# Innovation Network Marries Production Network, and Business Cycle

Wu Zhu*

Department of Econ, University of Pennsylvania
(with Yucheng Yang, Applied Math, Princeton)

September 12, 2020

The speed at which the US economy recovers from recessions varies greatly, from months to years.

1. Recession of 1991 v.s. Recession of 2008, Recession of 2001 v.s. Recession of 2008
2. The speed of recovery:
   - Time to get back to the pre-recession growth trend.
   - Not the level of gross output or time between NBER trough and peak.

Key question:

What drives the speed of the recovery of the economy from recessions or adversarial shocks?

# Main Results

Focus on the role of production and innovation networks on business cycle.

1. Input-output network: sectors are linked through intermediate goods.
2. Innovation network: sectors are linked by technology - firms learn from each other about production technology.

Examine, theoretically and empirically, how cross-sectional shocks to technology innovation become persistent and amplified

- "Shock" - a vector of sectoral shocks to innovation, each component is a shock to a sector.

# Main Results

Intuitively, when an adverse multi-sector shock hits the economy, the effect on aggregate growth includes:

1. Amplification - initial dip in growth of the economy,
2. Persistence - length of time to recovery.

We decompose the effects of the shock on growth into several components, each has its amplification and persistence:

1. Amplification - fully captured by two sufficient statistics:
   1.1 correlation between the shock and vector of node centralities (centrality) in innovation network - "e.g. large shock to IT sector"
   1.2 correlation between centrality in innovation network and centrality in production network. - "e.g. large shock to IT v.s. large shock to OIL"
2. Persistence - depreciation effect v.s. spillover effect in innovation network, fully captured by the eigenvalue of the innovation network.

# Main Results

A significantly slow recovery occurs if

C1. Cross-sectional shock highly correlates with the eigenvector centrality of the innovation network (amplification)

C2. Eigenvector centrality in the innovation network highly correlates with the Katz centrality in the production network (amplification)

C3. Spillover effect roughly cancels out the depreciation effect (persistence)

Does this channel matter ?

1. Innovation network - google patent citation (1919-2018).
2. Production network - BEA input-output table (1951-2018).
3. Other parameters - state space model.

Empirically, we show:

1. Large time-variation on the amplification (for the first component). E.g. during the recession of 2008, the shock highly correlated with the eigenvector centrality of the innovation network.(amplification)
2. Innovation network is low-rank, for the first component, the spillover effect roughly cancels out the depreciation effect.(persistence)

# Literature

- Amplification and Persistence
  1. Financial Frictions: (Bernnake and Gertler, 1989;Kiyotaki and Moore,1997;Brunnermeier, Eisenhach, and Sannikov,2012)
  2. Endogeneous TFP: (Comin and Gertler,2006; Anzoategui, Comin, Gertler, and Martinez,2019;Bianchi, Kung, and Morales,2019)
  3. We show sectoral distribution of the shock matters, reveals useful information on future recovery.
- Technology diffusion and Innovation network:
  1. Technology diffusion dominates in growth, (Jaffe,1986;Bloom,Schankerman,Van Reenen,2013)
  2. Stable network structure and Slow diffusion, (Acemoglu,Akcigit, and Kerr,2016;Ahmadpoor and Jones,2017)
  3. Dynamic and macro implications.
- Idiosyncratic shocks in production network:
  1. (Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi,2012; Atalay, 2017;Baqaee and Farhi,2019)
  2. Innovation network marries production network, sectoral distribution of the shock matters for persistence.
- Long run risk:
  1. (Garleanu, Panageas, and Yu,2012; Garleanu, Kogan, and Panageas, 2012; Kogan, Papanikolaou, and Stoffman,2013; Kung and Schmid,2015)
  2. Endogenize the long-run risk in networking economy

# Outline

# Model

$J$ sectors, sector $i$ produce its output at time $t$ (Long and Plosser, 1983):

$$Y_{it} = \tilde{A}_{it} I_{it}^{\eta}, \;\; s.t. \;\; I_{it} = \prod_{j \in [J]} X_{ijt}^{\theta_{ij}} \tag{1}$$

where $\tilde{A}_{it} = A_{it} M_{it}$.
$A_{it}$: productivity driven by technology.
$M_{it}$: productivity beyond technology like managerial ability or agency conflicts etc.

Denote $a_{it} = \log(A_{it})$, $\Delta a_{it} = a_{it} - a_{it-1}$, model $\Delta \boldsymbol{a_t} = (\Delta a_{1t}, ..., \Delta a_{Jt})$ as a learning process. Think of $\Delta \boldsymbol{a_t}$ as new idea about production.

