Factor Augmented Network Vector Autoregression (FANVAR)

A. Furkan [1] J. Miranda-Pinto [2] J. Morley [3] V. Panchenko [4] C. Rose [5]

[1]Macquarie U. [2]IMF and UQ [3]U. of Sydney [4]UNSW [5]UQ

CEF, Santiago, Chile
July 8, 2025

Disclaimer: The views expressed herein are those of the authors and not necessarily those of the International Monetary Fund.

Motivation

- Research on networks in economics (macro/trade) has exploded recently
  - Networks considered: input-output networks, investment network, trade credit networks, banking networks, cross-country trade/financial networks

- Two main approaches to studying the importance of network structures

1. **Empirical**: VAR or FAVAR-like models to describe/decompose/forecast
   - Diebold and Yilmaz (2014), Mlikota (2023), Barigozzi et al. (2023), Grant and Yung (2024), Miao et al. (2023)
2. **Structural**: Quantitative multisector business cycle models with production networks
   - Horvath (1998, 2000), Foerester, Sarte and Watson (2011), Atalay (2017), Herskovic (2018), Pasten, Schoenle, and Weber (2020), vom Lehn and Winberry (2022)

We propose a semi-structural approach to jointly estimate spillovers—from **observed** networks—and **unobserved** common factor(s).

What We Do

- **Estimate the following model:**

$$y_{i,t} = \sum_{s=1}^{P} \sum_{j=1}^{N} \mathbf{A}_{s,ij} y_{j,t-s} + \Lambda_i \mathbf{F}_t + u_{i,t}.$$

This is a high-dimensional problem. Estimating N^2 spillover coefficients, R factors and RxN factor loadings.

Solution: guided by theory, use observed networks (B_t) to guide the estimation of spillovers $\mathbf{A}_{s,ij}$. Specifically, define $A = \beta B_t$. Estimate β (could be scalar or a vector) using observed B_t.

Applications: Today: Sectoral price dynamics and the role of input-output networks. In progress, role of financial linkages and/or trade linkages in cross-country GDP dynamics

What We Find

- From Montecarlo exercises, FANVAR provides a significant improvement with respect to VAR and Factor models
  - VAR models, omitting factors, overstate the role of spillovers
  - Factor models overstate the role of common factors in the presence of networks
- Application to US PPI inflation. Compare FANVAR to FAVAR from Boivin, Giannoni, and Mihov (2009)
  - What Boivin et al. (2009) characterized as sluggish response to aggregate shocks, in our approach are sluggish sectoral spillovers (Minton and Wheaton, 2024)
  - Using more recent sample, we show that spillovers played a key role during COVID-19

Contribution to the Literature

- **Methodological**: Factor models (Geweke, 1977; and Sargent and Sims, 1977), Vector Autoregression (VAR) models (Sims, 1980), Factor-VAR (Bernanke, Boivin, and Eliasz, 2005), Panel Models with Interactive FE (Moon and Weidner, 2023), High-dimensional VAR with factors (Mlikota, 2023; Barigozzi et al., 2023, Maio, Phillips, Su, 2023)
  - Our approach uses network information to our disposal to study spillovers while accounting for common factors
- **Semi-structural applications**: Causes of the Great Recession (Altinoglu, 2019; Li and Martin, 2019); Reduction in comovement post-1984 (Foerster, Sarte, and Watson, 2011, Garin et al., 2019, vom Lehm and Wimberry, 2022)
  - Our approach can accommodate many networks (or models) and freely estimate their relative importance

Our approach

Model

Let the dynamics of unit i at time t (e.g., firm i's output growth or country i's inflation rate) be expressed as

$$y_{i,t} = \sum_s \sum_j A_{s,ij} y_{j,t-s} + \Pi_t + u_{i,t},$$

$$\Pi_{it} = \Lambda_i \mathbf{F}_t$$

- $\sum_s \sum_j A_{s,ij} y_{j,t-s}$ accounts for dynamic spillovers between *units*
- $\Pi_{it} = \Lambda_i \mathbf{F}_t$ is the unobserved common factor structure
- u_{it} are unit-specific idiosyncratic *errors*, which are orthogonal to $y_{j,t-s}$ and Π_t. u_{it} can be mildly correlated across units.

Estimating Network Spillovers and Common Factors

Moon and Weidner (2023) estimator applies when $\mathbf{A}$ is known up to a fixed number of unknown constants (e.g., $\mathbf{A} = \beta\mathbf{B}$) where β captures the intensity of spillovers.

