Disentangling Structural Breaks in Factor Models for Macroeconomic Data

Bonsoo Koo
Monash University

Benjamin Wong
Monash University

Ze-Yu Zhong*
Monash University

July 2024

We acknowledge the financial support of the Australian Research Council under Grants DP210101440, DP240100970, and DE200100693.

In this talk...

Structural breaks in dynamic factor models

- Unique challenge within factor models
- Both factors f_t and loadings λ_i are unobserved and enter in a multiplicative way

In this talk...

Structural breaks in dynamic factor models

- Unique challenge within factor models
- Both factors f_t and loadings λ_i are unobserved and enter in a multiplicative way
- Great Moderation, Global Financial Crisis, COVID-19
- Unclear whether these are legitimate breaks in loadings, or simply factor heteroskedasticity

In this talk...

Structural breaks in dynamic factor models

- Unique challenge within factor models
- Both factors f_t and loadings λ_i are unobserved and enter in a multiplicative way
- Great Moderation, Global Financial Crisis, COVID-19
- Unclear whether these are legitimate breaks in loadings, or simply factor heteroskedasticity

Questions

- How to decompose structural breaks into those associated with factors vs loadings?
- How to test for evidence of these separately?
- Does this lead to a more nuanced interpretation of empirical events?

Contribution

Key Idea: Projection Decomposition

- Reparameterize structural change in the factor structure into
  - Rotational component: factor heteroskedasticity $\leftrightarrow$ Change in the composition of the factors, but still span the same space
  - Shift component: change in the factor loading $\leftrightarrow$ Breaks in the loadings should be orthogonal to the original factor space

Contribution

Key Idea: Projection Decomposition

- Reparameterize structural change in the factor structure into
  - Rotational component: factor heteroskedasticity $\leftrightarrow$ Change in the composition of the factors, but still span the same space
  - Shift component: change in the factor loading $\leftrightarrow$ Breaks in the loadings should be orthogonal to the original factor space

Empirical Work: Great Moderation

Applying our procedure to the FRED-QD dataset

- 60% reduction of the factor variance
- Perhaps unsurprising, but was previously considered as a break in the factor loading in the literature
- Suggests a more nuanced interpretation of breaks in factor models

Dynamic Factor Models

Consider the model of Stock and Watson (2006):

$$X_t = \Lambda f_t + e_t \tag{1.1}$$

$$f_t = \sum_{j=1}^{p} \Phi_j f_{t-j} + \eta_t, \quad \eta_t \sim (0, \Sigma_\eta), \tag{1.2}$$

- Equation (1.1) clarifies how factors are related to a set of variables
- Equation (1.2) describes the dynamics of the factors in a VAR form
- Reduced form innovations η_t with Σ_η variance

Dynamic Factor Models

$$X_t = \Lambda f_t + e_t$$

$$f_t = \sum_{j=1}^{p} \Phi_j f_{t-j} + \eta_t, \quad \eta_t \sim (0, \Sigma_\eta),$$

- Breaks in the λ_i's are isolated to breaks in the relationship between the different variables and factors
- Breaks in the factor variance necessitate a break in Φ_j and/or Σ_η

Dynamic Factor Models

$$X_t = \Lambda f_t + e_t$$

$$f_t = \sum_{j=1}^{p} \Phi_j f_{t-j} + \eta_t, \quad \eta_t \sim (0, \Sigma_\eta),$$

- Breaks in the λ_i's are isolated to breaks in the relationship between the different variables and factors
- Breaks in the factor variance necessitate a break in Φ_j and/or Σ_η

Conjecture: Great Moderation

- Marked by a general reduction of volatility across many macroeconomic variables
- More naturally accommodated as a reduction in the variance of f_t, rather than multiple (proportional) breaks in λ_i

Dynamic Factor Models: Estimation and Normalization

- Principal Components (PC) estimator consistently estimated the space spanned by the factors
- VAR specifications occurs as a separate step
- Normalization required: factors are identified up to a rotation:

$$\frac{1}{N}\tilde{\Lambda}^{\intercal}\tilde{\Lambda} = V_{NT}, \quad \frac{1}{T}\tilde{F}^{\intercal}\tilde{F} = I_r,$$

  - $\tilde{\Lambda}$ and $\tilde{F}$ are the PC estimators of loadings and factors
  - V_{NT} is a diagonal matrix of the first r eigenvalues associated with cov matrix of X