1. The process $a_{it}$ follows (Aghion and Howitt,1992).
$$\Delta a_{it} = \mu_{it} + \epsilon_{it}^{A} \tag{2}$$
$\mu_{it}$: arrival rate of the new technology or innovation for sector $i$ between $t$ and $t+1$, $\epsilon_{it}^{A}$: the shock to innovation realization.
2. The latent arrival rate is modeled as
$$\mu_{it+1} = \underbrace{(1-\rho)\mu_{it}}_{\text{mean-reverse effect}} + \underbrace{\sum_{j} W_{ij} \Delta a_{jt}}_{\text{knowledge learning from other firms}} + \epsilon_{it}^{u} \tag{3}$$

Write (2) and (3) in matrix notation

$$\begin{aligned} \Delta \boldsymbol{a}_t &= \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t^A \\ \boldsymbol{\mu}_{t+1} &= (1-\rho)\boldsymbol{\mu}_t + \boldsymbol{W}\Delta \boldsymbol{a}_t + \boldsymbol{\epsilon}_t^u \end{aligned} \tag{4}$$

with $\Delta \boldsymbol{a}_t = (\Delta a_{1t}, ..., \Delta a_{Jt})'$, $\boldsymbol{\epsilon}_t^A = (\epsilon_{1t}^A, ..., \epsilon_{Jt}^A)'$, $\boldsymbol{\epsilon}_t^u = (\epsilon_{1t}^u, ..., \epsilon_{Jt}^u)'$

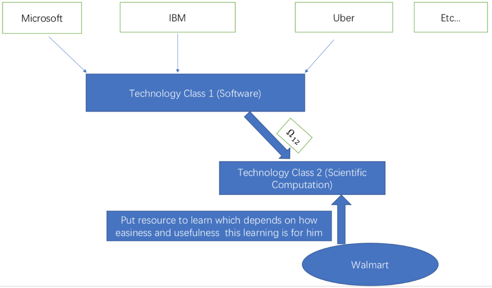
Special cases. Suppose $\boldsymbol{\epsilon}_t^A = 0$, the technology process is reduced to

$$\Delta \boldsymbol{a}_{t+1} = [(1-\rho)\boldsymbol{I} + \boldsymbol{W}]\Delta \boldsymbol{a}_t + \boldsymbol{\epsilon}_t^u \tag{5}$$

The usual VAR structure:

1. If $\boldsymbol{W}$ a diagonal matrix, no technology diffusion between sectors (Onatskia and Murcia,2013; Atalay,2017)
2. If $(1-\rho)\boldsymbol{I} + \boldsymbol{W} = 0$ and $\boldsymbol{\epsilon}_t^u = \Lambda_F \boldsymbol{F}_t + \boldsymbol{v}_t$ (Foerster,Sarte,and Watson,2011).
    2.1 $\boldsymbol{v}_t$ :idiosyncratic shock.
    2.2 $\boldsymbol{F}_t$ : Common shock.
    2.3 $\Lambda_F$ : the exposure matrix to the common shock.

Innovation Network - not a black box (micro foundation, see appendix), click back to page 24




1. $\boldsymbol{I} = (I_{i,1}, ..., I_{i,T})$: the resource distribution put by $i$ over technology space (drop subscript $t$ for simplicity), $T$: # of technology class.
2. $\boldsymbol{F}_q = (F_{q,1}, ..., F_{q,J}), q \in [\boldsymbol{T}]$: the distribution of innovations in technology field $q$ over industries $1, 2, ..., J$.
3. $\Omega_{\tau,q}$: the easiness of researchers with expertise in field $\tau$ learning from field $q$.
4. In micro foundation, we show $\tilde{W}_{i,j} = \sum_{q,\tau} I_{i\tau}\Omega_{\tau,q}F_{q,j}$.
5. $W_{i.} = (W_{i1}, ..., W_{ij}) = \xi_i(\tilde{W}_{i1}, ..., \tilde{W}_{iJ})$

# Model - Generalization

Generalization:

1. General CES production technology:

$$Y_{it} = A_{it} I_{it}^{\eta}, \;\; s.t. \;\; I_{it} = \Big[ \sum_{j \in [J]} \theta_{ij} X_{ijt}^{1-1/v_i} \Big]^{\frac{1}{1-1/v_i}}$$

2. General learning process:

$$\mu_{it+1} = (1-\rho)\mu_{it} + \sum_{j} W_{ij} \varphi(L) \Delta a_{jt} + \epsilon_{it}^{u}$$

- $\rho \in (0,1)$: depreciation rate of new idea (Bloom et al,2020AER).
- $\varphi(L)\Delta a_{jt} = \varphi_0 \Delta a_{jt} + \varphi_1 \Delta a_{jt-1} + \ldots$: learning from past innovation.
- We standardize $W_{i\cdot} = (W_{i1}, \ldots, W_{iJ}) = \xi_i(\tilde{W}_{i1}, \ldots, \tilde{W}_{iJ})$ such that

$$\sum_{j \in [J]} \tilde{W}_{ij} = 1$$

$\xi_i$ captures the learning efficiency.

