Identification and Estimation of Structural Factor Models with External Instruments

Ze-Yu Zhong[†]

[†]Monash University, Department of Econometrics and Business Statistics

May 2025

We acknowledge the financial support of the Australian Research Council under Grants LP160101038, DP210101440, and DE200100693.

Introduction

Structural Vector Autoregressions (SVARs) are popular workhorse models, but remain plagued by issues

- **Non-fundamentalness**: not enough variables to span, and therefore recover, primitive shocks
- **Singular-covariance**: no. of variables $>$ no. of primitive shocks
- **Identification Validity**: how to conduct valid inference?

Introduction

Structural Vector Autoregressions (SVARs) are popular workhorse models, but remain plagued by issues

- **Non-fundamentalness**: not enough variables to span, and therefore recover, primitive shocks
- **Singular-covariance**: no. of variables $>$ no. of primitive shocks
- **Identification Validity**: how to conduct valid inference?

Three issues form an **Iron Triangle** - solving one typically worsens the other:

- Non-fundamentalness: add more variables, but increases risk of singular-covariance
- Singular-covariance: reduce number of variables, or complex dimensional reduction
- Unclear interactions with identification strategy

Introduction - Existing Tools and Structural Factor Models

Non-fundamentalness: SVARs $\Rightarrow$ Factor Augmented Vector Autoregressions (FAVARs)

- Factors to summarise large information sets, (Bernanke et al., 2005)
- Existence of factor structure implies that SVARs are generally non-fundamental, (Forni et al., 2009)

Introduction - Existing Tools and Structural Factor Models

Non-fundamentalness: SVARs $\Rightarrow$ Factor Augmented Vector Autoregressions (FAVARs)

- Factors to summarise large information sets, (Bernanke et al., 2005)
- Existence of factor structure implies that SVARs are generally non-fundamental, (Forni et al., 2009)

Covariance Singularity: FAVARs $\Rightarrow$ Structural Factor Models (SFMs)

- FAVARs typically induce singular covariance, at odds with foundations of dynamic factor model (DFMs), (Stock & Watson, 2005)
- SFMs are a promising generalisation that combine a DFM with an SVAR

Introduction - Existing Tools and Structural Factor Models

Non-fundamentalness: SVARs $\Rightarrow$ Factor Augmented Vector Autoregressions (FAVARs)

- Factors to summarise large information sets, (Bernanke et al., 2005)
- Existence of factor structure implies that SVARs are generally non-fundamental, (Forni et al., 2009)

Covariance Singularity: FAVARs $\Rightarrow$ Structural Factor Models (SFMs)

- FAVARs typically induce singular covariance, at odds with foundations of dynamic factor model (DFMs), (Stock & Watson, 2005)
- SFMs are a promising generalisation that combine a DFM with an SVAR

Identification and Inference:

- Claimed that SVAR identification schemes carry over to SFMs, (Stock & Watson, 2016)
  - No theoretical justification for algorithms provided
  - Formalisations are typically strategy-specific, e.g. fast-slow (Han, 2015, 2018, 2024)

Introduction - Existing Tools and Structural Factor Models

Non-fundamentalness: SVARs $\Rightarrow$ Factor Augmented Vector Autoregressions (FAVARs)

- Factors to summarise large information sets, (Bernanke et al., 2005)
- Existence of factor structure implies that SVARs are generally non-fundamental, (Forni et al., 2009)

Covariance Singularity: FAVARs $\Rightarrow$ Structural Factor Models (SFMs)

- FAVARs typically induce singular covariance, at odds with foundations of dynamic factor model (DFMs), (Stock & Watson, 2005)
- SFMs are a promising generalisation that combine a DFM with an SVAR

Identification and Inference:

- Claimed that SVAR identification schemes carry over to SFMs, (Stock & Watson, 2016)
  - No theoretical justification for algorithms provided
  - Formalisations are typically strategy-specific, e.g. fast-slow (Han, 2015, 2018, 2024)

$\Rightarrow$ **Iron Triangle** remains unresolved even with SFMs

Contribution

Formalise Estimation and Inference of Structural Factor Models to resolve **Iron Triangle**:

Contribution

Formalise Estimation and Inference of Structural Factor Models to resolve **Iron Triangle**:

- **Non-fundamentalness**: factors summarise large datasets + detail nuisance parameters induced

Contribution

Formalise Estimation and Inference of Structural Factor Models to resolve **Iron Triangle**:

- **Non-fundamentalness**: factors summarise large datasets + detail nuisance parameters induced
- **Singular covariance**: allow for no. of static factors > no. of primitive shocks (not handled by FAVAR)

Contribution

Formalise Estimation and Inference of Structural Factor Models to resolve **Iron Triangle**:

- **Non-fundamentalness**: factors summarise large datasets + detail nuisance parameters induced
- **Singular covariance**: allow for no. of static factors > no. of primitive shocks (not handled by FAVAR)
- **Identification Validity**: external instruments using generalised method of moments (GMM)
  - Joint use of instruments, overidentification test, and instrument selection
- Compared with Local Projection, SFMs extend SVAR-IV, $\therefore$ less restrictive with instruments

Contribution

Formalise Estimation and Inference of Structural Factor Models to resolve **Iron Triangle**:

- **Non-fundamentalness**: factors summarise large datasets + detail nuisance parameters induced
- **Singular covariance**: allow for no. of static factors > no. of primitive shocks (not handled by FAVAR)
- **Identification Validity**: external instruments using generalised method of moments (GMM)
  - Joint use of instruments, overidentification test, and instrument selection
- Compared with Local Projection, SFMs extend SVAR-IV, $\therefore$ less restrictive with instruments

Monte Carlo study confirms the theoretical properties of the proposed estimator

Contribution

Formalise Estimation and Inference of Structural Factor Models to resolve **Iron Triangle**:

- **Non-fundamentalness**: factors summarise large datasets + detail nuisance parameters induced
- **Singular covariance**: allow for no. of static factors > no. of primitive shocks (not handled by FAVAR)
- **Identification Validity**: external instruments using generalised method of moments (GMM)
  - Joint use of instruments, overidentification test, and instrument selection
- Compared with Local Projection, SFMs extend SVAR-IV, $\therefore$ less restrictive with instruments

Monte Carlo study confirms the theoretical properties of the proposed estimator

Empirical Study with U.S. macroeconomic data:

- Evidence that all external instruments provided by the literature are valid
- Leveraging all instruments in a data-rich environment leads to more reasonable and efficient responses

Related Literature

SVARs and Instrument Identification

- Applications: Caldara and Kamps (2017), Gertler and Karadi (2015), Mertens and Ravn (2013), and Stock and Watson (2012)
- Theory: Cheng et al. (2021), Montiel Olea et al. (2021), and Stock and Watson (2018)
- Overidentification: Cheng et al. (2021), Montiel Olea et al. (2021), and Schlaak et al. (2023)

Identification in Structural Factor Models

- Fast/slow or zero restrictions: Han (2015, 2018) and Stock and Watson (2005)
- Sign restrictions (only possibility): Gafarov (2014)
- External Instrument (**one at a time**): Han (2024) and Stock and Watson (2012)

Model Setup

Consider the structural model for $N \times 1$ series X_t, $t = 1, \ldots, T$:

$$X_t = \Lambda F_t + e_t, \tag{1}$$

$$F_t = \sum_{j=1}^{p} \Phi_j F_{t-j} + G\eta_t, \tag{2}$$

$$\eta_t = A\zeta_t, \tag{3}$$

- F_t are $r \times 1$ unobserved factors, Λ are $N \times r$ loadings, e_t are $N \times 1$ idiosyncratic error
- G is $r \times q$, maps $q \times 1$ reduced form shocks η_t to r static factors ($q < r$ results in **covariance singularity**)
- ζ_t are $q \times 1$ structural shocks s.t. $E(\zeta_t \zeta_t^\intercal) = I_q$, A is $q \times q$ nonsingular matrix

Stationarity in F_t implies the MA representation

$$(I_r - \Phi_1 L - \cdots - \Phi_p L^p)^{-1} = \sum_{s=1}^{\infty} \Psi_s L^s. \tag{4}$$

Model Setup

Equations imply the following factor structure for the reduced form shocks

$$X_t = \Pi \mathcal{F}_t + \Theta \eta_t + e_t, \tag{5}$$

$$X = \mathcal{F} \Pi^\intercal + \eta \Theta^\intercal + e, \tag{6}$$

where $\Pi = \Lambda \Phi$, $\Phi = \Lambda G$ and $\Gamma = \Theta A$.

Model Setup

Equations imply the following factor structure for the reduced form shocks

$$X_t = \Pi \mathcal{F}_t + \Theta \eta_t + e_t, \tag{5}$$

$$X = \mathcal{F} \Pi^\intercal + \eta \Theta^\intercal + e, \tag{6}$$

where $\Pi = \Lambda\Phi$, $\Phi = \Lambda G$ and $\Gamma = \Theta A$.

Goal: identify the impulse responses to the first structural shock

1. Estimate the reduced form shocks η_t
2. Estimate A, recalling that $\eta_t = A\zeta_t$ with external instrument(s)
3. Map effect of shock to all N variables using factor structure

Model Setup

Equations imply the following factor structure for the reduced form shocks

$$X_t = \Pi \mathcal{F}_t + \Theta \eta_t + e_t, \tag{5}$$
$$X = \mathcal{F} \Pi^\intercal + \eta \Theta^\intercal + e, \tag{6}$$

where $\Pi = \Lambda\Phi$, $\Phi = \Lambda G$ and $\Gamma = \Theta A$.

Goal: identify the impulse responses to the first structural shock

1. Estimate the reduced form shocks η_t
2. Estimate A, recalling that $\eta_t = A\zeta_t$ with external instrument(s)
3. Map effect of shock to all N variables using factor structure

Notice that X with the lags of F projected out exhibits a factor structure:

$$X - \mathcal{F} \Pi^\intercal = \eta \Theta^\intercal + e.$$

Model Setup

Equations imply the following factor structure for the reduced form shocks

$$X_t = \Pi \mathcal{F}_t + \Theta \eta_t + e_t, \tag{5}$$

$$X = \mathcal{F} \Pi^\intercal + \eta \Theta^\intercal + e, \tag{6}$$

where $\Pi = \Lambda \Phi$, $\Phi = \Lambda G$ and $\Gamma = \Theta A$.

