Lag-augmented local projections and causality properties at different horizons.

Eric Renault
University of Warwick
Opening Conference
For the EcoDep project
September 9th, 2020
OUTLINE

1. Testing for Granger Non-Causality
   1.1. Noncausality at Different Horizons
   1.2. Estimation and Inference of Impulse Responses by Local Projections

2. Causality Measures
   2.1. Kullback Measure of Causality in Sims’ sense
   2.2. From Sims’ Causality to Granger’s Causality
   2.3. Causality and Markovianity
   2.4. Statistical Inference

3. Concluding remarks
1. Testing for Non-Causality

• 1.1. Noncausality at Different Horizons:


\[
X(t) = [X_1(t), \ldots, X_{m_1}(t)]' \\
Y(t) = [Y_1(t), \ldots, Y_{m_2}(t)]' \\
Z(t) = [Z_1(t), \ldots, Z_{m_3}(t)]' \\
I(t) = \text{increasing sequence of Hilbert subspaces of } L^2 \\
X(-\bar{\omega}, t] + Z(-\bar{\omega}, t] \subset I(t) \\
\text{Best linear forecast: } P\{X(t + h)|I(t)\} \\
\text{Y does not cause X at horizon h given I if } \forall t: \\
P\{X(t + h)|I(t)\} = P\{X(t + h)|I(t) + Y(-\bar{\omega}, t]\}
Causality in Linear Invertible Processes

• VAR of infinite order (possibly nonstationary):

\[ W(t) = [X(t)', Y(t)', Z(t)'], E[\varepsilon(t)\varepsilon(t)'] = \text{nonsingular} \]

\[ W(t) = c(t) + \sum_{j=1}^{\infty} \Pi_j W(t - j) + \varepsilon(t) \]

\[ I(t) = H + X(-\bar{\omega}, t] + Z(-\bar{\omega}, t] \]

\[ \Rightarrow P\{W(t + h)|H + W(-\bar{\omega}, t]\} \]

\[ = c^{(h)}(t) + \sum_{j=1}^{\infty} \pi_j^{(h)} W(t + 1 - j) \]

\[ \pi_j^{(1)} = \pi_j, \pi_j^{(h+1)} = \pi_j + h + \sum_{k=1}^{h} \pi_k \pi_j^{(h+1-k)} \]
Impulse Response Coefficients

- From Sims (1980), interpretation of Wold decomposition:

\[ W(t) = \mu(t) + \varepsilon(t) + \sum_{j=1}^{\infty} \psi_j \varepsilon(t - j) \]

\( \psi_j = j - \text{th impulse response coefficient} \)

→ solution of: \( \Pi(z)\Psi(z) = Id_m \)

\( \Psi(z) = Id_m + \sum_{j=1}^{\infty} \psi_j z^j, \Pi(z) = Id_m - \sum_{j=1}^{\infty} \pi_j z^j, \)

→ \( \psi_h = \sum_{k=1}^{h} \pi_k \psi_{h-k}, \forall h = 1, 2, \ldots \) (with \( \psi_0 = Id_m \))

→ to compare with: \( \pi^{(h+1)}_j = \pi_{j+h} + \sum_{k=1}^{h} \pi_k \pi_j^{(h+1-k)} \)
Impulse Response vs Causality

• Recursion formulas:

\[
\pi_1^{(h)} = \pi_h + \sum_{k=1}^{h-1} \pi_k \pi_1^{(h-k)} = \sum_{k=1}^{h} \pi_k \pi_1^{(h-k)} \quad (\pi_1^{(0)} = I_{d_m})
\]

\[\implies \pi_1^{(h)} = \psi_h \text{ (same recursion)}\]

\(R1.\) This identity = proved in:
Dufour and Renault (1998), Econometrica
→ revisited (with structural impulse) in:
"Local Projections and VARs
Estimate the Same Impulse Responses“

\(R2.\) Local Projection = the regression equation:

\[
W(t + h) = c^{(h)}(t) + \sum_{j=1}^{\infty} \pi_j^{(h)} W(t + 1 - j) + u^{(h)}(t)
\]
Impulse Response vs Causality (ctnd)

Y does not cause X at horizon h given I iff:

\[ \pi_{XYj}^{(h)} = 0, \forall j = 1,2, \ldots \]

(see impact of Y in \( P\{W(t + h)|H + W(-\bar{w}, t)\} \))

Impulse Response Coefficients:

\[ \psi_{XYh} = \pi_{XY1}^{(h)} \]

→ does not tell the complete story about causality at horizon h!

