

Lecture 4

Previsioni con sistemi multivariati

1

The aim

- Purpose of economics is to help decision makers make better decisions. All actions undertaken with the intention of affecting future outcomes involve forecasting. In fact, forecasts are either explicitly or implicitly at the heart of policy, while the methods for making a forecast often engender much less agreement.
- It is a difficult task, as it needs to estimate dynamic **causal effects** from macroeconomic data generated by possibly **instable/heterogeneous** relationships.
- Increasing the number of variables should increase the amount of information and forecast performance of the model. However, model instability often prevents structural macro models to improve with size.

2

Macroeconometric models

- Despite problems and criticisms, the macro-econometric model is the most largely used quantitative tool for policy analysis.
- Using the model, policy can incorporate the model's views about the "way the economy works" if the model is **theoretically coherent**.
- However, the model must also be able to replicate historical data, i.e. to be **empirically coherent**.
- It is hard to simultaneously achieve both these aims because a number of trade-offs leading to the graphical evidence below.
- Last, model-based forecasts end with a story; a model-based forecast has the ability to essentially explain itself.

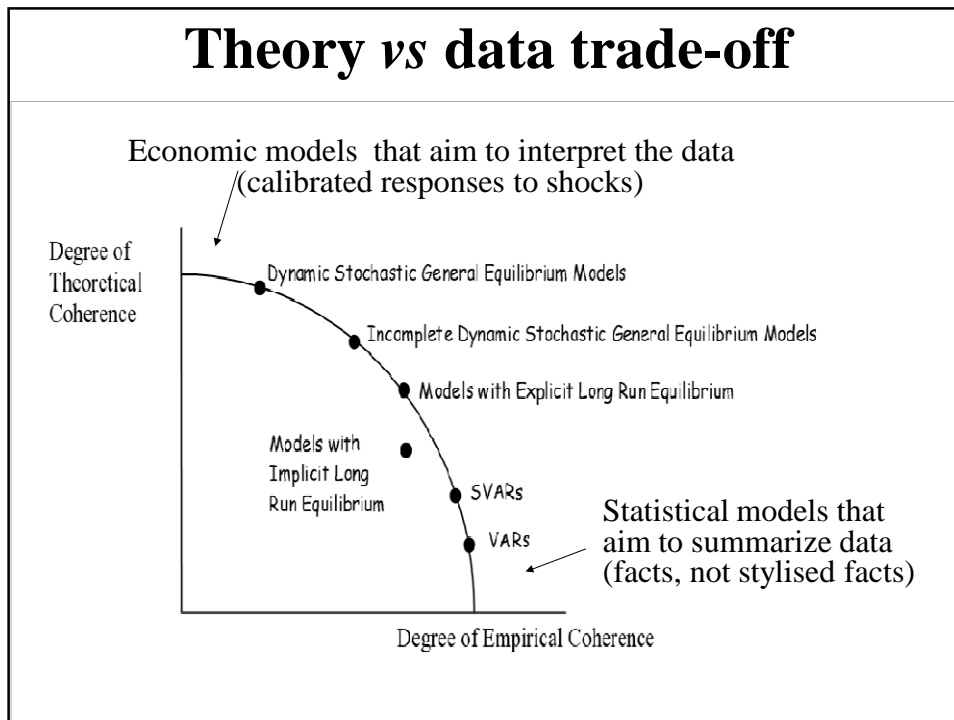
3

The empirical-theoretical coherence curve

- At one end of the curve there are models that fit everything to data but whose outcomes are impossible to interpret, while, at the other, there are theoretical models that have never been exposed to a data set.
- Being at either of these points is not attractive for the policymaker. Better models are located along the interior points on the curve.
- Over time the curve has shifted outward (same empirical coherence with stronger theory, e.g. thanks to more computing power).
- At any point in time there is a frontier of best practice models: combinations of the 2 coherences above.
- Where a model is located on the curve is a function of the constraints from the institutional structure in which models operate.

