (1981). See Bouhaddioui and Roy (2006) for further discussion. Lemma A.3 in Appendix A connects the one-sided statistic to Bouhaddioui and Roy (2006). For the equivalence between Euclidean norm and trace, see chapter 4 of Lütkepohl (1997).

A.2):

$$\mathcal{T}_{\omega} = \mathcal{T}_{1\omega} + \mathcal{T}_{2\omega} \tag{4}$$
$$\mathcal{T}_{1\omega} = \frac{1}{T^2}\sum_{j=1}^{T-1}\omega(j)\sum_{t=j+1}^{T}||X_t||^2||Z_{t-j}||^2$$
$$\mathcal{T}_{2\omega} = \frac{1}{T^2}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t}^{T} < X_t, X_s >< Z_{t-j}, Z_{s-j} >$$

The Portmanteau statistic, $\mathcal{T}_{\omega}$, consists of two components: the "sum of squares", $\mathcal{T}_{1\omega}$, and the "sum of cross-products", $\mathcal{T}_{2\omega}$. While the sum of squares preserves temporal ordering (from past omitted variables to present shocks), the sum of cross-products incorporates the interaction between two time indexes, $s$ and $t$, thus blending both directions of causality. This symmetry due to the norm suggests that when testing the null, Portmanteau statistics may not effectively distinguish between violations of weak exogeneity and dependence from past shocks to present omitted variables (inverse causality).

The distinction between components matters for understanding asymptotic properties of the testing strategies based on Hong (1996b) and subsequent work. Intuitively, the sum of cross-products, $\mathcal{T}_{2\omega}$, dominates under the null and thus controls test size, whereas the sum of squares, $\mathcal{T}_{1\omega}$, dominates under the alternatives, and so regulates power.[5]

The following proposition clarifies how cross-product terms incorporate both directions of dependence (from past omitted variables to present shocks and vice versa). We impose two simplifying assumptions: i) marginal independence of the joint process, i.e., $X_t \perp\!\!\!\perp X_k, Z_t \perp\!\!\!\perp Z_k, t \neq k$; ii) conditional homoskedasticity of the shocks, $\mathbb{E}[||X_t||^2|\mathcal{I}(t-1)] = \mathbb{E}[||X_t||^2]$.

**Proposition 1.** *Let $\{X_t, Z_t\}$ be marginally i.i.d. processes with finite fourth moments, such that the process $\{X_t\}$ be conditionally homoskedastic with respect to the joint information set, $\mathcal{I}(t-1)$. Under the null hypothesis in Eq.(1), the variance of the benchmark statistic, $\mathcal{T}_{\omega}$, depends on the inverse causality through the variance of the sum of cross-*

[5] In Proposition 2, the asymptotic properties of the test under the null are described in details. For the power of the test, refer to Theorem 3 and the discussion that follows.

*products, $\mathcal{T}_{2\omega}$. In particular:*

$$\mathbb{E}[(< X_t, X_s >< Z_{t-j}, Z_{s-j} >)^2] = \begin{cases} (d_1 d_2)^2, & s > t-j \\ d_1 \mathbb{E}[||X_s||^2 < Z_{t-j}, Z_{s-j} >^2], & s \leq t-j \end{cases}$$

*If $d_1 = d_2 = 1$, when $s \leq t-j$: $\mathbb{E}[||X_s||^2 < Z_{t-j}, Z_{s-j} >^2] = \mathbb{E}\Big[\mathbb{E}\big[Z_{t-j}^2 | \{X_s, Z_{s-1}\}_{s \leq t-j}\big] X_s^2 Z_{s-j}^2\Big]$. Under mutual independence of the two processes, $X_t \perp\!\!\!\perp Z_s, \forall s, t$ , we have:*

$$\mathbb{E}[(< X_t, X_s >< Z_{t-j}, Z_{s-j} >)^2] = (d_1 d_2)^2$$

Proposition 1 demonstrates that under the null hypothesis, the variance of the statistic $\mathcal{T}_\omega$ incorporates, through the cross-products in $\mathcal{T}_{2\omega}$, the dependence from past $X$ to present $Z$ captured by cross-moments of the joint process ( the inverse causality channel). This vanishes when either processes are independent, $X_t \perp\!\!\!\perp Z_s, \forall t, s$ (strict exogeneity between shocks and omitted variables at all lags/leads), or when a specific time ordering holds, $s > t-j$. Three remarks follow. First, in the univariate case $(d_1 = d_2 = 1)$, it is evident that inverse causality operates through the conditional variance of $Z$. Second, the result can be obtained also under past independence of shocks, $X_t \perp\!\!\!\perp Z_{s_1}, X_{s_2}$, with $s_1, s_2 < t$.[6] Third, the presence of such dependencies in the variances arise even under the restrictive assumption on the univariate processes, namely marginally i.i.d. time series.

Lemma A.4 in Appendix A.3 explicitly derives how dependencies from shocks to omitted variables affect the variance of the sum of cross-products, $\mathcal{T}_{2\omega}$, under three examples of general DGPs where inverse causality is present and the null of interest holds true.

[6] Similar assumptions appear in Hong et al., 2009, and Candelon and Tokpavi, 2016. This condition, together with the assumption of marginally i.i.d., is weaker than statistical independence, $X_t \perp\!\!\!\perp Z_s, \forall s, t$.

## 2.2 An Asymmetric Portmanteau Statistic

Motivated by the previous proposition, we introduce a modified statistic:

$$\mathcal{T}^c_\omega = \mathcal{T}_\omega - \frac{1}{T^2}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t,s\leq t-j}^{T} < X_t, X_s >< Z_{t-j}, Z_{s-j} >= \mathcal{T}_\omega - \mathcal{C}_\omega$$
$$= \mathcal{T}_{1\omega} + \frac{1}{T^2}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t,s>t-j}^{T} < X_t, X_s >< Z_{t-j}, Z_{s-j} >= \mathcal{T}_{1\omega} + \mathcal{T}^c_{2\omega} \quad (5)$$

The correction term $\mathcal{C}_\omega$ eliminates the influence of inverse causality captured by a subset of cross-products. By breaking the symmetry inherent in quadratic forms, the proposed statistic constitutes an *asymmetric* Portmanteau test. Proposition 2 and subsequent results formalize the benefits of this correction. To further motivate the correction term, we offer two complementary perspectives: one based on the jackknife debiasing approach and another based on predictive regressions.

The first perspective interprets a subset of the cross-products as introducing bias into the variance of sample covariance estimators, under weak exogeneity and conditional homoskedasticity (Proposition 1). A solution in the style of jackknife would employ then a block-deletion scheme targeting observations associated with the temporal ordering $s \leq t - j$ (for a given $j$), precisely those observations linked to the inverse causality channel.

The second perspective provides deeper insight by viewing moments of cross-products as coefficients in predictive regressions. Consider without loss of generality the bivariate process $\{X_t, Z_t\}$ for time indexes $t > s$. After applying the correction, the first moment of remaining cross-products, $\mathcal{T}^c_{2\omega}$, is proportional to coefficients in regressions of the form: (for a fixed $s > 0$)

$$X_t X_s = \sum_{j=t-s}^{s-1} \phi^{(1)}_j Z_{t-j} Z_{s-j} + e_t$$

where $e_t$ is an error term, $Z_{t-j}, Z_{s-j} \in \mathcal{I}(t-j)$ and $X_s \notin \mathcal{I}(t-j)$. Conversely, the first moment of cross-products constituting the correction term, $\mathcal{C}_\omega$, is proportional to coefficients in autoregressions: (for a fixed $j > 0$)

$$X_t Z_{t-j} = \sum_{l=1}^{s} \varphi_l^{(1)} X_l Z_{l-j} + \epsilon_t$$

where $\epsilon_t$ is an error term, $X_l, Z_{l-j} \in \mathcal{I}(s)$ and $Z_{t-j} \notin \mathcal{I}(s)$. These coefficients assess distinct implications of the null. By law of iterated expectations, weak exogeneity implies: i) $\mathbb{E}[X_t X_s] = 0,\ \forall s < t$; ii) $\mathbb{E}[X_t Z_{t-j}] = 0,\ \forall j > 0$, captured by the first and second regression sets, respectively. The second moment of $\mathcal{T}_{2\omega}^{c}$ and $\mathcal{C}_\omega$ are proportional to coefficients in regressions:

$$X_t^2 X_s^2 = \sum_{j=t-s}^{s-1} \phi_j^{(2)} Z_{t-j}^2 Z_{s-j}^2 + e_t^{(2)}, \quad s > t-j, \quad \text{(for a fixed } s > 0\text{)}$$

$$X_t^2 Z_{t-j}^2 = \sum_{l=1}^{s} \varphi_l^{(2)} X_l^2 Z_{l-j}^2 + \epsilon_t^{(2)}, \quad l \leq t-j, \quad \text{(for a fixed } j > 0\text{)}$$

Conditional homoskedasticity of $X$ has the following implication for second moments:

$$\phi_j^{(2)} \propto \mathbb{E}[(X_t^2 X_s^2) M_Z (Z_{t-j}^2 Z_{s-j}^2)]/\text{Var}[Z_{t-j}^2 Z_{s-j}^2] = d_1^2 \mathbb{E}[M_Z Z_{t-j}^2 Z_{s-j}^2]/\text{Var}[Z_{t-j}^2 Z_{s-j}^2]$$

for an appropriate projection operator, $M_Z \in \mathcal{I}(s)$, by a Frisch–Waugh–Lovell argument. Coefficients $\{\phi_j^{(2)}\}$ from the first regression set depend solely on higher moments of the marginal process $\{Z_t\}$. In contrast, coefficients from the second set depend on higher-order moments of the joint process $\{X_t, Z_t\}$. Consistent with Proposition 1, this distinction motivates the correction term, ensuring the testing procedure is robust to higher-order dependencies operating in the causal direction opposite to that being tested.

# 3 Asymptotic Theory

Section 3.1 establishes the asymptotic properties of the proposed statistic under the null for observed processes. Section 3.2 extends these results to estimated processes. Section 3.3 derives asymptotic properties under a general class of alternatives and discusses settings where the test has limited power. Section 3.4 discuss a set of Monte Carlo experiments. Proofs of the Proposition and Theorems appear in Appendix B. The results of the simulation study are in Appendix C.

### 3.1 Asymptotics of the Statistic under the Null

This paper considers the following regularity condition on the weighting scheme in Eq.(2):

**Assumption 1.** *Let the sequence of weights $\{\omega(j)\}$ be a function of some sequence of integers $M = M(T)$ for which there exists an appropriate square-integrable kernel $k(\cdot) : \mathbb{R} \to [-1, 1]$, continuous at 0 and at all points except for a finite number of points, such that: $\omega(j) = k^2(j/M)$, $k(0) = 1$.*

This assumption is standard in nonparametric spectral density estimation via kernel functions (Hong, 2001). The sequence of integers $M$, growing with sample size $T$, characterizes the kernel estimation window: larger $M$ incorporates more lags when summing cross-correlations in the Portmanteau statistic.[7] Define:

$$
\begin{aligned}
\mu_{\omega,T} &= \mu[\{\omega(j)\}, T] = d_1 d_2 \sum_{j=1}^{T-1} \left(1 - \frac{j}{T}\right) \omega(j) \\
D^{(Hete)}_{\omega,T} &= D^{(Hete)}[\{\omega(j)\}, T] \\
&= \frac{2d_1^2}{T^2} \sum_{j=1}^{T-2} \sum_{\ell=1}^{T-2} \omega(j)\omega(\ell) \sum_{s=\max\{j,\ell\}+1}^{T-1} \quad \sum_{t=s+1}^{\min\{T,\ s+\min(j,\ell)-1\}} \gamma_{t,s}(j,\ell) \\
D_{\omega,T} &= D[\{\omega(j)\}, T] = 2d_1^2 d_2^2 \sum_{j=1}^{T-2} \left(1 - \frac{j}{T}\right)\left(1 - \frac{j+1}{T}\right) \omega^2(j)
\end{aligned} \tag{6}
$$

with $\gamma_{t,s}(j,l) = \mathbb{E}[\langle Z_{t-j}, Z_{s-j}\rangle \langle Z_{t-\ell}, Z_{s-\ell}\rangle]$, when these last moments exist. The first two quantities approximately match the mean and the variance of the correct test statistic scaled by the sample size, $T \cdot \mathcal{T}^c_\omega$, under the null hypothesis (Proposition 2). The last quantity, $D_{\omega,T}$, represents the asymptotic variance of the benchmark statistic scaled by the sample size, $(T \cdot \mathcal{T}_\omega)$, under mutual independence of $X$ and $Z$ (Hong, 2001, pg.191-2).

Under Assumption 1, these quantities are of order $M$: $\mu_{\omega,T} = O(M)$, $D_{\omega,T} = O(M)$, and $D^{(Hete)}_{\omega,T} = O(M)$, for appropriate conditions on: $\{\gamma_{t,s}(j,l)\}$.

[7] See the discussion at the end of Section 3.1 regarding the smoothing parameter.

The following proposition characterizes the first moments of the corrected statistic, $\mathcal{T}_\omega^c$.

**Proposition 2.** *Suppose Assumption 1 holds, with* $\frac{M}{T} \to 0$*, as* $T, M \to \infty$*.*

*i) Suppose* $Z$ *has finite fourth moments, and* $|Cov[||Z_1||^2, ||Z_{1+h}||^2]| = O(h^{-1-\epsilon})$ *for* $\epsilon > 0$*.*
*If:* $\mathbb{E}[X_t X_t' | \mathcal{I}(t-1)] = \mathbb{E}[X_t X_t']$*, then:* $\mathbb{E}[T \cdot \mathcal{T}_{1\omega}] = \mu_{\omega,T}$*.*
*In addition, if:* $\mathbb{E}[(X_t X_t') \otimes (X_t X_t') | \mathcal{I}(t-1)] = \mathbb{E}[(X_t X_t') \otimes (X_t X_t')]$*,*
*then:* $Var[T \cdot \mathcal{T}_{1\omega}] = O(M^2/T)$*, implying mean-squared convergence:*
$\lim_{T\to\infty}(T \cdot \mathcal{T}_{1\omega} - \mu_\omega)\left(D_{\omega,T}^{(Hete)}\right)^{-1/2} = 0$

*ii) Under* $\mathcal{H}_0$ *stated in Eq.(1), we have:* $\mathbb{E}[T \cdot \mathcal{T}_{2\omega}^c] = 0$*.*
*If additionally:* $\mathbb{E}[X_t X_t' | \mathcal{I}(t-1)] = \mathbb{E}[X_t X_t']$*, then:* $Var[T \cdot \mathcal{T}_{2\omega}^c] = D_{\omega,T}^{(Hete)}$*.*

Proposition 2 formalizes the benefits of the correction term. Under conditional homoskedasticity and conditional homokurtosis of $X$ with respect to the joint past, the first two moments of the statistic do not incorporate dependencies running from past $X$ to present $Z$. Specifically, these moments depend only on the weighting scheme, $\{\omega(j)\}$ or, at most, on particular higher moments of the marginal process $Z$, $\{\gamma_{t,s}\}$, rather than on moments of the joint process that would reflect inverse causality. In fact, by breaking the symmetry of the quadratic form, the correction term restores a martingale structure that permits separating variances via law of iterated expectations. Consequently, the variance of the corrected statistic reduces to estimating the long-run variance of second-order moments of the process $Z$ (i.e., the cross-products $\{\langle Z_{t-j}, Z_{s-j}\rangle\}$).

Propositions 1-2 highlight a trade-off in terms of restrictions on marginal vs. joint processes: maintaining directional inference while avoiding to specify the dynamics between shocks and omitted variables. This paper prioritizes an agnostic stance toward inverse causality, imposing minimal restrictions on how past shocks influence present omitted variables at the cost of stronger moment restrictions on the structural shocks themselves.

Due to the Frobenius norm, these conditional moment restrictions serve to isolate *effects* rather than *causes*: they ensure the statistic correctly detects violations of weak exogeneity without constraining the inverse causality channel. Conditional homoskedasticity ensures proper centering by isolating weak exogeneity violations to the mean of cross-products, $\mathcal{T}_{2\omega}^c$, rather than the sum of squares, $\mathcal{T}_{1\omega}$.

Conditional homokurtosis bounds the variance of the latter sum, ensuring cross-products dominate under the null.

These moment restrictions on structural shocks offer practical advantages for empirical application. Since structural shocks are estimated rather than observed, conditional moment restrictions can guide towards sharper identification (e.g., Hafner et al. (2022), or related heteroskedastic identification schemes). Under additional parametric assumptions on inverse causality, these conditions could be relaxed.[8] Notably, analogous conditions have been proposed for testing in Proxy-SVAR frameworks (Bruns and Keweloh, 2024).

The following theorem formalizes these considerations, establishing asymptotic normality of the corrected statistic under $\mathcal{H}_0$. Define: $\Lambda_{s,t} = \sum_{j=t-s}^{s-1} \omega(j) X_s \langle Z_{t-j}, Z_{s-j} \rangle$, and consider the following assumptions on the process $Z$:

**Assumption 2.** *The process $\{Z_t\}$ is strictly stationary, has finite $(8+\delta)$-order moments, with $\alpha(h)$ such that: $\sum_{h=1}^{\infty} \alpha(h)^{\delta/(8+\delta)} < \infty$.*

**Theorem 1.** *Suppose the process $\{X_t\}$ is such that:*

$$\mathbb{E}[X_t X_t' | \mathcal{I}(t-1)] = \mathbb{E}[X_t X_t'], \quad \mathbb{E}[(X_t X_t') \otimes (X_t X_t') | \mathcal{I}(t-1)] = \mathbb{E}[(X_t X_t') \otimes (X_t X_t')]$$

*Suppose the time series $\{Z_t\}$ satisfies Assumption 2. Further, suppose the joint process $\{X_t, Z_t\}$ is strictly stationary, and Assumption 1 holds with $\frac{M^2}{T} \to 0$, as both $T, M \to \infty$. Under the null $\mathcal{H}_0$ in Eq.(1), we have:*

$$\frac{T \cdot \mathcal{T}_\omega^c - \mu_{\omega,T}}{\sqrt{D_{\omega,T}^{(Hete)}}} \xrightarrow{d} \mathcal{N}(0,1)$$

*If additionally the joint process $\{X_t, Z_t\}$ satisfies: $|\mathbb{E}\langle \Lambda_{1,t}, \Lambda_{1+i,t} \rangle| = O(i^{-2})$, for $i \to +\infty$, the asymptotic normality result holds with $\frac{M}{T} \to 0$, as both $T, M \to \infty$.*

Theorem 1 offers two notable improvements over existing testing strategies.

First, Portmanteau statistics following Hong (1996a,b) are typically studied under statistical independence. In the presence of inverse causality, benchmark tests based on squared cross-correlations may therefore exhibit size distortions under

[8] These moment restrictions are weaker than mutual independence (Hong, 1996b) or past independence (Candelon and Tokpavi, 2016), and comparable to approximate q-dependence (Hong and Lee, 2005).

weak exogeneity, as their higher moments incorporate dependencies from past $X$ to present $Z$ (Proposition 1). The corrected statistic $\mathcal{T}_\omega^c$ addresses this issue directly: Theorem 1 establishes that its asymptotic normality depends essentially on the martingale properties of $X$ with respect to the joint past, without requiring independence or parametric modeling of the inverse causal channel.

Second, tests for the martingale difference property (Hong and Lee, 2005; Escanciano and Velasco, 2006) could, in principle, be used to test $\mathcal{H}_0$, but at a high cost. These procedures would require either (i) the joint process $\{X_t, Z_t\}$ to be a martingale difference sequence (stronger than weak exogeneity), or (ii) explicit modeling of the joint conditional mean. The latter approach would involve additional high-level assumptions (e.g., Assumptions A2-A3 in Hong and Lee, 2005) whose implications are less transparent than the primitive moment conditions imposed in Theorem 1. By contrast, the proposed framework avoids modeling the joint dynamics altogether, while maintaining interpretable and testable restrictions on the moments of the structural shocks.

The main cost relative to Hong (1996b) is a more stringent rate condition on the smoothing parameter $M$, which must diverge slower than $\sqrt{T}$. This slower rate reflects the need to control the variance of the sum of squares under minimal assumptions on the marginal process $Z$ (i.e., finite eighth moments). However, the second part of Theorem 1 shows that, under mild additional dependence conditions (Dedecker et al., 2007), the standard rate $M/T \to 0$ suffices.

The smoothing parameter $M$ governs the number of lags considered in the test statistic and thus the rate at which the weighted sum of covariance terms converges to a Gaussian limit. For small/finite $M$, the limiting distribution is a weighted sum of chi-squared variables (Box and Pierce, 1970; Francq and Raïssi, 2007). As $M$ increases with $T$, the sum converges to normality by a standard central limit argument. The choice of $M$ involves a familiar trade-off: sufficiently fast growth ensures asymptotic normality under the null, while sufficiently slow growth preserves power against alternatives (see Theorem 3). In addition, since the correction term permits separating the variances, the smoothing parameter $M$ also governs the effective window used to estimate the long-run variance of second-order moments of $Z$, thus relating its choice to the size-power tradeoff in HAR inference (Lazarus et al., 2021).

### 3.2 Estimated Processes

In practice, structural shocks and omitted variables are seldom observed directly. Instead, practitioners fit models to observed data and conduct inference using estimated residuals. This practice therefore naturally aligns with the two-step approach of Haugh (1976) and Hong (1996b), where noncausality is tested after fitting separate models. We now extend the asymptotic theory to this setting.

Let $\{W_{1,t}, W_{2,t}; t = 1, ..., T\}$ denote observed processes of dimensions $d_1$ and $d_2$, respectively. We assume both admit causal conditional mean representations:

$$\begin{aligned} W_{1,t} &= \mu_X(\theta_1^0, \{X_s; s < t\}) + X_t \\ W_{2,t} &= \mu_Z(\theta_2^0, \{Z_s; s < t\}) + Z_t \end{aligned} \tag{7}$$

where, $\mu_X(\theta_1^0, \cdot) \in \mathbb{R}^{d_1}$, $\mu_Z(\theta_2^0, \cdot) \in \mathbb{R}^{d_2}$ are known measurable functions parameterized by finite-dimensional time-invariant parameters $\theta_1^0$ and $\theta_2^0$. The innovation $\{X_t\}$ is defined by the condition: $\mathbb{E}[X_t|\{X_s; s < t\}] = 0$, meaning that the shock $X$ forms a martingale difference sequence with respect to its own history. The process $W_2$ is specified as a function of its own past rather than the joint past $\{W_{1,s}, W_{2,s}\}_{s<t}$ to avoid implicitly modeling inverse causality. The representation in Eq.(7) encompasses a general class of multivariate time series models for conditional means (and variance), requiring only that innovations are correctly captured via finite-dimensional time-invariant parameters.

Suppose the practitioner has $\sqrt{T}-$consistent estimators, $\{\widehat{\theta}_i\}_{i=1,2}$, of the true parameters (e.g., least squares or maximum likelihood). Given sample data, estimated structural shocks and omitted variables, $\widehat{X}_t(\widehat{\theta}_1)$ and $\widehat{Z}_t(\widehat{\theta}_2)$, depend on the limited observed past, $\widehat{\mathcal{I}}(t-1)$, and on the estimators.[9] Define the standardized innovations and their estimated counterparts:

$$U_t = (\Gamma_X)^{-1/2} X_t, \quad V_t = (\Gamma_Z)^{-1/2} Z_t, \quad \text{s.t.} \quad \mathbb{E}[||U_t||^2] = d_1,\ \mathbb{E}[||V_t||^2] = d_2 \tag{8}$$

$$\widehat{U}_t = \left(\widehat{\Gamma}_X\right)^{-1/2} \widehat{X}_t, \quad \widehat{V}_t = \left(\widehat{\Gamma}_Z\right)^{-1/2} \widehat{Z}_t, \qquad t = 1, ..., T \tag{9}$$

where $\Gamma_X = \mathbb{E}[X_t X_t']$ and $\Gamma_Z = \mathbb{E}[Z_t Z_t']$ have empirical counterparts $\widehat{\Gamma}_X$ and $\widehat{\Gamma}_Z$.

Parallel to Eq.(5), the test statistic based on estimated standardized processes

[9] These are the estimated pseudo-version of the innovations with arbitrary starting values, since we do not observe the infinite past of the time series, i.e., $\mathcal{I}(t-1)$.

is:

$$\widehat{\mathcal{T}}_{\omega}^{c} = \widehat{\mathcal{T}}_{\omega} - \frac{1}{T^2}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s<t-j}^{T} < \widehat{U}_t, \widehat{U}_s >< \widehat{V}_{t-j}, \widehat{V}_{s-j} >$$
$$= \widehat{\mathcal{T}}_{\omega} - \widehat{\mathcal{C}}_{\omega} = \widehat{\mathcal{T}}_{1\omega} + \widehat{\mathcal{T}}_{2\omega}^{c} \quad (10)$$

where $\widehat{\mathcal{T}}_{\omega} = \sum_{j=1}^{T-1}\omega(j)||\widehat{\Gamma}_{UV}(j)||_F^2$ is the empirical counterpart of the statistic in Eq.(4) (see Lemma A.3). Similar to Eq.(6), define the variance estimator:

$$\widehat{D}_{\omega,T}^{(Hete)} = \frac{2d_1^2}{T^2}\sum_{j=1}^{T-2}\sum_{\ell=1}^{T-2}\omega(j)\omega(\ell)\sum_{s=\max\{j,\ell\}+1}^{T-1}\sum_{t=s+1}^{\min\{T,\ s+\min(j,\ell)-1\}}\widehat{\gamma}_{t,s}(j,\ell)$$

where: $\widehat{\gamma}_{t,s}(j,\ell) = \langle \widehat{V}_{t-j}, \widehat{V}_{s-j}\rangle\langle \widehat{V}_{t-\ell}, \widehat{V}_{s-\ell}\rangle$.

**Theorem 2.** *Suppose the processes $\{W_{i,t}\}_{t=1,\ldots,T}^{i=1,2}$ admit the causal representation of Eq.(7). Suppose the assumptions of Theorem 1, Assumption 3 and the regularity conditions of Eq.(12)-(13) in Appendix B.3 hold true. Let $\{\widehat{\theta}_i\}_{i=1,2}$ be $\sqrt{T}-$consistent estimators of the true parameters $\{\theta_i^0\}_{i=1,2}$. Under the null $\mathcal{H}_0$ in Eq.(1), we have:*

$$\frac{T\cdot\widehat{\mathcal{T}}_{\omega}^{c} - \mu_{\omega,T}}{\sqrt{\widehat{D}_{\omega,T}^{(Hete)}}} \xrightarrow{d} \mathcal{N}(0,1)$$

Theorem 2 shows that estimation error does not affect the limiting distribution under two broad sets of conditions. First, correct model specification delivers parametric-rate estimation of the innovations, so the approximation error is asymptotically negligible. Second, appropriate smoothness: differentiability of $\mu_X(\theta_1^0,\cdot)$ and $\mu_Z(\theta_2^0,\cdot)$ with respect to parameters, ensuring uniform $\ell_2$-convergence and bounded derivatives (Appendix B.3, Eq.(12)–(13)). These assumptions mirror those commonly imposed in the conditional mean specification literature (Wang et al., 2022) and in Hong-type procedures (e.g. Hong and Lee, 2005; Leong and Urga, 2023). Consequently, the plug-in statistic constructed from standardized estimated residuals retains the same standard normal asymptotic distribution as in the infeasible case. Other conditions may be imposed to reach similar conclusions (Domínguez and Lobato, 2020).

Theorem 2 applies beyond testing weak exogeneity. To illustrate the point, we show how the framework extends to second-order (variance) noncausality (Granger, 1969; Comte and Lieberman, 2000). Suppose $\mathcal{H}_0$ in Eq.(1) holds and the specifications in Eq.(7) are correct. The null of variance noncausality from $W_2$ to $W_1$ is:

$$\mathcal{H}_0^{VNC} : \mathbb{E}[X_t X_t' | \mathcal{I}(t-1)] = \mathbb{E}[X_t X_t' | \{X_s; s < t\}]$$

If additionally: $\mathbb{E}[X_t X_t' | \{X_s; s < t\}] = \Gamma_X$ (conditional homoskedasticity), then defining the centered squared process: $X_t^\dagger = \text{vech}[X_t X_t'] - \text{vech}[\Gamma_X]$, rewrites $\mathcal{H}_0^{VNC}$ into:

$$\mathbb{E}[X_t^\dagger | \{X_s^\dagger, Z_s\}_{s<t}] = 0$$

which reads same as Eq.(1). Theorem 2 thus applies directly to test variance noncausality by replacing $X_t$ with $X_t^\dagger$ in the test statistic. This connects to Portmanteau-type tests studied by Cheung and Ng (1996), Hong (2001), Aguilar and Hill (2015), and Leong and Urga (2023), which sum squared cross-correlations of centered squared residuals across lags.

## 3.3 Consistency under the Alternatives

This section establishes the asymptotic power under a general class of alternatives. Let $\kappa_{mrmr,XY}(t,j,k,l)$ denote the fourth-order cumulant of $\{X_{m,t}, Z_{r,t-j}, X_{m,t-k}, Z_{r,t-l}\}$, where $X_{m,t}$ and $Z_{r,t}$ are the $m^{th}$ and $r^{th}$ entries of $X_t$ and $Z_t$. We require absolute summability of fourth-order cumulants: $\sum_{m,r=1}^{d_1,d_2} \sum_{j,k,l=-\infty}^{\infty} \kappa_{mrmr,XZ}(0,j,k,l) < \infty$.

**Theorem 3.** *Suppose the assumptions of Theorem 2 hold. Suppose $\{X_t, Z_t\}$ is a jointly fourth-order stationary process with absolute summability of the fourth-order cumulants. Suppose further: $\exists j > 0$ , such that $||\Gamma_{XZ}(j)|| \neq 0$ , with $\sum_{j=1}^{\infty} ||\Gamma_{XZ}(j)||^2 < \infty$,*
*We have:*

$$\frac{M^{1/2}}{T} \left( \frac{T\widehat{\mathcal{T}}_\omega^c - \mu_{\omega,T}}{\sqrt{\widehat{D}_{\omega,T}^{(Hete)}}} \right) \xrightarrow{p} \Delta \sum_{j=1}^{\infty} \left|\left| vec \left[ \Gamma_X^{-1/2} \Gamma_{XZ}(j) \Gamma_Z^{-1/2} \right] \right|\right|^2$$

*for a finite* $\Delta > 0$. *By the asymptotic rates, equivalently:*

$$\lim_{T,M\to\infty} Pr\left(\left|\frac{T\widehat{\mathcal{T}}^{c}_{\omega} - \mu_{\omega,T}}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}}\right| > K\right) \to 1, \quad \forall K \in \mathbb{R}$$

Theorem 3 establishes a consistency result: the normalized statistic converges in probability to the sum of squared cross-correlations across lags (up to a positive scale). Under fixed alternatives with nonzero cross-correlation, the test statistic diverges at rate $T/M^{1/2}$. Slower growth of $M$ yields faster divergence and higher power, consistent with the discussion concluding Section 3.1 (see also p.517 Bouhaddioui and Roy, 2006). Fourth-order stationarity and absolute summability of joint cumulants are standard conditions imposed when studying the power of tests following Hong (1996b), as accommodate a wide class of processes (p.846 Hong, 1996a).

Two technical remarks clarify the role of fourth-order cumulants in our framework.

First, the proof exploits that under alternatives the sum of squares, $\widehat{\mathcal{T}}_{1\omega}$, stochastically dominates the sum of cross-products, $\widehat{\mathcal{T}}^{c}_{2\omega}$, as anticipated in Section 2.1. This follows from Theorem 6 in Hannan (1970) (p.210) via an Isserlis-type argument (see also Eq.(5.3.20) Priestley, 1981) establishing $\ell_2$-convergence of covariance estimators to their population counterparts under fourth-order stationarity and absolute summability. These conditions ensure the process is sufficiently close to multivariate normality for mean-square convergence to take effect. Crucially, while fourth-order cumulants are asymptotically negligible under alternatives (Theorem 3), they govern the asymptotic behavior under the null (Theorems 1–2). When $X$ is a martingale in higher moments, cumulants drive the distribution of the test statistic under $\mathcal{H}_0$ (Proposition 2). Hence, the correction term in Eq.(5) specifically targets the subset of cumulants associated with inverse causality.

Second, our asymptotic approach differs fundamentally from Escanciano and Velasco (2006). Their strategy, based on Cramér–von Mises norms similar to Hong and Lee (2005), establishes that sample autocovariances converge weakly in $\ell_2$-norm to a Gaussian process under the martingale difference null (their Theorem 1).[10] Since covariances enter their statistic squared, it converges in distribution to

[10] They generalize autocovariances to measure conditional mean dependence nonparametrically

a weighted sum of independent $\chi_1^2$ variables. In contrast, Hong (1996a) focuses on quadratic forms directly, showing that sums of cumulants converge to normality. The distinction is one of convergence order: Escanciano and Velasco (2006) achieve convergence of covariances *before* squaring, while Hong (1996a)'s work with cross-products *after* squaring. Our framework follows the latter approach, underscoring the importance of fourth-order cumulants when inverse causality is left unrestricted, precisely the setting where practitioners lack information about how omitted variables $Z$ interact with the structural shock and the structural dynamics.

The test has no power against alternatives with $\mathbb{E}[X_t|\{X_s, Z_s; s < t\}] \neq 0$ but zero cross-correlation (i.e., uncorrelated but non-martingale processes). This reflects a well-known limitation of Portmanteau tests: nonlinear dependencies from past $Z$ to present $X$ that leave linear associations unaffected cannot be detected (Hong, 2001; Bouhaddioui and Roy, 2006). One potential extension addresses this limitation through generalized spectral analysis (Hong and Lee, 2005). Rather than summing squared covariances between $X$ and $Z$, their approach considers squared covariances between $X$ and the empirical characteristic function of $Z$, thereby capturing nonlinear dependencies. Since their statistic involves quadratic forms (Eq.3.11-3 Hong and Lee, 2005), an analogous correction term can be constructed. Adapting their framework, define:

$$\mathcal{T}_\omega^{(HL)} = \sum_{j=1}^{T-1} \omega(j) \int_{\mathbb{R}} \left| \widehat{\Psi}_{XZ}(j;v) \right|^2 Q(dv), \qquad \widehat{\Psi}_{XZ}(j;v) = \frac{1}{T} \sum_{t=j+1}^{T} X_t \breve{Z}_{t-j}(v)$$

$$\breve{Z}_s(v) = e^{ivZ_s} - T^{-1} \sum_{m=1}^{T} e^{ivZ_s}$$

where $i$ denotes the imaginary unit and $Q : \mathbb{R} \to \mathbb{R}^+$ a nondecreasing weighting function symmetric about zero. Parallel to Eq.(5), the corrected Hong and Lee

(p.155). Convergence occurs in the Hilbert space of square-integrable functions; see pp.158–159 for details.

(2005)'s statistic is:

$$\mathcal{T}_{\omega}^{(HL),c} = \mathcal{T}_{\omega}^{(HL)} - \left( \frac{1}{T^2} \sum_{j=1}^{T-2} \omega(j) \int_{\mathbb{R}} \left( \sum_{s,t=j+1, s \leq t-j}^{T} X_t X_s \breve{Z}_{t-j}(v) \breve{Z}_{s-j}(v) \right) W(dv) \right)$$

A second limitation concerns the nature of detectable dependencies. Portmanteau statistics capture temporal dependence in a pairwise manner, detecting relationships of the form $\mathbb{E}[X_t|Z_{t-j}]$ but not necessarily non-pairwise interactions such as, for instance,
$\mathbb{E}[X_t|Z_{t-j}, Z_{t-j-1}]$. A standard solution is to augment $Z$ with its first $R$ lags: $\{Z_t^{\ddagger} = (Z_t', Z_{t-1}', ..., Z_{t-R}')'\}$, and then testing the null using cross-covariances between $X$ and $Z^{\ddagger}$ (see Domínguez and Lobato, 2004; Kuan and Lee, 2004; Wang et al., 2022). Albeit potentially applicable here, this approach complicates inference: the variance of the resulting statistic involves substantially more cross-product terms, requiring more elaborate correction terms to difference out the inverse causality effects. Developing such extensions remains an avenue for future research.

### 3.4 Simulation Study

This section summarizes the evidence from the simulation study in Appendix C that investigates the finite-sample properties of the corrected test statistics. All the statistics are formally defined in Appendix C.1. The study compares three finite-sample Portmanteau-type statistics: two versions of the finite-sample corrected statistics (*Hete* and *Hete2*) and the finite-sample benchmark statistic (*Hong*):

$$\frac{\mathcal{T}_{\omega}^{(f),c} - \mu_{\omega,T}^{(f)}}{\sqrt{\widehat{D}_{\omega,T}^{(f),(Hete)}}}, \quad \frac{\mathcal{T}_{1\omega}^{(f)} - \mu_{\omega,T}^{(f)}}{\sqrt{D_{\omega,T}^{(f)}}} + \frac{\mathcal{T}_{2\omega}^{(f),c}}{\sqrt{\widehat{D}_{\omega,T}^{(f),(Hete)}}}, \quad \frac{\mathcal{T}_{\omega}^{(f)} - \mu_{\omega,T}^{(f)}}{\sqrt{D_{\omega,T}^{(f)}}}$$

The difference between the previous statistics and these finite-sample versions is that the latter are scaled by effective sample, $T - j$ (or $T - \ell$), rather than the full sample, $T$. The difference between the two versions of the corrected statistics (*Hete* and *Hete2*, respectively) lies in the scaling applied to the sum of squares, which does not affect asymptotic size but matters for the finite-sample power of the testing procedures. Three additional statistics are also considered: i) the second version of the finite-sample corrected statistic (*Hete2*) applied to cross-correlations between residuals from an AR(1) fitted to $X$ and the process $Z$, as well as Wald statistics from a VAR(1) fitted to the joint process that test either a single zero-restriction (past $Z$ on present $X$) or a double zero-restriction (past $Z$ *and* past $X$ on present $X$). All Portmanteau-type statistics use the Bartlett kernel with smoothing parameters $M \in \{12, 36\}$, corresponding to 1 and 3 years of lags for monthly series. Regarding $\widehat{D}_{\omega,T}^{(Hete)}$, in place of $\widehat{\gamma}_{t,s}(j,\ell)$, a robustified variance estimator, $\widehat{\gamma}_{t,s}^{(rob)}(j,\ell)$, is adopted:

$$\widehat{\gamma}_{t,s}^{(rob)}(j,\ell) = \begin{cases} \langle \widehat{V}_{t-j}, \widehat{V}_{s-j} \rangle \langle \widehat{V}_{t-\ell}, \widehat{V}_{s-\ell} \rangle, & j = \ell \\ \langle \widehat{V}_{t-j}^{(w)}, \widehat{V}_{s-j}^{(w)} \rangle \langle \widehat{V}_{t-\ell}^{(w)}, \widehat{V}_{s-\ell}^{(w)} \rangle, & j \neq \ell \end{cases}$$

where $\{\widehat{V}_{l,s}^{(w)}\}$ denotes the winsorized version of the series $\{\widehat{V}_s\}$. The winsorization device applies only to off-diagonal elements ($j \neq \ell$) and aims to mitigate the influence of extreme realizations on the estimation of the long-run variance of cross-terms under non-normality. The baseline simulations consider the limiting case in which off-diagonal elements are set to zero. The main qualitative results remain unchanged under milder forms of winsorization.

Since the asymptotic theory emphasizes the correction term's importance under weak exogeneity, the simulations focus on empirical rejection rates for DGPs where the null holds but inverse causality is present. The main Monte Carlo experiments consider bivariate DGPs where the structural shock $X$ is i.i.d., while the omitted variable $Z$ follows an autoregressive process depending on past $X$ through either its first or second moments. Four DGP families capture different forms of inverse causality: linear and nonlinear dependence in the conditional mean, and ARCH and GARCH-type dependence in the conditional variance. The baseline simulations employ $T = 1000$ with $1,000$ replications at a 5% nominal significance level.

Under the null, the corrected statistics and Wald tests maintain rejection rates close to (or below) 5% across all specifications, while the benchmark suffers from severe size distortions reaching up to three times the nominal level. The benchmark's distortions are most pronounced when inverse causality operates through the conditional mean, particularly under nonlinearities, though they remain substantial even under inverse causality in the variances. When inverse causality is absent or $Z$ exhibits no serial dependence, the corrected statistics tend toward conservative inference, with rejection rates systematically below 5%. As inverse causality strengthens, however, the corrected statistics maintain empirical size near the nominal level, confirming that the correction term helps to control size.

The difference between Wald and corrected statistics is generally small, though under nonlinearities the Wald tests become more conservative (generally below 4%) while the corrected statistics remain better aligned with the nominal size. This pattern reflects that the Wald approach requires correct parametric specification of the joint system, whereas the corrected statistics avoid parametric assumptions about inverse causality. Among the corrected statistics, the AR residual version (*Hete2F*) exhibits tighter size control. This comes as unsurprising given that the null hypothesis jointly tests whether past $Z$ and past $X$ influence present $X$, and so incorporating information about the autoregressive structure of $X$ might be read as an analogy to the double zero-restriction tested by the Wald statistic. Nonetheless, all versions of the corrected statistic demonstrate substantially improved size properties relative to the benchmark.

