Techniques of Empirical Econometrics

Oscar Jordà
UC Davis, FRB San Francisco

Karel Mertens
FRB Dallas, CEPR

AEA Continuing Education, January 8-10, 2022

The views expressed do not necessarily reflect the views of the Federal Reserve Bank of Dallas, San Francisco or the Federal Reserve System.

Overview

1 **Time Series Representations of Dynamic Macro Models**
Structural State Space Models, MA, VARMA and VAR representations; Estimating Dynamic Causal Effects; Misspecification: Nonfundamentalness, Nonlinearities, and Time Aggregation

2 **State-Space Models and the Kalman Filter**
State Space Models, Kalman Filter, Forecasting, Maximum Likelihood Estimation

3 **Local Projections**
Impulse Responses as Dynamic Treatment Effects, LP Estimation and Basic Inference, VAR-LP Impulse Response Equivalence

4 **Identification of Dynamic Causal Effects**
Identification with Covariance Restrictions or Higher Moments. Proxy SVAR/SVAR-IV, Internal instrument SVAR

5 **Inference for Impulse Responses**
Inference methods for VAR/LP impulse responses. Detecting weak instruments; Robust Inference Methods; Joint inference for VAR and LP impulse responses

6 **Impulse Response Heterogeneity**
Kitagawa Decomposition, Time Varying Impulse Responses

7 **Other Uses of Impulse Responses**
Impulse Response Matching and Indirect Inference; Estimating Structural Single Equations using Impulse Responses, SP-IV; Counterfactuals with Impulse Responses, Optimal Policy Perturbations

1. Time Series Representations of Dynamic Macro Models

1.1 Structural Time Series Representations

- Structural State Space Model
- SMA Representation
- SVARMA and SVAR Representations

1.2 Estimating Dynamic Causal Effects

- When the Shock Realizations are Observed
- When the Contemporaneous Impact is Observed

1.3 What Can Go Wrong? Sources of Misspecification

- Nonfundamentalness
- Nonlinearities
- Time Aggregation

State Space Representation of Structural Models

Structural State Space Model

$$
\begin{aligned}
s_t &= \mathcal{G} s_{t-1} + \mathcal{F}\epsilon_t \\
z_t &= \mathcal{A} s_{t-1} + \mathcal{D}\epsilon_t
\end{aligned}
$$

s_t is a $N_s \times 1$ vector of (unobserved) state variables

ϵ_t is an $N_\epsilon \times 1$ vector of orthogonal white noise, or **structural shocks**,

$$
E\left[\epsilon_t\right] = 0\ ,\ E\left[\epsilon_t \epsilon_t'\right] = \mathcal{I}\ ,\ E\left[\epsilon_t \epsilon_s'\right] = 0 \text{ for } s \neq t
$$

z_t is an $N_z \times 1$ vector of (observed) variables of interest

$\mathcal{G}$ is $N_s \times N_s$, $\mathcal{F}$ is $N_s \times N_\epsilon$, $\mathcal{A}$ is $N_z \times N_s$ and $\mathcal{D}$ is $N_z \times N_\epsilon$

The solution of linear dynamic models can generally be written as an SS model.

In structural models, $\{\mathcal{G}, \mathcal{F}, \mathcal{A}, \mathcal{D}\}$ are specific functions of deep structural parameters (preferences, technology, ...)

One-Step-Ahead Forecasts:

$z_t^{\perp} = z_t - E[z_t | s_{t-1}] = \mathcal{D}\epsilon_t$ with $E[z_t^{\perp} z_t^{\perp}{}'] = \mathcal{D}\mathcal{D}'$

Example: New Keynesian Model

Simple New Keynesian model, e.g. Galí (2015)

$$
\begin{aligned}
R_t &= \phi_\pi \pi_t && \text{(Taylor Rule/LM curve)} \\
E_t \Delta gap_{t+1} &= R_t - E_t \pi_{t+1} - s_t^d && \text{(Eq. Cons Euler/IS curve)} \\
\pi_t &= \kappa gap_t + \beta E_t \pi_{t+1} + s_t^s && \text{(Phillips curve/AS curve)}
\end{aligned}
$$

where $\kappa > 0$, $\phi_\pi > 1$, $0 \leq \beta < 1$

gap_t : output gap , π_t : inflation

s_t^s: cost push factor

s_t^d: demand factor

s_t^s and s_t^d are stationary exogenous processes

Example: New Keynesian Model

Iterating forward,

$$\left[\begin{array}{c} gap_t \\ \pi_t \end{array}\right] = E_t \sum_{i=0}^{\infty} \mathcal{C}^{-(i+1)} \left[\begin{array}{c} s^d_{t+i} \\ s^s_{t+i} \end{array}\right]$$

where

$$\mathcal{C}^{-1} = \frac{1}{1+\phi_\pi \kappa} \left[\begin{array}{cc} 1 & 1-\beta\phi_\pi \\ \kappa & \beta+\kappa \end{array}\right]$$

$\mathcal{C}^{-1}$ has eigenvalues strictly less than one in modulus if $\phi_\pi > 0$, i.e. the Taylor principle holds. check

Example: New Keynesian Model

Suppose the exogenous factors follow the (stable) process:

$$\underbrace{\begin{bmatrix} s_t^d \\ s_t^s \end{bmatrix}}_{s_t} = \underbrace{\Lambda}_{\mathcal{G}} \begin{bmatrix} s_{t-1}^d \\ s_{t-1}^s \end{bmatrix} + \underbrace{\Sigma_\epsilon}_{\mathcal{F}} \epsilon_t$$

Then,

$$\begin{aligned} \underbrace{\begin{bmatrix} gap_t \\ \pi_t \end{bmatrix}}_{z_t} &= \sum_{j=0}^{\infty} \mathcal{C}^{-(j+1)} \Lambda^j \begin{bmatrix} s_t^d \\ s_t^s \end{bmatrix} \\ &= \underbrace{(\mathcal{C} - \Lambda)^{-1} \Lambda}_{\mathcal{A}} \underbrace{\begin{bmatrix} s_{t-1}^d \\ s_{t-1}^s \end{bmatrix}}_{s_{t-1}} + \underbrace{(\mathcal{C} - \Lambda)^{-1} \Sigma_\epsilon}_{\mathcal{D}} \epsilon_t \end{aligned}$$

This describes the (rational expectations) solution to the model in state space form.

Highlighting Some Key Assumptions

Assumption: Linearity

$$\begin{aligned} s_t &= \mathcal{G} s_{t-1} + \mathcal{F} \epsilon_t \\ z_t &= \mathcal{A} s_{t-1} + \mathcal{D} \epsilon_t \end{aligned}$$

Assumption: Stationarity

$N_s \times N_s$ matrix $\mathcal{G}$ has all eigenvalues strictly less than one in modulus, i.e.

- $det(\mathcal{G} - \lambda \mathcal{I}) \neq 0$ for $\mid \lambda \mid \geq 1$, or equivalently
- $det(\mathcal{I} - \mathcal{G} \lambda) \neq 0$ for $\mid \lambda \mid \leq 1$

As a result, s_t follows a stable VAR(1) process and $s_t = \sum_{i=0}^{\infty} \mathcal{G}^i \mathcal{F} \epsilon_{t-i}$.

s_t and z_t are **covariance stationary**, i.e. the first and second moments are time invariant.