4


Theory vs data trade-off



(Almost) statistical models: VAR

- VAR: n series in y_t vector: $A(L)y_t = v_t$
- Corresponds to the reduced-form in the simultaneous equation literature if the system is closed (no exogenous).
- When $A(1)$ is rank deficient, i.e. $0 < r < n$, the variables in y_t share $n-r$ common stochastic trends, and co-evolve (cointegration).
- **Problem:** how to select the variables included in y_t vector and the lag-order p ? Variables are selected with vague reference to “theory”, but the (blind?) choice has also implication on p and to move up the curve.

6

- Since y_t is selected from a larger universe of variables in w_t , even though w_t follows a VAR the subsystem y_t is bound to follow a VARMA. If we want to proxy MA with AR components, the order p must be larger than what typically set in order to appropriately approximate the impulse-response IR from: $y_t = D(L)v_t$. 

Exercise:

1. use a top-of-the-curve model to derive IR to a number of shocks for a subset of variables of interest;
2. fit VAR for the latter variables with alternative p in order to proxy previous “true” IR;
3. difficult representation [e.g. VAR(50)]! Probably, better fit results with careful selection of variables in VAR or by passing to VARMA.

7

- In case of reduced-rank $r < n$, $A(1)$ matrix is decomposed in $\alpha\beta'$, where β ($n \times r$) is supposed to reflect economic (too simple?) “long run relations”. Of course, unless the $r=1$, β matrix need (awkward!) identification.

Exercise:

1. permanent shocks v_t are given to a top-of-the-curve model to obtain permanent increases $D(L)$ –through IR– in the variables of interest: $y_t = D(L) v_t$;
 2. given that β must satisfy $\beta' D(1) = 0$ we can obtain cointegrating vectors’ parameters;
 3. compare these β estimates with β estimates from true data VAR-modeled and applying Johansen’s estimator.
- Again, conflicts between theory and unstructured modeling methodology (i.e. cointegration) emerge. 8

A step upwards: SVAR

- Moving up towards the economic theoretical models, we encounter the Structural VAR representations by re-expressing the statistical model in a form that enables us to tell an economic story.
- From the VAR $A(L)y_t = v_t$ to the SVAR $B(L)y_t = \varepsilon_t$ where the shocks ε_t are now given economic names such as “demand” or “supply” shocks, and are supposed to be uncorrelated with one another.
- The naming of the shocks introduces economic content. To do so, a number of alternative identifying restrictions must be imposed to recover B parameters from A and $vcov(v_t)$.
- The uncorrelated shock assumption is the weaker point, as it can be shown that the omission of a VAR member leads to common factors → biased SVARs. 9

About the need of going up

- Though prevalently statistical, VARs entail choices which have effects on the following steps towards more economic interpretation, namely:
 1. the n variables to be modeled in the vector;
 2. the lag length p .
- Omitting variables entails biased IR of SVAR due to invalid uncorrelated shocks assumptions.
- Of course, if the aim is only that of forecasting, the problems above are downsized. However, it remains the problem of the high number n ($np+1$) of parameters to be estimated in VAR models: inefficient overfitting (escapes: BVAR or F-VAR ▷).
- This fact weakens the power of in-sample Granger causality about predictability. ▷ 10

The history of macro-modelling

- From Tinbergen (1939), through Klein (1950) and Klein-Goldberger (1955), till Duesenberry et al (1965, 1969). Construction of structural macro-econometric models in the field of the Cowles Commission approach.
- Two events in the 1970s
 1. Commercialization of models: from basic research to day-to-day updating (supply side?). Adjusting forecasts to “reasonable” outcomes. Statistical extrapolations were as good as large-scale models (see Nelson, 1972).
 2. Lucas’s (1976) critique: models are not useful for policy purposes. Real-business-cycle theories (not seriously interested in testing models) and the new Keynesian counter-response (macroeconomics away from its econometric base).
- What happened during the first 50 years?