Minimize the convex objective function with respect to β and Π,

$$argmin_{\beta,\Pi}\left(\underbrace{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{i,t} - \sum_{s=1}^{P}\sum_{j=1}^{N}\mathbf{A_{s,ij}}y_{j,t-s} - \Pi_{it}\right)^2}_{LS} + \underbrace{\lambda p(\Pi)}_{Rank}\right),$$

$$A_{s,ij} = \sum_{k=1}^{K}\sum_{l=1}^{K}\mathbf{1}(k(i)=k)\mathbf{1}(k(j)=l)\quad\underbrace{\beta_{kl}}_{Unobserved}\quad\cdot\quad\underbrace{B_{s,ij}}_{Observed},$$

where $\lambda \geq 0$ is a tuning (penalty) parameter selected in a data driven manner and $p(\Pi)$ is the sum of the singular values of Π. K is the number of clusters assumed to estimate the spillover coefficients

Practical implementation

- Using *pilot* OLS estimate of A choose the optimal number of factors R
- Iterative algorithm
  - given R, estimate the common factor structure using PCA
  - Use the factors to estimate the spillover coefficients
- Moon and Weidner (2023) establish $\sqrt{NT}$ consistency of β using this iterative estimation procedure

Montecarlo Simulations

The Montecarlo Experiment

The data-generating process follows

$$y_{i,t} = \sum_j A_{1,ij} y_{j,t-1} + \Pi_t + u_{i,t},$$

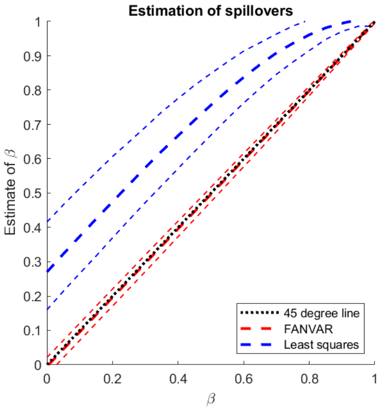
where the true network $A = \beta B$ in which B is drawn from a Bernoulli distribution (for each sample t), Erdos-Renyi with $p = log(N)/N$. We assume there is one common factor. We set $N = 100$, $T = 100$.

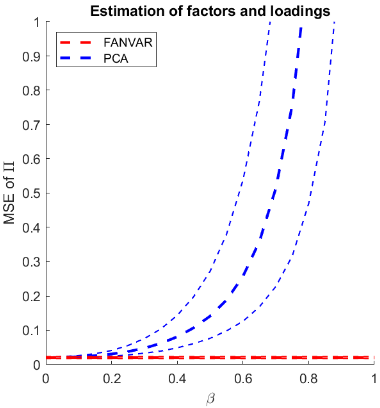
We vary the degree to which the observed network B generates spillovers (β). Then, using B, we estimate β and Π_t

We also estimate β assuming we do not fully observe B

True vs Estimated Spillover

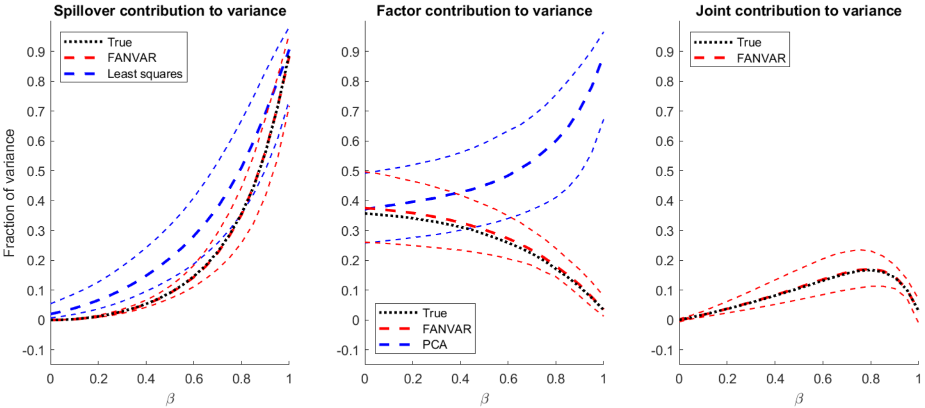
$$y_{i,t} = \sum_s \sum_j \beta B_{s,ij} y_{j,t-s} + \Pi_t + u_{i,t},$$



