# Empirical Time Series Methods for Macroeconomic Analysis

Luca Gambetti[1]

[1]Universitat Autònoma de Barcelona and Barcelona GSE

BMaSS BGSE

# Preliminaries

# Preliminary Definitions and Concepts

- *Lag operator*: the lag operator $L$ is an operator such that $LX_t = X_{t-1}$. Notice that $L^2X_t = X_{t-2}$ and $L^{-1} = X_{t+1}$, if $X_t = c$, $LX_t = Lc = c$.
- *Polynomial in the lag operator*: $\phi(L)$;

  $$\phi(L) = \phi_0 + \phi_1 L + \phi_2 L^2$$

  with $\phi_1, \phi_2$ constant.
- Lag polynomials can also be inverted. For a polynomial $\phi(L)$, we are looking for the values of the coefficients $\alpha_i$ of $\phi(L)^{-1} = \alpha(L) = \alpha_0 + \alpha_1 L + \alpha_2 L^2 + ...$ such that $\phi(L)^{-1}\phi(L) = 1$.

# Preliminary Definitions and Concepts

*Example.* Let $\phi(L) = (1 - \phi L)$ with $|\phi| < 1$. To find the inverse write

$$(1 - \phi L)(\alpha_0 + \alpha_1 L + \alpha_2 L^2 + ...) = 1$$

note that all the coefficients of the non-zero powers of $L$ must be equal to zero. This gives

$$\begin{aligned} \alpha_0 &= 1 & & \\ -\phi + \alpha_1 &= 0 & &\Rightarrow \alpha_1 = \phi \\ -\phi\alpha_1 + \alpha_2 &= 0 & &\Rightarrow \alpha_2 = \phi^2 \\ -\phi\alpha_2 + \alpha_3 &= 0 & &\Rightarrow \alpha_3 = \phi^3 \end{aligned}$$

and so on.

# Preliminary Definitions and Concepts

In general $\alpha_k = \phi^k$, so $(1-\phi L)^{-1} = \sum_{j=0}^{\infty} \phi^j L^j$ provided that $|\phi| < 1$. It is easy to check this because

$$(1-\phi L)(1+\phi L+\phi^2 L^2+...+\phi^k L^k) = 1-\phi^{k+1}L^{k+1}$$

so

$$(1+\phi L+\phi^2 L^2+...+\phi^k L^k) = \frac{1-\phi^{k+1}L^{k+1}}{(1-\phi L)}$$

and $k \to \infty$ $\sum_{j=0}^{k} \phi^j L^j \to \frac{1}{(1-\phi L)}$.

# Preliminary Definitions and Concepts

- *Matrix of polynomials in the lag operator*: $\Phi(L)$ if its elements are polynomial in the lag operator, i.e.

$$\Phi(L) = \Phi_0 L^0 + \Phi_1 L^1 + \Phi_2 L^2 + ... + \Phi_p L^p$$

and by definition of $L$ when applied to vector the $X_t$

$$\Phi(L)x_t = \Phi_0 X_t + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + ... + \Phi_p X_{t-p}$$

Note that $\Phi(0) = \Phi_0$ and $\Phi(1) = \Phi_0 + \Phi_1 + \Phi_2 + ... + \Phi_p$

# Preliminary Definitions and Concepts

*Example*

$$\Phi(L) = \begin{pmatrix} 1 & -0.5L \\ L & 1+L \end{pmatrix} = \Phi_0 + \Phi_1 L$$

where

$$\Phi_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Phi_1 = \begin{pmatrix} 0 & -0.5 \\ 1 & 1 \end{pmatrix}, \quad \Phi_{j>1} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

When applied to a vector $X_t$ we obtain

$$\Phi(L)X_t = \begin{pmatrix} 1 & -0.5L \\ L & 1+L \end{pmatrix} \begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} X_{1t} - 0.5X_{2t-1} \\ X_{1t-1} + X_{2t} + X_{2t-1} \end{pmatrix}$$

# Covariance Stationarity

- Let $Y_t$ be a $n$-dimensional random vector, $Y_t' = [Y_{1t}, ..., Y_{nt}]$. Then $Y_t$ is covariance (weakly) stationary if $E(Y_t) = \mu$, and the *autocovariance* matrix $E(Y_t - \mu)(Y_{t-j} - \mu)' = \Gamma_j$ for all $t, j$, that is are independent of $t$ and both finite.
- Stationarity of each of the components of $Y_t$ does not imply stationarity. of the vector $Y_t$. Stationarity in the vector case requires that the components of the vector are stationary and costationary.
- Although $\gamma_j = \gamma_{-j}$ for a scalar process, the same is not true for a vector process. The correct relation is

$$\Gamma_j' = \Gamma_{-j}$$

# Stationary processes

- *Vector White Noise*: A n-dimensional vector white noise $\epsilon_t' = [\epsilon_{1t}, ..., \epsilon_{nt}] \sim WN(0, \Omega)$ is such if $E(\epsilon_t) = 0$ and $E(\epsilon_t \epsilon_\tau') = \Omega$ ($\Omega$ a symmetric positive definite matrix) if $t = \tau$ and 0 otherwise. If also $\epsilon_t \sim N$ the process is a Gaussian WN.

  Important: A vector whose components are white noise is not necessarily a white noise. Example: let $u_t$ be a scalar white noise and define $\epsilon_t = (u_t, u_{t-1})'$. Then $E(\epsilon_t \epsilon_t') = \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_u^2 \end{pmatrix}$ and $E(\epsilon_t \epsilon_{t-1}') = \begin{pmatrix} 0 & 0 \\ \sigma_u^2 & 0 \end{pmatrix}$.

# Stationary processes

- *VMA(q)*: Given the $n$-dimensional vector White Noise $\epsilon_t$ a vector moving average of order $q$ is defined as

$$Y_t = \mu + \epsilon_t + C_1\epsilon_{t-1} + ... + C_q\epsilon_{t-q}$$

where $C_j$ are $n \times n$ matrices of coefficients and $\mu$ is the mean of $Y_t$.

- *VMA(1)* Let us consider the VMA(1)

$$Y_t = \mu + \epsilon_t + C_1\epsilon_{t-1}$$

with $\epsilon_t \sim WN(0, \Omega)$. The variance of the process is given by

$$\begin{aligned} \Gamma_0 &= E[(Y_t - \mu)(Y_t - \mu)'] \\ &= \Omega + C_1\Omega C_1' \end{aligned}$$

with autocovariances

$$\Gamma_1 = C_1\Omega, \;\; \Gamma_{-1} = \Omega C_1', \;\; \Gamma_j = 0 \;\text{ for }\; |j| > 1$$

# Stationary processes

- *The VMA($\infty$)* : A useful process, as we will see, is the VMA($\infty$)

$$Y_t = \mu + \sum_{j=0}^{\infty} C_j \varepsilon_{t-j} \tag{1}$$

If $\sum_{j=0}^{\infty} |C_{m,n,j}| < \infty$ ( for all $m, n$ elements of $C_j$), then the infinite sequence above generates a well defined (mean square convergent) process.

# Stationary processes

- *Invertibility* A MA(q) process defined by the equation $Y_t = C(L)\varepsilon_t$ is said to be invertible is there exists a sequence of absolutely summable matrices of constants $\{A_j\}_{j=0}^{\infty}$ such that $\sum_{j=0}^{\infty} A_j Y_{t-j} = \varepsilon_t$.
- *Result* A MA process defined by the equation $Y_t = C(L)\varepsilon_t$ is invertible if and only if the determinant of $C(L)$ vanishes only outside the unit circle, i.e. if $det(C(z)) \neq 0$ for all $|z| \leq 1$.