The representative household maximizes

$$max_{C_t,\phi_{jt}} U_t := \sum_{s\geq t} \frac{C_t^{1-\gamma}}{1-\gamma} \tag{6}$$

subject to

$$C_t + \sum_j \phi_{jt}(V_{jt} - D_{jt}) = \sum_j \phi_{jt-1} V_{jt}$$

$\phi_{jt}$ is share holding on $j, V_{jt}$: the market value of firm j, $D_{jt}$: the dividend of firm $j$.

$C_t$: a representative firm that purchases intermediate goods at price $P_{jt}$ and combines them to a final good that is being consumed. The price of the final good is normalized to 1.

$$max_{c_{jt}} C_t - \sum_j P_{jt} c_{jt} \quad s.t. \ C_t = \prod_j c_{jt}^{\alpha_j} \tag{7}$$

# Theoretical Results - General Equilibrium

[1.] Markets for intermediates are competitive, Firm $i$'s optimization implies

$$I_{it} : [s_{it} Y_t]^{1-\eta} = (P^I_{it})^{-\eta} P_{it} A_{it} \eta^{\eta}, i \in [J] \Longrightarrow \tag{8}$$
$$(1-\eta)[\log(\boldsymbol{s_t}) + 1 \log(Y_t)] = -\eta \log(\boldsymbol{P}^I_t) + \log(\boldsymbol{P_t}) + \log(\boldsymbol{A_t}) + \eta \log(\eta) 1$$

$$X_{ijt} : P^I_{jt} = \prod_i [\frac{P_{it}}{\theta_{ji}}]^{\theta_{ji}} \Longrightarrow \log(\boldsymbol{P}^I_t) = \Theta \log(\boldsymbol{P}_t) - \boldsymbol{N}_t \tag{9}$$

[2.] The optimization for final consumption producer,

$$\prod [\frac{P_{it}}{\alpha_i}]^{\alpha_i} = 1 \Longrightarrow \boldsymbol{\alpha}' \log(\boldsymbol{P}_t) = \boldsymbol{\alpha}' \log(\boldsymbol{\alpha}) \tag{10}$$

with,
$Y_t$ : $Y_t = \sum_{j \in [J]} P_{jt} Y_{jt}$,
$P^I_{it}$: shadow price of composite good $I_{it}$,
$N_{it}$: $N_{it} = \sum_i \theta_{ij} \log(\theta_{ij})$, input sparsity of sector $i$,
$s_{it}$ : $s_{it} = \frac{P_{it} Y_{it}}{\sum_{j \in [J]} P_{jt} Y_{jt}}$, sale share of sector $i$

Market clearing:

[3.] Market clears for product $i$,

$$
\begin{aligned}
c_{it} + \sum_{j\in[J]} X_{jit} = Y_{it} &\Longrightarrow \alpha_i(1-\eta)Y_t + \sum_j \theta_{ji} P_{jt} Y_{jt} = P_{it} Y_{it} \\
&\Longrightarrow s_{it} = \sum_j s_{jt}\theta_{ji} + \alpha_i(1-\eta) \Longrightarrow \boldsymbol{s}_t = (1-\eta)(1-\Theta')^{-1}\boldsymbol{\alpha}
\end{aligned}
\tag{11}
$$

[4.] Market clears for stock market $i$ implies,

$$
\sum_i D_{it} = C_t \Longrightarrow C_t = (1-\eta)Y_t \tag{12}
$$

Note: slightly different from traditional definition that $Y_t = C_t$, but differs by a constant fraction $1-\eta$.