Assumption: No Deterministic Terms

s_t and z_t are purely nondeterministic processes

Time Series Representations

Section 2 of the course will discuss the estimation of state space models

Under appropriate assumptions, z_t can also be represented in other ways

- **Structural Moving Average Process (SMA)**
- **Structural Vector Autoregressive Moving Average Process (SVARMA)**
- **Structural Vector Autoregressive Process (SVAR)**

Moving Average (MA) Process

Moving Average (MA) Process

$$\text{MA(q)} \quad : \quad z_t = M(L)\upsilon_t = \sum_{i=0}^{q} M_i \upsilon_{t-i}$$

where $M(L) = M_0 + M_1 L + ... + M_q L^q$, L is the lag operator L (i.e. $L^k x_t = x_{t-k}$), and υ_t is white noise, i.e.

$$E\left[\upsilon_t\right] = 0 \ , \ E\left[\upsilon_t \upsilon_t'\right] = \Sigma_\upsilon \ , \ E\left[\upsilon_t \upsilon_s'\right] = 0 \text{ for } s \neq t$$

The MA process above is **causal**, i.e. $E[z_t \upsilon_{t-i}'] = 0$ for all $i < 0$

SMA(∞) Representation

Structural MA(∞) representation

$$z_t = M(L)\epsilon_t = \sum_{h=0}^{\infty} M_h \epsilon_{t-h}$$

where $M_0 = \mathcal{D}$ and $M_h = \mathcal{A}\mathcal{G}^{h-1}\mathcal{F}$ for $h \geq 1$

$$\begin{aligned} z_t &= \underbrace{\mathcal{A}\sum_{i=1}^{\infty} \mathcal{G}^{i-1}\mathcal{F}\epsilon_{t-i}}_{s_{t-1}} + \mathcal{D}\epsilon_t = \mathcal{A}(\mathcal{I} - \mathcal{G}L)^{-1}\mathcal{F}\epsilon_{t-1} + \mathcal{D}\epsilon_t \\ &= \left(\mathcal{D} + \mathcal{A}(\mathcal{I} - \mathcal{G}L)^{-1}\mathcal{F}L\right)\epsilon_t \end{aligned}$$

z_t is a linear function of all current and past realizations of ϵ_t

Note it is not necessarily the case that $N_z = N_\epsilon$

One-Step-Ahead Forecasts:

$z_t^{\perp} = z_t - E[z_t | \epsilon_{t-1}, \epsilon_{t-2}, \ldots] = \mathcal{D}\epsilon_t$ with $E[z_t^{\perp} z_t^{\perp\prime}] = \mathcal{D}\mathcal{D}'$

SMA(∞) and Dynamic Causal Effects

Dynamic Causal Effect

The dynamic causal effect of a unit intervention in $\epsilon_{j,t} \in \epsilon_t$ on z_{t+h} is $E[z_{t+h}|\epsilon_{j,t} = 1, \epsilon_{t-1}, ...] - E[z_{t+h}|\epsilon_{j,t} = 0, \epsilon_{t-1}, ...]$

Also known as **'structural' impulse response function (IRF) coefficients** and equal to $\partial z_{t+h}/\partial \epsilon_{j,t}$ for $h = 0, 1, ...$ in linear models.

The structural impulse response coefficients of z_t to shock $\epsilon_{j,t}$ at horizon h are the elements in the j-th column of M_h in the SMA(∞) representation

$$z_t \quad = \quad M(L)\epsilon_t = \sum_{h=0}^{\infty} M_h \epsilon_{t-h}$$

The SMA(∞) representation contains the dynamic causal effects/structural impulse responses to all shocks ϵ_t

Example: New Keynesian Model

State Space Representation:

$$\begin{aligned}
\begin{bmatrix} s_t^d \\ s_t^s \end{bmatrix} &= \Lambda \begin{bmatrix} s_{t-1}^d \\ s_{t-1}^s \end{bmatrix} + \Sigma_\epsilon \epsilon_t \\
\begin{bmatrix} gap_t \\ \pi_t \end{bmatrix} &= (\mathcal{C} - \Lambda)^{-1} \Lambda \begin{bmatrix} s_{t-1}^d \\ s_{t-1}^s \end{bmatrix} + (\mathcal{C} - \Lambda)^{-1} \Sigma_\epsilon \epsilon_t
\end{aligned}$$

SMA Representation

$$\begin{bmatrix} gap_t \\ \pi_t \end{bmatrix} = \sum_{h=0}^{\infty} (\mathcal{C} - \Lambda)^{-1} \Lambda^h \Sigma_\epsilon \epsilon_{t-h}$$

The impulse response coefficients are

$$\underbrace{(\mathcal{C} - \Lambda)^{-1} \Sigma_\epsilon}_{h=0}, \underbrace{(\mathcal{C} - \Lambda)^{-1} \Lambda \Sigma_\epsilon}_{h=1}, \underbrace{(\mathcal{C} - \Lambda)^{-1} \Lambda^2 \Sigma_\epsilon}_{h=2}, \ldots, \underbrace{(\mathcal{C} - \Lambda)^{-1} \Lambda^h \Sigma_\epsilon}_{h}, \ldots, \underbrace{\mathbf{0}_{2\times 2}}_{h=\infty}$$

Vector Autoregressive Moving Average Process

Vector Autoregressive Moving Average (VARMA) Process

$$\text{VARMA(p,q)} \quad : \quad A(L)z_t = C(L)v_t$$

where

$$\begin{aligned} A(L) &= \mathcal{I} - A_1 L - ... - A_p L^p \\ C(L) &= C_0 + C_1 L + ... + C_q L^q \end{aligned}$$

$$E\left[v_t\right] = 0 \ , \ E\left[v_t v_t'\right] = \Sigma_v \ , \ E\left[v_t v_s'\right] = 0 \text{ for } s \neq t$$

To arrive at a finite-order structural VARMA representation, consider

Nonsingularity of $\mathcal{A}$:

$N_z = N_s$ and $\mathcal{A}$ is an $N_z \times N_z$ nonsingular matrix

SVARMA Representation

Structural VARMA(1,1) Representation

Under **nonsingularity of** $\mathcal{A}$,

$$z_t = A_1 z_{t-1} + C_0 \epsilon_t + C_1 \epsilon_{t-1}$$

where $A_1 = \mathcal{A}\mathcal{G}\mathcal{A}^{-1}$, $C_0 = \mathcal{D}$ and $C_1 = \mathcal{A}(\mathcal{F} - \mathcal{G}\mathcal{A}^{-1}\mathcal{D})$

Recall the SMA(∞) representation, $z_t = \mathcal{A}(\mathcal{I} - \mathcal{G}L)^{-1}\mathcal{F}\epsilon_{t-1} + \mathcal{D}\epsilon_t$.
Under **nonsingularity of** $\mathcal{A}$,

$$\begin{aligned}
\mathcal{A}^{-1} z_t &= (\mathcal{I} - \mathcal{G}L)^{-1}\mathcal{F}\epsilon_{t-1} + \mathcal{A}^{-1}\mathcal{D}\epsilon_t \\
(\mathcal{I} - \mathcal{G}L)\mathcal{A}^{-1} z_t &= \mathcal{F}\epsilon_{t-1} + (\mathcal{I} - \mathcal{G}L)\mathcal{A}^{-1}\mathcal{D}\epsilon_t \\
z_t &= \mathcal{A}\mathcal{G}\mathcal{A}^{-1} z_{t-1} + \mathcal{D}\epsilon_t + \mathcal{A}(\mathcal{F} - \mathcal{G}\mathcal{A}^{-1}\mathcal{D})\epsilon_{t-1}
\end{aligned}$$

z_t is a linear function of z_{t-1}, ϵ_t and ϵ_{t-1}

Note it is not necessarily the case that $N_z = N_\epsilon$

One-Step-Ahead Forecasts:

$z_t^\perp = z_t - E[z_t | z_{t-1}, \epsilon_{t-1}] = \mathcal{D}\epsilon_t$ with $E[z_t^\perp z_t^{\perp\prime}] = \Sigma_v = \mathcal{D}\mathcal{D}'$

Vector Autoregressive Process

Vector Autoregressive Process

$$\text{VAR(p)}: \quad B(L)z_t = u_t$$

where

$$B(L) \quad = \quad I - B_1 L - ... - B_p L^p$$

$$E\left[u_t\right] = 0 \ , \ E\left[u_t u_t'\right] = \Sigma_u \ , \ E\left[u_t u_s'\right] = 0 \text{ for } s \neq t$$

To arrive at a structural VAR representation, consider

Assumption: Stochastic Nonsingularity

$N_z = N_\epsilon$ and $\mathcal{D}$ is an $N_z \times N_z$ nonsingular matrix

Assumption: Eigenvalue Condition for Invertibility in the Past (IP)

The eigenvalues of $\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}$ are strictly less then one in modulus.

