TECHNICAL APPENDIX FOR "*EFFICIENT TESTS OF LONG-RUN CAUSATION IN TRIVARIATE VAR PROCESSES WITH A ROLLING WINDOW STUDY OF THE MONEY-INCOME RELATIONSHIP*"

JONATHAN B. HILL

In this appendix we expand upon causality chains (Appendix B.1), the parametric bootstrap procedure for deriving p-values (Appendix B.2), we perform a simulation study (Appendix B.3), we present tables and figures (for rolling windows) for the empirical study, including all test results based on standard and bootstrap methods (Appendices B.4 and B.5), and prove all results (Appendix B.6).

Appendix B.1: Causality Chains

In order to understand what is required for non-causation to occur through some arbitrary horizon $h \geq 2$, consider $h = 2$. If $Y \overset{1}{\nrightarrow} X|I_{XZ}$ and (2.1) hold, then the orthogonal 1-step ahead projection of X_{t+2} is exactly

$$\hat{X}_{t+2}|I_W(t+1) = \sum_{i=1}^{\infty} \pi_{XX,i} X_{t+2-i} + \sum_{i=1}^{\infty} \pi_{XZ,i} Z_{t+2-i}.$$

Whether Y causes X at any other horizon $h \geq 2$ depends on a causal chain through Z (Theorem 2.1.*iii,iv*), and therefore on the coefficients $\pi_{XZ,i}$. Projecting both sides onto $I_{XZ}(t) + Y(-\infty, t]$, we obtain the best 2-step ahead forecast of X_{t+2} by iterated projections and $Y \overset{1}{\nrightarrow} X|I_{XZ}$

$$\begin{aligned}\hat{X}_{t+2}|I_W(t) &= \pi_{XX,1}\hat{X}_{t+1}|I_{XZ}(t) + \pi_{XZ,1}\hat{Z}_{t+1}|I_W(t) \\ &\quad + \sum_{i=2}^{\infty} \pi_{XX,i} X_{t+2-i} + \sum_{i=2}^{\infty} \pi_{XZ,i} Z_{t+2-i}.\end{aligned}$$

Clearly $\hat{X}_{t+2}|I_W(t) \in I_{XZ}(t)$ such that $Y \overset{2}{\nrightarrow} X|I_{XZ}$ *if and only if* $\pi_{XZ,1}\hat{Z}_{t+1}|I_W(t) \in I_{XZ}(t)$ with probability one for all t. If Z is vector-valued, then $\pi_{XZ,1}\hat{Z}_{t+1}|I_W(t) \in I_{XZ}(t)$ is feasible simply via nonlinear row-column combinations and the cancellation of Y-components.

Because the span $I_{XZ}(t) + Y(-\infty, t]$ can be written as $X(-\infty, t] + Y(-\infty, t] + Z(-\infty, t]$, we may write $\hat{Z}_{t+1}|I_W(t) = a_{zx}(t) + a_{zy}(t) + a_{zz}(t)$ for some elements $a_{zx}(t) \in X(-\infty, t]$, $a_{zy}(t) \in Y(-\infty, t]$ and $a_{zz}(t) \in Z(-\infty, t]$. Hence, $\hat{X}_{t+2}|I_W(t) \in I_{XZ}(t)$ *if and only if*

$$\pi_{XZ,1}\hat{Z}_{t+1}|I_W(t) \in I_{XZ}(t) \Rightarrow \pi_{XZ,1} a_{zy}(t) \in I_{XZ}(t).$$

Dept. of Economics, Florida International University, Miami, Fl; jonathan.hill@fiu.edu; www.fiu.edu/∼hilljona.

If the element $a_{zy}(t) \in I_{XZ}(t)$ for all t[1], then $\hat{Z}_{t+1}|I_W(t) \in I_{XZ}(t)$ and $Y \overset{1}{\nrightarrow} Z|I_{XZ}$. Conversely, if $Y \overset{1}{\rightarrow} Z$ then $a_{zy}(t) \neq 0$ with probability one for some t, hence $\pi_{XZ,1}a_{zy}(t) \in I_{XZ}(t)$ for all t *if and only if* $\pi_{XZ,1}a_{zy}(t) = 0$ with probability one for all t. If Z is scalar-valued then $\pi_{XZ,1}a_{zy}(t) = 0$ *if and only if* $\pi_{XZ,1} = 0$[2].

Appendix B.2: Parametric Bootstrap

We briefly outline the standard procedure for deriving p-values by parametric bootstrap. See, also, Dufour *et al* (2003) and Dufour (2005), and see Proposition 6.1 of Dufour (2005) for a proof of first-order asymptotic validity under standard assumptions.

The parametric bootstrap[3] is performed as follows for an arbitrary hypothesis: *i*. obtain estimated VAR coefficients, $\hat{\pi} = (\hat{\pi}_1, ..., \hat{\pi}_p)$, where p minimizes the AIC; *ii*. derive the test statistic, denoted T_n; *iii*. simulate J series $W_{t,j}$, $j = 1...J$, $t = 1...n$, based on the the estimated parameters $\hat{\pi}$ with the null hypothesis restrictions imposed (for example, a test of $Y \overset{1}{\nrightarrow} X$ imposes $\hat{\pi}_{XY,i} = 0$, $i = 1...p$): the process $W_{t,j}$ is simulated as $W_{t,j} = \sum_{i=1}^{p} \hat{\pi}_i W_{t-1} + \epsilon_t$ where ϵ_t are 3-vector *iid* draws from a standard normal distribution; *iv*. use the double-array $\{W_{t,j}\}_{t,j=1}^{n,J}$ to estimate J separate VAR(p) models, and generate J test statistics $T_{n,j}$ for the hypothesis in question; *v*. the approximate p-value is simply the percent frequency of the event $T_{n,j} > T_n$.

Appendix B.3: Simulation Study

In order to analyze the performance of the above test procedure we perform a controlled experiment for derivation of empirical test size and power for various VAR and VARMA processes. We perform Wald tests and compute p-values based on the asymptotic distribution and based on the parametric bootstrap method.

Set Up

For our study, we generated VAR(6) and vector MA(1) processes under the null of non-causation at all horizons, and under alternatives of causation at horizons $h = 1, 2,$ and 3. In all cases, $m_x = m_y = m_z = 1$ such that $m = 3$, sample sizes are restricted to $T \in \{100, 200, 300, 400, 500\}$ and 1000 series are generated for each test.

VAR(6) Construction and Hypotheses

For the VAR(6) process we simulate $W_t = \sum_{i=1}^{6} \pi_i W_{t-i} + \epsilon_t$, where ϵ_t denotes an *iid* 3-vector with mutually independent components $\epsilon_t = (\epsilon_{x,t}, \epsilon_{y,t}, \epsilon_{zt,})'$ drawn from a standard normal distribution. The matrix coefficients π_i are generated as uniform *iid* random numbers from the cube $[-.5, .5]^3$: we use $\pi = (\pi_1, ..., \pi_6)$ only

---

[1] Under the maintained assumptions this is possible only if $a_{zy}(t) = 0$ with probability one for all t.

[2] If $a_{zy}(t) = 0$ with probability one for all t, then $\hat{Z}_{t+1}|I_W(t) \in I_{XZ}(t)$, a contradiction of the assumption $Y \overset{1}{\rightarrow} Z|I_{XZ}$.

[3] This has also been refered to as an "asymptotic Monte Carlo test based on a consistent point estimate": see Dufour (2005).

if the resulting characteristic polynomial $I_3 - \pi_1 z - ... - \pi_6 z^6$ has all roots outside the unit circle, ensuring stability.

During the simulation process we impose the following restrictions (or lack, thereof), depending upon the hypothesis to be tested:

$$\begin{aligned}
H_0^{\infty} &: \pi_{XY,i} = \pi_{XZ,i} = 0, i = 1...6 \\
H_1^1 &: \pi_{XY,i} \neq 0, i = 1...6 \\
H_1^2 &: \pi_{XY,i} = 0, i = 1...6,\ \pi_{ZY,i} \neq 0,\ \pi_{XZ,i} \neq 0,\ i = 1...6 \\
H_1^3 &: \pi_{XY,i} = 0, i = 1...6,\ \pi_{ZY,i} \neq 0,\ i = 1...6,\ \pi_{XZ,i} \neq 0,\ i = 2...6
\end{aligned}$$

Under H_0^{∞}, we deduce $Y \overset{1}{\nrightarrow} (X, Z)|I_{XZ}$, cf. Theorem 2.2, and therefore Y never causes X, $Y \overset{(\infty)}{\nrightarrow} X|I_{XZ}$, cf. Theorem 2.1. Under H_1^1, Y causes X at horizon $h = 1$. Under H_1^2, non-causation $Y \overset{1}{\nrightarrow} X|I_{XZ}$, and causation $Y \overset{1}{\rightarrow} Z \overset{1}{\rightarrow} X$ are true, with $\pi_{XZ,1} \neq 0$, thus $Y \overset{2}{\rightarrow} X|I_{XZ}$ is true, cf. Theorem 3.2. Finally, under H_1^3, $Y \overset{1}{\nrightarrow} X|I_{XZ}$, $Y \overset{1}{\rightarrow} Z \overset{1}{\rightarrow} X$, $\pi_{XZ,1} = 0$ and $\pi_{XZ,2} \neq 0$, thus $Y \overset{(2)}{\nrightarrow} X|I_{XZ}$ and $Y \overset{3}{\rightarrow} X|I_{XZ}$ are true, cf. Theorem 3.2.

