Macro Shocks and Impulse Response Functions

Shuowen Chen

Spring 2020 Econometrics Reading Group

Outline

- Macro shocks and causal effects
- Structural Vector Autoregression (SVAR)
- Local projection (LP)
- Equivalence of impulse responses between SVAR and LP
- Innovations and productivity dispersions (with Ming)

Acknowledgement

The first part is based on the following papers:

- Jordà (2005)
- Ramey (2016)
- Stock and Watson (2018)
- Montiel Olea, Stock and Watson (2018)
- Plagborg-Møller and Wolf (2019)

Shocks in Macroeconomics

A **primitive**, unanticipated (**exogenous**) economic force, or driving impulse, that is unforecastable, uncorrelated with other shocks and economically meaningful.

Dynamic causal effects: responses of a macroeconomic aggregate to a shock over time (impulse response functions)

Dynamic Causal Effects

Denote $\varepsilon_{1,t}$ as a mean-zero random treatment at time t. Causal effect on variable of interest Y_2, h periods forward is

$$\mathbb{E}_t(Y_{2,t+h}|\varepsilon_{1,t} = 1; X_t) - \mathbb{E}_t(Y_{2,t+h}|\varepsilon_{1,t} = 0; X_t).$$

Assuming linearity and stationarity, h lag treatment effect regression:

$$Y_{2,t+h} = \Theta_{h,21}\varepsilon_{1,t} + \beta X_t + u_{t+h}$$

- Dynamic causal effect: $\Theta_{h,21}$ for $h = 0, 1, 2, ...$
- Structural shock ε_1 often unobserved
  - Key difference from microeconometrics
  - Need to identify shocks[1]
  - VAR (Sims, 1980): a link between innovations to a linear system and structural shocks

[1] In the R&D application we construct such a shock.

Slutzky–Frisch Paradigm

Path of observed macro variables arises from current and past shocks and measurement errors.

$$Y_t = \sum_{i=0}^{\infty} \Theta_i L^i \varepsilon_t,$$

- Y_t: $n \times 1$ macro variables
- ε_t : $m \times 1$ mutually uncorrelated structural shocks and measurement errors
- L: lag operator
- Θ_i: $n \times m$ matrix of coefficients

This is a structural moving average representation of Y_t

Structural Vector Autoregression

$$
\begin{aligned}
Y_t &= A_1 Y_{t-1} + A_2 Y_{t-2} + \ldots + A_p Y_{t-p} + \eta_t \\
\eta_t &= \Gamma \varepsilon_t,
\end{aligned}
$$

- η_t : VAR innovations, Γ non-singular (invertibility assumption)
- Structural shocks ε_t are serially and mutually uncorrelated:

$$
\mathbb{E}(\varepsilon_t) = 0, \quad \mathbb{E}(\varepsilon_t \varepsilon_t') = D = diag(\sigma_1^2, \ldots, \sigma_n^2)
$$

- Structural moving average representation:

$$
Y_t = \sum_{k=0}^{\infty} C_k(A) \Gamma \varepsilon_{t-k},
$$

where $C_k(A) = \sum_{m=1}^{k} C_{k-m}(A) A_m$, $k = 1, 2, \ldots$ with $C_0(A) = I_n$ and $A_m = 0$ for $m > p$.,

Structural Impulse Response

Response of $Y_{i,t+k}$ to a one-unit change in shock j in interest $\varepsilon_{j,t}$:

$$\frac{\partial Y_{i,t+k}}{\partial \varepsilon_{j,t}} = e_i' C_k(A) \Gamma e_j,$$

where e_j denotes the jth column of I_n. W.L.O.G,

- Target shock: $\varepsilon_{1,t}$
- Scale normalization[2]:
$$\frac{\partial Y_{1,t}}{\partial \varepsilon_{1,t}} = 1$$

Depends on VAR coefficients A and Γe_1.

- Need identification and VAR estimation
- Then invert the estimate to obtain IRF

A recent alternative: local projection (Jordà, 2005)

- Linear regressions of a future outcome on current covariates

[2] This is due to the shock being unobservable.