SVAR Representation

Structural VAR(∞) Representation Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007)

Under **stochastic nonsingularity** and the **IP eigenvalue condition**,

$$z_t = \sum_{i=1}^{\infty} B_i z_{t-i} + \mathcal{D}\epsilon_t$$

where $B_i = \mathcal{A}\left(\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}\right)^{i-1}\mathcal{F}\mathcal{D}^{-1}$

Recall

$$z_t = \mathcal{A}s_{t-1} + \mathcal{D}\epsilon_t$$

Under **stochastic nonsingularity**

$$\epsilon_t = \mathcal{D}^{-1}(z_t - \mathcal{A}s_{t-1})$$

Substituting into $s_t = \mathcal{G}s_{t-1} + \mathcal{F}\epsilon_t$ yields

$$s_t = \left(\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}\right)s_{t-1} + \mathcal{F}\mathcal{D}^{-1}z_t$$

SVAR Representation

Under the **IP eigenvalue condition**,

$$s_t = \sum_{i=0}^{\infty} \left(\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}\right)^i \mathcal{F}\mathcal{D}^{-1} z_{t-i}$$

Substituting into $z_t = \mathcal{A}s_{t-1} + \mathcal{D}\epsilon_t$ leads to

$$z_t \quad = \quad \sum_{i=1}^{\infty} \mathcal{A} \left(\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}\right)^{i-1} \mathcal{F}\mathcal{D}^{-1} z_{t-i} + \mathcal{D}\epsilon_t$$

One-Step-Ahead Forecasts:

$z_t^{\perp} = z_t - E[z_t | z_{t-1}, z_{t-2}, ...] = u_t = \mathcal{D}\epsilon_t$ with $E[z_t^{\perp} z_t^{\perp\prime}] = \Sigma_u = \mathcal{D}\mathcal{D}'$

Example: New Keynesian Model

State Space Representation:

$$\begin{aligned}\begin{bmatrix} s_t^d \\ s_t^s \end{bmatrix} &= \Lambda \begin{bmatrix} s_{t-1}^d \\ s_{t-1}^s \end{bmatrix} + \Sigma_\epsilon \epsilon_t \\ \begin{bmatrix} gap_t \\ \pi_t \end{bmatrix} &= (\mathcal{C} - \Lambda)^{-1} \Lambda \begin{bmatrix} s_{t-1}^d \\ s_{t-1}^s \end{bmatrix} + (\mathcal{C} - \Lambda)^{-1} \Sigma_\epsilon \epsilon_t\end{aligned}$$

SMA Representation

$$\begin{bmatrix} gap_t \\ \pi_t \end{bmatrix} = \sum_{h=0}^{\infty} (\mathcal{C} - \Lambda)^{-1} \Lambda^h \Sigma_\epsilon \epsilon_{t-h}$$

SVAR/SVARMA Representation

$$\begin{bmatrix} gap_t \\ \pi_t \end{bmatrix} = (\mathcal{C} - \Lambda)^{-1} \Lambda (\mathcal{C} - \Lambda) \begin{bmatrix} gap_{t-1} \\ \pi_{t-1} \end{bmatrix} + (\mathcal{C} - \Lambda)^{-1} \Sigma_\epsilon \epsilon_t$$

Note that $\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A} = 0$ such that **eigenvalue condition for IP** holds.

In this example, there is an exact finite-order SVAR representation.

Example: New Keynesian Model

$$\begin{bmatrix} gap_t \\ \pi_t \end{bmatrix} = \underbrace{(\mathcal{C} - \Lambda)^{-1} \Lambda (\mathcal{C} - \Lambda)}_{B_1} \begin{bmatrix} gap_{t-1} \\ \pi_{t-1} \end{bmatrix} + \underbrace{(\mathcal{C} - \Lambda)^{-1} \Sigma_\epsilon}_{\mathcal{D}} \epsilon_t$$

Structural impulse responses by repeated substitution

$$\begin{aligned}
h = 0 &: \mathcal{D} = (\mathcal{C} - \Lambda)^{-1} \Sigma_\epsilon \\
h = 1 &: B_1 \mathcal{D} = (\mathcal{C} - \Lambda)^{-1} \Lambda \Sigma_\epsilon \\
h = 2 &: B_1^2 \mathcal{D} = (\mathcal{C} - \Lambda)^{-1} \Lambda^2 \Sigma_\epsilon \\
&\vdots \\
&\quad B_1^h \mathcal{D} = (\mathcal{C} - \Lambda)^{-1} \Lambda^h \Sigma_\epsilon \\
&\vdots \\
h = \infty &: B_1^\infty \mathcal{D} = \mathbf{0}_{2\times 2}
\end{aligned}$$

Alternative Derivation: From SMA to SVAR

Rewrite SMA(∞) representation under **stochastic nonsingularity**:

$$z_t \quad = \quad \left(I + \mathcal{A}(I - \mathcal{G}L)^{-1}\mathcal{F}\mathcal{D}^{-1}L\right) \upsilon_t \ , \ \upsilon_t = \mathcal{D}\epsilon_t$$

From SMA(∞) to SVAR(∞)

Under the **eigenvalue condition for IP**, υ_t is fundamental for z_t

Using the matrix determinant lemma,

$$det\left(I + \mathcal{A}(I - \mathcal{G}L)^{-1}\mathcal{F}\mathcal{D}^{-1}L\right) = det\left(I - (\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A})L\right)/det(I - \mathcal{G}L)$$

Fundamentalness of v_t for z_t requires that

$$det\left(I - (\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A})\lambda\right) \neq 0 \text{ for } \mid \lambda \mid \leq 1$$

which is equivalent to

$$det\left((\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}) - I\lambda\right) \neq 0 \text{ for } \mid \lambda \mid \geq 1$$

which holds if $\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}$ has all eigenvalues strictly less then one in modulus.

Equivalently ε_t is fundamental for z_t.

Alternative Derivation: From SVARMA to SVAR

Rewrite the SVARMA(1,1) representation under **stochastic nonsingularity**:

$$(I - \mathcal{A}\mathcal{G}\mathcal{A}^{-1}L)z_t \quad = \quad \left(I - \mathcal{A}\left(\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}\right)\mathcal{A}^{-1}L\right)v_t \ , \ v_t = \mathcal{D}\epsilon_t$$

From SVARMA(1,1) to SVAR(∞)

Under the **eigenvalue condition for IP**, v_t is fundamental for z_t

Fundamentalness of v_t for z_t requires that

$$det\left(I - \mathcal{A}\left(\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}\right)\mathcal{A}^{-1}\lambda\right) \neq 0 \text{ for } \mid \lambda \mid \leq 1$$

which is equivalent to

$$det\left(\left(\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}\right) - I\lambda\right) \neq 0 \text{ for } \mid \lambda \mid \geq 1$$

which holds if $\mathcal{G} - \mathcal{F}\mathcal{D}^{-1}\mathcal{A}$ has all eigenvalues strictly less then one in modulus.

Equivalently ε_t is fundamental for z_t.

From S(VAR)MA to SVAR in General: Fundamentalness

Fundamentalness

Consider a VARMA process $A(L)z_t = C(L)v_t$ with $\dim(z_t) = \dim(v_t)$. The innovations v_t are **fundamental** for observables z_t if there is an $S(L)$ in

$$S(L)A(L)z_t = S(L)C(L)v_t = v_t$$

such that $S(L)C(L) = \mathcal{I}$ and $S(L)$ only has nonnegative powers of L.

This requires that $\det(C(\lambda)) \neq 0$ for $|\lambda| < 1$, i.e. the zeroes of $\det(C(\lambda))$ do not lie strictly within the unit circle of the complex plane.

See e.g. Hansen and Sargent (1980), Hansen and Sargent (1991).

Fundamentalness implies a (one-sided) VAR representation $B(L)z_t = u_t$, where $B(L) = S(L)A(L)$ and $v_t = u_t$.

Invertibility in the past:
The inverse of $C(L)$ only involves non-negative powers of L (i.e. no leads).

Estimating Dynamic Causal Effects

Suppose the economic structure $\{\mathcal{G}, \mathcal{F}, \mathcal{A}, \mathcal{D}\}$ is unknown but we observe $\{z_t\}_{t=1}^T$

How can we estimate the causal effects of $\epsilon_{j,t}$ on z_{t+h} for $h = 0, 1, 2...$ with minimal assumptions on $\{\mathcal{G}, \mathcal{F}, \mathcal{A}, \mathcal{D}\}$?

Knowledge of dynamic covariances of z_t is not enough for identification of structural impulse responses.

We will discuss identification later in Section 4.

For now, we will assume either of the following:

- A sample of the shock of interest $\{\epsilon_{j,t}\}_{t=0}^T$ is observed
- The impact response coefficients $\mathcal{D}_j$ are known

Estimating Dynamic Causal Effects

Suppose we observe $\{\epsilon_{j,t}\}_{t=0}^{T}$.