  *Example* Consider the process

$$\begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} 1 & L \\ 0 & \theta - L \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

  $det(C(z)) = \theta - z$ which is zero for $z = \theta$. The process is invertible if and only if $|\theta| > 1$.

# Stationary processes

- *Fundamentalness* The VMA is *fundamental* if and only if the $det(C(z)) \neq 0$ for all $|z| < 1$.

  In the previous example the process is fundamental if and only if $|\theta| \geq 1$. In the case $|\theta| = 1$ the process is fundamental but noninvertible.

  Provided that $|\theta| > 1$ the MA process can be inverted and the shock can be obtained as a combination of present and past values of $Y_t$. In fact

$$\begin{pmatrix} 1 & -\frac{L}{\theta - L} \\ 0 & \frac{1}{\theta - L} \end{pmatrix} \begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix} = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

# Lecture 2: Structural VARs

# Theory

- The study of the dynamic effects of economic shocks is one of the key applications in macroeconometrics.
- For instance: what are the effects of monetary policy shocks, fiscal policy shocks or technology shocks?
- This type of analysis is important to understand the transmission mechanisms and to provide policymakers with information about the consequences of their policies.
- Structural Vector Autoregressions models (SVAR, Sims, 1980 Econometrica) are the most popular models to conduct this type of analysis aimed at investigating the effects of economic shocks.

# Theory

Economics behind SVAR models:

- The economy is driven by exogenous orthogonal structural shocks.
- The shocks are dynamically propagated on the economy through the impulse response functions, coefficients which are the outcome of agents decisions.
- The economy is the result of these exogenous shocks plus the response of economic agents.

# Theory

More formally

- Vector of stationary economic variables: $Y_t$.
- Structural economic shocks $u_t$ are white noise:

$$\begin{aligned} E(u_{t-i}u'_{t-j}) &= I \text{ for } i = j \\ E(u_{t-i}u'_{t-j}) &= 0 \text{ for } i \neq j \end{aligned}$$

- Impulse response functions:

$$B(L) = B_0 + B_1 L + B_2 L^2 + ...$$

where $B_j$ are matrix of coefficients and $L$ is the lag operator.

- So the economy can be written as

$$Y_t = B(L)u_t \tag{2}$$

# Theory

Key question: how can the structural shocks $u_t$ be obtained?

# Theory

If (2) is invertible then the economy has a VAR representation

$$A_0 Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + ... + u_t.$$

Normalizing the contemporaneous effects

$$\begin{aligned} Y_t &= A_0^{-1} A_1 Y_{t-1} + A_0^{-1} A_2 Y_{t-2} + ... + A_0^{-1} u_t \\ &= D_1 Y_{t-1} + D_2 Y_{t-2} + ... + A_0^{-1} u_t \\ &= D_1 Y_{t-1} + D_2 Y_{t-2} + ... + \varepsilon_t \end{aligned} \tag{3}$$

where $\varepsilon_t \sim WN(0, \Sigma)$ with $\Sigma = A_0^{-1} A_0^{-1\prime}$.

In compact notation

$$D(L) Y_t = \varepsilon_t$$

where

$$D(L) = I - D_1 L - D_2 L^2 - ... = A_0^{-1} A(L) = A_0^{-1} B(L)^{-1}.$$

# Theory

- Equation (3) is the reduced-form VAR representation of the economy.
- The reduced form MA representation of the economy is

$$Y_t = C(L)\varepsilon_t$$

  where

$$C(L) = D(L)^{-1} = B(L)A_0.$$

- The vector $\varepsilon_t = A_0^{-1}u_t$ is the innovation.

# Theory

SVAR main idea:

1. Obtain an estimate of the reduced form representation $C(L)$.

2. Fix the matrix $A_0$ (identification) and get the structural impulse response functions $B(L) = C(L)A_0^{-1}$ and the structural shocks $u_t = A_0\varepsilon_t$.

# Steps of SVAR analysis

Four main steps:

1. Specification.
2. Estimation.
3. Impulse response functions.
4. Inference.

# VAR Specification

Two issues:

1. Econometrician information set: deciding which variables use to estimate the VAR; $Y_t$ is the whole economy while $Z_t$ is the variables the econometrician uses.

2. Number of lags (theoretically VAR are of infinite order). Statistical criteria:
   - AIC: Akaike information criterion. Choosing the $p$ that minimizes the following
   $$AIC(p) = \ln|\hat{\Omega}| + 2(n^2 p)/T$$
   - BIC: Bayesian information criterion. Choosing the $p$ that minimizes the following
   $$BIC(p) = \ln|\hat{\Omega}| + (n^2 p)\ln T/T$$
   - HQ: Hannan- Quinn information criterion Choosing the $p$ that minimizes the following
   $$HQ(p) = \ln|\hat{\Omega}| + 2(n^2 p)\ln\ln T/T$$

# Estimation

Let $Z_t$ the vector of variable the econometrician uses. The VAR($p$) model is then

$$Z_t = D_1 Z_{t-1} + ... + D_p Z_{t-p} + \varepsilon_t \tag{4}$$

Rewrite the model as

$$Z = XD + \varepsilon$$

where

$$Z = \begin{pmatrix} Z'_{p+1} \\ \vdots \\ Z'_t \\ \vdots \\ Z'_T \end{pmatrix} \quad X = \begin{pmatrix} Z'_p & \dots & Z'_1 \\ \vdots & \vdots & \vdots \\ Z'_{t-1} & \dots & Z'_{t-p} \\ \vdots & \vdots & \vdots \\ Z'_{T-1} & \dots & Z'_{T-p} \end{pmatrix} \quad \varepsilon = \begin{pmatrix} \varepsilon'_{p+1} \\ \vdots \\ \varepsilon'_t \\ \vdots \\ \varepsilon'_T \end{pmatrix} \quad D = \begin{pmatrix} D'_1 \\ \vdots \\ D'_k \\ \vdots \\ D'_p \end{pmatrix}$$

# Estimation

A consistent estimator the matrix of VAR parameters $D$ is the OLS equation by equation. Formally

$$\hat{D} = (X'X)^{-1}X'Z$$

The elements on the $j$th column of $\hat{D}$ are the OLS estimates of the parameters in the $j$th VAR equation.

# Impulse response functions

The structural impulse response functions are obtained in two steps.

1. Derive the reduced-for IRF

2. Fix $A_0^{-1}$.

# Reduced-form (Wold) impulse response functions

To derive the reduced-for IRF, notice that we can rewrite (4) as a VAR(1) as follows.

Define $\mathbf{e}_t' = [\varepsilon', 0, ..., 0]$, $\mathbf{Z}_t' = [Z_t', Z_{t-1}', ..., Z_{t-p+1}']$

$$\mathbf{D} = \begin{pmatrix} D_1 & D_2 & ... & D_{p-1} & D_p \\ I_n & 0 & ... & 0 & 0 \\ 0 & I_n & ... & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & ... & ... & I_n & 0 \end{pmatrix}$$

Therefore we can rewrite the VAR(p) as a VAR(1)

$$\mathbf{Z}_t = \mathbf{D}\mathbf{Z}_{t-1} + \mathbf{e}_t$$

This is known as the companion form of the VAR(p)

# Reduced-form impulse response functions

*Example:* Suppose $n = 2$ and $p = 2$. The VAR will be

$$\begin{pmatrix} z_{1t} \\ z_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} D^1_{11} & D^1_{12} \\ D^1_{21} & dD^1_{22} \end{pmatrix} \begin{pmatrix} z_{1t-1} \\ z_{2t-1} \end{pmatrix} + \begin{pmatrix} D^2_{11} & D^2_{12} \\ D^2_{21} & D^2_{22} \end{pmatrix} \begin{pmatrix} z_{1t-2} \\ z_{2t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

We can rewrite the above model as

$$\begin{pmatrix} z_{1t} \\ z_{2t} \\ z_{1t-1} \\ z_{2t-1} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} D^1_{11} & D^1_{12} & D^2_{11} & D^2_{12} \\ D^1_{21} & D^1_{22} & D^2_{21} & D^2_{22} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} z_{1t-1} \\ z_{2t-1} \\ z_{1t-2} \\ z_{2t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ 0 \\ 0 \end{pmatrix}$$

Setting

$$\mathbf{Z}_t = \begin{pmatrix} z_{1t} \\ z_{2t} \\ z_{1t-1} \\ z_{2t-1} \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} D^1_{11} & D^1_{12} & D^2_{11} & D^2_{12} \\ D^1_{21} & D^1_{22} & D^2_{21} & D^2_{22} \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{e}_t = \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ 0 \\ 0 \end{pmatrix}, \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ 0 \\ 0 \end{pmatrix}$$

we obtain the previous VAR(1) representation.