Regarding power, the simulations consider four DGP families, similar to the previous ones, where weak exogeneity fails. The corrected statistics achieve power comparable to the benchmark under all alternatives, confirming the asymptotic equivalence in Theorem 3. As expected, power is reduced under nonlinear alternatives and second-order moment violations, reflecting that these tests primarily capture pairwise cross-correlations (Hong and Lee, 2005). The key finding is that introducing the correction does not meaningfully sacrifice power.

Taken together, the simulations corroborate the asymptotic theory: the correction term eliminates size distortions from inverse causality while preserving power. For practitioners, the Monte Carlo evidence suggests using the corrected test especially when: (a) there is no prior information on how past omitted variables enter the structural model, (b) the omitted variables exhibit temporal dependence, and (c) the structural shocks are non-Gaussian. Regarding the smoothing parameter $M$, the size correction prioritizes larger lag lengths, suggesting $\ln T \ll M < \sqrt{T}$. An informal rule of thumb for the lower bound is: $\underline{M} = 0.75T^{1/3}$ (Lazarus et al., 2018), and for the upper bound is: $\overline{M} = \sqrt{T} - 1$. More formally, one should appeal to the bandwidth rule in Hong and Lee (2005) using their plug-in method with a preliminary bandwidth $M^{pre} = c(10T)^{1/5}$, with $c = 2, 4, 6$ (Wang et al., 2022). Under the same design, extensive additional simulations (available upon request) adapt Escanciano and Lobato (2009)'s automatic rule for the optimal bandwidth when applied to the corrected statistics, however it does not convergence to a small range of values, and rather occasionally suggests values close to the upper bound $\overline{M}$.

# 4 Empirical application

This section presents an empirical application of the proposed testing procedures. Section 4.1 introduces the concept of fundamentalness of structural shocks and relates it to the null of interest. Section 4.2 reports the results of testing weak exogeneity for a broad set of popular structural shock measures. Section 4.3 focuses on the uncertainty shock of Baker et al. (2016) by revisiting the empirical analysis of Diercks et al. (2024). All the results are available in Appendix D.

## 4.1 Testing Fundamentalness of Structural Shocks

Since Sims (1980), Structural Vector AutoRegressive (SVAR) models have become the standard framework for empirical macroeconomic analysis. Common practice is to assume that the macroeconomic multivariate time series, $\{W_t\}$, admits a Moving Average (MA) representation driven by mutually orthogonal structural shocks, $\{\epsilon_t\}$:

$$W_t = B(L)\epsilon_t = \sum_{j=0}^{\infty} B_j \epsilon_{t-j}, \qquad \epsilon_t \sim (0, I)$$

where $B(L)$ captures the propagation of the structural disturbances.[11] When $W$ is causal and invertible, structural shocks can be recovered from current and lagged values of $W$, up to a rotation matrix which governs the instantaneous relationships among the components of $W$, thus in turn motivating the use of SVAR models. In-

[11] This rationale is supported by a twofold motivation: i) by the Wold Representation theorem, if the time series is covariance-stationary then it admits a MA($\infty$) representation (Brockwell and Davis (1987)), with the Wold innovations being the reduced-form residuals of the linear projection of $W$ onto its infinite past; ii) the linear (or linearized) dynamic stochastic economic model, based on the variables $W$, usually admits a VARMA solution, whose structural shocks are assumed to be mutually orthogonal (Fernández-Villaverde et al., 2007).

vertibility may however fail when economic agents' information differs from the econometrician's (Hansen and Sargent, 2019) and, in such cases, the MA representation is said to be non-fundamental (Lippi and Reichlin, 1994; Nakamura and Steinsson, 2018). In practice, the issue of non-fundamentalness/invertibility spells out as a problem of VAR misspecification, due to omitted variables or insufficient set of lagged controls (Chen et al., 2017; Miranda-Agrippino and Ricco, 2023).

To fix ideas, following Giannone and Reichlin (2006), partition $W$ into two blocks $W_1$ and $W_2$ of dimensions $d_1$ and $d_2$, with reduced-form residuals $(X_t, Z_t)'$ and structural shocks $(\epsilon_{1,t}, \epsilon_{2,t})'$, where the process $Z$ collects the innovations of the block $W_2$. Suppose for simplicity that $B_0 = I$, so that structural shocks coincide with reduced-form residuals:

$$\begin{pmatrix} W_{1,t} \\ W_{2,t} \end{pmatrix} = \begin{pmatrix} A_{1,1}(L) & A_{1,2}(L) \\ A_{2,1}(L) & A_{2,2}(L) \end{pmatrix} \begin{pmatrix} X_t \\ Z_t \end{pmatrix}$$

Suppose the econometrician is interested in recovering the structural shocks, $\{\epsilon_{1,t} = X_t\}$, but omits from the empirical analysis the block $W_2$ (and so the space spanned by $Z$). The MA representation is fundamental only if $A_{1,2}(L) = 0$ holds or, equivalently, whether there is Granger noncausality from the omitted variables $W_2$ to $X$. Vice versa, if $A_{1,2}(L) \neq 0$, recovering the structural shock requires the enlarged information set $\{W_{1,t}, W_{2,t}\}$ or, equivalently, $\{W_{1,t}, Z_t\}$.

As the example makes clear, testing fundamentalness reduces to testing Granger non-causality or conditional lagged exogeneity (Giannone and Reichlin, 2006; Forni and Gambetti, 2014; Plagborg-Møller and Wolf, 2022; Miranda-Agrippino and Ricco, 2023). Complementarily, Chen et al. (2017) show that, when the DGP is a VARMA process generated by non-Gaussian i.i.d. shocks, fundamentalness holds if and only if the reduced-form innovations are m.d.s. (Chen et al., 2017, Theorem 1). Maintaining this assumption, the shocks $X$ are fundamental/invertible if jointly: i) $\{X_t\}$ is m.d.s. with respect to its own past, $\sigma(X_{t-1}, ...)$, that is invertibility of $W_1$ alone, and ii) $\{X_t\}$ is m.d.s. with respect to the past of the omitted innovations, $\sigma(Z_{t-1}, ...)$. Taken together, these two conditions amount to the structural shocks $X$ having zero conditional mean given the past of both internal and external variables, $\sigma(X_{t-1}, Z_{t-1}, ...)$, which is the null of weak exogeneity in Eq.(1).

## 4.2 Weak Exogeneity of Macroeconomic Shocks

This paper investigates the exogeneity properties of eleven popular measures of macroeconomic shocks at monthly frequency, $\{X_t\}$ spanning three uncertainty (UN) and eight monetary policy (MP) shock series. The UN structural shocks are: Baker et al. (2016)'s Economic Policy Uncertainty (EPU) shocks (discussed in detail in Section 4.3); Ludvigson et al. (2021)'s financial uncertainty shocks, identified via event restrictions and external variable constraints; Berger et al. (2020)'s shocks derived from realized market volatility under a recursive identification scheme. The MP shocks are: Aruoba and Drechsel (2024)'s orthogonalized changes in the Federal Funds Rate using natural language processing and machine learning techniques; Bu et al. (2021)'s monetary policy shock series estimated via Fama-MacBeth procedure from the common component of outcome variables and FOMC announcement day changes in interest rate across the full maturity spectrum; Bauer and Swanson (2023)'s monetary policy surprise (MPS) and its version orthogonalized with respect to pre-announcement macroeconomic and financial data (MPS Ort); Jarociński and Karadi (2020)'s three shock decompositions, i.e., the first principal component of interest rate derivative surprises (PC), and its Central Bank Information (CBI) and Monetary Policy Information (MPI) components obtained via "*Poor Man's*" sign restrictions; Miranda-Agrippino and Ricco (2021)'s conventional Fed monetary policy instrument.[12]

As candidates for the omitted variables, $\{Z_t\}$, this paper follows Forni and Gambetti (2014) and considers the first principal components of large macroeconomic and financial datasets, motivated by the idea that, under a state space representation of the economic system, these estimated factors should approximate the relevant state variables. Specifically, four sets of factors are considered: (i) 8 macroeconomic factors from McCracken and Ng (2016) (McK Ng); (ii) 8 macroeconomic factors from Rapach and Zhou (2021) (RZ); (iii) 5 financial factors from Giglio and Xiu (2021) (GX); (iv) 5 financial factors from Lettau and Pelger (2020b) (LP). The first two sets are extracted from FRED-MD, a database of 134 monthly U.S. macroeconomic indicators, and control for the state of the macroeconomy. The latter two, extracted from a panel of 647 portfolios spanning equities, corporate

[12] About the data, the first three series are provided by the replication package of Diercks et al. (2024). Regarding the latter shock series, I thank the authors, Thomas Drechsel, John Rogers, Michael D. Bauer, Marek Jarocinski and Silvia Miranda-Agrippino, for publicly sharing the data.

bonds, and currencies, principally control for financial market conditions.[13] All four sets of factors display nonnegligible persistence, therefore aligning with the simulation evidence about the magnitude of the correction term when the process $Z$ is autoregressive.[14]

The testing procedures employed are the finite-sample versions of the benchmark and corrected Portmanteau statistics in the baseline simulation study (Section 3.4):

$$Hong : \frac{\mathcal{T}_{\omega}^{(f)} - \mu_{\omega,T}^{(f)}}{\sqrt{D_{\omega,T}^{(f)}}}, \qquad Hete : \frac{\mathcal{T}_{\omega}^{(f),c} - \mu_{\omega,T}^{(f)}}{\sqrt{\widehat{D}_{\omega,T}^{(f),(Hete)}}}$$

with detailed definitions in Appendix C.1. When discussing the departure from the null, Table 1 further decomposes the statistics into the three relevant components: the sum of squares and the corrected part of the sum of cross-products (see Eq.(4-5)). The kernel used is Andrews (1991)'s quadratic spectral kernel.

The results, summarized in Table 1 and detailed in Appendix D, reveal a clear pattern. Weak exogeneity is rejected at 5% for all three UN shock series: Baker et al. (2016)'s with respect to McK Ng (Section 4.3) and GX factors ($M > 20$ months); Ludvigson et al. (2021)'s with respect to GX ($M < 30$ months) and LP ($M < 9$ months); Berger et al. (2020)'s with respect to GX ($M < 25$ months). Among MP shocks, we fail to reject the null for Aruoba and Drechsel (2024), Bu et al. (2021), and Miranda-Agrippino and Ricco (2021) at 10%, while Bauer and Swanson (2023)'s series (MPS and MPS Ort) are rejected with respect to GX (at 10% level, when $6 < M < 20$ months). Weak exogeneity is rejected at 5% for: Bauer and Swanson (2023)'s series (MPS and MPS Ort) with respect to LP ($M > 4$ months); Jarociński and Karadi (2020)'s CBI series with respect to LP ($M > 18$ months); Jaro-

[13]Regarding the first set of factors, using the FRED-MD dataset spanning until June 2021, I estimate 8 static factors by principal component analysis (PCA) adapted to allow for missing values (Stock and Watson (2002)'s EM algorithm). Regarding the second set of factors, using the same dataset, I estimate 8 static factors by sparse PCA following Rapach and Zhou (2021). Regarding the third set of factors, I estimate 5 static factors by PCA, using the dataset in their replication package. Covering from Jan. 1976 to Nov. 2009, their dataset consists in a large panel of 647 portfolios that include US equities as well as treasury bonds, corporate bonds, and currencies. Regarding the last set of factors, I estimate 5 latent factors following Lettau and Pelger (2020a)'s methodology. In the Appendix D, this paper considers a more recent vintage ofMcCracken and Ng (2016)'s.

[14]Preliminary analysis of VAR information criteria indicates optimal lag lengths of around 4 (AIC) and 2 (BIC) for the macroeconomic factors, and 3 (AIC) and 1 (BIC) for the financial factors.

Table 1: Testing Shock Exogeneity: Summary. This table presents a summary of the results regarding (weak) exogeneity properties of popular measures of macroeconomic structural shocks. The focus is on UNcertainty (UN) shocks (Baker et al., 2016; Ludvigson et al., 2021; Berger et al., 2020) and Monetary Policy (MP) shocks (Aruoba and Drechsel, 2024; Bu et al., 2021; Bauer and Swanson, 2023; Jarociński and Karadi, 2020; Miranda-Agrippino and Ricco, 2021). The set of omitted factors are: i) Macroeconomic factors: a) McCracken and Ng (2016) (McK Ng), b) Rapach and Zhou (2021) (RZ); ii) Financial factors: c) Giglio and Xiu (2021) (GX), Lettau and Pelger (2020b) (LP). The columns display whether we fail to reject the null of weak exogeneity ($\surd$) at 10% significance level. In case we reject the null at 5% ($\times$) with respect to a set of omitted factors, the columns display whether it is mainly due to the sum of squares ($\mathcal{T}_{1\omega}$) and/or the corrected sum of cross-products ($\mathcal{T}^c_{2\omega}$). For more details, refer to Appendix D.

| | | Fail to Reject | Reject Weak Exogeneity | | |
|---|---|---|---|---|---|
| | | Weak Exogeneity | Sum of Cross-terms | Sum of Squares | Omitted factors |
| UN | Baker et al. (2016) | $\times$ | $\star$ | | Macro (McK Ng) |
| | | | $\star$ | $\star$ | Finance (GX) |
| | Ludvigson et al. (2021) | $\times$ | $\star$ | $\star$ | Finance (GX, LP) |
| | Berger et al. (2020) | $\times$ | $\star$ | | Finance (GX) |
| MP | Aruoba and Drechsel (2024) | $\surd$ | | | |
| | Bu et al. (2021) | $\surd$ | | | |
| | Bauer and Swanson (2023) | $\times$ | $\star$ | $\star$ | Finance (GX) |
| | MPS and MPS Ort | | | $\star$ | Finance (LP) |
| | Jarociński and Karadi (2020) | $\times$ | | | |
| | i) Central Bank Inf. (CB) | | | $\star$ | Finance (LP) |
| | ii) Monetary Policy Inf. (MP) | | $\star$ | | Macro (RZ) |
| | | | | $\star$ | Finance (LP) |
| | iii) Principal Component (PC) | | $\star$ | | Macro (RZ) |
| | Miranda-Agrippino and Ricco (2021) | $\surd$ | | | |

ciński and Karadi (2020)'s MPI series with respect to RZ ($M > 18$ months) and LP factors ($M > 18$ months); c) Jarociński and Karadi (2020)'s PC series with respect to RZ ($M > 12$ months). Across all rejections, both procedures point to a relatively large number of relevant lags, generally between one and two years (12 months $\approx 2 \ln T$ to 24 months $\approx \sqrt{T}$, with $T \approx 500$).

These findings suggest that several popular measures of shock series may not represent truly unanticipated movements in the macroeconomic system, as they are forecastable by macroeconomic or financial factors (i.e., departures from weak exogeneity). This constitutes evidence in favor of augmenting the structural dynamics by those factors to recover a more exogenous shock series. However, due their nonparametric nature, the proposed procedures do not prescribe the functional form or lag specification of such parametric augmentation, therefore should be interpreted rather as pre-testing strategies. By revisiting Dierckx et al. (2024), the next section illustrates how these findings can be used to improve inference on impulse responses and to understand the economic consequences of departures from weak exogeneity.

### 4.3 Baker et al. (2016)'s EPU Shocks

Baker et al. (2016) construct a measure of economic policy uncertainty (EPU) based on the frequency of newspaper articles that simultaneously contain terms related to uncertainty, the economy, and policy. Following their Section IV.D, this paper estimates the EPU structural shock at monthly frequency by fitting a VAR(6) to five monthly U.S. time series from January 1985 to December 2019, imposing a Cholesky ordering with the EPU index first, followed by the log of the S&P 500 index, the federal funds rate, log employment and log industrial production. As noted by Diercks et al. (2024), the estimated structural shocks are serially uncorrelated.

The small VAR system, however, arguably controls for all relevant macroeconomic conditions. This turns to be critical when questioning the exogeneity of the estimated shock, especially since the EPU index should also capture *"uncertainties related to the economic ramifications of "noneconomic" policy matters [...] both near-term concerns [...] and longer term concerns"* (Baker et al., 2016). To address this, this paper tests whether the EPU shock is weakly exogenous with respect to its own past and the past of McCracken and Ng (2016)'s macroeconomic factors, using both Portmanteau-type testing procedures.

The left panel of Figure 1 reports the results. Both testing procedures reject the null of weak exogeneity at 10% significance level, though they deliver a somewhat different picture across horizons. The corrected statistic points to rejection from long horizons ($M > 24$ months) onwards. The benchmark statistic, instead, rejects only over a narrow window of horizons ($19 < M < 34$ months), failing to reject for longer ones, producing an inverse U-shaped curve visible in the left panel.

The right panel of Figure 1 sheds light on this discrepancy by decomposing the benchmark statistic as in eq.(5). The relative importance of the corrected causality channel ($\mathcal{T}^c_{2\omega}$) remains stable across horizons, while the inverse causality channel ($C_\omega$) contributes more heavily at shorter horizons and fades at longer ones, thus generating the ambiguous conclusions in terms of inference using the benchmark statistic. The corrected statistic, being robust to the inverse causality channel, is not affected by this issue and delivers a more coherent picture across horizons. In both cases, the contribution of the sum of squares ($\mathcal{T}_{1\omega}$) across horizons is mostly negligible.

In light of these results, the EPU shocks cannot be deemed fundamental/invertible

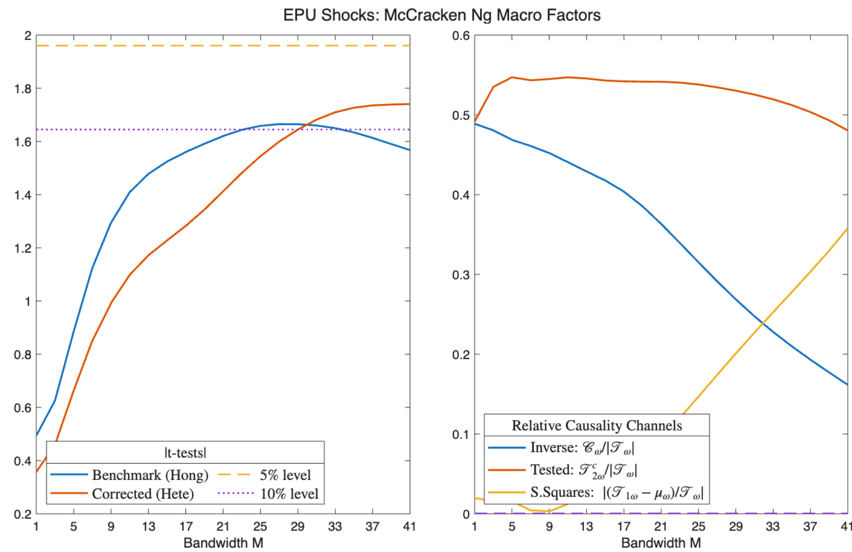


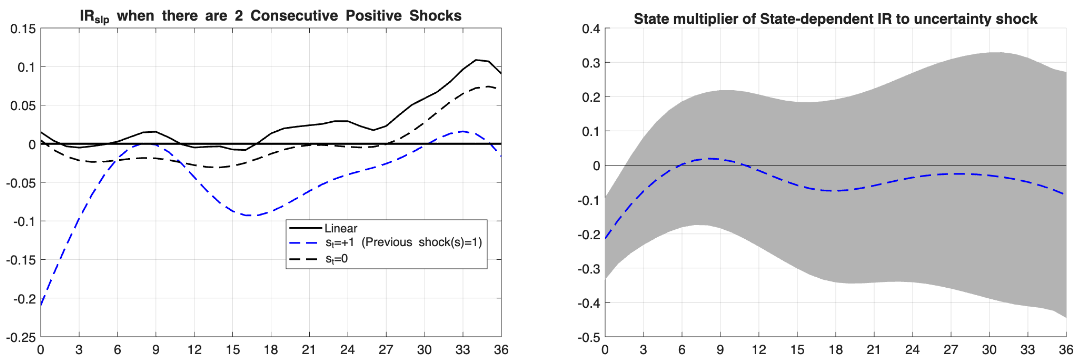

Figure 1: Baker et al. (2016)'s shocks and McCracken and Ng (2016)'s factors. Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

when considering McCracken and Ng (2016)'s macroeconomic factors. Thus, impulse response analysis can therefore benefit from augmenting the information set with McCracken and Ng (2016)'s factors. To illustrate this, I revisit Diercks et al. (2024)'s analysis of the superadditive effects of uncertainty shocks, focusing on the impulse response of inflation (PCE) and stock market (S&P500) to EPU shocks. Diercks et al. (2024) estimate the following set of state-dependent local projections:

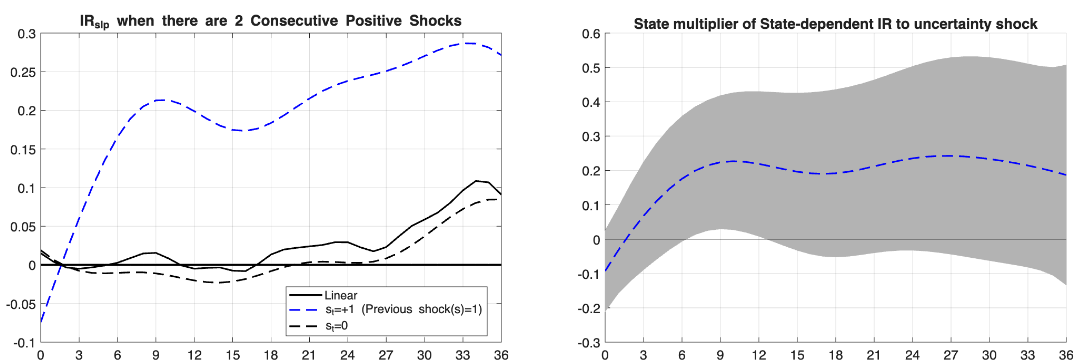
$$y_{t+h} = \alpha_h + \big(\beta_{0,h} + \beta_{1,h}\mathbb{1}\{\epsilon_{unc,t-1} > 0, .., \epsilon_{unc,t-L} > 0\}\big)\epsilon_{unc,t} + \sum_{i=1}^{p} \gamma_{i,h} w_{t-i} + u_{t+h}$$

where $h$ sets the predictive horizon, ranging from 0 to 36 months, $p$ is the number lags used for the control variables, $\{w_t\}$, and the indicator function takes value 1 if each one of the previous $L$ uncertainty shocks $\{\epsilon_{unc,t-1}, .., \epsilon_{unc,t-L}\}$ has been positive. The coefficient $\beta_{1,h}$ (i.e., the "state multiplier") captures the superadditive effect of uncertainty: the impact of a cascade of positive uncertainty shocks is more severe than the isolated sum of them. In the application of their Appendix B.1 (Figure B.2), the shock of interest $\epsilon_{unc}$ is EPU shock, the outcome variable $y$ is inflation, the number of consecutive positive shocks is 2 ($L = 1$), the number of lags for the controls is 6, and the control set follows Baker et al. (2016) (augmented by lagged values of the outcome variable). Figure 2a reproduces their baseline results.

Upon adding two lags of McCracken and Ng (2016)'s macroeconomic factors to the controls in the local projections, the unconditional linear impulse response remains largely unchanged (dashed vs. solid black lines in Figure 2), while the state-dependent response, $\{\beta_{1,h}\}_{h=0,...,H}$, becomes significantly positive. Strengthening Diercks et al. (2024)'s conclusions, this corroborates the superadditivity of EPU shocks: a sequence of consecutive positive uncertainty shocks leads to a noticeable increase of inflation, substantially more severe than the case when there is no such consecutive sequence. This finding connects to two strands of the literature. First, it aligns with Fernández-Villaverde et al. (2015), who show that unexpected changes in fiscal policy uncertainty can lead to increased inflation within a standard New Keynesian model. Second, it relates closely to Ascari et al. (2023), who demonstrate that, in a rich DSGE model with firm dynamics, an 'expectational' shock that raises short-term inflation expectations results in negative macroeconomic effects, causing inflation to rise while output declines (stagflationary response).




(a) Replication of Figure B.2.(e)-(f) in Diercks et al. (2024)'s Appendix B.1



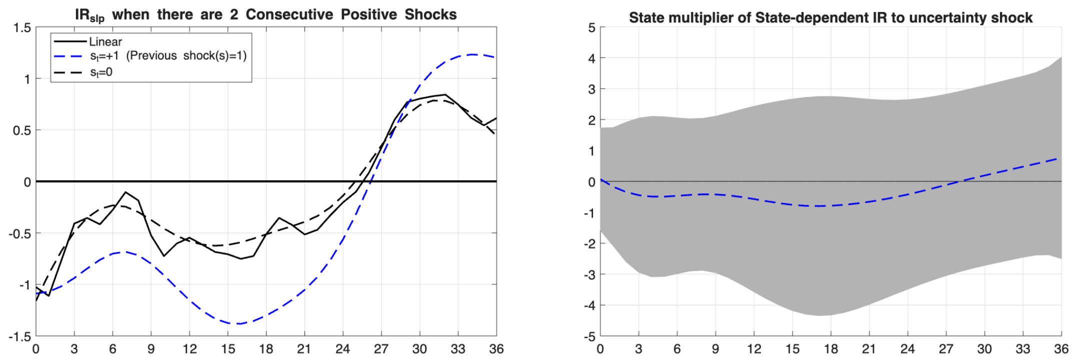

(b) Inclusion of two lags of McCracken and Ng (2016)'s factors to the set of controls

Figure 2: Response of price level to consecutive positive EPU uncertainty shocks: LEFT PANELS: the empirical state-dependent impulse responses (estimated with LPs as in Diercks et al. (2024)) to two consecutive positive uncertainty shocks (dashed blue line) and contrast it to the response to a single shock in the state-dependent model (dashed black line), and in the linear model (solid black line). RIGHT PANELS: the incremental effect of the second shock, i.e. $\{\beta_{1,h}\}_{h=1,..,H}$, with 90% confidence intervals (shaded area). In both panels, on the y-axes, the level of impulse responses; on the x-axes, the horizons, $h$. Price level is measured by the Personal Consumption Expenditures (PCE) Price Index.

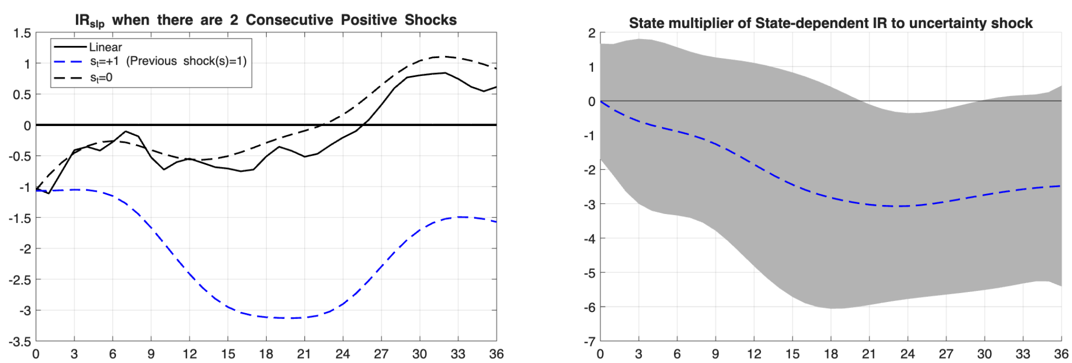
Turning to the remaining variables of the system, the impulse responses are broadly consistent with Diercks et al. (2024): industrial production and short rates respond negatively (Figures 26-27 in Appendix D.4). The notable exception is the stock market. The state-dependent response of the real S&P 500 to consecutive positive EPU shocks becomes significantly more negative at longer horizons (Figure 3), consistent with Berger et al. (2020)'s finding that uncertainty shocks are linked to declines in stock returns (refer to their discussion in Section 5.2).

A comparison across uncertainty shocks reveals an interesting heterogeneity. The inflation response to shocks from the other two considered measures (Ludvigson et al. (2021)'s and Berger et al. (2020)'s) is strongly negative, in contrast to the positive inflationary effect documented for EPU shocks. Meanwhile, the response of industrial production is negative across all three shock series, confirming their countercyclical nature (Diercks et al., 2024). This pattern suggests that EPU shocks might operate primarily through a supply-side channel, whereas the financial uncertainty and realized volatility uncertainty shocks might be better characterized as demand-side disturbances.

In conclusion, after accounting for additional macroeconomic factors, Baker et al. (2016)'s uncertainty shocks tend to behave as superadditive 'expectational' shocks: a cascade of positive EPU shocks raises inflation and rather depresses industrial production and the stock markets, with more severe effects the longer the sequence persists.




(a) Replication of Figure B.3(e)-(f) in Diercks et al. (2024)




(b) Inclusion of two lags of McCracken and Ng (2016)'s factors to the set of controls

Figure 3: Response of stock market to consecutive positive EPU uncertainty shocks:
LEFT PANELS: the empirical state-dependent impulse responses (estimated with LPs as in Diercks et al. (2024)) to two consecutive positive uncertainty shocks (dashed blue line) and contrast it to the response to a single shock in the state-dependent model (dashed black line), and in the linear model (solid black line). RIGHT PANELS: the incremental effect of the second shock, i.e. $\{\beta_{1,h}\}_{h=1,..,H}$, with 90% confidence intervals (shaded area). In both panels, on the y-axes, the level of impulse responses; on the x-axes, the horizons, $h$. Stock market is measured by real S&P 500.

# 5 Conclusion

This paper studies specification testing in dynamic linear models in the presence of omitted variables, focusing on tests of weak exogeneity of structural shocks. Standard Portmanteau tests based on quadratic forms of serial cross-correlations might confound violations of weak exogeneity with dependence running from past shocks to current omitted variables (inverse causality), potentially producing misleading rejections. To address this issue, the paper proposes an asymmetric Portmanteau statistic that introduces a correction term which removes the influence of inverse causality. The corrected statistic isolates the directional restriction implied by weak exogeneity without requiring parametric modeling of the joint dynamics. Under mild conditional moment restrictions, it is asymptotically normal under the null, both for observed and estimated processes, and achieves asymptotic power comparable to the benchmark under fixed alternatives. An empirical application revisits the exogeneity of widely used macroeconomic shock measures. The proposed test finds that Baker et al. (2016)'s Economic Policy Uncertainty shocks violate weak exogeneity. By revisiting Diercks et al. (2024), enlarging the information set accordingly leads to a positive inflation response and contractionary effects elsewhere, consistent with a supply-side interpretation.

# References


AGUILAR, M. AND J. B. HILL (2015): "Robust score and portmanteau tests of volatility spillover," *Journal of Econometrics*, 184, 37–61.

ANDREWS, D. W. (1991): "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica: Journal of the Econometric Society*, 817–858.

ARUOBA, S. B. AND T. DRECHSEL (2024): "Identifying monetary policy shocks: A natural language approach," Tech. rep., National Bureau of Economic Research.

ASCARI, G., S. FASANI, J. GRAZZINI, AND L. ROSSI (2023): "Endogenous uncertainty and the macroeconomic impact of shocks to inflation expectations," *Journal of Monetary Economics*, 140, S48–S63.

BAKER, S. R., N. BLOOM, AND S. J. DAVIS (2016): "Measuring economic policy uncertainty," *The quarterly journal of economics*, 131, 1593–1636.

BAUER, M. D. AND E. T. SWANSON (2023): "A reassessment of monetary policy surprises and high-frequency identification," *NBER Macroeconomics Annual*, 37, 87–155.

BERGER, D., I. DEW-BECKER, AND S. GIGLIO (2020): "Uncertainty shocks as second-moment news shocks," *The Review of Economic Studies*, 87, 40–76.

BOUHADDIOUI, C. AND R. ROY (2006): "A generalized portmanteau test for independence of two infinite-order vector autoregressive series," *Journal of Time Series Analysis*, 27, 505–544.

BOX, G. E. AND D. A. PIERCE (1970): "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models," *Journal of the American statistical Association*, 65, 1509–1526.

BROCKWELL, P. J. AND R. A. DAVIS (1987): "Time Series: Theory and Methods," *Springer Series in Statistics*.

BROWN, B. M. (1971): "Martingale central limit theorems," *The Annals of Mathematical Statistics*, 59–66.

BRUNNERMEIER, M., D. PALIA, K. A. SASTRY, AND C. A. SIMS (2021): "Feedbacks: financial markets and economic activity," *American Economic Review*, 111, 1845–1879.

BRUNS, M. AND S. A. KEWELOH (2024): "Testing for strong exogeneity in Proxy-VARs," *Journal of Econometrics*, 245, 105876.

BU, C., J. ROGERS, AND W. WU (2021): "A unified measure of Fed monetary policy shocks," *Journal of Monetary Economics*, 118, 331–349.

CANDELON, B. AND S. TOKPAVI (2016): "A nonparametric test for granger causality in distribution with application to financial contagion," *Journal of Business & Economic Statistics*, 34, 240–253.

CHEN, B., J. CHOI, AND J. C. ESCANCIANO (2017): "Testing for fundamental vector moving average representations," *Quantitative Economics*, 8, 149–180.

CHEUNG, Y.-W. AND L. K. NG (1996): "A causality-in-variance test and its application to financial market prices," *Journal of econometrics*, 72, 33–48.

COMTE, F. AND O. LIEBERMAN (2000): "Second-order noncausality in multivariate GARCH processes," *Journal of time series analysis*, 21, 535–557.

DEDECKER, J., P. DOUKHAN, G. LANG, L. R. JOSÉ RAFAEL, S. LOUHICHI, C. PRIEUR, J. DEDECKER, P. DOUKHAN, G. LANG, L. R. JOSÉ RAFAEL, ET AL. (2007): *Weak dependence*, Springer.

DIERCKS, A. M., A. HSU, AND A. TAMONI (2024): "When it rains it pours: Cascading uncertainty shocks," *Journal of Political Economy*, 132, 694–720.

DOMÍNGUEZ, M. A. AND I. N. LOBATO (2004): "Consistent estimation of models defined by conditional moment restrictions," *Econometrica*, 72, 1601–1615.

——— (2020): "Specification testing with estimated variables," *Econometric Reviews*, 39, 476–494.

DURLAUF, S. N. (1991): "Spectral based testing of the martingale hypothesis," *Journal of Econometrics*, 50, 355–376.

ESCANCIANO, J. C. AND I. N. LOBATO (2009): "An automatic portmanteau test for serial correlation," *Journal of Econometrics*, 151, 140–149.

ESCANCIANO, J. C. AND C. VELASCO (2006): "Generalized spectral tests for the martingale difference hypothesis," *Journal of Econometrics*, 134, 151–185.

FERNÁNDEZ-VILLAVERDE, J., P. GUERRÓN-QUINTANA, K. KUESTER, AND J. RUBIO-RAMÍREZ (2015): "Fiscal volatility shocks and economic activity," *American Economic Review*, 105, 3352–3384.

FERNÁNDEZ-VILLAVERDE, J., J. F. RUBIO-RAMÍREZ, T. J. SARGENT, AND M. W. WATSON (2007): "ABCs (and Ds) of understanding VARs," *American economic review*, 97, 1021–1026.

FORNI, M. AND L. GAMBETTI (2014): "Sufficient information in structural VARs," *Journal of Monetary Economics*, 66, 124–136.

FRANCQ, C. AND H. RAÏSSI (2007): "Multivariate portmanteau test for autoregressive models with uncorrelated but nonindependent errors," *Journal of Time Series Analysis*, 28, 454–470.

GIANNONE, D. AND L. REICHLIN (2006): "Does information help recovering structural shocks from past observations?" *Journal of the European Economic Association*, 4, 455–465.

GIGLIO, S. AND D. XIU (2021): "Asset pricing with omitted factors," *Journal of Political Economy*, 129, 1947–1990.

GRANGER, C. W. (1969): "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: journal of the Econometric Society*, 424–438.

HAFNER, C. M., H. HERWARTZ, AND S. MAXAND (2022): "Identification of structural multivariate GARCH models," *Journal of Econometrics*, 227, 212–227.

HALL, P. AND C. C. HEYDE (2014): *Martingale limit theory and its application*, Academic press.

HANNAN, E. J. (1970): *Multiple time series*, vol. 38, John Wiley & Sons.

HANSEN, L. P. AND T. J. SARGENT (2019): "Two difficulties in interpreting vector autoregressions," in *Rational expectations econometrics*, CRC Press, 77–119.

HAUGH, L. D. (1976): "Checking the independence of two covariance-stationary time series: a univariate residual cross-correlation approach," *Journal of the American Statistical Association*, 71, 378–385.

HONG, Y. (1996a): "Consistent testing for serial correlation of unknown form," *Econometrica: Journal of the Econometric Society*, 837–864.

——— (1996b): "Testing for independence between two covariance stationary time series," *Biometrika*, 83, 615–625.

——— (1999): "Hypothesis testing in time series via the empirical characteristic function: a generalized spectral density approach," *Journal of the American Statistical Association*, 94, 1201–1220.

——— (2001): "A test for volatility spillover with application to exchange rates," *Journal of Econometrics*, 103, 183–224.

HONG, Y. AND Y.-J. LEE (2005): "Generalized spectral tests for conditional mean models in time series with conditional heteroscedasticity of unknown form," *The Review of Economic Studies*, 72, 499–541.

HONG, Y., Y. LIU, AND S. WANG (2009): "Granger causality in risk and detection of extreme risk spillover between financial markets," *Journal of Econometrics*, 150, 271–287.

HOSKING, J. R. (1980): "The multivariate portmanteau statistic," *Journal of the American Statistical Association*, 75, 602–608.

JAROCIŃSKI, M. AND P. KARADI (2020): "Deconstructing monetary policy surprises—the role of information shocks," *American Economic Journal: Macroeconomics*, 12, 1–43.

JORDÀ, Ò. (2023): "Local projections for applied economics," *Annual Review of Economics*, 15, 607–631.

JURADO, K., S. C. LUDVIGSON, AND S. NG (2015): "Measuring uncertainty," *American Economic Review*, 105, 1177–1216.

KILIAN, L. AND H. LÜTKEPOHL (2017): *Structural vector autoregressive analysis*, Cambridge University Press.

KUAN, C.-M. AND W.-M. LEE (2004): "A New Test of the Martingale Difference Hypothesis." *Studies in Nonlinear Dynamics & Econometrics*, 8.

LAZARUS, E., D. J. LEWIS, AND J. H. STOCK (2021): "The Size-Power Tradeoff in HAR Inference," *Econometrica*, 89, 2497–2516.

LAZARUS, E., D. J. LEWIS, J. H. STOCK, AND M. W. WATSON (2018): "HAR inference: Recommendations for practice," *Journal of Business & Economic Statistics*, 36, 541–559.

LEONG, S. H. AND G. URGA (2023): "A practical multivariate approach to testing volatility spillover," *Journal of Economic Dynamics and Control*, 104694.

LETTAU, M. AND M. PELGER (2020a): "Estimating latent asset-pricing factors," *Journal of Econometrics*, 218, 1–31.

——— (2020b): "Factors that fit the time series and cross-section of stock returns," *The Review of Financial Studies*, 33, 2274–2325.

LI, W. K. AND A. I. MCLEOD (1981): "Distribution of the residual autocorrelations in multivariate ARMA time series models," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 43, 231–239.

LIPPI, M. AND L. REICHLIN (1994): "VAR analysis, nonfundamental representations, Blaschke matrices," *Journal of Econometrics*, 63, 307–325.

LJUNG, G. M. AND G. E. BOX (1978): "On a measure of lack of fit in time series models," *Biometrika*, 65, 297–303.

LOBATO, I. N., J. C. NANKERVIS, AND N. E. SAVIN (2002): "Testing for zero autocorrelation in the presence of statistical dependence," *Econometric Theory*, 18, 730–743.

LUDVIGSON, S. C., S. MA, AND S. NG (2021): "Uncertainty and business cycles: exogenous impulse or endogenous response?" *American Economic Journal: Macroeconomics*, 13, 369–410.

LÜTKEPOHL, H. (1997): *Handbook of matrices*, John Wiley & Sons.

——— (2013): *Introduction to multiple time series analysis*, Springer Science & Business Media.

MCCRACKEN, M. W. AND S. NG (2016): "FRED-MD: A monthly database for macroeconomic research," *Journal of Business & Economic Statistics*, 34, 574–589.

MIKUSHEVA, A. AND M. SØLVSTEN (2025): "Linear regression with weak exogeneity," *Quantitative Economics*, 16, 367–403.

MIRANDA-AGRIPPINO, S. AND G. RICCO (2021): "The transmission of monetary policy shocks," *American Economic Journal: Macroeconomics*, 13, 74–107.

——— (2023): "Identification with external instruments in structural VARs," *Journal of Monetary Economics*, 135, 1–19.

NAKAMURA, E. AND J. STEINSSON (2018): "Identification in macroeconomics," *Journal of Economic Perspectives*, 32, 59–86.

PLAGBORG-MØLLER, M. AND C. K. WOLF (2021): "Local projections and VARs estimate the same impulse responses," *Econometrica*, 89, 955–980.

——— (2022): "Instrumental variable identification of dynamic variance decompositions," *Journal of Political Economy*, 130, 2164–2202.

