Is there the Technology Shock? Confronting Sign Restrictions with the properties of the Data

Aleksei Netšunajev [1]

[1]Free University of Berlin

July 11, 2014

Motivation

- Have technology shocks anything to do with the aggregate fluctuations?
- Long run identification: Gali(1999): persistent fall of hours worked after a technology shock;
- Christiano et al (2003): first difference per capita hours worked and obtain increase of hours worked;
- Fernald (2007), Francis and Ramey(2009): low frequency movements in per capita hours worked distort results of Christiano et al (2003).
- Francis and Ramey(2005, 2009): persistent fall of hours worked after a technology shock.

Motivation

- Canova et al (2010): disentangle investment specific and neutral technology shocks.
- Dedola and Neri (2007): identify technology shocks via sign restrictions: technology improvement tend to drive hours worked up
- Sign restrictions not without problems. Fry and Pagan (2011): set identified model, interpretation of IR, FEVD difficult. No possibility to check against the data.
- Possible solution: use heteroskedasticity

SVAR

Let time evolution of an $n \times 1$ vector $\mathbf{y}_t$ of endogenous variables be given by the following SVAR model:

$$\mathbf{y}_t = \mathbf{k}_0 + \mathbf{A}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{A}_p \mathbf{y}_{t-p} + \mathbf{B}\epsilon_t(s_t)\,, \tag{1}$$

where $\mathbf{k}_0$ is an intercept term, $\mathbf{B}$ is the instantaneous impact matrix, $\mathbf{A}_1, \ldots, \mathbf{A}_p$ are autoregressive matrices, and $\epsilon_t(s_t)$ is the vector of uncorrelated structural innovations that depends on the hidden state parameter $s_t \in \{1, \ldots, m\}$. We assume the following distribution of $\epsilon_t(s_t)$:

$$\epsilon_t(s_t) \sim \text{Normal}(\,\mathbf{0}, \mathbf{\Lambda}(s_t)\,)\,,$$

where $\{\mathbf{\Lambda}(s) : s = 1, ..., m\}$ is a family of distinct $n \times n$ diagonal matrices and $\mathbf{\Lambda}(1)$ is normalized to identity matrix.

SVAR

The model can be written in the usual reduced form VAR with time–varying volatility of innovations:

$$\mathbf{y}_t = \mathbf{k}_0 + \mathbf{A}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{A}_p \mathbf{y}_{t-p} + \mathbf{u}_t(s_t)\,, \qquad (2)$$

with

$$\mathbf{u}_t(s_t) \sim \text{Normal}(\,\mathbf{0}, \boldsymbol{\Sigma}(s_t)\,)\,.$$

Covariance matrices have the structure:

$$\boldsymbol{\Sigma}(s) = \mathbf{B}\boldsymbol{\Lambda}(s)\mathbf{B}'$$

for each volatility state $s \in \{1, \ldots, m\}$. Thus, given a sufficient number of distinct volatility regimes m, enough statistical information about $\mathbf{B}$ can be learned from the family of reduced–form variance–covariance matrices $\{\boldsymbol{\Sigma}(s) : s = 1, ..., m\}$.

Inference

We apply Bayesian methods for inference on all relevant model parameters using Gibbs sampling for reduced–form VAR parameters. We assume that for each $t = 1, \ldots, T$:

$$s_t \mid s_{t-1} \sim \text{Markov}(\,\mathbf{P}, \boldsymbol{\eta}_0\,)\,,$$

where $m \times m$ matrix $\mathbf{P}$ gives conditional distributions of state transitions, and s_0 is distributed according to the m-dimensional vector $\boldsymbol{\eta}_0$.