# Reduced-form impulse response functions

Substituting backward one period in the companion form we obtain

$$\begin{aligned} \mathbf{Z}_t &= \mathbf{D}\mathbf{Z}_{t-1} + \mathbf{e}_t \\ &= \mathbf{D}^2\mathbf{Z}_{t-2} + \mathbf{D}\mathbf{e}_{t-1} + \mathbf{e}_t \end{aligned} \tag{5}$$

Substituting infinite many periods, if the eigenvalues of $\mathbf{D}$ are less than one in absolute value we find

$$\mathbf{Z}_t = \mathbf{e}_t + \mathbf{D}\mathbf{e}_{t-1} + \mathbf{D}^2\mathbf{e}_{t-2} + ...$$

The first upper-left $(n \times n)$ matrices of $\mathbf{D}^j$ represent the effects of $\varepsilon_t$ on $Z_{t+j}$ and are the impulse response functions of $\varepsilon_t$ (the coefficients of the reduced-for matrix $C(L)$).

Replacing the parameters with their OLS estimates one can obtain the estimates of the impulse response functions.

# Reduced-form impulse response functions: Example

Consider the simple VAR(1)

$$Z_t = DZ_{t-1} + \varepsilon_t$$

Substituting backward we find (and the eigenvalues condition holds)

$$Z_t = \varepsilon_t + D\varepsilon_{t-1} + D^2\varepsilon_{t-2} + ...$$

$D^j$ represents the effects of $\varepsilon_t$ on $Z_{t+j}$. So the coefficient matrices of $C(L) = I + C_1L + C_2L^2 + ...$ are $C_j = D^j$.

# Structural Impulse Response Functions

Once the estimated Wold coefficients are available the structural IRF, i.e. the response of economic variable to economic shocks, are obtained as

$$B(L) = C(L)A_0^{-1}$$

Problem: what is $A_0^{-1}$?

# Identification

Identification entails the choice of $A_0^{-1}$ based on the economic theory.

Commonly used identifications:

1. Short-run recursive.
2. Long-run recursive.
3. Sign restrictions.
4. Other.

# Short-run recursive

- The idea is to have a model where the first shock is the only shock affecting the first variable contemporaneously, the first and the second shock the only two affecting the second variable and so on.
- Implementation is easy using the Cholesky factor of $\Sigma$, the covariance matrix of $\varepsilon$.
- Definition: the Cholesky factor of $S$ is the unique lower triangular matrix such that $SS' = \Sigma$
- In this case $A_0^{-1} = S$.
- Remark: $A_0$ will also be lower triangular which implies a recursive scheme also among variables in $Z_t$.

# Long-run recursive

- Introduced by Blanchard and Quah (1989) the idea is to have a model where the first shock is the only shock driving the first variable in the long-run, the first and the second the only driving the second variable and so on.
- Implementation is easy using the Cholesky factor of $C(1)\Sigma C(1)$ where recall $C(1) = I + C_1 + C_2 + ...$
- Let $K$ be the the Cholesky factor of $C(1)\Sigma C(1)$.
- In this case $A_0^{-1} = C(1)^{-1}K$

# Sign restrictions

- Introduced by Uhlig (2004) the idea is to impose qualitative restrictions (in terms on signs) on the effects of structural shocks on the economic variables.
- Implementation as follows. Let $S$ be the Cholesky factor of $\Sigma$.
- Let $H$ be an orthogonal matrix $H$ ($HH' = I$).
- Compute $B_0 = A_0^{-1} = SH$. If the impulse response functions satisfy the restrictions keep the $H$.
- There is no point estimate but rather a set of estimates (a set of $H$) satisfying the restrictions.
- If you search for a single shock you can simply draw one column of the matrix $H$.

# Sign restrictions: drawing $H$

- Using the parametrization

$$H = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}$$

  where $\theta$ is uniform in $[0\ 2\pi]$.

- Using the QR decomposition: take a $n \times n$ matrix $M$ whose entries are $N(0, 1)$. Apply the QR decomposition $M = QR$, then set $H = Q$.

- For a single vector h: either use QR or. Take a $n \times 1$ vector $V$ whose entries are $N(0, 1)$. Then set $h = V/||V||$.

# Maximization

- Suppose you want to identify a shock by maximizing the effects on some variable $j$ at some horizon $k$. Let $C_k h$ be the effect at horizon $k$ of the shock of interest on variable $j$. The elements of $h$ can be obtained as
$$\max_{h_1, h_2, \ldots, h_n} C_j h$$
- If there is more than one variable then the elements of $h$ can be obtained as
$$\max_{h_1, h_2, \ldots, h_n} \sum_{j \in J} C_j h \sigma_{y_j}^{-1}$$
where $J$ is the set of variables whose effects are maximized and $\sigma_{y_j}$ is the standard deviation of the series $j$ (this makes the effects across variables comparable).

# Narrative approach

- Suppose you have an exogenous series $e_t$ representing the shock of interest. The IRF $\phi_k$, $k = 1, ..., K$ can be obtained in three ways:
  1. From the $K$ regressions

     $$Y_{i,t+k} = \alpha + \phi_{i,k} e_t + \eta_t$$

  2. Using a VARX

     $$D(L)Y_t \quad = \quad G(L)e_t + \varepsilon_t \tag{6}$$

     where the IRF are $D(L)^{-1}G(L)$
  3. Augmenting VAR with $e_t$ and then using a Cholesky identification with $e_t$ ordered first and where the first shock is the shock of interest.
- If the shock is exogenous and White Noise the last two identifications should provide the same IRF. This last identification can be useful if the shock is not WN.

# Stock and Watson approach

- Suppose $e_t$ is not exactly the shock (reasonable) but is close to the shock and satisfies the following assumption:

$$E(e_t) = 0, \; E(e_t^2) = 1$$

$$E(e_t u_{it}) = b$$

$$E(e_t u_{-it}) = 0$$

  where $u_{it}$ is the shock you want to identify and $u_{-it}$ is the vector of the remaining structural shocks.

- Recall that $\varepsilon_t = A_0^{-1} u_t$. Then

$$E(\varepsilon_t e_t) = b\alpha$$

  where $\alpha$ is the first column of $A_0^{-1}$.

# Stock and Watson approach

- The IRF can be computed as:
  1. Regress $\varepsilon_t$ on $e_t$ and obtain $b\alpha$
  2. Obtain the rescaled IRF $b\beta(L) = bC(L)\alpha$ where $\beta(L)$ refers to the column of $B(L)$ associated to $u_{it}$
  3. Normalize the impact effect of one variable, say variable $m$, and obtain $\beta(L)$ as $\beta(L) = b\beta(L)/(b\beta_m(0))$.