VMA(1) Construction and Hypotheses

For the VMA(1) processes, we simulate $W_t = \theta\epsilon_{t-1} + \epsilon_t$ by drawing *iid* uniform numbers θ from the cube $[-.9, .9]^3$, retaining only those matrices θ with characteristic roots outside the unit circle, ensuring invertability. We employ the following restrictions:

$$\begin{aligned}
H_0 &: \theta_{XY} = \theta_{XZ} = 0 \\
H_1 &: \theta_{XY} = 0.
\end{aligned} \tag{0.1}$$

In order to deduce that nature of multiple horizon non-causation for invertible VMA processes in VAR form, we require necessary and sufficient conditions for VAR noncausality in terms of VARMA coefficients. Boudjellaba *et al* (1992) derive reasonably simple necessary and sufficient conditions for non-causality at horizon $h = 1$ for such processes. Consider the general VARMA(p, q) process in lag form

$$\Phi(L)W(t) = \Theta(L)\epsilon(1), \tag{0.2}$$

where $\Phi(L)$ and $\Theta(L)$ denote the associated p^{th} and q^{th}-order lag $m \times m$ matrix-polynomials

$$\Phi(L) = I_m - \sum_{i=1}^{p} \phi_i L^i, \quad \Theta(L) = I_m + \sum_{i=1}^{q} \theta_i L^i \tag{0.3}$$

It is assumed that the polynomials do not have common roots, and all roots lie outside the unit circle. By Theorem 1 of Boudjellaba *et al* (1992, 1994), non-causality from scalar W_i to scalar W_j exists *if and only if*

$$\det\left(\Phi_i(z), \Theta_{(j)}(z)\right) = 0,\ |z| < \delta, \tag{0.4}$$

for some $\delta > 0$, where $\Phi_i(z)$ denotes the i^{th} column of $\Phi(z)$ and $\Theta_{(j)}(z)$ denotes the matrix $\Theta(z)$ with the j^{th} column removed. In the 3-vector MA(1) case with $W = (W_1, W_2, W_3)' = (X, Y, Z)'$, it follows that $\Phi(z) = I_m$, and $Y \overset{1}{\nrightarrow} X$ holds *if and* *only if*

$$\begin{aligned} & \det\left(\Phi_2(z), \Theta_{(1)}(z)\right) \\ = & \det\begin{pmatrix} 0 & \theta_{12}z & \theta_{13}z \\ 1 & 1+\theta_{22}z & \theta_{23}z \\ 0 & \theta_{32}z & 1+\theta_{33}z \end{pmatrix} \\ = & \theta_{13}\theta_{32}z^2 - \theta_{12}z - \theta_{12}\theta_{33}z^2 \\ = & 0, \ |z| < \delta. \end{aligned} \tag{0.5}$$

This occurs for *every* complex $|z| < \delta$, $\delta > 0$, *if and only if* $\theta_{12} = \theta_{13}\theta_{32} = 0$. Similarly, $Y \overset{1}{\nrightarrow} Z$ *if and only if* $\theta_{32} = \theta_{31}\theta_{12} = 0$, and $Z \overset{1}{\nrightarrow} X$ *if and only if* $\theta_{13} = \theta_{12}\theta_{23} = 0$. Consult Boudjellaba *et al* (1992, 1994), Dufour and Tessier (1993), and Dufour and Renault (1998) for further details on parametric conditions of non-causation at $h = 1$ for VARMA processes.

From the above details and Theorem 2.1, we deduce the hypothesis of non-causation at all horizons $Y \overset{(\infty)}{\nrightarrow} X|I_{XZ}$ is true *if and only if* $\theta_{12} = \theta_{13}\theta_{32} = 0$ and either $\theta_{32} = \theta_{31}\theta_{12} = 0$ and/or $\theta_{13} = \theta_{12}\theta_{23} = 0$. Therefore, the VMA(1) coefficients in (14) under H_0 in fact satisfy $Y \overset{(\infty)}{\nrightarrow} X|I_{XZ}$: the identity $\theta_{12} = \theta_{13} = 0$ (i.e. $\theta_{XY} = \theta_{XZ} = 0$) implies $Y \overset{1}{\nrightarrow} X$ and $Z \overset{1}{\nrightarrow} X$, therefore $Y \overset{(\infty)}{\nrightarrow} X|I_{XZ}$.

It is interesting to point out in the 3-vector MA(1) case that either non-causation at all horizons $Y \overset{(\infty)}{\nrightarrow} X|I_{XZ}$ or standard causation $Y \overset{1}{\rightarrow} X|I_{XZ}$ must be true, similar to the bivariate VAR case. Consider if non-causation is true $Y \overset{1}{\nrightarrow} X|I_{XZ}$, then either $\theta_{12} = \theta_{32} = 0$ and/or $\theta_{12} = \theta_{13} = 0$ must be true: in the former case $Y \overset{1}{\nrightarrow} Z$ follows, and in the latter case $Z \overset{1}{\nrightarrow} X$ follows. In either case, a causal chain does not exist, and Theorem 2.1 implies $Y \overset{(\infty)}{\nrightarrow} X|I_{XZ}$. Therefore, for 3-vector MA(1) vector-processes, $Y \overset{1}{\nrightarrow} X|I_{XZ}$ *if and only if* $Y \overset{(\infty)}{\nrightarrow} X|I_{XZ}$, which implies causation lags and causal neutralization are impossible. Thus, we deduce under H_1 in (6.1) that causation $Y \overset{1}{\rightarrow} X|I_{XZ}$ is true.

For each simulated series $\{W_t\}_{t=1}^n$ a minimum AIC method is employed for VAR order p selection, the VAR coefficients are estimated, and standard Wald tests are implemented for the linear compound hypotheses. All tests are performed at the 5%-level. We compute p-values by using the asymptotic chi-squared distribution and by parametric bootstrap.

Simulation Results

Tables A and B below contain all simulation results. Columns in each table contain empirical rejection frequencies based on p-values derived from the asymptotic chi-squared distribution, and the empirical bootstrap method [in brackets]. Tests at horizon $h = 0$ are tests of noncausation at all horizons: we fail to reject $Y \overset{(\infty)}{\nrightarrow} X|I_{XZ}$ for some series $\{W_t\}_{t=1}^n$ when we fail to reject $Y \overset{1}{\nrightarrow} X$, and fail to reject either $Y \overset{1}{\nrightarrow} Z$ and/or $Z \overset{1}{\nrightarrow} X$. For tests at individual horizons $h \geq 2$ we detect causation $Y \overset{h}{\rightarrow} X|I_{XZ}$ when we reject the compound hypothesis $Y \overset{1}{\nrightarrow} X$, $\pi_{XZ,i} = 0, i = 1...h-1$.

VAR Simulations

Consider the results for VAR processes based on p-values derived from the asymptotic distribution. For processes that satisfy $Y \overset{(\infty)}{\nrightarrow} X | I_{XZ}$ and for the sample size $n = 500$, rejection frequencies at horizons $h \geq 1$ are not far from the nominal level of 5% for tests of noncausation at all horizons: empirical sizes at $h \geq 1$ ranged from .066 to .076.

When causation occurs at horizons $h \geq 1$, tests rarely suggest noncausation at all horizons occur: evidence for noncausation at all horizons in such cases occurred in 7.2% or fewer of simulated series for $n \geq 300$, and for $n = 500$ in 2% or fewer of such series.

Moreover, when causation occurs exactly one-step ahead, rejection frequencies at $h \geq 1$ reach above 90% for sample sizes $n \geq 300$. For the same sample size range noncausation in all horizons is detected in fewer than 5% of all such series.

When a one-period causal delay exists such that $Y \overset{1}{\nrightarrow} X$ and $Y \overset{2}{\rightarrow} X$, again standard tests work reasonably well, generating empirical sizes at $h = 1$ near the 5%-level (.064 with $n = 500$), and producing reasonable empirical powers at subsequent horizons $h \geq 2$ (.812 with $n = 500$ at $h = 3$).

However, when causation occurs at horizon $h = 3$ (i.e. $Y \overset{(2)}{\nrightarrow} X$ and $Y \overset{3}{\rightarrow} X$), noticeable size distortions occur for tests at lower horizons 1 and 2. For such tests, empirical sizes approach .20 for nominal levels of 5% and $n \geq 300$. This implies we are more likely to detect causation at low horizons when in fact true causal delays are longer.

Bootstrapped p-values clearly provide better size approximations to the null distribution than standard p-values. However, even the bootstrapped p-values lead to over-rejections of the fundamental null of noncausality when non-causality occurs at all horizons: for sample sizes under 400, rejection rates reached 10% for tests at the 5%-level. Encouragingly, however, for sample sizes $n \geq 400$, rejection rates were very close to the nominal level.

When a causal lag exists, empirical sizes are again near the nominal level, however empirical powers are noticeably low. For example, with a sample size of 500 and a true causal lag of 2 periods such that $Y \overset{(2)}{\nrightarrow} X$ and $Y \overset{3}{\rightarrow} X$ are true, the bootstrap test detected causation at $h = 3$ in under 48% of simulated series.

For both asymptotic and bootstrap tests, however, empirical power diminishes severely as the horizon of causation increases. When causation occurs at $h = 1$, powers reach above 90% even for small n. However, when causation occurs at $h = 3$, powers drop to under 70% for the standard tests, and below 50% for bootstrap tests.

We argue that this evidence alone portrays a far more complicated picture of the relative merits of standard and bootstrap tests than typically argued in the literature. Neither method generates both competitive empirical sizes and powers in a benchmark Gaussian VAR environment in which model coefficients are randomly generated. Which method we favor in practice depends on whether we favor a conservative test with low power (bootstrap test), or a liberal test with excessive probability of a Type I error for some hypotheses (conventional test).

VMA Simulations

Next, consider test results for VMA(1) processes. For series in which $Y \overset{(\infty)}{\nrightarrow} X$ and for small sample sizes, standard asymptotic tests produce large empirical sizes for the fundamental tests of non-causation, in particular for tests at horizons $h \geq 2$. For $n \geq 400$, however, erroneous detection of causality dropped to frequencies of 5.1%-8.9% for tests of noncausation at horizons $h = 1...3$.