R1. \( W \sim VAR(p) \Rightarrow \pi_{j}^{(h)} = 0, \forall j > p \)

R2. Y does not cause X up to horizon h iff \( \forall j = 1,2, \ldots \)

\[ \pi_{XYj} = 0 \text{ and } \psi_{XZk}\pi_{ZYj} = 0 \forall k = 1,2, \ldots, h - 1 \]

(indirect causality through Z)
1.2. Estimation and Inference of Impulse Responses by Local Projections

- Jorda (2005), *American Economic Review*

$$\psi_{XYh} = \pi_{XY1}^{(h)} = \text{polynomial functions of VAR coefficients}$$

**Testing** $$\psi_{XZk} \pi_{ZYj} = 0, \forall k = 1,2, ..., h - 1$$

**R. Testing** $$g(\theta) = \theta_1 \theta_2 = 0$$

$$\text{Rank} \begin{bmatrix} \frac{\partial g}{\partial \theta'} \end{bmatrix} = 1 \text{ or } 0 \text{ under the null}$$

_Dufour, Renault, Zinde – Walsh (2019):_ "Wald Tests when Restrictions are Locally Singular"
Local Projections for VAR(p)

- "(p,h)-autoregression":

\[ W(t + h) = c^{(h)}(t) + \sum_{j=1}^{p} \pi_j^{(h)} W(t + h - j) + u^{(h)}(t) \]

\[ u^{(h)}(t) = \sum_{j=0}^{h-1} \psi_j \varepsilon(t + h - j) \]

\textbf{R. OLS = consistent in the stationary case but efficiency loss w.r.t. estimating VAR coefficients } \pi_1, \ldots, \pi_p

and computing nonlinear functions \( \pi_j^{(h)} \)

\( \rightarrow \text{Delta Theorem for Inference (singularity issue)} \)

Error Term = MA(h - 1) process
Local Projections and Tests for Non-Causality

R1. Due to Indirect Causality:
Y does not cause X at all horizons iff:
Y does not cause X at horizons h=1,2,..., \( pm_3 + 1 \)

R2. Another “Lag-Augmentation”
→ Regression on \((p+1)\) lags for hedging against unit roots:
Sims, Stock and Watson (1990), *Econometrica*, “Inference in linear time series models with some unit roots”
Dolado and Lutkepohl, Econometric Reviews, “Making Wald tests work for cointegrated VAR systems”
→ Extended to Local Projections by
Dufour, Pelletier and Renault (2006)
Benefits from Lag-Augmentation

• Montiel Olea and Plagborg-Moller, (2020), “Local Projection Inference is Simpler and More Robust Than You Think”

→ Intuition from univariate AR(1):

\[ Y(t + 1) = \pi Y(t) + \varepsilon(t + 1) \]

\[ \rightarrow Y(t + h) = \pi^{(h)} Y(t) + u^{(h)}(t), \pi^{(h)} = \pi^h, \]

\[ u^{(h)}(t) = \sum_{j=0}^{h-1} \pi^j \varepsilon(t + h - j) \]

→ By plugging \( Y(t) = \pi Y(t - 1) + \varepsilon(t) \):

\[ Y(t + h) = \pi^{(h)} \varepsilon(t) + \pi^{(h+1)} Y(t - 1) + u^{(h)}(t) \]
Benefits from Lag-Augmentation (ctnd)

