Granger-Causal-Priority and Choice of Variables in Vector Autoregressions

Marek Jarociński, *European Central Bank*

Bartosz Maćkowiak, *European Central Bank and CEPR*

ESSIM, Izmir, 22 May 2013

The questions that this paper studies

- A researcher wants to forecast y_i or compute impulse responses of y_i with a VAR.

- The researcher has access to data on y, where $y_i \subset y$.

- Let $y_J \equiv y \setminus y_i$ and consider $y_j \subseteq y_J$.
  - Does y_j belong in the VAR to be used to forecast y_i?
  - Does y_j belong in the VAR to be used to compute impulse responses of y_i?

Why study choice of variables in VARs?

- Most applications of VARs involve choice of variables.
  - Most economists have a preference for using the minimal means to get their points across.
  - Most audiences want to understand in simplest terms "where results come from."
- Typically, choice of variables occurs *informally*.
  - We show that choice of variables can occur *formally* in a straightforward way.
  - Even when choice of variables is informal, it is useful to know what assumptions are implicit and to what extent these assumptions are supported by the data.

How we answer the questions

- We propose a methodology relying on two ingredients:

  1. a restriction on the data generating process followed by *all variables in the dataset*

  2. tools for Bayesian inference.

- We apply the methodology to the case when the variables of interest are output, price level, and short-term interest rate.

  – We obtain remarkably similar findings for the euro area and the United States.

Granger-noncausality and Granger-causal-priority

- Consider $y(t) = B(L)\, y(t-1) + u(t)$
  - y_j *does not Granger-cause* y_i if the coefficients on all lags of y_j in the equations for y_i are equal to zero, $B_{ij}(L) = 0$
  - y_i *is Granger-causally-prior to* y_j if it is possible to partition all the variables in y into two subsets, y_1 and y_2, such that $y_i \subseteq y_1$, $y_j \subseteq y_2$, and y_2 does not Granger-cause y_1. That is, we have $B_{12}(L) = 0$ in

$$\begin{array}{l} y_i \rightarrow \\ y_j \rightarrow \end{array} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} B_{11}(L) & B_{12}(L) \\ B_{21}(L) & B_{22}(L) \end{pmatrix} \begin{pmatrix} y_1(t-1) \\ y_2(t-1) \end{pmatrix} + \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}$$

Granger-causal-priority is stronger than Granger-noncausality

- Consider $y = \{x, w, z\}$, $y_i = x$, $y_j = z$ and suppose we have $B_{xz} = 0$ in

$$\begin{pmatrix} x(t) \\ w(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} B_{xx} & B_{xw} & B_{xz} \\ B_{wx} & B_{ww} & B_{wz} \\ B_{zx} & B_{zw} & B_{zz} \end{pmatrix} \begin{pmatrix} x(t-1) \\ w(t-1) \\ z(t-1) \end{pmatrix} + u(t)$$

  - $x(t+1) = B_{xx}x(t) + B_{xw}w(t) + u_1(t+1)$, so z does not Granger-cause x.
  - However, $x(t+2) = \ldots + B_{xw}B_{wz}z(t) + \ldots$.
  - x is Granger-causally-prior to z if *either* $B_{xz} = B_{xw} = 0$ *or* $B_{xz} = B_{wz} = 0$. All indirect effects from z to x are ruled out. The VAR is block recursive.

Granger-causal-priority and choice of variables for forecasting

- If y_i is Granger-causally-prior to y_j, y_j does not belong in the VAR to be used to forecast y_i.
  - The reason is that the following two VARs give the same forecasts:

$$\begin{array}{l} y_i \rightarrow \\ y_j \rightarrow \end{array} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} B_{11}(L) & 0 \\ B_{21}(L) & B_{22}(L) \end{pmatrix} \begin{pmatrix} y_1(t-1) \\ y_2(t-1) \end{pmatrix} + \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}$$

and

$$y_1(t) = B_{11}(L) y_1(t-1) + u_1(t)$$

Granger-causal-priority and choice of variables for impulse response analysis

- If y_i is Granger-causally-prior to y_j *and* the number of structural shocks that affect y_1 contemporaneously is equal to the number of variables in y_1, y_j does not belong in the VAR to be used to infer impulse responses of y_i.
  - The reason is that the following two structural VARs yield the same impulse responses of y_1:

$$\begin{pmatrix} A_{11}(L) & 0 \\ A_{21}(L) & A_{22}(L) \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{pmatrix}$$

    and

$$A_{11}(L) y_1(t) = \varepsilon_1(t)$$

How to evaluate the posterior probability that y_i is GCP to y_j?