**Proposition (First Main Result)**

*Under Cobb-Douglas*

$$y_t := \log(Y_t) = \boldsymbol{s}_t'\Big[-\log(\boldsymbol{s}_t) + \frac{\eta}{1-\eta}\boldsymbol{N_t} + \frac{1}{1-\eta}\boldsymbol{a}_t\Big] + \frac{\eta}{1-\eta}log(\eta) + \boldsymbol{\alpha}'\boldsymbol{\alpha} \quad (13)$$

- *Concentration effect:* $-\boldsymbol{s}_t'\log(\boldsymbol{s}_t)$*, resource allocation across sectors*
- *Sparsity effect:* $\boldsymbol{s}_t'\boldsymbol{N_t} = \sum_{j\in[J]} s_{jt}N_{jt}$*, allocation within sector across inputs.*
- *Hulten effect:* $\boldsymbol{s}_t'\boldsymbol{a}_t$*, captures first order effect of technology (Hulten 1976)*

*Under Cobb-Douglas case,* $\boldsymbol{s}_t$ *and* $\boldsymbol{N}_t$ *are constant over time.*

$$g_t := \frac{1}{1-\eta}\boldsymbol{s}'\Delta\boldsymbol{a}_t \quad (14)$$

Under general CES, the equation 14 still true with an adjustment on the sparsity (see the paper)

We examine how sectoral shock to innovation affect future growth. Under Cobb-Douglas, the system is log-linear. (Note: not usual impulse response function)

1. Consider a shock $\epsilon_t$ to the arrival rate:

$$\boldsymbol{\mu}_t \to \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t, \text{ where } \boldsymbol{\epsilon}_t = \boldsymbol{\mu}_t - \mathbb{E}_{t-1}\boldsymbol{\mu}_t$$

2. The associated effect on growth and arrival rate at period $t+\tau$ :

$$\boldsymbol{\mu}_{t+\tau} \to \boldsymbol{\mu}_{t+\tau} + \delta\boldsymbol{\mu}_{t+\tau}, \text{ where } \delta\boldsymbol{\mu}_{t+\tau} = \boldsymbol{\mu}_{t+\tau} - \mathbb{E}_{t-1}\boldsymbol{\mu}_{t+\tau}$$
$$g_{t+\tau} \to g_{t+\tau} + \delta g_{t+\tau}, \text{ where } \delta g_{t+\tau} = g_{t+\tau} - \mathbb{E}_{t-1}g_{t+\tau}$$

Proposition

$$\begin{aligned}\mathbb{E}_t\delta\boldsymbol{\mu}_{t+\tau} &= \mathbb{E}_t[\boldsymbol{\mu}_{t+\tau} - \mathbb{E}_{t-1}\boldsymbol{\mu}_{t+\tau}] = [(1-\rho)\boldsymbol{I} + \boldsymbol{W}]^\tau \boldsymbol{\epsilon}_t \\ \mathbb{E}_t\delta g_{t+\tau} &= \mathbb{E}_t[g_{t+\tau} - \mathbb{E}_{t-1}g_{t+\tau}] = \frac{1}{1-\eta}\boldsymbol{\epsilon}_t'[(1-\rho)\boldsymbol{I} + \boldsymbol{W'}]^\tau \boldsymbol{s}_t\end{aligned} \tag{15}$$

PROPOSITION (SECOND MAIN RESULT)

*Assume* $\boldsymbol{W}$ *to be diagonalizable, we can decompose the effect of the shock on future growth into:*

$$\begin{aligned}\mathbb{E}_t\, \delta g_{t+\tau} &= \frac{1}{1-\eta}\boldsymbol{\epsilon}_t'[(1-\rho)\boldsymbol{I}+\boldsymbol{W'}]^{\tau}\boldsymbol{s}_t \\ &= \frac{1}{1-\eta}\sum_{k=1}^{J}[1-(\rho-\lambda_k(\boldsymbol{W'}))]^{\tau}(\boldsymbol{s}_t,\boldsymbol{v}_k)(\boldsymbol{\epsilon}_t,\boldsymbol{v}_k)\end{aligned} \tag{16}$$

Where,

1. $(\boldsymbol{x},\boldsymbol{y})$ the inner-product of $\boldsymbol{x}$ and $\boldsymbol{y}$
2. $\boldsymbol{v}_k'\boldsymbol{W'} = \lambda_k(W)\boldsymbol{v}_k'$, with $\lambda_1(\boldsymbol{W}) > ... > \lambda_J(\boldsymbol{W})$. $\lambda_k(\boldsymbol{W})$ the $k$th largest eigenvalue of $\boldsymbol{W}$ and $\boldsymbol{v}_k$ the associated eigenvector.

$$\mathbb{E}_t \delta g_{t+\tau} = \frac{1}{1-\eta} \sum_{k=1}^{J} \underbrace{[1-(\rho-\lambda_k(\boldsymbol{W}'))]^{\tau}}_{persistence} \underbrace{(\boldsymbol{s}_t, \boldsymbol{v}_k)(\boldsymbol{\epsilon}_t, \boldsymbol{v}_k)}_{amplification}$$

Intuition on main results.