Local Projection

Recall impulse response definition:

$$IR(t, s, \varepsilon_1) = \mathbb{E}(y_{t+s} | \varepsilon_{1,t} = 1; X_t) - \mathbb{E}(y_{t+s} | \varepsilon_{1,t} = 0; X_t)$$

Project y_{t+s} onto the linear space generated by $(y_{t-1}, ..., y_{t-p})'$:

$$y_{t+s} = \alpha^s + B_1^{s+1} y_{t-1} + B_2^{s+1} y_{t-2} + ... + B_p^{s+1} y_{t-p} + u_{t+h}^s, s = 0, 1, ..., h$$

- α^s: $n \times 1$ vector of constants
- B_i^{s+1}: matrices of coefficients for lag i and horizon $s + 1$
- Impulse responses: $\widehat{IR}(t, s, \varepsilon_1) = \widehat{B}_1^s$

If the structural shock ε_1 is directly observable or constructed:

$$Y_{i,t+h} = \theta_{i,h} \cdot \varepsilon_{1,t} + \text{ controls } + \xi_{t+h}$$

Do SVAR and LP estimate the same IRFs?

Conventional wisdom:

SVAR more efficient, LP more robust to model misspecifications

Empirical Evidence:

Different results (Ramey, 2016; Nakamura & Steinsson, 2018)

Theory:

- Jordà (2005): if DGP is $VAR(p)$, equivalent estimations
- Plagborg-Møller and Wolf (2019):
  1. Exact same IRF in population under weak stationarity and unrestricted lag structure
  2. For fixed number of lags p, the two agree out to horizon p

When DGP is VAR (Jordà, 2005)

Denote $\boldsymbol{X}_t = (y_{t-1}, y_{t-2}, ..., y_{t-p})'$ and consider the $VAR(p)$ model:

$$y_t = \mu + \boldsymbol{\Pi}'\boldsymbol{X}_t + v_t,$$

where $\boldsymbol{\Pi}' = [\Pi_1, \Pi_2, ..., \Pi_p]$. Rewrite to have, for $s = 0, 1, \cdots, h$

$$\begin{aligned} y_{t+s} &= (I - \boldsymbol{F}_1^s - ... - \boldsymbol{F}_p^s)\mu + \boldsymbol{F}_1^{s+1} y_{t-1} + ... + \boldsymbol{F}_p^{s+1} y_{t-p} \\ &\quad + (v_{t+s} + \boldsymbol{F}_1^1 v_{t+s-1} + ... + \boldsymbol{F}_1^s v_t) \end{aligned}$$

where $\boldsymbol{F}_i^s$ is the i–th upper, $n \times n$ blaock of the matrix $\boldsymbol{F}^s$:

$$\boldsymbol{F} = \begin{pmatrix} \Pi_1 & \Pi_2 & \cdots & \Pi_{p-1} & \Pi_p \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{pmatrix}$$

When DGP is VAR (Jordà, 2005)

Recall local projection form:

$$y_{t+s} = \alpha^s + B_1^{s+1} y_{t-1} + B_2^{s+1} y_{t-2} + \cdots + B_p^{s+1} y_{t-p} + u_{t+s}^s, \quad s = 0, 1, \cdots, h$$

Therefore we have

$$\begin{aligned} \alpha^s &= (I - \boldsymbol{F}_1^s - \ldots - \boldsymbol{F}_p^s)\mu \\ B_1^{s+1} &= F_1^{s+1} \\ u_{t+s}^s &= (v_{t+s} + \boldsymbol{F}_1^1 v_{t+s-1} + \ldots + \boldsymbol{F}_1^s v_t) \end{aligned}$$

This suggests using HAC for inference

Equivalent IRF Estimation: LP

Suppose we have the time series data $w_t = (r_t', x_t, y_t, q_t')'$, where r_t and q_t are $n_r \times 1$ and $n_q \times 1$ while x_t and y_t are scalars. Interested in dynamic response of y_t after an impulse in x_t.

Data assumptions:

1. Data $\{w_t\}$ are covariance stationary and purely non–deterministic, with an everywhere nonsingular spectral density matrix and absolutely summable Wold decomposition coefficients
2.