The SMA(∞) representation motivates the following two approaches:

Distributed Lag Model

$$z_t = \sum_{h=0}^{H-1} M_h^j \epsilon_{j,t-h} + w_t \quad , \quad w_t = \sum_{h=H}^{\infty} M_h^j \epsilon_{j,t-h} + \sum_{h=0}^{\infty} M_h^{-j} \epsilon_{-j,t-h}$$

Since $E[\epsilon_{j,t-h} w_t] = 0$, the H structural IRF coefficients $M_0^j, \ldots M_{H-1}^j$ can be estimated by OLS in a regression of z_t on a distributed lag $\epsilon_{j,t}, \ldots, \epsilon_{j,t-H+1}$

Jordà (2005) Local Projections

$$z_{t+h} = M_h^j \epsilon_{j,t} + w_{h,t} \quad , \quad w_{h,t} = M_h^{-j} \epsilon_{-j,t} + \sum_{i=0, h\neq i}^{\infty} M_i \epsilon_{t+h-i}, h = 0, \ldots H-1$$

Since $E[\epsilon_{j,t} w_{h,t}] = 0$, the structural IRF coefficients M_h^j at horizon h can be estimated by OLS in a regression of z_{t+h} on $\epsilon_{j,t}$

Estimating Dynamic Causal Effects

Suppose we do not observe $\epsilon_{j,t}$, but we know $M_0^j = \mathcal{D}_j$ (j-th column in $\mathcal{D}$).

Assume ϵ_t is fundamental for z_t

VAR Projection

$z_t = \sum_{i=0}^{\infty} B_i z_{t-i} + u_t, \quad B(L) = \mathcal{I} - \sum_{i=1}^{\infty} B_i L^i$

$z_t = G(L) u_t$

where $G(L) = B(L)^{-1} = \sum_{h=0}^{\infty} G_h L^h$ are the **reduced form IRF coefficients**

Linear projection (population analogue of OLS regression) of z_t on $z_{t-1}, z_{t-2}, \ldots$

Local Projections on Current and Lagged Observables

$z_{t+h} = \mu_h z_t + \sum_{i=1}^{\infty} \delta_i z_{t-i} + w_{h,t}\ , \ \ h = 1, 2, ...$

Linear projection of z_{t+h} on z_t and $z_{t-1}, z_{t-2}, \ldots$

Equivalence of VAR and LPs Plagborg-Møller and Wolf (2021)

In population, $\mu_h = G_h$

In population, we can equivalently obtain the reduced form IRF coefficients $G(L)$ by VAR projection or LPs on current and lagged observables

Since ϵ_t is fundamental for z_t, $z_t = G(L)\mathcal{D}\epsilon_t$

The structural IRF coefficients associated with shock $\epsilon_{j,t}$ are in $G(L)\mathcal{D}_j$.

These are just linear combinations of the reduced form IRF coefficients in $G(L)$.

Fundamentalness and knowledge of $\mathcal{D}_j$ is equivalent to observing $\epsilon_{j,t}$

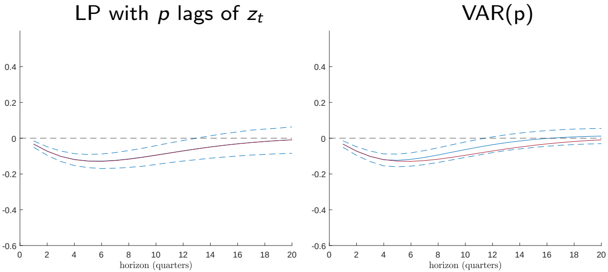
Estimating Dynamic Causal Effects

In finite samples, we cannot include infinite lags of z_t. Truncate the lag length to p.

Near-Equivalence of Truncated VAR and LPs Plagborg-Møller and Wolf (2021)

In population, $\mu_h \approx G_h$ for $h \leq p$

Estimated IRF of **output gap** to interest rate shock, $p = 4$, $N_z = 7$, $\mathbf{T} = \mathbf{5000}$:



Mean and 95% percentiles across 5000 Monte Carlo samples generated from Smets and Wouters (2007) model.

Estimating Dynamic Causal Effects

At short forecast horizons, LPs and VARs are close equivalents in population

At longer horizons ($h > p$), IRFs from finite-order VARs show lag truncation bias

VAR's extrapolate from the first p autocovariances of z_t by $G(L) = B(L)^{-1}$

$$G_0 = \mathcal{I} \ , \ G_h = \sum_{i=1}^{h} G_{h-i} B_i \quad , B_i = 0 \text{ for } h > p$$

Severity of extrapolation error depends on DGP

Other options to estimate $G(L)$ and avoid lag truncation bias:

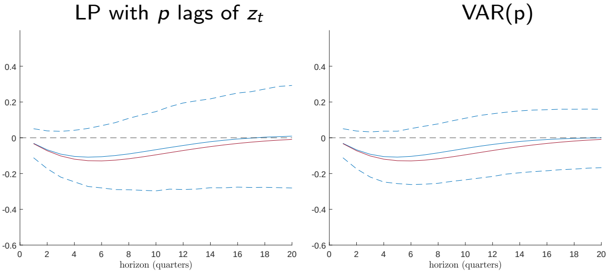
- Reduced Form State Space Models: Aoki (1987), Durbin and Koopman (2001)
- Reduced Form VARMA Models: Lütkepohl (2005), Kascha and Mertens (2009)

Estimating Dynamic Causal Effects

In finite samples, there is a bias-variance trade-off between LPs and VARs.

See Schorfheide (2005), Li, Plagborg-Møller, and Wolf (2022)

Estimated IRF of **output gap** to interest rate shock, $p = 4$, $N_z = 7$, $\mathbf{T} = \mathbf{250}$:



Mean and 95% percentiles across 5000 Monte Carlo samples generated from Smets and Wouters (2007) model.

Estimating a VAR

Sample of $T + p$ observations of an $N_z \times 1$ vector z_t:

$$\{z_{t-p+1}, z_{t-p+2}, \ldots, z_{T-1}, z_T\}$$

Define the $N_z \times T$ matrix z such that:

$$z = \left[\begin{array}{cccc} z_1 & z_2 & \cdots & z_T \end{array}\right]$$

Define a $pN_z \times 1$ vector Z_t:

$$Z_t = \left[\begin{array}{c} z_{t-1} \\ z_{t-2} \\ \vdots \\ z_{t-p} \end{array}\right]$$

Let Z be a $pN_z \times T$ matrix collecting T observations of Z_t:

$$Z = \left[\begin{array}{cccc} Z_1 & Z_2 & \cdots & Z_p \end{array}\right]$$

Estimating a VAR

Let u be a $N_z \times T$ matrix of $N_z \times 1$ residuals u_t:

$$u = \left[\begin{array}{cccc} u_1 & u_2 & \cdots & u_T \end{array}\right]$$

Let B be a $N_z \times pN_z$ matrix of coefficients:

$$B = \left[\begin{array}{cccc} B_1 & B_2 & \cdots & B_p \end{array}\right]$$

Introduce the vectorization operator:

$$\mathbf{z} = \text{vec}(z)$$
$$\mathbf{u} = \text{vec}(u)$$

Where $\mathbf{z}$ is a $N_z T \times 1$ vector of the stacked columns of z. The variance-covariance matrix of $\mathbf{u}$ is:

$$\mathrm{Var}(\mathbf{u}) \equiv \mathbf{\Sigma_u} = I_T \otimes \Sigma_u$$

Estimating a VAR

Re-write the VAR(p) as:

$$z = BZ + u$$

Or as:

$$\mathbf{z} = (Z' \otimes \mathcal{I}_{N_z})\beta + \mathbf{u}$$

Where $\otimes$ is the Kronecker product.

We can estimate β with Generalized Least Squares (GLS) by minimizing

$$\begin{aligned}
\mathbf{u}'\boldsymbol{\Sigma}_{\mathbf{u}}{}^{-1}\mathbf{u} &= (\mathbf{z} - (Z' \otimes \mathcal{I}_{N_z})\beta)'\boldsymbol{\Sigma}_{\mathbf{u}}{}^{-1}(\mathbf{z} - (Z' \otimes \mathcal{I}_{N_z})\beta) \\
&= \mathbf{z}'\boldsymbol{\Sigma}_{\mathbf{u}}{}^{-1}\mathbf{z} + \beta'(Z \otimes \mathcal{I}_{N_z})\boldsymbol{\Sigma}_{\mathbf{u}}{}^{-1}(Z' \otimes \mathcal{I}_{N_z})\beta - 2\beta'(Z \otimes \mathcal{I}_{N_z})\boldsymbol{\Sigma}_{\mathbf{u}}{}^{-1}\mathbf{z} \\
&= \mathbf{z}'(\mathcal{I}_T \otimes \Sigma_u^{-1})\mathbf{z} + \beta'(ZZ' \otimes \Sigma_u^{-1})\beta - 2\beta'(Z \otimes \Sigma_u^{-1})\mathbf{z}
\end{aligned}$$

Estimating a VAR

First order condition:

$$2(ZZ' \otimes \Sigma_u^{-1})\beta - 2(Z \otimes \Sigma_u^{-1})\mathbf{z} = 0$$

The GLS estimator is therefore:

$$\hat{\beta} = ((ZZ')^{-1}Z \otimes \mathcal{I}_{N_z})\mathbf{z}$$

Equivalent to 'system' OLS, equation-by-equation OLS, or Maximum Likelihood under Gaussianity.