PRIESTLEY, M. B. (1981): *Spectral analysis and time series: Univariate series*, vol. 1, Academic press.

RAMEY, V. A. (2016): "Macroeconomic shocks and their propagation," *Handbook of macroeconomics*, 2, 71–162.

RAPACH, D. AND G. ZHOU (2021): "Sparse macro factors," *Available at SSRN 3259447*.

SIMS, C. A. (1980): "Macroeconomics and reality," *Econometrica: journal of the Econometric Society*, 1–48.

STOCK, J. H. AND M. W. WATSON (2002): "Forecasting using principal components from a large number of predictors," *Journal of the American statistical association*, 97, 1167–1179.

WANG, G., K. ZHU, AND X. SHAO (2022): "Testing for the martingale difference hypothesis in multivariate time series models," *Journal of Business & Economic Statistics*, 40, 980–994.

# A Appendix

## A.1 Formulations of the statistics

**Lemma A.1.**

$$2\pi \int_{2\pi} vec[\overline{f(\hat{\lambda})}]' vec[f(\hat{\lambda})] d\lambda = T_\omega$$

*Proof.* By using the property $\text{tr}(A'C) = \text{vec}(A)'\text{vec}(C)$, together with the properties of the conjugates[15] and the interchangeability of trace and integral operator, we can rewrite the former as:

$$2\pi \int_{2\pi} \text{vec}[\overline{f(\hat{\lambda})}]' \text{vec}[f(\hat{\lambda})] d\lambda = 2\pi \int_{2\pi} \textbf{tr}\Big( [\overline{f(\hat{\lambda})}]' [f(\hat{\lambda})] \Big) d\lambda = 2\pi \textbf{tr}\left( \int_{2\pi} \overline{[f(\hat{\lambda})]'} [f(\hat{\lambda})] d\lambda \right)$$

By the definition of an appropriate kernel estimator of the cross-spectrum,

$$\hat{f}(\lambda) = \frac{1}{2\pi} \sum_{j=1}^{T-1} (\omega(j))^{1/2} \widehat{\Gamma}_{XZ}(j) e^{-ij\lambda}$$

with $i$ being the imaginary unit, and by virtue of Parseval's theorem, we rewrite the test statistic as follows:

$$\begin{aligned} 2\pi \textbf{tr}\left( \int_{2\pi} \overline{[f(\hat{\lambda})]'} [f(\hat{\lambda})] d\lambda \right) &= 2\pi \textbf{tr}\left( 2\pi \sum_{j=1}^{\infty} \frac{1}{4\pi^2} \omega(j) \left( \overline{\widehat{\Gamma}_{XZ}(j)'} \right) \left( \widehat{\Gamma}_{XZ}(j) \right) \right) \\ &= \sum_{j=1}^{\infty} \omega(j) \textbf{tr}\left( \widehat{\Gamma}_{XZ}(j)' \widehat{\Gamma}_{XZ}(j) \right) = \sum_{j=1}^{T-1} \omega(j) Q(j) \end{aligned}$$

where the last equality follows from the trace property. □

[15] Namely: the sum of the conjugate is the conjugate of the sum, the conjugate of the transpose is the transpose of the conjugate, the conjugate of a real matrix is the real matrix itself.

**Lemma A.2.**

$$\mathcal{T}_\omega = \sum_{j=1}^{T-1} \omega(j) Q(j) = \mathcal{T}_{1\omega} + \mathcal{T}_{2\omega}$$
$$= \sum_{j=1}^{T-1} \omega(j) \left( \frac{1}{T^2} \sum_{t=j+1}^{T} ||X_t||^2 ||Z_{t-j}||^2 + \frac{1}{T^2} \sum_{s,t=j+1, s\neq t}^{T} < X_t, X_s >< Z_{t-j}, Z_{s-j} > \right)$$

*Proof.* It suffices to show that:

$$\begin{aligned} Q(j) &= \text{Tr}\left[\widehat{\Gamma}_{XZ}(j)'\widehat{\Gamma}_{XZ}(j)\right] \\ &= \frac{1}{T^2}\text{Tr}\left[\left(\sum_{t=j+1}^{T} Z_{t-j}X_t'\right)\left(\sum_{t=j+1}^{T} X_t Z_{t-j}'\right)\right] \\ &= \frac{1}{T^2}\sum_{s,t=j+1}^{T} \text{Tr}\left[Z_{t-j}X_t'X_sZ_{s-j}'\right] \\ &= \frac{1}{T^2}\sum_{s,t=j+1}^{T} < X_t, X_s >< Z_{t-j}, Z_{s-j} > \end{aligned}$$

□

**Lemma A.3.** *Define:* $\widehat{\Gamma}_{\hat{X}\hat{Z}}(j) = \frac{1}{T}\sum_{t=j+1}^{T} \widehat{X}_t \widehat{Z}_{t-j}'$ *(see Appendix B.3). We have:*

$$\widehat{Q}(j) = vec\left(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j)\right)'\left[\left(\widehat{\Gamma}_Z\right)^{-1} \otimes \left(\widehat{\Gamma}_X\right)^{-1}\right] vec\left(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j)\right)$$

*Proof.* We have:

$$\begin{aligned} \widehat{Q}(j) &= \text{vec}\left(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j)\right)'\left[\left(\widehat{\Gamma}_Z\right)^{-1} \otimes \left(\widehat{\Gamma}_X\right)^{-1}\right] \text{vec}\left(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j)\right) \\ &= \text{vec}\left[(\widehat{\Gamma}_X)^{-1/2}\widehat{\Gamma}_{\hat{X}\hat{Z}}(j)(\widehat{\Gamma}_Z^{-1/2})'\right]' \text{vec}\left[(\widehat{\Gamma}_X)^{-1/2}\widehat{\Gamma}_{\hat{X}\hat{Z}}(j)(\widehat{\Gamma}_Z^{-1/2})'\right] \\ &= \left|\left|\text{vec}\left[(\widehat{\Gamma}_X)^{-1/2}\widehat{\Gamma}_{\hat{X}\hat{Z}}(j)(\widehat{\Gamma}_Z^{-1/2})'\right]\right|\right|^2 = \left|\left|\text{vec}\left[\widehat{\Gamma}_{UV}(j)\right]\right|\right|^2 \end{aligned}$$

where the last equality is due to the definition in eq.(9). Clearly, we see that if the estimated innovations are the actual population-standardized innovation, the statistic is equal to the one defined in eq.(3). □

## A.2 Proof of Proposition 1

To show that the variance of test statistic, $\mathcal{T}_\omega$, depends on the inverse causality in the second moments, it is sufficient to study the first two moments of the two components of the quadratic forms, $Q(j)$, as described in Eq.(4).
By the conditional homoskedasticity of $\{X_t\}$ and the law of iterated expectations, we have that the first two moments of the first component, i.e. the squared products, $||X_t||^2||Z_{t-j}||^2$, are:

$$\mathbb{E}[||X_t||^2||Z_{t-j}||^2] = d_1 d_2, \quad \mathbb{E}[(||X_t||^2||Z_{t-j}||^2)^2] = \mathbb{E}[||X_t||^4]\mathbb{E}[||Z_t||^4] = \kappa_1\kappa_2$$

which means that, for a fixed $j \geq 1$, the squared products that characterizes the first term of the test statistic, $\mathcal{T}_{1\omega}$, are not influenced by the dependence running from past $X$ to present $Z$ (inverse causality).
To study the moments of the cross-products, without loss of generality, assume $t > s$.
We have:

$$\mathbb{E}[(< X_t, X_s >< Z_{t-j}, Z_{s-j} >)] = \mathbb{E}[X_t']\mathbb{E}[X_s < Z_{t-j}, Z_{s-j} >)] = 0$$
$$\mathbb{E}[(< X_t, X_s >< Z_{t-j}, Z_{s-j} >)^2] = d_1\mathbb{E}[||X_s||^2 < Z_{t-j}, Z_{s-j} >^2]$$

where the first equality is due the null in Eq.(1), and the third equality is because of the conditional homoskedasticity of the process $X$.
Without additional assumptions on the dependence structure, we can apply the law of iterated expectations only when $s > t - j$:

$$\mathbb{E}[||X_s||^2 < Z_{t-j}, Z_{s-j} >^2] = \mathbb{E}[||X_s||^2]\mathbb{E}[< Z_{t-j}, Z_{s-j} >^2] = d_1(d_2)^2, \quad s > t - j$$

which implies that, for a fixed $j \geq 1$, the cross-products that characterizes the second term of the test statistic, $\mathcal{T}_{2\omega}$, are independent with respect to the inverse causality.
Similarly, if the two processes are independent, then regardless of the time indexes:

$$\mathbb{E}[||X_s||^2 < Z_{t-j}, Z_{s-j} >^2] = \mathbb{E}[||X_s||^2]\mathbb{E}[||Z_{t-j}||^2]\mathbb{E}[||Z_{s-j}||^2] = d_1(d_2)^2$$

which is aligned to the case when $s > t - j$.

## A.3 Inverse Causality in Conditional Heteroskedastic DGPs

**Lemma A.4.** *Consider three measurable real functions,* $f(\cdot) : \mathbb{R} \to \mathbb{R}$*,* $g(\cdot) : \mathbb{R} \to \mathbb{R}$*,* $h(\cdot) : \mathbb{R} \to \mathbb{R}$*, that are non-trivial (i.e.,* $f(x) = 0, \forall x \in \mathbb{R}$*).*

*a) Let* $\{X_t, Z_t\}$ *be a bivariate mean-zero and marginally i.i.d. process, with:*

$$Z_t^2 = f(X_{t-1}) + u_t, \quad u_t \sim i.i.d.(1, \sigma_u^2), \quad X_t \perp\!\!\!\perp u_s, \forall s, t$$

*Denote:* $\sigma_z^2 = Var[Z_t]$ *and* $\varphi = (\mathbb{E}[X_t^2 f(X_t)] + 1)/\sigma_z^2$*. We have:*

$$Var[\mathcal{T}_{2\omega}] = \frac{1}{T^4} \sum_{j=1}^{T-2} \omega^2(j) \big( \Sigma_1(j) + \Delta_1(j) \big)$$

*with:* $\Sigma_1(j) = \sigma_z^4 (T-j)(T-j-1)$*,* $\quad \Delta_1(j) = \sigma_z^4 \sum_{s,t=j+1, s \neq t}^{T} (\varphi - 1)\, \mathbf{1}\{s = t - j\}$*.*
*An identical result holds for the scenarios where:*

$$Z_t = g(X_{t-1}) + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0, 1), \quad X_t \perp\!\!\!\perp \epsilon_s, \forall s, t$$

*when denoting:* $\varphi = (\mathbb{E}[X_t^2 g(X_t)^2] + 1)/\sigma_z^2$*.*

*b) For some* $|\alpha| \in (0, 1)$*, let* $\{X_t, Z_t\}$ *be a bivariate mean-zero process characterized by:*

$$Z_t^2 = \alpha Z_{t-1}^2 + h(X_{t-1}) + u_t, \quad X_t \sim i.i.d.(0, 1), \quad u_t \sim i.i.d.(1, \sigma_u^2), \quad X_t \perp\!\!\!\perp u_s, \forall s, t$$

*Denote:* $\varrho_h = \mathbb{E}[h(X_s) X_s^2]$*,* $\mu_e = \mathbb{E}[h(X_{t-1}) + u_t]$*,* $\mu_z = \mathbb{E}[Z_t^2]$*, and* $\sigma_z^2 = Var[Z_t^2]$*.*
*We have:*

$$Var[\mathcal{T}_{2\omega}] = \frac{1}{T^4} \sum_{j=1}^{T-2} \omega^2(j) \big( \Sigma_2(j) + \Delta_2(j) \big)$$

*with:*

$$\Sigma_2(j) = \sum_{\upsilon_1=1}^{\tau(j)} \left( \alpha^{|j-\upsilon_1|} \sigma_z^2 + \mu_z^2 \right) + \sum_{\upsilon_2=0}^{\Upsilon(j)} \left( \alpha^{j+\upsilon_2} \sigma_z^2 + \alpha^{\upsilon_2+1} \mu_z^2 + \sum_{l=0}^{\upsilon_2} \alpha^l \mu_z \right)$$

$$\Delta_2(j) = \sum_{\upsilon_2=0}^{\Upsilon(j)} \left( \mu_z \mu_e (1 + \alpha^{\upsilon_2} - \mathbb{1}[\upsilon_2 = 1]) - \sum_{l=0}^{\upsilon_2} \alpha^l \mu_z \right) + \sum_{\upsilon_2=0}^{\Upsilon(j)} \alpha^{\upsilon_2 - 1} \mu_z (1 + \varrho_h)$$

*where: i) $\tau(j)$ is the number of times that $s > t - j$ up to $T$, at a given $j \geq 1$, such that $s \neq t$, and $s, t = j + 1$; ii) $\Upsilon(j)$ is the number of times that $s \leq t - j$, with respect to the same conditions. In particular:*

$$\tau(j) = \begin{cases} (T^2 - 2j^2 - 3T + 4j)/2, & j < T/2 \\ (T - j)(T - j - 1), & j \geq T/2 \end{cases}$$
$$\Upsilon(j) = (T - j)(T - j - 1) - \tau(j)$$

Notice that, in all scenarios, the variance comprises two components: the sigmas, $\{\Sigma_i\}_{i=1,2}$, and the deltas, $\{\Delta_i\}_{i=1,2}$. The sigmas represent the variance as if there is no dependence from past $X$ to present $Z$. The deltas capture dependence from $\{X_{t-1}\}$ to $\{Z_t\}$ operating through the conditional variance (and/or conditional mean) of $Z$, despite weak exogeneity holding, since for all scenarios: $X_t \perp\!\!\!\perp \{X_s, Z_s\}_{s<t}$.

*Proof.* a) For the first part, consider the scenario where the DGPs are such that:

$$Z_t^2 = f(X_{t-1}) + u_t, \quad u_t \sim i.i.d.(1, \sigma_u^2), \quad X_t \perp\!\!\!\perp u_s, \forall s, t$$

where $f(\cdot) : \mathbb{R} \to \mathbb{R}$ is a measurable function.
Since $\{X_t\}$ is an i.i.d. process that is independent of $u_t$, we also have:

$$\mathbb{E}[X_t | \mathcal{I}(t-1)] = \mathbb{E}[X_t] = 0$$
$$\mathbb{E}[X_t^2 | \mathcal{I}(t-1)] = \mathbb{E}[X_t^2] = 1$$

By the same logic of Proposition 1, we have: (without loss of generality: $t > s$)

$$\mathbb{E}[X_t X_s Z_{t-j} Z_{s-j}] = 0$$
$$\mathbb{E}[X_t^2 X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2]$$
$$\mathbb{E}[(X_t, X_s Z_{t-j}, Z_{s-j})(X_{t+h}, X_{s+h} Z_{t-j+h}, Z_{s-j+h})] = 0, \quad \forall h \neq 0$$

Denote: $\sigma_f^2 = \mathbb{E}[f(X_t)]$, $\varrho_f = \mathbb{E}[X_t^2 f(X_t)]$, $\sigma_z^2 = \text{Var}(Z_t) = \sigma_f^2 + 1$.

Under the assumption about the DGP:

$$\mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \sigma_z^4, \quad \text{for } s > t-j$$
$$\mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2] = (\varrho_f + 1)\sigma_z^2, \quad \text{for } s = t-j$$
$$\mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \sigma_z^4, \quad \text{for } s < t-j$$

Thus:

$$\begin{aligned}
\text{Var}[\mathcal{T}_{2\omega}] &= \frac{1}{T^4}\sum_{j=1}^{T-1}\omega^2(j)\sum_{s,t=j+1,s\neq t}^{T}\mathbb{E}[X_t^2 X_s^2 Z_{t-j}^2 Z_{s-j}^2] \\
&= \frac{\sigma_z^4}{T^4}\sum_{j=1}^{T-1}\omega^2(j)\sum_{s,t=j+1,s\neq t}^{T} 1 + \left(\frac{\varrho_f + 1}{\sigma_z^2} - 1\right)\mathbb{1}\{s = t-j\}
\end{aligned}$$

Now I consider the scenario where the DGPs are such that:

$$Z_t = g(X_{t-1}) + \epsilon_t, \quad \epsilon_t \sim i.i.d.(0,1), \quad X_t \perp\!\!\!\perp \epsilon_s, \forall s,t$$

with $g(\cdot): \mathbb{R} \to \mathbb{R}$ being a measurable function such that: $\mathbb{E}[g(X_{t-1})] = 0$. Denote $\varrho_g = \mathbb{E}[X_t^2 g(X_t)^2]$. Under this class DGPs for $Z$, we have:

$$\text{Var}(Z_t) = \mathbb{E}[Z_t^2] = \sigma_z^2$$

Similar to the previous case:

$$\mathbb{E}[X_t|\mathcal{I}(t-1)] = \mathbb{E}[X_t] = 0$$
$$\mathbb{E}[X_t^2|\mathcal{I}(t-1)] = \mathbb{E}[X_t^2] = 1$$

Again, by direct application of Proposition 1, we have:
(without loss of generality: $t > s$)

$$\mathbb{E}[X_t X_s Z_{t-j} Z_{s-j}] = 0$$
$$\mathbb{E}[X_t^2 X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2]$$
$$\mathbb{E}[(X_t, X_s Z_{t-j}, Z_{s-j})(X_{t+h}, X_{s+h} Z_{t-j+h}, Z_{s-j+h})] = 0, \quad \forall h \neq 0$$

Thus:

$$\mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \sigma_z^4, \qquad \text{for } s > t-j$$
$$\mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2] = (\varrho_g + 1)\sigma_z^2, \qquad \text{for } s = t-j$$
$$\mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \sigma_z^4, \qquad \text{for } s < t-j$$

Finally:

$$\text{Var}[\mathcal{T}_{2\omega}] = \frac{\sigma_z^4}{T^4} \sum_{j=1}^{T-1} \omega^2(j) \sum_{s,t=j+1, s\neq t}^{T} 1 + \left( \frac{\varrho_g + 1}{\sigma_z^2} - 1 \right) \mathbb{1}\{s = t-j\}$$

b) Finally, consider the scenario where the DGPs are such that:

$$Z_t^2 = \alpha Z_{t-1}^2 + h(X_{t-1}) + u_t, \quad X_t \sim i.i.d.(0,1), \quad u_t \sim i.i.d.(1, \sigma_u^2), \quad X_t \perp\!\!\!\perp u_s, \forall s, t$$

for some measurable function $h(\cdot) : \mathbb{R} \to \mathbb{R}$ and $|\alpha| \in (0,1)$.
Denote: $\mu_h = \mathbb{E}[h(X_s)]$ and $\varrho_h = \mathbb{E}[h(X_s)X_s^2]$, $\mu_e = \mathbb{E}[h(X_{s-1}) + u_s] = \mu_h + 1$.
Denote $\sigma_e^2 = \mathbb{E}[(h(X_{t-1}) + u_t)^2]$, so that we finally denote:

$$\mu_z = \mu_e/(1-\alpha), \quad \sigma_z^2 = (\sigma_e^2 - \mu_e^2)/(1-\alpha^2)$$

Similar to the previous case, we have:

$$\mathbb{E}[X_t | \mathcal{I}(t-1)] = \mathbb{E}[X_t] = 0$$
$$\mathbb{E}[X_t^2 | \mathcal{I}(t-1)] = \mathbb{E}[X_t^2] = 1$$

By the same logic of Proposition 1, we have:

$$\mathbb{E}[X_t X_s Z_{t-j} Z_{s-j}] = 0$$
$$\mathbb{E}[X_t^2 X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2]$$
$$\mathbb{E}[(X_t, X_s Z_{t-j}, Z_{s-j})(X_{t+h}, X_{s+h} Z_{t-j+h}, Z_{s-j+h})] = 0, \quad \forall h \neq 0$$

Note that the process $\{Z_t^2\}$ is a causal AR(1) process with i.i.d. noise, since we have:

$$Z_{t-1} \perp\!\!\!\perp X_{t-1}, \quad u_t \perp\!\!\!\perp X_{t-1}$$

Notice that, for $h \geq 0$:

$$\mathbb{E}[Z_t^2 Z_{t-h}^2] = \alpha^h \sigma_z^2 + \mu_z^2 = \alpha^h \frac{\sigma_e^2 - \mu_e^2}{1-\alpha^2} + \left(\frac{\mu_e}{1-\alpha}\right)^2 = \alpha^h \frac{\sigma_e^2}{1-\alpha^2} + \mu_e^2 \left(\frac{1}{(1-\alpha)^2} - \alpha^h\right)$$

Consider the two scenarios:

i) When $s > t-j$, we have:

$$\mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \mathbb{E}[X_s^2]\mathbb{E}[Z_{t-j}^2 Z_{s-j}^2] = \mathbb{E}[Z_{t-j}^2 Z_{s-j}^2] = \alpha^{|j-\upsilon_1|}\sigma_z^2 + \mu_z^2$$

where $\upsilon_1 = s - t + j > 0$.
To see the last equality, note that we have the following:

- For $s = t - j + 1$, that is $\upsilon_1 = 1$:

$$\mathbb{E}[Z_{t-j}^2 Z_{s-j}^2] = \mathbb{E}[Z_{s-1}^2 Z_{s-j}^2] = \alpha^{|j-1|}\sigma_z^2 + \mu_z^2$$

since:

$$\mathbb{E}[Z_{s-1}^2 Z_{s-j}^2] = \sigma_z^2 + \mu_z^2, \qquad \text{for } j = 1$$
$$\dots$$
$$\mathbb{E}[Z_{s-1}^2 Z_{s-j}^2] = \mathbb{E}[Z_{s-1}^2 Z_{m+1}^2] = \alpha^m \sigma_z^2 + \mu_z^2, \quad \text{for } j = m+1$$

- For $s = t - j + 2$, that is $\upsilon_1 = 2$:

$$\mathbb{E}[Z_{t-j}^2 Z_{s-j}^2] = \mathbb{E}[Z_{s-2}^2 Z_{s-j}^2] = \alpha^{|j-2|}\sigma_z^2 + \mu_z^2$$

since:

$$
\begin{aligned}
\mathbb{E}[Z_{s-2}^2 Z_{s-j}^2] &= \alpha\sigma_z^2 + \mu_z^2, \qquad \text{for } j = 1 \\
\mathbb{E}[Z_{s-2}^2 Z_{s-j}^2] &= \sigma_z^2 + \mu_z^2, \qquad \text{for } j = 2 \\
&\cdots \\
\mathbb{E}[Z_{s-1}^2 Z_{s-j}^2] &= \mathbb{E}[Z_{s-1}^2 Z_{m+1}^2] = \alpha^m\sigma_z^2 + \mu_z^2, \quad \text{for } j = m+2
\end{aligned}
$$

Summing up the terms, we have:

$$
\sum_{s>t-j, s\neq t}^{T} \mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \sum_{v_1=1}^{\tau(j)} \left(\alpha^{|j-v_1|}\sigma_z^2 + \mu_Z^2\right)
$$

where $\tau(j)$ is the number of times that $s > t - j$ up to $T$, at a given $j > 0$, such that: $s \neq t$, $s, t = j + 1$. In particular, we have:

$$
\begin{aligned}
\tau(j) = &\frac{(T-j)(T-j-1)}{2} \\
&+ \mathbb{1}(2j - T < 0) \cdot \big((T-2j)(j-1) + j(j-1)/2\big) \\
&+ \mathbb{1}(2j - T \geq 0) \cdot (T-j)(T-j-1)/2
\end{aligned}
$$

To see this, note that the summation is taken over two indices, $s$ and $t$. The total number of (admissible) pairs $(s, t)$ can be decomposed into lower and upper triangular parts of the following table. The entries take value one if $s > t - j$. Because of the additional condition that $s \neq t$, the diagonal entries are zero.

| | $t = j+1$ | $t = j+2$ | $t = j+3$ | $t = j+4$ | $j+4 < t < 2j+2$ | $t = 2j+2$ | $t = 2j+3$ | $t = 2j+4$ |
|---|---|---|---|---|---|---|---|---|
| $s = j+1$ | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| $s = j+2$ | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 |
| $s = j+3$ | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| $s = j+4$ | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 |
| $\cdots$ | . | . | . | . | . | . | . | . |
| $s = 2j+2$ | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 |
| $s = 2j+3$ | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| $s = 2j+4$ | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |

i. The first term, $\dfrac{(T-j)(T-j-1)}{2}$, counts the number of pairs in the lower triangular region $s > t$ (bottom left), where the condition $s > t - j$ is always satisfied for any $j > 0$.

ii. The second term, $(T-2j)(j-1) + j(j-1)/2$, counts the number of pairs in the upper triangular region when $2j < T$. The constraint $s > t-j$ zeroes pairs in the top right corner of the table. Specifically, for $t > 2j$, only $j-1$ values of $s$ satisfy the inequality, while for $t \leq 2j$ we have a triangular block of size $\frac{j(j-1)}{2}$.

iii. The third term, $\dfrac{(T-j)(T-j-1)}{2}$, counts the number of pairs in the upper triangular region when $2j \geq T$. In this latter case, as the condition $s > t - j$ does not exclude any pair, the number of upper triangular pairs coincides with the lower triangular ones.

- When $s \leq t - j$, define $\upsilon_2 = t - s - j \geq 0$. In this case,

$$\begin{aligned}\mathbb{E}[X_s^2 Z_{t-j}^2 Z_{s-j}^2] = \alpha^{\upsilon_2+1}\left(\alpha^{j-1}\sigma_z^2 + \mu_z^2\right) + \mu_z\mu_e(1+\alpha^{\upsilon_2}) - \mu_z\mu_e\,\mathbb{1}[\upsilon_2 = 1] \\ + \alpha^{\upsilon_2-1}\mu_z(1+\varrho_h)\,\mathbb{1}[\upsilon_2 > 0],\end{aligned}$$

where $\mu_e = \mathbb{E}[h(X_{t-1}) + u_t] = \mu_h + 1$. To see the equality, note that we have the following:

  - For $s = t - j$, that is $\upsilon_2 = 0$:

$$\begin{aligned}\mathbb{E}[Z_s^2 X_s^2 Z_{s-j}^2] &= \alpha\mathbb{E}[Z_{s-1}^2 X_s^2 Z_{s-j}^2] + \mathbb{E}[u_s X_s^2 Z_{s-j}^2] + \mathbb{E}[h(X_{s-1}) X_s^2 Z_{s-j}^2] \\ &= \alpha\mathbb{E}[Z_{s-1}^2 Z_{s-j}^2] + \mathbb{E}[Z_{s-j}^2] + \mathbb{E}[h(X_{s-1})]\mathbb{E}[X_s^2]\mathbb{E}[Z_{s-j}^2] \\ &= \alpha\left(\alpha^{j-1}\sigma_z^2 + \mu_z^2\right) + \mu_z(1+\mu_h).\end{aligned}$$

  - For $s = t - j - 1$, that is $\upsilon_2 = 1$:

$$\begin{aligned}\mathbb{E}[Z_{s+1}^2 X_s^2 Z_{s-j}^2] &= \alpha\mathbb{E}[Z_s^2 X_s^2 Z_{s-j}^2] + \mathbb{E}[u_{s+1} X_s^2 Z_{s-j}^2] + \mathbb{E}[h(X_s) X_s^2 Z_{s-j}^2] \\ &= \alpha^2\left(\alpha^{j-1}\sigma_z^2 + \mu_z^2\right) + \alpha\mu_z(1+\mu_h) + \mu_z(1+\varrho_h).\end{aligned}$$

– For $s = t - j - 2$, that is $\upsilon_2 = 2$:

$$\begin{aligned}\mathbb{E}[Z_{s+2}^2 X_s^2 Z_{s-j}^2] &= \alpha\mathbb{E}[Z_{s+1}^2 X_s^2 Z_{s-j}^2] + \mathbb{E}[u_{s+2} X_s^2 Z_{s-j}^2] + \mathbb{E}[h(X_{s+1}) X_s^2 Z_{s-j}^2] \\ &= \alpha^3 \left(\alpha^{j-1}\sigma_z^2 + \mu_z^2\right) + \alpha\mu_z(1+\varrho_h) + \mu_z(1+\mu_h)(1+\alpha^2),\end{aligned}$$

and the generalization follows by induction.

Thus, reconciling both cases:

$$\begin{aligned}T^4\text{Var}(\mathcal{T}_{2\omega}) &= \sum_{j=1}^{T-2}\omega^2(j)\left(\sum_{\upsilon_1=1}^{\tau(j)}\left(\alpha^{|j-\upsilon_1|}\sigma_z^2 + \mu_z^2\right) + \sum_{\upsilon_2=0}^{(T-j)(T-j-1)-\tau(j)}\left(\alpha^{\upsilon_2+j}\sigma_z^2 + \alpha^{\upsilon_2+1}\mu_z^2\right)\right) \\ &+ \sum_{j=1}^{T-2}\omega^2(j)\sum_{\upsilon_2=0}^{(T-j)(T-j-1)-\tau(j)}\Big(\mu_z\mu_e(1+\alpha^{\upsilon_2}) - \mu_z\mu_e\mathbb{1}[\upsilon_2=1]\Big) \\ &+ \sum_{j=1}^{T-2}\omega^2(j)\sum_{\upsilon_2=1}^{(T-j)(T-j-1)-\tau(j)}\alpha^{\upsilon_2-1}\mu_z(1+\varrho_h).\end{aligned}$$

When $h(\cdot) = 0$, meaning $Z_t^2 = \alpha Z_{t-1}^2 + u_t$, notice that:

$$\begin{aligned}T^4\text{Var}(\mathcal{T}_{2\omega})|_{h(\cdot)=0} &= \sum_{j=1}^{T-2}\omega^2(j)\left(\sum_{\upsilon_1=1}^{\tau(j)}\left(\alpha^{|j-\upsilon_1|}\sigma_z^2 + \mu_z^2\right) + \sum_{\upsilon_2=0}^{(T-j)(T-j-1)-\tau(j)}\left(\alpha^{\upsilon_2+j}\sigma_z^2 + \alpha^{\upsilon_2+1}\mu_z^2\right) \right. \\ &+ \sum_{j=1}^{T-2}\omega^2(j)\sum_{\upsilon_2=0}^{(T-j)(T-j-1)-\tau(j)}\left(\sum_{l=0}^{\upsilon_2}\alpha^l\mu_z\right) \\ &= \sum_{j=1}^{T-2}\omega^2(j)\Sigma_2(j)\end{aligned}$$

so that we finally define the difference as:

$$T^4\text{Var}(\mathcal{T}_{2\omega}) - \sum_{j=1}^{T-2}\omega^2(j)\Sigma_2(j) = \sum_{j=1}^{T-2}\omega^2(j)\Delta_2(j)$$

□

# B Appendix B: Proof of Proposition 2, Theorem 1 to 3

## B.1 Proof of Proposition 2

i) The proof consists of three parts:

A. From its definition in eq.(4):

$$\mathbb{E}[T \cdot \mathcal{T}_{1\omega}] = \frac{1}{T}\sum_{j=1}^{T-1}\omega(j)\sum_{t=j+1}^{T}\mathbb{E}[||X_t||^2||Z_{t-j}||^2]$$

The expectations are such that, by law of iterated expectations:

$$\mathbb{E}[||X_t||^2||Z_{t-j}||^2] = \mathbb{E}[\mathbb{E}[||X_t||^2|\mathcal{I}(t-1)]||Z_{t-j}||^2] = d_1\mathbb{E}[||Z_{t-j}||^2] = d_1 d_2$$

where the second equality is because of the assumption of conditional homoskedasticity, $\mathbb{E}[X_t X_t'|\mathcal{I}(t-1)] = \mathbb{E}[X_t X_t']$, and the processes being standardized. Substituting above:

$$\mathbb{E}[T \cdot \mathcal{T}_{1\omega}] = \frac{1}{T}\sum_{j=1}^{T-1}\omega(j)(T-j)d_1 d_2 = \mu_{\omega,T}$$

B. By definition:

$$\text{Var}[T \cdot \mathcal{T}_{1\omega}] = \mathbb{E}[(T \cdot \mathcal{T}_{1\omega} - \mu_{\omega,T})^2]$$

For simplicity, denote:

$$T \cdot \mathcal{T}_{1\omega} - \mu_{\omega,T} = \frac{1}{T}\sum_{j=1}^{T-1}\omega(j)\sum_{t=j+1}^{T}\left(||X_t||^2||Z_{t-j}||^2 - d_1 d_2\right) = \frac{1}{T}\sum_{j=1}^{T-1}\omega(j)\sum_{t=j+1}^{T}\Upsilon_{t,j}^{(1)}$$

Define: $Y_{1,t} = ||X_t||^2 - d_1$, and $Y_{2,t} = ||Z_t||^2 - d_2$, so that we write:

$$\Upsilon_{t,j}^{(1)} = Y_{1,t}||Z_{t-j}||^2 + d_1 Y_{2,t-j}$$

Hence, we have:

$$\Big\| \sum_{t=j+1}^{T} \Upsilon_{t,j}^{(1)} \Big\|_2^2 \leq 2\Big\| \sum_{t=j+1}^{T} Y_{1,t}||Z_{t-j}||^2 \Big\|_2^2 + 2d_1^2 \Big\| \sum_{t=j+1}^{T} Y_{2,t-j} \Big\|_2^2.$$

Let us consider the two summations:

- Regarding the first term: $\Big\| \sum_{t=j+1}^{T} Y_{1,t}||Z_{t-j}||^2 \Big\|_2^2$.
  Under the assumption of conditional homoskedasticity, we have that:

  $$\mathbb{E}[Y_{1,t}||Z_{t-j}||^2|\mathcal{I}(t-1)] = 0$$

  so that $\{Y_{1,t}||Z_{t-j}||^2, \mathcal{I}(t-1)\}$ constitutes a martingale difference sequence. This means we can write:

  $$\Big\| \sum_{t=j+1}^{T} Y_{1,t}||Z_{t-j}||^2 \Big\|_2^2 = \sum_{t=j+1}^{T} \mathbb{E}[||Y_{1,t}||^2||Z_{t-j}||^4]$$

  By the assumptions on the conditional moments:

  $$\mathbb{E}[||X_t||^4||Z_{t-j}||^4] = \mathbb{E}[\mathbb{E}[X_t'X_tX_t'X_t|\mathcal{I}(t-1)]||Z_{t-j}||^4] = \mathbb{E}[||X_t||^4]\mathbb{E}[||Z_{t-j}||^4]$$

  which means that, together with the finiteness of fourth moments, we have: $\sup_t \mathbb{E}[||Y_{1,t}||^2||Z_{t-j}||^4] < \infty$.
  Hence, we conclude:

  $$\Big\| \sum_{t=j+1}^{T} Y_{1,t}||Z_{t-j}||^2 \Big\|_2^2 = O(T)$$

- Regarding the second term: $\Big\| \sum_{t=j+1}^{T} Y_{2,t} \Big\|_2^2$.
  By stationarity, Var $\left(\sum_{t=j+1}^{T} Y_{2,t-j}\right)$ = Var $\left(\sum_{t=1}^{T-j} Y_{2,t}\right)$.