# Variance decomposition

- Decompose the total variance of a time series into the percentages attributable to each structural shock.
- Variance decomposition analysis is useful in order to address questions like "What are the sources of the business cycle?" or "Is the shock important for economic fluctuations?".

# Variance decomposition

Consider the MA representation of an identified SVAR

$$Z_t = B(L)u_t$$

The variance of $Y_{it}$ is given by

$$\begin{aligned} \mathrm{var}(Z_{it}) &= \sum_{k=1}^{n}\sum_{j=0}^{\infty} B_{ik}^{j2}\mathrm{var}(u_{kt}) \\ &= \sum_{k=1}^{n}\sum_{j=0}^{\infty} B_{ik}^{j2} \end{aligned}$$

where $\sum_{j=0}^{\infty} B_{ik}^{j2}$ is the variance of $Y_{it}$ generated by the $k$th shock. This implies that

$$\frac{\sum_{j=0}^{\infty} B_{ik}^{j2}}{\sum_{k=1}^{n}\sum_{j=0}^{\infty} B_{ik}^{j2}}$$

is the percentage of variance of $Y_{it}$ explained by the $k$th shock if $Z_t$ is stationary.

# Variance decomposition

It is also possible to study the of the series explained by the shock at different horizons, i.e. short vs. long run. Consider the forecast error in terms of structural shocks.

Best forecast $t+h$ periods ahead given the information at time $t$

$$\mathbf{Z}_{t+h} = \mathbf{D}^h \mathbf{Z}_t$$

# Variance decomposition

The forecast error is given by

$$Z_{t+h} - Z_{t+h|t} = B_0 u_{t+h} + B_1 u_{t+h-1} + ... + B_{h-1} u_{t+1}$$

the variance of the forecast error is thus

$$\begin{aligned} \text{var}(Z_{t+h} - Z_{t+h|t}) &= \sum_{k=1}^{n} \sum_{j=0}^{h-1} B_{ik}^{j2} \text{var}(u_{kt}) \\ &= \sum_{k=1}^{n} \sum_{j=0}^{h-1} B_{ik}^{j2} \end{aligned}$$

Thus the percentage of variance of $Z_{it}$ explained by the $k$th shock is

$$\frac{\sum_{j=0}^{h-1} B_{ik}^{j2}}{\sum_{k=1}^{n} \sum_{j=0}^{h-1} B_{ik}^{j2}}$$

# Inference with bootstrap

Main idea (Runkle, 1987): to obtain estimates of the small sample distribution for the impulse response functions without assuming that the shocks are Gaussian.

# Inference with bootstrap

Steps:

1. Estimate the VAR and save the OLS fitted residuals $\{\hat{\varepsilon}_1, \hat{\varepsilon}_2, ..., \hat{\varepsilon}_T\}$.
2. Draw uniformly from $\{\hat{\varepsilon}_1, \hat{\varepsilon}_2, ..., \hat{\varepsilon}_T\}$ and set $\tilde{\varepsilon}_1^{(1)}$ equal to the selected realization and use this to construct

$$Z_1^{(1)} = \hat{D}_1 Z_0 + \hat{D}_2 Z_{-1} + ... + \hat{D}_p Z_{-p+1} + \tilde{\varepsilon}_1^{(1)} \tag{7}$$

3. Taking a second draw (with replacement) $\tilde{\varepsilon}_2^{(1)}$ generate

$$Z_2^{(1)} = \hat{D}_1 Z_1^{(1)} + \hat{D}_2 Z_0 + ... + \hat{D}_p Z_{-p+2} + \tilde{\varepsilon}_2^{(1)} \tag{8}$$

4. Proceeding in this fashion generate a sample of length $T$ $\{Z_1^{(1)}, Z_2^{(1)}, ..., Z_T^{(1)}\}$ and use the sample to compute $\hat{\pi}^{(1)}$ and the implied impulse response functions $C^{(1)}(L)$.
5. For each draw repeat the identification.
6. Repeat steps $(3) - (4)$ $M$ times and collect M realizations of $C^{(l)}(L)$, $l = 1, ...M$ and takes for all the elements of the impulse response functions and for all the horizons the $\alpha$th and $1 - \alpha$th percentile to construct confidence bands.

# Lecture 2: SVAR applications

# Monetary Policy Shocks, Christiano Eichenbaum and Evans, 1998

- Monetary policy shocks is the unexpected part of the equation for the monetary policy instrument ($S_t$).

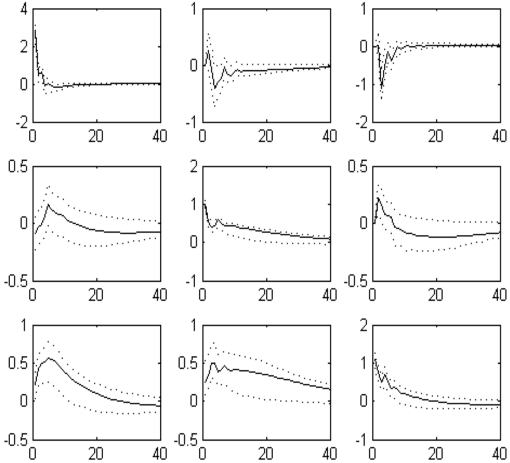
$$S_t = f(\mathcal{I}_t) + w_t^{mp}$$

- $f(\mathcal{I}_t)$ represents the systematic response of the monetary policy to economic conditions, $\mathcal{I}_t$ is the information set at time $t$ and $w_t^{mp}$ is the monetary policy shock.

# Monetary Policy Shocks, Christiano Eichenbaum and Evans, 1998

- The "standard" way to identify monetary policy shock is through zero contemporaneous restrictions. Using the standard trivariate monetary VAR (a simplified version of the CEE 98 VAR) including output growth, inflation and the federal funds rate we identify the monetary policy shock using the following restrictions:
  1. Monetary policy shocks do not affect output within the same quarter
  2. Monetary policy shocks do not affect inflation within the same quarter
  3. Monetary policy responds contemporaneously to inflation and GDP.
- A simple implementation: Cholesky decomposition the federal funds rate is ordered after GDP and inflation. The third column of the impulse response functions is the column of the monetary policy shock.

# Monetary Policy Shocks, Christiano Eichenbaum and Evans, 1998




Cholesky impulse response functions of a system with GDP inflation and the federal funds rate. Monetary shock is in the third column.

# Monetary Policy Shocks, Christiano Eichenbaum and Evans, 1998

- *Price puzzle*: after a contractionary shock inflation goes up. Why is this the case?.
- *"Sims (1992) conjectured that prices appeared to rise after certain measures of a contractionary policy shock because those measures were based on specifications of $\mathcal{I}_t$ that did not include information about future inflation that was available to the Fed. Put differently, the conjecture is that policy shocks which are associated with substantial price puzzles are actually confounded with non-policy disturbances that signal future increases in prices."* (CEE 98)
- Including commodity prices (signaling future inflation increases) may solve the puzzle.

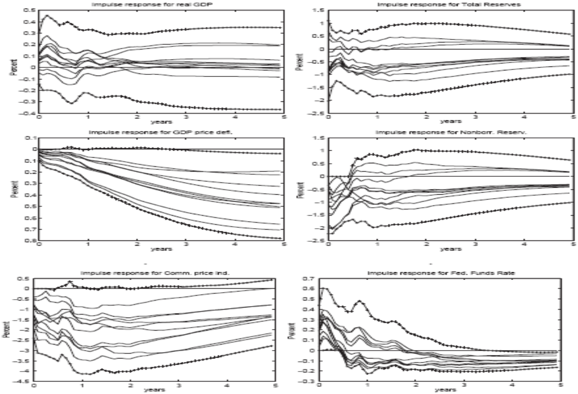
# Uhlig (2005) JME's monetary policy shocks

- Uhlig (2005 JME) proposes a very different method to identify monetary policy shocks. Instead of using zero restrictions as in CEE he uses sign restrictions.
- A contractionary monetary policy shock:
  1. does not increase prices for k periods after the shock
  2. does not increase money or monetary aggregates (i.e. reserves) for k periods after the shock
  3. does not reduce short term interest rate for k periods after the shock.