- Intuition from AR(1):
  \[ Y(t + h) = \pi^{(h)}\varepsilon(t) + \pi^{(h+1)}Y(t - 1) + u^{(h)}(t) \]
  \[ \rightarrow \text{By Frisch – Waugh, infeasible OLS} \]
  \[ Y(t + h) \text{ on } (\varepsilon(t), Y(t - 1)) \]
  coincides for estimator of \( \pi^{(h)} \) with
  feasible OLS of \( Y(t + h) \text{ on } (Y(t), Y(t - 1)) \)

1st benefit: Hedge against unit root

2nd benefit: Hedge against
  serial correlation in \( u^{(h)}(t), t = 1,2, \ldots \)
  \[ \rightarrow \text{due to non–correlation} \]
  in regression scores \( (\text{see } u^{(h)}(t)\varepsilon(t)) \)

under maintained assumption (restrictive??):
\[ E\{\varepsilon(t)|\varepsilon(s), s \neq t\} = 0 \]
2. Causality Measures

• Geweke, J. (1982), JASA, “Measurement of linear dependence and feedback between multiple time series”.

→ while the non-causality hypothesis may not be literally entertained, we need to be able to measure the actual degree of causality


Renault, Scida, (2016), Advances in Econometrics, “Causality and Markovianity: Information Theoretic Measures”
2.1. Kullback Measure of Causality in Sims’ sense

• Sims: “Money, Income and Causality”, 1972:
Is there statistical evidence that money (X) is exogenous in some sense in the money-income (Y) relationship? In Sims’ words:
“If and only if causality runs one way from current and past values of some list of exogenous variables (X) to a given endogenous variable (Y), then in a regression of the endogenous variable on past, current, and future values of the exogenous variables, the future values of the exogenous variable should have zero coefficients”
Sims Non-Causality

- Chamberlain (1982) = two additions to Sims (1972):
  (i) In terms of probability distributions (not only regressions) → p.d.f. with respect to product measure
  (ii) Conditioning information = also includes past values of Y

Y does not cause X in the Sims sense (Y.NCS.X) if:

\[ f_t(y_t | x(-\omega, T), y(-\omega, t - 1)) = f_t(y_t | x(-\omega, t), y(-\omega, t - 1)) \]

\[ \forall t = 1, \ldots, T \ (given \ \omega > 0) \]
Kullback Measure

- Gourieroux-Monfort-Renault (1987)

Data Generating Process (DGP): Characterized by joint p.d.f.:

\[
f_0[x(-\omega, T), y(-\omega, T)]
\]

\[
HS[Y \leftrightarrow X] = \{f[\ldots]; Y.\ NCS.\ X\}
\]

\[
CS_{Y \rightarrow X} = \min_{f} \frac{1}{T} \min \left\{ E_0 \log \left( \frac{f_0[x(-\omega, T), y(-\omega, T)]}{f[x(-\omega, T), y(-\omega, T)]} \right); f \in HS[Y \leftrightarrow X] \right\}
\]
Kullback Measure of Sims’ Causality

• Value of the minimization program:

\[ CS_{Y \rightarrow X} = \frac{1}{T} \sum_{t=1}^{T} E_0 \log \left( \frac{f_{0t}\{y_t|x(-\omega, T), y(-\omega, t - 1)\}}{f_{0t}\{y_t|x(-\omega, t), y(-\omega, t - 1)\}} \right) \]

\[ CS_{Y \rightarrow X} \geq 0 \]

\[ CS_{Y \rightarrow X} = 0 \iff Y. NCS. X \]
2.2. From Sims’ Causality to Granger’s Causality

- Decomposition of the joint (conditional) density:

\[
\frac{f_0[x(-\omega, T), y(-\omega, T)]}{f_0[x(-\omega, 0), y(-\omega, 0)]} =
\]

\[
f_0\{x(1, T)|x(-\omega, 0), y(-\omega, 0)\} \prod_{t=1}^{T} f_{0_t}\{y_t|x(-\omega, T), y(-\omega, t - 1)\}
\]

\[
= (Initiation Y \rightarrow X) \times (Sims Causality Y \rightarrow X)
\]
Kullback Measure of Initiation