Inference: Gibbs sampling

Estimation algorithm includes four steps:

1. $\mathbf{S}_T$ is generated by drawing in reverse time order from the posterior distribution:

$$p(s_t \,|\, \mathbf{Y}_T) \propto p(s_t \,|\, \mathbf{Y}_t) \cdot p(s_{t+1} \,|\, s_t)\,,$$

where the first term in the expression is generated recursively using Chib (1996) simulation algorithm for hidden Markov models. It involves the prediction:

$$p(s_t \,|\, \mathbf{Y}_{t-1}) = \sum_{j=1}^{m} p(s_t \,|\, s_{t-1} = j) \cdot p(s_{t-1} = j \,|\, \mathbf{Y}_{t-1})\,,$$

and update steps:

$$p(s_t \,|\, \mathbf{Y}_t) \propto p(s_t \,|\, \mathbf{Y}_{t-1}) \cdot \ell(\mathbf{y}_t \,|\, \mathbf{Y}_{t-1}, \boldsymbol{\beta}, \boldsymbol{\Sigma}(s_t))\,,$$

where $\mathbf{Y}_t$ denotes data up to $1 \leq t \leq T$, $\boldsymbol{\beta}$ is the vector of VAR parameters, and $\ell(\mathbf{y}_t \,|\, \mathbf{Y}_{t-1}, \boldsymbol{\beta}, \boldsymbol{\Sigma}(s_t))$ is the likelihood of $\mathbf{y}_t$ in state s_t.

2. Simulations of the Markov transition matrix $\mathbf{P}$, where posterior probability of leaving state $s \in \{1, \ldots, m\}$ is given by:

$$p_s \,|\, \mathbf{S}_T, \mathbf{Y}_T \sim \text{Dirichlet}(\, \alpha_{s1} + n_{s1}, \ldots, \alpha_{sm} + n_{sm} \,) \,,$$

where $\{\alpha_{sk} : 1 \leq s, k \leq m\}$ are hyper–parameters of Dirichlet prior for $\mathbf{P}$, and n_{sk} is the number of transitions from state s to state k in $\mathbf{S}_T$.

Inference: Gibbs sampling

3. Posterior distribution of the variance–covariance matrix of reduced–form innovations in state $s \in \{1, \dots, m\}$ is following:

$$\mathbf{\Sigma}(s) \,|\, \mathbf{S}_T, \mathbf{Y}_T \sim \mathrm{IW}(\, [\underline{\mathbf{C}}(s) + \mathbf{C}(s)]\, \omega(s) + T(s)\,)\,,$$

where the family of $n \times n$ non-singular matrices $\{\underline{\mathbf{C}}(s) : s = 1, ..., m\}$ and scalars $\{\omega(s) : s = 1, ..., m\}$ are hyper–parameters of Wishart prior for $\{\mathbf{\Sigma}(s) : s = 1, ..., m\}$, and $\mathbf{C}(s)$ are sums of residual cross products for observations belonging to each volatility state s, where the number of such observations is given by $T(s)$.

4. Given the prior Gaussian distribution of the VAR parameters:

$$\beta \sim N(\underline{\mathbf{b}}, \underline{\mathbf{B}}),$$

posterior distribution is Gaussian:

$$\beta \,|\, \mathbf{Y}_T, \mathbf{S}_T, \{\mathbf{\Sigma}(s) : 1 \leq s \leq m\} \sim \text{Normal}(\,\bar{\mathbf{b}}, \bar{\mathbf{B}}\,)\,, \qquad (3)$$

where parameters of this distribution are given by the expressions:

$$\bar{\mathbf{b}} = \bar{\mathbf{B}}\,(\underline{\mathbf{B}}^{-1}\underline{\mathbf{b}} + (\mathbf{X}'\otimes\mathbf{I}_n)\,\mathbf{\Omega}(S_T)\,\mathbf{y})\,, \bar{\mathbf{B}} = \left[\underline{\mathbf{B}}^{-1} + (\mathbf{X}'\otimes\mathbf{I}_n)\,\mathbf{\Omega}(S_T)\,(\mathbf{X}\otimes\mathbf{I}_n)\right]^{-1},$$

where $nT \times nT$ block–diagonal matrix $\mathbf{\Omega}(S_T)$ is defined as follows:

$$\mathbf{\Omega}(S_T) := \begin{pmatrix} \mathbf{\Sigma}^{-1}(s_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mathbf{\Sigma}^{-1}(s_T) \end{pmatrix},$$