Asymptotic normality:

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{\mathrm{d}} \mathcal{N}(0, \Gamma^{-1} \otimes \Sigma)$$

where $ZZ'/T \xrightarrow{p} \Gamma$ and the estimators are

$$\hat{\Gamma} = \frac{ZZ'}{T}$$
$$\hat{\Sigma} = \frac{1}{T - pN_z - 1} z(\mathcal{I}_T - Z'(ZZ')^{-1}Z)z'$$

Estimating Dynamic Causal Effects using a VAR

Note that $\hat{\beta} = \text{vec}(\hat{B})$ where $\hat{B} = (ZZ')^{-1}Zu' = \begin{bmatrix} \hat{B}_1 & \hat{B}_2 & \cdots & \hat{B}_p \end{bmatrix}$

$$z_t = \hat{B}_1 z_{t-1} + \ldots + \hat{B}_p z_{t-p} + \hat{u}_t$$

The estimates of the structural impulse responses to $\epsilon_{j,t}$ are

$$\begin{aligned}
h = 0 &: \hat{M}_0^j = \mathcal{D}_j \\
h = 1 &: \hat{M}_1^j = \hat{B}_1 \hat{M}_0^j \\
h = 2 &: \hat{M}_2^j = \hat{B}_1 \hat{M}_1^j + \hat{B}_2 \hat{M}_0^j \\
h = 3 &: \hat{M}_3^j = \hat{B}_1 \hat{M}_2^j + \hat{B}_2 \hat{M}_1^j + \hat{B}_3 \hat{M}_0^j \\
&\vdots \\
\text{any } h > 0 &: \hat{M}_h = \sum_{i=1}^{h} \hat{B}_i \hat{M}_{h-i}^j = \hat{G}_h M_0^j = \sum_{i=1}^{h} \hat{G}_{h-i} \hat{B}_i M_0^j
\end{aligned}$$

What Can Go Wrong? Sources of Misspecification

We will discuss three potential sources of misspecification

1. Nonfundamentalness
2. Nonlinearities
3. Time Aggregation

Nonfundamentalness

Consider the SMA representation $z_t = M(L)\epsilon_t$ with $\dim(z_t) = \dim(\epsilon_t)$

Nonfundamentalness and Noninvertibility

If $det(M(\lambda)) = 0$ for at least one $|\lambda| < 1$, ϵ_t is **nonfundamental** for z_t

If $det(M(\lambda)) = 0$ for at least one $|\lambda| = 1$, the MA process is **noninvertible**

If ϵ_t is nonfundamental for z_t, but $det(M(\lambda))$ has no zeroes on the unit circle, $M(L)$ is still invertible, just not invertible in the past only.

There is an $F(L) = \sum_{i=-\infty}^{\infty} F_i L^i$ with $F(L)M(L) = \mathcal{I}$ such that $F(L)z_t = \epsilon_t$

ϵ_t can only be recovered from all past, current, and *future* observations of z_t.

z_t therefore does not admit an SVAR representation

If only $\mathcal{D}_j$ is known, fundamentalness is also required for estimating structural IRFs with local projections

Note: it is common to use 'nonfundamentalness' and 'noninvertibility' interchangeably

Nonfundamentalness and Omitted Variables

One-step ahead forecast errors by economic agents in theoretical models are

$$z_t - E[z_t|s_{t-1}] = z_t - E[z_t|\epsilon_{t-1}, \epsilon_{t-2}, ...]$$

Fundamentalness implies that

$$z_t - E[z_t|\epsilon_{t-1}, \epsilon_{t-2}, ...] = z_t - E[z_t|z_{t-1}, z_{t-2}, ...]$$

Observing the history of ϵ_t is equivalent to observing the history of z_t

If z_t is the data available to the econometrician, than nonfundamentalness

$$z_t - E[z_t|\epsilon_{t-1}, \epsilon_{t-2}, ...] \neq z_t - E[z_t|z_{t-1}, z_{t-2}, ...]$$

means that the observables z_t available to the econometrician are **informationally insufficient**

Nonfundamentalness is mostly an **omitted variables problem**

Trivially, $\epsilon_{j,t}$ is fundamental for $\tilde{z}_t = [\epsilon_{j,t}\ z_t]$.

But $dim(z_t) = dim(\epsilon_t)$ does not guarantee fundamentalness

Example: News Shocks in the New Keynesian Model

Consider again the simple NK model

$$\begin{bmatrix} gap_t \\ \pi_t \end{bmatrix} = E_t \sum_{i=0}^{\infty} \mathcal{C}^{-(i+1)} \begin{bmatrix} s^s_{t+i} \\ s^d_{t+i} \end{bmatrix}$$

where $\kappa > 0$, $\phi_\pi > 1$, $0 \leq \beta < 1$ and $\mathcal{C}^{-1} = \frac{1}{1+\phi_\pi \kappa} \begin{bmatrix} 1 & 1 - \beta\phi_\pi \\ \kappa & \beta + \kappa \end{bmatrix}$

If the shocks follow a VAR(1) process $\begin{bmatrix} s^s_t \\ s^d_t \end{bmatrix} = \Lambda \begin{bmatrix} s^s_{t-1} \\ s^d_{t-1} \end{bmatrix} + \Sigma_\epsilon \epsilon_t$,

there is a VAR(1) representation for gap_t and π_t

$$\begin{bmatrix} gap_t \\ \pi_t \end{bmatrix} = (\mathcal{C} - \Lambda)^{-1} \Lambda (\mathcal{C} - \Lambda) \begin{bmatrix} gap_{t-1} \\ \pi_{t-1} \end{bmatrix} + (\mathcal{C} - \Lambda)^{-1} \Sigma_\epsilon \epsilon_t$$

The assumed shock process delivers fundamentalness

Example: News Shocks in the New Keynesian Model

But now suppose $s_t^d = \epsilon_t^d$ is iid white noise and

$$s_t^s = \epsilon_{t-1}^s \text{ , } \epsilon_t^s \text{ is iid}$$

The solution is :

$$\begin{aligned} \begin{bmatrix} gap_t \\ \pi_t \end{bmatrix} &= \left(\mathcal{C}^{-1} \begin{bmatrix} L & 0 \\ 0 & 1 \end{bmatrix} + \mathcal{C}^{-2} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) \begin{bmatrix} \epsilon_t^s \\ \epsilon_t^d \end{bmatrix} \\ &= \mathcal{C}^{-1} \begin{bmatrix} L + \frac{1}{1+\phi_\pi \kappa} & \frac{\kappa}{1+\phi_\pi \kappa} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_t^s \\ \epsilon_t^d \end{bmatrix} \end{aligned}$$

The MA term looses rank at $L = -\frac{1}{1+\phi_\pi \kappa}$, which is inside the unit circle.

Hence there is no SVAR representation for $[gap_t, \pi_t]$ for this shock process.

Nonfundamentalness

Nonfundamentalness is **not** a theoretical curiosity.

Many examples in the literature:

- News Shocks — Leeper, Walker, and Yang (2013)
- Noninvertible MA shock processes — Lippi and Reichlin (1993)
- Multistep prediction errors as observables — Hansen and Hodrick (1980)
- Heterogeneous Information — Blanchard, L'Huillier, and Lorenzoni (2013), Chahrour and Jurado (2021)
- etc.

Nonfundamentalness

So if ϵ_t is nonfundamental for z_t, what is the residual u_t in the VAR(∞) projection $z_t = \sum_{i=1}^{\infty} B_i z_{t-i} + u_t$?