- Y does not initiate X (Y.NI.X) if

\[
\mathbf{f}_0\{x(1, T) | x(-\omega, 0), y(-\omega, 0)\} \\
= \mathbf{f}_0\{x(1, T) | x(-\omega, 0)\} \\
HIN[Y \nrightarrow X] = \{\mathbf{f}[., ., .] ; Y. NI. X\}
\]

\[
CIN_{Y \rightarrow X} = \min_\mathbf{f} \left\{ \frac{1}{T} \min \left\{ E_0 \log \left( \frac{\mathbf{f}_0[x(-\omega, T), y(-\omega, T)]}{\mathbf{f}[x(-\omega, T), y(-\omega, T)]} \right) ; \mathbf{f} \in HIN[Y \nrightarrow X] \right\}
\]
Kullback Measure of Initiation (ctnd)

• Value of the minimization program:

$$CIN_{Y \to X} = \frac{1}{T} E_0 \log \left( \frac{f_0\{x(1,T) \mid x(-\omega,0), y(-\omega,0)\}}{f_0\{x(1,T) \mid x(-\omega,0)\}} \right)$$

$$CIN_{Y \to X} = \frac{1}{T} \sum_{t=1}^{T} E_0 \log \left( \frac{f_{0t}\{x_t \mid x(-\omega,t-1), y(-\omega,0)\}}{f_{0t}\{x_t \mid x(-\omega,t-1)\}} \right)$$

$$CIN_{Y \to X} \geq 0$$

$$CIN_{Y \to X} = 0 \iff Y \text{ does not initiate } X$$

*May not be negligible even when } T \to \infty*
Granger Causality

• Pythagoras’ theorem on causality measures:

\[ CS_{Y \rightarrow X} + CIN_{Y \rightarrow X} = CG_{Y \rightarrow X} = \]

\[
\frac{1}{T} \sum_{t=1}^{T} E_0 \log \left( \frac{f_{0t} \{ x_t | x(-\omega, t - 1), y(-\omega, t - 1) \}}{f_{0t} \{ x_t | x(-\omega, t - 1) \}} \right) \]

\[ CG_{Y \rightarrow X} \geq 0 \]
\[ CG_{Y \rightarrow X} = 0 \iff Y \text{ does not cause } X \text{ in the Granger sense}(Y.NCG.X) \]
\[ \iff Y.NCS.X \text{ and } Y.NI.X \]
From Sims to Granger

- Granger measure:

\[ HG[Y \leftrightarrow X] = \{ f[.,.]; Y. NCG. X \} \]

\[ CG_{Y \rightarrow X} = \]

\[ \frac{1}{T} \min \left\{ E_0 \log \left( \frac{f_0[x(-\omega,T), y(-\omega,T)]}{f[x(-\omega,T), y(-\omega,T)]} \right); f \in HG[Y \leftrightarrow X] \right\} \]

- Florens-Mouchart (1982): “Granger’s noncausality still implies, but is not equivalent to, Sims’ noncausality. Some care has to be taken (...) in order to handle the initial condition properly”.

- Chamberlain (1982): Maintains the so-called regularity condition (R), precisely to make asymptotically negligible the role of initial condition.
2.3. Causality and Markovianity

• We revisit the previous causality measures under the maintained assumption: the process \((X,Y)\) is Markov of order \(p\).

**Example 1:** Gaussian VAR\((p)\) (Geweke (1982), *JASA*).

**Example 2:** Qualitative Panel data (Bouissou, Laffont, Vuong (1986), *Econometrica*).

Markov simplification

• The three causality measures can be rewritten as follows

\[ CS_{Y \rightarrow X} = \frac{1}{T} \sum_{t=1}^{T} E_0 \log \left( \frac{f_{0t}\{y_t|x(t-p,T),y(t-p,t-1)\}}{f_{0t}\{y_t|x(t-p,t),y(t-p,t-1)\}} \right) \]

\[ CIN_{Y \rightarrow X} = \frac{1}{T} \sum_{t=1}^{T} E_0 \log \left( \frac{f_{0t}\{x_t|x(1-p,t-1),y(1-p,0)\}}{f_{0t}\{x_t|x(\omega,t-1)\}} \right) \]

\[ CG_{Y \rightarrow X} = \frac{1}{T} \sum_{t=1}^{T} E_0 \log \left( \frac{f_{0t}\{x_t|x(t-p,t-1),y(t-p,t-1)\}}{f_{0t}\{x_t|x(\omega,t-1)\}} \right) \]
Causality and Lag Augmentation