Dynamic Factor Models: Estimation and Normalization

- Principal Components (PC) estimator consistently estimated the space spanned by the factors
- VAR specifications occurs as a separate step
- Normalization required: factors are identified up to a rotation:

$$\frac{1}{N}\tilde{\Lambda}^{\intercal}\tilde{\Lambda} = V_{NT}, \quad \frac{1}{T}\tilde{F}^{\intercal}\tilde{F} = I_r,$$

  - $\tilde{\Lambda}$ and $\tilde{F}$ are the PC estimators of loadings and factors
  - V_{NT} is a diagonal matrix of the first r eigenvalues associated with cov matrix of X

Inability of differentiating between these different break types

Necessary routine normalization therefore subsumes breaks in the factor variance into the loadings, or vice versa.

Earlier Literature

Tests for breaks:

- Breitung and Eickmeier (2011), Han and Inoue (2015), and Stock and Watson (2009)

Estimation of breaks:

- Baltagi et al. (2017, 2021), Duan et al. (2022), and Ma and Su (2018)

Consequence of breaks:

- Overestimation of no. of factors if ignored, Breitung and Eickmeier (2011)
- Inconsistent estimation of factor space, Bates et al. (2013)
- Forecast failure, Banerjee et al. (2008) and Bates et al. (2013)

Earlier Literature

Tests for breaks:

- Breitung and Eickmeier (2011), Han and Inoue (2015), and Stock and Watson (2009)

Estimation of breaks:

- Baltagi et al. (2017, 2021), Duan et al. (2022), and Ma and Su (2018)

Consequence of breaks:

- Overestimation of no. of factors if ignored, Breitung and Eickmeier (2011)
- Inconsistent estimation of factor space, Bates et al. (2013)
- Forecast failure, Banerjee et al. (2008) and Bates et al. (2013)

Placement of our work

- Highlights and addresses limitations of earlier works on estimation and testing of breaks
- Reconciles structural breaks in factor models with macroeconomic intuition
- Closest to Massacci (2021), Pelger and Xiong (2022), and Wang and Liu (2021)

Model Setup

Suppose x_{it} is subject to a structural break at k, for some indexing variable k_t:

$$X_t = \begin{cases} \Lambda_1 f_t + e_t, & \text{for} \quad k_t \leq k, \\ \Lambda_2 f_t + e_t, & \text{for} \quad k_t > k, \end{cases}$$

- f_t is a $r \times 1$ vector of factors
- $\Lambda_1 = (\lambda_{1,1}, \ldots, \lambda_{1,N})^\intercal$ and $\Lambda_2 = (\lambda_{2,1}, \ldots, \lambda_{2,N})^\intercal$ are corresponding $N \times r$ pre and post-break loading matrices
- e_t is idiosyncratic shock w/ mild serial and cross sectional correlation

Estimation of r and k

Both r and k can be consistently estimated, and hence can be treated as known *a priori* without affecting the asymptotic theory.

Projection-Based Decomposition (Reparameterization)

$$\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}. \tag{1.3}$$

Projection-Based Decomposition (Reparameterization)

$$\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}. \tag{1.3}$$

- Z represents an $r \times r$ rotational change
  - Factor heteroskedasticity: some suitable twisting/stretching of factor space corresponding to rotation Z, Massacci (2021), Pelger and Xiong (2022), and Wang and Liu (2021)

Projection-Based Decomposition (Reparameterization)

$$\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}. \quad (1.3)$$

- Z represents an $r \times r$ rotational change
  - Factor heteroskedasticity: some suitable twisting/stretching of factor space corresponding to rotation Z, Massacci (2021), Pelger and Xiong (2022), and Wang and Liu (2021)
- W represents an $N \times r$ near *orthogonal* shift, s.t. $\Lambda_1^{\intercal} W = O_p(1)$
  - Break in loadings: outside the space spanned by the factors, some leftover shift operation, Massacci (2021)

Projection-Based Decomposition (Reparameterization)

$$\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}. \tag{1.3}$$

- Z represents an $r \times r$ rotational change
  - Factor heteroskedasticity: some suitable twisting/stretching of factor space corresponding to rotation Z, Massacci (2021), Pelger and Xiong (2022), and Wang and Liu (2021)
- W represents an $N \times r$ near *orthogonal* shift, s.t. $\Lambda_1^{\intercal} W = O_p(1)$
  - Break in loadings: outside the space spanned by the factors, some leftover shift operation, Massacci (2021)

Change in the number of factors

Focus on case of nonsingular $r \times r$ dimensional Z.
Change in the number of factors can be accommodated with an $r_2 \times r_1$ "rectangular" Z.