Let us denote $\gamma_{Y_2}(h) := \text{Cov}(Y_2, Y_{2+h})$, then:

$$\begin{aligned}\text{Var}\left(\sum_{t=1}^{T-j} Y_{2,t}\right) &= (T-j)\gamma_{Y_2}(0) + 2\sum_{h=1}^{T-j-1}(T-j-h)\gamma_{Y_2}(h) \\ &\leq (T-j)\gamma_{Y_2}(0) + 2(T-j)\sum_{h=1}^{\infty}|\gamma_{Y_2}(h)|\end{aligned}$$

Under the assumption $|\text{Cov}[||Z_1||^2, ||Z_{1+h}||^2]| = O(h^{-1-\epsilon})$, we have a standard argument for the absolute summability of the covariances, which in turn implies:

$$\Big\| \sum_{t=j+1}^{T} Y_{2,t-j} \Big\|_2^2 = O(T)$$

Thus, we have the following:

$$\Big\| \sum_{t=j+1}^{T} \Upsilon_{t,j}^{(1)} \Big\|_2 = O(T^{1/2})$$

so that the variance can be bounded as follows:

$$\begin{aligned}\text{Var}[T\mathcal{T}_{1\omega}] &= \frac{1}{T^2}\mathbb{E}\left[\left(\sum_{j=1}^{T-1}\omega(j)\sum_{t=j+1}^{T}\Upsilon_{t,j}^{(1)}\right)^2\right] = \frac{1}{T^2}\left\|\sum_{j=1}^{T-1}\omega(j)\sum_{t=j+1}^{T}\Upsilon_{t,j}^{(1)}\right\|_2^2 \\ &\leq \frac{1}{T^2}\left(\sum_{j=1}^{T-1}\omega(j)\left\|\sum_{t=j+1}^{T}\Upsilon_{t,j}^{(1)}\right\|_2\right)^2\end{aligned}$$

by virtue of Minkowski inequality. By the previous results:

$$\text{Var}[T\mathcal{T}_{1\omega}] \leq \frac{\Delta}{T}\left(\sum_{j=1}^{T-1}\omega(j)\right)^2 = \frac{\Delta}{T}M^2\left(M^{-1}\sum_{j=1}^{T-1}\omega(j)\right)^2 = O(M^2/T)$$

for finite $\Delta > 0$, where the last equality comes by realizing that, under Assumption 1, the following convergence holds asymptotically:

$$M^{-1}\sum_{j=1}^{T-1}\omega(j)\to\int_0^\infty k^2(z)dz<\infty$$

C. By the same reason as above, under Assumption 1, the following convergence holds asymptotically:

$$M^{-1}D_{\omega,T}\to\frac{1}{2}\int_0^\infty k^4(z)dz<\infty$$

so that:

$$D_{\omega,T}=O(M)$$

Given the finiteness of the fourth moments of $Z$, and the result in Lemma A.4, we also have that:

$$\begin{aligned}D_{\omega,T}^{(Hete)}&=\frac{4d_1^2}{T^2}\sum_{t=2}^{T}\sum_{s=2}^{t-1}\sum_{j\in\mathfrak{J}_{ts}}\sum_{\ell\in\mathfrak{J}_{ts}}\omega(j)\omega(\ell)\,\gamma_{t,s}(j,\ell)\\&=\frac{4d_1^2}{T^2}\sum_{t=2}^{T}\sum_{s=2}^{t-1}\sum_{j=\ell,j,\ell\in\mathfrak{J}_{ts}}\omega^2(j)\,\gamma_{t,s}(j,j)+\frac{4d_1^2}{T^2}\sum_{t=2}^{T}\sum_{s=2}^{t-1}\sum_{j\neq\ell,j,\ell\in\mathfrak{J}_{ts}}\omega(j)\omega(\ell)\,\gamma_{t,s}(j,\ell)\\&\asymp O\left(\frac{1}{T^2}\sum_{j=1}^{T-1}\omega^2(j)\sum_{l=1}^{\tau(j)}\right)=O(M)\end{aligned}$$

where $a\asymp b$ stands for the joint condition: $a=O(b)$ and $b=O(a)$. Together with the previous findings, we have:

$$\mathbb{E}\left[\frac{T\cdot\mathcal{T}_{1\omega}-\mu_{\omega,T}}{\sqrt{D_{\omega,T}^{(Hete)}}}\right]=0,\quad \text{Var}\left[\frac{T\cdot\mathcal{T}_{1\omega}-\mu_{\omega,T}}{\sqrt{D_{\omega,T}^{(Hete)}}}\right]=O(M/T)$$

so that, under the asymptotics of Assumption 1, when $M/T\to 0$ as $T\to\infty$, we have $\ell_2$ (or MSE) convergence to zero.

ii) The proof consists of two parts:

A. From its definition in eq.5):

$$\mathbb{E}[T \cdot \mathcal{T}^c_{2\omega}] = \frac{2}{T}\sum_{j=1}^{T-2}\omega(j)\sum_{t=j+1}^{T}\sum_{s=t-j+1}^{t-1}\mathbb{E}[< X_t, X_s >< Z_{t-j}, Z_{s-j} >]$$
$$= \frac{2}{T}\sum_{t=2}^{T}\sum_{s=2}^{t-1}\sum_{j=t-s+1}^{s-1}\omega(j)\mathbb{E}[< X_t, X_s >< Z_{t-j}, Z_{s-j} >]$$

Under the null hypothesis in eq.(1) and by law of iterated expectations:

$$\mathbb{E}[< X_t, X_s >< Z_{t-j}, Z_{s-j} >] = \mathbb{E}[\mathbb{E}[X_t'|\mathcal{I}(t-1)]X_s < Z_{t-j}, Z_{s-j} >] = 0$$

as the indexes are such that $t > s$. Thus:

$$\mathbb{E}[T \cdot \mathcal{T}^c_{2\omega}] = 0$$

B. Define:

$$J_t := \sum_{s=2}^{t-1}\sum_{j\in\mathfrak{I}_{ts}}\omega(j)\,\langle X_t, X_s\rangle\langle Z_{t-j}, Z_{s-j}\rangle, \qquad T\cdot\mathcal{T}^c_{2\omega} = \frac{2}{T}\sum_{t=2}^{T}J_t.$$

The process $\{(J_t, \mathcal{I}(t-1)), t \in \mathbb{Z}^+\}$ constitutes a martingale difference sequence, since: 1) we have $\mathbb{E}[J_t|\mathcal{I}(t-1)] = 0$ under the null hypothesis of eq.(1); 2) $\mathbb{E}[|J_t|] < \infty$ by the finiteness of the moments.
Hence:

$$\text{Var}[T \cdot \mathcal{T}^c_{2\omega}] = \frac{4}{T^2}\sum_{t=2}^{T}\mathbb{E}[J_t^2].$$

where:

$$J_t^2 = \sum_{s=2}^{t-1}\sum_{r=2}^{t-1}\sum_{j\in\mathfrak{I}_{ts}}\sum_{\ell\in\mathfrak{I}_{tr}}\omega(j)\omega(\ell)\langle X_t, X_s\rangle\langle X_t, X_r\rangle\langle Z_{t-j}, Z_{s-j}\rangle\langle Z_{t-\ell}, Z_{r-\ell}\rangle.$$

Under the null, only the cross-terms such that: $r = s$, will be nonzero. Thus:

$$J_t^2 = \sum_{s=2}^{t-1} \sum_{j \in \mathfrak{J}_{ts}} \sum_{\ell \in \mathfrak{J}_{tr}} \omega(j)\omega(\ell)\langle X_t, X_s\rangle^2 \langle Z_{t-j}, Z_{s-j}\rangle \langle Z_{t-\ell}, Z_{s-\ell}\rangle.$$

By virtue of the law of iterated expectations and the conditional homoskedasticity of $X$, then we write:

$$\begin{aligned}
&\mathbb{E}\left[\langle X_t, X_s\rangle^2 \langle Z_{t-j}, Z_{s-j}\rangle \langle Z_{t-\ell}, Z_{s-\ell}\rangle\right] \\
&= \mathbb{E}[\mathbb{E}[X_s' X_t X_t' X_s \langle Z_{t-j}, Z_{s-j}\rangle \langle Z_{t-\ell}, Z_{s-\ell}\rangle | \mathcal{I}(t-1)]] \\
&= \mathbb{E}[X_s' \mathbb{E}[X_t X_t' | \mathcal{I}(t-1)] X_s \langle Z_{t-j}, Z_{s-j}\rangle \langle Z_{t-\ell}, Z_{s-\ell}\rangle] \\
&= d_1 \mathbb{E}[||X_s||^2 \langle Z_{t-j}, Z_{s-j}\rangle \langle Z_{t-\ell}, Z_{s-\ell}\rangle] \\
&= d_1^2 \mathbb{E}\left[\langle Z_{t-j}, Z_{s-j}\rangle \langle Z_{t-\ell}, Z_{s-\ell}\rangle\right] = d_1^2 \gamma_{t,s}(j,\ell)
\end{aligned}$$

Meaning that:

$$\begin{aligned}
\text{Var}[T \cdot \mathcal{T}_{2\omega}^c] &= \frac{4}{T^2} \sum_{t=2}^{T} \sum_{s=2}^{t-1} \sum_{j \in \mathfrak{J}_{ts}} \sum_{\ell \in \mathfrak{J}_{tr}} \omega(j)\omega(\ell) d_1^2 \mathbb{E}\left[\langle Z_{t-j}, Z_{s-j}\rangle \langle Z_{t-\ell}, Z_{s-\ell}\rangle\right] \\
&= D_{\omega,T}^{(Hete)}
\end{aligned}$$

with $\mathfrak{J}_{ts} := \{j \in \mathbb{Z} : \; t-s+1 \leq j \leq s-1\}$, and $\gamma_{t,s}(j,l) = \mathbb{E}[\langle Z_{t-j}, Z_{s-j}\rangle \; \langle Z_{t-\ell}, Z_{s-\ell}\rangle]$, when these last moments exist.
Notice that we rewrite the summation in $j-\ell$ terms:

$$\begin{aligned}
D_{\omega,T}^{(Hete)} &= \frac{4d_1^2}{T^2} \sum_{t=2}^{T} \sum_{s=2}^{t-1} \sum_{j \in \mathfrak{J}_{ts}} \sum_{\ell \in \mathfrak{J}_{ts}} \omega(j)\omega(\ell)\; \gamma_{t,s}(j,\ell) \\
&= \frac{2d_1^2}{T^2} \sum_{j=1}^{T-2} \sum_{\ell=1}^{T-2} \omega(j)\omega(\ell) \sum_{s=\max\{j,\ell\}+1}^{T-1} \; \sum_{t=s+1}^{\min\{T,\, s+\min(j,\ell)-1\}} \gamma_{t,s}(j,\ell)
\end{aligned}$$

## B.2 Proof of Theorem 1 (Size)

The proof of asymptotic normality of the test statistic, $\mathcal{T}_\omega^c$, translates into proving the asymptotic normality of the dominating term $T \cdot \mathcal{T}_{2\omega}^c$, as because of application of Proposition 2 (together with Lemma B.5) we have asymptotically:

$$\frac{T\mathcal{T}_{1\omega} - \mu_{\omega,T}}{\sqrt{D_{\omega,T}^{(Hete)}}} = o_p(1)$$

when $M/T \to 0$ as $M, T \to \infty$. Note that, if we have $M^2/T \to 0$, the asymptotical negligibility holds as well.
For simplicity, define:

$$T \cdot \mathcal{T}_{2\omega}^c = \frac{2}{T}\sum_{t=2}^{T} J_t$$

$$J_t = \sum_{s=2}^{t-1}\sum_{j=t-s+1}^{s-1} \omega(j)\langle X_t, X_s\rangle\langle Z_{t-j}, Z_{s-j}\rangle = \sum_{s=2}^{t-1} X_t' X_s W_{ts}$$

such that: $W_{ts} = \sum_{j=t-s+1}^{s-1} \omega(j)\langle Z_{t-j}, Z_{s-j}\rangle$.
Note that, under Assumption 1, $\sum_{j=t-s+1}^{s-1} \omega(j) = O(M)$, as discussed in Proposition 2. As well, as previously discussed, the process $\{(J_t, \mathcal{I}(t-1)), t \in \mathbb{Z}^+\}$ constitutes a martingale difference sequence. Finally, to invoke Brown (1971)'s theorem, two conditions need to be verified.

i) The following Lindeberg condition needs to hold:

$$T^{-2}(D_{\omega,T}^{(Hete)})^{-1}\sum_{t=2}^{T}\mathbb{E}\Big[(J_t)^2 \cdot \mathbf{1}\big\{|J_t| > \epsilon(D_{\omega}^{(Hete)})^{1/2}\big\}\Big] \to 0$$

To prove it, it is sufficient to show that the Lyapunov condition holds:

$$T^{-4}(D_{\omega,T}^{(Hete)})^{-2}\sum_{t=2}^{T}\mathbb{E}[(J_t)^4] \to 0$$

For convenience, let us define:

$$S_t := \sum_{s=2}^{t-1} X_s W_{ts} = \sum_{s=2}^{t-1} U_{s,t}$$

By the conditional homokurtosis of $X$, we can write:

$$\mathbb{E}[J_t^4] = \mathbb{E}\big[\mathbb{E}[(X_t' S_t)^4 \mid \mathcal{I}(t-1)]\big] \leq \Delta \mathbb{E}\big[\|S_t\|^4\big].$$

for some $\Delta > 0$, due to the fact that $S_t \in \mathcal{I}(t-1)$.
Fix $t$ and define the filtration $\mathcal{G}_s := \sigma\{(X_u, Z_u) : u \leq s\}$.
Note that, due to the correction term: $W_{ts} \in G_{s-1}$. Thus, under $\mathcal{H}_0$, $\{X_s, \mathcal{G}_s\}$ is a martingale difference sequence, so $\mathbb{E}[X_s \mid \mathcal{G}_{s-1}] = 0$. Since $W_{ts} \in \mathcal{G}_{s-1}$, we obtain

$$\mathbb{E}[U_{s,t} \mid \mathcal{G}_{s-1}] = \mathbb{E}[X_s \mid \mathcal{G}_{s-1}]\, W_{ts} = 0.$$

Hence, for each fixed $t$, the sequence $\{(U_{s,t}, \mathcal{G}_s),\ s = 2, \ldots, t-1\}$ is a martingale difference sequence in $s$, implying $S_t$ is a martingale sum.
Next, by Theorem 2.11 of Hall and Heyde (2014), we apply the Burkholder-type inequality for martingales (for $p = 4$):

$$\mathbb{E}\|S_t\|^4 \leq \Delta \left\{ \mathbb{E}\left( \sum_{s=2}^{t-1} \mathbb{E}\big[||U_{s,t}||^2 | \mathcal{G}_{s-1}\big] \right)^2 + \sum_{s=2}^{t-1} \mathbb{E}\big[||U_{s,t}||^4\big] \right\}$$

for some finite $\Delta > 0$.

A. Regarding the second term: $\sum_{s=2}^{t-1} \mathbb{E}\big[||U_{s,t}||^4\big]$.
By law of iterated expectations and the conditional homokurtosis, we can write:

$$\mathbb{E}\big[||U_{s,t}||^4\big] = \mathbb{E}\big[||X_s||^4 ||W_{ts}||^4\big] \leq \Delta \mathbb{E}\big[||W_{ts}||^4\big]$$

for some finite $\Delta > 0$. By Cauchy-Schwarz, we have:

$$||W_{ts}|| \leq \sum_{j=t-s+1}^{s-1} |\omega(j)|\, \|Z_{t-j}\|\, \|Z_{s-j}\| \leq \Big( \sum_{j=t-s+1}^{s-1} |\omega(j)| \Big) \max_j \|Z_{t-j}\|\, \|Z_{s-j}\|$$

where, by Assumption 1 and the finiteness of moments, it follows that:

$$\sup_{t>s} \mathbb{E}||W_{ts}||^4 = O(M^4)$$

Hence, for some finite $\Delta > 0$:

$$\sum_{s=2}^{t-1} \mathbb{E}||U_{s,t}||^4 \leq \Delta \sum_{s=2}^{t-1} \mathbb{E}||W_{ts}||^4 = O(tM^4).$$

B. Regarding the first term: $\mathbb{E}\left(\sum_{s=2}^{t-1} \mathbb{E}\left[||U_{s,t}||^2|\mathcal{G}_{s-1}\right]\right)^2$.
Under the conditional homoskedasticity of $X$, we have:

$$\mathbb{E}[||U_{s,t}||^2|G_{s-1}] = \mathbb{E}[||X_s||^2|G_{s-1}]W_{ts}^2 = d_1 W_{ts}^2$$

Thus:

$$\mathbb{E}\left(\sum_{s=2}^{t-1} \mathbb{E}\left[||U_{s,t}||^2|\mathcal{G}_{s-1}\right]\right)^2 = d_1^2 \mathbb{E}\left(\sum_{s=2}^{t-1} W_{ts}^2\right)^2 \tag{11}$$

By Cauchy-Schwarz and the previous bounds:

$$\mathbb{E}\left(\sum_{s=2}^{t-1} W_{ts}^2\right)^2 \leq (t-2)\sum_{s=2}^{t-1} \mathbb{E}[W_{ts}^4] = O(t^2 M^4)$$

Combining the bounds, it gives:

$$\mathbb{E}[||S_t||^4] = O(t^2 M^4)$$

Plugging it back, we write:

$$\mathbb{E}[J_t^4] = O(t^2 M^4), \qquad \sum_{t=2}^{T} \mathbb{E}[J_t^4] = O(T^3 M^4)$$

Since $D_{\omega,T}^{(\text{Hete})} \asymp M$ (Assumption 1 and Proposition 2), we obtain:

$$T^{-4}\left(D_{\omega,T}^{(\text{Hete})}\right)^{-2} \sum_{t=2}^{T} \mathbb{E}[J_t^4] = O\left(T^{-4} M^{-2} \cdot T^3 M^4\right) = O\left(\frac{M^2}{T}\right)$$

This in turn verifies the Lyapunov (and Lindeberg) condition when $\frac{M^2}{T} \to 0$. To reach a sharper conclusion in terms of the ratio between $M$ and $T$, the alternative line is to appeal to Marcinkiewicz-Zygmund inequalities. Under the assumptions of the time series $\{\Lambda_{i,t}\}$:

$$|\text{Cov}[\Lambda_{1,t}, \Lambda_{1+i,t}]| = O(i^{-4/2}) = O(i^{-2}), \quad i \to +\infty$$

by application of Theorem 4.1 of Dedecker et al. (2007), we have that:

$$\mathbb{E}[(J_t)^4] = O(t^2 M^2)$$

Similar result can be obtained via Rosenthal type inequalities. For further details, please refer to Theorem 4.2 and Corollary 5.4 of Dedecker et al. (2007).

For this latter scenario, we have the sharper conclusion:

$$T^{-4}(D_{\omega}^{(Hete)})^{-2} \left( \sum_{t=2}^{T} \mathbb{E}[(J_t)^4] \right) = \Delta T^{-4} M^{-2} \left(T^3 M^2\right) = O(1/T) = o_p(1)$$

for finite $\Delta > 0$, as $T \to \infty$.

ii) The following condition needs to hold:

$$T^{-2}\left(D_{\omega,T}^{(\text{Hete})}\right)^{-1} \sum_{t=2}^{T} \mathbb{E}[J_t^2 \mid \mathcal{I}(t-1)] \xrightarrow{p} 1.$$

Because of the conditional homoskedasticity of $X$, we write:

$$\mathbb{E}[J_t^2 | \mathcal{I}(t-1)] = \mathbb{E}[(X_t' S_t)^2 | \mathcal{I}(t-1)] = d_1 \|S_t\|^2$$

For convenience, let us denote:

$$V_T^2 := \sum_{t=2}^{T} \mathbb{E}[J_t^2 \mid \mathcal{I}(t-1)] = d_1 \sum_{t=2}^{T} \|S_t\|^2.$$

Then we wish to prove the following convergence in probability:

$$\frac{4V_T^2}{T^2 D_{\omega,T}^{(Hete)}} \xrightarrow{p} 1.$$

since we have: $T\mathcal{T}_{2\omega}^c = (2/T)\sum_{t=2}^T J_t$. We write:

$$||S_t||^2 = \Big\|\sum_{s=2}^{t-1} X_s W_{ts}\Big\|^2 = \sum_{s=2}^{t-1}\sum_{r=2}^{t-1} W_{ts}W_{tr}\langle X_s, X_r\rangle.$$

Under the null hypothesis, we have:

$$\mathbb{E}[W_{ts}W_{tr}\langle X_s, X_r\rangle] = 0, \qquad \text{when} \quad r \neq s$$

Thus:

$$\mathbb{E}||S_t||^2 = \sum_{s=2}^{t-1}\mathbb{E}\big[||X_s||^2 W_{ts}^2\big] = d_1\sum_{s=2}^{t-1}\mathbb{E}[W_{ts}^2],$$

where the last equality is by conditional homoskedasticity. Hence,

$$\mathbb{E}[V_T^2] = d_1^2\sum_{t=2}^{T}\sum_{s=2}^{t-1}\mathbb{E}[W_{ts}^2] = \frac{T^2}{4}D_{\omega,T}^{(Hete)}$$

Note that: $\mathbb{E}[V_T^2] \asymp T^2 M$ due to $D_{\omega,T}^{(Hete)} = O(M)$.
Under Assumption 2, the products of the form $\langle Z_{t-j}, Z_{s-j}\rangle\langle Z_{t-\ell}, Z_{s-\ell}\rangle$ have absolutely summable covariances across $t$ and uniformly over $(j,\ell)$ (see Lemma B.5). In particular, we have the following bound for the covariance across $\{\|S_t\|^2\}$:

$$\sum_{h=1}^{\infty}\sup_t\left|\operatorname{Cov}\left(||S_t||^2, ||S_{t+h}||^2\right)\right| \leq \Delta M^2$$

for some finite $\Delta > 0$. Hence:

$$\begin{aligned}\text{Var}(V_T^2) &= \text{Var}\left(d_1 \sum_{t=2}^{T} \|S_t\|^2\right) \\ &\leq \Delta \left(T \sup_t \text{Var}(\|S_t\|^2) + 2T \sum_{h=1}^{\infty} \sup_t \left|\text{Cov}\left(\|S_t\|^2, \|S_{t+h}\|^2\right)\right|\right) \\ &= O(T^3 M^4)\end{aligned}$$

for some finite $\Delta > 0$ (see Lemma B.5). Since $\mathbb{E}[V_T^2] \asymp T^2 M$, we conclude:

$$\frac{\text{Var}(V_T^2)}{\mathbb{E}[V_T^2]^2} = O(M^2/T)$$

which means: $\text{Var}(V_T^2)/\mathbb{E}[V_T^2]^2 \to 0$, as $M, T \to \infty$, ultimately implying the desired convergence in probability.
Note that, by a similar argument using Marcinkiewicz-Zygmund inequalities, under the additional conditions on $\{\Lambda_{i,t}\}$, we have: $\text{Var}(V_T^2) = O(T^3 M^2)$, which in turn implies: $\text{Var}(V_T^2)/\mathbb{E}[V_T^2]^2 = O(1/T)$.

**Lemma B.5.** *We have the following:*

*i) For a finite $\Delta > 0$, we have:*

$$\sum_{h=1}^{\infty} \left|Cov(||Z_0||^2, ||Z_h||^2)\right| \leq \Delta \sum_{h=1}^{\infty} \alpha(h)^{\delta/(8+\delta)} < \infty$$

*ii) For a finite $\Delta > 0$ (independent of the indexes and lags $t, s, j, \ell$), we have for all $h \geq 1$,*

$$\sup_{t>s} \left|\text{Cov}\big(\langle Z_{t-j}, Z_{s-j}\rangle\langle Z_{t-\ell}, Z_{s-\ell}\rangle, \langle Z_{t+h-j}, Z_{s+h-j}\rangle\langle Z_{t+h-\ell}, Z_{s+h-\ell}\rangle\big)\right| \leq \Delta\alpha(h)^{\delta/(8+\delta)}$$

*and hence the summability:*

$$\sum_{h=1}^{\infty} \sup_{t>s} \left|Cov\big(\langle Z_{t-j}, Z_{s-j}\rangle\langle Z_{t-\ell}, Z_{s-\ell}\rangle, \langle Z_{t+h-j}, Z_{s+h-j}\rangle\langle Z_{t+h-\ell}, Z_{s+h-\ell}\rangle\big)\right| < \infty$$

*iii) Under the Assumptions of Theorem 1, for a finite $\Delta > 0$, we have:*

$$\sum_{h=1}^{\infty} \sup_t \left|\mathrm{Cov}(\|S_t\|^2, \|S_{t+h}\|^2)\right| \leq \Delta\, M^2,$$

*Proof.* Let be $p = (8+\delta)/4$. By Assumption 2, we have: $\mathbb{E}||Z_0||^{8+\delta} < \infty$, and so the finiteness of the other lower order moments. By standard covariance inequality for $\alpha$-mixing sequences (Dedecker et al., 2007), for some finite $\Delta > 0$:

$$|\mathrm{Cov}(U, V)| \leq \Delta\, \|U\|_p \|V\|_p\, \alpha(h)^{1-2/p},$$

where $U$ and $V$ are processes measurable with respect to the sigma of process $Z$.

i. By direct application of the inequality with $U = ||Z_0||^2$ and $V = ||Z_h||^2$, we have:

$$|\mathrm{Cov}(||Z_0||^2, ||Z_h||^2)| \leq \Delta ||||Z_0\|^2\|_p^2\, \alpha(h)^{\delta/(8+\delta)}$$

which proves the statement.

ii. We have:

$$\langle Z_{t-j}, Z_{s-j}\rangle\langle Z_{t-\ell}, Z_{s-\ell}\rangle \leq ||Z_{t-j}||||Z_{s-j}||||Z_{t-\ell}||||Z_{s-\ell}||$$

and so bounded. Fixing $(t, s, j, \ell)$ and looking over lags $h$ along time, for a finite $\Delta > 0$:

$$\sup_{t>s} \left|\mathrm{Cov}\big(\langle Z_{t-j}, Z_{s-j}\rangle\langle Z_{t-\ell}, Z_{s-\ell}\rangle, \langle Z_{t+h-j}, Z_{s+h-j}\rangle\langle Z_{t+h-\ell}, Z_{s+h-\ell}\rangle\big)\right| \leq \Delta \sup_{t>s} ||A_{t,s}(j,\ell)||_p^2 \alpha(h)^{\delta/(8+\delta)} \leq \Delta\alpha(h)^{\delta/(8+\delta)}$$

using a standard block-style argument for overlapping with small $h$.

iii. Recall:

$$||S_t||^2 = \sum_{s=2}^{t-1}\sum_{r=2}^{t-1} W_{ts}W_{tr}\, \langle X_s, X_r\rangle$$

Under the assumptions of Theorem 1, we have: $\sup_t \mathbb{E}||X_t||^4 < \infty$.
Let us consider the following:

$$W_{ts}^2 = \sum_{j\in\mathfrak{I}_{ts}} \sum_{\ell\in\mathfrak{I}_{ts}} \omega(j)\omega(\ell)\langle Z_{t-j}, Z_{s-j}\rangle\langle Z_{t-\ell}, Z_{s-\ell}\rangle$$

Note that, due to Assumption 1, there are at most $O(M)$ indices and so the total sum has at most $O(M^2)$ terms. Thus, by virtue of Cauchy-Schwarz and the covariance summability, the covariances across $\{\|S_t\|^2\}$ can be bounded as follows:

$$\sum_{h=1}^{\infty} \sup_t \left|\text{Cov}\left(||S_t||^2, ||S_{t+h}||^2\right)\right| \leq \Delta\Big(\sum_{j\in\mathfrak{I}} |\omega(j)|\Big)^2 \leq \Delta M^2$$

for some finite $\Delta > 0$. Note that, a similar argument can applied to $\sup_t \text{Var}(||S_t||^2)$, which leads to: $\sup_t \text{Var}(||S_t||^2) = O(T^2M^4)$.

□

## B.3 Proof of Theorem 2 (Size)

I introduce some additional notations and objects.
Given a parameter belonging to a parameter space, $\theta \in \Theta$, I define the gradient and Hessian operators with respect to $\theta$ to be $\nabla_\theta$ and $\nabla^2_\theta$, whenever have proper meaning. Define $\{\Theta_i\}_{i=1,2}$ the parameter spaces of the real parameters $\{\theta_i^0\}_{i=1,2}$, respectively. Following eq.(7), given $\{\theta_i \in \Theta_i\}_{i=1,2}$, I defined the processes:

$$\widetilde{X}_t(\theta_1) = W_{1,t} - \mu_X(\theta_1, \mathcal{I}_X(t-1))$$
$$\widetilde{Z}_t(\theta_2) = W_{2,t} - \mu_Z(\theta_2, \mathcal{I}_Z(t-1))$$

with corresponding standardized innovations of the unobservable infinite past:

$$\widetilde{U}_t = (\Gamma_X)^{-1/2}\,\widetilde{X}_t(\theta_1), \qquad \widetilde{V}_t = (\Gamma_Z)^{-1/2}\,\widetilde{Z}_t(\theta_2)$$

Under the parametrization of eq.(7), we have:

$$U_t = \widetilde{U}_t(\theta_1^0), \qquad V_t = \widetilde{V}_t(\theta_2^0)$$

Define further the population-standardized estimated innovations as follow:

$$\breve{U}_t = (\Gamma_X)^{-1/2}\,\widehat{X}_t, \qquad \breve{V}_t = (\Gamma_Z)^{-1/2}\,\widehat{Z}_t$$

where, using the previous notation, we have:

$$\widehat{X}_t = W_{1,t} - \mu_X(\widehat{\theta}_1, \widehat{\mathcal{I}}_X(t-1))$$
$$\widehat{Z}_t = W_{2,t} - \mu_Z(\widehat{\theta}_2, \widehat{\mathcal{I}}_Z(t-1))$$

where $\widehat{\mathcal{I}}_X(t-1)$ and $\widehat{\mathcal{I}}_Z(t-1)$ are the feasible information sets, i.e., the information sets constrained to the observable finite past of the time series, $\{W_{i,t}\}_{t=1,\dots,T}^{i=1,2}$.
Note that, generally: $\widehat{X}_t \neq \widetilde{X}_t(\widehat{\theta}_1),\ \widehat{Z}_t \neq \widetilde{Z}_t(\widehat{\theta}_2)$.

Recall that:

$$\widehat{U}_t = \left(\widehat{\Gamma}_X\right)^{-1/2} \widehat{X}_t, \qquad \widehat{V}_t = \left(\widehat{\Gamma}_Z\right)^{-1/2} \widehat{Z}_t$$

Denote the $k^{th}$ entry-wise element of $\widehat{U}_t, \breve{U}_t, \widetilde{U}_t, U_t$ with $\widehat{U}_{k,t}, \breve{U}_{k,t}, \widetilde{U}_{k,t}, U_{k,t}$, respectively. In a similar fashion, denote $\widehat{V}_{l,t-j}, \breve{V}_{l,t-j}, \widetilde{V}_{l,t-j}, V_{l,t-j}$ the $l^{th}$ element of, respectively, $\widehat{V}_{t-j}, \breve{V}_{t-j}, \widetilde{V}_{t-j}, V_{t-j}$.
We denote:

$$C_{UV}(j) = \frac{1}{T}\sum_{t=1}^{T} U_t(V_{t-j})', \;\; C_X = \frac{1}{T}\sum_{t=1}^{T} X_t(X_t)', \;\; C_Z = \frac{1}{T}\sum_{t=1}^{T} Z_t(Z_t)'$$

$$\widehat{\Gamma}_{UV}(j) = \frac{1}{T}\sum_{t=1}^{T} \widehat{U}_t(\widehat{V}_{t-j})'$$

$$\widehat{C}_{UV}(j) = \frac{1}{T}\sum_{t=j+1}^{T} \breve{U}_t(\breve{V}_{t-j})' = (\Gamma_X)^{-1/2}\widehat{\Gamma}_{\hat{X}\hat{Z}}(j)(\Gamma_Z)^{-1/2}$$

$$\widehat{\Gamma}_{\hat{X}\hat{Z}}(j) = \frac{1}{T}\sum_{t=j+1}^{T} \widehat{X}_t\widehat{Z}'_{t-j}$$

which are respectively: i) the sample covariance between standardized innovations and the sample variances of the innovations; ii) the sample covariance between feasible standardized residuals, iii) and the population-standardized sample covariance between estimated residuals, iv) the sample covariance between estimated residuals.

Additional to the assumptions listed in Theorem 2, this paper presume the following assumption:

**Assumption 3.** *For each $j \geq 1$ and each $(k, l)$, the process $\{(\breve{U}_{k,t} - \widetilde{U}_{k,t})V_{l,t-j}\}_{t\geq j+1}$ is a martingale difference sequence with respect to the feasible filtration, $\widehat{\mathcal{I}}(t-1)$, with:*

$$\sup_{t\geq j+1} \mathbb{E}\left[(\breve{U}_{k,t} - \widetilde{U}_{k,t})^2 V^2_{l,t-j} | \widehat{\mathcal{I}}(t-1)\right] = O(T^{-1})$$

and presume the following conditions:

$$\sup_{\theta_1\in\Theta_1}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}||U_t-\widetilde{U}_t||^2=O(T^{-1}),\quad \sup_{\theta_1\in\Theta_1}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}||\breve{U}_t-\widetilde{U}_t||^2=O(T^{-1})$$
$$\sup_{\theta_2\in\Theta_2}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}||V_t-\widetilde{V}_t||^2=O(T^{-1}),\quad \sup_{\theta_2\in\Theta_2}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}||\breve{V}_t-\widetilde{V}_t||^2=O(T^{-1}) \tag{12}$$

and

$$\begin{aligned}&\sup_{\theta_1\in\Theta_1}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}||\nabla_{\theta_1}\widetilde{U}_t(\theta_1)||^4=O(1),\quad \sup_{\theta_2\in\Theta_2}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}||\nabla_{\theta_2}\widetilde{V}_t(\theta_2)||^4=O(1)\\&\sup_{\theta_1\in\Theta_1}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}||\nabla^2_{\theta_1}\widetilde{U}_t(\theta_1)||^4=O(1),\quad \sup_{\theta_2\in\Theta_2}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}||\nabla^2_{\theta_2}\widetilde{V}_t(\theta_2)||^4=O(1)\end{aligned} \tag{13}$$

which are regularity conditions on the uniform $\ell_2-$convergence, the second-order differentiability and boundedness of the derivatives.

These conditions are comparably standard in the literature. For instance, see Assumption A3 in Hong and Lee (2005), Assumptions 3.2-3.4 in Wang et al. (2022), and Assumptions 2.3-2.4 in Leong and Urga (2023).

A remark is needed. Assumption 3 reads as a high-level orthogonality condition on the estimation error of the structural shock $X$, that ensures that the estimation error is (sufficiently) asymptotically uncorrelated, given the consistent $\sqrt{T}-$estimator. When admitting the following expansion: $\breve{U}_t-\widetilde{U}_t=\nabla_{\theta_1}\widetilde{U}_t(\theta_1^0)'(\widehat{\theta}_1-\theta_1^0)+r_t$, with: $\sup_t\mathbb{E}r_t^2=o(T^{-1})$, Assumption 3 essentially requires the estimation error term to be orthogonal to $V_{t-j}$, conditionally on the feasible past.

Notice that we can write the following:

$$\frac{T\cdot\widehat{\mathcal{T}}^c_\omega-\mu_{\omega,T}}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}}=\frac{T\left(\widehat{\mathcal{T}}_\omega-\widehat{\mathcal{T}}^\star_\omega\right)}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}}+\frac{T\left(\widehat{\mathcal{T}}^\star_\omega-\mathcal{T}^\star_\omega\right)}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}}-\frac{T\left(\widehat{\mathcal{C}}_\omega-\mathcal{C}^\star_\omega\right)}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}}$$
$$+\frac{\sqrt{D^{(Hete)}_{\omega,T}}}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}}\left(\frac{T\cdot\mathcal{T}^{\star c}_\omega-\mu_{\omega,T}}{\sqrt{D^{(Hete)}_{\omega,T}}}\right)$$

with:

$$
\begin{aligned}
\widehat{\mathcal{T}}_\omega &= \sum_{j=1}^{T-1} \omega(j) ||\widehat{\Gamma}_{UV}(j)||_F^2 \\
&= \frac{1}{T^2}\sum_{j=1}^{T-1}\omega(j)\sum_{t=j+1}^{T} ||\widehat{U}_t||^2||\widehat{V}_{t-j}||^2 + \frac{1}{T^2}\sum_{j=1}^{T-1}\omega(j)\sum_{s,t=j+1,s\neq t}^{T} <\widehat{U}_t,\widehat{U}_s><\widehat{V}_{t-j},\widehat{V}_{s-j}> \\
\widehat{\mathcal{T}}_\omega^\star &= \sum_{j=1}^{T-1} \omega(j) ||\widehat{C}_{UV}(j)||_F^2 \\
&= \frac{1}{T^2}\sum_{j=1}^{T-1}\omega(j)\sum_{t=j+1}^{T} ||\breve{U}_t||^2||\breve{V}_{t-j}||^2 + \frac{1}{T^2}\sum_{j=1}^{T-1}\omega(j)\sum_{s,t=j+1,s\neq t}^{T} <\breve{U}_t,\breve{U}_s><\breve{V}_{t-j},\breve{V}_{s-j}> \\
\mathcal{T}_\omega^\star &= \sum_{j=1}^{T-1} \omega(j) ||C_{UV}(j)||_F^2 \\
&= \frac{1}{T^2}\sum_{j=1}^{T-1}\omega(j)\sum_{t=j+1}^{T} ||U_t||^2||V_{t-j}||^2 + \frac{1}{T^2}\sum_{j=1}^{T-1}\omega(j)\sum_{s,t=j+1,s\neq t}^{T} <U_t,U_s><V_{t-j},V_{s-j}> \\
\widehat{\mathcal{C}}_\omega &= \frac{1}{T^2}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t,s\leq t-j}^{T} <\widehat{U}_t,\widehat{U}_s><\widehat{V}_{t-j},\widehat{V}_{s-j}> \\
\mathcal{C}_\omega^\star &= \frac{1}{T^2}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t,s\leq t-j}^{T} <U_t,U_s><V_{t-j},V_{s-j}> \\
\mathcal{T}_\omega^{\star c} &= \mathcal{T}_\omega^\star - \mathcal{C}_\omega^\star \\
\widehat{D}_{\omega,T}^{(Hete)} &= \frac{4d_1^2}{T^2}\sum_{t=2}^{T}\sum_{s=2}^{t-1}\sum_{j\in\mathfrak{J}_{ts}}\sum_{\ell\in\mathfrak{J}_{ts}}\omega(j)\omega(\ell)\widehat{\gamma}_{t,s}(j,\ell)
\end{aligned}
$$

where the correction term, $\mathcal{C}_\omega^\star$, is defined accordingly to eq.(5).
The proof of Theorem 2 follows from Propositions (3)-(6), and by direct application of Slutsky's theorem together with Theorem 1 which shows that: (adapting the notation)

$$
\frac{T\cdot\mathcal{T}_\omega^{\star c} - \mu_{\omega,T}}{\sqrt{D_{\omega,T}^{(Hete)}}} \xrightarrow{d} \mathcal{N}(0,1)
$$

**Proposition 3.** *Under the Assumptions of Theorem 2, we have:*

$$\frac{T\left(\widehat{\mathcal{T}}^{\star}_{\omega}-\mathcal{T}^{\star}_{\omega}\right)}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}}\xrightarrow{p}0$$

*Proof.* The proof is provided in Appendix B.3.1. □

**Proposition 4.** *Under the Assumptions of Theorem 2, we have:*

$$\frac{T\left(\widehat{\mathcal{T}}_{\omega}-\widehat{\mathcal{T}}^{\star}_{\omega}\right)}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}}\xrightarrow{p}0$$

*Proof.* The proof is provided in Appendix B.3.2. □

**Proposition 5.** *Under the Assumptions of Theorem 2, we have:*

$$\frac{T\left(\widehat{\mathcal{C}}_{\omega}-\mathcal{C}^{\star}_{\omega}\right)}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}}\xrightarrow{p}0$$

*Proof.* The proof is provided in Appendix B.3.3. □

**Proposition 6.** *Under the Assumptions of Theorem 2, we have:*

$$\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}\xrightarrow{p}\sqrt{D^{(Hete)}_{\omega,T}}$$

*Proof.* Given the consistency of the estimators, the boundedness of the moments and the conditions of eq.(12)-(13), showing that $\left(\widehat{D}^{(Hete)}_{\omega,T}-D^{(Hete)}_{\omega,T}\right)=o_p(1)$ is parallel to the first part of the proof of Proposition 4 (Appendix B.3.2). The proof concludes by virtue of the continuous mapping theorem. □

### B.3.1 Proof of Proposition 3

The proof is parallel to the one of Hong (2001)'s Lemma A.1-2, Bouhaddioui and Roy (2006)'s Lemma 2, and Leong and Urga (2023)'s Appendix

B.
The aim is to show that:

$$T\left(\widehat{\mathcal{T}}_\omega^\star - \mathcal{T}_\omega^\star\right) = o_p(M^{1/2})$$

since $\widehat{D}_{\omega,T}^{(Hete)} = O(M)$ by Proposition 2 and Proposition 6.
The initial section of the proof studies the difference:

$$\begin{aligned} T\left(\widehat{\mathcal{T}}_\omega^\star - \mathcal{T}_\omega^\star\right) &= T\sum_{j=1}^{T-1}\omega(j)\left(\left|\left|\text{vec}\left[\widehat{C}_{UV}(j)\right]\right|\right|^2 - ||\text{vec}\left[C_{UV}(j)\right]||^2\right) \\ &= T\sum_{j=1}^{T-1}\omega(j)\Big(||\text{vec}(\widehat{C}_{UV}(j)) - \text{vec}(C_{UV}(j))||^2 \\ &\qquad + 2\langle\text{vec}(C_{UV}(j)), \text{vec}(\widehat{C}_{UV}(j)) - \text{vec}(C_{UV}(j))\rangle\Big) \end{aligned}$$

The proof then consists of two parts:

i) We have:

$$\begin{aligned} &||\text{vec}(\widehat{C}_{UV}(j)) - \text{vec}(C_{UV}(j))|| \\ &= \sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\frac{1}{T}\left(\sum_{t=j+1}^{T}\breve{U}_{k,t}\breve{V}_{l,t-j} - U_{k,t}V_{l,t-j}\right) \\ &= \sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\Bigg(\frac{1}{T}\sum_{t=j+1}^{T}(\breve{U}_{k,t} - U_{k,t})V_{l,t-j} + U_{k,t}(\breve{V}_{l,t-j} - V_{l,t-j}) \\ &\qquad + (\breve{U}_{k,t} - U_{k,t})(\breve{V}_{l,t-j} - V_{l,t-j})\Bigg) \\ &= \sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\left(F_{Tj}^{(1)} + F_{Tj}^{(2)} + F_{Tj}^{(3)}\right) \end{aligned}$$

Thus, applying of Cauchy-Schwarz inequality:

$$||\text{vec}(\widehat{C}_{UV}(j)) - \text{vec}(C_{UV}(j))||^2 \leq \Delta\sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\left((F_{Tj}^{(1)})^2 + (F_{Tj}^{(2)})^2 + (F_{Tj}^{(3)})^2\right)$$

for some finite $\Delta > 0$. Let us study the three terms:

A. By Cauchy–Schwarz,

$$\sup_j (F_{Tj}^{(3)})^2 = \sup_j \left( \frac{1}{T} \sum_{t=j+1}^{T} (\breve{U}_{k,t} - U_{k,t})(\breve{V}_{l,t-j} - V_{l,t-j}) \right)^2$$
$$\leq \left( \frac{1}{T} \sum_{t=1}^{T} (\breve{U}_{k,t} - U_{k,t})^2 \right) \left( \frac{1}{T} \sum_{t=1}^{T} (\breve{V}_{l,t} - V_{l,t})^2 \right).$$

Next, by decomposing the difference as: $\breve{U} - U = (\breve{U} - \widetilde{U}) + (\widetilde{U} - U)$, we write:

$$\frac{1}{T} \sum_{t=1}^{T} (\breve{U}_{k,t} - U_{k,t})^2 \leq 2 \left( \frac{1}{T} \sum_{t=1}^{T} (\breve{U}_{k,t} - \widetilde{U}_{k,t})^2 \right) + 2 \left( \frac{1}{T} \sum_{t=1}^{T} (\widetilde{U}_{k,t} - U_{k,t})^2 \right) = O_p(T^{-1}),$$

where the last equality follows from (12) and by virtue of Markov inequality. The same holds for term: $\frac{1}{T} \sum_{t=1}^{T} (\breve{V}_{l,t} - V_{l,t})^2 = O_p(T^{-1})$. Hence, we have: $\sup_j (F_{Tj}^{(3)})^2 = O_p(T^{-2})$.

B. By the same decomposition, we have:

$$F_{Tj}^{(1)} = \frac{1}{T} \sum_{t=j+1}^{T} (\breve{U}_{k,t} - U_{k,t}) V_{l,t-j}$$
$$= \frac{1}{T} \sum_{t=j+1}^{T} (\breve{U}_{k,t} - \widetilde{U}_{k,t}) V_{l,t-j} + \frac{1}{T} \sum_{t=j+1}^{T} (\widetilde{U}_{k,t} - U_{k,t}) V_{l,t-j} = F_{Tj}^{(11)} + F_{Tj}^{(12)}.$$

Let us consider the two terms:

- Under Assumption 3, we can write:

$$\mathbb{E}\big[(F_{Tj}^{(11)})^2\big] = \mathbb{E}\left[ \left( \frac{1}{T} \sum_{t=j+1}^{T} (\breve{U}_{k,t} - \widetilde{U}_{k,t}) V_{l,t-j} \right)^2 \right]$$
$$= \frac{1}{T^2} \sum_{t=j+1}^{T} \mathbb{E}\big[(\breve{U}_{k,t} - \widetilde{U}_{k,t})^2 V_{l,t-j}^2\big] = O(T^{-2})$$

  By Markov inequality, we have: $F_{Tj}^{(11)} = O_p(T^{-1})$ and $(F_{Tj}^{(11)})^2 = O_p(T^{-2})$.