11

Chronology of papers

Keynes (1936), *General Theory*

Tinbergen (1939), *Statistical Testing of Business Cycle Theories*

Klein (1946), *The keynesian Revolution*

Klein (1950), *Economic Fluctuations in the US*

Klein-Goldberger (1955), *An Econometric Model for the US*

Duesenberry et al (1965), *The Brookings Quarterly Econometric Model of the US*

Duesenberry et al (1969), *The Brookings Model: Some Further Results*

Phelps et al (1970), *Microeconomic Foundations of Employment and Inflation Theory*

Nelson (1972), *The Prediction Performance of the FRB/MIT/Penn Model of the US*

Granger-Newbold (1974), *Spurious Regressions in Econometrics*

Sargent-Wallace (1975), *Rational’ Expectations, the Optimal Money Instrument, and the Optimal Money Supply Rule*

Lucas (1976), *Econometric Policy Evaluation: A Critique*

Box-Jenkins (1976), *Time Series Analysis, Forecasting and Control*

Sargent-Sims (1977), *Business Cycle Modelling Without Pretending to Have too Much a Priori Theory*

Sims (1980), *Macroeconomics and Reality*

12

Dickey-Fuller (1979), *Distribution of the Estimators for Autoregressive Time Series With a Unit Root*

Parke (1982), *An Algorithm for FIML and 3SLS Estimation of Large Non-Linear Models*

Kydland-Prescott (1982), *Time to Build and Aggregate Fluctuations*

Hansen (1982), *Large Sample Properties of GMM Estimators*

Engle (1982), *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance in the UK Inflation*

Fair-Taylor (1983), *Solution and ML Estimation of Dynamic RE Models*

Hendry-Pagan-Sargan (1984), *Dynamic Specifications*

Cooley-Leroy (1985), *Atheoretical Macroeconomics: A Critique*

Engle-Granger (1987), *Cointegration and Error-Correction: Representation, Estimation and Testing*

Hamilton (1989), *A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle*

Harvey (1989), *Forecasting Structural Time Series Models and the Kalman Filter*

Stock-Watson (1989), *New Indexes of Coincident and Leading Economic Indicators*

Christiano-Eichenbaum (1992), *Current Real Business Cycle Theories and Aggregate Labour Market Fluctuations*

13

Hybrid models (late 80s-early 90s)

- The core principle is articulated in two stages:
 1. there is an (implicit/explicit) equilibrium path along which the economy evolves;
 2. define the adjustment to the path.
- In type I models, the long run is implicit (fairly loose economic reasoning), as equilibrium path is not computed, but the relationship is only described. The existence of a (trending) long run is assessed through cointegration tests. The ECM parameters estimation is data-driven. Special attention is paid on the speed at which the disequilibrium gap is closed.
- Type II models explicitly describe the equilibrium path of the variables (as the DSGE model does) on the basis of an optimized framework with forward-looking behavior. Data still set the adjustment process.

14

- A further step towards upper theory is to acknowledge that the economic theory used to form the equilibrium paths can also be used to describe the adjustment path: the dynamics has to be intrinsic to the model instead of extrinsic (i.e. data-driven).
- However, to better fitting, models incorporate lags proxying some inertia in decisions. These models are termed in Figure above as Incomplete DSGE.
- IDSGE model builders feel the data might not accurately measure the variables of interest, and are quite cautious about the role of data.
- Basic idea: the econometric model should focus on capturing the underlying economics to describe how economy works, rather than overfitting idiosyncrasies.
- In this, calibration reflects staff judgment of appropriate properties and may be better than estimation.

15

Upper Structural Models

- Lucas paved the way for a new macroeconomics based on dynamic stochastic model economies with fully-articulated preferences, technologies, and rules of the game. Nothing new: neoclassical economics.
- Hence the name of Dynamic Stochastic General Equilibrium models (DSGE), built on a foundation of fully-specified stochastic dynamic optimiz. behaviors.
- Calibration: learning about the properties of the DSGE model and assess its agreement with data based on simulations. Non-probabilistic, but take misspecification seriously.
- ML estimates through a state-space representation using the Kalman filter. It is probabilistic, but may not take misspecification seriously enough.
- Bayesian approach? But what is the influence of priors?

The Global VAR approach

- **Need:** increased globalization of the world economy has important consequences for the conduct of monetary and financial policies by central bankers and risk management by commercial bankers. As a result, many different channels of transmission must be taken into account: national economic issues need to be considered from a global perspective.
- **Answer:** a global macroeconomic model that is able of making forecasts for a core set of variables for a set of regions and countries that explicitly allows for inter-connections and interdependencies between national and international factors in a coherent and consistent manner.