• The far past \((-\infty)\) still shows up in conditional density of \(X\) given its own past because:

Joint Markovianity of \((X,Y)\) does not imply Markovianity of \(X\) → an issue for measures of Granger causality and Initiation (not for Sims)

→ Issue ignored by Gourieroux, Monfort, Renault (1987) and Schreiber (2000) who compute Granger causality measure replacing \((-\infty)\) by \((t-p)\)

→ More generally, lag augmentation:

\[ t-q, \quad q \geq p \]
Joint hypothesis Granger/Markov

• We decide to only incorporate the knowledge that \((X,Y)\) is Markov of order \(q\), \(q \geq p\), and, with this maintained assumption, we look for the closest p.d.f. such that \(Y\) does not Granger cause \(X\).


\(Y. \text{NCG. } X \Rightarrow \) Marginal Markovianity of order \(q\) for \(X\)

\[ H_{Gq}[Y \leftrightarrow X] \]

\[ = \{ f[. . . ]; Y. \text{NCG. } X \text{ and } (Y, X) \sim \text{Mark}(q) \} \]
Measure (Granger+Markov)

• With joint hypothesis:

\[ CG_{Y \rightarrow X}(q) = \]

\[
\frac{1}{T} \min \left\{ E_0 \log \left( \frac{f_0 \{ x(-\omega, T), y(-\omega, T) \}}{f \{ x(-\omega, T), y(-\omega, T) \}} \right); f \in HG_q[Y \Rightarrow X] \right\}
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} E_0 \log \left( \frac{f_{0t} \{ x_t | x(t - p, t - 1), y(t - p, t - 1) \}}{f_{0t} \{ x_t | x(t - q, t - 1) \}} \right)
\]

\[ CG_{Y \rightarrow X}(q) - CG_{Y \rightarrow X} = \]

\[
\frac{1}{T} \sum_{t=1}^{T} E_0 \log \left( \frac{f_{0t} \{ x_t | x(-\omega, t - 1) \}}{f_{0t} \{ x_t | x(t - q, t - 1) \}} \right) = M_X(q) \geq 0
\]
Measure of the degree of Non-Markovianity

- Minimizing the Kullback “distance” to the set of p.d.f. such that $X \sim$ Markov (q):

\[ M_X(q) \geq 0 \]

\[ M_X(q) = 0 \iff X \sim \text{Markov}(q) \]

\[ CG_{Y \rightarrow X}(q) - CG_{Y \rightarrow X} = CIN_{Y \rightarrow X}(q) - CIN_{Y \rightarrow X} = M_X(q) \]

Pythagoras' theorem:

\[ CS_{Y \rightarrow X} + CIN_{Y \rightarrow X}(q) = CG_{Y \rightarrow X}(q) \]

\[ CG_{Y \rightarrow X}(p) \geq CG_{Y \rightarrow X}(q) \geq CG_{Y \rightarrow X}(q + 1) \geq CG_{Y \rightarrow X}, \forall q \geq p \]

R.Y.NCG.X $\iff$ All these measures are zero

Conclusion: If not enough "lag augmentation" in $X$ equation, $CG_{Y \rightarrow X}(q)$ overestimates the level of Granger Causality
Numerical Evidence

- Framework: \((X,Y) \sim \text{Gaussian VAR}(1)\)

\[
\begin{bmatrix} X_t \\ Y_t \end{bmatrix} = M \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} M^{-1} \begin{bmatrix} X_{t-1} \\ Y_{t-1} \end{bmatrix} + \begin{bmatrix} u_t \\ v_t \end{bmatrix}
\]

\[
M \sim \begin{bmatrix} 0.93 & -0.25 \\ 0.36 & 0.96 \end{bmatrix}
\]

\(\lambda_1 \sim \text{Persistence of } X\)