Goal: identify the impulse responses to the first structural shock

1. Estimate the reduced form shocks η_t
2. Estimate A, recalling that $\eta_t = A \zeta_t$ with external instrument(s)
3. Map effect of shock to all N variables using factor structure

Notice that X with the lags of F projected out exhibits a factor structure:

$$X - \mathcal{F} \Pi^\intercal = \eta \Theta^\intercal + e.$$

Estimate η by PC on $X - \mathcal{F} \Pi^\intercal$ (2 Stage PC of Han, 2018)

Estimation

Static factors and loadings are estimated via PC $\widehat{\Lambda}$ and $\widehat{F}$. Define

$$\widehat{X} = M_{\widehat{\mathcal{F}}} X, \tag{7}$$

where $M_{\widehat{\mathcal{F}}}$ the residual maker using $\widehat{\mathcal{F}}$.

Reduced form shocks η_t are estimated using a second PC fit on $\widehat{X}$, i.e. $\widehat{\eta}$ is $\sqrt{T-p}$ times the eigenvectors of the first q eigenvalues of $\widehat{X}\widehat{X}^\intercal$.

$\widehat{G}$ ($r \times q$ mapping) and VAR coefficients estimated using LS

$$\begin{aligned} \widehat{G} &= \widehat{F}^\intercal \widehat{\eta} \left(\widehat{\eta}^\intercal \widehat{\eta} \right)^{-1} \\ &= \frac{1}{T-p} \sum_{t=p+1}^{T} \widehat{F}_t \widehat{\eta}_t^\intercal, \end{aligned} \tag{8}$$

$$\widehat{\Phi} = \widehat{F}^\intercal \widehat{\mathcal{F}} \left(\widehat{\mathcal{F}}^\intercal \widehat{\mathcal{F}} \right)^{-1}. \tag{9}$$

$\widehat{\Psi}_s$ for $s = 1, 2, \ldots$ follow by inverting the lag polynomial.

Identification of Structural Parameters

Goal: identify impulse responses to the first structural shock

$$\frac{\partial X_t}{\partial \zeta_{1,t-s}} = \Lambda \Psi_s G a_1, \tag{10}$$

where a_1 is the first column of A.

- Estimators for Λ, Ψ_s and G described earlier.
- Estimator for η_t obtained by PC

Identification of Structural Parameters

Goal: identify impulse responses to the first structural shock

$$\frac{\partial X_t}{\partial \zeta_{1,t-s}} = \Lambda \Psi_s G a_1, \tag{10}$$

where a_1 is the first column of A.

- Estimators for Λ, Ψ_s and G described earlier.
- Estimator for η_t obtained by PC

However, PC based estimator of $\widehat{\eta}_t$ recovers $H_\eta^\intercal \eta$, where H_η is a rotational basis.

Identification of Structural Parameters

Goal: identify impulse responses to the first structural shock

$$\frac{\partial X_t}{\partial \zeta_{1,t-s}} = \Lambda \Psi_s G a_1, \tag{10}$$

where a_1 is the first column of A.

- Estimators for Λ, Ψ_s and G described earlier.
- Estimator for η_t obtained by PC

However, PC based estimator of $\widehat{\eta}_t$ recovers $H_\eta^\intercal \eta$, where H_η is a rotational basis.

- Careful adjustments necessary for asymptotic theory
- Instead identifying $H_\eta^\intercal a_1 = a_1^*$, but we show that this suffices for identification of the IRFs themselves

Identification of Structural Parameters - Instruments

Instrument Conditions

1. Relevance: $E(Z_t\zeta_{1t}) = \alpha \neq 0_k$
2. Exogeneity: $E(Z_t\zeta_{jt}) \neq 0_k$ for $j \neq 1$

$$\begin{aligned} E(H_\eta^\intercal \eta_t Z_t^\intercal) &= E(H_\eta^\intercal A\zeta_t Z_t^\intercal) \\ &= H_\eta^\intercal a_1 \alpha^\intercal \\ &= a_1^* \alpha^\intercal \in R^{1\times k}. \end{aligned} \tag{11}$$

Identification of Structural Parameters - Instruments

Instrument Conditions

1. Relevance: $E(Z_t \zeta_{1t}) = \alpha \neq 0_k$
2. Exogeneity: $E(Z_t \zeta_{jt}) \neq 0_k$ for $j \neq 1$

$$\begin{aligned} E(H_\eta^\top \eta_t Z_t^\top) &= E(H_\eta^\top A \zeta_t Z_t^\top) \\ &= H_\eta^\top a_1 \alpha^\top \\ &= a_1^* \alpha^\top \in R^{1\times k}. \end{aligned} \tag{11}$$

Normalize the first element of a_1^* to unity, and define

$$\delta = \left[a_{12}^*, \ldots, a_{1q}^* \right]^\top \in R^{q-1}. \tag{12}$$

Implied moment conditions (just-identified for $k = 1$, over-identified if $k > 1$)

$$E\left[(\eta_{-1t}^* - \delta \eta_{1t}^*) \otimes Z_t \right] = \mathbf{0} \in R^{k(q-1)}. \tag{13}$$