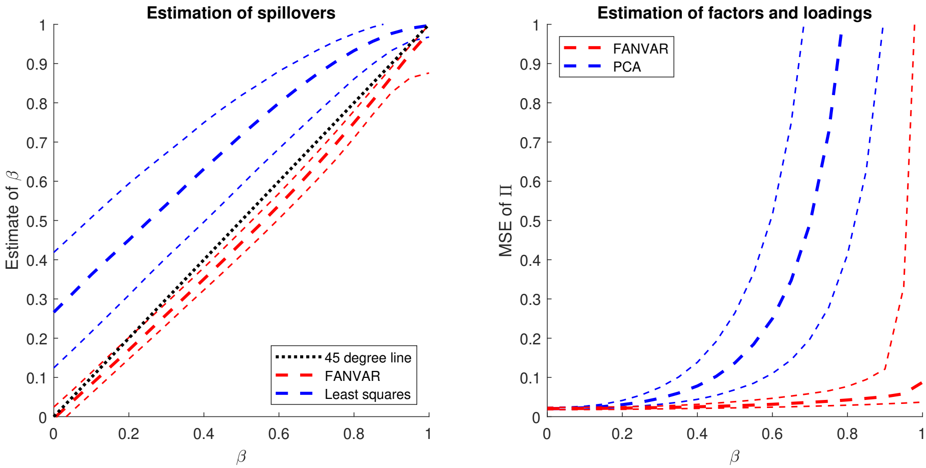


True vs Estimated Spillovers

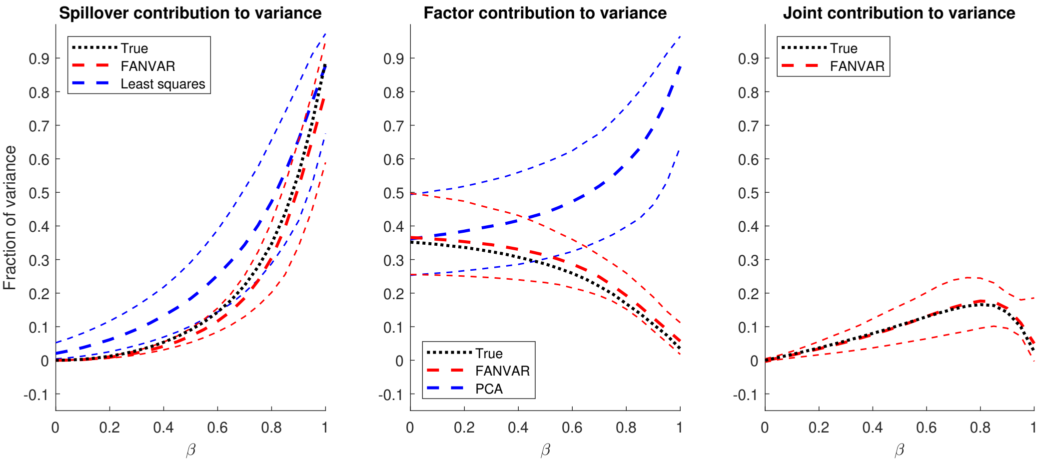
$$y_{i,t} = \sum_j \beta B_{1,ij} y_{j,t-1} + \Pi_t + u_{i,t},$$



True vs Estimated Spillover if B not fully observed (only 90 % of links)



True vs Estimated Spillover if B not fully observed (only 90 % of links)



Inference and Forecasting

Impulse Response Analysis

Rewrite the system and provide more structure to the factors

$$
\begin{aligned}
Y_t =& \Phi Y_{t-1} + \Lambda F_t + v_t, \\
F_t =& \psi(L) F_{t-1} + \Omega W_t + u_t,
\end{aligned}
$$

where W_t is a set of observable macro controls (potential factors). Now we can obtain the IRFs from aggregate shocks to W, which are observed factors such as US monetary policy, oil prices or global financial conditions.

$$
\frac{\partial Y_t}{\partial W_{it-K}} = \left(\sum_{j=0}^{K} \Phi^{K-j} \Lambda \psi^j \right) \Omega_i.
$$

IRF to an aggregate shock W

Figure: Impulse Response to aggregate shock: different underlying network

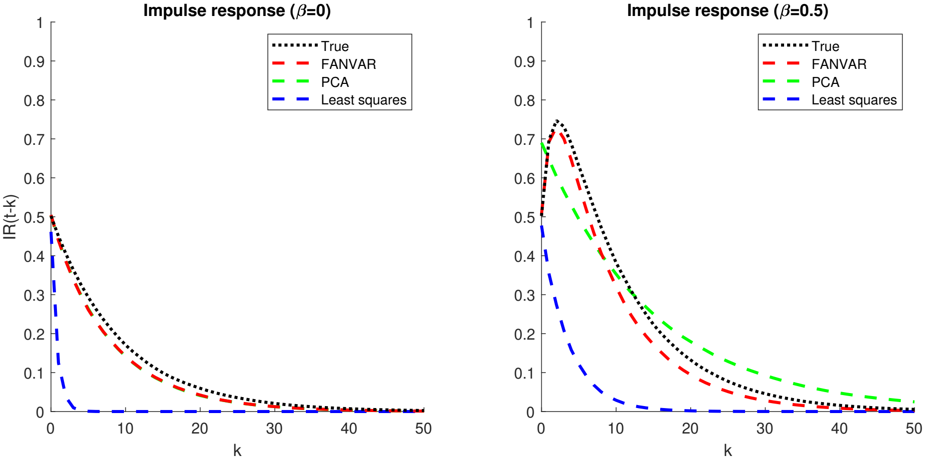


Figure: *

Application to sectoral inflation dynamics

FANVAR vs. FAVAR

- Bernanke, Boivin, and Eliasz (2005) proposed the Factor Augmented Vector Autoregression (FAVAR)
- FAVAR augments the information of a VAR using pre-estimated factor(s) from large disaggregated time-series data
- Boivin, Giannoni, and Mihov (2009), and many others recently, have used it to understand inflation dynamics and monetary policy transmission

Data

- 154 sectoral PPI series for the US
    - Sample 1: 1970m1-2005m6 (original Boivin et al., 2009)
    - Sample 2: 1990m1-2024m6 (our new data)
- Match these series to IO tables to obtain our measure of observed linkages

VAR (OLS) vs FANVAR

- Assume $\mathbf{A} = \beta \mathbf{B}^\top$ where $\mathbf{B}$ is the observed IO matrix.