Projection Based Equivalent Representation Theorem

Decompose Λ_2 via a *projection* as: $\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}$ to yield:

Projection Based Equivalent Representation Theorem

Decompose Λ_2 via a *projection* as: $\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}$ to yield:

$$X = \begin{bmatrix} F_1 & 0 \\ F_2 Z^\intercal & F_2 \end{bmatrix} \begin{bmatrix} \Lambda_1^\intercal \\ W^\intercal \end{bmatrix} + \begin{bmatrix} e_{(1)} \\ e_{(2)} \end{bmatrix}$$

$$X = G\Xi^\intercal + e. \tag{1.4}$$

Ignoring the break will cause PC to estimate *pseudo* factors G and loadings Ξ.

Projection Based Equivalent Representation Theorem

Decompose Λ_2 via a *projection* as: $\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}$ to yield:

$$X = \begin{bmatrix} F_1 & 0 \\ F_2 Z^\intercal & F_2 \end{bmatrix} \begin{bmatrix} \Lambda_1^\intercal \\ W^\intercal \end{bmatrix} + \begin{bmatrix} e_{(1)} \\ e_{(2)} \end{bmatrix}$$

$$X = G\Xi^\intercal + e. \tag{1.4}$$

Ignoring the break will cause PC to estimate *pseudo* factors G and loadings Ξ.

- Rotational changes corresponding to $Z \neq I$ are absorbed into the factors
- Shift changes corresponding to $W \neq 0$ result in augmentation of the factor space
- Broader applications and utility in modeling structural changes in factor models

Projection Based Equivalent Representation Theorem

Decompose Λ_2 via a *projection* as: $\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}$ to yield:

$$X = \begin{bmatrix} F_1 & 0 \\ F_2 Z^\intercal & F_2 \end{bmatrix} \begin{bmatrix} \Lambda_1^\intercal \\ W^\intercal \end{bmatrix} + \begin{bmatrix} e_{(1)} \\ e_{(2)} \end{bmatrix}$$

$$X = G\Xi^\intercal + e. \tag{1.4}$$

Ignoring the break will cause PC to estimate *pseudo* factors G and loadings Ξ.

- Rotational changes corresponding to $Z \neq I$ are absorbed into the factors
- Shift changes corresponding to $W \neq 0$ result in augmentation of the factor space
- Broader applications and utility in modeling structural changes in factor models

Existing literature at large uses estimate of G - unable to differentiate between breaks.

Hypothesis Tests to Disentangle Breaks

Existing literature cannot *differentiate* between the types of breaks.

$$\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}.$$

Hypothesis Tests to Disentangle Breaks

Existing literature cannot *differentiate* between the types of breaks.

$$\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}.$$

Disentanglement is possible via two hypothesis tests:

1. Test for evidence of rotations:

$$\mathcal{H}_0 : \Sigma_F = Z\Sigma_F Z^\intercal, \quad \mathcal{H}_1 : \Sigma_F \neq Z\Sigma_F Z^\intercal, \tag{1.5}$$

where $\Sigma_F = E(f_t f_t^\intercal)$

Hypothesis Tests to Disentangle Breaks

Existing literature cannot *differentiate* between the types of breaks.

$$\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}.$$

Disentanglement is possible via two hypothesis tests:

1. Test for evidence of rotations:

$$\mathcal{H}_0 : \Sigma_F = Z\Sigma_F Z^\intercal, \quad \mathcal{H}_1 : \Sigma_F \neq Z\Sigma_F Z^\intercal, \tag{1.5}$$

   where $\Sigma_F = E(f_t f_t^\intercal)$

2. Test for evidence of shifts:

$$\mathcal{H}_0 : W_{(N\times r)} = \mathbf{0}, \quad \mathcal{H}_1 : W_{(N\times r)} \neq \mathbf{0}. \tag{1.6}$$

Hypothesis Tests to Disentangle Breaks

Existing literature cannot *differentiate* between the types of breaks.