- By the conditions of eq.(13), the term $F_{Tj}^{(12)}$ can be expressed into two terms using its Taylor expansion (up to the second order):

$$
\begin{aligned}
F_{Tj}^{(12)} &= \frac{1}{T}\sum_{t=j+1}^{T}(\widetilde{U}_{k,t} - U_{k,t})V_{l,t-j} \\
&= \frac{1}{T}(\hat{\theta}_1 - \theta_1^0)' \sum_{t=j+1}^{T}(\nabla_{\theta_1}\widetilde{U}_{k,t}(\theta_1^0)V_{l,t-j}) \\
&\quad + \frac{1}{2T}(\hat{\theta}_1 - \theta_1^0)' \sum_{t=j+1}^{T}(\nabla^2_{\theta_1}\widetilde{U}_{k,t}(\breve{\theta}_1)V_{l,t-j})(\hat{\theta}_1 - \theta_1^0)
\end{aligned}
$$

where $\breve{\theta}_1 \in [\hat{\theta}_1, \theta_1^0]$. By virtue of Cauchy-Schwarz inequality:

$$
\begin{aligned}
\frac{1}{T}(\hat{\theta}_1 - \theta_1^0)' \sum_{t=j+1}^{T}(\nabla_{\theta_1}\widetilde{U}_{k,t}(\theta_1^0)V_{l,t-j}) &= \frac{1}{T}(\hat{\theta}_1 - \theta_1^0)' \sum_{t=j+1}^{T}((\nabla_{\theta_1}U_{k,t})(V_{l,t-j})) \\
&\leq \frac{1}{T^2}\mathbb{E}\left[||\hat{\theta}_1 - \theta_1^0||^2\right]\mathbb{E}\left[\left\{\sum_{t=j+1}^{T}||(\nabla_{\theta_1}U_{k,t})(V_{l,t-j})||\right\}^2\right] \\
&\leq \frac{1}{T^2}\mathbb{E}\left[||\hat{\theta}_1 - \theta_1^0||^2\right]\sum_{t=j+1}^{T}\mathbb{E}\left[||\nabla_{\theta_1}U_{k,t}||^4\right]\mathbb{E}\left[||V_{l,t-j}||^2\right] = O_p(T^{-2})
\end{aligned}
$$

where the last equality follows from the boundedness of the moments, the consistency of the estimator, and the conditions in eq.(13). A remark is needed. To prove the last inequality, there is a trade-off between bounding the moments of the derivatives $\{\nabla_{\theta_1}U_{k,t}, \nabla^2_{\theta_1}U_{k,t}\}$ and imposing additional orthogonality conditions. In fact, instead of imposing the boundedness of the fourth moment of $\{\nabla_{\theta_1}U_{k,t}\}$, one could reach the same conclusions by imposing orthogonality between $\{\nabla_{\theta_1}U_t\}$ and $\{V_{t-j}\}$. Since the purpose of the proposed statistic is to infer about weak exogeneity, we wish not to preclude possible causality channels by imposing extra restrictions on those.

By the same logic, the other term:

$$\frac{1}{2T}(\hat{\theta}_1-\theta_1^0)'\sum_{t=j+1}^{T}(\nabla^2_{\theta_1}\widetilde{U}_{k,t}(\breve{\theta}_1)V_{l,t-j})(\hat{\theta}_1-\theta_1^0)$$
$$\leq \frac{1}{4T^2}\mathbb{E}\left[||\hat{\theta}_1-\theta_1^0||^4\right]\mathbb{E}\left[\left|\left|\sum_{t=j+1}^{T}\nabla^2_{\theta_1}\widetilde{U}_{k,t}(\breve{\theta}_1)V_{l,t-j}\right|\right|^2\right]=O_p(T^{-2})$$

which means: $F_{Tj}^{(12)}=O_p(T^{-2})$. In conclusions:

$$(F_{Tj}^{(1)})^2=O_p(T^{-2})$$

C. By reasoning analogue to the one above, together with the null of interest and the conditional homoskedasticity of the series $U$, we have: $(F_{Tj}^{(2)})^2=O_p(T^{-2})$.

Since:

$$\sum_{j=1}^{T-1}\omega(j)||\text{vec}(\widehat{C}_{UV}(j))-\text{vec}(C_{UV}(j))||^2=O_p(MT^{-2})$$

we have:

$$T\sum_{j=1}^{T-1}\omega(j)||\text{vec}(\widehat{C}_{UV}(j))-\text{vec}(C_{UV}(j))||^2=O_p(M/T)=o_p(1)$$

where the last equality is due to $M/T\to 0$, as $M,T\to\infty$.

ii) We have:

$$\sum_{j=1}^{T-1}\omega(j)\langle \text{vec}(C_{UV}(j)), \text{vec}(\widehat{C}_{UV}(j)) - \text{vec}(C_{UV}(j))\rangle$$

$$= \sum_{j=1}^{T-1}\omega(j)\text{vec}(C_{UV}(j))'(\text{vec}(\widehat{C}_{UV}(j)) - \text{vec}(C_{UV}(j)))$$

$$= \sum_{j=1}^{T-1}\omega(j)\sum_{k=1}^{d_1d_2}\sum_{l=1}^{d_1d_2}C_{k,UV}(j)\left(\widehat{C}_{l,UV}(j) - C_{l,UV}(j)\right)$$

$$= \sum_{k=1}^{d_1d_2}\sum_{l=1}^{d_1d_2}\sum_{j=1}^{T-1}\omega(j)C_{k,UV}(j)\left(\widehat{C}_{l,UV}(j) - C_{l,UV}(j)\right)$$

$$\leq \sum_{k=1}^{d_1d_2}\sum_{l=1}^{d_1d_2}\left(\sum_{j=1}^{T-1}\omega(j)(C_{k,UV}(j))^2\right)^{1/2}\left(\sum_{j=1}^{T-1}\omega(j)\left(\widehat{C}_{l,UV}(j) - C_{l,UV}(j)\right)^2\right)^{1/2}$$

where the last inequality is by Cauchy-Schwarz inequality.
By the results of the previous parts:

$$\left(\sum_{j=1}^{T-1}\omega(j)\left(\widehat{C}_{l,UV}(j) - C_{l,UV}(j)\right)^2\right)^{1/2} = O_p(M^{1/2}T^{-1})$$

$$\left(\sum_{j=1}^{T-1}\omega(j)(C_{k,UV}(j))^2\right)^{1/2} = O_p(M^{1/2}T^{-1/2})$$

where the last equality is by:

$$\sum_{j=1}^{T-1}\omega(j)(C_{k,UV}(j))^2 \leq \sum_{j=1}^{T-1}\omega(j)||C_{k,UV}(j)||^2 = O(MT^{-3/2})$$

In conclusions:

$$T\sum_{j=1}^{T-1}\omega(j)\langle \text{vec}(C_{UV}(j)), \text{vec}(\widehat{C}_{UV}(j)) - \text{vec}(C_{UV}(j))\rangle = O_p(MT^{-1/2}) = o_p(M^{1/2})$$

As announced, since $M/T \to 0$, as $T, M \to \infty$:

$$\widehat{\mathcal{T}}^\star_\omega - \mathcal{T}^\star_\omega = o_p(M^{1/2})$$

which means that the proof is concluded by Proposition 2 and Proposition 6.

### B.3.2 Proof of Proposition 4

The proof is parallel to the one of Hong (2001)'s Lemma A.1, Bouhaddioui and Roy (2006)'s Proposition 3, and Leong and Urga (2023)'s Appendix B. Similar to the proof of Proposition 3, the aim is to show that:

$$T\left(\widehat{\mathcal{T}}_\omega - \widehat{\mathcal{T}}^\star_\omega\right) = O_p(MT^{-1/2})$$

since $\widehat{D}^{(Hete)}_{\omega,T} = O(M)$ by Proposition 6 and Proposition 2.
By Lemma A.3, we can write:

$$\begin{aligned}
T\left(\widehat{\mathcal{T}}_\omega - \widehat{\mathcal{T}}^\star_\omega\right) &= T\sum_{j=1}^{T-1}\left(||\widehat{\Gamma}_{UV}(j)||_F^2 - ||\widehat{C}_{UV}(j)||_F^2\right) \\
&= T\sum_{j=1}^{T-1} k^2\left(\frac{j}{M}\right)\left(\text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))'\left(\widehat{\Gamma}_Z^{-1}\otimes\widehat{\Gamma}_X^{-1}\right)\text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))\right. \\
&\qquad\qquad \left. - \text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))'\left(\Gamma_Z^{-1}\otimes\Gamma_X^{-1}\right)\text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))\right) \\
&= T\sum_{j=1}^{T-1} k^2\left(\frac{j}{M}\right)\text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))'\left(\widehat{\Gamma}_Z^{-1}\otimes\widehat{\Gamma}_X^{-1} - \Gamma_Z^{-1}\otimes\Gamma_X^{-1}\right)\text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))
\end{aligned}$$

Recall that $\widehat{X}_{k,t}, X_{k,t}$ are the $k^{th}$ element of $\widehat{X}_t(\hat{\theta}_1)$, $X_t$, respectively.
Consider the term $(\widehat{\Gamma}_X - \Gamma_X)$:

$$||\widehat{\Gamma}_X - \Gamma_X||_F \leq \sum_{k=1}^{d_1}\sum_{l=1}^{d_1}||\widehat{\Gamma}_{kl,X} - \Gamma_{kl,X}||$$

By triangular equality, it follows that:

$$||\widehat{\Gamma}_{kl,X} - \Gamma_{kl,X}|| = ||\widehat{\Gamma}_{kl,X} - C_{kl,X} + C_{kl,X} - \Gamma_{kl,X}|| \leq ||\widehat{\Gamma}_{kl,X} - C_{kl,X}|| + ||C_{kl,X} - \Gamma_{kl,X}||$$

For the first term, by virtue of Cauchy-Schwarz inequality:

$$\widehat{\Gamma}_{kl,X} - C_{kl,X} = \frac{1}{T}\sum_{t=1}^{T} 2(\widehat{X}_{k,t} - X_{k,t})X_{k,t} + (\widehat{X}_{k,t} - X_{k,t})^2$$
$$\leq 4\left(\frac{1}{T}\sum_{t=1}^{T}(X_{k,t})^2\right)^{1/2}\left(\frac{1}{T}\sum_{t=1}^{T}(\widehat{X}_{k,t} - X_{k,t})^2\right)^{1/2} + \frac{1}{T}\sum_{t=1}^{T}(\widehat{X}_{k,t} - X_{k,t})^2 = O_p(T^{-1/2})$$

since we have that the term $\frac{1}{T}\sum_{t=1}^{T}(X_{k,t})^2 = O_p(1)$ because of Chebyshev inequality and boundedness of moments, while the term $\frac{1}{T}\sum_{t=1}^{T}(\widehat{X}_{k,t} - X_{k,t})^2 = O_p(T^{-1})$ as consequence of what proved in Proposition 3.
By application of Chebyshev inequality: $||C_{kl,X} - \Gamma_{kl,X}|| = O_p(T^{-1/2}),\ \forall k, l$.
This in turn implies:

$$\widehat{\Gamma}_{kl,X} - \Gamma_{kl,X} = O_p(T^{-1/2})$$

By analogue reasoning:

$$\widehat{\Gamma}_{kl,Z} - \Gamma_{kl,Z} = O_p(T^{-1/2})$$

and $\frac{1}{T}\sum_{t=1}^{T}(Z_{k,t})^2 = O_p(1)$, as well as $\Gamma_{kl,X} = O(1)$ and $\Gamma_{kl,Z} = O(1)$.
Thus, by virtue of the continuous mapping theorem, I can conclude that:

$$\widehat{\Gamma}_X^{-1} \otimes \widehat{\Gamma}_Z^{-1} - \Gamma_X^{-1} \otimes \Gamma_Z^{-1} = O_p(T^{-1/2})$$

Now, since $\Gamma_X$ and $\Gamma_Z$ are bounded (i.e., $\Gamma_X = O(1), \Gamma_Z = O(1)$), studying the boundedness of the term, $\widehat{\Gamma}_{\hat{X}\hat{Z}}(j)$, is equivalent to studying the boundedness of the term, $\widehat{C}_{UV}(j)$. Thus, I direct my focus to the term that follows:

$$\begin{aligned}
&\sum_{j=1}^{T-1}\omega(j)\big(\text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))'\text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))\big)\\
&\quad\asymp\sum_{j=1}^{T-1}\omega(j)\big(\text{vec}(\widehat{C}_{UV}(j))'\text{vec}(\widehat{C}_{UV}(j))\big)\\
&\quad=\sum_{j=1}^{T-1}\omega(j)\Big(\text{vec}(\widehat{C}_{UV}(j))'\text{vec}(\widehat{C}_{UV}(j))-\text{vec}(C_{UV}(j))'\text{vec}(C_{UV}(j))\\
&\qquad\qquad+\text{vec}(C_{UV}(j))'\text{vec}(C_{UV}(j))\Big)
\end{aligned}$$

i) We have:

$$\begin{aligned}
&\sum_{j=1}^{T-1}\omega(j)\left(\text{vec}(\widehat{C}_{UV}(j))'\text{vec}(\widehat{C}_{UV}(j))-\text{vec}(C_{UV}(j))'\text{vec}(C_{UV}(j))\right)\\
&=\sum_{j=1}^{T-1}\omega(j)\sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\widehat{C}^2_{kl,UV}(j)-C^2_{kl,UV}(j)\\
&=\sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\sum_{j=1}^{T-1}\omega(j)\left\{\left(\widehat{C}_{kl,UV}(j)-C_{kl,UV}(j)\right)^2+2\widehat{C}_{kl,UV}(j)\left(\widehat{C}_{kl,UV}(j)-C_{kl,UV}(j)\right)\right\}\\
&=O_p(MT^{-1})
\end{aligned}$$

where the last equality is because of Proposition 3.

ii) We have:

$$\sum_{j=1}^{T-1}\omega(j)\big(\text{vec}(C_{UV}(j))'\text{vec}(C_{UV}(j))\big)=\sum_{r=1}^{d_1d_2}\sum_{j=1}^{T-1}\omega(j)(C_{r,UV}(j))^2=O_p(MT^{-1})$$

where the last equality follows from Proposition 3, since $C_{i,UV}(j)$ as the $i^{th}$ entry of the matrices $\text{vec}(C_{UV}(j))$.

This in turn means that:

$$\sum_{j=1}^{T-1}\omega(j)\text{vec}(\widehat{C}_{UV}(j))'\text{vec}(\widehat{C}_{UV}(j))=O_p(MT^{-1})$$

In conclusion:

$$T\left(\widehat{\mathcal{T}}_\omega - \widehat{\mathcal{T}}^\star_\omega\right) = T\sum_{j=1}^{T-1} k^2\left(\frac{j}{M}\right)\text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))'\left(\widehat{\Gamma}_Z^{-1}\otimes\widehat{\Gamma}_X^{-1} - \Gamma_Z^{-1}\otimes\Gamma_X^{-1}\right)\text{vec}(\widehat{\Gamma}_{\hat{X}\hat{Z}}(j))$$
$$= T\cdot O_p(MT^{-1})\cdot O_p(T^{-1/2}) = O_p(MT^{-1/2})$$

which concludes the proof as $\widehat{D}^{(Hete)}_{\omega,T} = O(M)$ by Proposition 6 and Proposition 2.

### B.3.3 Proof of Proposition 5

The proof is parallel to Proposition 3 and Proposition 4.
As previously done, I consider the following decomposition:

$$\frac{T\left(\widehat{\mathcal{C}}_\omega - \breve{\mathcal{C}}_\omega\right)}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}} + \frac{T\left(\breve{\mathcal{C}}_\omega - \mathcal{C}^\star_\omega\right)}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}} \xrightarrow{p} 0$$

with:

$$\breve{\mathcal{C}}_\omega = \frac{1}{T^2}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t,s\leq t-j}^{T} < \breve{U}_t, \breve{U}_s >< \breve{V}_{t-j}, \breve{V}_{s-j} >$$

By a similar argument of Proposition 4:

$$\frac{T\left(\widehat{\mathcal{C}}_\omega - \breve{\mathcal{C}}_\omega\right)}{\sqrt{\widehat{D}^{(Hete)}_{\omega,T}}} \xrightarrow{p} 0$$

since the convergence is uniquely driven by the distance between $\widehat{\Gamma}_X$ and $\Gamma_X$, and between $\widehat{\Gamma}_Z$ and $\Gamma_Z$. It remains to prove:

$$T\left(\breve{\mathcal{C}}_\omega - \mathcal{C}^\star_\omega\right) = o_p(M^{1/2})$$

We have:

$$
\begin{aligned}
&T\left(\breve{\mathcal{C}}_\omega - \mathcal{C}^\star_\omega\right) \\
&= \frac{1}{T}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t,s\leq t-j}^{T}\left(\breve{U}_t'\breve{U}_s\breve{V}_{t-j}'\breve{V}_{s-j} - (U_t)'U_s(V_{t-j})'V_{s-j}\right) \\
&= \sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\frac{1}{T}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t,s\leq t-j}^{T}\left(\breve{U}_{k,t}\breve{U}_{k,s}\breve{V}_{l,t-j}\breve{V}_{l,s-j} - U_{k,t}U_{k,s}V_{l,t-j}V_{l,s-j}\right) \\
&= \sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\frac{1}{T}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t,s\leq t-j}^{T}\Bigg(\left[\breve{U}_{k,t}\breve{U}_{k,s} - U_{k,t}U_{k,s}\right]V_{l,t-j}V_{l,s-j} \\
&\qquad + U_{k,t}U_{k,s}\left[\breve{V}_{l,t-j}\breve{V}_{l,s-j} - V_{l,t-j}V_{l,s-j}\right] + \left[\breve{U}_{k,t}\breve{U}_{k,s} - U_{k,t}U_{k,s}\right]\left[\breve{V}_{l,t-j}\breve{V}_{l,s-j} - V_{l,t-j}V_{l,s-j}\right]\Bigg) \\
&= \sum_{k=1}^{d_1}\sum_{l=1}^{d_2}\sum_{j=1}^{T-2}\omega(j)\sum_{s,t=j+1,s\neq t,s\leq t-j}^{T}\Bigg( \\
&\qquad T^{-1}\Big[(\breve{U}_{k,t} - U_{k,t})U_{k,s} + (\breve{U}_{k,s} - U_{k,s})U_{k,t} + (\breve{U}_{k,t} - U_{k,t})(\breve{U}_{k,s} - U_{k,s})\Big]V_{l,t-j}V_{l,s-j} \\
&\quad + T^{-1}U_{k,t}U_{k,s}\Big[(\breve{V}_{l,t-j} - V_{l,t-j})V_{l,s-j} + (\breve{V}_{l,s-j} - V_{l,s-j})V_{l,t-j} + (\breve{V}_{l,t-j} - V_{l,t-j})(\breve{V}_{l,s-j} - V_{l,s-j})\Big] \\
&\quad + T^{-1}\left[\breve{U}_{k,t}\breve{U}_{k,s} - U_{k,t}U_{k,s}\right]\left[\breve{V}_{l,t-j}\breve{V}_{l,s-j} - V_{l,t-j}V_{l,s-j}\right]\Bigg)
\end{aligned}
$$

Proposition 3 shows that all terms inside the brackets are at most $O_p(T^{-3/2})$, which concludes the proof as:

$$\breve{\mathcal{C}}_\omega - \mathcal{C}^\star_\omega = o_p(M^{1/2}T^{-1/2})$$

## B.4 Proof of Theorem 3 (Power)

For the definition of the objects and the regularity conditions, please refer to the first part of the proof of Theorem 2 in Appendix B.3. The proof is parallel to the one of Hong (2001)'s Theorem 2, Bouhaddioui and Roy (2006)'s Theorem 2, and Leong and Urga (2023)'s Appendix C.
As done for Theorem 2, the test statistic is decomposed as follows:

$$\left(\frac{M^{1/2}}{T}\right)\frac{T\cdot\widehat{\mathcal{T}}_{\omega}^{c}-\mu_{\omega,T}}{\sqrt{\widehat{D}_{\omega,T}^{(Hete)}}}=\frac{\left(\widehat{\mathcal{T}}_{\omega}-\widehat{\mathcal{T}}_{\omega}^{\star}\right)}{\sqrt{M^{-1}\widehat{D}_{\omega,T}^{(Hete)}}}+\frac{\left(\widehat{\mathcal{T}}_{\omega}^{\star}-\mathcal{T}_{\omega}^{\star}\right)}{\sqrt{M^{-1}\widehat{D}_{\omega,T}^{(Hete)}}}-\frac{\left(\widehat{\mathcal{C}}_{\omega}-\mathcal{C}_{\omega}^{\star}\right)}{\sqrt{M^{-1}\widehat{D}_{\omega,T}^{(Hete)}}}+\frac{\sqrt{M^{-1}D_{\omega,T}^{(Hete)}}}{\sqrt{M^{-1}\widehat{D}_{\omega,T}^{(Hete)}}}\left(\frac{\mathcal{T}_{\omega}^{\star c}-\mu_{\omega,T}}{\sqrt{M^{-1}D_{\omega,T}^{(Hete)}}}\right)$$

By Assumption 1, as $\mu_{\omega,T}$ and $D_{\omega,T}^{(Hete)}$ are of order $O(M)$, together with Proposition 6, I can conclude:

$$\left(\frac{M^{1/2}}{T}\right)\frac{T\cdot\widehat{\mathcal{T}}_{\omega}^{c}-\mu_{\omega,T}}{\sqrt{\widehat{D}_{\omega,T}^{(Hete)}}}\xrightarrow{p}\frac{\left(\widehat{\mathcal{T}}_{\omega}-\widehat{\mathcal{T}}_{\omega}^{\star}\right)}{\sqrt{M^{-1}\widehat{D}_{\omega,T}^{(Hete)}}}+\frac{\left(\widehat{\mathcal{T}}_{\omega}^{\star}-\mathcal{T}_{\omega}^{\star}\right)}{\sqrt{M^{-1}\widehat{D}_{\omega,T}^{(Hete)}}}-\frac{\left(\widehat{\mathcal{C}}_{\omega}-\mathcal{C}_{\omega}^{\star}\right)}{\sqrt{M^{-1}\widehat{D}_{\omega,T}^{(Hete)}}}+\left(\frac{\mathcal{T}_{\omega}^{\star c}}{\sqrt{\Delta\int_{0}^{\infty}k^{4}(z)dz}}\right)+o_{p}(1)$$

for some finite $\Delta > 0$.
The proof of Theorem 3 follows from: i) Corollary 2, which shows that the first three terms are of order $o_p(1)$, ii) Corollary 1 together with Slutsky's theorem, and iii) Lemma B.6.
In conclusions, we have the following:

$$\frac{M^{1/2}}{T}\frac{T\widehat{\mathcal{T}}_{\omega}^{c}-\mu_{\omega,T}}{\sqrt{\widehat{D}_{\omega,T}^{(Hete)}}}\xrightarrow{p}\frac{\mathcal{T}_{\omega}^{\star c}}{\sqrt{\Delta\int_{0}^{\infty}k^{4}(z)dz}}+o_{p}(1)$$

**Corollary 1.** *Suppose Assumptions of Theorem 3 hold. It follows that:*

$$M^{-1}\widehat{D}_{\omega,T}^{(Hete)} \xrightarrow{p} \Delta \int_0^\infty k^4(z)dz$$

*for some finite $\Delta > 0$.*

*Proof.* The proof follows directly from Proposition 6 and by having:

$$M^{-1}D_{\omega,T}^{(Hete)} \xrightarrow[M\to\infty]{} \Delta \int_0^\infty k^4(z)dz$$

for some finite $\Delta > 0$, as discussed in Appendix B.1. □

**Corollary 2.** *Suppose the assumptions of Theorem 3 hold. It follows that:*

$$\widehat{\mathcal{T}}_\omega - \widehat{\mathcal{T}}_\omega^\star = o_p(1)$$
$$\widehat{\mathcal{T}}_\omega^\star - \mathcal{T}_\omega^\star = o_p(1)$$
$$\widehat{\mathcal{C}}_\omega - \mathcal{C}_\omega^\star = o_p(1)$$

*Proof.* The proof follows from direct application of Proposition 3-4-5, since:

$$\widehat{\mathcal{T}}_\omega - \widehat{\mathcal{T}}_\omega^\star = o_p(M^{1/2}T^{-1}), \quad \widehat{\mathcal{T}}_\omega^\star - \mathcal{T}_\omega^\star = o_p(M^{1/2}T^{-1}), \quad \widehat{\mathcal{C}}_\omega - \mathcal{C}_\omega^\star = o_p(M^{1/2}T^{-1})$$

and concludes by Assumption 1, as $M^{1/2}T^{-1} \to 0$. □

**Lemma B.6.** *Suppose Assumptions of Theorem 3 hold. We have as $T \to \infty$:*

$$\mathcal{T}_\omega^{\star c} \xrightarrow{p} \sum_{j=1}^{\infty} \left|\left| vec\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right|\right|^2 + o_p(1)$$

*Proof.* The proof is focused on establishing the consistency of the statistic, so that the difference between the two quantities is of order $o_p(1)$. We have:

$$\mathcal{T}^{\star c}_{\omega} - \sum_{j=1}^{\infty} ||\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}||^2 = \mathcal{T}^{\star}_{\omega} - \mathcal{C}^{\star}_{\omega} - \sum_{j=1}^{\infty} ||\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}||^2$$

$$= \sum_{j=1}^{T-1} \omega(j) \left( ||\text{vec}\,[C_{UV}(j)]||^2 - \left|\left|\text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right|\right|^2 \right) - \mathcal{C}^{\star}_{\omega}$$

$$+ \sum_{j=1}^{T-1} (\omega(j) - 1) \left|\left|\text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right|\right|^2 + \sum_{T}^{\infty} \left|\left|\text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right|\right|^2$$

For the first term, we have:

$$\sum_{j=1}^{T-1} \omega(j) \left( ||\text{vec}\,[C_{UV}(j)]||^2 - \left|\left|\text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right|\right|^2 \right) - \mathcal{C}^{\star}_{\omega}$$

$$= \left[ \sum_{j=1}^{T-1} \omega(j) \left|\left|\text{vec}\,[C_{UV}(j)] - \text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right|\right|^2 - \mathcal{C}^{\star}_{\omega} \right] + 2\sum_{j=1}^{T-1} \omega(j)\lambda^{\star}(j)$$

$$= \left[ \sum_{k,l=1}^{d_1,d_2} \sum_{j=1}^{T-1} \omega(j)\text{Var}[C_{kl,UV}(j)] - \mathcal{C}^{\star}_{\omega} \right] + 2\sum_{j=1}^{T-1} \omega(j)\lambda^{\star}(j)$$

with $\lambda^{\star}(j) = \left\langle \text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right], \text{vec}\left[C_{UV}(j) - \Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right] \right\rangle$.
Denote $\rho_{kl}(j)$, the $(k,l)-$ entry of the matrix $\left(\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right)$.
If $\{X_t, Z_t\}$ is a fourth-order stationary process so $\{U_t, V_t\}$ is fourth-order stationary as well, by Hannan (1970) pg.209-210 or, equivalently, by Priestley (1981) pg. 325–26:

$$\text{Var}[C_{kl,UV}(j)] = T^{-1} \sum_{j=1}^{T-1} \omega(j) \sum_{i=-T+1}^{T-1} \left(1 - \frac{i}{T}\right) \rho_{kl}(i+j)\rho_{kl}(i-j)$$

$$+ T^{-1} \sum_{j=1}^{T-1} \omega(j) \sum_{i=-T+1}^{T-1} \left(1 - \frac{|i|}{T}\right) \kappa_{klkl,XZ}(0, j, i, j+i)$$

Following a similar argument in Hong (2001)'s Lemma A.6, Bouhaddioui and Roy (2006)'s Lemma A.7, or Leong and Urga (2023)'s Lemma C.3, as $\sum_{m,r=1}^{d_1,d_2}\sum_{j,k,l=-\infty}^{\infty}\kappa_{mrmr,XZ}(0,j,k,l)<\infty$, we have:

$$\sum_{k,l=1}^{d_1,d_2}\sum_{j=1}^{T-1}\omega(j)\text{Var}[C_{kl,UV}(j)]-\mathcal{C}^{\star}_{\omega}=O_p(1/T+M/T)=o_p(1)$$

where the last equality is because of the assumption on the asymptotic rates ( $\frac{M}{T}\to 0$, as $T,M\to\infty$) and by realizing that $\mathcal{C}^{\star}_{\omega}$ represents the fourth-order cumulants of: $\{U_{m,t},V_{r,t-j},U_{m,t-k},V_{r,t-l}\}$.
By the dominated convergence theorem, the boundedness condition on the covariances, $\sum_{j=1}^{\infty}||\Gamma_{XZ}(j)||^2<\infty$, and by the assumptions on the asymptotic rates, I can conclude that:

$$\sum_{j=1}^{T-1}(\omega(j)-1)\left|\left|\text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right|\right|^2=o_p(1)$$
$$\sum_{T}^{\infty}\left|\left|\text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right|\right|^2=o_p(1)$$

This in turn implies as well:

$$\sum_{j=1}^{T-1}\omega(j)\left\langle\text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right],\text{vec}\left[C_{UV}(j)-\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right\rangle=o_p(1)$$

Hence:

$$\mathcal{T}^{\star c}_{\omega}\xrightarrow{p}\sum_{j=1}^{\infty}\left|\left|\text{vec}\left[\Gamma_X^{-1/2}\Gamma_{XZ}(j)\Gamma_Z^{-1/2}\right]\right|\right|^2+o_p(1)$$

which concludes the proof. □

# C Online Appendix: Monte Carlo Experiments

This Appendix presents Monte Carlo evidence on the finite-sample properties of the proposed test statistic relative to several benchmark procedures. Considering the bivariate process $\{X_t, Z_t\}$, we examine rejection rates under Data-Generating Processes (DGPs) calibrated to reflect features common in macroeconomic applications.

## C.1 Statistics Under Comparison

The experiments compare six test statistics. The first three are the finite-sample Portmanteau-type statistics:

$$\textit{Hong}: \frac{\mathcal{T}_{\omega}^{(f)} - \mu_{\omega,T}^{(f)}}{\sqrt{D_{\omega,T}^{(f)}}}, \quad \textit{Hete}: \frac{\mathcal{T}_{\omega}^{(f),c} - \mu_{\omega,T}^{(f)}}{\sqrt{\widehat{D}_{\omega,T}^{(f),(Hete)}}}, \quad \textit{Hete2}: \frac{\mathcal{T}_{1\omega}^{(f)} - \mu_{\omega,T}^{(f)}}{\sqrt{D_{\omega,T}^{(f)}}} + \frac{\mathcal{T}_{2\omega}^{(f),c}}{\sqrt{\widehat{D}_{\omega,T}^{(f),(Hete)}}}$$

where:

$$\mathcal{T}_{\omega}^{(f)} = \mathcal{T}_{1\omega}^{(f)} + \mathcal{T}_{2\omega}^{(f)}, \quad \mathcal{T}_{\omega}^{(f),c} = \mathcal{T}_{1\omega}^{(f)} + \mathcal{T}_{2\omega}^{(f),c}$$

$$\mathcal{T}_{1\omega}^{(f)} = \sum_{j=1}^{T-1} \omega(j) \frac{1}{T-j} \sum_{t=j+1}^{T} ||X_t||^2 ||Z_{t-j}||^2$$

$$\mathcal{T}_{2\omega}^{(f)} = \sum_{j=1}^{T-2} \omega(j) \frac{1}{T-j} \sum_{s,t=j+1, s\neq t}^{T} < X_t, X_s >< Z_{t-j}, Z_{s-j} >$$

$$\mathcal{T}_{2\omega}^{(f),c} = \sum_{j=1}^{T-2} \omega(j) \frac{1}{T-j} \sum_{s,t=j+1, s\neq t, s>t-j} < X_t, X_s >< Z_{t-j}, Z_{s-j} >$$

with:

$$\mu_{\omega,T}^{(f)} = d_1 d_2 \sum_{j=1}^{T-1}\left(1-\frac{j}{T-j}\right)\omega(j)$$

$$D_{\omega,T}^{(f)} = 2d_1^2 d_2^2 \sum_{j=1}^{T-2}\left(1-\frac{j}{T-j}\right)\left(1-\frac{j+1}{T-j}\right)\omega^2(j)$$

$$\widehat{D}_{\omega,T}^{(f),(Hete)} = 2d_1^2 \sum_{j,\ell=1}^{T-2}\omega(j)\omega(\ell)\widehat{\Xi}_{j\ell}$$

$$\widehat{\Xi}_{j\ell} = \begin{cases} \frac{2}{(T-j)^2}\sum_{s=1}^{j-1}\frac{T-s-j}{T-s}\sum_{t=j+s+1}^{T}(\langle Z_{t-j}, Z_{t-j-s}\rangle)^2 & j=\ell \\ \frac{1}{(T-m)^2}\sum_{s=1}^{\min\{j,\ell\}-1}\frac{T-m-s-\delta}{T-m-s} \\ \qquad \sum_{t=m+\delta+1}^{T}\langle Z_{t-j}^{(w)}, Z_{t-j-s}^{(w)}\rangle & j\neq\ell \end{cases}$$

$$Z_{l,s}^{(w)} = \min\{\max\{Z_{l,s}, -\varpi\cdot 1.4826\cdot\widehat{\sigma}_l\}, \varpi\cdot 1.4826\cdot\widehat{\sigma}_l\}, \quad l=1,...,d_2$$

with $m=\max\{j,\ell\}$, $\delta=|j-\ell|$, and $\widehat{\sigma}_l$ being the sample median absolute deviation of $\{Z_{l,t}\}$. Alternative robustifications could be used without affecting the asymptotic arguments. The baseline simulations use $\varpi = 0$. Note that the finite-sample corrected statistics are already scaled by the (effective) sample size.

The first two are the finite-sample corrected statistics proposed in this paper (Eq.(5)); the third is the benchmark (Eq.(2)). The variants *Hete* and *Hete2* differ only in the standardization of $\mathcal{T}_{1\omega}^{(f)}$ (sum of squares), which affects asymptotic power but not size (Theorem 1-3). The remaining three statistics are:

- *Hete2F*: The corrected statistic (*Hete2*) applied to cross-correlations between residuals $\hat{u}_{x,t}$ and $Z_t$, where $\hat{u}_{x,t} = X_t - \hat{a}X_{t-1}$ with $\hat{a}$ the least squares estimator from an AR(1) regression of $X$ on its own lag.
- *Wald-single* and *Wald-double*.
  Heteroskedastic-consistent Wald statistics from a VAR(1) fitted to the joint process:

$$\begin{pmatrix} X_t \\ Z_t \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}\begin{pmatrix} X_{t-1} \\ Z_{t-1} \end{pmatrix} + \begin{pmatrix} u_{x,t} \\ u_{z,t} \end{pmatrix}$$

constructed following Lütkepohl (2013) (Section 3.6, Eq. 3.6.4). *Wald-single* tests $\mathcal{H}_0^{single}: b_{12} = 0$, *Wald-double* tests $\mathcal{H}_0^{double}: b_{11} = 0, b_{12} = 0$.

The weighting function for the Portmanteau-type statistics is the Bartlett kernel: $\omega(j) = |j/M|$, for $|j/M| \leq 1$, and $\omega(j) = 0$, for $|j/M| > 1$.

## C.2 DGPs under the Null

We consider four DGP families where $X$ is a standardized Strong White Noise (SWN):

$$X_t = \epsilon_x, \quad \epsilon_x \sim \text{i.i.d.}(0, 1)$$

ensuring weak exogeneity holds by construction. Following Lemma A.4, the DGPs vary along two dimensions: (i) the inverse causality channel from past $X$ to present $Z$, and (ii) the conditional mean and variance properties of $Z$. The four specifications for the univariate process $Z$ are:

a) DGP 1A (LINEAR-IN-MEAN):

$$Z_t = \alpha Z_{t-1} + \beta_1 X_{t-1} + \epsilon_z, \quad \epsilon_z \sim \text{i.i.d.}(0, 1)$$

b) DGP 2A (NONLINEAR-IN-MEAN):

$$Z_t = \alpha Z_{t-1} + \beta_1 X_{t-1} + \sum_{h=1}^{2} \frac{\beta_1}{2h} X_{t-1}^{2h} + \epsilon_z, \quad \epsilon_z \sim \text{i.i.d.}(0, 1)$$

c) DGP 3A (ARCH-TYPE INVERSE CAUSALITY):

$$Z_t = \alpha Z_{t-1} + \sigma_{z,t}\epsilon_z, \quad \epsilon_z \sim \text{i.i.d.}(0, 1)$$
$$\sigma_{z,t}^2 = 1 + \sum_{h=1}^{3} \frac{\beta_1}{h!} X_{t-h}^2$$

d) DGP 4A (GARCH-TYPE INVERSE CAUSALITY):

$$Z_t = \sigma_{z,t}\epsilon_z, \quad \epsilon_z \sim \text{i.i.d.}(0,1)$$
$$\sigma^2_{z,t} = 1 + \alpha\sigma^2_{z,t-1} + \beta_2 X^2_{t-1}$$

Parameter values span: $\alpha \in \{0, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7\}$, $\beta_1 \in \{0, 0.2, 0.4, 0.6, 1.0\}$, $\beta_2 \in \{0.05, 0.1, 0.2, 0.3\}$. When $\beta_1 = 0$ the two processes are mutually independent. When $\beta_1 \neq 0$ ($\beta_2 \neq 0$) inverse causality is present while weak exogeneity is maintained.

To examine the magnitude of the correction term, given that higher-order moments of $X$ are relevant (Proposition 2), innovations are drawn from a multivariate $t$-distribution with 6 degrees of freedom: $(\epsilon_x, \epsilon_z) \sim t_6(0, I_2)$.[16] This choice aligns with the underlying assumption of the empirical analysis of this paper (Section 4), and with empirical evidence from macroeconomic applications. For instance, Brunnermeier et al. (2021) estimate structural shocks identified through heteroskedasticity in SVARs as scaled $t$-variates with 5.7 degrees of freedom, thus motivating our calibration. The smoothing parameter $M$ takes value 12 or 36, which corresponds to lags up to 1 year or 3 years, when the data is observed at monthly frequency. For each design, 1,000 Monte Carlo simulations are run. All results report rejection rates at the 5% nominal significance level.

[16] Results under Gaussian innovations, $(\epsilon_x, \epsilon_z) \sim \mathcal{N}(0, I_2)$, are qualitatively similar and available upon request.

Table 2: Rejection frequencies for DGP1A (Baseline): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP1A; sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 0$; nominal significance level is 5%.

$M = 12 \quad \approx 2\ln T$

| | Hete | | | | | Hong | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.024 | 0.019 | 0.015 | 0.016 | 0.011 | 0.051 | 0.044 | 0.034 | 0.037 | 0.043 |
| $\alpha = 0.2$ | 0.018 | 0.019 | 0.019 | 0.024 | 0.010 | 0.044 | 0.051 | 0.044 | 0.050 | 0.034 |
| $\alpha = 0.3$ | 0.026 | 0.025 | 0.028 | 0.019 | 0.033 | 0.052 | 0.051 | 0.065 | 0.059 | 0.070 |
| $\alpha = 0.4$ | 0.024 | 0.034 | 0.041 | 0.030 | 0.025 | 0.052 | 0.066 | 0.060 | 0.062 | 0.054 |
| $\alpha = 0.5$ | 0.034 | 0.043 | 0.026 | 0.034 | 0.045 | 0.062 | 0.079 | 0.059 | 0.068 | 0.065 |
| $\alpha = 0.6$ | 0.034 | 0.048 | 0.060 | 0.048 | 0.048 | 0.068 | 0.079 | 0.089 | 0.085 | 0.081 |
| $\alpha = 0.7$ | 0.048 | 0.060 | 0.045 | 0.055 | 0.054 | 0.075 | 0.099 | 0.090 | 0.087 | 0.092 |

$M = 12 \quad \approx 2\ln T$

| | Hete2 | | | | | Hete2F | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.022 | 0.020 | 0.016 | 0.016 | 0.011 | 0.022 | 0.018 | 0.014 | 0.014 | 0.013 |
| $\alpha = 0.2$ | 0.018 | 0.019 | 0.018 | 0.024 | 0.010 | 0.018 | 0.020 | 0.019 | 0.021 | 0.009 |
| $\alpha = 0.3$ | 0.026 | 0.025 | 0.028 | 0.020 | 0.033 | 0.026 | 0.024 | 0.022 | 0.021 | 0.026 |
| $\alpha = 0.4$ | 0.025 | 0.035 | 0.040 | 0.028 | 0.024 | 0.021 | 0.033 | 0.041 | 0.028 | 0.022 |
| $\alpha = 0.5$ | 0.033 | 0.043 | 0.028 | 0.034 | 0.045 | 0.033 | 0.041 | 0.028 | 0.031 | 0.039 |
| $\alpha = 0.6$ | 0.033 | 0.048 | 0.058 | 0.049 | 0.046 | 0.033 | 0.043 | 0.054 | 0.051 | 0.041 |
| $\alpha = 0.7$ | 0.046 | 0.061 | 0.048 | 0.055 | 0.052 | 0.046 | 0.055 | 0.049 | 0.050 | 0.055 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete | | | | | Hong | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.013 | 0.016 | 0.013 | 0.016 | 0.013 | 0.046 | 0.049 | 0.043 | 0.045 | 0.050 |
| $\alpha = 0.2$ | 0.020 | 0.022 | 0.015 | 0.014 | 0.014 | 0.052 | 0.055 | 0.051 | 0.051 | 0.050 |
| $\alpha = 0.3$ | 0.021 | 0.016 | 0.022 | 0.025 | 0.029 | 0.058 | 0.055 | 0.058 | 0.054 | 0.068 |
| $\alpha = 0.4$ | 0.020 | 0.040 | 0.041 | 0.035 | 0.025 | 0.055 | 0.075 | 0.083 | 0.072 | 0.056 |
| $\alpha = 0.5$ | 0.040 | 0.041 | 0.031 | 0.040 | 0.040 | 0.084 | 0.094 | 0.079 | 0.090 | 0.087 |
| $\alpha = 0.6$ | 0.044 | 0.052 | 0.065 | 0.051 | 0.065 | 0.090 | 0.102 | 0.129 | 0.115 | 0.114 |
| $\alpha = 0.7$ | 0.052 | 0.075 | 0.058 | 0.068 | 0.075 | 0.135 | 0.151 | 0.133 | 0.131 | 0.147 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete2 | | | | | Hete2F | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.014 | 0.016 | 0.013 | 0.015 | 0.010 | 0.011 | 0.018 | 0.013 | 0.019 | 0.010 |
| $\alpha = 0.2$ | 0.019 | 0.021 | 0.015 | 0.013 | 0.013 | 0.018 | 0.020 | 0.016 | 0.014 | 0.013 |
| $\alpha = 0.3$ | 0.021 | 0.015 | 0.021 | 0.022 | 0.028 | 0.016 | 0.019 | 0.022 | 0.022 | 0.028 |
| $\alpha = 0.4$ | 0.019 | 0.036 | 0.039 | 0.031 | 0.024 | 0.020 | 0.033 | 0.040 | 0.030 | 0.016 |
| $\alpha = 0.5$ | 0.039 | 0.040 | 0.031 | 0.039 | 0.039 | 0.035 | 0.043 | 0.026 | 0.040 | 0.034 |
| $\alpha = 0.6$ | 0.043 | 0.050 | 0.062 | 0.050 | 0.062 | 0.043 | 0.051 | 0.052 | 0.051 | 0.058 |
| $\alpha = 0.7$ | 0.052 | 0.075 | 0.058 | 0.065 | 0.075 | 0.055 | 0.074 | 0.056 | 0.065 | 0.064 |

| | Wald Single | | | | | Wald Double | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.043 | 0.048 | 0.051 | 0.040 | 0.048 | 0.049 | 0.044 | 0.054 | 0.040 | 0.046 |
| $\alpha = 0.2$ | 0.046 | 0.048 | 0.059 | 0.045 | 0.047 | 0.050 | 0.040 | 0.042 | 0.048 | 0.051 |
| $\alpha = 0.3$ | 0.040 | 0.059 | 0.063 | 0.040 | 0.045 | 0.038 | 0.046 | 0.063 | 0.053 | 0.049 |
| $\alpha = 0.4$ | 0.055 | 0.048 | 0.050 | 0.050 | 0.048 | 0.056 | 0.052 | 0.054 | 0.046 | 0.040 |
| $\alpha = 0.5$ | 0.044 | 0.053 | 0.054 | 0.054 | 0.056 | 0.038 | 0.054 | 0.051 | 0.043 | 0.050 |
| $\alpha = 0.6$ | 0.061 | 0.051 | 0.060 | 0.055 | 0.047 | 0.045 | 0.053 | 0.066 | 0.048 | 0.042 |
| $\alpha = 0.7$ | 0.041 | 0.042 | 0.047 | 0.054 | 0.049 | 0.049 | 0.044 | 0.052 | 0.041 | 0.052 |

Table 3: Rejection frequencies for DGP2A (Baseline): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP2A; sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 0$; nominal significance level is 5%.