1. Amplification:
   1.1 $(\boldsymbol{\epsilon}_t, \boldsymbol{v}_k) = \sum_{i=j}^{J} v_{ki}\epsilon_{jt}$, weighted shock with weights $v_{ki}$.
   1.2 $(\boldsymbol{s}_t, \boldsymbol{v}_k) = \sum_{i=j}^{J} v_{ki}s_{jt}$, production network interacts with innovation network.
2. Persistence: $\rho - \lambda_k(\boldsymbol{W})$, the decline rate of the shock in innov-network when $\boldsymbol{\epsilon}_t \propto \boldsymbol{v}_k$ since
   $$[(1-\rho)\boldsymbol{I} + \boldsymbol{W}']^{\tau}\boldsymbol{\epsilon}_t = [1-(\rho-\lambda_k(\boldsymbol{W})]\boldsymbol{\epsilon}_t$$
   $\rho$ : depreciation rate, $\lambda_k(\boldsymbol{W})$ : promotion rate due to spillover.

Intuition: $\boldsymbol{v}_1$ is the eigenvector centrality (Bonacich etal.,2001; Jackson etal.,2020)

$$v_{1i} = \frac{1}{\lambda_1(\boldsymbol{W})} \sum_j v_{1j} W_{ji}$$

$v_{1i}$: $i$ is important if sectors who learn from $i$ are important.

# Theoretical Results - Structure Matters

$$\mathbb{E}_t \delta g_{t+\tau} = \frac{1}{1-\eta} \sum_{k=1}^{J} [(1-\rho) + \lambda_k(\boldsymbol{W'})]^{\tau} (\boldsymbol{s}_t, \boldsymbol{v}_k)(\boldsymbol{\epsilon}_t, \boldsymbol{v}_k)$$

Case1: If $\boldsymbol{W} = 0$, then $\mathbb{E}_t \delta g_{t+\tau} = \frac{1}{1-\eta}(1-\rho)^{\tau}(\boldsymbol{\epsilon}_t, \boldsymbol{s}_t) = (1-\rho)^{\tau} \delta g_t$.

Case2: If $\lambda_1(\boldsymbol{W}) = ... = \lambda_J(\boldsymbol{W}) := \lambda$, then $\mathbb{E}_t \delta g_{t+\tau} = (1-\rho+\lambda)^{\tau} \delta g_t$.

Case3: The network is low-rank, such that $\lambda_1(\boldsymbol{W}) >> \lambda_2(\boldsymbol{W})$.

- 3.1: if $\rho \approx \lambda_1(\boldsymbol{W})$, the effect will become very persistent if $\boldsymbol{\epsilon} \propto \boldsymbol{v}_1$. rapidly decline if $\boldsymbol{\epsilon} \propto \boldsymbol{v}_2$

Under cases 1 and 2, Sectoral distribution of the shock (direction of the vector) does not matter conditional on the initial aggregate effect on growth. However, under case 3, the direction of the shock matters for the recovery process.

# An illustrative example - structure meets shock

Consider an economy with 3 sectors:

1. Symmetric production network, $\boldsymbol{s}_t = (1/3, 1/3, 1/3)$
2. $\boldsymbol{W}$ matrix

$$\begin{bmatrix} 0.327 & 0.067 & 0.067 \\ 0.067 & 0.047 & 0.047 \\ 0.067 & 0.047 & 0.047 \end{bmatrix}$$

3. $\lambda_1(\boldsymbol{W}) = 0.36, \lambda_2(\boldsymbol{W}) = 0.06, \lambda_3(\boldsymbol{W}) = 0;$
   $\boldsymbol{v}_1 \propto (4, 1, 1), \boldsymbol{v}_2 \propto (-1, 2, 2), \boldsymbol{v}_3 \propto (0, 1, -1).$
4. $\rho = 0.4 \approx \lambda_1(\boldsymbol{W}) >> \lambda_2(\boldsymbol{W})$, $\eta = 0.35$.

Consider two scenarios of cross-sectional shocks

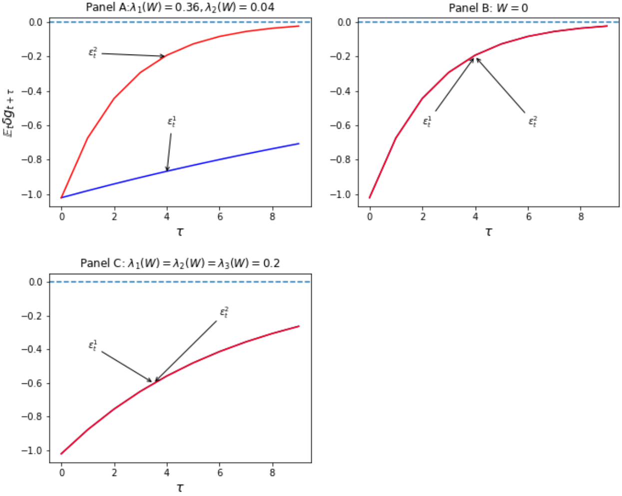
- Scenario 1: $\boldsymbol{\epsilon}_{\boldsymbol{t}}^1 = (-0.2, -0.05, -0.05) \propto \boldsymbol{v}_1$
- Scenario 2: $\boldsymbol{\epsilon}_{\boldsymbol{t}}^2 = (0.1, -0.2, -0.2) \propto \boldsymbol{v}_2$