- We consider a set Ω of models such that: (i) all models in Ω are Gaussian VARs with all variables y and at most one Granger-causal-priority restriction, and (ii) $y_i \subseteq y_1$ in each restricted model in Ω.

- Consider $y_j \subseteq y_J$. Let Ω^j be the subset of Ω that contains all the models in which $y_j \subseteq y_2$.
  - We evaluate the posterior probability of Ω^j, $p\left(\Omega^j | Y, \Omega\right)$.
  - Note: If y_j consists of a single variable, Ω^j and $\Omega \backslash \Omega^j$ have an equal number of elements.

Implementation

- We can evaluate *analytically* the marginal likelihood of any model in Ω.
  - The reason is that we derive an *analytical* Bayes factor in favor of a Granger-noncausality restriction, $p(Y|\omega^R)/p(Y|\omega^U)$.
  - We assume: $p(B, \Sigma|\omega^U)$ is conjugate, and

$$p(B_{(\beta\alpha)}, \Sigma|\omega^R) = p(B_{(\beta\alpha)}, \Sigma|\omega^U, B_{\beta\alpha} = \mathbf{0}).$$

- In our application Ω is large. We approximate the posterior probabilities of GCP using the Markov Chain Monte Carlo Model Composition algorithm.

Evaluating the probability of Granger-causal-priority in practice

- In the euro area and in the United States.

- GDP, price level, and short-term interest rate enter y_i.

- Thirty-eight macroeconomic and financial variables enter y_J: GDP components, government debt, labor market variables, interest rates, monetary aggregates, credit aggregates, exchange rates, commodity prices and other price indexes, housing market variables, stock market variables, survey-based leading indicators.

Prior, sample

- Prior with two components:
  - Sims-Zha prior: Minnesota prior + one-unit-root dummy observation + no-cointegration dummy observations (with hyperparameter values that maximize the marginal likelihood in a training sample).
  - Training sample prior: 1989Q1-1998Q4.
- Sample: 1999Q1-2012Q4.
- One lag.

Variables associated with low Granger-causal-priority probability

Euro area			U.S.	
Variable	*prob.*	*rank*	*Variable*	*prob.*
Change of Real Inventories	0.00	*1*	Oil Price	0.00
Industrial Confidence	0.00	*2*	Industrial Confidence	0.00
2-Year Bond Yield	0.00	*3*	Change of Real Inventories	0.00
PMI	0.02	*4*	Unemployment Rate	0.00
Lending Rate to NFCs	0.02	*5*	Bond Spread BBB 7-10 Years	0.00
Real Investment	0.05	*6*	2-Year Bond Yield	0.01
Real Exports	0.05	*7*	Capacity Utilization	0.03
Real Imports	0.07	*8*	Lending Rate to NFCs	0.04
Mortgage Interest Rate	0.08	*9*	Hours Worked**	0.08
Oil Price	0.08	*10*	Real Investment	0.11

Note: survey-based leading indicators, change in inventories, interest rates on public and private debt, oil price.

Variables associated with high Granger-causal-priority probability

Euro area			U.S.	
Variable	*prob.*	*rank*	*Variable*	*prob.*
M3	0.85	*29*	Foreign Consumer Prices*	0.74
M1	0.86	*30*	Loans to NFCs	0.76
Stock Market Volatility	0.90	*31*	Real Exports	0.86
Loans for House Purchase	0.91	*32*	Government Debt	0.88
Stock Market Index	0.92	*33*	M2	0.90
Nominal Effective Exchange Rate	0.95	*34*	Loans for House Purchase	0.95
Consumer Prices Excl. Energy, Food	0.98	*35*	Real Housing Investment	0.97
Government Debt	0.99	*36*	Consumer Loans	0.99
House Prices	1.00	*37*	House Prices	1.00
Dollar/Euro Exchange Rate	1.00	*38*	Dollar/Euro Exchange Rate	1.00

Note: house prices, nominal exchange rate, also government debt, broad money and credit aggregates.