17

Methodological foundations

- 1) *The analysis of cointegrating systems* would provide the global modeling framework, but usually confined to a single country covering only some of the key macro-economic variables (see, e.g., King, Plosser, Stock, and Watson, 1991; and Garratt, Lee, Pesaran, and Shin, 2000 and 2003).
- 2) *Impulse-response in unrestricted VAR* that (covering N regions, k key variables, and p lags) would require $p(kN - 1)$ unknown parameters [plus additional deterministic / exogenous variables' parameters] to be estimated for each equation. Therefore, it is unfeasible unless imposing restrictions. Further: shock identification issue...

18

3) **Linking up macroeconomic models:** traditionally developed for the national economies; see e.g. Lawrence Klein's Project Link adopted by the UN, the IMF and OECD national models' links (smaller scale), and NiGEM.

BROAD IDEA

Use the techniques (1) to model restricted VARs (2) by conditioning the domestic macro variables to foreign variables constructed to match the international trade pattern of each country (3).

In fact, these country-specific foreign variables can be treated as weakly exogenous (or long-run forcing) for most economies (except US), when N is sufficiently large and the idiosyncratic shocks are weakly correlated.

19

Intuition: start from a VAR(1) model for one variable and N=4 countries

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \\ y_{4t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \\ \epsilon_{4t} \end{pmatrix}$$

The generic equation for the i-th country is:

$$y_{it} = a_i + b_{i1}y_{1t-1} + b_{i2}y_{2t-1} + b_{i3}y_{3t-1} + b_{i4}y_{4t-1} + \epsilon_{it}$$

The corresponding i-th foreign variable is defined as ($w_{ii} = 0 \forall i$):

$$y_{it-1}^* \equiv w_{i1}y_{1t-1} + w_{i2}y_{2t-1} + w_{i3}y_{3t-1} + w_{i4}y_{4t-1}$$

In doing so, 2N parameters instead of N² because w_{ij} are numbers

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ y_{4t} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} + \begin{pmatrix} b_1 & & & \\ & b_2 & & \\ & & b_3 & \\ & & & b_4 \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \\ y_{3t-1} \\ y_{4t-1} \end{pmatrix} + \begin{pmatrix} c_1 & & & \\ & c_2 & & \\ & & c_3 & \\ & & & c_4 \end{pmatrix} \begin{pmatrix} y_{1t-1}^* \\ y_{2t-1}^* \\ y_{3t-1}^* \\ y_{4t-1}^* \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \\ \epsilon_{4t} \end{pmatrix}$$

20

N+1 country-specific VARX models

- Countries/regions are indexed by $i = 0, 1, 2, \dots, N$ (country 0, US, is the reference country)
- The k_i country-specific variables \mathbf{x}_{it} are related to the k_i^* global economy variables \mathbf{x}_{it}^* (country specific weighted averages of foreign variables plus deterministic and global weakly exogenous variables) by a first order dynamic specification.
- The remaining idiosyncratic shocks $\boldsymbol{\varepsilon}_{it}$ are assumed to be serially uncorrelated with zero mean and nonsingular covariance matrix (possibly switching between “normal” and “crisis” set of values), but mildly correlated across regions.

21

$$\mathbf{x}_{it} = \mathbf{a}_{i0} + \mathbf{a}_{i1}t + \boldsymbol{\Phi}_i \mathbf{x}_{i,t-1} + \boldsymbol{\Lambda}_{i0} \mathbf{x}_{it}^* + \boldsymbol{\Lambda}_{i1} \mathbf{x}_{i,t-1}^* + \boldsymbol{\varepsilon}_{it},$$

$$t = 1, 2, \dots, T, i = 0, 1, 2, \dots, N, \quad (1)$$

where $\boldsymbol{\Phi}_i$ is a $k_i \times k_i$ matrix of lagged coefficients, $\boldsymbol{\Lambda}_{i0}$ and $\boldsymbol{\Lambda}_{i1}$ are $k_i \times k_i^*$ matrices of coefficients associated with the foreign-specific variables, and $\boldsymbol{\varepsilon}_{it}$ is a $k_i \times 1$ vector of idiosyncratic country-specific shocks. In the special case where $\boldsymbol{\Lambda}_{i0} = \boldsymbol{\Lambda}_{i1} = \mathbf{0}$, this model reduces to a standard VAR process of order 1, VAR(1). However, in the presence of foreign-specific variables, (1) is an augmented VAR model, which we denote by VARX*(1, 1).