# Uhlig (2005) JME's monetary policy shocks

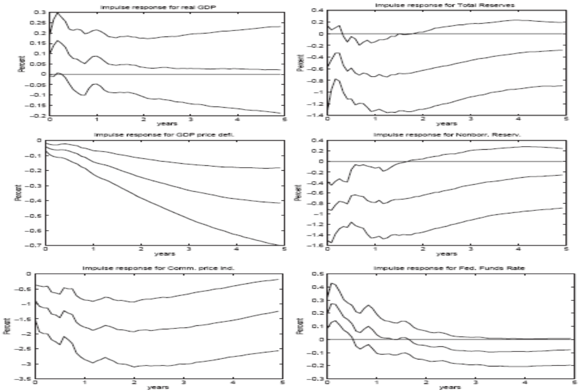
- To draw impulse response functions he applies the following algorithm:
  1. He assumes that the column of $H$, $H_1$, represents the coordinate of a point uniformly distributed over the unit hypersphere (in case of bivariate VAR it represents a point in a circle). To draw such point he draws from a $N(0, I)$ and divide by the norm of the vector.
  2. Compute the impulse response functions $C_j S H_1$ for j=1,..,k.
  3. If the draw satisfies the restrictions keep it and go to 1), otherwise discard it and go to 1). Repeat 1)-3) a big number of timef L.

# Uhlig (2005) JME's monetary policy shocks




Source: What are the effects of a monetary policy shock... JME H. Uhlig (2006)

# Uhlig (2005) JME's monetary policy shocks




Source: What are the effects of a monetary policy shock... JME H. Uhlig (2006)

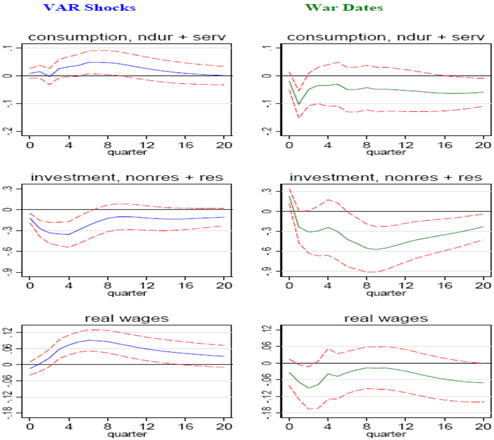
# The effects of government spending shocks

- Understanding the effects of government spending shocks is important for policy authorities but also to assess competing theories of the business cycle.
- Keynesian theory: government spending, GDP, consumption and real wage $\uparrow$ , (because of the government spending multiplier).
- RBC theory: government spending $\uparrow$, but consumption and the real wage $\downarrow$ because of a negative wealth effect.
- Disagreement from the empirical point of view.

# Government spending shocks: Blanchard and Perotti (2002)

- BP (originally) use a VAR for real per capita taxes, government spending, and GDP.
- The shock is identified assuming that government spending does not react to taxes and GDP contemporaneously, Cholesky identification with government spending ordered first. The government spending shock is the first one (quadratic trend four lags).
- When augmented with consumption consumption increases.
- When augmented with investment, investment reduces.
- In a more recent version Perotti (2007) uses a larger VAR but the results are confirmed. Consumption and real wage $\uparrow$ but investment $\downarrow$

# Government spending shocks: Blanchard and Perotti (2002)

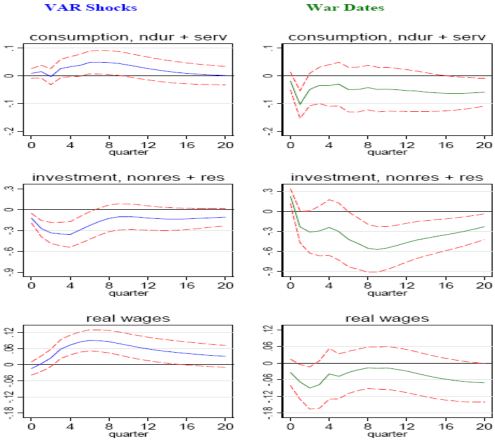



Source: IDENTIFYING GOVERNMENT SPENDING SHOCKS: IT'S ALL IN THE TIMING Valerie A. Ramey NBER Working Paper 15464 (2009)

# Government spending shocks: Ramey and Shapiro (1998)

- Ramey and Shapiro (1998) use a narrative approach to identify shocks to government spending.
- Focus on episodes where Business Week suddenly forecast large rises in defense spending induced by major political events that were unrelated to the state of the U.S. economy (exogenous episodes of government spending).
- Three of such episodes: Korean War, The Vietnam War and the Carter-Reagan Buildup + 9/11.
- The military date variable takes a value of unity in 1950:3, 1965:1, 1980:1, and 2001:3, and zeros elsewhere.
- To identify government spending shocks, the military date variable is embedded in the standard VAR, but ordered before the other variables.

# Government spending shocks: Ramey and Shapiro (1998)




Source: IDENTIFYING GOVERNMENT SPENDING SHOCKS: IT'S ALL IN THE TIMING Valerie A. Ramey NBER Working Paper 15464 (2009)

# Tax multipliers Mertens and Ravn 2014 JME

- Disagreement about the size of the tax multiplier.
- VAR identified a la BP find multiplier which are smaller that those obtained with narrative evidence.
- Mertens and Ravn 2014 JME reconcile the results: key difference BP impose a value for output elasticity of tax revenues which is too small.
- Use a proxy identification of the tax shock.

# Tax multipliers Mertens and Ravn 2014 JME

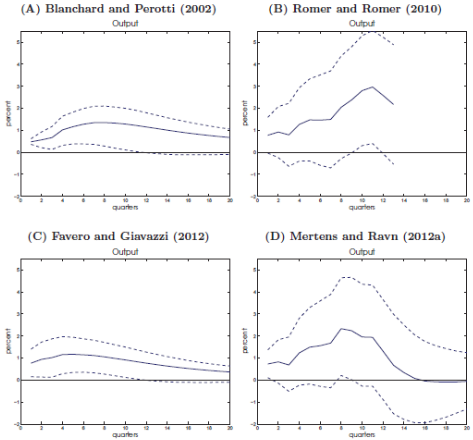



Figure 1: Replication of Existing Estimates of the Output Response to Tax Cuts. Broken lines in (A), (C) and (D) are 95% bootstrapped percentiles. Broken lines in (B) are ± 2 asymptotic standard error bands.

Source: A Reconciliation of SVAR and Narrative Estimates of

# Tax multipliers Mertens and Ravn 2014 JME

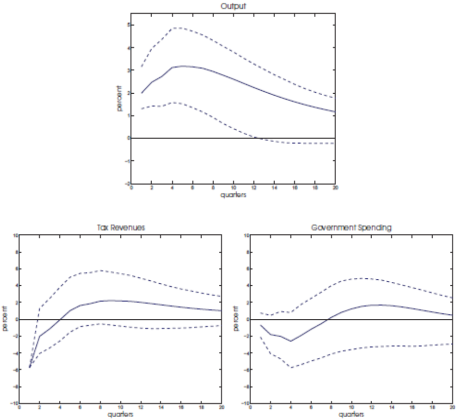
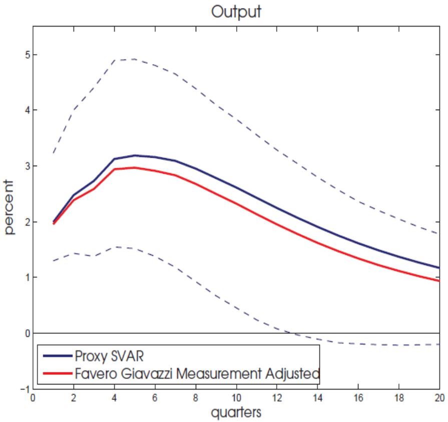



Figure 4: Proxy SVAR: Response to a Tax Cut of 1% of GDP. Broken lines are 95% bootstrapped percentiles.