It is important to point out that for tests of noncausation at all horizons, in 95.8% (95.9%) of all series with $n = 400$ (500) did tests *correctly* conclude noncausation occurred at all horizons, which implies an the effective empirical size is 4.2% (4.1%) based on this fundamental hypothesis. Bootstrap tests, by comparison, generated empirical sizes near the nominal 5%-level for tests at $n \geq 300$, with extreme accuracy at $n = 500$.

When causation occurs one-period ahead ($Y \overset{1}{\to} X$), both tests work well when judged by whether they detect causation at all, although standard tests uniformly perform better. However, both tests struggle with tests of noncausation at all horizons: for a large sample size $n = 500$, standard (bootstrap) tests incorrectly detect $Y \overset{(\infty)}{\nrightarrow} X$ in 36.7% (32.8%) of all series

Noticeable lags exist before either test method leads to the correct detection of causation. In general, tests are sensitive to causation at $h \geq 2$, and comparatively weak at $h = 1$. For example, with $n = 400$ (500) standard tests correctly detect causation one-step ahead in only 18.3% (31.5%) of all series, however causation is detected at $h = 2$ in 70.1% (81.9%) of all series. It seems that a relatively large sample size would be required in order for empirical rejection rates at $h = 1$ to reach a reasonable level, in particular for the bootstrap tests.

The characteristic that causal relationships in VMA processes may not be sufficiently detected using VAR models presents a clear case for the need to implement multiple-horizon causality tests. Beyond the obvious necessity for such tests when true causal lags exists, even when a causal lag is impossible Wald tests may not be able to detect causation in the classic sense of a one-step ahead VAR forecast improvement. If classic tests at $h = 1$ were the only tests performed in the present setting, empirical power would be a dismal .315 (.328) for a sample size $n = 500$ based on the asymptotic (bootstrap) distribution. However, if we generalize the concept of power to engross the probability of detecting causation at *any horizon* at or beyond the true horizon of causality (in this case, $h = 1$), power reaches 82.3% (78.9%).

Table A

Empirical Size and Power:VAR(6)

$H_0^{\infty}: Y \overset{(\infty)}{\nrightarrow} X \vert I_{XZ}$					$H_1^1: Y \overset{1}{\to} X \vert I_{XZ}$				
n	h=0[a]	h=1	h=2	h=3	n	h=0	h=1	h=2	h=3
100	.955	.092[b]	.136	.137	100	.596	.623	.833	.886
	[.946]	[.055][c]	[.063]	[.047]		[.535]	[.377]	[.395]	[.465]
200	.909	.149	.201	.209	200	.133	.945	.973	.992
	[.937]	[.104]	[.094]	[.086]		[.102]	[.906]	[.945]	[.945]
300	.913	.081	.145	.162	300	.050	.967	.983	.989
	[.939]	[.070]	[.064]	[.058]		[.028]	[.961]	[.978]	[.983]
400	.885	.076	.111	.124	400	.026	.992	.994	.998
	[.899]	[.065]	[.059]	[.061]		[.022]	[.986]	[.996]	[.998]
500	.946	.066	.070	.076	500	.003	.993	.993	.999
	[.941]	[.061]	[.061]	[.054]		[.001]	[.989]	[.992]	[.994]

$H_1^2: Y \overset{1}{\nrightarrow} X \vert I_{XZ}, Y \overset{2}{\to} X \vert I_{XZ}$					$H_1^3: Y \overset{(2)}{\nrightarrow} X \vert I_{XZ}, Y \overset{3}{\to} X \vert I_{XZ}$				
n	h=0	h=1	h=2	h=3	n	h=0	h=1	h=2	h=3
100	.569	.119	.390	.517	100	.612	.150	.284	.400
	[.468]	[.037]	[.093]	[.167]		[.598]	[.058]	[.052]	[.120]
200	.136	.098	.432	.604	200	.186	.124	.198	.466
	[.090]	[.062]	[.244]	[.364]		[.188]	[.066]	[.058]	[.220]
300	.030	.112	.540	.658	300	.071	.071	.176	.547
	[.021]	[.040]	[.402]	[.552]		[.065]	[.053]	[.041]	[.317]
400	.015	.087	.551	.777	400	.022	.108	.186	.592
	[.012]	[.053]	[.418]	[.671]		[.018]	[.051]	[.049]	[.401]
500	.003	.064	.582	.812	500	.020	.150	.177	.642
	[.004]	[.051]	[.487]	[.735]		[.015]	[.049]	[.048]	[.473]

Notes: a. Values below "h = 0" denote series frequencies for which we fail to reject H_0^{∞}.
b. Rejection rates based on p-values derived from the chi-squared distribution.
c. Rejection rates based on p-values derived from the parametric bootstrap.

Table B

Empirical Size and Power: VMA(1)

$H_0: Y \overset{(\infty)}{\nrightarrow} X \vert I_{XZ}$					$H_1: Y \overset{1}{\to} X \vert I_{XZ}$				
n	h=0	h=1	h=2	h=3	n	h=0	h=1	h=2	h=3
100	.964	.074	.172	.197	100	.850	.033	.392	.408
	[.947]	[.039]	[.066]	[.053]		[.683]	[.025]	[.242]	[.242]
200	.900	.108	.152	.183	200	.525	.117	.675	.652
	[.892]	[.097]	[.108]	[.083]		[.375]	[.118]	[.609]	[.567]
300	.917	.034	.076	.108	300	.467	.219	.758	.759
	[.932]	[.057]	[.059]	[.068]		[.264]	[.225]	[.717]	[.729]
400	.958	.076	.118	.132	400	.442	.183	.701	.708
	[.936]	[.058]	[.057]	[.066]		[.274]	[.207]	[.669]	[.643]
500	.959	.058	.051	.068	500	.367	.315	.819	.823
	[.942]	[.050]	[.025]	[.042]		[.199]	[.328]	[.820]	[.789]

Appendix B.4: Tables and Figures

Table 1

Auxiliary: Unemployment Rate					
		Differences		Levels	
Test #	Hypothesis	p-value[a]	p-boot[b]	p-value	p-boot
Test 0.1	$\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$	.0373	.080	.0000	.000
Test 0.2	$\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$	.0086	.040	.0012	.216
Test 1.0	$\Delta m1 \overset{1}{\nrightarrow} \Delta y$	.5693	.626	.0070	.148
Test 1.1	$\Delta m1 \overset{1}{\nrightarrow} \Delta u$	.0096	.000	.0000	.000
Test 1.2	$\Delta u \overset{1}{\nrightarrow} \Delta y$	.0055	.026	.4264	.376
Test 2.0	$\Delta m1 \overset{(2)}{\nrightarrow} \Delta y$	.3182	.450	.0000	.024
Test 3.0	$\Delta m1 \overset{(3)}{\nrightarrow} \Delta y$	.0194	.092	.0000	.012
Test 4.0	$\Delta m1 \overset{(4)}{\nrightarrow} \Delta y$	.0030	.032	.0000	.020
Test 5.0	$\Delta m1 \overset{(5)}{\nrightarrow} \Delta y$	.0024	.032[c]	.0000	.008[d]
Min. AIC VAR Order p		8		10	
Ljung-Box p-value		.045		.0089	

Auxiliary: M2					
		Differences		Levels	
Test #	Hypothesis	p-value	p-boot	p-value	p-boot
Test 0.1	$\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$	.0000	.000	.0000	.000
Test 0.2	$\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$	.0093	.198	.0000	.002
Test 1.0	$\Delta m1 \overset{1}{\nrightarrow} \Delta y$	.4938	.466	.0000	.008
Test 1.1	$\Delta m1 \overset{1}{\nrightarrow} \Delta o$	.0000	.000	.0000	.000
Test 1.2	$\Delta o \overset{1}{\nrightarrow} \Delta y$	.0383	.244	.0000	.000
Test 2.0	$\Delta m1 \overset{(2)}{\nrightarrow} \Delta y$	.0092	.082	.0000	.000
Test 3.0	$\Delta m1 \overset{(3)}{\nrightarrow} \Delta y$	.0170	.126	.0000	.000
Test 4.0	$\Delta m1 \overset{(4)}{\nrightarrow} \Delta y$	.0088	.128	.0000	.000
Test 5.0	$\Delta m1 \overset{(5)}{\nrightarrow} \Delta y$	.0022	.066	.0000	.000
...		….	….	….	….
Test 11.0	$\Delta m1 \overset{(11)}{\nrightarrow} \Delta y$	.0000	.006[e]	.0000	.000[f]
Min. AIC VAR Order p		10		12	
Ljung-Box p-value		.370		.183	

Notes: a. p-values based on the chi-squared distribution; b. p-values based on parametric bootstrap.

c. Reject $\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$ at 10%-level, and reject $\Delta m1 \overset{(4)}{\nrightarrow} \Delta y$ at bounded 13%-level.

d. Fail to reject $\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$ at 10%-level; or reject $\Delta m1 \overset{(2)}{\nrightarrow} \Delta y$ at bounded 5%-level.

e. Fail to reject $\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$ at 10%-level; or reject $\Delta m1 \overset{(11)}{\nrightarrow} \Delta y$ at bounded 11%-level

f. Reject $\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$ at 1%-level, and $\Delta m1 \overset{1}{\nrightarrow} \Delta y$ at 1%-level.

Table 1 - Cont.