$$d \ln Y_t = \beta \mathbf{B}^\top d \ln Y_{t-1} + \Pi_t + \mathbf{U}_t,$$

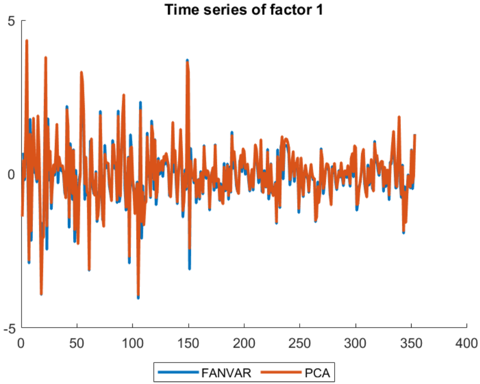
- Estimates of β (one common factor chosen)

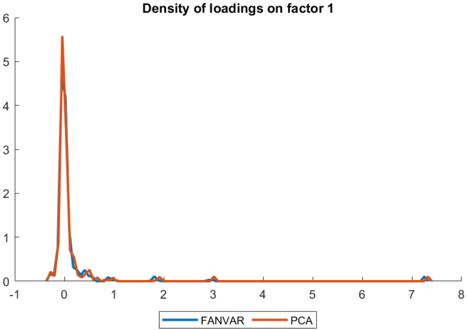
	OLS	M+W
β	0.71	0.69
	(0.021)	(0.021)
Obs	54,362	54,362

- Results suggest that dynamic price spillovers are 70% of those implied by the IO network (generally omitted, even in the theoretical/quantitative literature on production networks)

- $\beta = 0.11$ if use Leontief inverse for B, meaning delayed indirect effect might take more than one period to materialize

Factor Models vs FANVAR





Decomposing the Sources of Sectoral Inflation

	Standard deviation (in percent)			
	Inflation	Common Factors	Sector-Specific	Spillovers
FAVAR				
PPI Average	1.36	0.38	1.30	-
Median	0.92	0.31	0.88	-
Standard deviation	1.16	0.21	1.15	-
FANVAR				
PPI Average	1.36	0.17	1.28	0.23
Median	0.92	0.05	0.90	0.14
Standard deviation	1.16	0.66	0.99	0.25

Note: this table reports moments of sectoral inflation as well as its different components (common factors, idiosyncratic shocks, and spillovers). The FAVAR results are those from Boivin et al. (2009). The FANVAR results are obtained using our approach using the same PPI series and the US input-output tables in 2002 to inform the existence of network effects.

Sample 1990m1-2024m6

Update PPI series and IO tables. Re estimate:

$$d \ln Y_t = \beta \mathbf{B}^\top d \ln Y_{t-1} + \Pi_t + \mathbf{U}_t,$$

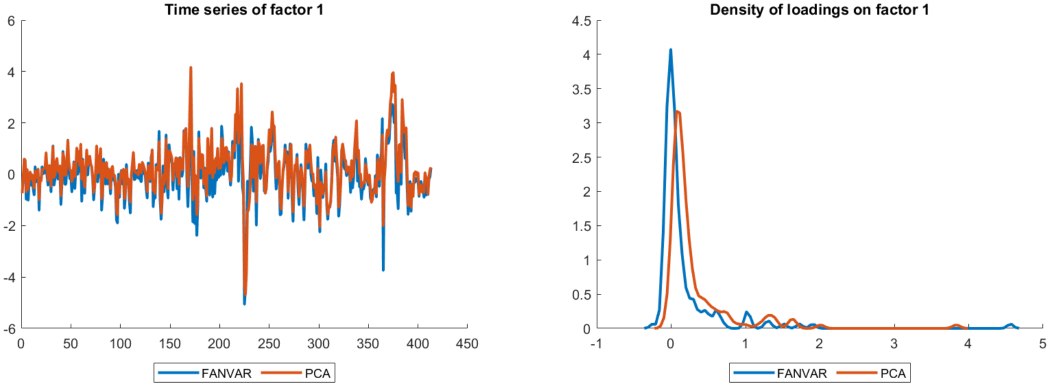
Estimates of β (one common factor chosen)

	OLS	M+W
β	0.93	0.77
	(0.017)	(0.019)
Obs	63,602	63,602

Increased importance of spillovers, increasing the overestimation from OLS.

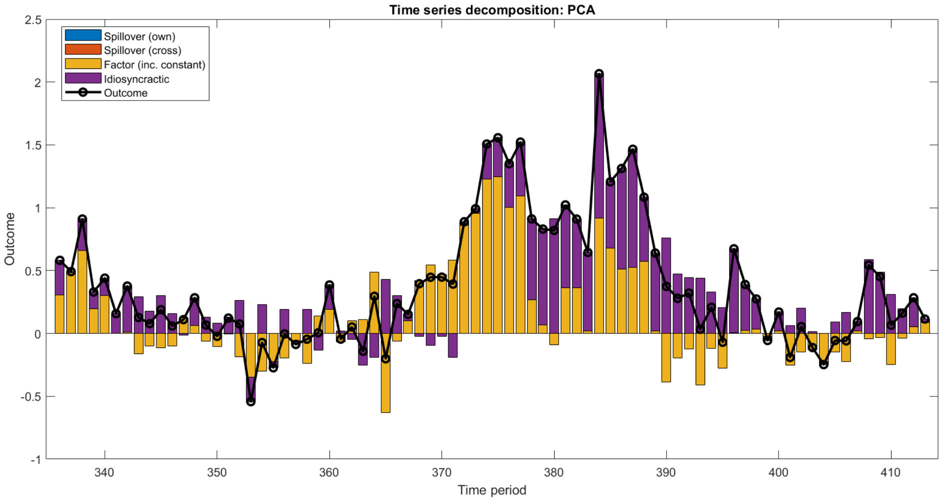
Using Leontief for B gives 0.17 ($>$0.11).