Source: A Reconciliation of SVAR and Narrative Estimates of

# Tax multipliers Mertens and Ravn 2014 JME




Source: A Reconciliation of SVAR and Narrative Estimates of

# Blanchard Quah (AER 1989) aggregate demand and supply shocks

- Blanchard and Quah proposed an identification scheme based on long run restrictions.
- In their model there are two shocks: an aggregate demand and an aggregate supply disturbance.
- The restriction used to identify is that aggregate demand shocks have no effects on the long run levels of output, i.e. demand shocks are transitory on output.

# Blanchard Quah (AER 1989) aggregate demand and supply shocks

- Let us consider the following bivariate VAR

$$\begin{pmatrix} \Delta Y_t \\ U_t \end{pmatrix} = \begin{pmatrix} B_{11}(L) & B_{12}(L) \\ B_{21}(L) & B_{22}(L) \end{pmatrix} \begin{pmatrix} u_t^s \\ u_t^d \end{pmatrix}$$

  where $Y_t$ is output, $U_t$ is the unemployment rate and $u_t^s, u_t^d$ are two aggregate supply and demand disturbances respectively.

- The identification restriction is given by $B_{12}(1) = 0$.

# Blanchard Quah (AER 1989) aggregate demand and supply shocks

- Implementation

$$\begin{pmatrix} \Delta Y_t \\ U_t \end{pmatrix} = \begin{pmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{pmatrix} \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

  where $E(\epsilon_t \epsilon_t') = \Omega$.

- Let $S = chol(C(1)\Omega C(1)')$ and $K = C(1)^{-1}S$. The identified shocks and IRF are $u_t = K^{-1}\epsilon_t$ and $B(L) = C(L)K$.

# Blanchard Quah (AER 1989) aggregate demand and supply shocks

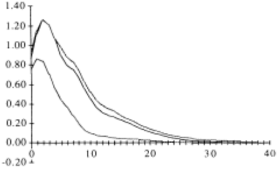



FIGURE 3. OUTPUT RESPONSE TO DEMAND

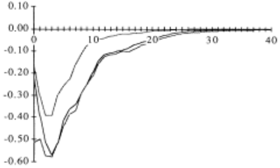



FIGURE 5. UNEMPLOYMENT RESPONSE TO DEMAND

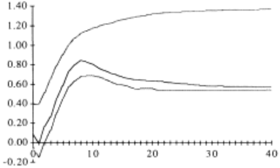



FIGURE 4. OUTPUT RESPONSE TO SUPPLY

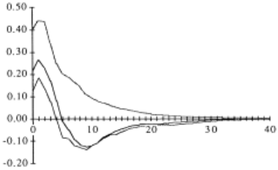



FIGURE 6. UNEMPLOYMENT RESPONSE TO SUPPLY

Source: The Dynamic Effects of Aggregate Demand and Supply Disturbances, (AER) Blanchard and Quah (1989):

# Blanchard Quah (AER 1989) aggregate demand and supply shocks

TABLE 2—VARIANCE DECOMPOSITION OF OUTPUT AND UNEMPLOYMENT
(CHANGE IN OUTPUT GROWTH AT 1973/1974; UNEMPLOYMENT DETRENDED)

| Percentage of Variance Due to Demand: | | |
|---|---|---|
| Horizon (Quarters) | Output | Unemployment |
| 1 | 99.0 | 51.9 |
| | (76.9, 99.7) | (35.8, 77.6) |
| 2 | 99.6 | 63.9 |
| | (78.4, 99.9) | (41.8, 80.3) |
| 3 | 99.0 | 73.8 |
| | (76.0, 99.6) | (46.2, 85.6) |
| 4 | 97.9 | 80.2 |
| | (71.0, 98.9) | (49.7, 89.5) |
| 8 | 81.7 | 87.3 |
| | (46.3, 87.0) | (53.6, 92.9) |
| 12 | 67.6 | 86.2 |
| | (30.9, 73.9) | (52.9, 92.1) |
| 40 | 39.3 | 85.6 |
| | (7.5, 39.3) | (52.6, 91.6) |

Source: The Dynamic Effects of Aggregate Demand and Supply Disturbances, (AER) Blanchard and Quah (1989):

# Blanchard Quah (AER 1989) aggregate demand and supply shocks

TABLE 2A—VARIANCE DECOMPOSITION OF OUTPUT AND UNEMPLOYMENT
(NO DUMMY BREAK, TIME TREND IN UNEMPLOYMENT)

Percentage of Variance Due to Demand:

| Horizon (Quarters) | Output | Unemployment |
|---|---|---|
| 1 | 83.8 | 79.7 |
| | (59.4, 93.9) | (55.3, 92.0) |
| 2 | 87.5 | 88.2 |
| | (62.8, 95.4) | (58.9, 95.2) |
| 3 | 83.4 | 93.5 |
| | (58.8, 93.3) | (61.3, 97.5) |
| 4 | 78.9 | 95.7 |
| | (53.5, 90.0) | (63.9, 98.2) |
| 8 | 52.5 | 88.9 |
| | (31.4, 68.6) | (63.5, 94.5) |
| 12 | 37.8 | 79.7 |
| | (21.3, 51.4) | (58.8, 90.3) |
| 40 | 18.7 | 75.9 |
| | (7.4, 23.5) | (56.9, 88.6) |

Source: The Dynamic Effects of Aggregate Demand and Supply Disturbances, (AER) Blanchard and Quah (1989):

# The technology shocks and hours debate

- This is a nice example of how SVAR models can be used in order to distinguish among competing models of the business cycles.
- RBC technology important source of business cycles.
- Other models (sticky prices) tech shocks not so important.
- Response of hours worked very important in distinguish among theories:
  - RBC hours increase.
  - Other hours fall

# Tech and Hours: The model

Technology shock: $z_t = z_{t-1} + \eta_t$ $\eta_t$ = technology shock

Monetary Policy: $m_t = m_{t-1} + \xi_t + \gamma\eta_t$ where $\xi_t$ = monetary policy shock.