Estimation of Structural Parameters - GMM

$$\text{GMM Criterion}: \mathcal{Q}_T(\delta) = \bar{g}_T(\delta)^{\intercal} W_T \bar{g}_T(\delta), \tag{14}$$

$$\text{Empirical Moments}: \bar{g}_T(\delta) = \frac{1}{T-p} \sum_{t=p+1}^{T} \left[(\widehat{\eta}_{-1t} - \delta \widehat{\eta}_{1t}) \otimes Z_t \right]. \tag{15}$$

Estimation of Structural Parameters - GMM

$$\text{GMM Criterion}: \mathcal{Q}_T(\delta) = \bar{g}_T(\delta)^\intercal W_T \bar{g}_T(\delta), \tag{14}$$

$$\text{Empirical Moments}: \bar{g}_T(\delta) = \frac{1}{T-p} \sum_{t=p+1}^{T} \left[(\widehat{\eta}_{-1t} - \delta \widehat{\eta}_{1t}) \otimes Z_t \right]. \tag{15}$$

FOC yields GMM estimator

$$\widehat{\delta} = \left(\mathcal{A}_T W_T \mathcal{A}_T^\intercal \right)^{-1} \mathcal{A}_T W_T \mathcal{G}_T, \tag{16}$$

$$\mathcal{A}_T = I_{q-1} \otimes \left(\frac{1}{T-p} \sum_{t=p+1}^{T} \widehat{\eta}_{1t} Z_t^\intercal \right), \text{ and } \mathcal{G}_T = \frac{1}{T-p} \sum_{t=p+1}^{T} \left(\widehat{\eta}_{-1t} \otimes Z_t \right). \tag{17}$$

Estimation of Structural Parameters - GMM

$$\text{GMM Criterion}: \mathcal{Q}_T(\delta) = \bar{g}_T(\delta)^\intercal W_T \bar{g}_T(\delta), \tag{14}$$

$$\text{Empirical Moments}: \bar{g}_T(\delta) = \frac{1}{T-p} \sum_{t=p+1}^{T} \left[(\widehat{\eta}_{-1t} - \delta \widehat{\eta}_{1t}) \otimes Z_t \right]. \tag{15}$$

FOC yields GMM estimator

$$\widehat{\delta} = \left(\mathcal{A}_T W_T \mathcal{A}_T^\intercal \right)^{-1} \mathcal{A}_T W_T \mathcal{G}_T, \tag{16}$$

$$\mathcal{A}_T = I_{q-1} \otimes \left(\frac{1}{T-p} \sum_{t=p+1}^{T} \widehat{\eta}_{1t} Z_t^\intercal \right), \text{ and } \mathcal{G}_T = \frac{1}{T-p} \sum_{t=p+1}^{T} \left(\widehat{\eta}_{-1t} \otimes Z_t \right). \tag{17}$$

Eqn by eqn 2SLS of Ramey (2016) corresponds to $W_T = I_{q-1} \otimes \left(\frac{1}{T-p} \sum_{t=p+1}^{T} Z_t Z_t^\intercal \right)^{-1}$ weights. Two-step optimal GMM estimator $\widehat{\delta}^o$ follows by typical GMM arguments.

Assumption 1. Factors satisfy Assumption 1 of Han (2018).
Assumption 2. Loadings and VAR coefficients, Assumption 2 of Han (2018):
Assumption 3. Moments of idiosyncratic errors in Bai (2003).
Assumption 4. $\{\lambda_i\}$, $\{\zeta_t\}$ and $\{e_{it}\}$ are mutually independent groups.
Assumption 5. Weak serial and cross sectional correlation in errors.
Assumption 6. Assumption F in Bai (2003).
Assumption 7. Central Limit Theorem. For $i = 1, \ldots, N$,

$$\frac{1}{\sqrt{T}} \begin{bmatrix} \text{vec}\left(Z^\intercal \eta - E(Z^\intercal \eta)\right) \\ F^\intercal e_i \\ \text{vec}\left(\mathcal{F}^\intercal \eta\right) \end{bmatrix} \xrightarrow{d} N\left(\mathbf{0}_{(qk+r+rpq)\times 1}, \Sigma_i\right).$$

Assumption 8. The structural shock ζ_t is linked the reduced form error by the linear transformation $\eta_t = A\zeta_t$, for some nonsingular matrix A, and $E(Z_t \zeta_t^\intercal) = [\alpha, 0_{k\times(q-1)}]$, where $\alpha \neq 0_k$.
Assumption 8 allows for Z_t to be correlated with lags of ζ_t, hence much looser than LP-IV, (Stock & Watson, 2018).

Asymptotic Results - Notations and Difficulty

Principal Components only estimates the true quantities up to an arbitrary rotation

$$\widehat{F} \to FH_F,$$
$$\widehat{\eta} \to \eta H_\eta.$$

Asymptotic Results - Notations and Difficulty

Principal Components only estimates the true quantities up to an arbitrary rotation

$$
\begin{aligned}
\widehat{F} &\to FH_F, \\
\widehat{\eta} &\to \eta H_\eta.
\end{aligned}
$$

Two Stage PC implies two arbitrary rotations:

$$
\begin{aligned}
H_F &= \left(\frac{\Lambda^\intercal \Lambda}{N}\right)\left(\frac{F^\intercal \widehat{F}}{T}\right)\widehat{V}_F^{-1}, \quad \operatorname{plim} H_F = \bar{H}_F, \\
H_\eta &= \left(\frac{\Theta^\intercal \Theta}{N}\right)\left(\frac{\eta^\intercal \widehat{\eta}}{T}\right)\widehat{V}_\eta^{-1}, \quad \operatorname{plim} H_\eta = \bar{H}_\eta
\end{aligned}
$$

$\widehat{V}_F$, $\widehat{V}_\eta$ are the eigenvalue matrices associated with $XX^\intercal/NT$ and $\widehat{X}\widehat{X}^\intercal/NT$.