Auxiliary: Oil Price					
		Differences		Levels	
Test #	Hypothesis	p-value	p-boot	p-value	p-boot
Test 0.1	$\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$	.7900	.856	.0021	.020
Test 0.2	$\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$	.0524	.560	.0000	.106
Test 1.0	$\Delta m1 \overset{1}{\nrightarrow} \Delta y$	.9590	.954	.0038	.026
Test 1.1	$\Delta m1 \overset{1}{\nrightarrow} \Delta o$	.3780	.000	.0379	.000
Test 1.2	$\Delta o \overset{1}{\nrightarrow} \Delta y$	.1925	.194	.3686	.516
Test 2.0	$\Delta m1 \overset{(2)}{\nrightarrow} \Delta y$	.9343	.890	.6230	.670
...		...	...	...	...
Test 5.0	$\Delta m1 \overset{(5)}{\nrightarrow} \Delta y$	.5664	.734[a]	.8361	.850[b]
Min. AIC VAR Order p		4		6	
Ljung-Box p-value		.004		.0014	

Auxiliary: Rate Spread					
		Differences		Levels	
Test #	Hypothesis	p-value	p-boot	p-value	p-boot
Test 0.1	$\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$	.7715	.718	.0000	.000
Test 0.2	$\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$	.0081	.166	.0000	.020
Test 1.0	$\Delta m1 \overset{1}{\nrightarrow} \Delta y$	.9769	.964	.0000	.032
Test 1.1	$\Delta m1 \overset{1}{\nrightarrow} rr$	.3240	.000	.0000	.000
Test 1.2	$rr \overset{1}{\nrightarrow} \Delta y$	.0009	.034	.8579	.526
Test 2.0	$\Delta m1 \overset{(2)}{\nrightarrow} \Delta y$	.9043	.912	.0847	.402
Test 3.0	$\Delta m1 \overset{(3)}{\nrightarrow} \Delta y$	.7937	.822	.1264	.396
Test 4.0	$\Delta m1 \overset{(4)}{\nrightarrow} \Delta y$	.8208	.892	.1012	.508
Test 5.0	$\Delta m1 \overset{(5)}{\nrightarrow} \Delta y$	.6289	.710[c]	.1145	.526[d]
Min. AIC VAR Order p		6		8	
Ljung-Box p-value		.009		.002	

Notes: a. Fail to reject $\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$ at 10%-level; or fail to reject $\Delta m1 \overset{(5)}{\nrightarrow} \Delta y$.

b. Reject $\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$ at 10%-level, and reject $\Delta m1 \overset{1}{\nrightarrow} \Delta y$ at 5%-level, or fail to reject $\Delta m1 \overset{(5)}{\nrightarrow} \Delta y$ at bounded 5%-level

c. Fail to reject $\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$ at 10%-level; or fail to reject $\Delta m1 \overset{(5)}{\nrightarrow} \Delta y$.

d. Reject $\Delta m1 \overset{(\infty)}{\nrightarrow} \Delta y$ at 5%-level, and reject $\Delta m1 \overset{1}{\nrightarrow} \Delta y$ at 5%-level or fail to reject $\Delta m1 \overset{(5)}{\nrightarrow} \Delta y$ at bounded 5%-level.

Table 2

Horizon Rejection Frequencies: First Differences

Increasing Width Rolling Windows					Fixed Width Rolling Windows				
Horizon	u	o	rr	$m2$	Horizon	u	o	rr	$m2$
0[a]	.2647	.6667	.7281	.8775	0	.5245	.9510	.8137	.9069
	[.0294][b]	[.6373]	[.7647]	[.9608]		[.0294]	[.2402]	[.8725]	[.9706]
1	.0000	.0000	.0000	.2255	1	.3872	.1176	.1520	.1324
	[.0000]	[.0000]	[.0000]	[.2304]		[.1716]	[.0490]	[.1569]	[.1275]
2	.2990	.2353	.0147	.3725	2	.0000	.3775	.1521	.0931
	[.0000]	[.1569]	[.0000]	[.3676]		[.0049]	[.3824]	[.1078]	[.1029]
3	.3137	.0000	.0733	.0000	3	.2108	.0000	.0343	.0196
	[.4510]	[.0147]	[.0392]	[.0000]		[.2010]	[.0000]	[.0343]	[.0049]
4	.1422	.0000	.0000	.0000	4	.3725	.00000	.0000	.0245
	[.1275]	[.0000]	[.0000]	[.0000]		[.5588]	[.0000]	[.0000]	[.0049]
5	.0000	.0000	.1618	.0049	5	.0000	.0000	.0000	.0833
	[.0049]	[.0000]	[.0000]	[.0049]		[.0049]	[.0000]	[.0000]	[.0000]
≥ 1[c]	.7549	.2353	.2498	.6029	≥ 1	.9705	.4951	.3384	.2696
	[.5785]	[.1716]	[.0392]	[.5980]		[.9363]	[.4314]	[.2990]	[.2402]
$0, \geq 1$[d]	.1324	.0392	.2500	.4804	$0, \geq 1$	.5245	.4559	.1666	.2598
	[.0098]	[.0294]	[.0392]	[.5735]		[.0294]	[.1078]	[.2304]	[.2157]

Notes: a. h = 0 denotes noncausation at all horizons: values are window frequencies for which we fail to reject H_0^{∞}.

b. Bracketed values denote window frequencies based on bootstrapped p-values;

c. Window frequencies for causation at any horizon h $\geq$ 1.

d. Window frequencies for noncausation at all horizons, h = 0, and causation at some horizon h $\geq$ 1.

Table 2 - Cont.

Horizon Rejection Frequencies: Levels with Excess Lags

Increasing Width Rolling Windows					Fixed Width Rolling Windows				
Horizon	u	o	rr	$m2$	Horizon	u	o	rr	$m2$
0	.0000	1.000	1.000	.00000	0	.2634	.2488	1.000	.0439
	[.0000]	[1.000]	[1.000]	[.0000]		[.4098]	[.3658]	[1.000]	[.0976]
1	.6049	.0293	.4634	1.000	1	.9122	.4049	.4195	.9561
	[.5317]	[.0585]	[.4049]	[1.000]		[.9024]	[.6146]	[.4585]	[.9561]
2	.3171	.6341	.0195	.0000	2	.0146	.4537	.1902	.0439
	[.2341]	[.6683]	[.0488]	[.0000]		[.0244]	[.1707]	[.0976]	[.0390]
3	.0000	.0244	.0000	.0000	3	.0049	.0732	.0000	.0000
	[.0098]	[.0000]	[.0000]	[.0000]		[.0439]	[.0000]	[.0049]	[.0000]
4	.0000	.0000	.0000	.0000	4	.0000	.0000	.0244	.0000
	[.0000]	[.0000]	[.0000]	[.0000]		[.0000]	[.0000]	[.0098]	[.0000]
5	.0000	.0000	.0000	.0000	5	.0390	.0000	.0000	.0000
	[.1756]	[.0000]	[.0000]	[.0000]		[.0000]	[.0000]	[.0000]	[.0000]
≥ 1	.9220	.6878	.4829	1.000	≥ 1	.9707	.9318	.6341	1.000
	[.9512]	[.7268]	[.4537]	[1.000]		[.9707]	[.7853]	[.5708]	[.9951]
$0, \geq 1$	.9220	.6878	.4878	.0000	$0, \geq 1$	.2341	.2488	.6341	.0439
	[.9512]	[.7268]	[.4537]	[.0000]		[.3805]	[.2927]	[.5707]	[.0927]

Appendix B.5
Rolling Window Figures

Figure 1: Z = ue (diff., inc.)

Figure 2: Z = rr (diff., inc.)

Figure 3: Z = o (diff., inc.)

Figure 4: Z = m2 (diff., inc.)

Figure 5: Z = ue (levels, inc.)

Figure 6: Z = rr (levels, inc.)

Figure 7: Z = o (levels, inc.)

Figure 8: Z = m2 (levels, inc.)

Figure 9: Z = ue (levels, fix)

Figure 10: Z = rr (levels, fix)

Figure 11: Z = o (levels, fix)

Figure 12: Z = m2 (levels, fix)

Appendix B.6 : Formal Proofs

Proof of Theorem 2.1. Consider (iv) and assume $Y \xrightarrow{1} Z \xrightarrow{1} X$ where Z is univariate. Then either $Y \xrightarrow{1} X|I_{XZ}$ or $Y \overset{1}{\nrightarrow} X|I_{XZ}$. Suppose $Y \overset{1}{\nrightarrow} X|I_{XZ}$: we will prove $Y \overset{(h-1)}{\nrightarrow} X|I_{XZ}$ and $Y \xrightarrow{h} X|I_{XZ}$ for some $h \geq 2$.

From Lemma 2.1.2, below, for any $h \geq 2$ given $Y \xrightarrow{1} Z$, if $Y \overset{(h-1)}{\nrightarrow} X|I_{XZ}$ then $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$ *if and only if* $c^{t,1}_{z,t-h+1,t-h+1} = 0$ with probability one for every t, where $c^{t,1}_{z,t-h+1,t-h+1}$ denotes that unique component of the subspace $Z(t-h+1, t-h+1]$ that enters into the orthogonal projection of X_{t+1} onto $I_{XZ}(t) + Y(-\infty, t]$. However, because $Z \xrightarrow{1} X$, it must be the case that $c^{t,1}_{z,t-h+1,t-h+1} \neq 0$ with probability one for *some* $h \geq 2$. Therefore, by Lemma 2.1.2, $Y \overset{(h-1)}{\nrightarrow} X|I_{XZ}$ and $c^{t,1}_{z,t-h+1,t-h+1} \neq 0$ for some $h \geq 2$ implies $Y \xrightarrow{h} X|I_{XZ}$. □

Lemma 2.1.1 *Let Z be scalar-valued. Denote by $c^{t-k_3,h}_{z,t-k_1,t-k_2}$ that unique element of the span $Z(t-k_1, t-k_2]$ that enters into the projection of X_{t+h} onto $I_{XZ}(t-k_3) + Z(-\infty, t-k_3]$, $k_3 \leq k_2 \leq k_1$. For any $h \geq 2$, if $Y \overset{(h-1)}{\nrightarrow} X|I_{XZ}$ and $Y \xrightarrow{1} Z$, then $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$ if and only if $c^{t,h-1}_{z,t,t} = 0$ with probability one for every t.*

Lemma 2.1.2 *Let Z be scalar-valued. For any $h \geq 1$, if $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$ and $Y \xrightarrow{1} Z$ then $Y \overset{(h+1)}{\nrightarrow} X|I_{XZ}$ if and only if $c^{t,1}_{z,t-h+1,t-h+1} = 0$ with probability one for every t.*