The aggregate effects on the current growth are the same since

$$\delta g_t = \frac{1}{1-\eta}(\boldsymbol{s}_t, \boldsymbol{\epsilon}^1) = \frac{1}{1-\eta}(\boldsymbol{s}_t, \boldsymbol{\epsilon}^2) = -0.1$$

# An illustrative example - structure meets shock

How shocks interact with the innovation network.



We estimate the innovation network as a state space model:

$$\begin{aligned}\Delta \boldsymbol{a}_t &= \boldsymbol{\mu}_t + \boldsymbol{\epsilon}_t^A \\ \boldsymbol{\mu}_{t+1} &= (1-\rho)\boldsymbol{\mu}_t + \boldsymbol{W}\varphi_A(L)\Delta \boldsymbol{a}_t + \boldsymbol{\epsilon}_t^u\end{aligned} \tag{17}$$

- Observable: $\Delta \boldsymbol{a}_t = \Delta \log(\boldsymbol{A_t})$ using Patent filing.
- Latent process: arrival rate $\boldsymbol{\mu}_t$.
- Estimated with the Expectation-Maximization (EM) algorithm in the state space model.

Parameters: $\Theta = (\rho, \boldsymbol{W}, \varphi_A(L), \Sigma_A, \Sigma_u)$. Without restrictions,

$$1 + J^2 + L + J(J+1) = 2J^2 + J + 1 + L$$

parameters, where $J$ : # sectors = 89 and $L$ : # lags in $\varphi_A(L)$.

Table: Model Setup and Restrictive Assumptions

| Model Parameter | Restrictions |
|---|---|
| $\Sigma_A$ | $\Sigma_A = \sigma_A^2 \boldsymbol{I}$ |
| $\Sigma_u$ | $\Sigma_u = \sigma_u^2 \boldsymbol{I}$ |
| $\varphi(L) = \sum_{j \geq 0} \varphi_j L^j$ | $\varphi(L) = \varphi_A \sum_{j \geq 0} (1 - \varphi_A)^j L^j$ with $\varphi_A = 0.05, 0.1, 0.2, 1$ |
| $\boldsymbol{W}_t = \Xi \tilde{\boldsymbol{W}}_t$, $\Xi = \text{diag}(\xi_1, ..., \xi_J)$ | $\tilde{\boldsymbol{W}}_t$ directly estimated with patent citations |

We examine:

1. Persistence: $1 - \rho + \lambda_1(\boldsymbol{W}) \approx 1$.
2. Distribution of $\lambda_k(\boldsymbol{W}), k \leq J$.
3. Amplification: $(\boldsymbol{s}_t, \boldsymbol{v}_k), (\epsilon_t, \boldsymbol{v}_k), k \leq J$.

1. Measure $\Delta \boldsymbol{a}_t$: we interpret $A_{it}$ as the innovation-driven productivity and proxy it using the patent stock of the sector.
   1.1 $N_{it}$: # of patents issued by firms in sector $i$ in year $t$, we proxy for $A_{it}$ as

$$A_{it} = \delta_A \sum_{s \geq 0} (1 - \delta_A)^s N_{it-s}$$

   We choose $\delta_A = 0.05$, the value of the patent on production depreciate to zero after 20 years.

   1.2 $\Delta a_{it} = \log(A_{it}) - \log(A_{it-1})$
2. Measure $\boldsymbol{W}$: we write

$$\boldsymbol{W} = \Xi \tilde{\boldsymbol{W}}$$

with $\sum_j \tilde{W}_{ij} = 1, \Xi = \text{diag}(\xi_1, ..., \xi_J)$, $\xi_i = \sum_j W_{ij}$. $\tilde{W}$ is estimated based on page 10