Equilibrium:

$$
\begin{aligned}
\Delta x_t &= \left(1 - \frac{1}{\varphi}\right)\Delta\xi_t + \left(\frac{1-\gamma}{\varphi} + \gamma\right)\eta_t + (1-\gamma)\left(1 - \frac{1}{\varphi}\right)\eta_{t-1} \\
n_t &= \frac{1}{\varphi}\xi_t - \frac{(1-\gamma)}{\varphi}\eta_t
\end{aligned}
$$

or

$$
\begin{pmatrix} \Delta x_t \\ n_t \end{pmatrix} = \begin{pmatrix} \left(\frac{1-\gamma}{\varphi} + \gamma\right) + (1-\gamma)\left(1 - \frac{1}{\varphi}\right)L & \left(1 - \frac{1}{\varphi}\right)(1-L) \\ \frac{-(1-\gamma)}{\varphi} & \frac{1}{\varphi} \end{pmatrix} \begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix} \qquad (9)
$$

# Tech and Hours: The model

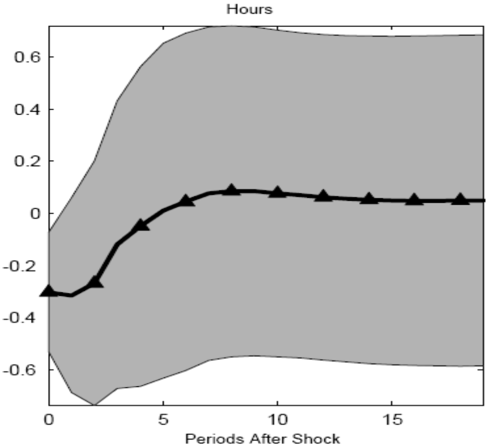
In the long run $L = 1$

$$\begin{pmatrix} \Delta x_t \\ n_t \end{pmatrix} = \begin{pmatrix} \left(\frac{1-\gamma}{\varphi} + \gamma\right) + (1-\gamma)\left(1 - \frac{1}{\varphi}\right) & 0 \\ \frac{-(1-\gamma)}{\varphi} & \frac{1}{\varphi} \end{pmatrix} \begin{pmatrix} \eta_t \\ \xi_t \end{pmatrix}$$

that is only the technology shocks affects labor productivity.

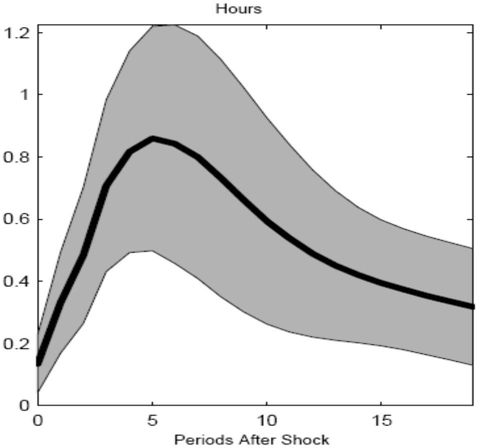
Prediction: if monetary policy is not completely accommodative $\gamma < 1$ then the response of hours to a technology shock $\frac{-(1-\gamma)}{\varphi}$ is negative.

# Tech and Hours: IRF (per-capita hours in growth rates)




Source: What Happens After a Technology Shock?... Christiano Eichenbaum and Vigfusson NBER WK (2003)

# Tech and Hours: IRF (per-capita hours in levels)




Source: What Happens After a Technology Shock?... Christiano Eichenbaum and Vigfusson NBER WK (2003)

# Variance decomposition

Table 1: Contribution of Technology Shocks to Variance, Bivariate System

Level Specification

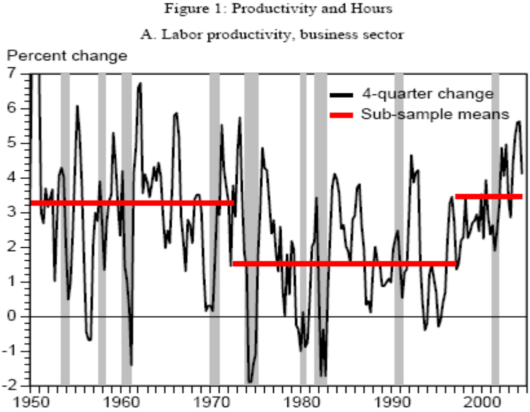
| | Forecast Variance at Indicated Horizon | | | | |
|---|---|---|---|---|---|
| Variable | 1 | 4 | 8 | 12 | 20 | 50 |
| Output | 81.1 | 78.1 | 86.0 | 89.1 | 91.8 | 96 |
| Hours | 4.5 | 23.5 | 40.7 | 45.4 | 47.4 | 48.3 |

Difference Specification

| | Forecast Variance at Indicated Horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 1 | 4 | 8 | 12 | 20 | 50 |
| Output | 16.5 | 11.7 | 17.9 | 20.7 | 22.3 | 23.8 |
| Hours | 21.3 | 6.4 | 2.3 | 1.6 | 1.0 | 0.5 |

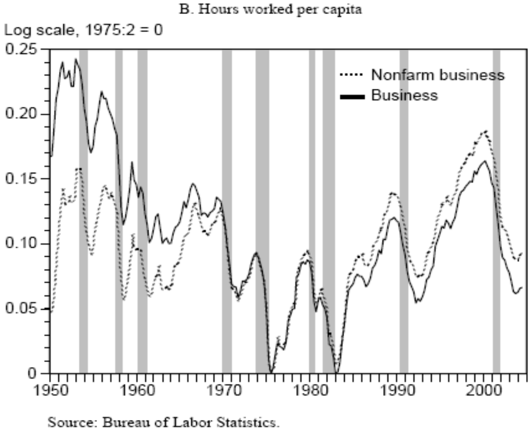
Source: What Happens After a Technology Shock?... Christiano Eichenbaum and Vigfusson NBER WK (2003)

# Tech and Hours: Labor productivity growth



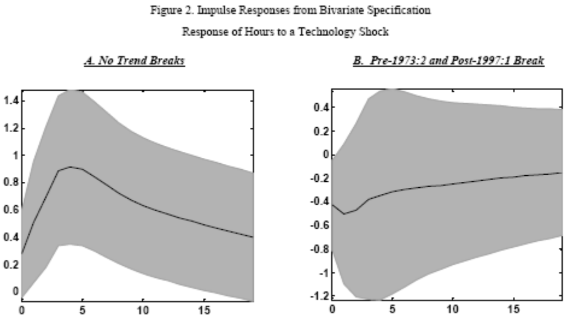

Source: Trend Breaks, Long-Run Restrictions, and Contractionary Technology Improvements, JME John Fernald (2007)

# Tech and Hours: Per-capita hours




Source: Trend Breaks, Long-Run Restrictions, and Contractionary Technology Improvements, JME John Fernald (2007)

# IRF (trend breaks)




Source: Trend Breaks, Long-Run Restrictions, and Contractionary Technology Improvements, JME John Fernald (2007)

# News and business cycles, Beaudry and Portier (2006 AER)

- Main idea back to Pigou: news about future productivity growth can generate business cycles since agents react to news by investing and consuming.
- Beaudry and Portier (2006 AER) finds news shocks are important for economic fluctuations. Output, investment, consumption and hours positively comove and the shocks explain a large fraction of the their variance.
- Use a VAR for TFP and stock prices.
- Standard model do not replicate the empirical finding since consumption comoves negatively with investment and hours.
- Big effort in building models where news shocks generate business cycles (Jaimovich and Rebelo, 2009, Den Haan and Kaltenbrunner, 2009, Schmitt-Grohe and Uribe, 2008).

# News and business cycles, Beaudry and Portier (2006 AER)

Two identification procedures:

1. Technology shocks is the only shock driving TFP in the long run.
2. News shocks raise stock prices on impact but not TFP (lagged adjustment.

# News and business cycle, Beaudry and Portier (2006 AER)

Main finding:

- the two identified shocks are the same;
- such shocks generate positive comovement in consumption, investment, output and hours (consistently with business cycles conmovements) and they explain a large portion of the variance of these series.

Conclusion: news shocks can generate business cycles.