$M = 12 \quad \approx 2 \ln T$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.150 | 0.166 | 0.169 | 0.165 | 0.061 | 0.066 | 0.052 | 0.044 |
| $\alpha = 0.3$ | 0.126 | 0.149 | 0.141 | 0.139 | 0.055 | 0.061 | 0.065 | 0.056 |
| $\alpha = 0.4$ | 0.106 | 0.126 | 0.136 | 0.121 | 0.056 | 0.069 | 0.065 | 0.061 |
| $\alpha = 0.5$ | 0.086 | 0.095 | 0.119 | 0.104 | 0.054 | 0.054 | 0.070 | 0.058 |
| $\alpha = 0.6$ | 0.090 | 0.083 | 0.106 | 0.091 | 0.084 | 0.068 | 0.099 | 0.077 |
| $\alpha = 0.7$ | 0.095 | 0.100 | 0.107 | 0.081 | 0.094 | 0.086 | 0.083 | 0.079 |

$M = 12 \quad \approx 2 \ln T$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.046 | 0.048 | 0.044 | 0.039 | 0.046 | 0.050 | 0.040 | 0.041 |
| $\alpha = 0.3$ | 0.048 | 0.050 | 0.055 | 0.045 | 0.049 | 0.050 | 0.055 | 0.041 |
| $\alpha = 0.4$ | 0.045 | 0.061 | 0.059 | 0.051 | 0.046 | 0.059 | 0.056 | 0.059 |
| $\alpha = 0.5$ | 0.045 | 0.045 | 0.061 | 0.039 | 0.044 | 0.039 | 0.058 | 0.041 |
| $\alpha = 0.6$ | 0.072 | 0.050 | 0.071 | 0.058 | 0.072 | 0.052 | 0.070 | 0.055 |
| $\alpha = 0.7$ | 0.077 | 0.079 | 0.072 | 0.060 | 0.079 | 0.075 | 0.075 | 0.062 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.169 | 0.196 | 0.191 | 0.219 | 0.076 | 0.070 | 0.060 | 0.060 |
| $\alpha = 0.3$ | 0.150 | 0.171 | 0.180 | 0.160 | 0.076 | 0.080 | 0.079 | 0.060 |
| $\alpha = 0.4$ | 0.139 | 0.150 | 0.158 | 0.163 | 0.087 | 0.081 | 0.076 | 0.085 |
| $\alpha = 0.5$ | 0.135 | 0.129 | 0.152 | 0.151 | 0.090 | 0.080 | 0.087 | 0.090 |
| $\alpha = 0.6$ | 0.134 | 0.152 | 0.158 | 0.135 | 0.111 | 0.116 | 0.116 | 0.120 |
| $\alpha = 0.7$ | 0.168 | 0.161 | 0.185 | 0.160 | 0.170 | 0.164 | 0.166 | 0.151 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.069 | 0.062 | 0.065 | 0.056 | 0.069 | 0.061 | 0.059 | 0.051 |
| $\alpha = 0.3$ | 0.061 | 0.064 | 0.060 | 0.055 | 0.065 | 0.068 | 0.064 | 0.052 |
| $\alpha = 0.4$ | 0.061 | 0.072 | 0.064 | 0.074 | 0.055 | 0.071 | 0.056 | 0.066 |
| $\alpha = 0.5$ | 0.066 | 0.070 | 0.091 | 0.074 | 0.074 | 0.065 | 0.086 | 0.074 |
| $\alpha = 0.6$ | 0.091 | 0.089 | 0.094 | 0.091 | 0.085 | 0.094 | 0.098 | 0.089 |
| $\alpha = 0.7$ | 0.131 | 0.126 | 0.130 | 0.121 | 0.126 | 0.124 | 0.126 | 0.121 |

| | Wald Single | | | | Wald Double | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.045 | 0.055 | 0.056 | 0.050 | 0.049 | 0.050 | 0.044 | 0.046 |
| $\alpha = 0.3$ | 0.040 | 0.053 | 0.050 | 0.046 | 0.038 | 0.041 | 0.038 | 0.056 |
| $\alpha = 0.4$ | 0.037 | 0.046 | 0.045 | 0.025 | 0.023 | 0.038 | 0.044 | 0.035 |
| $\alpha = 0.5$ | 0.044 | 0.049 | 0.046 | 0.037 | 0.037 | 0.044 | 0.034 | 0.038 |
| $\alpha = 0.6$ | 0.030 | 0.040 | 0.048 | 0.038 | 0.034 | 0.039 | 0.045 | 0.037 |
| $\alpha = 0.7$ | 0.039 | 0.036 | 0.042 | 0.037 | 0.035 | 0.031 | 0.035 | 0.036 |

Table 4: Rejection frequencies for DGP3A (Baseline): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP3A; sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 0$; nominal significance level is 5%.

$M = 12 \quad \approx 2 \ln T$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.020 | 0.026 | 0.022 | 0.018 | 0.060 | 0.050 | 0.056 | 0.050 |
| $\alpha = 0.3$ | 0.029 | 0.026 | 0.025 | 0.028 | 0.059 | 0.070 | 0.069 | 0.043 |
| $\alpha = 0.4$ | 0.036 | 0.030 | 0.030 | 0.025 | 0.069 | 0.068 | 0.061 | 0.058 |
| $\alpha = 0.5$ | 0.037 | 0.044 | 0.044 | 0.034 | 0.064 | 0.086 | 0.068 | 0.066 |
| $\alpha = 0.6$ | 0.048 | 0.045 | 0.035 | 0.045 | 0.072 | 0.079 | 0.068 | 0.092 |
| $\alpha = 0.7$ | 0.058 | 0.056 | 0.050 | 0.054 | 0.095 | 0.092 | 0.084 | 0.090 |

$M = 12 \quad \approx 2 \ln T$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.020 | 0.028 | 0.021 | 0.018 | 0.019 | 0.028 | 0.018 | 0.014 |
| $\alpha = 0.3$ | 0.024 | 0.025 | 0.022 | 0.025 | 0.028 | 0.020 | 0.024 | 0.026 |
| $\alpha = 0.4$ | 0.035 | 0.030 | 0.030 | 0.022 | 0.035 | 0.030 | 0.028 | 0.020 |
| $\alpha = 0.5$ | 0.037 | 0.040 | 0.043 | 0.033 | 0.034 | 0.036 | 0.041 | 0.030 |
| $\alpha = 0.6$ | 0.049 | 0.043 | 0.035 | 0.046 | 0.046 | 0.044 | 0.030 | 0.041 |
| $\alpha = 0.7$ | 0.059 | 0.052 | 0.050 | 0.056 | 0.059 | 0.059 | 0.046 | 0.059 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.014 | 0.018 | 0.020 | 0.020 | 0.061 | 0.043 | 0.046 | 0.058 |
| $\alpha = 0.3$ | 0.021 | 0.021 | 0.022 | 0.022 | 0.059 | 0.055 | 0.056 | 0.051 |
| $\alpha = 0.4$ | 0.034 | 0.034 | 0.022 | 0.041 | 0.080 | 0.075 | 0.061 | 0.081 |
| $\alpha = 0.5$ | 0.051 | 0.037 | 0.036 | 0.044 | 0.085 | 0.090 | 0.089 | 0.091 |
| $\alpha = 0.6$ | 0.055 | 0.061 | 0.045 | 0.060 | 0.101 | 0.115 | 0.117 | 0.119 |
| $\alpha = 0.7$ | 0.071 | 0.083 | 0.059 | 0.059 | 0.131 | 0.159 | 0.124 | 0.149 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.013 | 0.018 | 0.018 | 0.019 | 0.013 | 0.016 | 0.014 | 0.019 |
| $\alpha = 0.3$ | 0.021 | 0.021 | 0.022 | 0.022 | 0.021 | 0.020 | 0.025 | 0.025 |
| $\alpha = 0.4$ | 0.033 | 0.031 | 0.020 | 0.037 | 0.035 | 0.025 | 0.021 | 0.035 |
| $\alpha = 0.5$ | 0.050 | 0.037 | 0.035 | 0.041 | 0.045 | 0.040 | 0.031 | 0.039 |
| $\alpha = 0.6$ | 0.055 | 0.061 | 0.043 | 0.059 | 0.052 | 0.054 | 0.041 | 0.060 |
| $\alpha = 0.7$ | 0.071 | 0.083 | 0.056 | 0.059 | 0.075 | 0.085 | 0.051 | 0.055 |

| | Wald Single | | | | Wald Double | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.056 | 0.050 | 0.064 | 0.056 | 0.059 | 0.052 | 0.058 | 0.059 |
| $\alpha = 0.3$ | 0.052 | 0.048 | 0.055 | 0.050 | 0.057 | 0.048 | 0.055 | 0.048 |
| $\alpha = 0.4$ | 0.029 | 0.048 | 0.053 | 0.056 | 0.041 | 0.038 | 0.053 | 0.045 |
| $\alpha = 0.5$ | 0.038 | 0.038 | 0.044 | 0.044 | 0.048 | 0.043 | 0.051 | 0.042 |
| $\alpha = 0.6$ | 0.061 | 0.054 | 0.047 | 0.056 | 0.074 | 0.045 | 0.052 | 0.061 |
| $\alpha = 0.7$ | 0.052 | 0.035 | 0.040 | 0.056 | 0.051 | 0.040 | 0.032 | 0.050 |

Table 5: Rejection frequencies for DGP4A (Baseline): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP4A; sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 0$; nominal significance level is 5%.

$M = 12 \quad \approx 2\ln T$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.016 | 0.016 | 0.013 | 0.019 | 0.045 | 0.052 | 0.037 | 0.050 |
| $\alpha = 0.3$ | 0.025 | 0.018 | 0.016 | 0.013 | 0.058 | 0.049 | 0.044 | 0.036 |
| $\alpha = 0.4$ | 0.014 | 0.016 | 0.020 | 0.020 | 0.050 | 0.048 | 0.044 | 0.059 |
| $\alpha = 0.5$ | 0.022 | 0.014 | 0.025 | 0.016 | 0.046 | 0.036 | 0.051 | 0.040 |
| $\alpha = 0.6$ | 0.019 | 0.020 | 0.020 | 0.021 | 0.051 | 0.059 | 0.052 | 0.051 |
| $\alpha = 0.7$ | 0.016 | 0.018 | 0.016 | 0.024 | 0.043 | 0.045 | 0.051 | 0.051 |

$M = 12 \quad \approx 2\ln T$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.013 | 0.019 | 0.013 | 0.020 | 0.013 | 0.019 | 0.011 | 0.019 |
| $\alpha = 0.3$ | 0.024 | 0.018 | 0.015 | 0.013 | 0.024 | 0.020 | 0.015 | 0.014 |
| $\alpha = 0.4$ | 0.014 | 0.016 | 0.020 | 0.020 | 0.013 | 0.015 | 0.021 | 0.022 |
| $\alpha = 0.5$ | 0.022 | 0.014 | 0.026 | 0.015 | 0.021 | 0.016 | 0.026 | 0.014 |
| $\alpha = 0.6$ | 0.018 | 0.020 | 0.020 | 0.021 | 0.016 | 0.020 | 0.020 | 0.021 |
| $\alpha = 0.7$ | 0.016 | 0.018 | 0.016 | 0.024 | 0.015 | 0.016 | 0.015 | 0.025 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.013 | 0.019 | 0.011 | 0.013 | 0.050 | 0.050 | 0.040 | 0.049 |
| $\alpha = 0.3$ | 0.015 | 0.022 | 0.011 | 0.010 | 0.049 | 0.046 | 0.039 | 0.033 |
| $\alpha = 0.4$ | 0.009 | 0.011 | 0.015 | 0.010 | 0.048 | 0.051 | 0.059 | 0.050 |
| $\alpha = 0.5$ | 0.015 | 0.011 | 0.015 | 0.015 | 0.050 | 0.041 | 0.059 | 0.049 |
| $\alpha = 0.6$ | 0.015 | 0.015 | 0.013 | 0.019 | 0.035 | 0.048 | 0.046 | 0.050 |
| $\alpha = 0.7$ | 0.006 | 0.013 | 0.011 | 0.016 | 0.039 | 0.046 | 0.043 | 0.049 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.013 | 0.016 | 0.011 | 0.011 | 0.011 | 0.013 | 0.010 | 0.015 |
| $\alpha = 0.3$ | 0.015 | 0.022 | 0.013 | 0.009 | 0.015 | 0.022 | 0.011 | 0.009 |
| $\alpha = 0.4$ | 0.009 | 0.011 | 0.015 | 0.010 | 0.011 | 0.010 | 0.015 | 0.009 |
| $\alpha = 0.5$ | 0.015 | 0.011 | 0.015 | 0.015 | 0.015 | 0.011 | 0.014 | 0.014 |
| $\alpha = 0.6$ | 0.015 | 0.015 | 0.013 | 0.016 | 0.014 | 0.015 | 0.014 | 0.018 |
| $\alpha = 0.7$ | 0.006 | 0.011 | 0.011 | 0.016 | 0.007 | 0.011 | 0.013 | 0.018 |

| | Wald Single | | | | Wald Double | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.058 | 0.060 | 0.040 | 0.070 | 0.046 | 0.048 | 0.044 | 0.050 |
| $\alpha = 0.3$ | 0.046 | 0.048 | 0.042 | 0.042 | 0.044 | 0.054 | 0.038 | 0.048 |
| $\alpha = 0.4$ | 0.036 | 0.030 | 0.040 | 0.072 | 0.032 | 0.050 | 0.058 | 0.068 |
| $\alpha = 0.5$ | 0.050 | 0.046 | 0.050 | 0.054 | 0.042 | 0.038 | 0.056 | 0.072 |
| $\alpha = 0.6$ | 0.046 | 0.044 | 0.042 | 0.058 | 0.042 | 0.054 | 0.048 | 0.050 |
| $\alpha = 0.7$ | 0.048 | 0.050 | 0.064 | 0.042 | 0.054 | 0.050 | 0.068 | 0.056 |

Table 6: Rejection frequencies for DGP1A ($\varpi = 3$): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP1A; sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 3$; nominal significance level is 5%.

$M = 12 \quad \approx 2\ln T$

| | Hete | | | | | Hong | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.024 | 0.019 | 0.015 | 0.016 | 0.011 | 0.051 | 0.044 | 0.034 | 0.037 | 0.043 |
| $\alpha = 0.2$ | 0.018 | 0.019 | 0.019 | 0.024 | 0.010 | 0.044 | 0.051 | 0.044 | 0.050 | 0.034 |
| $\alpha = 0.3$ | 0.026 | 0.025 | 0.028 | 0.019 | 0.033 | 0.052 | 0.051 | 0.065 | 0.059 | 0.070 |
| $\alpha = 0.4$ | 0.024 | 0.034 | 0.041 | 0.030 | 0.025 | 0.052 | 0.066 | 0.060 | 0.062 | 0.054 |
| $\alpha = 0.5$ | 0.034 | 0.043 | 0.026 | 0.034 | 0.045 | 0.062 | 0.079 | 0.059 | 0.068 | 0.065 |
| $\alpha = 0.6$ | 0.034 | 0.048 | 0.060 | 0.048 | 0.048 | 0.068 | 0.079 | 0.089 | 0.085 | 0.081 |
| $\alpha = 0.7$ | 0.048 | 0.060 | 0.045 | 0.055 | 0.054 | 0.075 | 0.099 | 0.090 | 0.087 | 0.092 |

$M = 12 \quad \approx 2\ln T$

| | Hete2 | | | | | Hete2F | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.022 | 0.020 | 0.016 | 0.016 | 0.011 | 0.022 | 0.018 | 0.014 | 0.014 | 0.013 |
| $\alpha = 0.2$ | 0.018 | 0.019 | 0.018 | 0.024 | 0.010 | 0.018 | 0.020 | 0.019 | 0.021 | 0.009 |
| $\alpha = 0.3$ | 0.026 | 0.025 | 0.028 | 0.020 | 0.033 | 0.026 | 0.024 | 0.022 | 0.021 | 0.026 |
| $\alpha = 0.4$ | 0.025 | 0.035 | 0.040 | 0.028 | 0.024 | 0.021 | 0.033 | 0.041 | 0.028 | 0.022 |
| $\alpha = 0.5$ | 0.033 | 0.043 | 0.028 | 0.034 | 0.045 | 0.033 | 0.041 | 0.028 | 0.031 | 0.039 |
| $\alpha = 0.6$ | 0.033 | 0.048 | 0.058 | 0.049 | 0.046 | 0.033 | 0.043 | 0.054 | 0.051 | 0.041 |
| $\alpha = 0.7$ | 0.046 | 0.061 | 0.048 | 0.055 | 0.052 | 0.046 | 0.055 | 0.049 | 0.050 | 0.055 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete | | | | | Hong | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.013 | 0.016 | 0.013 | 0.016 | 0.013 | 0.046 | 0.049 | 0.043 | 0.045 | 0.050 |
| $\alpha = 0.2$ | 0.020 | 0.022 | 0.015 | 0.014 | 0.014 | 0.052 | 0.055 | 0.051 | 0.051 | 0.050 |
| $\alpha = 0.3$ | 0.021 | 0.016 | 0.022 | 0.025 | 0.029 | 0.058 | 0.055 | 0.058 | 0.054 | 0.068 |
| $\alpha = 0.4$ | 0.020 | 0.040 | 0.041 | 0.035 | 0.025 | 0.055 | 0.075 | 0.083 | 0.072 | 0.056 |
| $\alpha = 0.5$ | 0.040 | 0.041 | 0.031 | 0.040 | 0.040 | 0.084 | 0.094 | 0.079 | 0.090 | 0.087 |
| $\alpha = 0.6$ | 0.044 | 0.052 | 0.065 | 0.051 | 0.065 | 0.090 | 0.102 | 0.129 | 0.115 | 0.114 |
| $\alpha = 0.7$ | 0.052 | 0.075 | 0.058 | 0.068 | 0.075 | 0.135 | 0.151 | 0.133 | 0.131 | 0.147 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete2 | | | | | Hete2F | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.014 | 0.016 | 0.013 | 0.015 | 0.010 | 0.011 | 0.018 | 0.013 | 0.019 | 0.010 |
| $\alpha = 0.2$ | 0.019 | 0.021 | 0.015 | 0.013 | 0.013 | 0.018 | 0.020 | 0.016 | 0.014 | 0.013 |
| $\alpha = 0.3$ | 0.021 | 0.015 | 0.021 | 0.022 | 0.028 | 0.016 | 0.019 | 0.022 | 0.022 | 0.028 |
| $\alpha = 0.4$ | 0.019 | 0.036 | 0.039 | 0.031 | 0.024 | 0.020 | 0.033 | 0.040 | 0.030 | 0.016 |
| $\alpha = 0.5$ | 0.039 | 0.040 | 0.031 | 0.039 | 0.039 | 0.035 | 0.043 | 0.026 | 0.040 | 0.034 |
| $\alpha = 0.6$ | 0.043 | 0.050 | 0.062 | 0.050 | 0.062 | 0.043 | 0.051 | 0.052 | 0.051 | 0.058 |
| $\alpha = 0.7$ | 0.052 | 0.075 | 0.058 | 0.065 | 0.075 | 0.055 | 0.074 | 0.056 | 0.065 | 0.064 |

| | Wald Single | | | | | Wald Double | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.043 | 0.048 | 0.051 | 0.040 | 0.048 | 0.049 | 0.044 | 0.054 | 0.040 | 0.046 |
| $\alpha = 0.2$ | 0.046 | 0.048 | 0.059 | 0.045 | 0.047 | 0.050 | 0.040 | 0.042 | 0.048 | 0.051 |
| $\alpha = 0.3$ | 0.040 | 0.059 | 0.063 | 0.040 | 0.045 | 0.038 | 0.046 | 0.063 | 0.053 | 0.049 |
| $\alpha = 0.4$ | 0.055 | 0.048 | 0.050 | 0.050 | 0.048 | 0.056 | 0.052 | 0.054 | 0.046 | 0.040 |
| $\alpha = 0.5$ | 0.044 | 0.053 | 0.054 | 0.054 | 0.056 | 0.038 | 0.054 | 0.051 | 0.043 | 0.050 |
| $\alpha = 0.6$ | 0.061 | 0.051 | 0.060 | 0.055 | 0.047 | 0.045 | 0.053 | 0.066 | 0.048 | 0.042 |
| $\alpha = 0.7$ | 0.041 | 0.042 | 0.047 | 0.054 | 0.049 | 0.049 | 0.044 | 0.052 | 0.041 | 0.052 |

Table 7: Rejection frequencies for DGP2A ($\varpi = 3$): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP2A; sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 3$; nominal significance level is 5%.

$M = 12$ $\approx 2 \ln T$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.150 | 0.168 | 0.169 | 0.166 | 0.061 | 0.066 | 0.052 | 0.044 |
| $\alpha = 0.3$ | 0.126 | 0.149 | 0.141 | 0.140 | 0.055 | 0.061 | 0.065 | 0.056 |
| $\alpha = 0.4$ | 0.106 | 0.126 | 0.136 | 0.121 | 0.056 | 0.069 | 0.065 | 0.061 |
| $\alpha = 0.5$ | 0.083 | 0.095 | 0.119 | 0.105 | 0.054 | 0.054 | 0.070 | 0.058 |
| $\alpha = 0.6$ | 0.086 | 0.080 | 0.105 | 0.090 | 0.084 | 0.068 | 0.099 | 0.077 |
| $\alpha = 0.7$ | 0.089 | 0.099 | 0.106 | 0.080 | 0.094 | 0.086 | 0.083 | 0.079 |

$M = 12$ $\approx 2 \ln T$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.045 | 0.048 | 0.044 | 0.039 | 0.046 | 0.050 | 0.040 | 0.041 |
| $\alpha = 0.3$ | 0.048 | 0.050 | 0.055 | 0.045 | 0.049 | 0.050 | 0.055 | 0.041 |
| $\alpha = 0.4$ | 0.045 | 0.061 | 0.059 | 0.051 | 0.046 | 0.059 | 0.056 | 0.059 |
| $\alpha = 0.5$ | 0.045 | 0.045 | 0.061 | 0.039 | 0.044 | 0.039 | 0.058 | 0.041 |
| $\alpha = 0.6$ | 0.071 | 0.049 | 0.071 | 0.056 | 0.070 | 0.051 | 0.070 | 0.055 |
| $\alpha = 0.7$ | 0.074 | 0.077 | 0.071 | 0.060 | 0.076 | 0.074 | 0.072 | 0.061 |

$M = 36$ $\approx \sqrt{T}$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.169 | 0.196 | 0.193 | 0.220 | 0.076 | 0.070 | 0.060 | 0.060 |
| $\alpha = 0.3$ | 0.150 | 0.171 | 0.184 | 0.164 | 0.076 | 0.080 | 0.079 | 0.060 |
| $\alpha = 0.4$ | 0.133 | 0.150 | 0.159 | 0.166 | 0.087 | 0.081 | 0.076 | 0.085 |
| $\alpha = 0.5$ | 0.128 | 0.129 | 0.154 | 0.158 | 0.090 | 0.080 | 0.087 | 0.090 |
| $\alpha = 0.6$ | 0.124 | 0.149 | 0.158 | 0.135 | 0.111 | 0.116 | 0.116 | 0.120 |
| $\alpha = 0.7$ | 0.144 | 0.158 | 0.184 | 0.164 | 0.170 | 0.164 | 0.166 | 0.151 |

$M = 36$ $\approx \sqrt{T}$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.069 | 0.062 | 0.065 | 0.056 | 0.069 | 0.061 | 0.059 | 0.051 |
| $\alpha = 0.3$ | 0.061 | 0.064 | 0.061 | 0.055 | 0.065 | 0.069 | 0.064 | 0.052 |
| $\alpha = 0.4$ | 0.060 | 0.071 | 0.064 | 0.075 | 0.054 | 0.071 | 0.059 | 0.066 |
| $\alpha = 0.5$ | 0.065 | 0.069 | 0.094 | 0.076 | 0.068 | 0.064 | 0.086 | 0.074 |
| $\alpha = 0.6$ | 0.089 | 0.091 | 0.094 | 0.091 | 0.081 | 0.094 | 0.098 | 0.089 |
| $\alpha = 0.7$ | 0.111 | 0.128 | 0.128 | 0.124 | 0.105 | 0.119 | 0.121 | 0.122 |

| | Wald Single | | | | Wald Double | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.045 | 0.055 | 0.056 | 0.050 | 0.049 | 0.050 | 0.044 | 0.046 |
| $\alpha = 0.3$ | 0.040 | 0.053 | 0.050 | 0.046 | 0.038 | 0.041 | 0.038 | 0.056 |
| $\alpha = 0.4$ | 0.037 | 0.046 | 0.045 | 0.025 | 0.023 | 0.038 | 0.044 | 0.035 |
| $\alpha = 0.5$ | 0.044 | 0.049 | 0.046 | 0.037 | 0.037 | 0.044 | 0.034 | 0.038 |
| $\alpha = 0.6$ | 0.030 | 0.040 | 0.048 | 0.038 | 0.034 | 0.039 | 0.045 | 0.037 |
| $\alpha = 0.7$ | 0.039 | 0.036 | 0.042 | 0.037 | 0.035 | 0.031 | 0.035 | 0.036 |

Table 8: Rejection frequencies for DGP3A ($\varpi = 3$): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP3A; sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 3$; nominal significance level is 5%.

$M = 12 \quad \approx 2 \ln T$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.019 | 0.026 | 0.021 | 0.015 | 0.060 | 0.050 | 0.056 | 0.050 |
| $\alpha = 0.3$ | 0.024 | 0.019 | 0.020 | 0.022 | 0.059 | 0.070 | 0.069 | 0.043 |
| $\alpha = 0.4$ | 0.029 | 0.019 | 0.021 | 0.018 | 0.069 | 0.068 | 0.061 | 0.058 |
| $\alpha = 0.5$ | 0.024 | 0.021 | 0.022 | 0.024 | 0.064 | 0.086 | 0.068 | 0.066 |
| $\alpha = 0.6$ | 0.025 | 0.025 | 0.011 | 0.020 | 0.072 | 0.079 | 0.068 | 0.092 |
| $\alpha = 0.7$ | 0.022 | 0.033 | 0.022 | 0.025 | 0.095 | 0.092 | 0.084 | 0.090 |

$M = 12 \quad \approx 2 \ln T$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.019 | 0.026 | 0.020 | 0.013 | 0.016 | 0.026 | 0.016 | 0.014 |
| $\alpha = 0.3$ | 0.022 | 0.019 | 0.021 | 0.022 | 0.024 | 0.016 | 0.021 | 0.025 |
| $\alpha = 0.4$ | 0.029 | 0.020 | 0.021 | 0.018 | 0.030 | 0.016 | 0.020 | 0.016 |
| $\alpha = 0.5$ | 0.024 | 0.021 | 0.021 | 0.024 | 0.020 | 0.021 | 0.024 | 0.028 |
| $\alpha = 0.6$ | 0.022 | 0.025 | 0.011 | 0.020 | 0.022 | 0.026 | 0.015 | 0.024 |
| $\alpha = 0.7$ | 0.022 | 0.033 | 0.022 | 0.025 | 0.026 | 0.033 | 0.021 | 0.026 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.011 | 0.015 | 0.019 | 0.020 | 0.061 | 0.043 | 0.046 | 0.058 |
| $\alpha = 0.3$ | 0.019 | 0.018 | 0.016 | 0.019 | 0.059 | 0.055 | 0.056 | 0.051 |
| $\alpha = 0.4$ | 0.025 | 0.020 | 0.016 | 0.020 | 0.080 | 0.075 | 0.061 | 0.081 |
| $\alpha = 0.5$ | 0.033 | 0.028 | 0.016 | 0.025 | 0.085 | 0.090 | 0.089 | 0.091 |
| $\alpha = 0.6$ | 0.019 | 0.026 | 0.014 | 0.029 | 0.101 | 0.115 | 0.117 | 0.119 |
| $\alpha = 0.7$ | 0.019 | 0.028 | 0.021 | 0.021 | 0.131 | 0.159 | 0.124 | 0.149 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.010 | 0.014 | 0.016 | 0.019 | 0.010 | 0.014 | 0.013 | 0.015 |
| $\alpha = 0.3$ | 0.018 | 0.018 | 0.015 | 0.019 | 0.015 | 0.016 | 0.016 | 0.016 |
| $\alpha = 0.4$ | 0.021 | 0.021 | 0.016 | 0.018 | 0.021 | 0.016 | 0.016 | 0.024 |
| $\alpha = 0.5$ | 0.034 | 0.026 | 0.018 | 0.026 | 0.022 | 0.020 | 0.022 | 0.021 |
| $\alpha = 0.6$ | 0.019 | 0.026 | 0.014 | 0.028 | 0.021 | 0.026 | 0.018 | 0.024 |
| $\alpha = 0.7$ | 0.019 | 0.029 | 0.020 | 0.022 | 0.021 | 0.024 | 0.018 | 0.026 |

| | Wald Single | | | | Wald Double | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.056 | 0.050 | 0.064 | 0.056 | 0.059 | 0.052 | 0.058 | 0.059 |
| $\alpha = 0.3$ | 0.052 | 0.048 | 0.055 | 0.050 | 0.057 | 0.048 | 0.055 | 0.048 |
| $\alpha = 0.4$ | 0.029 | 0.048 | 0.053 | 0.056 | 0.041 | 0.038 | 0.053 | 0.045 |
| $\alpha = 0.5$ | 0.038 | 0.038 | 0.044 | 0.044 | 0.048 | 0.043 | 0.051 | 0.042 |
| $\alpha = 0.6$ | 0.061 | 0.054 | 0.047 | 0.056 | 0.074 | 0.045 | 0.052 | 0.061 |
| $\alpha = 0.7$ | 0.052 | 0.035 | 0.040 | 0.056 | 0.051 | 0.040 | 0.032 | 0.050 |

Table 9: Rejection frequencies for DGP4A ($\varpi = 3$): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP4A; sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 3$; nominal significance level is 5%.

$M = 12 \quad \approx 2\ln T$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.016 | 0.018 | 0.013 | 0.019 | 0.045 | 0.052 | 0.037 | 0.050 |
| $\alpha = 0.3$ | 0.025 | 0.018 | 0.016 | 0.013 | 0.058 | 0.049 | 0.044 | 0.036 |
| $\alpha = 0.4$ | 0.014 | 0.016 | 0.020 | 0.020 | 0.050 | 0.048 | 0.044 | 0.059 |
| $\alpha = 0.5$ | 0.022 | 0.014 | 0.025 | 0.016 | 0.046 | 0.036 | 0.051 | 0.040 |
| $\alpha = 0.6$ | 0.019 | 0.020 | 0.020 | 0.021 | 0.051 | 0.059 | 0.052 | 0.051 |
| $\alpha = 0.7$ | 0.016 | 0.018 | 0.016 | 0.024 | 0.043 | 0.045 | 0.051 | 0.051 |

$M = 12 \quad \approx 2\ln T$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.014 | 0.019 | 0.013 | 0.020 | 0.013 | 0.019 | 0.011 | 0.019 |
| $\alpha = 0.3$ | 0.024 | 0.018 | 0.015 | 0.013 | 0.024 | 0.020 | 0.015 | 0.014 |
| $\alpha = 0.4$ | 0.014 | 0.016 | 0.020 | 0.020 | 0.013 | 0.015 | 0.021 | 0.021 |
| $\alpha = 0.5$ | 0.022 | 0.014 | 0.026 | 0.015 | 0.021 | 0.016 | 0.026 | 0.014 |
| $\alpha = 0.6$ | 0.016 | 0.020 | 0.020 | 0.021 | 0.016 | 0.020 | 0.020 | 0.021 |
| $\alpha = 0.7$ | 0.016 | 0.018 | 0.016 | 0.024 | 0.015 | 0.016 | 0.015 | 0.025 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.013 | 0.019 | 0.011 | 0.013 | 0.050 | 0.050 | 0.040 | 0.049 |
| $\alpha = 0.3$ | 0.015 | 0.022 | 0.011 | 0.010 | 0.049 | 0.046 | 0.039 | 0.033 |
| $\alpha = 0.4$ | 0.009 | 0.011 | 0.015 | 0.010 | 0.048 | 0.051 | 0.059 | 0.050 |
| $\alpha = 0.5$ | 0.015 | 0.013 | 0.015 | 0.015 | 0.050 | 0.041 | 0.059 | 0.049 |
| $\alpha = 0.6$ | 0.015 | 0.015 | 0.013 | 0.019 | 0.035 | 0.048 | 0.046 | 0.050 |
| $\alpha = 0.7$ | 0.006 | 0.013 | 0.011 | 0.016 | 0.039 | 0.046 | 0.043 | 0.049 |

$M = 36 \quad \approx \sqrt{T}$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.013 | 0.016 | 0.011 | 0.011 | 0.011 | 0.013 | 0.010 | 0.015 |
| $\alpha = 0.3$ | 0.015 | 0.022 | 0.013 | 0.009 | 0.015 | 0.024 | 0.011 | 0.009 |
| $\alpha = 0.4$ | 0.009 | 0.011 | 0.015 | 0.010 | 0.011 | 0.010 | 0.015 | 0.009 |
| $\alpha = 0.5$ | 0.014 | 0.011 | 0.015 | 0.015 | 0.014 | 0.011 | 0.014 | 0.014 |
| $\alpha = 0.6$ | 0.015 | 0.015 | 0.013 | 0.016 | 0.014 | 0.014 | 0.014 | 0.016 |
| $\alpha = 0.7$ | 0.006 | 0.011 | 0.011 | 0.016 | 0.007 | 0.011 | 0.013 | 0.018 |

| | Wald Single | | | | Wald Double | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.058 | 0.060 | 0.040 | 0.070 | 0.046 | 0.048 | 0.044 | 0.050 |
| $\alpha = 0.3$ | 0.046 | 0.048 | 0.042 | 0.042 | 0.044 | 0.054 | 0.038 | 0.048 |
| $\alpha = 0.4$ | 0.036 | 0.030 | 0.040 | 0.072 | 0.032 | 0.050 | 0.058 | 0.068 |
| $\alpha = 0.5$ | 0.050 | 0.046 | 0.050 | 0.054 | 0.042 | 0.038 | 0.056 | 0.072 |
| $\alpha = 0.6$ | 0.046 | 0.044 | 0.042 | 0.058 | 0.042 | 0.054 | 0.048 | 0.050 |
| $\alpha = 0.7$ | 0.048 | 0.050 | 0.064 | 0.042 | 0.054 | 0.050 | 0.068 | 0.056 |

## C.3 DGPs under the Alternatives

We consider four DGP families where weak exogeneity does not hold such that past $Z$ impacts present $X$, meaning that, in macroeconometric terms, the process $X$ cannot be considered as structural shocks. The four specifications for the bivariate process $\{X_t, Z_t\}$ are:

a) DGP 1B (LINEAR-IN-MEAN):

$$X_t = \gamma_1 Z_{t-1} + \epsilon_x, \quad Z_t = 0.5Z_{t-1} + \beta_3 X_{t-1} + \epsilon_z, \quad (\epsilon_x, \epsilon_z) \sim \text{i.i.d.}(0, I_2)$$

b) DGP 2B (NONLINEAR-IN-MEAN):

$$X_t = \gamma_1 Z_{t-1}^2/8 + \epsilon_x, \quad Z_t = 0.4Z_{t-1} - \beta_3 X_{t-1}/2 + \epsilon_z, \quad (\epsilon_x, \epsilon_z) \sim \text{i.i.d.}(0, I_2)$$

c) DGP 3B (ARCH-TYPE CAUSALITY):

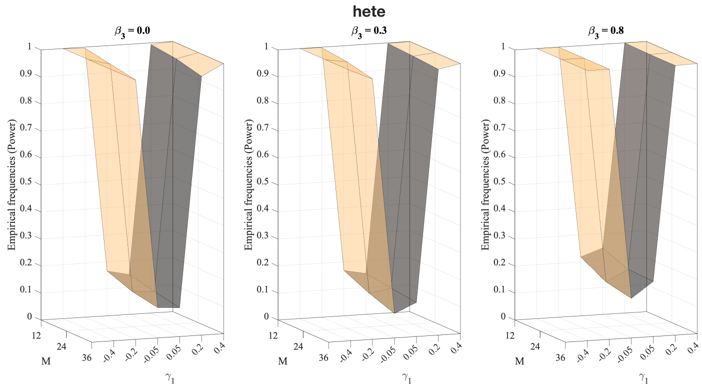
$$X_t = \beta_3 Z_{t-1} + \epsilon_x, \quad Z_t = 0.4Z_{t-1} + \epsilon_z, \quad (\epsilon_x, \epsilon_z) \sim \mathcal{N}(0, \Sigma_t)$$
$$\Sigma_t = (\sigma_{x,t}^2, \sigma_{xz,t}; \sigma_{xz,t}, 1), \quad \sigma_{x,t}^2 = 0.4 + \gamma_2 Z_{t-1}^2, \quad \sigma_{xz,t} = 0.4|\sigma_{x,t}|$$

d) DGP 4A (GARCH-TYPE CAUSALITY):

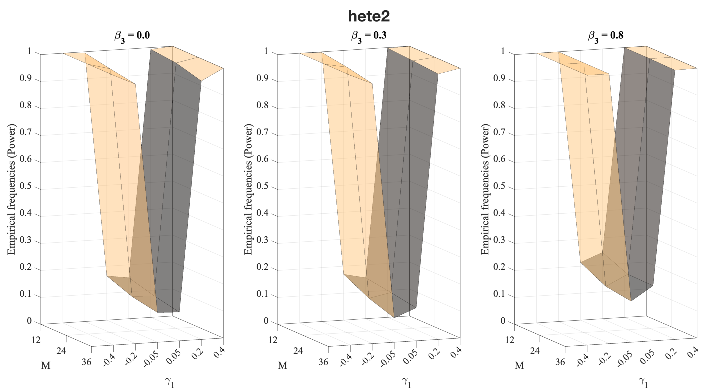
$$X_t = \epsilon_x, \quad Z_t = \epsilon_z, \quad (\epsilon_x, \epsilon_z) \sim \mathcal{N}(0, \Sigma_t), \quad \Sigma_t = (\sigma_{x,t}^2, 0; 0, \sigma_{z,t}^2)$$
$$\sigma_{x,t}^2 = 0.5 + \gamma_2 \sigma_{x,t-1}^2 + 0.1Z_{t-1}^2, \quad \sigma_{z,t} = 0.5 + \beta_3 \sigma_{z,t-1}^2 + 0.1Z_{t-1}^2$$

Parameter values span: $\gamma_1 \in \{-0.4, -0.2, -0.05, 0.05, 0.2, 0.4\}$,
$\gamma_2 \in \{0.05, 0.1, 0.2, 0.4, 0.6, 0.8\}$, $\beta_3 \in \{0, 0.3, 0.8\}$.