$\mathbf{y} := (\mathbf{y}_1', \dots, \mathbf{y}_T')'$ is a $nT \times 1$ data vector, and where each row of a $T \times (1 + np)$ data matrix $\mathbf{X}$ contains the following elements:

$$(1, \mathbf{y}_{t-1}', \dots, \mathbf{y}_{t-p}')\,.$$

Inference: Gibbs sampling

Recursive iteration on steps 1 to 4 produces a sequence of posterior draws of the reduced–form variance–covariance matrices $\{\boldsymbol{\Sigma}(s) : s = 1, ..., m\}$ and other model parameters after a certain number of burn-in steps. Sampling from the posterior distribution of the structural matrix $\mathbf{A}_0$ in the SVAR model (1) relies on the decomposition:

$$\boldsymbol{\Sigma}(1) = \mathbf{B}\,\mathbf{I}\,\mathbf{B}', \quad \boldsymbol{\Sigma}(s) = \mathbf{B}\,\boldsymbol{\Lambda}(s)\,\mathbf{B}' \ \text{ for } s = 1, ..., m,$$

which is unique up to column signs of $\mathbf{B}$ when $m = 2$:

$$\boldsymbol{\Sigma}(1) = \mathbf{B}\,\mathbf{I}\,\mathbf{B}', \quad \boldsymbol{\Sigma}(2) = \mathbf{B}\,\boldsymbol{\Lambda}(2)\,\mathbf{B}'$$

For the general case $m > 2$, this decomposition exists only when certain restrictions are imposed on the reduced–form variance–covariance matrices $\boldsymbol{\Sigma}(s)$. Empirical application in this paper is limited to the case of two volatility states.

Data

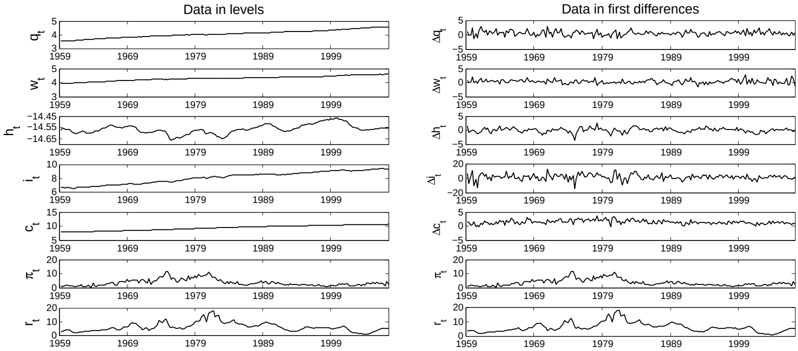
- Labor productivity (q_t): log of hourly labor productivity (Bureau of Labor Statistics (BLS));
- Real wage (w_t): log of real hourly compensation, non-farm business sector (BLS);
- Log of per capita hours worked (h_t);
- Per capita investment (i_t): log of per capita Gross Private Domestic Investment (Bureau of Economic Analysis (BEA));
- Per capita consumption (c_t): log of per capita Personal Consumption Expenditures (BEA);
- Inflation (π_t): is annualized percent changes in the implicit GDP deflator (BEA);
- Short-term interest rate (r_t): Federal Funds rate (Federal Reserve Bank of St. Louis).

Additional series: (i) hours worked is an index of total hours worked, non-farm business sector (BLS); (ii) population is civilian noninstitutional population 16+ (FED). Quarterly data for 1959Q1-2007Q6 is used.

Data

Data: $\mathbf{y}_t = [q_t, w_t, h_t, i_t, c_t, \pi_t, r_t]'$. Quarterly data for 1959Q1-2007Q2

Figure: US data 1959-2007



Priors

- Dirichlet prior on the elements of the Markov transition kernel $\mathbf{P}$ are $\alpha_{sk} = 10$ for $s = k$ and $\alpha_{sk} = 1$ for $s \neq k$, where $s, k \in \{1, 2\}$.
- The inverse Wishart prior hyper–parameters on the reduced–form variance–covariance matrices $\mathbf{\Sigma}(s)$ are given by $\underline{\mathbf{C}}(s) = \omega(s)\,\mathbf{I}$, $\omega(s) = 9$, $s = 1, 2$.
- Minnesota type priors using dummy observations on $\boldsymbol{\beta}$ are used. These are centered on random walk. Implementation as shown by Del Negro and Schorfheide (2011).