Proof of Lemma 2.1.1. Assume $Y \overset{(h-1)}{\nrightarrow} X|I_{XZ}$ for some $h \geq 2$. By iterated projections for Hilbert space metric projection operators

$$\begin{aligned} &P(X_{t+h}|I_{XZ}(t) + Y(-\infty, t]) \\ &\quad = P(P(X_{t+h}|I_{XZ}(t+1) + Y(-\infty, t+1])\,|I_{XZ}(t) + Y(-\infty, t]) \\ &\quad = P(P(X_{t+h}|I_{XZ}(t+1))\,|I_{XZ}(t) + Y(-\infty, t]) \,. \end{aligned}$$

Notice $I_{XZ}(t+1)$ decomposes into

$$\begin{aligned} I_{XZ}(t+1) &= H + X(-\infty, t+1] + Z(-\infty, t+1] \\ &= H + X(-\infty, t] + X(t+1, t+1] + Z(-\infty, t] + Z(t+1, t+1] \\ &= I_{XZ}(t) + X(t+1, t+1] + Z(t+1, t+1]. \end{aligned}$$

Hence, we may write $P(X_{t+h}|I_{XZ}(t+1))$ as

$$P(X_{t+h}|I_{XZ}(t+1)) = a^{t+1,h}_{xz,t} + b^{t+1,h}_{x,t+1,t+1} + c^{t+1,h}_{z,t+1,t+1},$$

where

$$a^{t+1,h}_{xz,t} \in I_{XZ}(t), \quad b^{t+1,h}_{x,t+1,t+1} \in X(t+1, t+1], \quad c^{t+1,h}_{z,t+1,t+1} \in Z(t+1, t+1].$$

We obtain from projection operator linearity, the assumption $Y \overset{1}{\nrightarrow} X|I_{XZ}$ and $b^{t+1,h}_{x,t+1,t+1} \in X(t+1,t+1]$,

$$\begin{aligned}
&P\left(X_{t+h}|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\quad = P\left(P\left(X_{t+h}|I_{XZ}(t+1)\right)|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\quad = P\left(a^{t+1,h}_{xz,t} + b^{t+1,h}_{x,t+1,t+1} + c^{t+1,h}_{z,t+1,t+1}|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\quad = a^{t+1,h}_{xz,t} + P\left(b^{t+1,h}_{x,t+1,t+1}|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\quad + P\left(c^{t+1,h}_{z,t+1,t+1}|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\quad = a^{t+1,h}_{xz,t} + P\left(b^{t+1,h}_{x,t+1,t+1}|I_{XZ}(t)\right) + P\left(c^{t+1,h}_{z,t+1,t+1}|I_{XZ}(t)+Y(-\infty,t]\right).
\end{aligned}$$

Because $Y \overset{1}{\rightarrow} Z$, $c^{t+1,h}_{z,t+1,t+1} \in Z(t+1,t+1]$, and $Z(t+1,t+1]$ is a scalar-valued Hilbert space, we deduce $P(c^{t+1,h}_{z,t+1,t+1}|I_{XZ}(t)+Y(-\infty,t]) = P(c^{t+1,h}_{z,t+1,t+1}|I_{XZ}(t))$ with probability one *if and only if* $c^{t+1,h}_{z,t+1,t+1} = 0$ with probability one for all t[4]. Therefore if $Y \overset{(h-1)}{\nrightarrow} X|I_{XZ}$, $Y \overset{1}{\rightarrow} Z$, and Z is scalar-valued, then $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$ *if and only if* $c^{t+1,h}_{z,t+1,t+1} = 0$ with probability one for every t. Because t and h are arbitrary, we conclude $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$ *if and only if* $c^{t,h-1}_{z,t,t} = 0$ with probability one for every t. □

Proof of Lemma 2.1.2. We prove the claim by induction exploiting Lemma 2.1.1. Let $Y \overset{1}{\nrightarrow} X|I_{XZ}$. By Lemma 2.1.1, $Y \overset{(2)}{\nrightarrow} X|I_{XZ}$ *if and only if* $c^{t,1}_{z,t,t} = 0$ with probability one for every t. This proves the claim for $h=1$.

For any $h \geq 2$ assume $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$ *if and only if* $c^{t,1}_{z,t-k+2,t-k+2} = 0$ with probability one for every t and each $k = 2...h$. We will prove $Y \overset{(h+1)}{\nrightarrow} X|I_{XZ}$ *if and only if* $c^{t,1}_{z,t-h+1,t-h+1} = 0$.

By iterated projections, the assumption $Y \overset{1}{\nrightarrow} X|I_{XZ}$, and the decomposition

$$I_{XZ}(t+h) = I_{XZ}(t) + X(t+1,t+h] + Z(t+1,t+h],$$

we obtain

$$\begin{aligned}
&P\left(X_{t+h+1}|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\quad = P\left(P\left(X_{t+h+1}|I_{XZ}(t+h)+Y(-\infty,t+h]\right)|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\quad = P\left(P\left(X_{t+h+1}|I_{XZ}(t+h)\right)|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\quad = P\left(a^{t+h,h+1}_{xz,t} + b^{t+h,h+1}_{x,t+1,t+h} + c^{t+h,h+1}_{z,t+1,t+h}|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\quad = a^{t+h,h+1}_{xz,t} + P\left(b^{t+h,h+1}_{x,t+1,t+h}|I_{XZ}(t)+Y(-\infty,t]\right) \\
&\qquad + P\left(c^{t+h,h+1}_{z,t+1,t+h}|I_{XZ}(t)+Y(-\infty,t]\right).
\end{aligned}$$

[4]If Z is multivariate, then elements of the single-period span $Z(t+1,t+1]$ may contain linear combinations of the multiple $Z_{t,i}$-components, hence $c^{t+1,h}_{z,t+1,t+1} \neq 0$ would be possible while also $P(c^{t+1,h}_{z,t+1,t+1}|I_{XZ}(t)+Y(-\infty,t]) = P(c^{t+1,h}_{z,t+1,t+1}|I_{XZ}(t) = 0$ due to linearity of the projection operator and causal neutralization.

By $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$, $b^{t+h,h+1}_{x,t+1,t+h} \in X(t+1,t+h]$ and projection operator linearity, we deduce

$$P\left(b^{t+h,h+1}_{x,t+1,t+h}|I_{XZ}(t)+Y(-\infty,t]\right) = P\left(b^{t+h,h+1}_{x,t+1,t+h}|I_{XZ}(t)\right).$$

Moreover, the element $c^{t+h,h+1}_{z,t+1,t+h}$ denotes that unique component of the subspace $Z(t+1,t+h]$ that enters into the orthogonal projection of X_{t+h+1} onto $I_{XZ}(t+h)+Y(-\infty,t+h]$. Because $Z(t+1,t+h]$ decomposes into

$$Z(t+1,t+h] = Z(t+1,t+1] + ... + Z(t+h,t+h]$$

we deduce $c^{t+h,h+1}_{z,t+1,t+h}$ satisfies for every t

$$c^{t+h,h+1}_{z,t+1,t+h} = c^{t+h,h+1}_{z,t+1,t+1} + ... + c^{t+h,h+1}_{z,t+h,t+h},$$

hence, because t and h are arbitrary,

$$c^{t,1}_{z,t-h+1,t} = c^{t,1}_{z,t-h+1,t-h+1} + ... + c^{t,1}_{z,t,t}.$$

By the induction assumption $c^{t,1}_{z,t-k+2,t-k+2} = 0$ with probability one for every t and each $k = 2...h$, thus

$$c^{t+h,h+1}_{z,t+1,t+h} = c^{t+h,h+1}_{z,t+1,t+1}.$$

This implies

$$\begin{aligned} &P\left(X_{t+h+1}|I_{XZ}(t)+Y(-\infty,t]\right) \\ &\quad = a^{t+h,h+1}_{xz,t} + P\left(b^{t+h,h+1}_{x,t+1,t+h}|I_{XZ}(t)+Y(-\infty,t]\right) + P\left(c^{t+h,h+1}_{z,t,t}|I_{XZ}(t)+Y(-\infty,t]\right) \\ &\quad = a^{t+h,h+1}_{xz,t} + P\left(b^{t+h,h+1}_{x,t+1,t+h}|I_{XZ}(t)\right) + P\left(c^{t+h,h+1}_{z,t+1,t+1}|I_{XZ}(t)+Y(-\infty,t]\right). \end{aligned}$$

We deduce if $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$ then $Y \overset{(h+1)}{\nrightarrow} X|I_{XZ}$ *if and only if* $P(c^{t+h,h+1}_{z,t+1,t+1}|I_{XZ}(t) + Y(-\infty,t]) = P(c^{t+h,h+1}_{z,t+1,t+1}|I_{XZ}(t)])$ for all t. Using the logic from the line of proof of Lemma 2.1.1, because $Z(t+1,t+1]$ is a scalar-valued Hilbert space, $c^{t+h,h+1}_{z,t+1,t+1} \in Z(t+1,t+1]$, and $Y \overset{1}{\rightarrow} Z$, we deduce $Y \overset{(h+1)}{\nrightarrow} X|I_{XZ}$ *if and only if* $c^{t+h,h+1}_{z,t+1,t+1} = 0$ with probability one for all t, or $c^{t,1}_{z,t-h+1,t-h+1} = 0$ with probability one for all t. □

Proof of Lemma 3.1. Recall we assume $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$ and $Y \overset{1}{\rightarrow} Z|I_{XZ}$. From (3.3), and the assumption $m_z = 1$, we know $Y \overset{(h+1)}{\nrightarrow} X|I_{XZ}$ *if and only if* $\pi^{(h)}_{XZ,1} = 0$. Therefore, if $Y \overset{1}{\nrightarrow} X|I_{XZ}$, then $Y \overset{(2)}{\nrightarrow} X|I_{XZ}$ *if and only if* $\pi_{XZ,1} = 0$.