Wold Representation Theorem

If z_t is covariance stationary, there is an MA(∞) represention $z_t = G(L)u_t$ where the Wold innovation process u_t is white noise

$$E\left[u_t\right] = 0\ ,\ E\left[u_t u_t'\right] = \Sigma_u\ ,\ E\left[u_t u_s'\right] = 0 \text{ for } s \neq t$$

Note that u_t is fundamental for z_t

But, we cannot rotate $G(L)$ by $\mathcal{D}$ and obtain the structural IRFs, since $u_t \neq \mathcal{D}\epsilon_t$

Connection between SMA and Wold Representation

$z_t = M(L)\epsilon_t = M(L)\mathbb{B}(L)^{-1}\mathcal{D}^{-1}\mathcal{D}\mathbb{B}(L)\epsilon_t = G(L)u_t$

$G(L)\mathcal{D}\mathbb{B}(L) = M(L)$ and $u_t = \mathcal{D}\mathbb{B}(L)\epsilon_t$ where $\mathbb{B}(L)$ is a **Blaschke matrix**

The Blaschke matrix $\mathbb{B}(L)$ ‘flips’ the roots of $G(L)\mathcal{D}$ to obtain the SMA

Nonfundamentalness

Blaschke Matrix Lippi and Reichlin (1994)

All (rational) Blaschke matrices have the form

$$\mathbb{B}(L) = R(\lambda_1, L)K_1 R(\lambda_2, L)K_2 \ldots R(\lambda_r, L)K_r$$

where r is an integer, K_i is an orthogonal matrix (i.e. $K_i K_i' = \mathcal{I}$) and

$$R(\lambda_i, L) = \begin{bmatrix} \frac{L-\lambda_i}{1-\lambda_i^* L} & 0 \\ 0 & \mathcal{I} \end{bmatrix}$$

where λ_i^* is the complex conjugate of λ_i

If ϵ_t is iid white noise than $u_t = \mathcal{D}\mathbb{B}(L)\epsilon_t$ is white noise, but u_t is generally not iid unless ϵ_t is Gaussian

Let λ_i^{-1} be the roots of $G(L)$, where we know that $|\lambda_i| < 1$ since u_t is fundamental for z_t.

To construct the true nonfundamental SMA representation we need to know the K_i's and which roots of $G(L)$ to flip.

To recover ϵ_t, we need to know not only $\mathcal{D}$ but also the right Blaschke matrix.

Example: Government Spending News Shocks in RBC Model

News shocks are typically nonfundamental for 'conventional' z_t Leeper, Walker, and Yang (2013)

Example from Mertens and Ravn (2010) on anticipated government spending shocks

$$\mathcal{U} = E_0 \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma} \left(1 - \psi N_t^{\theta}\right)^{1-\sigma} - 1}{1-\sigma} \qquad \text{(preferences)}$$

$$C_t + K_{t+1} + G_t = (1-\delta) K_t + K_t^{\alpha} (X_t N_t)^{1-\alpha} \qquad \text{(budget constraint)}$$

Exogenous processes:

$$\mu_g(L) \ln (G_t) = \mu_g(1) \ln(\bar{g}) + \sigma_g \epsilon_{0,t}^g + \sigma_g \lambda \epsilon_{q,t-q}^g \qquad \text{(govt. spending)}$$

$\epsilon_{q,t-q}$: Anticipated shocks enter the information set at date $t-q$ but affect government spending only in period t

λ parametrizes the variance contribution of anticipated shocks

Example: Government Spending News Shocks in RBC Model

The solution to a log-linearized approximation of the model in terms of detrended variables can be formulated as

$$
\begin{aligned}
k_{t+1} &= \phi_{kk} k_t + \phi_{kg} g_t + \sum_{i=0}^{q-1} \phi_{k,q-i} \epsilon_{q,t-i} \quad , \\
c_t &= \phi_{ck} k_t + \phi_{cg} g_t + \sum_{i=0}^{q-1} \phi_{c,q-i} \epsilon_{q,t-i} \quad ,
\end{aligned}
$$

The coefficients relating to the impact of anticipated government spending shocks can be expressed as

$$
\begin{aligned}
\phi_{c,q-i} &= \omega^{q-1-i} \phi_{c,1} \quad \text{for } i = 0, \ldots, q-1 \quad , \\
\phi_{k,q-i} &= \omega^{q-1-i} \phi_{k,1} \quad \text{for } i = 0, \ldots, q-1 \quad .
\end{aligned}
$$

where ω is the inverse of the unstable eigenvalue of the RE system.

$\mid \omega \mid < 1$ is the **anticipation rate** Ljungqvist and Sargent (2012))

Constant discounting of news

Example: Government Spending News Shocks in RBC Model

Vector of observables $z_t = [g_t \;\; c_t]'$ has SMA representation

$$z_t = \Upsilon(L)\Sigma_\epsilon \epsilon_t \quad , \mathcal{D} = \Upsilon(0)\Sigma_\epsilon \quad , \quad \epsilon_t = \begin{bmatrix} \epsilon^g_{0,t} \\ \epsilon^g_{q,t} \end{bmatrix} \quad \Sigma_\epsilon = \sigma_g \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix}$$

Fundamentalness requires that all roots of $det(\Upsilon(L))$ are outside the unit circle

$$det(\Upsilon(L)) = \left(\frac{\phi_{ck}\phi_{k,1}}{1 - \phi_{kk}L} L + \phi_{c,1} \right) \mu_g(L)\Theta(L)$$

$\Upsilon(L)$ inherits the roots of $\Theta(L) = \omega^{q-1} + \omega^{q-2}L + \cdots + \omega L^{q-2} + L^{q-1}$

For $q > 1$ (but not $q = 1$!), roots lie on a circle in the complex plane with radius $\omega < 1$

Example: Government Spending News Shocks in RBC Model

The Blaschke matrix that recovers the shocks from the Wold innovations in $\epsilon_t = \mathbb{B}(L)^{-1}\mathcal{D}^{-1}u_t$ is

$$
\begin{aligned}
\mathbb{B}(L)^{-1} &= KR(\omega_1, L)R(\omega_2, L)\ldots R(\omega_{q-1}, L) \quad , \\
R(\omega_i, L) &= \begin{bmatrix} 1 & 0 \\ 0 & \frac{1-\omega_i L}{L-\omega_i^*} \end{bmatrix} \quad , \\
K &= \frac{1}{\sqrt{1+(\lambda\omega^q)^2}} \begin{bmatrix} 1 & -\lambda\omega^q \\ \lambda\omega^q & 1 \end{bmatrix} \quad ,
\end{aligned}
$$

where ω_i, $i = 1, ..., q-1$ are the roots of the polynomial $\Theta(L)$ and ω_i^* denotes the complex conjugate.

Depends only on three additional parameters λ, ω and q.

Partial Invertibility/Fundamentalness

Assuming fundamentalness is assuming that ϵ_t is a linear combination of $u_t = B(L)z_t$

In practice, it is only required that the shock of interest is fundamental for z_t

Partial Invertibility/Fundamentalness

$\epsilon_{j,t}$ is fundamental for z_t if $\epsilon_{j,t}$ is a linear combination of u_t in $z_t = G(L)u_t$

Semi-Structural VAR Representation Stock and Watson (2018), Miranda-Agrippino and Ricco (2019)

Let $\epsilon_{j,t}$ be fundamental for z_t such that $\epsilon_{j,t} = \lambda_j' u_t$. There exists a $\Lambda = [\lambda_j \ \lambda_{-j}]$ where λ_{-j} is $N_z \times (N_z - 1)$ and $\Lambda' \Sigma_u \Lambda = \mathcal{I}$ such that

$$B(L)z_t = \mathcal{D}_j \epsilon_{j,t} + \xi_t \quad , \quad \text{where } \mathcal{D}_j = \Sigma_u \lambda_j, \ \xi_t = \Sigma_u \lambda_{-j} \lambda_{-j}' u_t, \ E[\epsilon_{j,t} \xi_t'] = 0$$

Detecting Nonfundamentalness

Checking nonfundamentalness given an economic model

- Given a DGP $\{\mathcal{G}, \mathcal{F}, \mathcal{A}, \mathcal{D}\}$, fundamentalness of u_t for z_t can be verified by checking the eigenvalue condition for IP Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007).
- A measure of the informational deficiency for $\epsilon_{j,t}$ is the unexplained variation in regression of $\epsilon_{j,t}$ on u_t Sims and Zha (2006), Forni, Gambetti, and Sala (2019)

Testing the null of fundamentalness with additional macroeconomic variables x_t

- Test whether any $x_{j,t} \in x_t$ Granger-causes z_t Giannone and Reichlin (2006)
- Test whether Wold innovations u_{t+j} for $j \geq 1$ predict x_t Canova and Hamidi Sahneh (2017)

Testing the null of fundamentalness under non-Gaussianity

- For iid non-Gaussian ϵ_t, test whether u_t is a martingale difference sequence Chen, Choi, and Escanciano (2017)

Adressing Nonfundamentalness

- Think carefully about selection of variables in z_t, and include variables that are predictive for the main outcome variables of interest

- Measure the shock of interest directly

  Romer and Romer (1989), Romer and Romer (2010), Ramey (2011)

- Large information sets (e.g. Factor models)

  Giannone and Reichlin (2006), Alessi, Barigozzi, and Capasso (2011)

- Impose additional structure (beyond $\mathcal{D}_j$) Mertens and Ravn (2010), Chahrour and Jurado (2021)

What Can Go Wrong? Sources of Misspecification

1. Nonfundamentalness
2. Nonlinearities
3. Time Aggregation

Nonlinearities

Linearity imposes some strong properties on structural impulse responses:

- responses scale linearly with the size of the shocks ϵ_t
- responses are symmetric with respect to sign of the shocks
- responses are independent of t (e.g. state of the economy)

What if the true DGP is nonlinear?