$$\Lambda_2 = \Lambda_1 \underbrace{Z}_{\text{rotation}} + \underbrace{W}_{\text{shift}}.$$

Disentanglement is possible via two hypothesis tests:

1. Test for evidence of rotations:

$$\mathcal{H}_0 : \Sigma_F = Z\Sigma_F Z^\intercal, \quad \mathcal{H}_1 : \Sigma_F \neq Z\Sigma_F Z^\intercal, \tag{1.5}$$

   where $\Sigma_F = E(f_t f_t^\intercal)$

2. Test for evidence of shifts:

$$\mathcal{H}_0 : W_{(N\times r)} = \mathbf{0}, \quad \mathcal{H}_1 : W_{(N\times r)} \neq \mathbf{0}. \tag{1.6}$$

Structural Break Tests

Focus on structural break setup, $k_t = t$ indexing variable, break fraction s.t. $k = \lfloor \pi T \rfloor$.

Estimation

1. Estimate $\tilde{F}_1$ and $\tilde{F}_2$ via PC

$$\frac{1}{T_1}\tilde{F}_1^\intercal \tilde{F}_1 = \frac{1}{T_2}\tilde{F}_2^\intercal \tilde{F}_2 = I_r$$

where $\tilde{F}_m$ are $\sqrt{T_m}$ times the first r eigenvectors of $X_m X_m^\intercal$ for $m = 1, 2$.

Estimation

1. Estimate $\tilde{F}_1$ and $\tilde{F}_2$ via PC

$$\frac{1}{T_1}\tilde{F}_1^\intercal \tilde{F}_1 = \frac{1}{T_2}\tilde{F}_2^\intercal \tilde{F}_2 = I_r$$

   where $\tilde{F}_m$ are $\sqrt{T_m}$ times the first r eigenvectors of $X_m X_m^\intercal$ for $m = 1, 2$.

2. Conditional factors, estimate $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$ via OLS

$$\tilde{\Lambda}_1 = X_1^\intercal \tilde{F}_1 \left(\tilde{F}_1^\intercal \tilde{F}_1\right)^{-1} = \frac{1}{T_1} X_1^\intercal \tilde{F}_1, \qquad \tilde{\Lambda}_2 = X_2^\intercal \tilde{F}_2 \left(\tilde{F}_2^\intercal \tilde{F}_2\right)^{-1} = \frac{1}{T_2} X_2^\intercal \tilde{F}_2.$$

Estimation

1. Estimate $\tilde{F}_1$ and $\tilde{F}_2$ via PC

$$\frac{1}{T_1}\tilde{F}_1^\intercal \tilde{F}_1 = \frac{1}{T_2}\tilde{F}_2^\intercal \tilde{F}_2 = I_r$$

where $\tilde{F}_m$ are $\sqrt{T_m}$ times the first r eigenvectors of $X_m X_m^\intercal$ for $m = 1, 2$.

2. Conditional factors, estimate $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$ via OLS

$$\tilde{\Lambda}_1 = X_1^\intercal \tilde{F}_1 \left(\tilde{F}_1^\intercal \tilde{F}_1\right)^{-1} = \frac{1}{T_1} X_1^\intercal \tilde{F}_1, \qquad \tilde{\Lambda}_2 = X_2^\intercal \tilde{F}_2 \left(\tilde{F}_2^\intercal \tilde{F}_2\right)^{-1} = \frac{1}{T_2} X_2^\intercal \tilde{F}_2.$$

3. Estimate rotation and shift as

$$\tilde{Z} = \left(\tilde{\Lambda}_1^\intercal \tilde{\Lambda}_1\right)^{-1} \tilde{\Lambda}_1^\intercal \tilde{\Lambda}_2,$$
$$\tilde{W} = \tilde{\Lambda}_1 - \tilde{\Lambda}_1 \tilde{Z}$$

Estimation - Assumptions

For $m = 1, 2$ regimes,

Assumption 1. $E\|f_t\|^4 < \infty$, $E(f_t f_t^\top) = \Sigma_F$ for some $\Sigma_F > 0$.

Assumption 2. $E\|\lambda_{m,i}\|^4 \leq M$, $\left\|\frac{\Lambda_m^\top \Lambda_m}{N} - \Sigma_{\Lambda_m}\right\| \xrightarrow{p} 0$ for some $\Sigma_{\Lambda_m} > 0$

Assumption 3. Moments of idiosyncratic errors

Assumption 4. $\{\lambda_{m,i}\}$, $\{f_t\}$ and $\{e_{it}\}$ are mutually independent groups.