# News and business cycle, Beaudry and Portier (2006 AER)

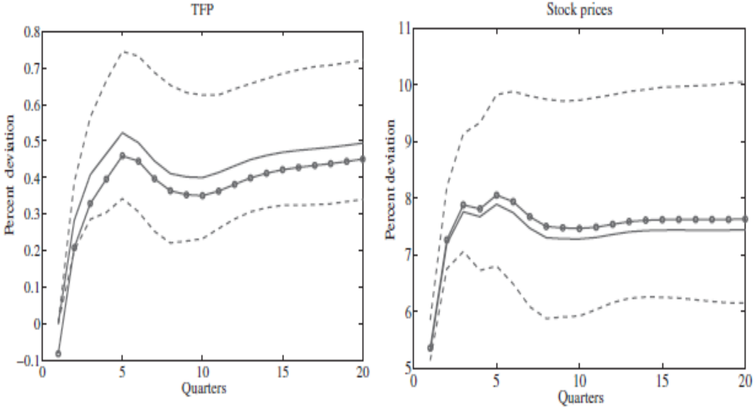



FIGURE 1. IMPULSE RESPONSES TO SHOCKS $\varepsilon_2$ AND $\tilde{\varepsilon}_1$ IN THE (*TFP*, *SP*) VECM

Source: Beaudry and Portier (AER 2006)

# News and business cycle, Beaudry and Portier (2006 AER)

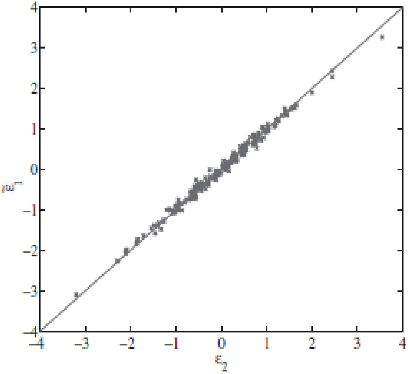



FIGURE 2. PLOT OF $\varepsilon_2$ AGAINST $\tilde{\varepsilon}_1$ IN THE (*TFP*, *SP*) VECM

Source: Beaudry and Portier (AER 2006)

# News and business cycle, Beaudry and Portier (2006 AER)

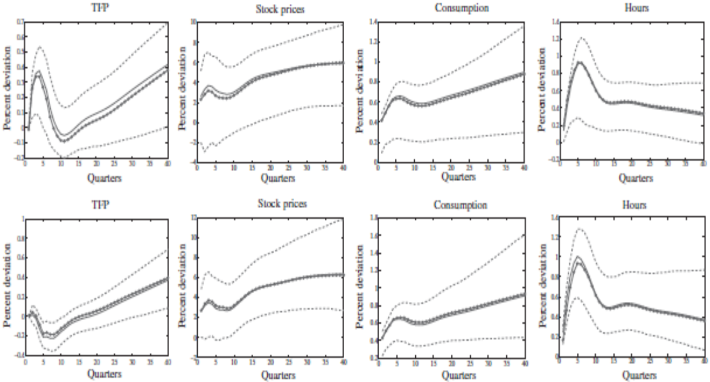



FIGURE 9. IMPULSE RESPONSES TO $\varepsilon_2$ AND $\tilde{\varepsilon}_1$ IN THE (*TFP*, *SP*, *C*, *H*) VECM, WITHOUT (UPPER PANELS) OR WITH (LOWER PANELS) ADJUSTING TFP FOR CAPACITY UTILIZATION

Source: Beaudry and Portier (AER 2006)

# News and business cycle, Beaudry and Portier (2006 AER)

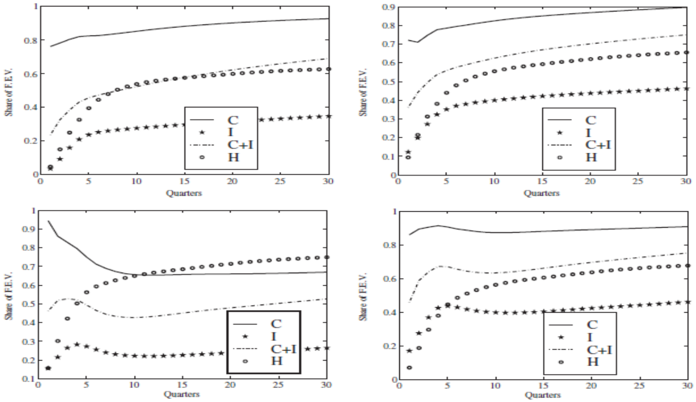
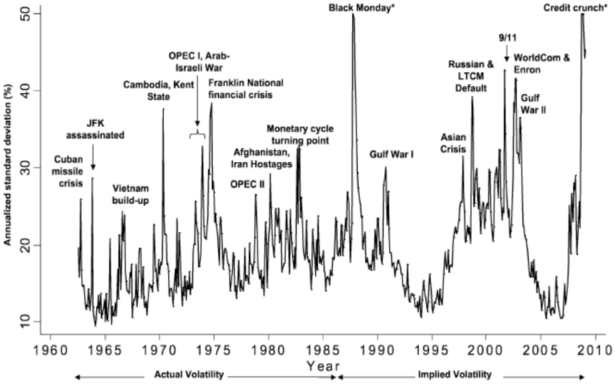



Figure 10. Share of the Forecast Error Variance (F.E.V.) of Consumption ($C$), Investment $I$, Output ($C + I$), and Hours ($H$) Attributable to $\varepsilon_2$ (left panels) and to $\tilde{\varepsilon}_1$ (right panels) in VECMs, with Nonadjusted TFP (top panels) or Adjusted TFP (bottom panels)

Source: Beaudry and Portier (AER 2006)

# Uncertainty Shocks, Bloom (Econometrica 2009)




Source: The impact of uncertainty shocks Bloom (ECON 2009)

# Uncertainty Shocks, Bloom (Econometrica 2009)

TABLE A.1

MAJOR STOCK-MARKET VOLATILITY SHOCKS

| Event | Max Volatility | First Volatility | Type |
|---|---|---|---|
| Cuban missile crisis | October 1962 | October 1962 | Terror |
| Assassination of JFK | November 1963 | November 1963 | Terror |
| Vietnam buildup | August 1966 | August 1966 | War |
| Cambodia and Kent State | May 1970 | May 1970 | War |
| OPEC I, Arab–Israeli War | December 1973 | December 1973 | Oil |
| Franklin National | October 1974 | September 1974 | Economic |
| OPEC II | November 1978 | November 1978 | Oil |
| Afghanistan, Iran hostages | March 1980 | March 1980 | War |
| Monetary cycle turning point | October 1982 | August 1982 | Economic |
| Black Monday | November 1987 | October 1987 | Economic |
| Gulf War I | October 1990 | September 1990 | War |
| Asian Crisis | November 1997 | November 1997 | Economic |
| Russian, LTCM default | September 1998 | September 1998 | Economic |
| 9/11 terrorist attack | September 2001 | September 2001 | Terror |
| Worldcom and Enron | September 2002 | July 2002 | Economic |
| Gulf War II | February 2003 | February 2003 | War |
| Credit crunch | October 2008 | August 2007 | Economic |

Source: The impact of uncertainty shocks Bloom (ECON 2009)

# Uncertainty Shocks, Bloom (Econometrica 2009)

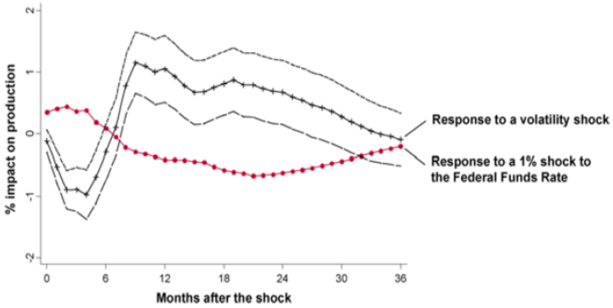



FIGURE 2.—VAR estimation of the impact of a volatility shock on industrial production. *Notes*: Dashed lines are 1 standard-error bands around the response to a volatility shock.

Source: The impact of uncertainty shocks Bloom (ECON 2009)