Now, from (2.3) we deduce

$$\pi^{(h)}_{XZ,1} = \pi^{(h-1)}_{XZ,2} + \pi^{(h-1)}_{XX,1}\pi_{XZ,1} + \pi^{(h-1)}_{XY,1}\pi_{YZ,1} + \pi^{(h-1)}_{XZ,1}\pi_{ZZ,1}. \tag{B.1}$$

For $h = 2$, $Y \overset{(2)}{\nrightarrow} X|I_{XZ}$ implies $\pi^{(1)}_{XZ,1} = 0$ from above, hence

$$\begin{aligned} \pi^{(2)}_{XZ,1} &= \pi_{XZ,2} + \pi_{XX,1}\pi_{XZ,1} + \pi_{XY,1}\pi_{YZ,1} + \pi_{XZ,1}\pi_{ZZ,1} \\ &= \pi_{XZ,2} + \pi_{XX,1} \times 0 + 0 \times \pi_{YZ,1} + 0 \times \pi_{ZZ,1} \\ &= \pi_{XZ,2}. \end{aligned}$$

For $h=3$, $Y\overset{(3)}{\nrightarrow}X|I_{XZ}$ implies $\pi_{XZ,1}^{(2)}=\pi_{XZ,1}^{(1)}=0$ from above, hence

$$\begin{aligned}\pi_{XZ,1}^{(3)} &= \pi_{XZ,2}^{(2)}+\pi_{XX,1}^{(2)}\pi_{XZ,1}+\pi_{XY,1}^{(2)}\pi_{YZ,1}+\pi_{XZ,1}^{(2)}\pi_{ZZ,1}\\ &= \pi_{XZ,2}^{(2)}+\pi_{XX,1}^{(2)}\times 0+0\times\pi_{YZ,1}+0\times\pi_{ZZ,1}\\ &= \pi_{XZ,2}^{(2)},\end{aligned}$$

where

$$\begin{aligned}\pi_{XZ,2}^{(2)} &= \pi_{XZ,3}+\pi_{XX,1}\pi_{XZ,2}+\pi_{XY,1}\pi_{YZ,2}+\pi_{XZ,1}\pi_{ZZ,2}\\ &= \pi_{XZ,3}+\pi_{XX,1}\pi_{XZ,2}+0\times\pi_{YZ,2}+0\times\pi_{ZZ,2}\\ &= \pi_{XZ,3}+\pi_{XX,1}\pi_{XZ,2}.\end{aligned}$$

Thus,

$$\pi_{XZ,1}^{(3)}=\pi_{XZ,2}^{(2)}=\pi_{XZ,3}+\pi_{XX,1}\pi_{XZ,2}.$$

Recursively we deduce for $h\geq 2$, $Y\overset{(h)}{\nrightarrow}X|I_{XZ}$ implies

$$\pi_{XZ,1}^{(h)}=\pi_{XZ,h}+\sum_{i=1}^{h-1}\pi_{XX,1}^{(h-i)}\pi_{XZ,i}. \tag{B.2}$$

Notice in (B.2) we include the term $\pi_{XX,1}^{(h-1)}\pi_{XZ,1}=0$ which follows from above: $Y\overset{(h)}{\nrightarrow}X|I_{XZ}$ for $h\geq 2$ implies $Y\overset{(2)}{\nrightarrow}X|I_{XZ}$, hence $\pi_{XZ,1}=0$. □

Proof of Theorem 3.2. *i.* Consider $h=2$ and assume $Y\overset{1}{\nrightarrow}X|I_{XZ}$. By the assumptions $Y\overset{1}{\nrightarrow}X|I_{XZ}$ and $Y\overset{1}{\rightarrow}Z|I_{XZ}$, and $m_z=1$, from (3.3) we deduce $Y\overset{(2)}{\nrightarrow}X|I_{XZ}$ *if and only if* $\pi_{XZ,1}=0$.

For the remainder of the proof, assume $h>2$. Let $Y\overset{(h)}{\nrightarrow}X|I_{XZ}$. Then $Y\overset{k}{\nrightarrow}X|I_{XZ}$, $k=1...h$, and we deduce from (B.1) that $\pi_{XZ,1}^{(k)}=0$, $k=1...h-1$. Using the zero identities $\pi_{XZ,1}^{(k)}=0$, $k=1...h-1$, we deduce from formula (B.2)

$$\begin{aligned}0 &= \pi_{XZ,1}^{(1)}=\pi_{XZ,1}\\ 0 &= \pi_{XZ,1}^{(2)}=\pi_{XZ,2}+\left(\pi_{XX,1}^{(1)}\pi_{XZ,1}\right)=\pi_{XZ,2}\\ 0 &= \pi_{XZ,1}^{(3)}=\pi_{XZ,3}+\left(\pi_{XX,1}^{(2)}\pi_{XZ,1}\right)+\left(\pi_{XX,1}^{(1)}\pi_{XZ,2}\right)\\ &= \pi_{XZ,3}+\left(\pi_{XX,1}^{(2)}\times 0\right)+\left(\pi_{XX,1}^{(1)}\times 0\right)=\pi_{XZ,3}\\ &\ldots\\ 0 &= \pi_{XZ,1}^{(h-1)}\\ &= \pi_{XZ,h-1}+\sum_{i=1}^{h-2}\left(\pi_{XX,1}^{(h-1-i)}\times 0\right)=\pi_{XZ,h-1},\end{aligned}$$

which gives $\pi_{XZ,k}=0$, $k=1...h-1$. This proves the first direction.

Conversely, let $Y\overset{1}{\nrightarrow}X|I_{XZ}$ and $\pi_{XZ,i}=0$, $i=1...h-1$. From (2) we have

$$\pi_{XZ,1}^{(k)}=\pi_{XZ,2}^{(k-1)}+\left(\pi_{XX,1}^{(k-1)}\pi_{XZ,1}+\pi_{XY,1}^{(k-1)}\pi_{YZ,1}+\pi_{XZ,1}^{(k-1)}\pi_{ZZ,1}\right).$$

If $k=1$, then trivially $\pi_{XZ,1}^{(1)}=\pi_{XZ,1}=0$, and by (B.1) $Y\overset{2}{\nrightarrow}X|I_{XZ}$ follows. Thus, along with $Y\overset{1}{\nrightarrow}X|I_{XZ}$ by assumption, we obtain $Y\overset{(2)}{\nrightarrow}X|I_{XZ}$. However, by Theorem 1 and Lemma 3.1, $Y \overset{(2)}{\nrightarrow} X|I_{XZ}$ implies $\pi^{(k)}_{XY,j} = 0$, $k = 1, 2$, $j \geq 1$ and

$$\pi^{(2)}_{XZ,1} = \pi_{XZ,2} + \sum_{i=1}^{2-1} \left(\pi^{(2-i)}_{XX,1}\pi_{XZ,i}\right) = \pi_{XZ,2} = 0.$$

Thus, by (3.3), $Y \overset{3}{\nrightarrow} X|I_{XZ}$ follows immediately, and in conjunction with $Y \overset{(2)}{\nrightarrow} X|I_{XZ}$, we deduce $Y \overset{(3)}{\nrightarrow} X|I_{XZ}$. Recursively, for each $k = 1...h-1$ that we have $Y \overset{(k)}{\nrightarrow} X|I_{XZ}$, from Lemma 3.1 we deduce

$$\pi^{(k)}_{XZ,1} = \pi_{XZ,k} + \sum_{i=1}^{k-1} \left(\pi^{(k-i)}_{XX,1}\pi_{XZ,i}\right),$$

which reduces to

$$\begin{aligned}\pi^{(k)}_{XZ,1} &= \pi_{XZ,k} + \sum_{i=2}^{k-1} \left(\pi^{(k-i)}_{XX,1}\pi_{XZ,i}\right) \\ &= 0 + \sum_{i=1}^{k-1} \left(\pi^{(k-i)}_{XX,1} \times 0\right) = 0.\end{aligned}$$

By (B.1) and $\pi^{(k)}_{XZ,1} = 0$, we deduce $Y \overset{k+1}{\nrightarrow} X|I_{XZ}$ for $k = 1...h-1$. Combined with the assumption $Y \overset{1}{\nrightarrow} X|I_{XZ}$, it follows that $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$. This proves claim (i).

ii. The result is a direct consequence of claim (i): $Y \overset{(h+1)}{\nrightarrow} X|I_{XZ}$ *if and only if* $Y \overset{1}{\nrightarrow} X|I_{XZ}$ and $\pi_{XZ,i} = 0$, $i = 1...h$. Thus, given $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$, we have $\pi_{XZ,i} = 0$, $i = 1...h-1$, hence $Y \overset{(h+1)}{\nrightarrow} X|I_{XZ}$ follows *if and only if* $\pi_{XZ,h} = 0$. □

Proof of Lemma 3.3. From (2.1), due to error orthogonality $\epsilon_t \perp W(-\infty, t-1]$

$$\begin{aligned}P\left(X_{t+h}|I_{XZ} + Y(-\infty, t]\right) &= \sum_{k=1}^{\infty} \pi^{(h)}_{XX,k} X_{t+1-k} \\ &+ \sum_{k=1}^{\infty} \pi^{(h)}_{XY,k} Y_{t+1-k} + \sum_{k=1}^{\infty} \pi^{(h)}_{XZ,k} Z_{t+1-k}.\end{aligned}$$