Asymptotic Results - Notations and Difficulty

Principal Components only estimates the true quantities up to an arbitrary rotation

$$\widehat{F} \to FH_F,$$
$$\widehat{\eta} \to \eta H_\eta.$$

Two Stage PC implies two arbitrary rotations:

$$H_F = \left(\frac{\Lambda^\intercal \Lambda}{N}\right)\left(\frac{F^\intercal \widehat{F}}{T}\right)\widehat{V}_F^{-1}, \quad \text{plim}\, H_F = \bar{H}_F,$$
$$H_\eta = \left(\frac{\Theta^\intercal \Theta}{N}\right)\left(\frac{\eta^\intercal \widehat{\eta}}{T}\right)\widehat{V}_\eta^{-1}, \quad \text{plim}\, H_\eta = \bar{H}_\eta$$

$\widehat{V}_F$, $\widehat{V}_\eta$ are the eigenvalue matrices associated with $XX^\intercal/NT$ and $\widehat{X}\widehat{X}^\intercal/NT$.

Let $\mathbb{S}_\delta = \left[\delta \vdots I_{q-1}(1:q-1)\right]$, and $\mathbb{S}_{\widehat{\delta}}$ its feasible counterpart, s.t.

$$\mathbb{S}_\delta \eta_t^* = \eta_{-1t}^* - \delta \eta_{1t}^*$$

Asymptotic Results - Preliminary

Theorem 1. Under Assumptions 1 to 8, and the conditions that $W_T \xrightarrow{p} W$, and $\sqrt{T}/N \to 0$ as $N, T \to \infty$,

1. $\widehat{\delta}$ is a consistent estimator of δ, and

$$\sqrt{T}\left(\widehat{\delta}-\delta\right) \xrightarrow{d} (\mathcal{A}W\mathcal{A}^{\intercal})^{-1}\mathcal{A}WN\left(0_{kq\times 1}, \left(\mathbb{S}_{\delta}\bar{H}_{\eta}^{\intercal}\otimes I_k\right)\Sigma_i^{(1)}\left(\mathbb{S}_{\delta}\bar{H}_{\eta}^{\intercal}\otimes I_k\right)^{\intercal}\right),$$

   where $\mathcal{A} = I_{q-1}\otimes \mathbb{S}_1\bar{H}_{\eta}^{\intercal}E(\eta_{1t}Z_t^{\intercal})$, $\mathbb{S}_1 = [1, 0_{1\times(q-1)}]$, and $\Sigma_i^{(1)}$ is the upper left block of Σ_i.
2. The optimal choice of the weighting matrix is V_{δ}^{-1}, where $V_{\delta} = \mathcal{C}\Sigma_i^{(1)}\mathcal{C}^{\intercal}$ and $\mathcal{C} = \left[\mathbb{S}_{\delta}\bar{H}_{\eta}^{\intercal}\otimes I_k\right]$.
3. $\widehat{V}_{\delta} \xrightarrow{p} V_{\delta}$.
4. $\sqrt{T}\left(\widehat{\delta}^{o}-\delta\right) \xrightarrow{d} N\left(0, [\mathcal{A}V_{\delta}^{-1}\mathcal{A}^{\intercal}]^{-1}\right)$.

Consistency for the IRFs additionally rely on consistency and distributions of $\widehat{\lambda}_i$, $\widehat{G}$ and $\widehat{\Psi}_s$ (details omitted for brevity).

Asymptotic Results - Main

Theorem 2. Under Assumptions 1 to 8, and the conditions that $W_T \xrightarrow{p} W$ and $\sqrt{T}/N \to 0$ as $N, T \to \infty$, the IRFs of X_{it} to $\zeta_{1,t-s}, (s \geq 1)$ satisfy:

$$\sqrt{T}\left(\widehat{\lambda}_i^\intercal \widehat{\Psi}_s \widehat{G}\widehat{a}_1 - \lambda_i^\intercal \Psi_s G a_1\right) = \sqrt{T}\bar{Q}_{2,i}\begin{bmatrix} \widehat{a}_1 - a_1^* \\ \widehat{\lambda}_i - H_F^{-1}\lambda_i \\ \text{vec}\left(\widehat{\Psi}_s^\intercal - H_F^{-1}\Psi_s H_F\right) \end{bmatrix} + o_p(1)$$
$$\xrightarrow{d} N\left(0, \bar{Q}_{2,i}B_s\Sigma_i B_s^\intercal \bar{Q}_{2,i}^\intercal\right),$$

where

$$\bar{Q}_{2,i} = \lambda_i^\intercal \Psi_s G \Sigma_\eta \bar{H}_\eta C_3 + a_1^\intercal G^\intercal \Psi_s^\intercal \bar{H}_F C_4 + \left(\lambda_i^\intercal \bar{H}_F^{-\intercal} \otimes a_1^\intercal G^\intercal \bar{H}_F\right) C_5,$$

using $C_3 = [I_{qk} \vdots 0_{qk\times r} \vdots 0_{qk\times r^2}]$, $C_4 = [0_{r\times qk} \vdots I_r \vdots 0_{r\times r^2}]$ and $C_5 = [0_{r^2\times qk} \vdots 0_{r^2\times r} \vdots I_{r^2}]$.