Check: $1-\rho+\lambda_1(\boldsymbol{W})\approx 1$

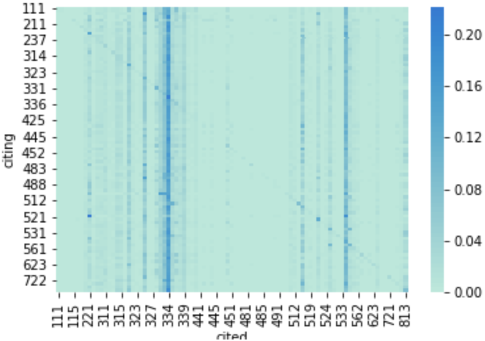
Table: EM Estimates of the Innovation Networks

| Panel A: EM estimates with assumption that $\Xi=\xi\boldsymbol{I}$ | | | | |
|---|---|---|---|---|
| | $\varphi_A=0.05$ | $\varphi_A=0.1$ | $\varphi_A=0.2$ | $\varphi_A=1.0$ |
| $1-\rho$ | 0.823 | 0.818 | 0.814 | 0.817 |
| $\lambda_1$ | 0.163 | 0.149 | 0.140 | 0.130 |
| $\sigma_u$ | 0.0322 | 0.0324 | 0.0325 | 0.0316 |
| $\sigma_A$ | 0.0237 | 0.0235 | 0.0234 | 0.0243 |
| $1-\rho+\lambda_1$ | 0.986 | 0.967 | 0.958 | 0.947 |
| **Panel B: EM estimates with general $\Xi$** | | | | |
| | $\varphi_A=0.05$ | $\varphi_A=0.1$ | $\varphi_A=0.2$ | $\varphi_A=1.0$ |
| $1-\rho$ | 0.791 | 0.779 | 0.770 | 0.780 |
| Mean of $\xi_j,\ j\in[J]$ | 0.198 | 0.187 | 0.182 | 0.162 |
| Standard Dev of $\xi_j,\ j\in[J]$ | 0.160 | 0.141 | 0.131 | 0.113 |
| 25th percentile of $\xi_j,\ j\in[J]$ | 0.099 | 0.090 | 0.093 | 0.095 |
| 75th percentile of $\xi_j,\ j\in[J]$ | 0.298 | 0.264 | 0.266 | 0.238 |
| $\sigma_u$ | 0.0332 | 0.0336 | 0.0339 | 0.0324 |
| $\sigma_A$ | 0.0227 | 0.0222 | 0.0220 | 0.0235 |
| $1-\rho+\xi(\xi=$ Mean of $\xi_j,\ j\in[J])$ | 0.989 | 0.966 | 0.952 | 0.942 |

This table presents the parameter estimates using EM algorithm (for details, please see the appendix). In Panel A, we impose an assumption that $\Xi=\xi\boldsymbol{I}$ - all sectors share the same parameter $\xi$. In Panel B, we remove this restriction and allow for heterogeneity in $\xi$ across sectors. For the general case that $\Xi=diag(\xi_1,\ldots,\xi_J)$ in Panel B, we also report the mean, standard deviation, 25th and 75th percentiles. In both panels, columns 1-4 report the results with $\varphi_A=0.05$, 0.1,0.2, and 1.0.

# Sparsity of Innovation Network (year = 2014)

X-axis: sectors with knowledge flow out, Y-axis: sectors learning from others. Some sectors are dominant in generating new knowledge, like sector 334 (Computer and Electronic Product Manufacturing) and sector 541 (Professional, Scientific, and Technical Services).

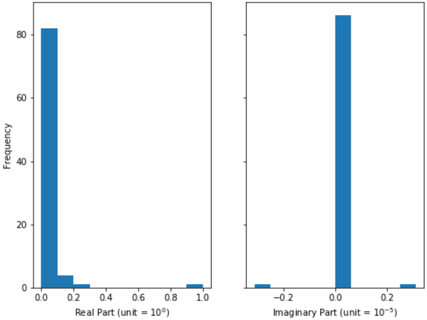


# Low Rank of Innovation Netowork (year = 2014)

The innovation network is low rank:

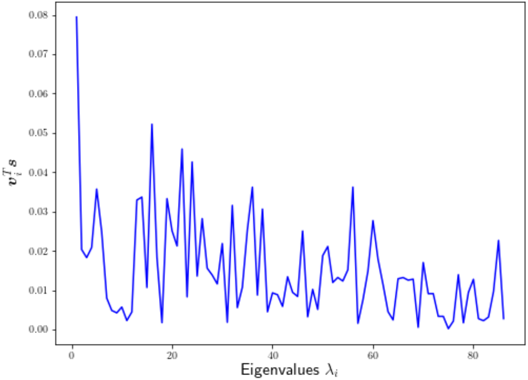
1. largest eigenvalue is much larger than the others in magnitude;
2. eigenvalues are approximately real, the imaginary parts are negligible.