Nonlinearities

Suppose z_t is generated by a nonlinear (causal) process

$$z_t = M^*(\epsilon_t, \epsilon_{t-1}, ...)$$

where the nonlinear function $M^*(\cdot)$ generates the covariance stationary process z_t.

Best Linear Approximation Plagborg-Møller and Wolf (2021)

The nonlinear process $z_t = M^*(\epsilon_t, \epsilon_{t-1}, ...)$ has an SMA(∞) representation

$$z_t = M(L)\epsilon_t + N(L)\zeta_t$$

where $M(L) = M_0 + M_1 L + M_2 L^2 + ...$ and $N(L) = N_0 + N_1 L + N_2 L^2 + ...$, ζ_t is N_z-dimensional white noise with $E\zeta_t \epsilon_s' = 0$ for all s, t and

$$\{M_0, M_1, ...\} = \underset{\tilde{M}_0, \tilde{M}_1, ...}{\operatorname{argmin}} \left\{ E\left[\left(M^*(\epsilon_t, \epsilon_{t-1}, ...) - \sum_{i=0}^{\infty} \tilde{M}_i \epsilon_{t-i} \right)^2 \right] \right\}$$

$[\epsilon_t' \; \zeta_t']'$ is in general nonfundamental for z_t, so the Wold innovations $u_t = B(L) z_t$ are not linear combinations of ϵ_t

Nonlinearities

Conditional Dynamic Causal Effect

In nonlinear models, the conditional dynamic causal effect of a δ intervention in $\epsilon_{j,t} \in \epsilon_t$ on z_{t+h} is

$$E[M^*(\epsilon_{t+h}, \epsilon_{t+h-1}, \ldots)|\epsilon_{j,t} = \delta, \epsilon_{t-1}, \ldots] \\ - E[M^*(\epsilon_{t+h}, \epsilon_{t+h-1}, \ldots)|\epsilon_{t-1}, \ldots] \neq M_h$$

Expectations integrate over $\epsilon_{-j,t}$ and all shocks from $t+1$ to $t+h$

Unconditional Dynamic Causal Effect

In nonlinear models, the unconditional dynamic causal effect of a δ intervention in $\epsilon_{j,t} \in \epsilon_t$ on z_{t+h} is

$$E[M^*(\epsilon_{t+h}, \epsilon_{t+h-1}, \ldots)|\epsilon_{j,t} = \delta] \\ - E[M^*(\epsilon_{t+h}, \epsilon_{t+h-1}, \ldots)] \neq M_h$$

Expectations additionally integrate over all shocks (but $\epsilon_{j,t}$) from $-\infty$ to t

Both objects depend on δ (sign and size of the shock)
The 'best linear approximation' M_h can differ substantially from either object

Nonlinearities

There are various ways to introduce nonlinearities in VAR models

- Include nonlinear transformations of individual series in z_t
- Markov-Switching VARs
- Threshold/Smooth Transition VARs
- Time-Varying Parameter VARs
- Nonparametric VARs

Curse of dimensionality, use with care

What Can Go Wrong? Sources of Misspecification

1. Nonfundamentalness
2. Nonlinearities
3. Time Aggregation

Time Aggregation

What if the true DGP has higher frequency than that of data collection?

Suppose the model for high frequency data is a VAR(p)

$$N(\mathcal{L})z_\tau^* \quad = \quad \mathcal{D}\epsilon_\tau \ , \ \tau = 1, ..., kT$$

where $N(\mathcal{L}) = I_n - N_1\mathcal{L} - ... - N_p\mathcal{L}^p$, $\mathcal{L}$ is the lag operator $\mathcal{L}^j x_\tau = x_{\tau-j}$.

Suppose the econometrician observes **average sampled data**

$$z_t = (\mathcal{I} + \mathcal{L} + ...\mathcal{L}^{k-1})z_{tk}^* \ , \quad t = 1, .., T$$

where t indexes the lower frequency. For concreteness, assume that $p \geq k - 1$.

Can we fit a VAR for z_t and do structural VAR analysis?

Time Aggregation

Generally no.

The time aggregated data has a VARMA representation

$$A(L)z_t \quad = \quad C(L)v_t \ , \ \ t = 1, ..., T$$

where

$$E\left[v_t\right] = 0 \ , \ \ E\left[v_t v_t'\right] = \Sigma_v \ , \ \ E\left[v_t v_s'\right] = 0 \text{ for } s \neq t$$

Some Results on Time Aggregation Marcellino (1999)

- $A(L)$ has order p or less – generally the same order as $N(\mathcal{L})$
- The order of $C(L)$ is bounded by p if $p = k - 1$, or otherwise by $p + 1 + q$ where q is the smallest positive integer that satisfies $qk \leq k - p - 2 < (q + 1)k$
- $v_t \neq \mathcal{D}(\mathcal{I} + \mathcal{L} + ...\mathcal{L}^{k-1})\epsilon_{tk}$. Instead v_t is a linear combination of current and up to $(p + 1)(k - 1)$ lags of shocks.

Similar results hold for **point-in-time sampling**.

Time Aggregation

We cannot expect to uncover high frequency dynamics with low frequency data.

Possible solutions:

- High frequency macro data

  Lewis, Mertens, Stock, and Trivedi (2022), Baumeister, Leiva-León, and Sims (2022), Jacobson, Matthes, and Walker (2022)

- Mixed Frequency models
  - Mixed-Frequency VARs and factor models

    Mariano and Murasawa (2010), Bańbura, Giannone, Modugno, and Reichlin (2013), Schorfheide and Song (2015)
  - MIxed DAta Sampling (MIDAS)-VARs: Ghysels (2016)

Footnote: Local Uniqueness

In matrix form

$$E_t \begin{bmatrix} \hat{y}^{gap}_{t+1} \\ \pi_{t+1} \end{bmatrix} = \mathcal{C} \begin{bmatrix} \hat{y}^{gap}_{t} \\ \pi_{t} \end{bmatrix} - \begin{bmatrix} u_t \\ v_t \end{bmatrix} \quad , \quad \mathcal{C} \equiv \frac{1}{\beta} \begin{bmatrix} \beta + \kappa & \beta\phi_\pi - 1 \\ -\kappa & 1 \end{bmatrix}$$

The companion matrix $\mathcal{C}$ has the characteristic polynomial

$$\begin{aligned} \mathcal{P}(\varphi) &= \varphi^2 - \mathrm{tr}(\mathcal{C})\varphi + \det(\mathcal{C}) \\ \mathrm{tr}(\mathcal{C}) &= 1 + 1/\beta + \kappa/\beta > 1 \\ \det(\mathcal{C}) &= (1 + \kappa\phi_\pi)/\beta > 1 \end{aligned}$$

which has roots outside the unit circle if $\mathrm{tr}(\mathcal{C}) < 1 + \det(\mathcal{C})$ or

$$\phi_\pi > 1 \qquad \text{(Taylor Principle)}$$

Back

Alessi, Lucia, Matteo Barigozzi, and Marco Capasso (2011). "Non-Fundamentalness in Structural Econometric Models: A Review". In: *International Statistical Review / Revue Internationale de Statistique* 79.1, pp. 16–47.

Aoki, Masanao (1987). *State Space Modeling of Time Series*. Universitext. Springer-Verlag.