Asymptotic Theory: Covariance Matrix Estimation

Define $\widehat{e}_{it}$ as the ith element of $\widehat{e}_t = X_t - \widehat{\Lambda}\widehat{F}_t$. Naive estimators for the variance defined as

$$\widehat{B_s \Sigma_i B_s^\intercal} = \frac{1}{T-p} \sum_{t=p+1}^{T} \xi_{it}\xi_{it}^\intercal, \tag{18}$$

$$\xi_{it} = \begin{bmatrix} \mathrm{vec}\left(Z_t^\intercal \widehat{\eta}_t - \frac{1}{T-p}(Z^\intercal \widehat{\eta})\right) \\ \widehat{F}_t^\intercal \widehat{e}_{it} \\ \widehat{R}_s \left(\widehat{G} \otimes \left(\frac{\widehat{\mathcal{F}}^\intercal \widehat{\mathcal{F}}}{T-p}\right)^{-1}\right) \mathrm{vec}\left(\widehat{\mathcal{F}}_t^\intercal \widehat{\eta}_t\right) \end{bmatrix}, \quad \widehat{R}_s = \sum_{j=1}^{s} \left(\widehat{\Psi}_{j-1} \otimes \left[\widehat{\Psi}_{s-j}^\intercal, \widehat{\Psi}_{s-j-1}^\intercal, \ldots, \widehat{\Psi}_{s-j-p+1}^\intercal\right]\right),$$

with $\widehat{\Phi}_0 = I_r$ and $\widehat{\Psi}_s = 0_{r\times r}$ for $s < 0$.

- HAC of Bai (2003), CS-HAC of Bai and Ng (2006) or Bootstrap of Gonçalves and Perron (2020) also possible
- $\bar{Q}_{s,i}$ for $s = 1, 2$ can estimated using $\widehat{\lambda}_i$, $\widehat{\Psi}_s$, $\widehat{G}$ and $\widehat{a}_1$

Asymptotic Theory: Overidentification and Instrument Selection

GMM framework naturally allows for J-test:

$$J_T \equiv T\mathcal{Q}_T(\widehat{\delta}) \xrightarrow{d} \chi^2_{(k-1)(q-1)}. \tag{19}$$

Additionally, instrument selection using either information criterion or Downwards Testing possible, following Andrews (1999).

Asymptotic Theory: Overidentification and Instrument Selection

GMM framework naturally allows for J-test:

$$J_T \equiv T\mathcal{Q}_T(\widehat{\delta}) \xrightarrow{d} \chi^2_{(k-1)(q-1)}. \tag{19}$$

Additionally, instrument selection using either information criterion or Downwards Testing possible, following Andrews (1999).

Assumption 9. Define $\mathscr{Z} = \{c \in \mathcal{C} : c = c^0(\delta) \text{ for some } \delta\}$, the set of selection vectors in $\mathcal{C}$ which select only moment conditions that are zero asymptotically, and $\mathscr{M}\mathscr{Z} = \{c \in \mathscr{Z} : |c| \geq |c^*| \forall c^* \in \mathscr{Z}\}$, the set of selection vectors in $\mathscr{Z}$ that maximize the selected moments out of selection vectors in $\mathscr{Z}$.

- $\mathscr{M}\mathscr{Z}$ contains a single element c^0,
- $\bar{g}_{T,c}(\delta)$ has a unique solution δ.

Instrument Selection & Downwards Testing

Information Criteria: Select instruments using $\widehat{c}_{BIC}, \widehat{c}_{AIC}$, and $\widehat{c}_{HQIC}$, respectively minimize:

$$
\begin{aligned}
GMM_{BIC} &= J_T(c) - (|c| - 1)(q - 1) log T \\
GMM_{AIC} &= J_T(c) - 2(|c| - 1)(q - 1) \\
GMM_{HQIC} &= J_T(c) - Q(|2| - 1)(q - 1) loglog T, \quad Q > 2
\end{aligned}
$$

Instrument Selection & Downwards Testing

Information Criteria: Select instruments using $\widehat{c}_{BIC}, \widehat{c}_{AIC}$, and $\widehat{c}_{HQIC}$, respectively minimize:

$$GMM_{BIC} = J_T(c) - (|c| - 1)(q - 1)logT$$
$$GMM_{AIC} = J_T(c) - 2(|c| - 1)(q - 1)$$
$$GMM_{HQIC} = J_T(c) - Q(|2| - 1)(q - 1)loglogT, \quad Q > 2$$

Downwards Testing: Sequentially test all instrument combos until non-rejection

- Starting with all instruments, carry out J-tests with smaller $|c|$ until we do not reject
- Let $\widehat{k}_{DT} = |c|$ for the first J-test that does not reject
- Downwards-testing estimator $\widehat{c}_{DT}$ is instrument set that minimize $J_T(c)$ s.t. $|c| = \widehat{k}_{DT}$

Asymptotic Results - Overidentification and Moment Selection

Theorem 3. Under Assumptions 1 to 8 and the condition that $\frac{\sqrt{T}}{N} \to 0$ as $N, T \to \infty$,

1. $J_T \equiv T\mathcal{Q}_T(\widehat{\delta}) \xrightarrow{d} \chi^2_{(k-1)(q-1)}$,
2. Additionally under Assumption 9, for $MSC \in \{GMM_{BIC}, GMM_{AIC}, GMM_{HQIC}\}$, $\widehat{c}_{MSC} = c^0$ wp $\to 1$,
3. Additionally under Assumption 9, $\widehat{c}_{DT} = c^0$ wp $\to 1$.