22

$\mathbf{x}_{it} = (y_{it}, p_{it}, q_{it}, e_{it}, \rho_{it}, m_{it})'$, with $k_i = 6$.

$$y_{it} = \ln(GDP_{it}/CPI_{it}), \quad p_{it} = \ln(CPI_{it}),$$

$$q_{it} = \ln(EQ_{it}/CPI_{it}), \quad m_{it} = \ln(M_{it}/CPI_{it}),$$

$$e_{it} = \ln(E_{it}), \quad \rho_{it} = .25 * \ln(1 + R_{it}/100),$$

$$\mathbf{x}_{it}^* = (y_{it}^*, p_{it}^*, q_{it}^*, e_{it}^*, \rho_{it}^*, m_{it}^*)'$$

$$y_{it}^* = \sum_{j=0}^N w_{ij}^y y_{jt}, \quad p_{it}^* = \sum_{j=0}^N w_{ij}^p p_{jt},$$

$$q_{it}^* = \sum_{j=0}^N w_{ij}^q q_{jt}, \quad e_{it}^* = \sum_{j=1}^N w_{ij}^e e_{jt},$$

$$\rho_{it}^* = \sum_{j=0}^N w_{ij}^{\rho} \rho_{jt}, \quad m_{it}^* = \sum_{j=0}^N w_{ij}^m m_{jt}.$$

The weights w_{ij} are based on trade shares (namely, the share of country j in the total trade of country i measured in US \$) in the case of real variables, and capital flows in the case of equity price indices and interest rates; $w_{ii} = 0 \forall i$

23

VARX identification-estimation-testing

- Integration tests for \mathbf{x}_{it} and \mathbf{x}_{it}^* being $\sim I(1)$
- Cointegration rank & identification/estimation of the long-run relationship between \mathbf{x}_{it} and \mathbf{x}_{it}^* variables
- Weak exogeneity of \mathbf{x}_{it}^* (i.e. no feedback of \mathbf{x}_{it} on \mathbf{x}_{it}^*) that allows causal simultaneity going from \mathbf{x}_{it}^* to \mathbf{x}_{it}
- Tests for structural breaks based on equation residuals. Being \mathbf{x}_{it} conditional on \mathbf{x}_{it}^* , co-breaking should prevent generalized breaking
- Robustness: check for cross-sectionally weakly correlated idiosyncratic shocks, so that $\text{cov}(\mathbf{x}_{it}^* \boldsymbol{\epsilon}_{it}) \rightarrow 0$ as $N \rightarrow \infty$ (additional evidence of weak exogeneity, crucial assptn)

24

Putting the VARX together in the GVAR

After all above steps are accomplished, the $N + 1$ country-specific VARX models, together with the relations linking the (weakly) exogenous variables of the country specific models to the variables in the rest of the global model, provide the complete GVAR system.

In general, the GVAR model allows for interactions among the different economies through three separate but interrelated channels:

1. Contemporaneous dependence of \mathbf{x}_{it} on \mathbf{x}_{it}^* and on its lagged values.
2. Dependence of the country-specific variables on common global exogenous variables, such as oil prices (see Sec. 5).
3. Nonzero contemporaneous dependence of shocks in country i on the shocks in country j , measured via the cross-country covariances, Σ_{ij} ,

$$\Sigma_{ij} = \text{cov}(\boldsymbol{\varepsilon}_{it}, \boldsymbol{\varepsilon}_{jt}) = E(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{jt}') \quad \text{for } i \neq j, \quad (6) \quad 25$$

Structure and solution of the GVAR model

To construct the GVAR model from the country-specific models, we first define the $(k_i + k_i^*) \times 1$ vector

$$\mathbf{z}_{it} = \begin{pmatrix} \mathbf{x}_{it} \\ \mathbf{x}_{it}^* \end{pmatrix}, \quad (7)$$

and then rewrite (1) as

$$\mathbf{A}_i \mathbf{z}_{it} = \mathbf{a}_{i0} + \mathbf{a}_{i1}t + \mathbf{B}_i \mathbf{z}_{i,t-1} + \boldsymbol{\varepsilon}_{it}, \quad (8)$$

where

$$\mathbf{A}_i = (\mathbf{I}_{k_i}, -\mathbf{A}_{i0}) \quad \text{and} \quad \mathbf{B}_i = (\Phi_i, \mathbf{A}_{i1}). \quad (9)$$