Factor Models vs FANVAR



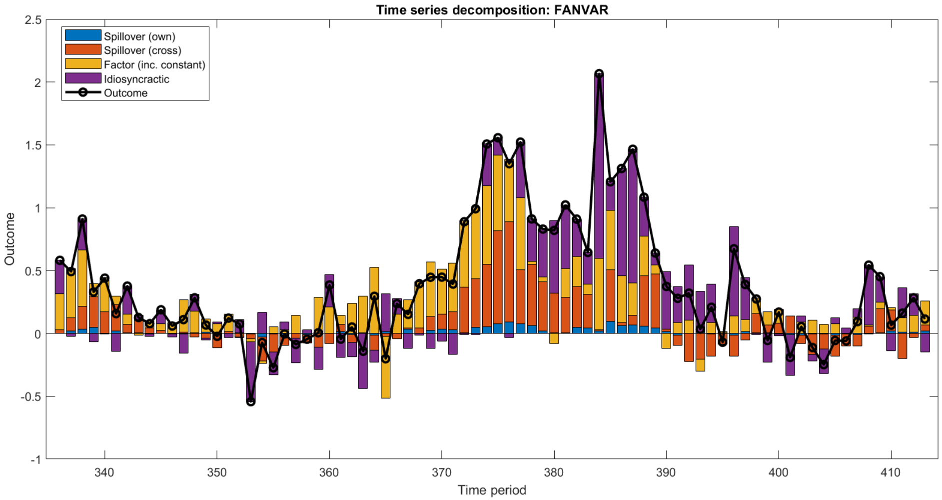
Factor models overstate the role of common shocks (factor loadings)

Inflation during 2018-2024: factor models



Note: This figure plots the Factor Model decomposition of average PPI for the period 2018m01-2024m6.

Inflation during 2018-2024: FANVAR

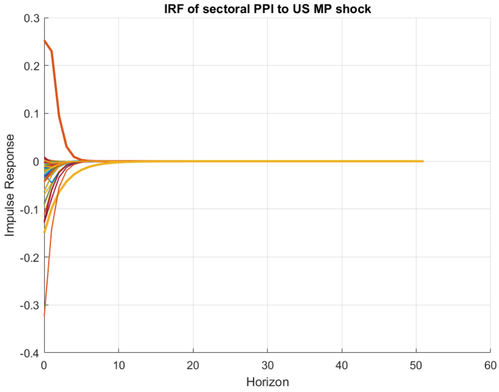


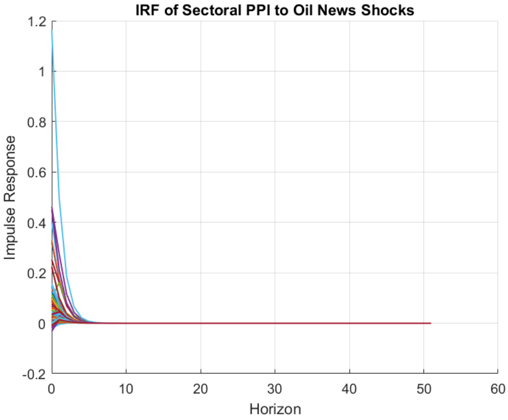
Note: This figure plots the FANVAR decomposition of average PPI (monthly) growth for the period 2018m06-2024m6.

Sectoral IRFs from Monetary Policy and Oil price shock

Using externally identified shocks to monetary policy and oil prices

$$\frac{\partial Y_t}{\partial W_{it-K}} = \left(\sum_{j=0}^{K} \Phi^{K-j} \Lambda \psi^j \right) \Omega_i$$





Conclusion

- We propose an econometric approach that jointly identifies general spillovers, using observed networks and unobserved common factor(s)
- Useful to understand the sources of macroeconomic fluctuations (prices, GDP, etc) from disaggregated and high-dimensional data
- Applied to US sectoral inflation, our approach highlights the relevance of linkages and dynamic spillovers

Thank you

Variance decomposition from PCA, OLS and joint estimator

	PCA	**OLS**	**Joint**
Common Factors	58%	NA	36%
Idiosyncratic shocks	42%	68%	34%
Spillovers	NA	32%	30%

Application II:
Common factor vs linkages in commodity prices

Commodity price comovement

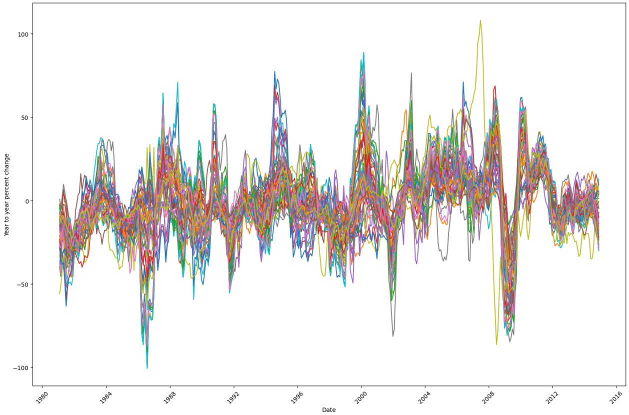


Figure: Monthly real commodity price index for 60 EMEs, Fernández, González, and Rodriguez (2018)