Baumeister, Christiane, Danilo Leiva-León, and Eric Sims (Jan. 2022). "Tracking Weekly State-Level Economic Conditions". In: *The Review of Economics and Statistics*, pp. 1–45. DOI: 10.1162/rest_a_01171.

Bańbura, Marta, Domenico Giannone, Michele Modugno, and Lucrezia Reichlin (2013). "Chapter 4 - Now-Casting and the Real-Time Data Flow". In: *Handbook of Economic Forecasting*. Ed. by Graham Elliott and Allan Timmermann. Vol. 2. Handbook of Economic Forecasting. Elsevier, pp. 195–237. DOI: https://doi.org/10.1016/B978-0-444-53683-9.00004-9.

Blanchard, Olivier J., Jean-Paul L'Huillier, and Guido Lorenzoni (2013). "News, Noise, and Fluctuations: An Empirical Exploration". In: *American Economic Review* 103.7, pp. 3045–70. DOI: 10.1257/aer.103.7.3045. URL: https://www.aeaweb.org/articles?id=10.1257/aer.103.7.3045.

Canova, Fabio and Mehdi Hamidi Sahneh (2017). "Are Small-Scale SVARs Useful for Business Cycle Analysis? Revisiting Nonfundamentalness". In: *Journal of the European Economic Association* 16.4, pp. 1069–1093. DOI: 10.1093/jeea/jvx032.

Chahrour, Ryan and Kyle Jurado (2021). "Recoverability and Expectations-Driven Fluctuations". In: *The Review of Economic Studies* 89.1, pp. 214–239. DOI: 10.1093/restud/rdab010.

Chen, Bin, Jinho Choi, and Juan Carlos Escanciano (2017). "Testing for fundamental vector moving average representations". In: *Quantitative Economics* 8.1, pp. 149–180. DOI: 10.3982/QE393.

Durbin, J. and S.J. Koopman (2001). *Time series analysis by state space methods*. Oxford; New York: Oxford University Press.

Fernández-Villaverde, Jesús, Juan F. Rubio-Ramírez, Thomas J. Sargent, and Mark W. Watson (2007). "ABCs (and Ds) of Understanding VARs". In: *American Economic Review* 97.3, pp. 1021–1026. DOI: 10.1257/aer.97.3.1021.

Forni, Mario, Luca Gambetti, and Luca Sala (2019). "Structural VARs and noninvertible macroeconomic models". In: *Journal of Applied Econometrics* 34.2, pp. 221–246. DOI: 10.1002/jae.2665.

Galí, Jordi (2015). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications Second edition*. Economics Books. Princeton University Press.

Ghysels, Eric (2016). "Macroeconomics and the reality of mixed frequency data". In: *Journal of Econometrics* 193.2, pp. 294–314. DOI: https://doi.org/10.1016/j.jeconom.2016.04.008.

Giannone, Domenico and Lucrezia Reichlin (2006). "Does Information Help Recovering Structural Shocks from Past Observations?" In: *Journal of the European Economic Association* 4.2/3, pp. 455–465.

Hansen, Lars Peter and Robert J Hodrick (1980). "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis". In: *Journal of Political Economy* 88.5, pp. 829–853. DOI: 10.1086/260910.

Hansen, Lars Peter and Thomas J. Sargent (1980). "Formulating and estimating dynamic linear rational expectations models". In: *Journal of Economic Dynamics and Control* 2, pp. 7–46. DOI: 10.1016/0165-1889(80)90049-4.

— (1991). "Chapter 4: Two Difficulties in Interpreting Structural Vector Autogressions". In: Rational Expectations Econometrics. Westview Press, pp. 77–119.

Jacobson, Margaret M., Christian Matthes, and Todd B. Walker (Aug. 2022). *Inflation Measured Every Day Keeps Adverse Responses Away: Temporal Aggregation and Monetary Policy Transmission*. Finance and Economics Discussion Series 2022-054. Board of Governors of the Federal Reserve System (U.S.) DOI: 10.17016/FEDS.2022.054.

Jordà, Òscar (2005). "Estimation and Inference of Impulse Responses by Local Projections". In: *American Economic Review* 95.1, pp. 161–182. DOI: 10.1257/0002828053828518.

Kascha, Christian and Karel Mertens (2009). "Business cycle analysis and VARMA models". In: *Journal of Economic Dynamics and Control* 33.2, pp. 267–282. DOI: 10.1016/j.jedc.2008.05.006.

Leeper, Eric M., Todd B. Walker, and Shu-Chun Susan Yang (2013). "Fiscal Foresight and Information Flows". In: *Econometrica* 81.3, pp. 1115–1145. DOI: https://doi.org/10.3982/ECTA8337.

Lewis, Daniel J., Karel Mertens, James H. Stock, and Mihir Trivedi (2022). "Measuring real activity using a weekly economic index". In: *Journal of Applied Econometrics* 37.4, pp. 667–687. DOI: https://doi.org/10.1002/jae.2873.

Li, Dake, Mikkel Plagborg-Møller, and Christian K Wolf (2022). *Local Projections vs. VARs: Lessons From Thousands of DGPs*. Working Paper 30207. National Bureau of Economic Research. DOI: 10.3386/w30207.

Lippi, Marco and Lucrezia Reichlin (1993). "The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment". In: *American Economic Review* 83.3, pp. 644–52.

Lippi, Marco and Lucrezia Reichlin (1994). “VAR analysis, nonfundamental representations, blaschke matrices”. In: *Journal of Econometrics* 63.1, pp. 307–325. DOI: https://doi.org/10.1016/0304-4076(93)01570-C.

Ljungqvist, Lars and Thomas J. Sargent (2012). *Recursive Macroeconomic Theory, Third Edition*. The MIT Press.

Lütkepohl, Helmut (2005). *New introduction to multiple time series analysis*. Berlin: Springer. DOI: 10.1007/978-3-540-27752-1.

Marcellino, Massimiliano (1999). “Some Consequences of Temporal Aggregation in Empirical Analysis”. In: *Journal of Business & Economic Statistics* 17.1, pp. 129–136. DOI: 10.2307/1392244.

Mariano, Roberto S. and Yasutomo Murasawa (2010). “A Coincident Index, Common Factors, and Monthly Real GDP*”. In: *Oxford Bulletin of Economics and Statistics* 72.1, pp. 27–46. DOI: https://doi.org/10.1111/j.1468-0084.2009.00567.x.

Mertens, Karel and Morten O. Ravn (2010). “Measuring the Impact of Fiscal Policy in the Face of Anticipation: A Structural VAR Approach”. In: *The Economic Journal* 120.544, pp. 393–413. DOI: 10.1111/j.1468-0297.2010.02361.x.

Miranda-Agrippino, Silvia and Giovanni Ricco (2019). *Identification with external instruments in structural VARs under partial invertibility*. Working Papers 13853. CEPR.

Plagborg-Møller, Mikkel and Christian K. Wolf (2021). “Local Projections and VARs Estimate the Same Impulse Responses”. In: *Econometrica* 89.2, pp. 955–980. DOI: 10.3982/ECTA17813.

Ramey, Valerie A. (2011). “Identifying Government Spending Shocks: It’s all in the Timing”. In: *The Quarterly Journal of Economics* 126.1, pp. 1–50.

Romer, Christina D. and David H. Romer (1989). “Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz”. In: *NBER Macroeconomics Annual* 4, pp. 121–170. DOI: 10.1086/654103.

— (2010). “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks”. In: *American Economic Review* 100.3, pp. 763–801. DOI: 10.1257/aer.100.3.763.

Schorfheide, Frank (2005). “VAR forecasting under misspecification”. In: *Journal of Econometrics* 128.1, pp. 99–136. DOI: 10.1016/j.jeconom.2004.08.009.

Schorfheide, Frank and Dongho Song (2015). "Real-Time Forecasting With a Mixed-Frequency VAR". In: *Journal of Business and Economic Statistics* 33.3, pp. 366–380.

Sims, Christopher A. and Tao Zha (2006). "Does Monetary Policy Generate Recessions?" In: *Macroeconomic Dynamics* 10.2, pp. 231–272.

Smets, Frank and Rafael Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach". In: *American Economic Review* 97.3, pp. 586–606. DOI: 10.1257/aer.97.3.586.

Stock, James H. and Mark W. Watson (2018). "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments". In: *The Economic Journal* 128.610, pp. 917–948. DOI: doi.org/10.1111/ecoj.12593.