The dimensions of \mathbf{A}_i and \mathbf{B}_i are $k_i \times (k_i + k_i^*)$, and \mathbf{A}_i has a full row rank, namely $\text{rank}(\mathbf{A}_i) = k_i$.

Collect all the country-specific variables in the k -vector $\mathbf{x}_t = (\mathbf{x}'_{0t}, \mathbf{x}'_{1t}, \mathbf{x}'_{2t}, \dots, \mathbf{x}'_{Nt})$ and, with $w_{ij} : \mathbf{z}_{it} = \mathbf{W} \mathbf{x}_t$ 26

$$\mathbf{G}\mathbf{x}_t = \mathbf{a}_0 + \mathbf{a}_1 t + \mathbf{H}\mathbf{x}_{t-1} + \boldsymbol{\varepsilon}_t,$$

where

$$\mathbf{a}_0 = \begin{pmatrix} \mathbf{a}_{00} \\ \mathbf{a}_{10} \\ \vdots \\ \mathbf{a}_{N0} \end{pmatrix}, \quad \mathbf{a}_1 = \begin{pmatrix} \mathbf{a}_{01} \\ \mathbf{a}_{11} \\ \vdots \\ \mathbf{a}_{N1} \end{pmatrix}, \quad \boldsymbol{\varepsilon}_t = \begin{pmatrix} \varepsilon_{0t} \\ \varepsilon_{1t} \\ \vdots \\ \varepsilon_{Nt} \end{pmatrix}$$

and

$$\mathbf{G} = \begin{pmatrix} \mathbf{A}_0 \mathbf{W}_0 \\ \mathbf{A}_1 \mathbf{W}_1 \\ \vdots \\ \mathbf{A}_N \mathbf{W}_N \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} \mathbf{B}_0 \mathbf{W}_0 \\ \mathbf{B}_1 \mathbf{W}_1 \\ \vdots \\ \mathbf{B}_N \mathbf{W}_N \end{pmatrix}.$$

which can be solved recursively forward (i.e. iterating) to obtain the future values of \mathbf{x}_t

$$\mathbf{x}_t = \mathbf{G}^{-1} \mathbf{a}_0 + \mathbf{G}^{-1} \mathbf{a}_1 t + \mathbf{G}^{-1} \mathbf{H} \mathbf{x}_{t-1} + \mathbf{G}^{-1} \boldsymbol{\varepsilon}_t,$$


27

Forecasting with GVAR models

Motivation: “Suppose one was interested in forecasting growth and inflation across a number of different countries; how would one go about it? What additional variables might help in such forecasting (the oil price comes to mind), and should one also consider adding financial variables such as equity returns and the long-term interest rate? Should they be treated separately (two isolated equations) or together, say in a VAR? Should one consider only domestic, or also foreign variables? If foreign variables are included, should they be endogenised as well? How important are cointegrating relationships, either across variables within a country or across countries (PPP relationships come to mind)? And how should one address the ever-present problem of structural breaks, which may happen multiple times in any one or several of the relationships in the forecasting model under consideration?” GVAR is a possible answer.

28

Paper: Pesaran, Schuermann, Smith (2009), IJF

- Out-of-sample 1- and 4-step ahead forecasts over 2004q1-2005q4 (8 quarters), based on a fixed model searched over the 1979q1-2003q4 period and estimated over 10 estimation windows (W1 starts 1979q1, W2 1979q2, ..., W10 1981q2)
- Benchmarks: univariate AR(1) and RW (with/without trend/drift)
- Premise: all modelling exercises face the trade-off between bias and efficiency, and model averaging serves to increase the latter
- The present situation is it not an exception to the rule: huge size of the problem (134 variables, 26 regions), heterogeneity of the economies, and structural breaks
- Therefore, AveAve forecasts (i.e. averaging forecasts from different GVAR models AveM estimated over different windows AveW) of inflation, growth and equity prices outperform benchmarks 
- However, performance is mixed for interest & exchange rates

Impulse-response: IR

Simplest case: the AR(1) model:

$y_t = a + b y_{t-1} + v_t \rightarrow (1 - bL) y_t = a + v_t$; then if the model is stationary, the Wold representation:

$y_t = y^* + \sum_{i=0, \infty} b^i v_{t-i}$ [where $y^* = a(1-b)^{-1}$], i.e.