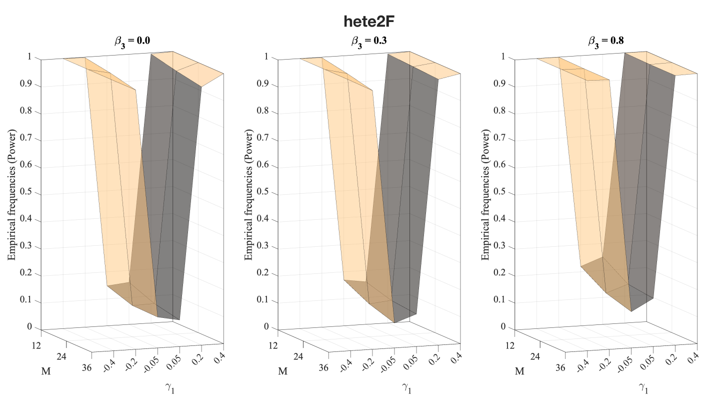
Given that higher-order moments of $X$ are relevant for the correction term, this set of Monte Carlo experiments consider the innovations to be drawn from a multivariate normal distribution. The smoothing parameter $M$ takes value $\{12, 24, 36\}$, which corresponds to lags up to 1 year, 2 years and 3 years, when the data is observed at monthly frequency. For each design, $1,000$ Monte Carlo simulations are run. All results report rejection rates at the 5% nominal significance level.




(a) DGP 1B, with $\beta = \{0, 0.3, 0.8\}$

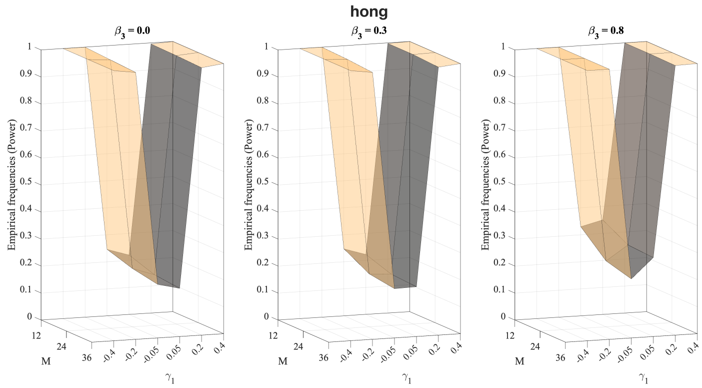



(b) DGP 1B, with $\beta = \{0, 0.3, 0.8\}$

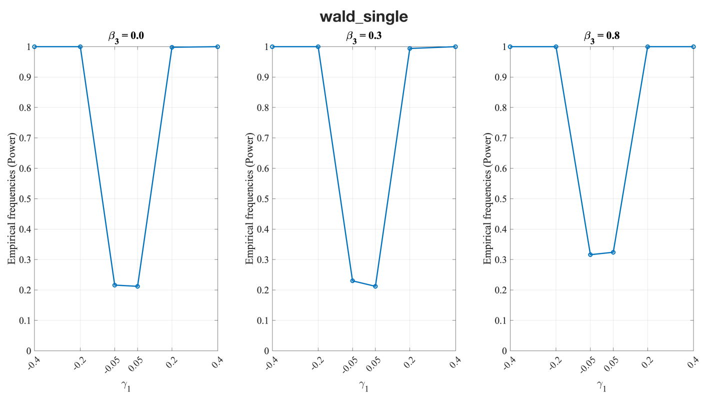



(c) DGP 1B, with $\beta = \{0, 0.3, 0.8\}$

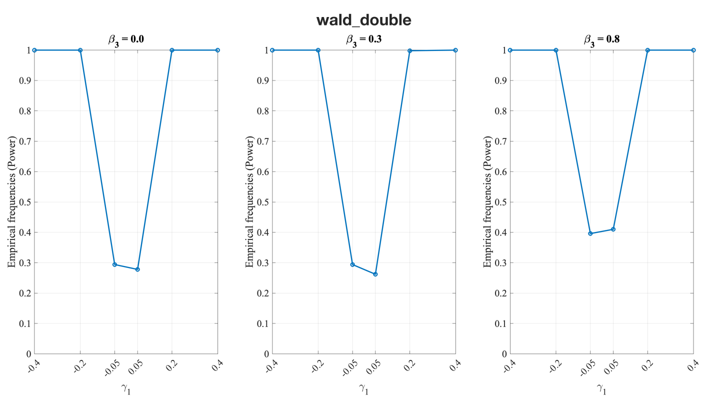
Figure 4: Power curves of the corrected tests: These figures present the rejection rates of the testing procedure associated to three corrected statistics (*Hete*, *Hete2*, *Hete2F*), under the alternatives (empirical power); sample size, $T = 1000$; $1000$ iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 24, 36\}$; nominal significance level is 5%.




(a) DGP 1B, with $\beta = \{0, 0.3, 0.8\}$

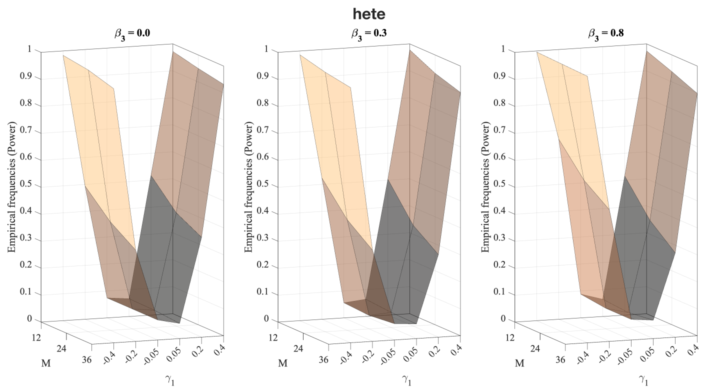



(b) DGP 1B, with $\beta = \{0, 0.3, 0.8\}$

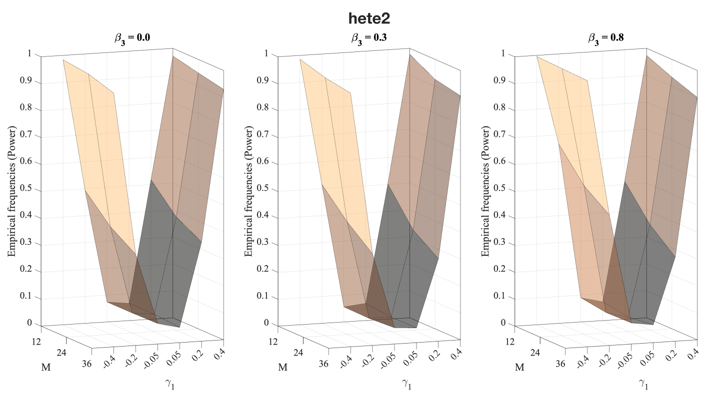



(c) DGP 1B, with $\beta = \{0, 0.3, 0.8\}$

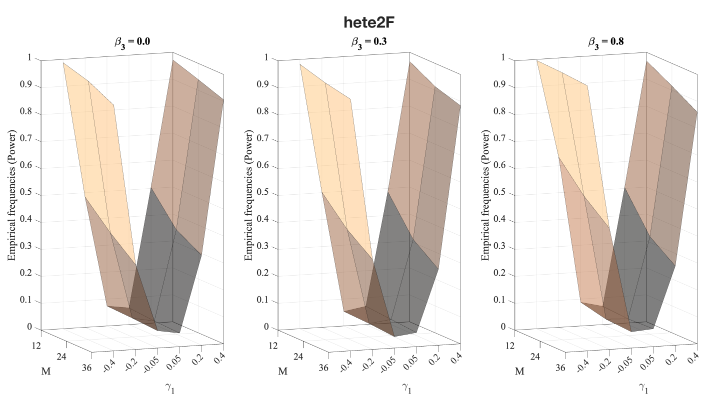
Figure 5: Power curves of the benchmark tests: These figures present the rejection rates of the testing procedure associated to three benchmark statistics (*Hong, Wald Single, Wald Double*), under the alternatives (empirical power); sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 24, 36\}$; nominal significance level is 5%.




(a) DGP 2B, with $\beta = \{0, 0.3, 0.8\}$

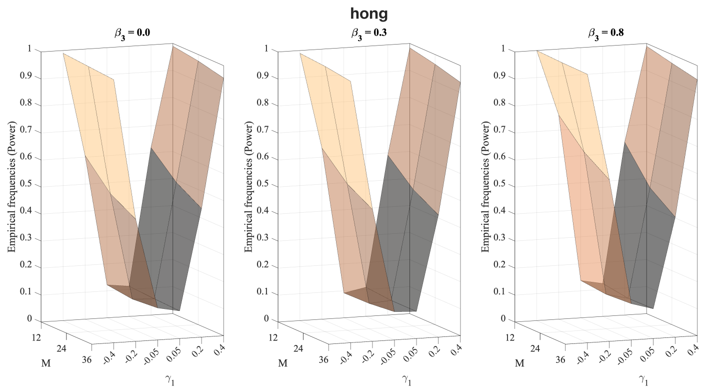



(b) DGP 2B, with $\beta = \{0, 0.3, 0.8\}$

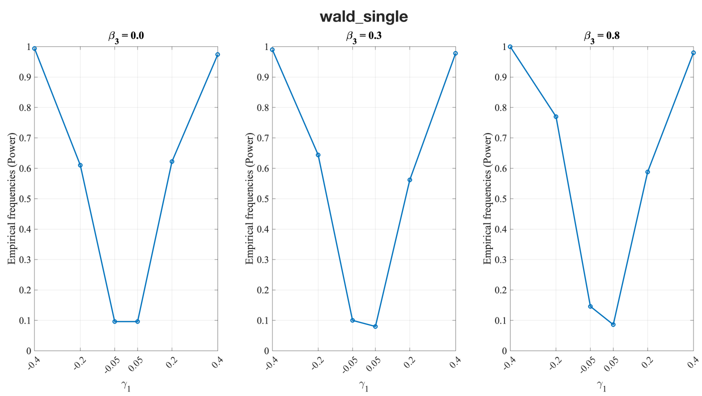



(c) DGP 2B, with $\beta = \{0, 0.3, 0.8\}$

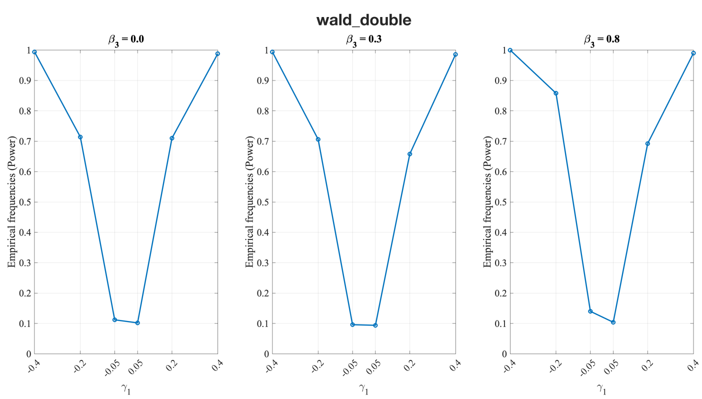
Figure 6: Power curves of the corrected tests: These figures present the rejection rates of the testing procedure associated to three corrected statistics (*Hete*, *Hete2*, *Hete2F*), under the alternatives (empirical power); sample size, $T = 1000$; $1000$ iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 24, 36\}$; nominal significance level is 5%.




(a) DGP 2B, with $\beta = \{0, 0.3, 0.8\}$

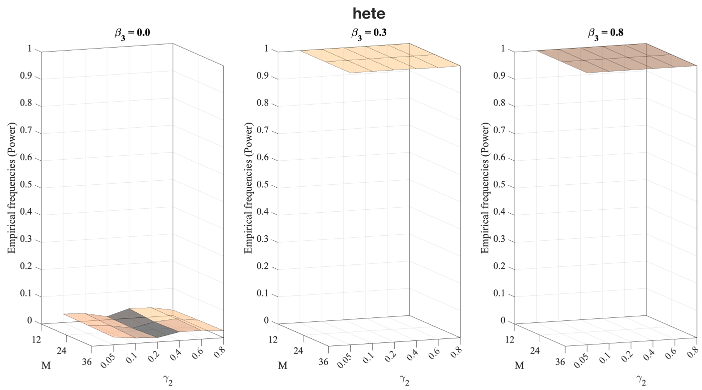



(b) DGP 2B, with $\beta = \{0, 0.3, 0.8\}$

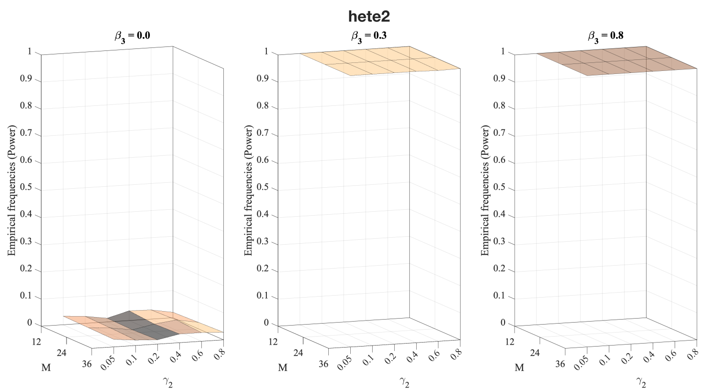



(c) DGP 2B, with $\beta = \{0, 0.3, 0.8\}$

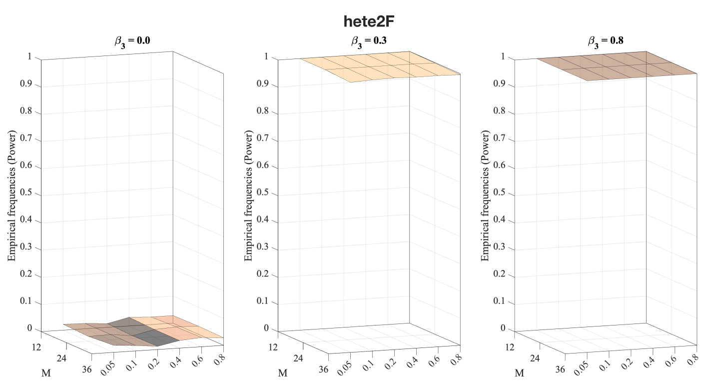
Figure 7: Power curves of the benchmark tests: These figures present the rejection rates of the testing procedure associated to three benchmark statistics (*Hong, Wald Single, Wald Double*), under the alternatives (empirical power); sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 24, 36\}$; nominal significance level is 5%.




(a) DGP 3B, with $\beta = \{0, 0.3, 0.8\}$

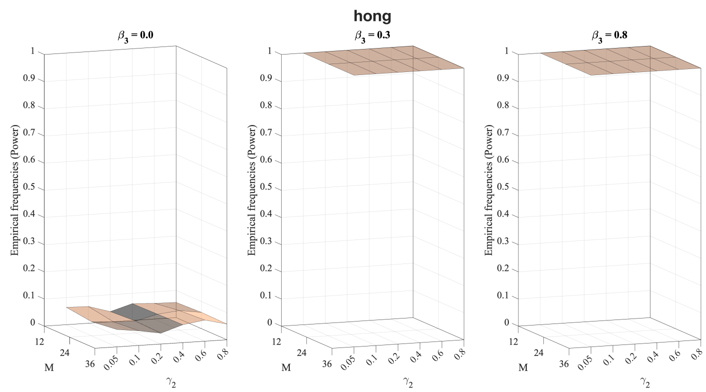



(b) DGP 3B, with $\beta = \{0, 0.3, 0.8\}$

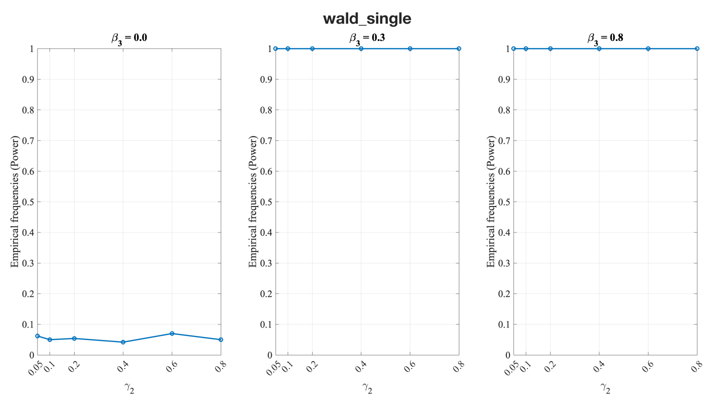



(c) DGP 3B, with $\beta = \{0, 0.3, 0.8\}$

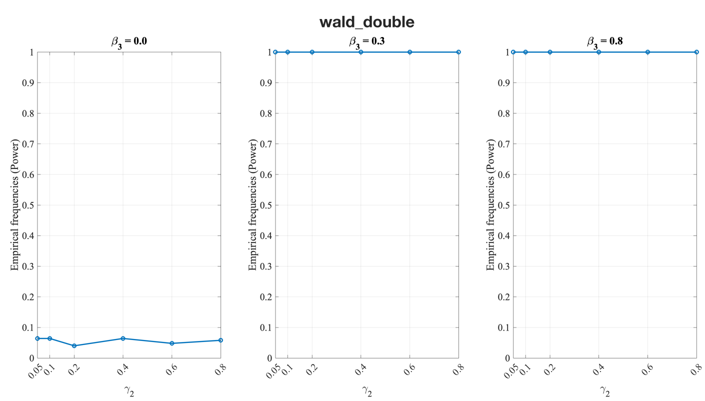
Figure 8: Power curves of the corrected tests: These figures present the rejection rates of the testing procedure associated to three corrected statistics (*Hete*, *Hete2*, *Hete2F*), under the alternatives (empirical power); sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 24, 36\}$; nominal significance level is 5%.




(a) DGP 3B, with $\beta = \{0, 0.3, 0.8\}$

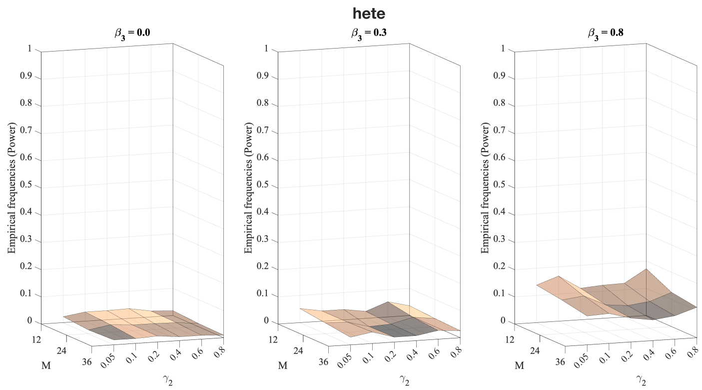



(b) DGP 3B, with $\beta = \{0, 0.3, 0.8\}$

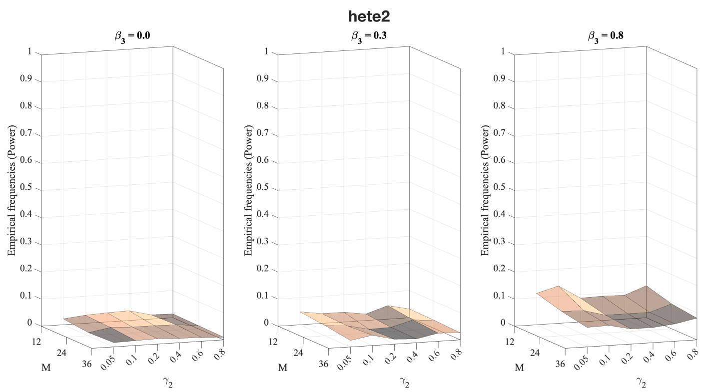



(c) DGP 3B, with $\beta = \{0, 0.3, 0.8\}$

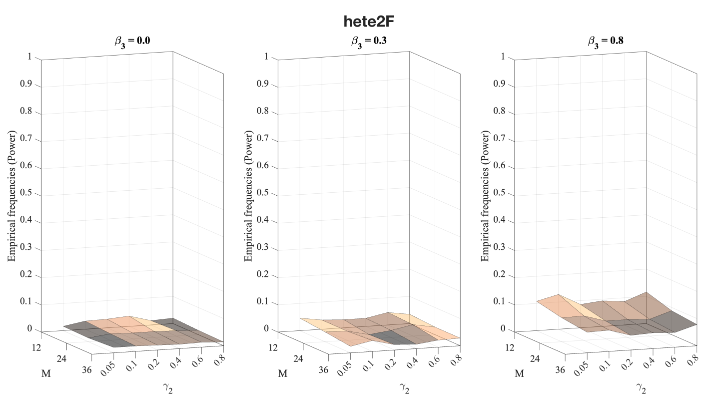
Figure 9: Power curves of the benchmark tests: These figures present the rejection rates of the testing procedure associated to three benchmark statistics (*Hong*, *Wald Single*, *Wald Double*), under the alternatives (empirical power); sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 24, 36\}$; nominal significance level is 5%.




(a) DGP 4B, with $\beta = \{0, 0.3, 0.8\}$

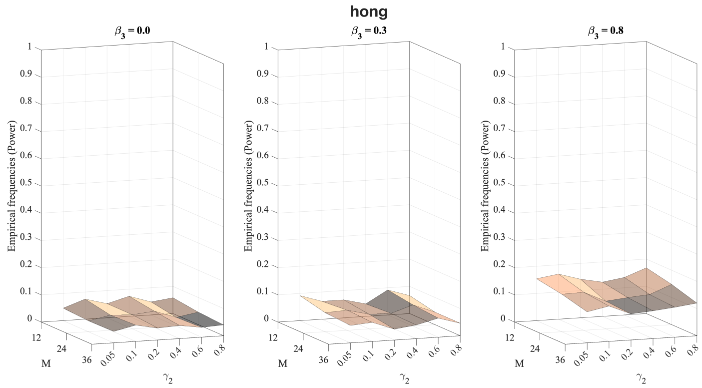



(b) DGP 4B, with $\beta = \{0, 0.3, 0.8\}$

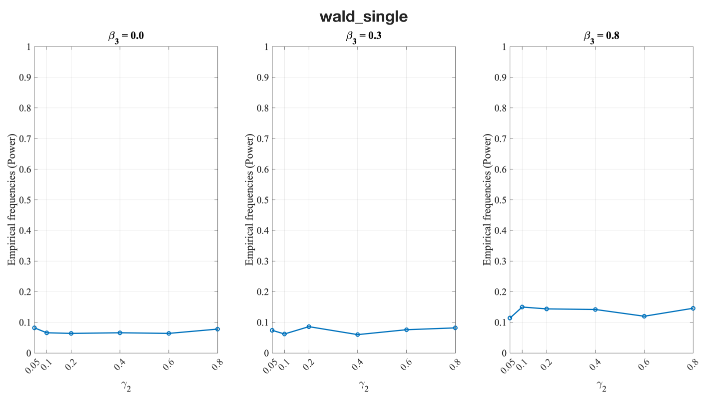



(c) DGP 4B, with $\beta = \{0, 0.3, 0.8\}$

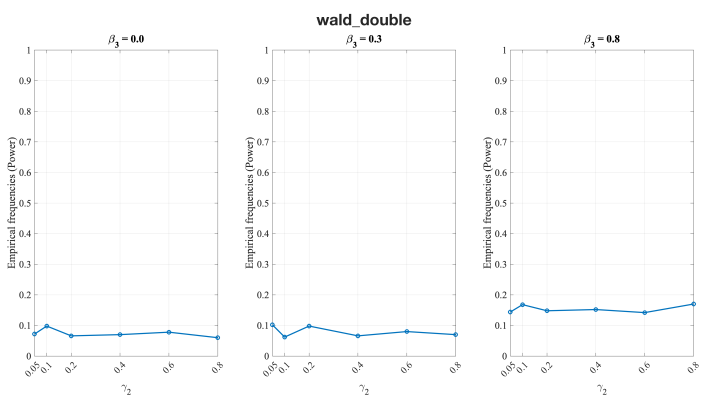
Figure 10: Power curves of the corrected tests: These figures present the rejection rates of the testing procedure associated to three corrected statistics (*Hete*, *Hete2*, *Hete2F*), under the alternatives (empirical power); sample size, $T = 1000$; $1000$ iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 24, 36\}$; nominal significance level is 5%.




(a) DGP 4B, with $\beta = \{0, 0.3, 0.8\}$




(b) DGP 4B, with $\beta = \{0, 0.3, 0.8\}$




(c) DGP 4B, with $\beta = \{0, 0.3, 0.8\}$

Figure 11: Power curves of the benchmark tests: These figures present the rejection rates of the testing procedure associated to three benchmark statistics (*Hong, Wald Single, Wald Double*), under the alternatives (empirical power); sample size, $T = 1000$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 24, 36\}$; nominal significance level is 5%.

## C.4 DGPs under the Null: Small Sample $T = 150$

Table 10: Rejection frequencies for DGP1A (Baseline): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP1A; sample size, $T = 150$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 0$; nominal significance level is 5%.

$M = 12 \quad \approx \sqrt{T}$

| | Hete | | | | | Hong | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.040 | 0.018 | 0.024 | 0.028 | 0.028 | 0.060 | 0.048 | 0.048 | 0.050 | 0.052 |
| $\alpha = 0.2$ | 0.022 | 0.034 | 0.018 | 0.016 | 0.034 | 0.052 | 0.052 | 0.042 | 0.044 | 0.054 |
| $\alpha = 0.3$ | 0.028 | 0.026 | 0.026 | 0.034 | 0.028 | 0.050 | 0.050 | 0.050 | 0.074 | 0.052 |
| $\alpha = 0.4$ | 0.038 | 0.026 | 0.034 | 0.038 | 0.038 | 0.052 | 0.054 | 0.064 | 0.056 | 0.066 |
| $\alpha = 0.5$ | 0.052 | 0.054 | 0.036 | 0.040 | 0.038 | 0.082 | 0.074 | 0.068 | 0.050 | 0.064 |
| $\alpha = 0.6$ | 0.042 | 0.056 | 0.066 | 0.038 | 0.044 | 0.052 | 0.088 | 0.094 | 0.062 | 0.064 |
| $\alpha = 0.7$ | 0.064 | 0.062 | 0.042 | 0.050 | 0.066 | 0.096 | 0.092 | 0.064 | 0.078 | 0.094 |

$M = 12 \quad \approx \sqrt{T}$

| | Hete2 | | | | | Hete2F | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.036 | 0.016 | 0.018 | 0.022 | 0.024 | 0.032 | 0.014 | 0.018 | 0.014 | 0.028 |
| $\alpha = 0.2$ | 0.020 | 0.030 | 0.018 | 0.016 | 0.030 | 0.018 | 0.024 | 0.018 | 0.018 | 0.022 |
| $\alpha = 0.3$ | 0.030 | 0.026 | 0.028 | 0.032 | 0.028 | 0.024 | 0.022 | 0.020 | 0.034 | 0.022 |
| $\alpha = 0.4$ | 0.032 | 0.022 | 0.034 | 0.034 | 0.038 | 0.018 | 0.026 | 0.028 | 0.028 | 0.038 |
| $\alpha = 0.5$ | 0.050 | 0.046 | 0.036 | 0.038 | 0.038 | 0.050 | 0.044 | 0.024 | 0.030 | 0.038 |
| $\alpha = 0.6$ | 0.038 | 0.060 | 0.060 | 0.026 | 0.046 | 0.034 | 0.056 | 0.050 | 0.032 | 0.034 |
| $\alpha = 0.7$ | 0.062 | 0.062 | 0.042 | 0.052 | 0.068 | 0.054 | 0.066 | 0.050 | 0.042 | 0.072 |

$M = 36$

| | Hete | | | | | Hong | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.018 | 0.036 | 0.026 | 0.026 | 0.016 | 0.040 | 0.048 | 0.032 | 0.042 | 0.042 |
| $\alpha = 0.2$ | 0.032 | 0.018 | 0.016 | 0.028 | 0.040 | 0.042 | 0.044 | 0.034 | 0.042 | 0.054 |
| $\alpha = 0.3$ | 0.024 | 0.036 | 0.034 | 0.050 | 0.032 | 0.034 | 0.056 | 0.050 | 0.078 | 0.058 |
| $\alpha = 0.4$ | 0.048 | 0.066 | 0.032 | 0.046 | 0.040 | 0.062 | 0.082 | 0.076 | 0.072 | 0.080 |
| $\alpha = 0.5$ | 0.068 | 0.066 | 0.052 | 0.054 | 0.056 | 0.108 | 0.100 | 0.092 | 0.086 | 0.088 |
| $\alpha = 0.6$ | 0.074 | 0.100 | 0.068 | 0.050 | 0.072 | 0.112 | 0.124 | 0.102 | 0.080 | 0.100 |
| $\alpha = 0.7$ | 0.134 | 0.100 | 0.102 | 0.084 | 0.122 | 0.138 | 0.136 | 0.118 | 0.120 | 0.160 |

$M = 36$

| | Hete2 | | | | | Hete2F | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.020 | 0.026 | 0.018 | 0.018 | 0.014 | 0.020 | 0.022 | 0.016 | 0.020 | 0.014 |
| $\alpha = 0.2$ | 0.020 | 0.016 | 0.014 | 0.026 | 0.040 | 0.020 | 0.014 | 0.016 | 0.026 | 0.028 |
| $\alpha = 0.3$ | 0.020 | 0.024 | 0.026 | 0.046 | 0.028 | 0.018 | 0.024 | 0.016 | 0.040 | 0.018 |
| $\alpha = 0.4$ | 0.040 | 0.058 | 0.026 | 0.040 | 0.040 | 0.026 | 0.044 | 0.026 | 0.034 | 0.038 |
| $\alpha = 0.5$ | 0.052 | 0.054 | 0.048 | 0.048 | 0.056 | 0.042 | 0.048 | 0.038 | 0.024 | 0.050 |
| $\alpha = 0.6$ | 0.072 | 0.094 | 0.062 | 0.040 | 0.068 | 0.062 | 0.084 | 0.058 | 0.036 | 0.046 |
| $\alpha = 0.7$ | 0.108 | 0.096 | 0.090 | 0.082 | 0.116 | 0.086 | 0.096 | 0.066 | 0.046 | 0.096 |

| | Wald Single | | | | | Wald Double | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0$ | 0.048 | 0.042 | 0.064 | 0.064 | 0.054 | 0.058 | 0.060 | 0.072 | 0.068 | 0.054 |
| $\alpha = 0.2$ | 0.042 | 0.054 | 0.052 | 0.030 | 0.058 | 0.048 | 0.054 | 0.066 | 0.038 | 0.060 |
| $\alpha = 0.3$ | 0.040 | 0.044 | 0.044 | 0.060 | 0.056 | 0.044 | 0.050 | 0.046 | 0.074 | 0.050 |
| $\alpha = 0.4$ | 0.052 | 0.032 | 0.034 | 0.064 | 0.066 | 0.040 | 0.048 | 0.036 | 0.060 | 0.058 |
| $\alpha = 0.5$ | 0.052 | 0.048 | 0.046 | 0.062 | 0.048 | 0.050 | 0.056 | 0.044 | 0.046 | 0.044 |
| $\alpha = 0.6$ | 0.044 | 0.046 | 0.062 | 0.054 | 0.042 | 0.036 | 0.054 | 0.060 | 0.054 | 0.042 |
| $\alpha = 0.7$ | 0.056 | 0.054 | 0.040 | 0.050 | 0.046 | 0.040 | 0.050 | 0.040 | 0.056 | 0.054 |

Table 11: Rejection frequencies for DGP2A (Baseline): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP2A; sample size, $T = 150$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 0$; nominal significance level is 5%.

$M = 12 \quad \approx \sqrt{T}$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.124 | 0.188 | 0.178 | 0.202 | 0.042 | 0.038 | 0.048 | 0.026 |
| $\alpha = 0.3$ | 0.132 | 0.172 | 0.140 | 0.166 | 0.044 | 0.048 | 0.040 | 0.048 |
| $\alpha = 0.4$ | 0.128 | 0.154 | 0.120 | 0.132 | 0.054 | 0.060 | 0.044 | 0.046 |
| $\alpha = 0.5$ | 0.094 | 0.116 | 0.120 | 0.104 | 0.052 | 0.064 | 0.060 | 0.054 |
| $\alpha = 0.6$ | 0.072 | 0.092 | 0.074 | 0.098 | 0.066 | 0.064 | 0.048 | 0.058 |
| $\alpha = 0.7$ | 0.068 | 0.092 | 0.086 | 0.092 | 0.068 | 0.060 | 0.066 | 0.068 |

$M = 12 \quad \approx \sqrt{T}$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.032 | 0.038 | 0.038 | 0.032 | 0.028 | 0.036 | 0.036 | 0.036 |
| $\alpha = 0.3$ | 0.038 | 0.046 | 0.030 | 0.036 | 0.032 | 0.040 | 0.028 | 0.034 |
| $\alpha = 0.4$ | 0.050 | 0.060 | 0.034 | 0.038 | 0.042 | 0.066 | 0.032 | 0.038 |
| $\alpha = 0.5$ | 0.040 | 0.052 | 0.056 | 0.056 | 0.034 | 0.052 | 0.046 | 0.040 |
| $\alpha = 0.6$ | 0.056 | 0.050 | 0.042 | 0.048 | 0.046 | 0.048 | 0.044 | 0.054 |
| $\alpha = 0.7$ | 0.054 | 0.048 | 0.048 | 0.054 | 0.052 | 0.042 | 0.056 | 0.050 |

$M = 36$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.158 | 0.236 | 0.212 | 0.260 | 0.054 | 0.052 | 0.056 | 0.062 |
| $\alpha = 0.3$ | 0.172 | 0.212 | 0.234 | 0.208 | 0.064 | 0.068 | 0.084 | 0.044 |
| $\alpha = 0.4$ | 0.174 | 0.188 | 0.214 | 0.196 | 0.066 | 0.090 | 0.086 | 0.066 |
| $\alpha = 0.5$ | 0.152 | 0.184 | 0.188 | 0.174 | 0.072 | 0.098 | 0.104 | 0.108 |
| $\alpha = 0.6$ | 0.174 | 0.198 | 0.166 | 0.198 | 0.118 | 0.114 | 0.116 | 0.132 |
| $\alpha = 0.7$ | 0.186 | 0.218 | 0.184 | 0.206 | 0.148 | 0.162 | 0.128 | 0.128 |

$M = 36$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.048 | 0.044 | 0.072 | 0.072 | 0.046 | 0.032 | 0.064 | 0.068 |
| $\alpha = 0.3$ | 0.060 | 0.076 | 0.090 | 0.054 | 0.054 | 0.058 | 0.078 | 0.048 |
| $\alpha = 0.4$ | 0.064 | 0.076 | 0.084 | 0.056 | 0.058 | 0.086 | 0.086 | 0.054 |
| $\alpha = 0.5$ | 0.084 | 0.102 | 0.104 | 0.102 | 0.056 | 0.086 | 0.078 | 0.078 |
| $\alpha = 0.6$ | 0.108 | 0.112 | 0.106 | 0.126 | 0.088 | 0.104 | 0.100 | 0.120 |
| $\alpha = 0.7$ | 0.152 | 0.150 | 0.126 | 0.132 | 0.138 | 0.108 | 0.128 | 0.120 |

| | Wald Single | | | | Wald Double | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.052 | 0.030 | 0.058 | 0.034 | 0.044 | 0.052 | 0.036 | 0.026 |
| $\alpha = 0.3$ | 0.036 | 0.036 | 0.042 | 0.056 | 0.030 | 0.036 | 0.034 | 0.034 |
| $\alpha = 0.4$ | 0.068 | 0.056 | 0.038 | 0.038 | 0.048 | 0.050 | 0.026 | 0.020 |
| $\alpha = 0.5$ | 0.050 | 0.046 | 0.038 | 0.026 | 0.034 | 0.040 | 0.032 | 0.026 |
| $\alpha = 0.6$ | 0.034 | 0.030 | 0.034 | 0.032 | 0.026 | 0.024 | 0.024 | 0.022 |
| $\alpha = 0.7$ | 0.034 | 0.020 | 0.032 | 0.020 | 0.022 | 0.020 | 0.018 | 0.008 |

Table 12: Rejection frequencies for DGP3A (Baseline): This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP3A; sample size, $T = 150$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$ and $\varpi = 0$; nominal significance level is 5%.

$M = 12 \quad \approx \sqrt{T}$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.024 | 0.042 | 0.030 | 0.032 | 0.048 | 0.056 | 0.060 | 0.056 |
| $\alpha = 0.3$ | 0.032 | 0.038 | 0.026 | 0.040 | 0.050 | 0.052 | 0.044 | 0.062 |
| $\alpha = 0.4$ | 0.048 | 0.040 | 0.028 | 0.050 | 0.072 | 0.068 | 0.054 | 0.066 |
| $\alpha = 0.5$ | 0.032 | 0.034 | 0.028 | 0.030 | 0.054 | 0.050 | 0.052 | 0.060 |
| $\alpha = 0.6$ | 0.050 | 0.056 | 0.052 | 0.032 | 0.080 | 0.088 | 0.066 | 0.064 |
| $\alpha = 0.7$ | 0.054 | 0.054 | 0.046 | 0.054 | 0.082 | 0.070 | 0.078 | 0.078 |

$M = 12 \quad \approx \sqrt{T}$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.018 | 0.034 | 0.028 | 0.028 | 0.018 | 0.036 | 0.030 | 0.022 |
| $\alpha = 0.3$ | 0.026 | 0.030 | 0.026 | 0.032 | 0.032 | 0.026 | 0.024 | 0.032 |
| $\alpha = 0.4$ | 0.048 | 0.040 | 0.028 | 0.048 | 0.038 | 0.040 | 0.026 | 0.042 |
| $\alpha = 0.5$ | 0.030 | 0.034 | 0.028 | 0.026 | 0.020 | 0.032 | 0.024 | 0.024 |
| $\alpha = 0.6$ | 0.048 | 0.056 | 0.052 | 0.032 | 0.042 | 0.054 | 0.036 | 0.034 |
| $\alpha = 0.7$ | 0.056 | 0.052 | 0.042 | 0.050 | 0.052 | 0.050 | 0.040 | 0.042 |

$M = 36$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.018 | 0.036 | 0.040 | 0.028 | 0.038 | 0.064 | 0.066 | 0.052 |
| $\alpha = 0.3$ | 0.028 | 0.026 | 0.026 | 0.034 | 0.052 | 0.046 | 0.050 | 0.066 |
| $\alpha = 0.4$ | 0.038 | 0.052 | 0.030 | 0.044 | 0.084 | 0.076 | 0.048 | 0.076 |
| $\alpha = 0.5$ | 0.046 | 0.052 | 0.050 | 0.060 | 0.074 | 0.074 | 0.076 | 0.094 |
| $\alpha = 0.6$ | 0.114 | 0.110 | 0.070 | 0.066 | 0.122 | 0.138 | 0.092 | 0.074 |
| $\alpha = 0.7$ | 0.112 | 0.094 | 0.116 | 0.102 | 0.136 | 0.128 | 0.130 | 0.126 |

$M = 36$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.014 | 0.028 | 0.036 | 0.028 | 0.014 | 0.026 | 0.020 | 0.012 |
| $\alpha = 0.3$ | 0.028 | 0.022 | 0.028 | 0.028 | 0.026 | 0.022 | 0.026 | 0.020 |
| $\alpha = 0.4$ | 0.040 | 0.040 | 0.028 | 0.040 | 0.036 | 0.038 | 0.024 | 0.028 |
| $\alpha = 0.5$ | 0.044 | 0.044 | 0.042 | 0.050 | 0.034 | 0.022 | 0.042 | 0.044 |
| $\alpha = 0.6$ | 0.098 | 0.098 | 0.058 | 0.052 | 0.082 | 0.078 | 0.046 | 0.048 |
| $\alpha = 0.7$ | 0.104 | 0.086 | 0.106 | 0.096 | 0.072 | 0.076 | 0.070 | 0.074 |

| | Wald Single | | | | Wald Double | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ | $\beta_1 = 0.2$ | $\beta_1 = 0.4$ | $\beta_1 = 0.6$ | $\beta_1 = 1$ |
| $\alpha = 0.2$ | 0.058 | 0.060 | 0.054 | 0.062 | 0.060 | 0.052 | 0.058 | 0.058 |
| $\alpha = 0.3$ | 0.062 | 0.056 | 0.052 | 0.058 | 0.048 | 0.054 | 0.042 | 0.034 |
| $\alpha = 0.4$ | 0.040 | 0.052 | 0.050 | 0.068 | 0.058 | 0.058 | 0.056 | 0.056 |
| $\alpha = 0.5$ | 0.046 | 0.040 | 0.052 | 0.052 | 0.056 | 0.036 | 0.044 | 0.038 |
| $\alpha = 0.6$ | 0.060 | 0.068 | 0.044 | 0.050 | 0.052 | 0.068 | 0.048 | 0.048 |
| $\alpha = 0.7$ | 0.042 | 0.044 | 0.050 | 0.054 | 0.054 | 0.050 | 0.066 | 0.058 |

Table 13: Rejection frequencies for DGP4A: This table presents the rejection frequencies of six testing procedure, when the time series are generated by DGP4A; sample size, $T = 150$; 1000 iterations; the weighting function is the Bartlett kernel; the smoothing parameter range is: $M = \{12, 36\}$; nominal significance level is 5%.