Choosing the best VAR

- Evaluating the posterior probabilities yields a unique choice of variables if the probabilities are either zero or one.
  - Asymptotically, the posterior probabilities *are* either zero or one.

- We assume a zero-one loss function on models, and we search for the VAR with highest marginal likelihood, ω^*.

Euro area			United States	
Variable	prob.	Rank	Variable	prob
Change of Real Inventories	**0.00**	1	**Oil Price**	**0.00**
Industrial Confidence	**0.00**	2	**Industrial Confidence**	**0.00**
2-Year Bond Yield	**0.00**	3	**Change of Real Inventories**	**0.00**
PMI	**0.02**	4	**Unemployment Rate**	**0.00**
Lending Rate to NFCs	**0.02**	5	**Bond Spread BBB 7-10 Years**	**0.00**
Real Investment	**0.05**	6	**2-Year Bond Yield**	**0.01**
Real Exports	**0.05**	7	**Capacity Utilization**	**0.03**
Real Imports	**0.07**	8	**Lending Rate to NFCs**	**0.04**
Mortgage Interest Rate	**0.08**	9	**Hours Worked**	**0.08**
Oil Price	**0.08**	10	**Real Investment**	**0.11**
Real Consumption	**0.12**	11	**10-Year Bond Yield**	**0.13**
Real Housing Investment	**0.18**	12	**Foreign Short Term Interest Rate***	**0.13**
Bond Spread BBB 7-10 Years	0.21	13	**PMI**	**0.14**
Unit Labor Cost	**0.23**	14	**Mortgage Interest Rate**	**0.16**
Foreign Short Term Interest Rate*	0.26	15	**Consumer Confidence**	**0.17**
Unemployment Rate	0.29	16	**Foreign Real GDP***	**0.22**
Total Employment	0.40	17	**Total Employment**	**0.28**
Consumer Confidence	0.48	18	**PPI**	**0.30**
Foreign Consumer Prices*	**0.53**	19	**Real Imports**	**0.30**
10-Year Bond Yield	0.54	20	**Real Consumption**	**0.33**

Log marginal likelihoods for comparing different models in Ω

	Euro area	U.S.
The best model in Ω: ω^*	4825	3516
The unrestricted model ω^U	4774	3488
The model "opposite" to ω^*	4740	3380

Lessons:

- The VAR with the best GCP restriction is much better than the unrestricted VAR
- The unrestricted VAR is much better than a VAR with a poorly chosen GCP restriction

Robustness

- Subsamples.
- Prior hyperparameter values.
- Lag length.

Conclusions

- Choice of variables in VARs, for forecasting and for impulse response analysis, connects with the concept of Granger-causal-priority.

- The posterior probability of Granger-causal-priority can be evaluated analytically.

- The following variables belong in the VAR with output, price level, and short-term interest rate:

  – survey-based leading indicators, change in inventories, interest rates on public and private debt, oil price.

  – The same finding obtains in the euro area and in the United States.

Marginal likelihood vs predictive density score

$$p(Y|\omega) = p(y_i(1,...,T), y_J(1,...,T)|y_i(0), y_J(0), \omega) =$$

$$= \prod_{t=0}^{T-1} p(y_i(t+1), y_J(t+1)|y_i(0,...,t), y_J(0,...,t), \omega)$$

$$g(y_i, h=1|\psi) = \prod_{t=0}^{T-1} p(y_i(t+1)|y_i(0,...,t), y_J(0,...,t), \omega)$$

- cannot be related to posterior odds on models, *wastes information*

- $g(y_i, h=1|\psi)$ depends on h

- costly to compute (requires loop over t)