Projecting both sides onto the subspace $I_{XZ_1} + Y(-\infty, t] \subseteq I_{XZ} + Y(-\infty, t]$, and invoking iterated projections and projection operator linearity, we obtain

$$\begin{aligned}&P\left(P\left(X_{t+h}|I_{XZ} + Y(-\infty, t]\right)|I_{XZ_1} + Y(-\infty, t]\right) \\ &= P\left(X_{t+h}|I_{XZ_1} + Y(-\infty, t]\right) \\ &= \sum_{i=1}^{\infty} \pi^{(h)}_{XX,i} X_{t+1-i} + \sum_{i=1}^{\infty} \pi^{(h)}_{XY,i} Y_{t+1-i} \\ &+ \sum_{i=1}^{\infty} \pi^{(h)}_{XZ,i} P\left(Z_{t+1-i}|I_{XZ_1} + Y(-\infty, t]\right) \\ &= \sum_{i=1}^{\infty} \pi^{(h)}_{XX,i} X_{t+1-i} + \sum_{i=1}^{\infty} \pi^{(h)}_{XY,i} Y_{t+1-i} + \sum_{i=1}^{\infty} \pi^{(h)}_{XZ_1,i} Z_{1,t+1-i} \\ &+ \sum_{i=1}^{\infty} \pi^{(h)}_{XZ_2,i} P\left(Z_{2,t+1-i}|I_{XZ_1} + Y(-\infty, t]\right) \\ &= \sum_{i=1}^{\infty} \pi^{(h)}_{XX,i} X_{t+1-i} + \sum_{i=1}^{\infty} \pi^{(h)}_{XY,i} Y_{t+1-i} + \sum_{i=1}^{\infty} \pi^{(h)}_{XZ_1,i} Z_{1,t+1-i} \\ &+ \sum_{i=1}^{\infty} \pi^{(h)}_{XZ_2,i} \left[\sum_{k=1}^{\infty} \beta^{j,t+1-i}_{Z_2X,k} X_{t+1-k}\right] \\ &+ \sum_{i=1}^{\infty} \pi^{(h)}_{XZ_2,i} \left[\sum_{k=1}^{\infty} \beta^{j,t+1-i}_{Z_2Y,k} Y_{t+1-k} + \sum_{k=1}^{\infty} \beta^{j,t+1-i}_{Z_2Z_1,k} Z_{1,t+1-k}\right]\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{\infty} \pi_{XX,i}^{(h)} X_{t+1-i} + \sum_{i=1}^{\infty} \pi_{XY,i}^{(h)} Y_{t+1-i} + \sum_{i=1}^{\infty} \pi_{XZ_1,i}^{(h)} Z_{1,t+1-i} \\
&\quad + \sum_{k=1}^{\infty} X_{t+1-k} \left(\sum_{i=1}^{\infty} \pi_{XZ_j,i}^{(h)} \beta_{Z_2X,k}^{j,1-i} \right) \\
&\quad + \sum_{k=1}^{\infty} Y_{t+1-k} \left(\sum_{i=1}^{\infty} \pi_{XZ_j,i}^{(h)} \beta_{Z_2Y,k}^{j,1-i} \right) \\
&\quad + \sum_{k=1}^{\infty} Z_{1,t+1-k} \left(\sum_{i=1}^{\infty} \pi_{XZ_j,i}^{(h)} \beta_{Z_2Z_1,k}^{j,1-i} \right) \\
&= \sum_{k=1}^{\infty} X_{t+1-k} \left(\pi_{XX,k}^{(h)} + \sum_{i=1}^{\infty} \pi_{XZ_j,i}^{(h)} \beta_{Z_2X,k}^{j,1-i} \right) \\
&\quad + \sum_{i=k}^{\infty} Y_{t+1-k} \left(\pi_{XY,k}^{(h)} + \sum_{i=1}^{\infty} \pi_{XZ_j,i}^{(h)} \beta_{Z_2Y,k}^{j,1-i} \right) \\
&\quad + \sum_{i=k}^{\infty} Z_{1,t+1-k} \left(\pi_{XZ_1,k}^{(h)} + \sum_{i=1}^{\infty} \pi_{XZ_j,i}^{(h)} \beta_{Z_2Z_1,k}^{j,1-i} \right) \\
&= \sum_{i=k}^{\infty} \delta_{XX,k}^{(h)} X_{t+1-k} + \sum_{k=1}^{\infty} \delta_{XY,k}^{(h)} Y_{t+1-k} + \sum_{k=1}^{\infty} \delta_{XZ_1,k}^{(h)} Z_{1,t+1-k}
\end{aligned}$$

hence

$$\begin{aligned}
P\left(X_{t+h} | I_{XZ_1} + Y(-\infty, t]\right) &= \sum_{k=1}^{\infty} \delta_{XX,k}^{(h)} X_{t+1-k} \\
&\quad + \sum_{k=1}^{\infty} \delta_{XY,k}^{(h)} Y_{t+1-k} + \sum_{k=1}^{\infty} \delta_{XZ_1,k}^{(h)} Z_{1,t+1-k}
\end{aligned}$$

where

$$\delta_{XY,k}^{(h)} \equiv \pi_{XY,k}^{(h)} + \sum_{i=1}^{\infty} \pi_{XZ_2,i}^{(h)} \beta_{Z_2Y,k}^{1-i},$$

and $\beta_{Z_2Y,k}^{1-i}$ denotes the Y-specific coefficients in the projection of each vector $Z_{2,t+1-i}$ onto $I_{XZ_1} + Y(-\infty, t]$, $i \geq 1$. □

Proof of Theorem 3.4. *i.* Let $Z_2 \overset{1}{\nrightarrow} X | I_{XZ_1}$. Then $\pi_{XZ_2,j} = 0$, $\forall j \geq 1$, hence from Lemma 3.3, cf. (3.7), we obtain $\delta_{XY,j} = \pi_{XY,j}$. Therefore $\delta_{XY,j} = 0$ *if and only if* $\pi_{XY,j} = 0$, which implies $Y \overset{1}{\nrightarrow} X | I_{XZ_1}$ *if and only if* $Y \overset{1}{\nrightarrow} X | I_{XZ}$.

ii. Let $(Y, Z_2) \overset{1}{\nrightarrow} X | I_{XZ_1}$ and $Y \overset{(h)}{\nrightarrow} X | I_{XZ_1}$ for any $h \geq 1$, and recall $\delta_j^{(h)}$ denote the VAR coefficients in the projection of W_{t+h} onto $I_{XZ_1}(t) + Y(-\infty, t]$. By the assumption $Z_2 \overset{1}{\nrightarrow} X | I_{XZ_1}$, we just proved $Y \overset{(h)}{\nrightarrow} X | I_{XZ_1}$ *if and only if* $Y \overset{(h)}{\nrightarrow} X | I_{XZ}$ for $h = 1$. Therefore let $h \geq 2$. By Lemma 3.1 and Theorem 3.2, $Y \overset{(h)}{\nrightarrow} X | I_{XZ_1}$ implies $\delta_{XZ_1,1}^{(k)} = \delta_{XZ_1,k} = 0$, $k = 1...h-1$, and using Lemma 3.3 for $\delta_{XZ_1,j}^{(h)}$ and $\delta_{XZ_1,1}^{(k)}$ we deduce

$$\begin{aligned}
&\delta_{XY,j}^{(h)} \equiv \pi_{XY,j}^{(h)} + \sum_{i=1}^{\infty} \pi_{XZ_2,i}^{(h)} \beta_{Z_2Y,j}^{1-i} \\
&\delta_{XZ_1,1}^{(k)} = \delta_{XZ_1,k} = \pi_{XZ_1,k} + \sum_{i=1}^{\infty} \pi_{XZ_2,i} \beta_{Z_2Z_1,k}^{1-i} = 0.
\end{aligned}$$

The assumption $Z_2 \overset{1}{\nrightarrow} X | I_{XZ_1}$ implies $\pi_{XZ_2,i} = 0$, $i \geq 1$, hence $\pi_{XY,j} = \delta_{XY,j} = 0$ $\forall j \geq 1$, giving $Y \overset{1}{\nrightarrow} X | I_{XZ}$, cf. Theorem 2.2. Additionally, the zeros $\delta_{XZ_1,k} = 0$, $k = 1...h-1$, imply $\delta_{XZ_1,k} = \pi_{XZ_1,k} = 0$, $k = 1...h-1$, hence $\pi_{XZ,1} = [\pi_{XZ_1,1}, \pi_{XZ_2,1}] = 0$. From (3.3) we deduce $Y \overset{(2)}{\nrightarrow} X | I_{XZ}$ given $\pi_{XZ,1} \pi_{ZY,j} = 0$.

Similarly, using (2.3) non-causation $Y \overset{(2)}{\nrightarrow} X|I_{XZ}$ and $\pi_{XZ,1} = 0$ imply

$$\begin{aligned}\pi_{XZ,1}^{(2)} &= \pi_{XZ,2} + \pi_{XX,1}\pi_{XZ,1} + \pi_{XY,1}\pi_{YZ,1} + \pi_{XZ,1}\pi_{ZZ,1} \\ &= \pi_{XZ,2} + \pi_{XX,1} \times 0 + 0 \times \pi_{YZ,1} + 0 \times \pi_{ZZ,1} \\ &= \pi_{XZ,2}.\end{aligned}$$