Figure: Distribution of Eigenvalues of $\lambda_k(\tilde{\boldsymbol{W}})$ (year = 2014)



Remember, amplification of the $k^{th}$ component depends on

1. $(\boldsymbol{s}_t, \boldsymbol{v}_k)$: we find that $(\boldsymbol{s}_t, \boldsymbol{v}_1) >> (\boldsymbol{s}_t, \boldsymbol{v}_k), k \geq 2$
2. $(\boldsymbol{v}_k, \boldsymbol{\epsilon_t})$

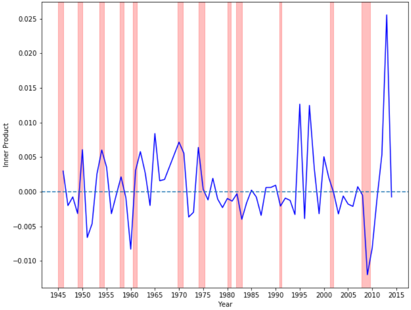


# Correlation between $\boldsymbol{v}_1$ and $\epsilon_t$

Remember, amplification of the $k^{th}$ component depends on

1. $(\boldsymbol{s}_t, \boldsymbol{v}_k)$
2. $(\boldsymbol{v}_k, \boldsymbol{\epsilon_t})$

Figure: Correlation between $\boldsymbol{v}_1$ and $\epsilon_t$(Patent data)

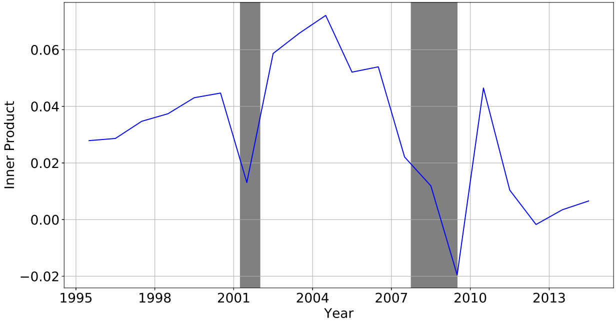


How about the $(\boldsymbol{v}_1, \boldsymbol{\epsilon_t})$? if we use sectoral TFP data rather than patents. We write

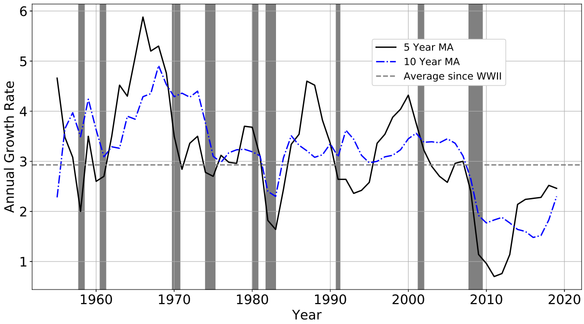
$$\log(TFP_{it}) = a_{it} + m_{it} + e_{it} \tag{18}$$

$a_{it}$: productivity driven by technology; $m_{it}$: productivity driven beyond technology, follow AR(1); $e_{it}$: measure error.

Figure: Correlation between $\boldsymbol{v}_1$ and $\epsilon_t$(TFP data)



How long the economy takes to recover from recessions in U.S. ?



# Conclusion

A model where innovation network marries production network:

1. Theoretically, the shock direction and inn-network structure matters in amplification and persistence:
   1.1 Persistence: captured by the structure of the inn-network, $\rho - \lambda_k(\boldsymbol{W})$
   1.2 When the inn-network is low-rank, the sectoral distribution of the shock reveals useful information on future recover process.
   1.3 Amplification: captured by $(\boldsymbol{v}_k, \boldsymbol{s}_t), (\boldsymbol{\epsilon}_t, \boldsymbol{v}_k)$
2. Empirically, we show
   2.1 Persistence: $\rho \approx \lambda_1(\boldsymbol{W})$, the shock becomes very persistent when the shock is parallel to the eigenvector centrality of the inn-network.
   2.2 The inn-network is low-rank for U.S.
   2.3 Amplification: $(\boldsymbol{v}_1, \boldsymbol{s}_t) >> (\boldsymbol{v}_1, \boldsymbol{s}_t), k \geq 2$, $(\boldsymbol{\epsilon}_t, \boldsymbol{v}_1)$ is much lower in Great Recession than others.
3. Policy implication: to avoid long persistent recession, policy should target at firms in the center of the innovation network.
4. Future work:
   4.1 Endogenize the long-run risk in networks - puzzles in asset pricing.
   4.2 General non-linear effect due to endogenized R&D.
   4.3 What is the implication of Covid-19 on persistent recession?