Simulation Study: Model Setup

VAR in dynamic factors with autocorrelation ($\zeta_t \sim N(0_q, I_q)$)

$$f_t = \phi f_{t-1} + A\zeta_t, \quad \phi = 0.7, \quad A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}. \tag{20}$$

Observable series generated using static factors which stack lags of dynamic factors, with $r = 5, q = 3$ to induce singularity:

$$X_t = \Lambda F_t + e_t, \quad F_t = [f_t^{\mathsf{T}}, f_{1,t-1}, f_{r-q,t-1}]^{\mathsf{T}} \tag{21}$$

$$\text{DGP 1} : Z_{jt} = \sqrt{1-a^2} w_{jt} + a\zeta_{1t} + \zeta_{q,t-1}, j = 1, \ldots, k = 4 \tag{22}$$

$$\text{DGP 2} : Z_{jt} = \sqrt{1-a^2} w_{jt} + a\zeta_{1t} + \zeta_{q,t-1}, j = 1, \ldots, k = 4 \tag{23}$$

Table 1: Coverage Probabilities

T	N	$h=0$, $k=1$	$h=0$, $k=2$	$h=0$, $k=3$	$h=0$, $k=4$	$h=3$, $k=1$	$h=3$, $k=2$	$h=3$, $k=3$	$h=3$, $k=4$
250	125	0.919	0.897	0.889	0.881	0.954	0.949	0.947	0.944
	250	0.920	0.904	0.895	0.888	0.954	0.949	0.945	0.944
500	125	0.903	0.890	0.883	0.880	0.946	0.943	0.941	0.940
	250	0.909	0.898	0.892	0.889	0.948	0.945	0.943	0.942

Note:
Entries report the coverage probabilities of the IRFs using the proposed asymptotic distributions (nominal 95%).

Table 2: RMSE ratios

T	N	$h=0$, $k=2$	$h=0$, $k=3$	$h=0$, $k=4$	$h=0$, SVAR-IV ($k=4$)	$h=3$, $k=2$	$h=3$, $k=3$	$h=3$, $k=4$	$h=3$, SVAR-IV ($k=4$)
250	125	0.923	0.911	0.892	6.150	0.968	0.956	0.953	3.688
	250	0.924	0.919	0.889	6.340	0.963	0.951	0.935	3.706
500	125	0.940	0.914	0.902	8.293	0.973	0.969	0.966	4.874
	250	0.941	0.913	0.901	8.711	0.979	0.970	0.963	4.934

Note:
Entries report the RMSE of the estimated IRFs of the overidentified system, compared to the RMSE of the IRFs of the just-identified system. The SVAR-IV is estimated with $k=4$ variables, the first of which is X_{1t} for normalisation, and

Table 3: Power of J-test

T	N	Rejection Frequency
250	125	1.000
	250	1.000
500	125	1.000
	250	1.000

Note:
Entries report the rejection frequency of J-test for overidentification, with $k = 4$ instruments.

Table 4: Accuracy of Moment Selection Procedures

		Information Criteria			Testing	
T	N	GMM_{BIC}	GMM_{AIC}	GMM_{HQIC}	DT	UT
250	125	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000
500	125	1.000	1.000	1.000	1.000	1.000
	250	1.000	1.000	1.000	1.000	1.000

Note:
Entries report the frequencies of correct instrument selection. DT and UT denote Downwards and Upwards Testing respectively. Correct instruments are Z_1 and Z_2.

Empirical Study - Monetary Policy Shocks

Quarterly Dataset of Stock and Watson (2012), 1980Q1 - 2007Q2 due to data availability

- Narrative instrument of Romer and Romer (2004)
- Model-based instruments of Bernanke and Mihov (1998)
- High-frequency instruments of Barakchian and Crowe (2013), Gertler and Karadi (2015), and Miranda-Agrippino and Ricco (2021)

Empirical Study - Monetary Policy Shocks

Quarterly Dataset of Stock and Watson (2012), 1980Q1 - 2007Q2 due to data availability

- Narrative instrument of Romer and Romer (2004)
- Model-based instruments of Bernanke and Mihov (1998)
- High-frequency instruments of Barakchian and Crowe (2013), Gertler and Karadi (2015), and Miranda-Agrippino and Ricco (2021)

Benchmark Model:

- $r = 9$ static factors, Bai and Ng (2002)
- $q = 3$ dynamic factors, Bai and Ng (2007)
- $p = 2$ lags