$M = 12 \quad \approx \sqrt{T}$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.018 | 0.024 | 0.032 | 0.024 | 0.058 | 0.050 | 0.062 | 0.048 |
| $\alpha = 0.3$ | 0.040 | 0.024 | 0.020 | 0.016 | 0.066 | 0.054 | 0.038 | 0.040 |
| $\alpha = 0.4$ | 0.008 | 0.018 | 0.020 | 0.022 | 0.014 | 0.044 | 0.044 | 0.040 |
| $\alpha = 0.5$ | 0.026 | 0.006 | 0.024 | 0.024 | 0.056 | 0.032 | 0.058 | 0.058 |
| $\alpha = 0.6$ | 0.018 | 0.006 | 0.026 | 0.020 | 0.038 | 0.042 | 0.050 | 0.054 |
| $\alpha = 0.7$ | 0.026 | 0.006 | 0.026 | 0.020 | 0.036 | 0.046 | 0.060 | 0.036 |

$M = 12 \quad \approx \sqrt{T}$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.018 | 0.020 | 0.032 | 0.026 | 0.018 | 0.016 | 0.034 | 0.020 |
| $\alpha = 0.3$ | 0.040 | 0.022 | 0.020 | 0.016 | 0.036 | 0.024 | 0.020 | 0.016 |
| $\alpha = 0.4$ | 0.008 | 0.018 | 0.020 | 0.022 | 0.006 | 0.020 | 0.020 | 0.020 |
| $\alpha = 0.5$ | 0.024 | 0.008 | 0.026 | 0.024 | 0.022 | 0.010 | 0.026 | 0.026 |
| $\alpha = 0.6$ | 0.016 | 0.006 | 0.022 | 0.016 | 0.016 | 0.004 | 0.024 | 0.016 |
| $\alpha = 0.7$ | 0.024 | 0.006 | 0.026 | 0.020 | 0.020 | 0.008 | 0.022 | 0.020 |

$M = 36$

| | Hete | | | | Hong | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.024 | 0.016 | 0.018 | 0.018 | 0.046 | 0.042 | 0.054 | 0.054 |
| $\alpha = 0.3$ | 0.030 | 0.018 | 0.010 | 0.014 | 0.068 | 0.044 | 0.046 | 0.040 |
| $\alpha = 0.4$ | 0.006 | 0.016 | 0.016 | 0.014 | 0.024 | 0.042 | 0.044 | 0.040 |
| $\alpha = 0.5$ | 0.018 | 0.016 | 0.020 | 0.024 | 0.042 | 0.036 | 0.052 | 0.062 |
| $\alpha = 0.6$ | 0.020 | 0.012 | 0.026 | 0.024 | 0.048 | 0.044 | 0.054 | 0.048 |
| $\alpha = 0.7$ | 0.012 | 0.002 | 0.022 | 0.014 | 0.034 | 0.026 | 0.044 | 0.052 |

$M = 36$

| | Hete2 | | | | Hete2F | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.024 | 0.016 | 0.018 | 0.020 | 0.024 | 0.016 | 0.018 | 0.016 |
| $\alpha = 0.3$ | 0.030 | 0.018 | 0.008 | 0.014 | 0.026 | 0.016 | 0.008 | 0.014 |
| $\alpha = 0.4$ | 0.006 | 0.016 | 0.012 | 0.014 | 0.004 | 0.016 | 0.014 | 0.014 |
| $\alpha = 0.5$ | 0.014 | 0.010 | 0.018 | 0.020 | 0.014 | 0.010 | 0.024 | 0.022 |
| $\alpha = 0.6$ | 0.020 | 0.014 | 0.026 | 0.024 | 0.022 | 0.014 | 0.024 | 0.024 |
| $\alpha = 0.7$ | 0.010 | 0.002 | 0.022 | 0.014 | 0.014 | 0.006 | 0.018 | 0.016 |

| | Wald Single | | | | Wald Double | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ | $\beta_2 = 0.05$ | $\beta_2 = 0.1$ | $\beta_2 = 0.2$ | $\beta_2 = 0.3$ |
| $\alpha = 0.2$ | 0.052 | 0.060 | 0.060 | 0.054 | 0.052 | 0.054 | 0.064 | 0.054 |
| $\alpha = 0.3$ | 0.058 | 0.046 | 0.046 | 0.056 | 0.058 | 0.056 | 0.042 | 0.058 |
| $\alpha = 0.4$ | 0.038 | 0.052 | 0.046 | 0.054 | 0.040 | 0.052 | 0.050 | 0.056 |
| $\alpha = 0.5$ | 0.074 | 0.036 | 0.040 | 0.050 | 0.064 | 0.042 | 0.054 | 0.062 |
| $\alpha = 0.6$ | 0.058 | 0.046 | 0.054 | 0.056 | 0.048 | 0.046 | 0.068 | 0.042 |
| $\alpha = 0.7$ | 0.054 | 0.070 | 0.044 | 0.048 | 0.052 | 0.058 | 0.062 | 0.046 |

# D Online Appendix Empirical Application

## D.1 Testing Weak Exogeneity: Measures of UN Shocks

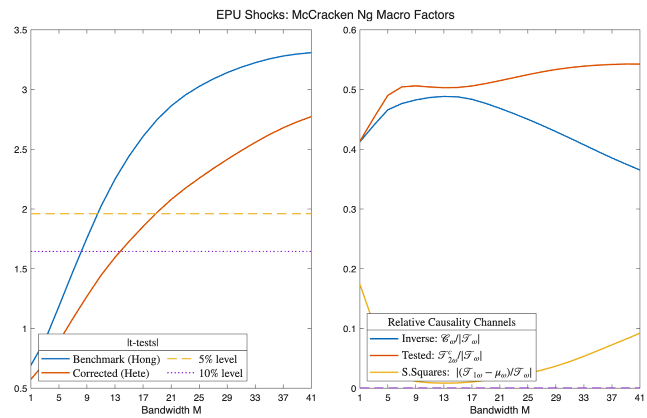



Figure 12: Baker et al. (2016)'s shocks and McCracken and Ng (2016)'s factors (2024/06 vintage). Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

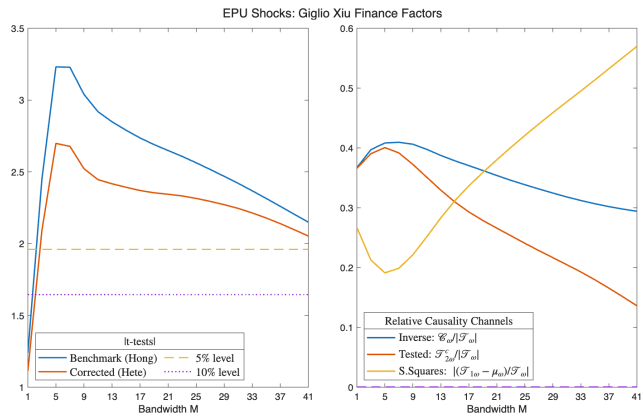



Figure 13: Baker et al. (2016)'s shocks and Giglio and Xiu (2021)'s factors.
Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

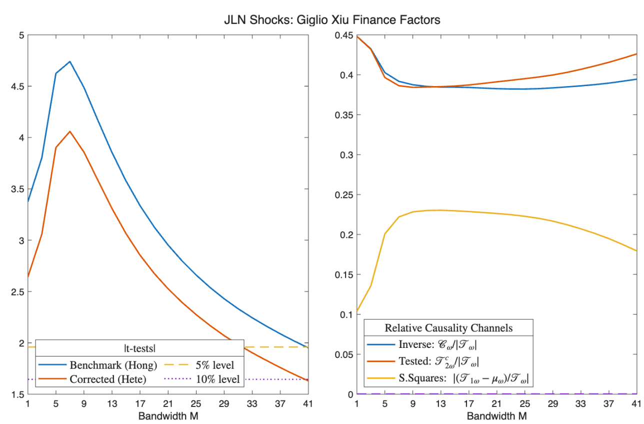



Figure 14: Ludvigson et al. (2021)'s shocks and Giglio and Xiu (2021)'s factors.
Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

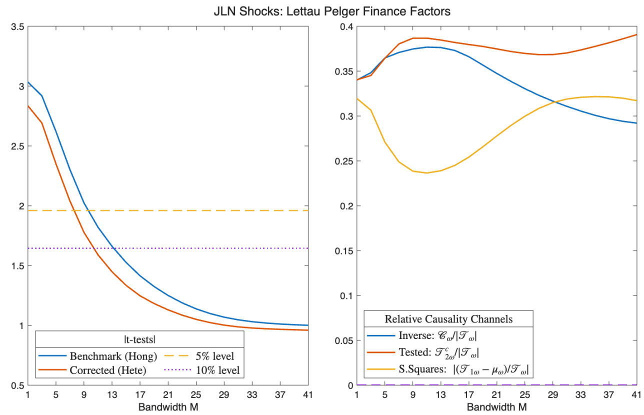



Figure 15: Ludvigson et al. (2021)'s shocks and Lettau and Pelger (2020b)'s factors.
Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 ($\sim$3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

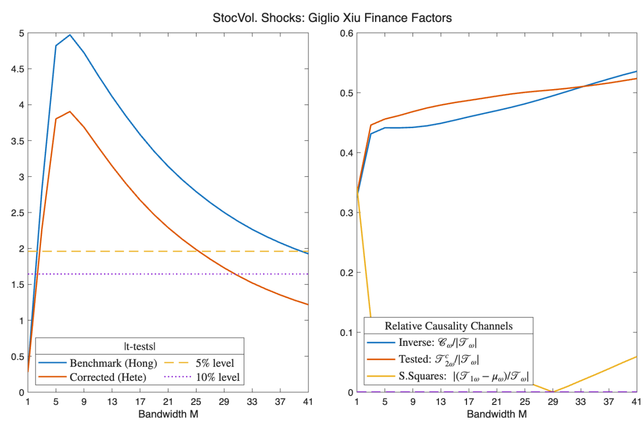



Figure 16: Berger et al. (2020)'s shocks and Giglio and Xiu (2021)'s factors.
Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 ($\sim$3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

## D.2 Testing Weak Exogeneity: Measures of MP Shocks

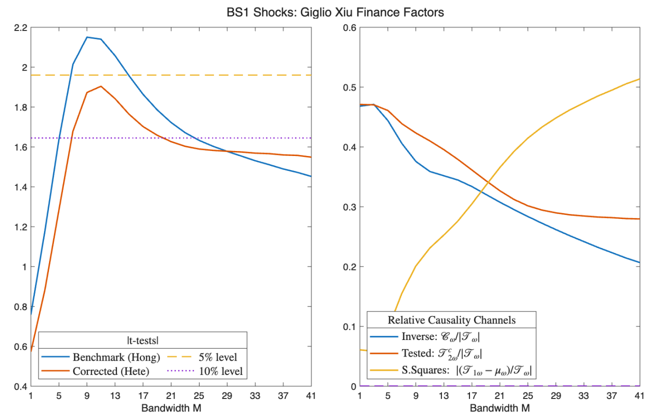



Figure 17: Bauer and Swanson (2023)'s shocks (MPS) and Giglio and Xiu (2021)'s factors.
Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

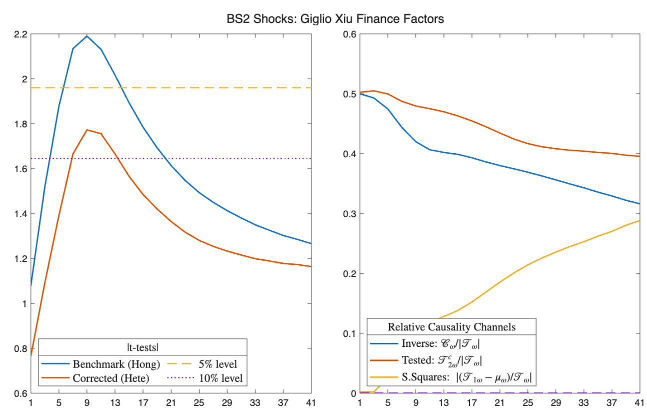



Figure 18: Bauer and Swanson (2023)'s orthogonalized shocks (MPS Ort) and Giglio and Xiu (2021)'s factors.
Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

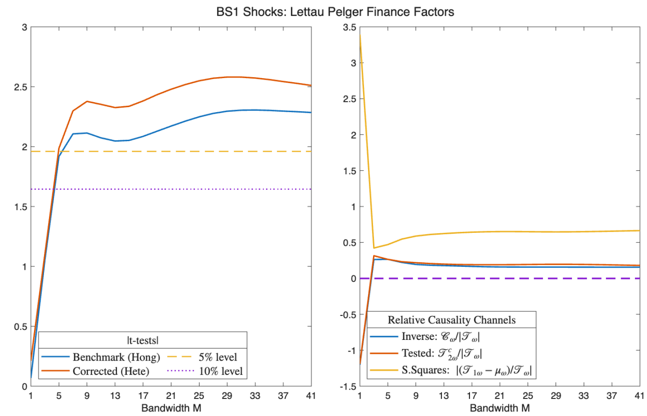



Figure 19: Bauer and Swanson (2023)'s shocks (MPS) and Lettau and Pelger (2020b)'s factors.
Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

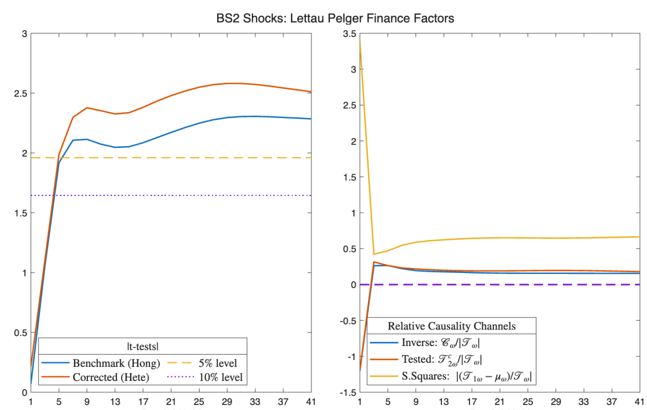



Figure 20: Bauer and Swanson (2023)'s orthogonalized shocks (MPS Ort) and Lettau and Pelger (2020b)'s factors.
Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

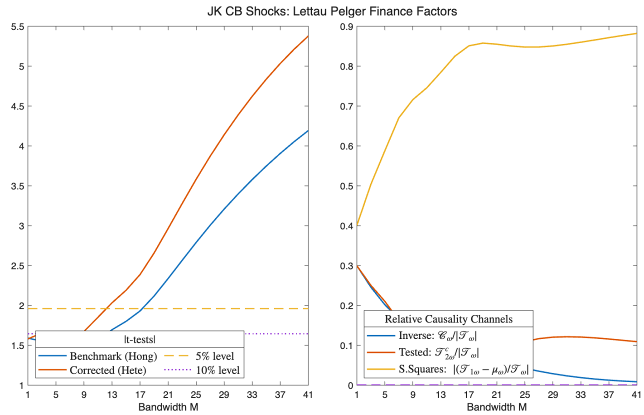



Figure 21: Jarociński and Karadi (2020)'s CBI shocks and Lettau and Pelger (2020b)'s factors. Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

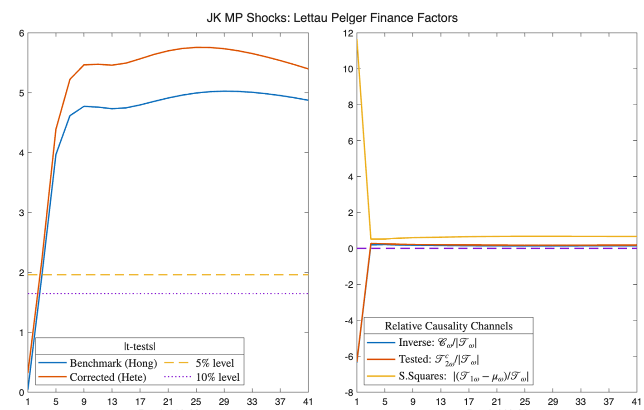



Figure 22: Jarociński and Karadi (2020)'s MPI shocks and Lettau and Pelger (2020b)'s factors. Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

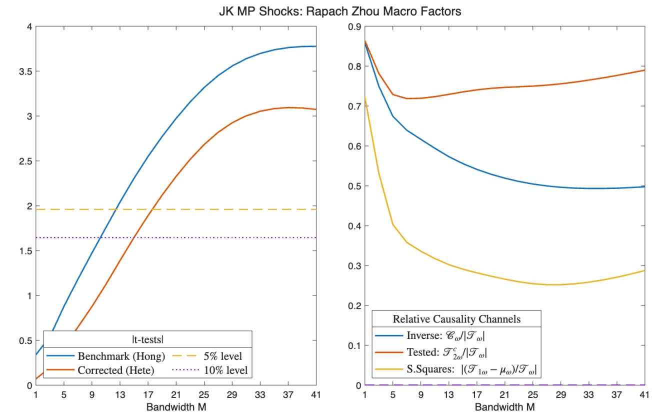



Figure 23: Jarociński and Karadi (2020)'s MPI shocks and Rapach and Zhou (2021)'s factors. Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

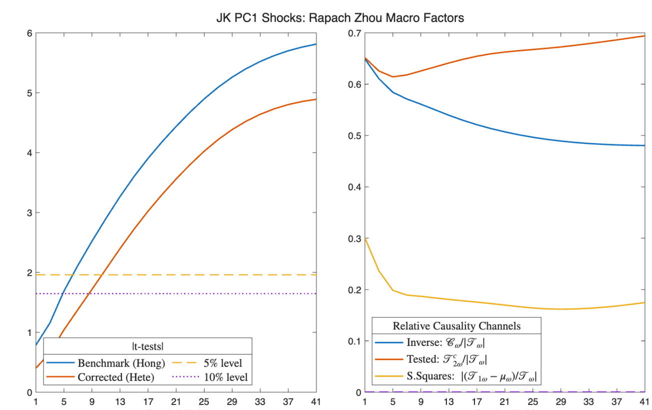
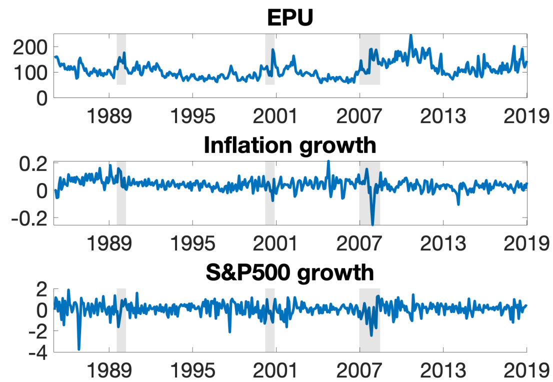
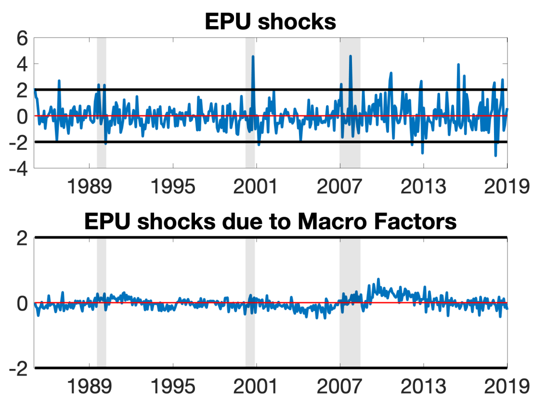



Figure 24: Jarociński and Karadi (2020)'s PC shocks and Rapach and Zhou (2021)'s factors. Comparison between the two testing strategies. LEFT PANEL: on the y-axes, the level of the benchmark Portmanteau statistic (blue solid) and corrected test statistic (orange solid) at different horizons $M$; on the x-axis, the smoothing parameter range from 1 to 41 (∼3y+1q); the weighting function is the quadratic spectral kernel (Andrews, 1991); nominal significance levels are 5% (dashed yellow), and 10% (dashed purple). RIGHT PANEL: on the y-axes, the contributions of the channels relative to the absolute value of the centered benchmark statistic, as decomposed in eq. (5): the inverse causality $\mathcal{C}_\omega$ (blue solid), the corrected sum of cross-products $\mathcal{T}^c_{2\omega}$ (orange solid) and, in absolute value, the sum of squares $T_{1\omega}$ (yellow solid) after being centered (see Proposition 2). All reported statistics are in their finite-sample versions (Appendix C.1) with $\varpi = 0$.

## D.3 Diercks et al. (2024): Baker et al. (2016)'s shocks



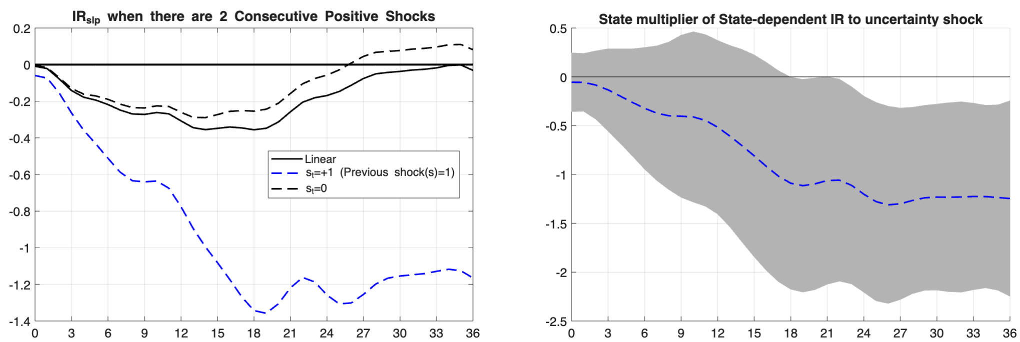

(a) Time series of Baker et al. (2016)'s EPU, PCE and S&P 500 price indexes in percentage growth.



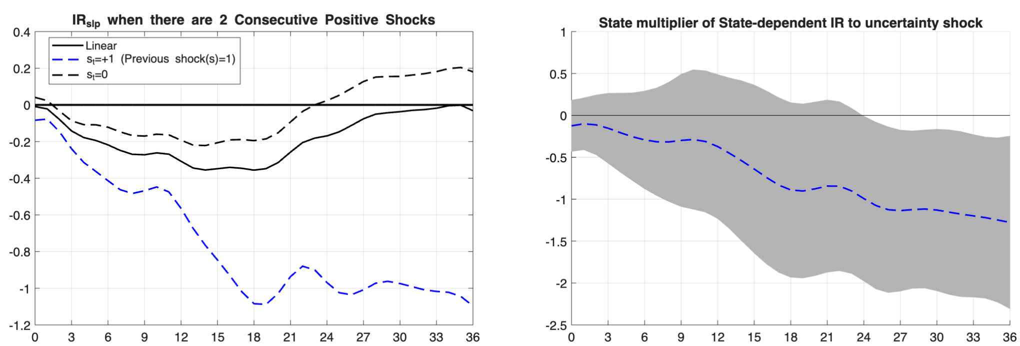

(b) Time series of the estimated EPU shock and the estimated projection of the shock series on the two lags of McCracken and Ng (2016)'s macroeconomic factors.

Figure 25: Time series and shocks:
Parallel to Figure 1 in Diercks et al. (2024), the upper panel displays the time series of EPU, together with the time series associated to inflation and stock market in percentage growth (i.e, (current/previous $-$ 1) $\times$ 100). The lower panel displays the estimated EPU shock series and its part that correlates with the past of the macroeconomic factors. The shaded areas represent NBER (National Bureau of Economic Research) recessions.

## D.4 Responses of the other variables to Baker et al. (2016)'s shocks

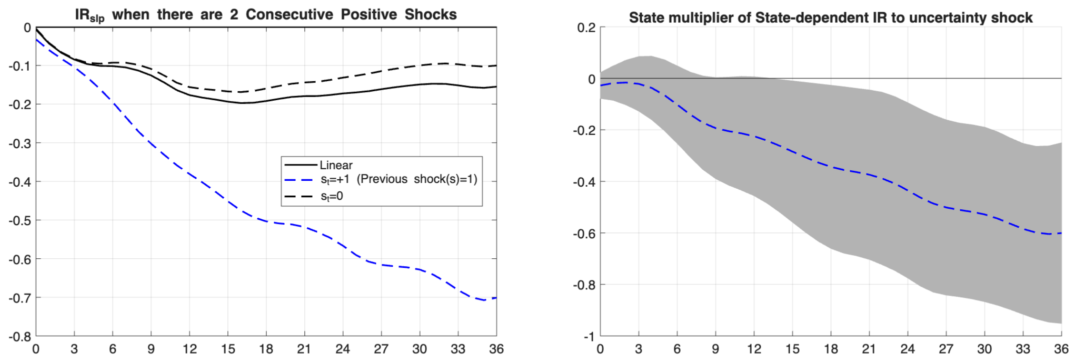



(a) Replication of Figure 2.E-F in Diercks et al. (2024)

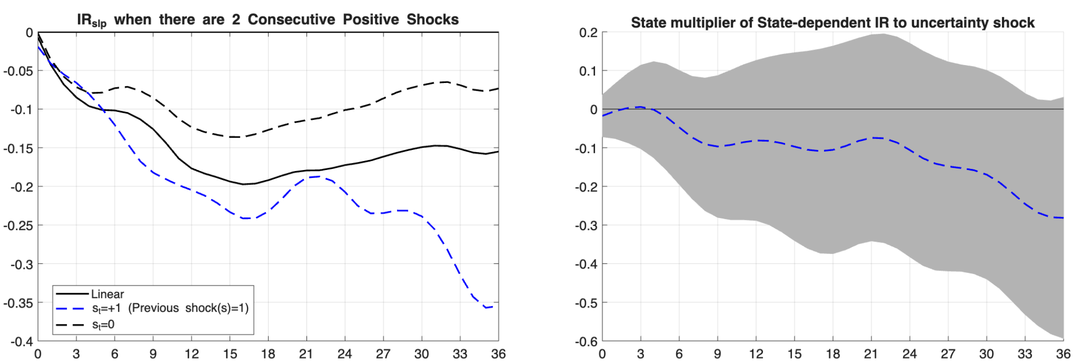



(b) Inclusion of two lags of McCracken and Ng (2016)'s factors to the set of controls

Figure 26: Response of industrial production to consecutive positive EPU uncertainty shocks:
LEFT PANELS: the empirical state-dependent impulse responses (estimated with LPs as in Diercks et al. (2024)) to two consecutive positive uncertainty shocks (dashed blue line) and contrast it to the response to a single shock in the state-dependent model (dashed black line), and in the linear model (solid black line). RIGHT PANELS: the incremental effect of the second shock, i.e. $\{\beta_{1,h}\}_{h=1,..,H}$, with 90% confidence intervals (shaded area). In both panels, on the y-axes, the level of impulse responses; on the x-axes, the horizons, $h$.



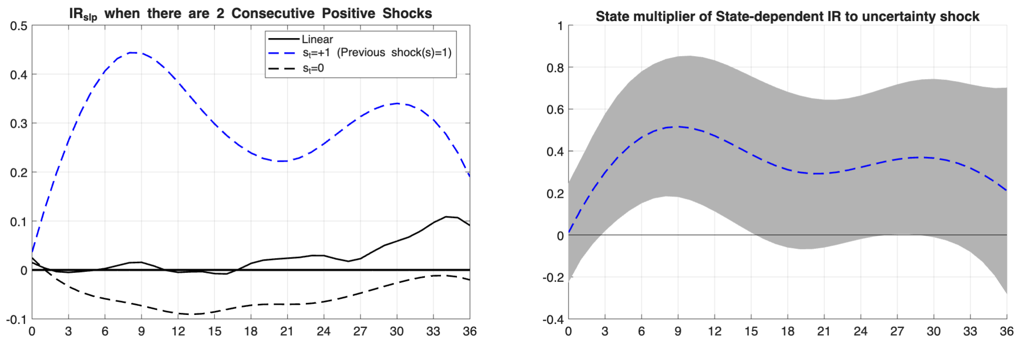

(a) Replication of Figure B.1(e)-(f) in Diercks et al. (2024)



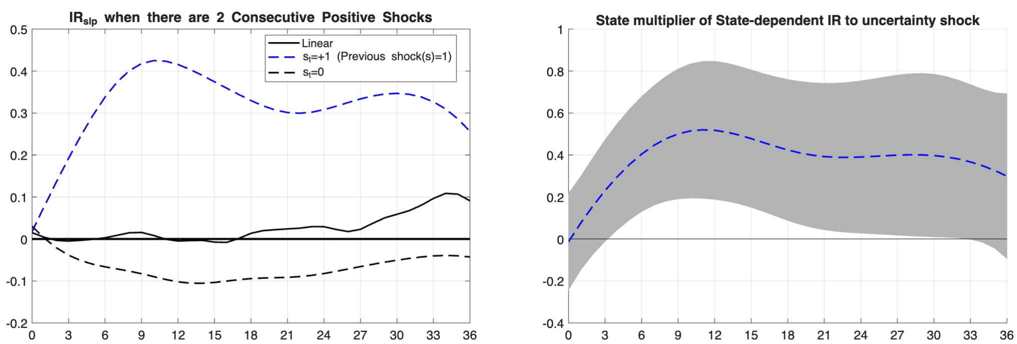

(b) Inclusion of two lags of McCracken and Ng (2016)'s factors to the set of controls

Figure 27: Response of short rate to consecutive positive EPU uncertainty shocks: LEFT PANELS: the empirical state-dependent impulse responses (estimated with LPs as in Diercks et al. (2024)) to two consecutive positive uncertainty shocks (dashed blue line) and contrast it to the response to a single shock in the state-dependent model (dashed black line), and in the linear model (solid black line). RIGHT PANELS: the incremental effect of the second shock, i.e. $\{\beta_{1,h}\}_{h=1,..,H}$, with 90% confidence intervals (shaded area). In both panels, on the y-axes, the level of impulse responses; on the x-axes, the horizons, $h$.

## D.5 Responses to Baker et al. (2016)'s shocks: Giglio and Xiu (2021) and McCracken and Ng (2016)'s factors

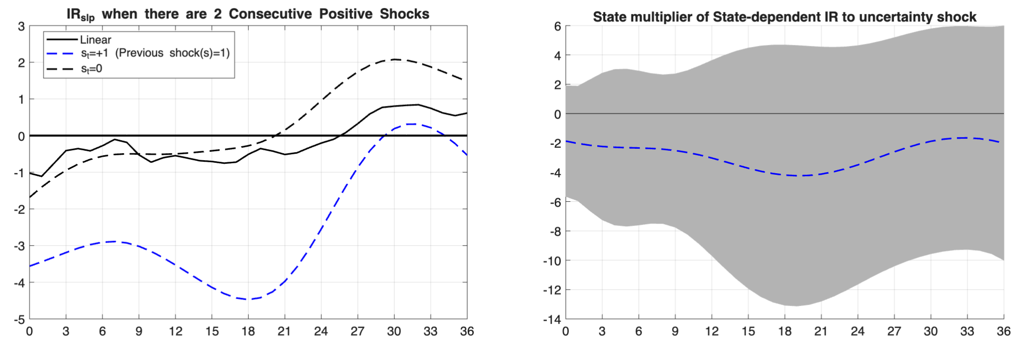



(a) Diercks et al. (2024)'s baseline specification (1986/01-2017/12).

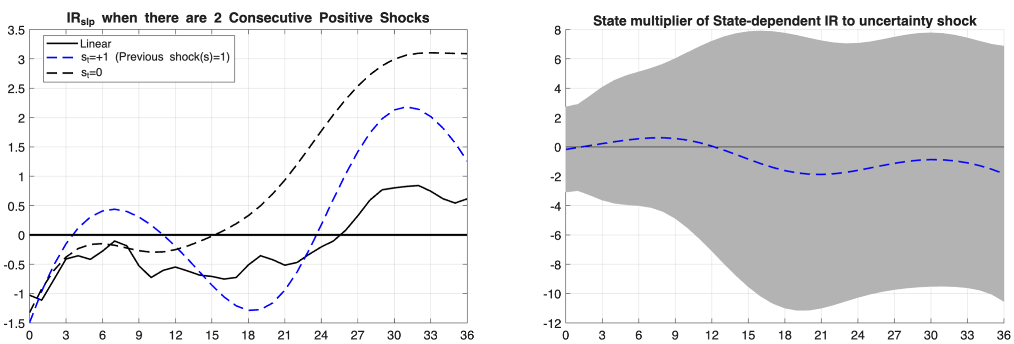



(b) Inclusion of one lag of McCracken and Ng (2016)'s and Giglio and Xiu (2021)'s factors to the controls.

Figure 28: Response of price level to consecutive positive EPU uncertainty shocks (1986/01-2017/12):
LEFT PANELS: the empirical state-dependent impulse responses (estimated with LPs as in Diercks et al. (2024)) to two consecutive positive uncertainty shocks (dashed blue line) and contrast it to the response to a single shock in the state-dependent model (dashed black line), and in the linear model (solid black line). RIGHT PANELS: the incremental effect of the second shock, i.e. $\{\beta_{1,h}\}_{h=1,..,H}$, with 90% confidence intervals (shaded area). In both panels, on the y-axes, the level of impulse responses; on the x-axes, the horizons, $h$.

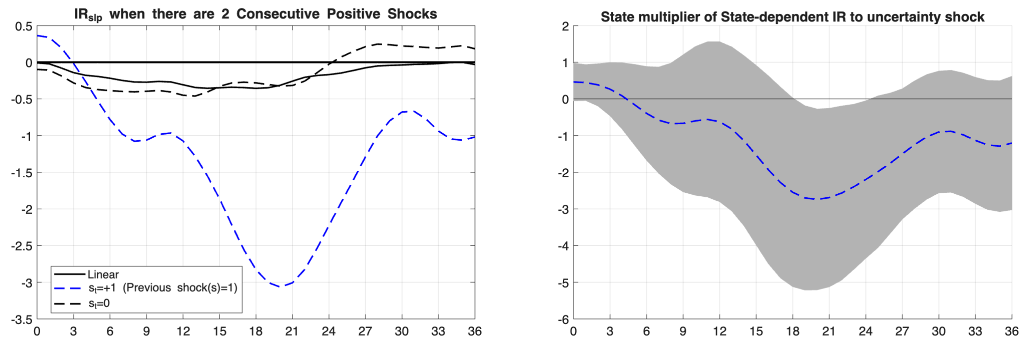



(a) Diercks et al. (2024)'s baseline specification (1986/01-2017/12).

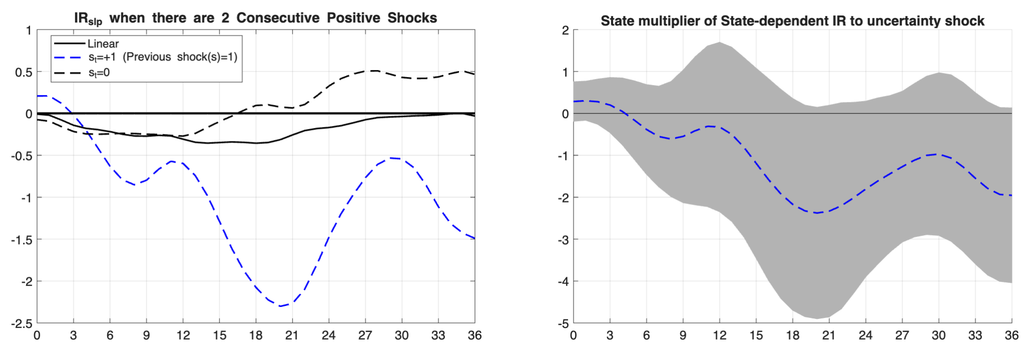



(b) Inclusion of one lag of McCracken and Ng (2016)'s and Giglio and Xiu (2021)'s factors to the controls.

Figure 29: Response of stock market to consecutive positive EPU uncertainty shocks (1986/01-2017/12):
LEFT PANELS: the empirical state-dependent impulse responses (estimated with LPs as in Diercks et al. (2024)) to two consecutive positive uncertainty shocks (dashed blue line) and contrast it to the response to a single shock in the state-dependent model (dashed black line), and in the linear model (solid black line). RIGHT PANELS: the incremental effect of the second shock, i.e. $\{\beta_{1,h}\}_{h=1,..,H}$, with 90% confidence intervals (shaded area). In both panels, on the y-axes, the level of impulse responses; on the x-axes, the horizons, $h$.

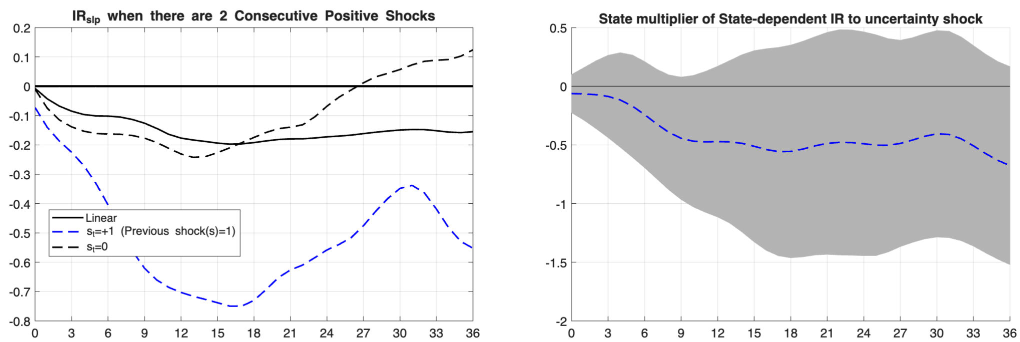



(a) Diercks et al. (2024)'s baseline specification (1986/01-2017/12).

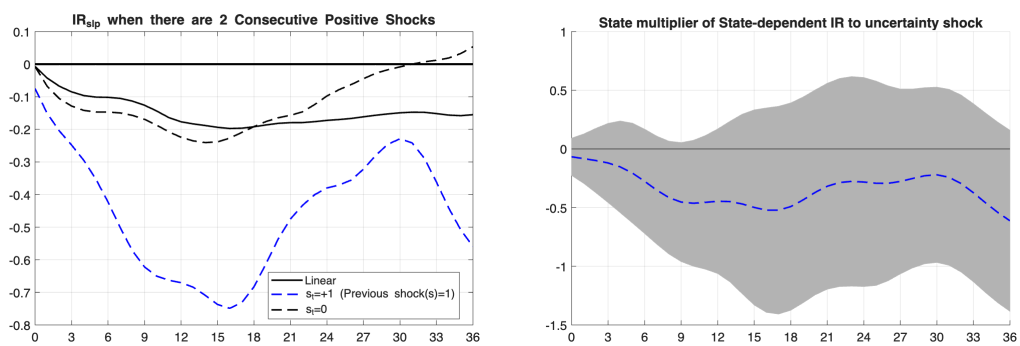



(b) Inclusion of one lag of McCracken and Ng (2016)'s and Giglio and Xiu (2021)'s factors to the controls.

Figure 30: Response of industrial production to consecutive positive EPU uncertainty shocks (1986/01-2017/12):
LEFT PANELS: the empirical state-dependent impulse responses (estimated with LPs as in Diercks et al. (2024)) to two consecutive positive uncertainty shocks (dashed blue line) and contrast it to the response to a single shock in the state-dependent model (dashed black line), and in the linear model (solid black line). RIGHT PANELS: the incremental effect of the second shock, i.e. $\{\beta_{1,h}\}_{h=1,..,H}$, with 90% confidence intervals (shaded area). In both panels, on the y-axes, the level of impulse responses; on the x-axes, the horizons, $h$.




(a) Diercks et al. (2024)'s baseline specification (1986/01-2017/12).




(b) Inclusion of one lag of McCracken and Ng (2016)'s and Giglio and Xiu (2021)'s factors to the controls.

Figure 31: Response of short rate to consecutive positive EPU uncertainty shocks (1986/01-2017/12):
LEFT PANELS: the empirical state-dependent impulse responses (estimated with LPs as in Diercks et al. (2024)) to two consecutive positive uncertainty shocks (dashed blue line) and contrast it to the response to a single shock in the state-dependent model (dashed black line), and in the linear model (solid black line). RIGHT PANELS: the incremental effect of the second shock, i.e. $\{\beta_{1,h}\}_{h=1,..,H}$, with 90% confidence intervals (shaded area). In both panels, on the y-axes, the level of impulse responses; on the x-axes, the horizons, $h$.