Algeria, Argentina, Australia, Austria, Belarus, Bolivia, Botswana, Brazil, Bulgaria, Cameroon, Canada, Chile, Colombia, Costa Rica, Ivory Coast (Côte d'Ivoire), Croatia, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Gabon, Georgia, Ghana, Guatemala, Honduras, India, Indonesia, Iran, Jamaica, Jordan, Kazakhstan, Kuwait, Lithuania, Malaysia, Mexico, Mongolia, Morocco, New Zealand, Niger, Nigeria, Norway, Oman, Pakistan, Panama, Paraguay, Peru, Philippines, Russia, Saudi Arabia, South Africa, Thailand, Trinidad and Tobago, Tunisia, Ukraine, United Arab Emirates, Uruguay, Venezuela, Vietnam.

Spillovers in commodity price fluctuations

Use three versions of observed linkages based on bilateral trade linkages from the Direction of Trade Statistics (DOTS)

Table: Importance of common factors in cross-country commodity price indexes

Model	**Factors**	**Idiosync.**	**Spillovers**
PCA	77%	23%	NA
Joint using M_1	49%	23%	28%
Joint using M_2	49%	23%	28%
Joint using M_3	42%	16%	42%

Note: M_1 uses trade linkages between country i and country j as follows. M_{ij} represents the observed links between country i and country j. In particular, $M_{1_{ij}}$ is the ratio of the sum between exports from i to j and imports of i from j ($\text{exports}_{ij} + \text{imports}_{ij}$) and total trade of countries i and j. $M_{2_{ij}}$ is the ratio of the sum between exports from i to j and imports of i from j ($\text{exports}_{ij} + \text{imports}_{ij}$) and the GDP of country i. $M_{3_{ij}}$ is the ratio of the sum between exports from i to j and imports of i from j ($\text{exports}_{ij} + \text{imports}_{ij}$) and the GDP of country j.

Application III:

Drivers of global inflation

Drivers of global inflation

$$i_t = \mathbf{A}^T i_{t-1} + \Pi_t + \mathbf{U}_t$$

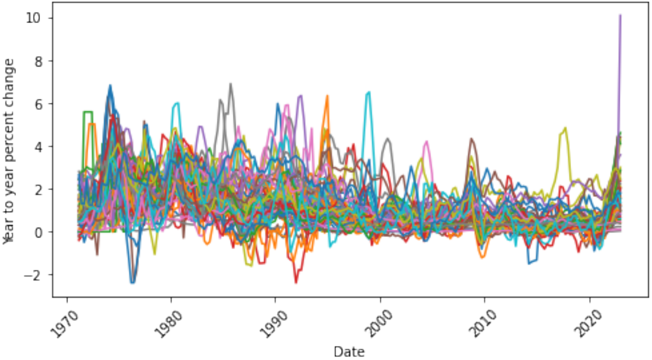


Figure: Quarterly headline CPI inflation 60 countries from Ha, Kose, and Ohnsorge. (2023)

Countries: Argentina, Australia, Austria, Burundi, Belgium, Burkina Faso, Bahamas, Bolivia, Canada, Switzerland, China, Ivory Coast (Côte d'Ivoire), Cameroon, Colombia, Cyprus, Germany, Denmark, Dominican Republic, Ecuador, Egypt, Spain, Finland, Fiji, France, Gabon, United Kingdom, Greece, Guatemala, Honduras, Haiti, Indonesia, India, Ireland, Iceland, Italy, Jamaica, Japan, South Korea, Luxembourg, Morocco, Mauritius, Malaysia, Niger, Netherlands, Norway, New Zealand, Pakistan, Peru, Philippines, Portugal, Paraguay, Singapore, El Salvador, Sweden, Thailand, Tunisia, Turkey, Tanzania, Uruguay, United States of America, Samoa, and South Africa.