$y_t = y^* + b^0 v_t + b^1 v_{t-1} + b^2 v_{t-2} + b^3 v_{t-3} + \dots$

impact effect $\frac{\partial y_t}{\partial v_t} = b^0 = 1$

one period later $\frac{\partial y_t}{\partial v_{t-1}} = b^1 = b$ Being stationary,
the variable(s)

many periods later $\frac{\partial y_t}{\partial v_{t-\infty}} = b^\infty \approx 0$ converge to y^*



The Fair VAR model

- When testing the forecasting ability of complete macro-structural models, VAR models provide useful benchmarks. This is the Fair-Shiller test context.
- The curse of dimensionality with VAR models may be reduced when we are interested in forecasting a variable for which a number of disaggregated items are provided (e.g. the demand side bridge models for GDP).
- In this context, the GDP may be predicted by more accurate autoregressive components, AC, models in which each GDP component is regressed on its own lagged values and lagged values of GDP. GDP is then determined from the GDP identity. In this way, AC does not have the VAR problem of many parameters as the number of variables (components in the AC case) is increased.

32

Granger causality

- Causality statement embodies 2 components:
 1. the cause occurs before the effect;
 2. the cause contains information about the effect that is unique, i.e. it is in no other variable.
- As consequence, the causal variable can help forecast the effect variable after other data have first been used.
- G-causality entails out-of-sample predictability, not mere better fitting-in sample.
- Make two models with/without the possible causing variable.
- Using a forecasts comparison test, ask which model actually forecasts the better out-of-sample.

Causation vs correlation

- The temporal ordering of variables is relevant but not conclusive, as correlation does not imply causation. Examples
 - Weather forecasts and sun/rain
 - Shoe size and reading abilityboth above do not cause, but predict (in a stable structural environment). Remember the issue of predictive ability testing.
- The word “cause” is a reference to some hypothetical intervention. Experiment: put a gun to the head of the weather forecaster and making her say “sunny”. Collecting experimental evidence to assess causation.
- Social scientists (as detectives) very often work with available evidence, so how to make causal inferences from non-experimental data?
- We must rely on causal models (i.e. theoretical assumptions) that are in our heads, not in the data. The model tells how/why some facts/intervention explain (could cause) actual outcomes.



Keynes (1939, EJ) objections to Tinbergen

1. **Omitted variables.** Tinbergen says “the main purpose of the econometrician is to discover with what strength (i.e. the quantitative importance of) each of the correctly listed causes operates”, but this does not need the mere list of the causes of interest to the researcher/policymaker, but a complete list!
2. **Presence of non-measurable phenomena.** Tinbergen: claims that it can be supplemented by other information, but how?
3. **Spurious correlation and simultaneity.** Are we confident that the different factors are substantially independent of one another, and no feedback?
4. **Linearity assumption.** What implications?
5. **Ad hoc time-lags and trends.** Trial-and-error method: he fidgets about until he finds something not fitting too badly with theory.
6. **Inductive claim on the future or simple curve fitting?** He would be better to check if the same relationships are stable over time.

Bayesian averaging with unstable models

Model averaging and forecast combination have a rich history in statistics and forecasting. An early survey of the literature on forecast combination was provided by Bates and Granger (1969), while Timmermann (2006) is the more recent survey.

Given a set of admissible models, **AveM** is the average of their forecasts based on the same estimation window

On the other side, each model is likely to be unstable; for this **AveW** averages the forecasts of the same model estimated over different sampling windows (from minimum to the largest)

Choice of the weighting scheme: Timmermann (2006) reports that across many different empirical applications, the equal weighting scheme tends to beat other forecast combination procedures.