Local projection: for $h = 0, 1, 2, \ldots$

$$y_{t+h} = \mu_h + \beta_h x_t + \gamma_h' r_t + \sum_{l=1}^{\infty} \delta_{h,l}' w_{t-l} + \xi_{h,t}$$

LP impulse response coefficients: $\{\beta_h\}_{h \geq 0}$

Equivalent IRF Estimation: VAR

VAR(∞):

$$w_t = c + \sum_{l=1}^{\infty} A_l w_{t-l} + u_t,$$

where $u_t = w_t - \mathbb{E}(w_t | \{w_\tau\}_{-\infty<\tau<t})$ is the projection residual. Let $\Sigma_u = \mathbb{E}(u_t u_t')$ and define Cholesky decomposition $\Sigma_u = BB'$. Consider the following SVAR representation

$$A(L) w_t = c + B\eta_t,$$

where $A(L) = I - \sum_{l=1}^{\infty} A_l L^l$ and $\eta_t = B^{-1} u_t$.
Define $\sum_{l=1}^{\infty} C_l L^l = C(L) = A(L)^{-1}$, then VAR impulse of y_t w.r.t x_t is

$$\theta_h = C_{n_r+2,\cdot,h} B_{\cdot,n_r+1}$$

Plagborg-Møller and Wolf: Proposition 1 & 2

1. Define $\tilde{x}_t = x_t - \mathbb{E}(x_t | r_t, \{w_\tau\}_{-\infty<\tau<t})$. Under **Data assumptions**, for all $h = 0, 1, 2,$

$$\theta_h = \sqrt{\mathbb{E}(\tilde{x}_t^2)} \times \beta_h$$

2. Now consider finiate lag structure (p instead of ∞).

$$y_{t+h} = \mu_h + \beta_h x_t + \gamma_h' r_t + \sum_{l=1}^{p} \delta_{h,l}' w_{t-l} + \xi_{h,t}$$

$$w_t = c + \sum_{l=1}^{p} A_l w_{t-l} + u_t$$

Define $\tilde{x}_t(l) = x_t - \mathbb{E}(x_t | r_t, \{w_\tau\}_{t-l\leq\tau<t})$ for all $l = 0, 1, 2, \ldots$ Let nonnegative integers h, p satisfy $h \leq p$. If $\tilde{x}_t(p) = \tilde{x}_t(p-h)$, then

$$\theta_h(p) = \sqrt{\mathbb{E}[\tilde{x}_t(p)^2]} \times \beta_h(p)$$

Project: R&D Responses to Innovations

Research Question:

After a firm has accomplished a highly valuable innovation, how does its competitors respond differently in terms of R&D?

Why should we care?

- Important to understand: how destructive is innovation?
- Heterogenous responses can explain productivity dispersions

What do we do?

- A quantitative measure of high–valued innovations.
- Such innovations are desctructive to competitors' market shares
- Competitors' R&D responses dependent on previous levels

Data

We merge the following three datasets:

1. CRSP/Compustat Merged Database (1957 – 2011)
2. NBER-CES Manufacturing Industry Database (1958 – 2011)
3. Measure of firm-level innovation dataset (Kogan et al., 2017)
    - Unbalanced panel with patent grants and CRSP stock market data (1926 – 2010)
    - Scientific (Tcw) and economic (Tsm) values of patents

The data we work with

- Unbalanced panel (1970 – 2010) with 4062 U.S manufacturing firms and 135 industries by 4–digit SIC

High–Valued Innovation Index

A destructive innovation occurs in industry j, year t if a within industry firm i has a **high valued** new patent $(Tsm_{i,j,t})$

- Pool all firm–year patent values within industry j
- Take the $95th$-percentile as a cutoff: c^j_{95}

Firm level index:

$$MI\_firm_{i,j,t} = \begin{cases} 1 & \text{if } Tsm_{i,j,t} \geq c^j_{95}, \\ 0 & \text{otherwise} \end{cases}$$

Industry level index:

$$MI_{j,t} = \begin{cases} 1 & \text{if } \max_i \{MI\_firm_{i,j,t}\} = 1, \\ 0 & \text{otherwise} \end{cases}$$

High–Valued Innovations and Market Shares

Local projection:

$$S_{j,t+h} = \alpha_j^h + \delta_t^h + \beta_t^h MI_{j,t} + \boldsymbol{\lambda_{j,t}^h X_{j,t}} + \varepsilon_{j,t}^h,$$