Results with homogeneous spillovers: the UK

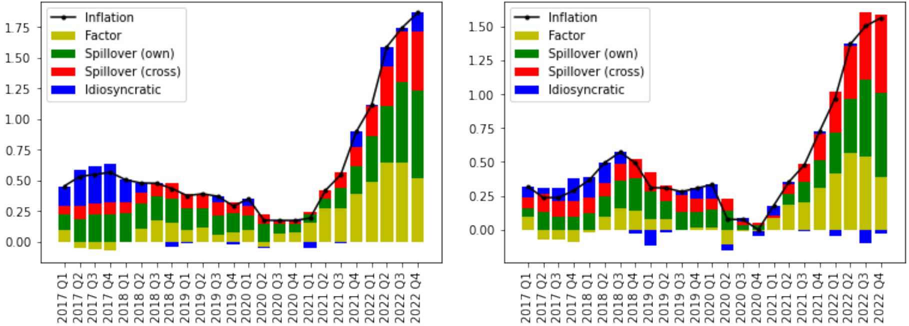


Figure: Decomposition of UK (left) and France (right) inflation

Estimate heterogeneous spillovers: USA-Europe-ROW

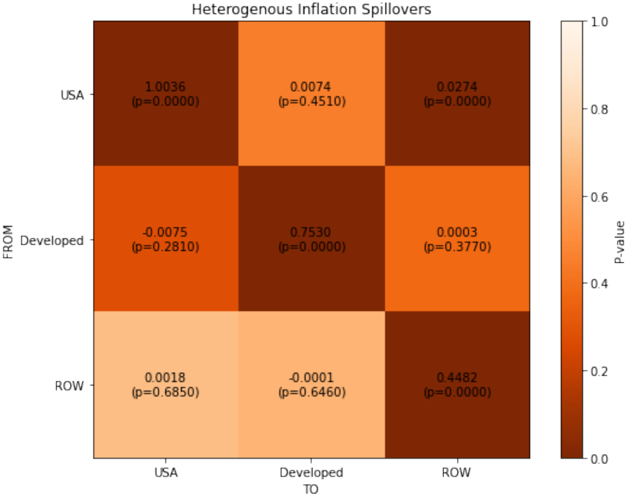


Figure: Heterogeneous spillovers

Results with heterogeneous spillovers: the case of Korea

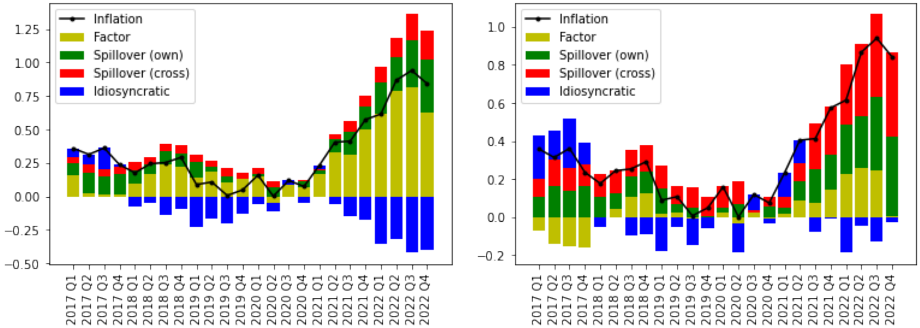


Figure: Decomposition of Korea's inflation. Homogeneous (left) and heterogeneous (right) spillovers

Counterfactual analysis

Suppose we have our estimated model

$$\mathbf{y_t} = \widehat{\mathbf{A}}\mathbf{y_{t-1}} + \hat{\Pi}_t + \hat{\mathbf{U}}_t,$$

We can analyse responses to

- another network structure $\widetilde{\mathbf{A}}$ that could prevail
- or the propagation of an idiosyncratic shock (e.g., wildfires in Canada) or a common shock (e.g., COVID-19). In this case we could use $\widehat{\mathbf{A}}$ to construct impulse response functions arising from some counterfactual shocks $\widetilde{\mathbf{U}}_t, ..., \widetilde{\mathbf{U}}_s$

.

A Well-Known Structural Model: Carvalho (2007)

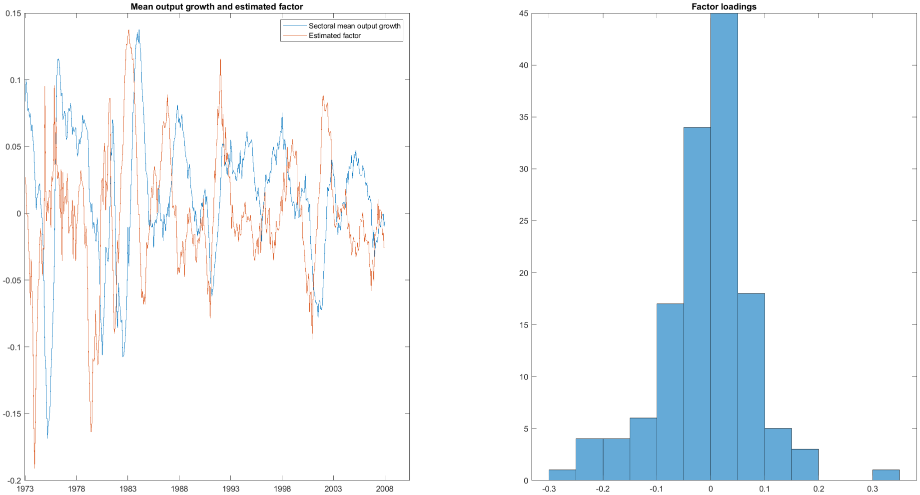
There is labor, capital, and intermediate inputs. Capital depreciates in one period and is produced using same sector's output. We have

$$d \ln Y_t = (I - \mathbf{A})^{-1} \alpha_d \ln Y_{t-1} + (I - \mathbf{A})^{-1} \epsilon_t,$$

where α_d is a matrix with sectoral capital shares in its diagonal and zero otherwise.