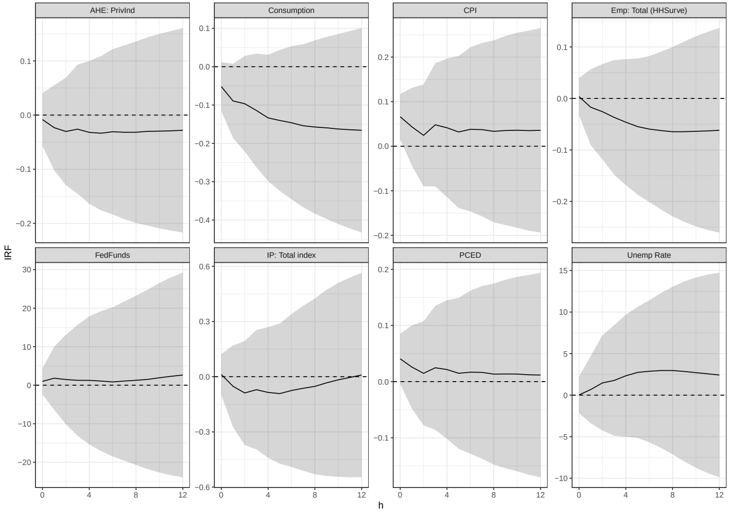

Figure 1: Cumulative IRFs after a contractionary monetary policy shock for the over-identified benchmark (solid curve), and just identified setups each using one instrument at a time, normalised to a 100 basis point movement in the Federal Funds rate.

Table 5: Results of J-test for Overidentification and GMM_{MSC} Criteria.

	GMM_{BIC}	GMM_{HQIC}	J_T	J_{crit}	RR04	GK15	MR21	BM98	BK13
1	-8.57	-5.50	0.638	5.99	1	1	0	0	0
2	-8.82	-5.75	0.386	5.99	1	0	1	0	0
3	-8.84	-5.76	0.375	5.99	0	1	1	0	0
4	-15.29	-9.14	3.135	9.49	1	1	1	0	0
5	-9.20	-6.13	0.011	5.99	1	0	0	1	0
11	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
22	-23.04	-13.82	4.596	12.59	1	1	0	1	1
23	-14.78	-8.64	3.641	9.49	0	0	1	1	1
24	-23.04	-13.83	4.587	12.59	1	0	1	1	1
25	-21.32	-12.11	6.311	12.59	0	1	1	1	1
26	-28.95	-16.67	7.887	15.51	1	1	1	1	1

Note:
RR04, GK15, MR21, BM98, and BC13 refer to the external instruments of Romer and Romer (2004), Gertler and Karadi (2015), Miranda-Agrippino and Rossi (2021), Bernanke and Mihov (1998), and Barakchian and Crowe (2013), where 1 denotes that the instrument was used.

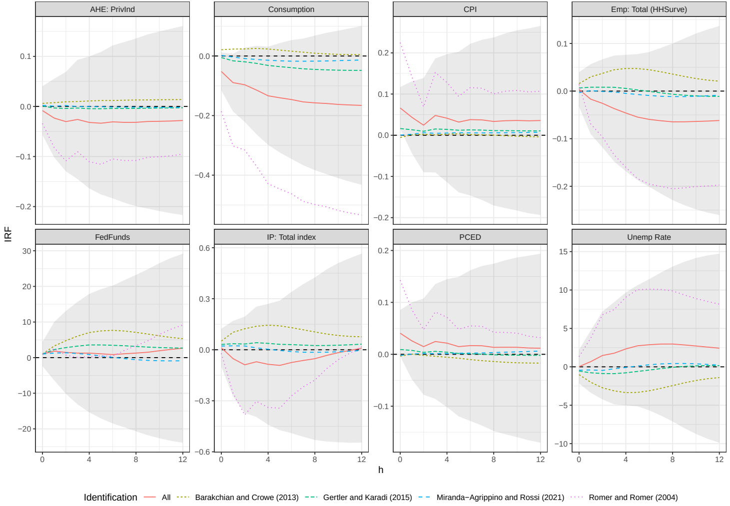

Figure 2: IRFs to a 100 basis point shock in the Federal Funds Rate, using Benchmark Model (solid) compared to using each instrument one at a time (broken).

Conclusion

Propose new estimators for impulse responses functions using a Structural Factor Model framework

- Factors parsimoniously incorporate large datasets to deal with non-fundamentalness
- Allow for no. of static factors $>$ no. of primitive shocks, to deal with singular covariance matrix
- GMM approach to use multiple instruments, allows for testing and selection of instruments

Conclusion

Propose new estimators for impulse responses functions using a Structural Factor Model framework

- Factors parsimoniously incorporate large datasets to deal with non-fundamentalness
- Allow for no. of static factors $>$ no. of primitive shocks, to deal with singular covariance matrix
- GMM approach to use multiple instruments, allows for testing and selection of instruments

Simulation study shows that proposed estimators produce more accurate impulse responses.

Conclusion

Propose new estimators for impulse responses functions using a Structural Factor Model framework

- Factors parsimoniously incorporate large datasets to deal with non-fundamentalness
- Allow for no. of static factors $>$ no. of primitive shocks, to deal with singular covariance matrix
- GMM approach to use multiple instruments, allows for testing and selection of instruments

Simulation study shows that proposed estimators produce more accurate impulse responses.

Empirical study on U.S. macroeconomic data shows all instruments are valid, and their joint use can lead to more efficient and reasonable estimates of monetary policy responses.

- High-frequency instruments by themselves seem to produce puzzling responses

Thank You!