Table: Marginal data density for VAR models

Model	Lag order 1	2	3	4
VAR	-1992.373	-1964.837	-1957.466	-1969.328
MS(2)-VAR	-1794.931	-1752.385	-1762.509	-1775.822
MS(3)-VAR	-1796.895	-1766.741	-1785.320	-1825.455

Sign restrictions

Table: Sign restrictions of Dedola and Neri (2007)

Variables	$q_t \geq 0$	$w_t \geq 0$	$i_t \geq 0$	$c_t \geq 0$	$Y_t \geq 0$
Periods	0,...,19	2,...,19	0,...,9	0,...,4	0,...,9

S_T trajectories

Figure: Marginal posterior state probabilities

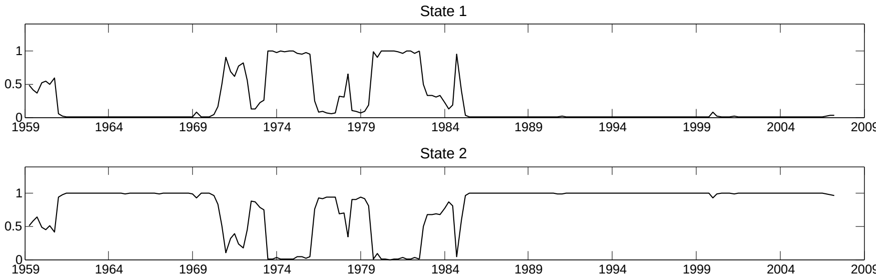


Figure: Prior-Posterior of **B**

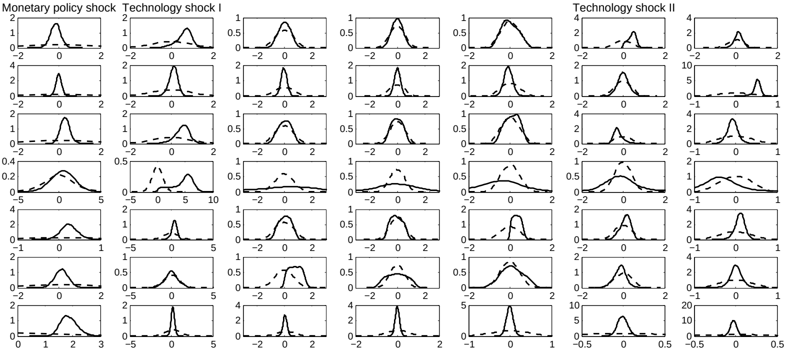


Figure: Technology shock I

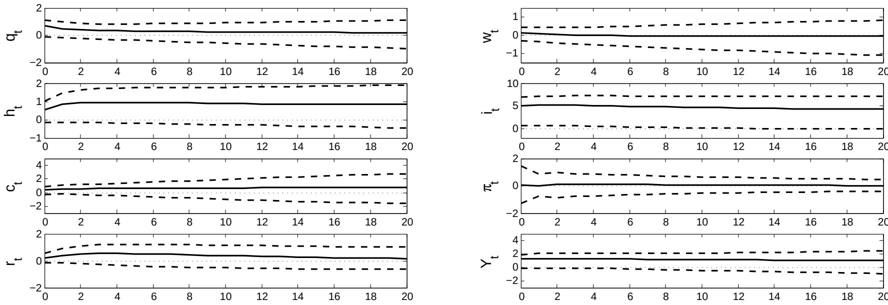


Figure: Technology shock II

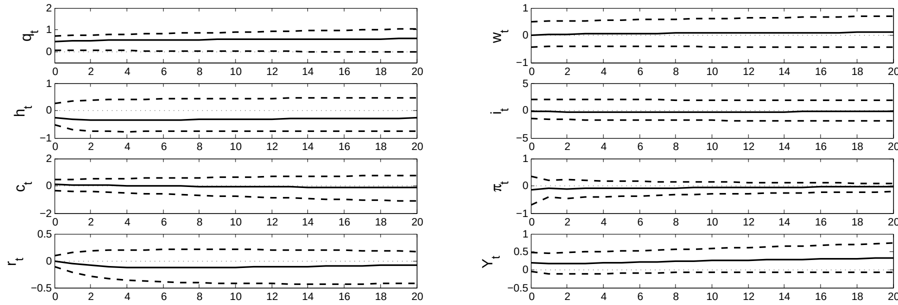


Figure: Monetary policy shock

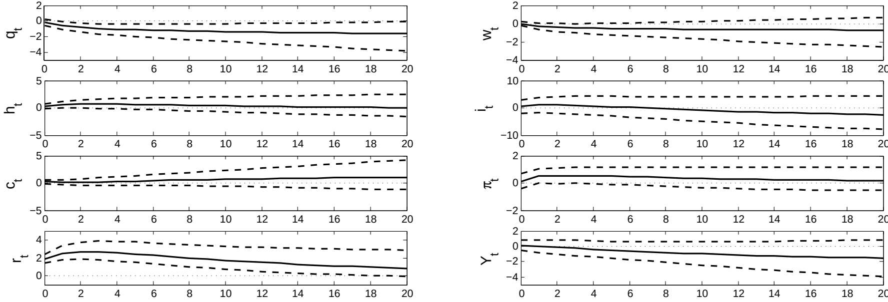