\(\lambda_2 \sim \text{Persistence of } Y\)

Other parameters:

\[
\text{Var}(u_t) = (\sigma_u)^2, \text{Var}(v_t) = (\sigma_v)^2
\]

\[
\text{Corr}[u_t, v_t] = \rho
\]
Numerical Evidence (ctnd)

• Goal: Study $CG_{Y \rightarrow X}$ as computed by a researcher who
  – “Wrongly” sticks to GMR (1987) and Schreiber (2000) using $q$ lags for the marginal $X_t$ instead of the correct measure which takes ϖ
  – Does not misspecify Markov ($q \geq p$)

• Then, the marginal autoregression for $X_t$ is:

$$X_t = \sum_{j=1}^{q} a_j X_{t-j} + \varepsilon_t, E(\varepsilon_t) = 0, Var(\varepsilon_t) = [\sigma_\varepsilon(q)]^2$$

• This leads to the following Kullback measure of Granger Causality

$$CG_{Y \rightarrow X}(q) = \frac{1}{2} \log \left( \frac{[\sigma_\varepsilon(q)]^2}{[\sigma_u]^2} \right)$$
Comparison of $CG_{Y \rightarrow X}(q)$ across $q$ (for fixed values of the parameters)

**Figure 1:** Comparison of $CG_{Y \rightarrow X}(q)$ across $q$ for fixed values of the parameters: $\lambda_1 = -0.6$, $\lambda_2 = 0.9$, $\rho = 0.5$, $\sigma_u = 1$, $\sigma_v = 1$. 
Comparison of $CG_{Y \rightarrow X}(q)$:
Wrong conclusions...

**Figure 2:** Comparison of $CG_{Y \rightarrow X}(q)$ for $q = 1, 2, \ldots, 10$:

- $\lambda_1 = -0.8$, $\lambda_2 = 0.8$
- $\rho = -0.99$, $\sigma_u = 1$, $\sigma_v = 1$

<table>
<thead>
<tr>
<th>$q$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CG_{Y \rightarrow X}(q)$</td>
<td>0.2011</td>
<td>0.0146</td>
<td>0.0026</td>
<td>0.0016</td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0015</td>
<td>0.0015</td>
</tr>
</tbody>
</table>
2.4. Statistical Inference

• Geweke (1982) for Gaussian VAR(p) (see also Gourieroux-Monfort-Renault (1987) for general Markov (p)) note that “the estimated measures of feedback are the usual likelihood ratio test statistics for the hypothesis of no feedback deflated by sample size”. In both cases:

(i) Maintained hypothesis = Stationary Markov (p)

(ii) Granger Causality measure:

\[ CG_{Y \rightarrow X}(p) = E_0 \log \left( \frac{f_0\{x_t|x(t-p, t-1), y(t-p, t-1)\}}{f_0\{x_t|x(t-p, t-1)\}} \right) \]
Likelihood Ratio Test of Non-Causality

• Obviously:

\[ LR_T(Y, NCG, X) = 2T \hat{C}_G_{Y \rightarrow X} \]

Under the null hypothesis:

\[ LR_T(Y, NCG, X) = 2T \hat{C}_G_{Y \rightarrow X}(p; \hat{\theta}) + o_P(1) \xrightarrow{d} \chi^2 \]

In the Gaussian VAR(p) case:

\[ LR_T(Y, NCG, X) = T \log \left( \frac{\det(\hat{\Sigma}_\varepsilon(p))}{\det(\hat{\Sigma}_u)} \right) \]

\[ X_t = \sum_{i=1}^{p} a_i X_{t-i} + \sum_{i=1}^{p} b_i Y_{t-i} + u_t, \quad X_t = \sum_{i=1}^{p} \alpha_i X_{t-i} + \varepsilon_t \]

*(Geweke, 1982)*
Asymmetric Markov

• **Example**: Asymmetric VAR

Hsiao (1981): *JME*

Keating (2000): *Journal of Macroeconomics*

→ We want to consider:

\[ CG_{Y \rightarrow X}(q), q > p \]