If $h \geq 3$, then non-causation $Z_2 \overset{1}{\nrightarrow} X|I_{XZ_1}$ and the identity $\delta_{XZ_1,k} = \pi_{XZ_1,k} = 0$, $k = 1...h-1$ again imply $\pi_{XZ,2} = 0$, thus $\pi_{XZ,1}^{(2)}\pi_{ZY,j} = 0$ for all j, giving $Y \overset{(3)}{\nrightarrow} X|I_{XZ}$, cf. (3.3). Repeating this logic,

$$\begin{aligned}\pi_{XZ,1}^{(3)} &= \pi_{XZ,2}^{(2)} + \pi_{XX,1}^{(2)}\pi_{XZ,1} + \pi_{XY,1}^{(2)}\pi_{YZ,1} + \pi_{XZ,1}^{(2)}\pi_{ZZ,1} \\ &= \pi_{XZ,2}^{(2)} + \pi_{XX,1}^{(2)} \times 0 + 0 \times \pi_{YZ,1} + 0 \times \pi_{ZZ,1} \\ &= \pi_{XZ,2}^{(2)} \\ &= \pi_{XZ,3} + \pi_{XX,1}\pi_{XZ,2} + \pi_{XY,1}\pi_{YZ,2} + \pi_{XZ,1}\pi_{ZZ,2} \\ &= \pi_{XZ,3} + \pi_{XX,1} \times 0 + 0 \times \pi_{YZ,2} + 0 \times \pi_{ZZ,2} \\ &= \pi_{XZ,3},\end{aligned}$$

and so on. Recursively we deduce $(Y, Z_2) \overset{1}{\nrightarrow} X|I_{XZ_1}$ and $Y \overset{(h)}{\nrightarrow} X|I_{XZ_1}$ imply $\delta_{XZ_1,k} = \pi_{XZ_1,k} = 0$, $k = 1...h-1$, $\pi_{XZ_2,i} = 0, \forall i \geq 1$, and $\pi_{XZ,k} = \pi_{XZ,1}^{(k)}$, $k = 1...h-1$, hence $\pi_{XZ,1}^{(k)} = 0$, $k = 1...h-1$, giving $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$.

iii. and *iv*. For any $h \geq 1$, if $(Y, Z_2) \overset{1}{\nrightarrow} X|I_{XZ_1}$, $Y \overset{(h)}{\nrightarrow} X|I_{XZ_1}$ and $Y \overset{h+1}{\rightarrow} X|I_{XZ_1}$, then $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$ also holds by (*ii*). Moreover, by Theorem 2.1 causation $Y \overset{h+1}{\rightarrow} X|I_{XZ_1}$ implies a causal chain $Y \overset{1}{\rightarrow} Z_1 \overset{1}{\rightarrow} X$ must exist. Using the above logic, we recursively deduce $\pi_{XZ,1}^{(h)} = \pi_{XZ,h}$. Because $Z_2 \overset{1}{\nrightarrow} X|I_{XZ_1}$, $Y \overset{(h)}{\nrightarrow} X|I_{XZ_1}$, and $Y \overset{h+1}{\rightarrow} X|I_{XZ_1}$, by Theorems 2.2 and 3.2 it must be the case that $\pi_{XZ_2,i} = 0, \forall i \geq 1$, $_{XZ_1,h} \neq 0$, hence

$$\pi_{XZ,1}^{(h)} = \pi_{XZ,h} = [\pi_{XZ_1,h}, 0].$$

Therefore, by (3.3), $Y \overset{(h+1)}{\nrightarrow} X|I_{XZ}$ *if and only if*

$$\pi_{XZ,1}^{(h)}\pi_{ZY,j} = \pi_{XZ,h}\pi_{ZY,j} = \pi_{XZ_1,h}\pi_{Z_1Y,j} = 0$$

for all $j \geq 1$. Because $Y \overset{1}{\rightarrow} Z_1|I_{XZ_1}$ it must be the case that $\pi_{Z_1Y,j} \neq 0$ for at least one j, therefore $Y \overset{(h+1)}{\nrightarrow} X|I_{XZ}$ *if and only if* $\pi_{XZ_1,h} = 0$. Because $\pi_{XZ_1,h} \neq 0$ and $Y \overset{(h)}{\nrightarrow} X|I_{XZ}$, we conclude $Y \overset{h+1}{\rightarrow} X|I_{XZ}$. □

Proof of Lemma 4.1. We reject $Y \overset{1}{\nrightarrow} X$ if we reject Tests 0.1- 0.2 ($Y \overset{(\infty)}{\nrightarrow} X$) and Test 1.0 ($Y \overset{1}{\nrightarrow} X$) :

$$\begin{aligned}&P\left(rej.\ H_0^{(1)}|H_0^{(1)} \text{ is true}\right) \\ &\quad= P\left(rej.\ 0.1 \cap rej.\ 0.1 \cap rej.\ 1.0|H_0^{(1)} \text{ is true}\right) \\ &\quad\leq P\left(rej.\ 1.0|H_0^{(1)} \text{ is true}\right) = \alpha_{1.0}.\end{aligned}$$

We reject $Y \overset{(2)}{\nrightarrow} X$ if we reject Tests 0.1- 0.2 ($Y \overset{(\infty)}{\nrightarrow} X$), and either reject Test 1.0 ($Y \overset{1}{\nrightarrow} X$), or fail to reject Test 1.0 and reject Tests 1.1 and 1.2 ($Y \overset{1}{\nrightarrow} Z$ and $Z \overset{1}{\nrightarrow} X$) and Test 2.0 ($Y \overset{1}{\nrightarrow} X$, $\pi_{XZ,1} = 0$). We have

$$P\left(rej.\ H_0^{(2)} | H_0^{(2)} \text{ is true}\right)$$
$$= P\Big((rej0.1 \cap rej0.2) \cap (rej1.0 \cup [fail1.0 \cap rej1.1 \cap rej1.2 \cap rej2.0])\ |H_0^{(1)} \text{ is true}\Big).$$

There are four cases to consider: $Y \overset{1}{\nrightarrow} Z \overset{1}{\rightarrow} X$, $Y \overset{1}{\rightarrow} Z \overset{1}{\nrightarrow} X$, $Y \overset{1}{\nrightarrow} Z \overset{1}{\nrightarrow} X$ and $Y \overset{1}{\rightarrow} Z \overset{1}{\rightarrow} X$. Under $H_0^{(2)}$, if $Y \overset{1}{\nrightarrow} Z \overset{1}{\rightarrow} X$ then $Y \overset{1}{\nrightarrow} (X, Z)$ is true, $(Y, Z) \overset{1}{\nrightarrow} X$ is false, and Test 2.0 represents only a sufficient condition for non-causation $Y \overset{(2)}{\nrightarrow} X$. Hence

$$P\left(rej.\ H_0^{(2)} | H_0^{(2)} \text{ is true}\right)$$
$$= P\Big((rej0.1 \cap rej0.2) \cap (rej1.0 \cup [fail1.0 \cap rej1.1 \cap rej1.2 \cap rej2.0])\ |H_0^{(1)} \text{ is true}\Big)$$
$$\leq P\Big(rej0.1 \cap (rej1.0 \cup [fail1.0 \cap rej1.1 \cap rej1.2 \cap rej2.0])\ |H_0^{(1)} \text{ is true}\Big)$$
$$\leq \min\left\{\alpha_{0.1}, P\left(rej1.0 \cup [fail1.0 \cap rej1.1 \cap rej1.2 \cap rej2.0]\ |H_0^{(1)} \text{ is true}\right)\right\}$$
$$\leq \min\left\{\alpha_{0.1}, \alpha_{1.0} + P\left((rej1.1 \cap rej1.2 \cap rej2.0)\ |H_0^{(1)} \text{ is true}\right)\right\}$$
$$\leq \min\{\alpha_{0.1}, \alpha_{1.0} + \alpha_{1.1}\}$$

If $Y \overset{1}{\rightarrow} Z \overset{1}{\nrightarrow} X$, then $Y \overset{1}{\nrightarrow} (X, Z)$ is false, $(Y, Z) \overset{1}{\nrightarrow} X$ is true, and Test 2.0 represents a valid necessary and sufficient condition for non-causation $Y \overset{(2)}{\nrightarrow} X$, hence

$$P\left(rej.\ H_0^{(2)} | H_0^{(2)} \text{ is true}\right)$$
$$= P\Big((rej0.1 \cap rej0.2) \cap (rej1.0 \cup [fail1.0 \cap rej1.1 \cap rej1.2 \cap rej2.0])\ |H_0^{(1)} \text{ is true}\Big)$$
$$\leq P\Big(rej0.2 \cap (rej1.0 \cup [fail1.0 \cap rej1.1 \cap rej1.2 \cap rej2.0])\ |H_0^{(1)} \text{ is true}\Big)$$
$$\leq \min\left\{\alpha_{0.2}, \alpha_{1.0} + P\left(rej1.1 \cap rej1.2 \cap rej2.0 | H_0^{(1)} \text{ is true}\right)\right\}$$
$$\leq \min[\alpha_{0.2}, \alpha_{1.0} + \min\{\alpha_{1.2}, \alpha_{2.0}\}]$$

If both $Y \overset{1}{\nrightarrow} Z \overset{1}{\nrightarrow} X$, then both $Y \overset{1}{\nrightarrow} (X, Z)$ and $(Y, Z) \overset{1}{\nrightarrow} X$ are true, and Test 2.0 does not present a necessary condition, and

$$P\left(rej.\ H_0^{(2)} | H_0^{(2)} \text{ is true}\right)$$
$$= P\Big((rej0.1 \cap rej0.2) \cap (rej1.0 \cup [fail1.0 \cap rej1.1 \cap rej1.2 \cap rej2.0])\ |H_0^{(1)} \text{ is true}\Big)$$
$$\leq \min[\alpha_{0.1}, \alpha_{0.2}, \alpha_{1.0} + \min\{\alpha_{1.1}, \alpha_{1.2}\}]$$

Finally, if $Y \overset{1}{\to} Z \overset{1}{\to} X$ then both $Y \overset{1}{\nrightarrow} (X, Z)$ and $(Y, Z) \overset{1}{\nrightarrow} X$ are false, and

$$
\begin{aligned}
&P\left(rej.\ H_0^{(2)}|H_0^{(2)} \text{ is true}\right) \\
&\quad = P\left((rej0.1 \cap rej0.2) \cap (rej1.0 \cup [fail1.0 \cap rej1.1 \cap rej1.2 \cap rej2.0])\ |H_0^{(1)} \text{ is true}\right) \\
&\quad \leq \alpha_{1.0} + \alpha_{2.0}.
\end{aligned}
$$

Repeating the above for each $H_0^{(h)}$, $h \geq 2$, gives the case-specific size bounds. □

DEPT. OF ECONOMICS, FLORIDA INTERNATIONAL UNIVERSITY