Back

Moon and Weidner (2023): First estimated factor and loadings



Practical implementation

Estimate number of factors R (assume $s = 1$ for exposition only)

$$\hat{R} = \sum_{r=1}^{\min(N,T)} \mathbf{1}\left\{s_r\left(\mathbf{Y}_t - \hat{\mathbf{A}}^*\mathbf{Y}_{t-1}\right) > \hat{\lambda}\right\}$$

where $\hat{\lambda}$ is a thresholding term. The function $s_r(\cdot)$ returns the r^{th} singular value of a matrix and $\hat{\mathbf{A}}^*$ is based on the pilot estimator $\hat{\beta}^*$ is defined by:

$$\hat{\beta}^* = \arg\min_{\beta \in \mathbb{R}} \left\|\mathbf{Y}_t - \mathbf{A}\mathbf{Y}_{t-1}\right\|_1,$$

where $\|.\|_1$ is the nuclear norm.

Practical implementation

The penalty term $(\hat{\lambda})$ is chosen as follows:

1. Choose an upper threshold for the R: R_{max}.
2. Given the estimator $\hat{\beta}^*$,
$$(\hat{\Lambda}, \hat{\mathbf{F}}) \in \underset{\Lambda \in \mathbb{R}^{N \times R_{max}}, \mathbf{F} \in \mathbb{R}^{T \times R_{max}}}{\arg\min} \left\| \mathbf{Y}_t - \hat{\mathbf{A}}^* \mathbf{Y}_{t-1} - \Lambda \mathbf{F}^T \right\|_2^2 .$$
3. Compute residuals,
$$\hat{\mathbf{E}}_* = \mathbf{Y}_t - \hat{\mathbf{A}}^* \mathbf{Y}_{t-1} - \hat{\Lambda} \hat{\mathbf{F}}^T.$$
4. Choose,
$$\hat{\lambda} = K \| \hat{\mathbf{E}}_* \|_\infty, K > 2 \,(\text{use } 4).$$

Practical implementation

Once we have estimated the number of factors (R_0) following the steps above, the next step is an iterative approximation algorithm taking the number of factors as known:

1. Set $\hat{\beta}^{(s)} = \hat{\beta}^*$ for $s = 0$.
2. Estimate the common factor structure by the principal component method given $\mathbf{Y}_t - \hat{\mathbf{A}}^{(s)}\mathbf{Y}_{t-1}$:

$$\left(\hat{\Lambda}^{(s+1)}, \hat{\mathbf{F}}^{(s+1)}\right) \in \underset{\Lambda \in \mathbb{R}^{N \times \hat{R}}, \mathbf{F} \in \mathbb{R}^{T \times \hat{R}}}{\arg\min} \left\| \mathbf{Y}_t - \hat{\mathbf{A}}^{(s)}\mathbf{Y}_{t-1} - \Lambda \mathbf{F}^T \right\|_2^2.$$

3. Estimate $\hat{\beta}^{(s+1)}$:

$$\hat{\beta}^{(s+1)} = \arg\min_{\beta} \min_{g \in \mathbb{R}^{T \times \hat{R}}, h \in \mathbb{R}^{N \times \hat{R}}} \left\| \mathbf{Y}_t - \mathbf{A}\mathbf{Y}_{t-1} - \hat{\Lambda}^{(s+1)} g^T + h \hat{\mathbf{F}}^{(s+1)T} \right\|_2^2.$$

MW show that this estimator ($\hat{\beta}^{(s)}$) converges to β_0 at a rate of $\sqrt{NT}$ under certain assumptions.

Estimating Unobserved Networks and Common Factors - High Dimensional

If $\mathbf{A}$ is not known up to a fixed number of unknown constants we have a high dimensional problem.

Propose to minimize a convex objective function such as the following with respect to $(A_s)_{s=1,\dots,P}$ and Π,

$$argmin_{A,\Pi}\left(\underbrace{\sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{i,t}-\sum_{s=1}^{P}\sum_{j=1}^{N}A_{s,ij}y_{j,t-s}-\Pi_{it}\right)^2}_{OLS}+\lambda_1\underbrace{\sum_{i=1}^{N}\sum_{j=1}^{N}p_1(i,j)}_{Sparsity}+\underbrace{\lambda_2 p_2(\Pi)}_{Rank}\right),$$

where $\lambda_1 \geq 0, \lambda_2 \geq 0$ are tuning parameters.

The sparsity function is $p_1(i,j)=\left(\sum_{s=1}^{P}\left(A_{s,ij}-A_{0,ij}\right)^2\right)^{1/2}$. The matrix $\mathbf{A}_0$ could be zeros or a prior network up to a fixed number of constants. Additional convex restrictions can be added to $\mathbf{A}$.

More complex economy

Assume VARMA (1,1) structure in Foerster, Sarte, and Watson (2011)

$$d \ln \mathbf{Y}_t = \varrho d \ln \mathbf{Y}_{t-1} + \Theta \mathbf{U}_{t-1} + \Pi_a \mathbf{U}_t,$$

where ϱ, Θ, and Π_a are functions of IO, investment network, factor shares, and consumption-capital policy rules. Assume that capital fully depreciates after a period to get

$$d \ln \mathbf{Y}_t = (I - \mathbf{A}^T)^{-1} \alpha_d \tilde{\Theta} d \ln \mathbf{Y}_{t-1} + (I - \mathbf{A}^T)^{-1} \epsilon_t,$$

More generally, we could express a VARMA(1,1) as a $VAR(\infty)$ in which

$$d \ln \mathbf{Y}_t = \sum_l^{\infty} \Xi_i d \ln \mathbf{Y}_{t-l} + \mathbf{U}_t,$$

Back