**Model 1: VAR(\(p\))**

\[
X_t = \sum_{i=1}^{p} a_i X_{t-i} + \sum_{i=1}^{p} b_i Y_{t-i} + u_t, \quad Y_t = \sum_{i=1}^{p} c_i X_{t-i} + \sum_{i=1}^{p} d_i Y_{t-i} + \nu_t
\]

**Model 2: \( q \) lags for \( X \), \( q > p \):**

\[
X_t = \sum_{i=1}^{q} \alpha_i X_{t-i} + \sum_{i=1}^{p} \beta_i Y_{t-i} + \varepsilon_t, \quad Y_t = \sum_{i=1}^{q} \gamma_i X_{t-i} + \sum_{i=1}^{p} \delta_i Y_{t-i} + \eta_t
\]
Asymmetric VAR

• For each variable, same number of lags in each equation $\rightarrow$ OLS = Gaussian MLE.

• Hsiao (1981): “although (...) the test of causal relationship is a direct test equivalent to the LR test when no a priori information is used, the results are not independent of the order of AR operators chosen for the model”

• Example: Keating (2000): “each equation in a bivariate VAR could have 3 lags of output and 6 lags of money” (larger number $q$ of lags for candidate exogenous variable $X$).
Non Nested Hypotheses (Cox LR test)

- Asymmetric VAR + Non-Causality Hypothesis → Non Nested Hypotheses.
- Asymmetric VAR = modified to have the same regressors in the two equations for Y.

**Hypothesis VAR(\(p\))**:

\[
X_t = \sum_{i=1}^{p} a_i X_{t-i} + \sum_{i=1}^{p} b_i Y_{t-i} + u_t, Y_t = \sum_{i=1}^{p} c_i X_{t-i} + \sum_{i=1}^{p} d_i Y_{t-i} + v_t
\]

**Hypothesis NC(\(q\))**:

\[
X_t = \sum_{i=1}^{q} a_i X_{t-i} + 0 + \varepsilon_t, Y_t = \sum_{i=1}^{p} \gamma_i X_{t-i} + \sum_{i=1}^{p} \delta_i Y_{t-i} + \eta_t
\]
Testing for Granger Non-Causality

• If our focus of interest is significance testing of the Granger Non-Causality property, the null hypothesis should be NC(q) for some given q much larger than p (q.(dimX) > p.(dimX+dimY))

⇒ Definition of an alternative causality measure computed by minimization of a Kullback distance under the maintained hypothesis NC(q):

\[ \tilde{CG}_{Y \rightarrow X}(q; \theta^0) \]
\[ = E_0 \log \left( \frac{f\{x_t|x(t - q, t - 1); \theta^0\}}{f\{x_t|x(t - p, t - 1), y(t - p, t - 1); \tilde{\theta}(\theta^0)\}} \right) \]
3. Concluding remarks


“Model selection is more appropriate when the objective is decision making. Hypothesis testing is better suited to inferential problems where the empirical validity of a theoretical prediction is the primary objective”.

1. Initial causality measure (under the null of VAR(p)) = to compare to the average LR statistic when objective is forecasting (decision making).

2. New causality measure (under the null of NC) = better suited when objective is validity of theoretical prediction of non-causality (money exogenous w.r.t. output).

3. The difference is due to non-nested hypotheses due to the need of lag augmentation in the Granger Causality measure.
Advantage of Sims’ causality concept

• Pesaran and Weeks (2001): The dual “aspect of non-nested hypotheses has been criticized by some commentators, pointing out the test outcome can lead to ambiguities”

• See e.g. Granger, King and White (1995).

• Sims concept = immune to this criticism because does not imply considering non-nested hypotheses (no joint assumption with Markov)

• Sims’ concept = plays a similar role to a third comprehensive model, constructed so that each of the non-nested models can be obtained from it as a special case (Atkinson, 1970)
Causality Measures at Different Horizons


→ Simulation-based inference from complex functions of model parameters in VAR and VARMA models

→ Need to extend local projection analysis to causality measures (and not only causality tests)

→ Need to revisit causality measures with indirect causality

→ Several motivations for lag augmentation