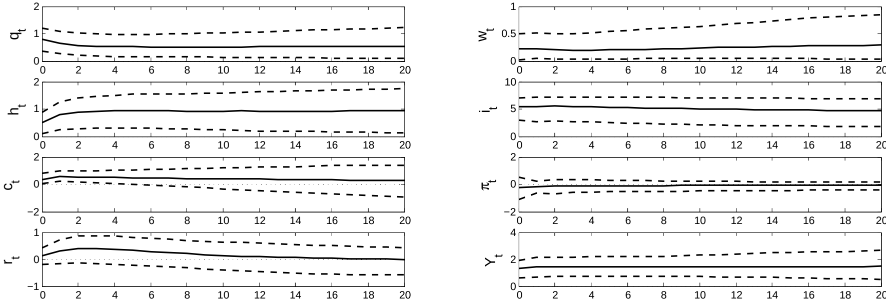


Figure: Technology shock I after truncation



Sign restrictions II

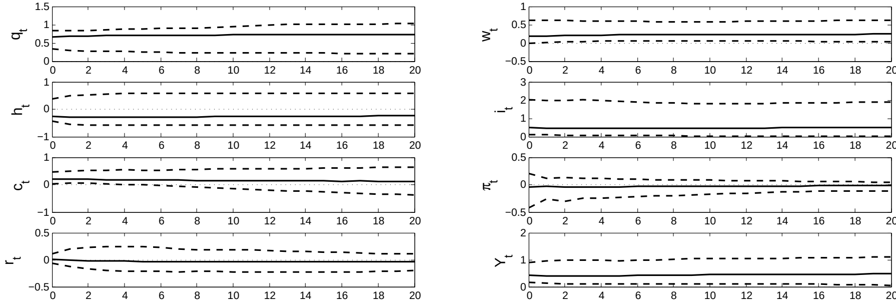


Figure: Technology shock II after truncation

Conclusions

- Heteroskedasticity is useful for checking sign restrictions in the medium scale SVARs
- The data is informative about the technology shocks as well as monetary policy shocks
- Two types of technology shocks are found in the dataset. The sign restrictions of Dedola and Neri (2007) are supported by the data
- Sign restrictions are not enough to disentangle two types of technology shocks.
- Heteroskedasticity and sign restrictions allow to label shocks as investment specific and neutral technology shocks. Hours worked tend to rise and fall respectively after those shocks.