- $h = 0, 1, \cdots, 10$
- Placebo: $h = -5, \cdots, -1$
- $S_{j,t}$: competitors' market shares

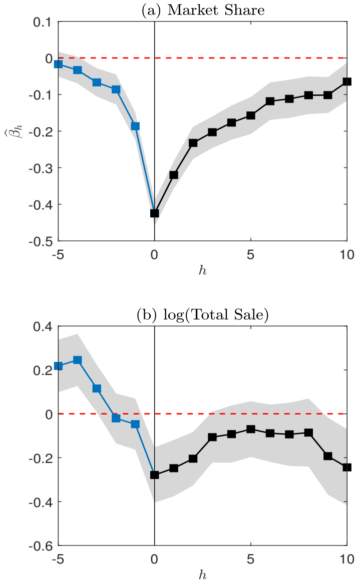
$$S_{j,t} = \sum_i ms_{i,j,t} \times \mathbb{1}\left[\max_{s=t-10,\cdots,t} MI\_firm_{i,j,s} = 0\right]$$

- $\boldsymbol{X_{j,t}}$: control variables
  - lagged firm numbers within industry
  - lagged Herfindahl-Hirschman index (HHI)
  - lagged dependent variable
- Alternative dependent variable:

  total sale of non–MI firms (in logs)

mkt share and HHI definitions

High-Valued Innovations Are Destructive



- Standard errors clustered at industry level (90% CI)
- Downward pretrend:

  Innovation has effect before patent granted: firms don't wait until getting granted to utilize new technology
- Destructive Innovation:

  Decreases competitors' market shares and log total sales
- **Next question:**

  How do competitors' R&D intensities (RDI) respond to destructive innovations?

R&D Responses to Destructive Innovations

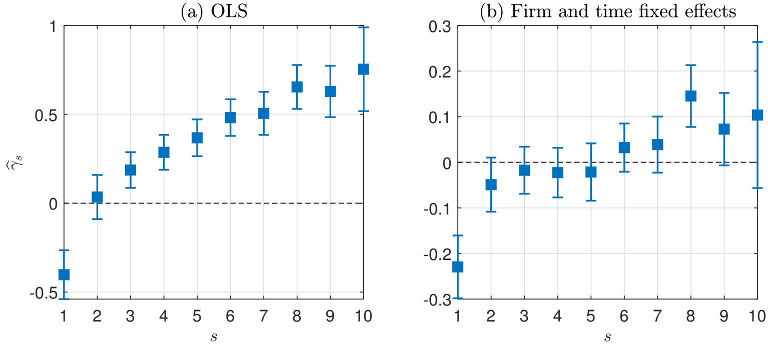
Sort and partition $\log(RDI_{i,j,t-1})$ into 10 groups; let $g^s_{i,j,t-1}$ denote group index, $s = 1, ..., 10$

Regression specification:

$$\log(RDI_{i,j,t}) = \alpha_i + \delta_t + \sum_{s=2}^{10} \beta_s g^s_{i,j,t-1} + \sum_{s=1}^{10} \gamma_s MI_{j,t-1} \times g^s_{i,j,t-1} + \boldsymbol{\lambda X_{i,j,t}} + \varepsilon_{i,j,t},$$

- Fully saturated model (no need to add $\sum_{s=2}^{10} \beta^M_s MI_{j,t-1}$)
- γ_s : group-wise responses to destructive innovations

Heterogeneity Across Deciles in Log-RDI



In response to destructive innovations:

- Previously high-RDI firms become higher in RDI
- Previously low-RDI firms become lower in RDI

Wrapping Up

What we have done

- A quantitative measure of high–valued innovations.
- Such innovations are desctructive to competitors' market shares
- Competitors' R&D responses dependent on previous levels

What we are doing

1. Build up a model to explain (almost done)
2. Quantitative analysis (33% done in Matlab and Julia) yy

Varibale Constructions

- Market share: $$ms_{i,j,t} = \frac{sale_{i,j,t}}{\sum_i sale_{i,j,t}} \times 100,$$
- Herfindahl-Hirschman index:

$$HHI_{j,t} = \sum_i ms^2_{i,j,t}$$

◂ Back