Assumption 5. Weak serial and cross sectional correlation in errors

Assumption 6. Subsample version of Assumption F in Bai (2003)

Assumption 7. The eigenvalues of $(\Sigma_{\Lambda_1}\Sigma_F)$ and $(\Sigma_{\Lambda_2}\Sigma_F)$ are distinct.

Assumption 8. Break fraction π is bounded away from 0 and 1.

Notes

Assumption 1 assumes "strict" stationarity, but this is not restrictive because changes in factor variance are characterised by Z.

Notation: $\delta_{NT} = min(\sqrt{N}, \sqrt{T})$.

Estimation - Results

Theorem 1. Under Assumptions 1 to 8, as $N, T \to \infty$

$$\left\| \tilde{Z} - H_1^{\intercal} Z H_2^{-\intercal} \right\| = O_p\left(\delta_{NT}^{-2}\right),$$

$$\frac{1}{N}\left\| \tilde{W} - W H_2^{-\intercal} \right\|^2 = O_p\left(\delta_{NT}^{-2}\right).$$

Estimation - Results

Theorem 1. Under Assumptions 1 to 8, as $N, T \to \infty$

$$\left\| \tilde{Z} - H_1^{\intercal} Z H_2^{-\intercal} \right\| = O_p\left(\delta_{NT}^{-2}\right),$$
$$\frac{1}{N} \left\| \tilde{W} - W H_2^{-\intercal} \right\|^2 = O_p\left(\delta_{NT}^{-2}\right).$$

Principal components estimates $\tilde{F}_2$ which is consistent for $F_2 H_2$.

- Define combined series $\widehat{F} = [\tilde{F}_1^{\intercal}, \tilde{Z}\tilde{F}_2^{\intercal}]^{\intercal}$
  - Combined series $\widehat{F} = (\widehat{f}_1, \ldots, \widehat{f}_T)^{\intercal}$ is on the same rotational basis both before and after the break, is also free from the effects of W.

Estimation - Results

Theorem 1. Under Assumptions 1 to 8, as $N, T \to \infty$

$$\left\| \tilde{Z} - H_1^{\intercal} Z H_2^{-\intercal} \right\| = O_p\left(\delta_{NT}^{-2}\right),$$

$$\frac{1}{N}\left\| \tilde{W} - W H_2^{-\intercal} \right\|^2 = O_p\left(\delta_{NT}^{-2}\right).$$

Principal components estimates $\tilde{F}_2$ which is consistent for $F_2 H_2$.

- Define combined series $\widehat{F} = [\tilde{F}_1^{\intercal}, \tilde{Z}\tilde{F}_2^{\intercal}]^{\intercal}$
  - Combined series $\widehat{F} = (\widehat{f}_1, \ldots, \widehat{f}_T)^{\intercal}$ is on the same rotational basis both before and after the break, is also free from the effects of W.

If $W = 0_{N\times r}$, then $\tilde{W}$ should also be close to zero.

Test Statistic for $\mathcal{H}_0 : \Sigma_F = Z\Sigma_F Z^\intercal$ vs $\mathcal{H}_1 : \Sigma_F \neq Z\Sigma_F Z^\intercal$

Recall combined series $\widehat{F} = \left[\widehat{f}_t, \ldots, \widehat{f}_T\right]^\intercal$:

$$\widehat{f}_t = \begin{cases} \tilde{f}_{1,t} & \text{for} \quad t = 1, \ldots, \lfloor \pi T \rfloor, \\ \tilde{Z}\tilde{f}_{2,t} & \text{for} \quad t = \lfloor \pi T \rfloor + 1, \ldots, T. \end{cases}$$

Test Statistic for $\mathcal{H}_0 : \Sigma_F = Z\Sigma_F Z^\intercal$ vs $\mathcal{H}_1 : \Sigma_F \neq Z\Sigma_F Z^\intercal$

Recall combined series $\widehat{F} = \left[\widehat{f}_t, \ldots, \widehat{f}_T\right]^\intercal$:

$$\widehat{f}_t = \begin{cases} \tilde{f}_{1,t} & \text{for} \quad t = 1, \ldots, \lfloor \pi T \rfloor, \\ \tilde{Z}\tilde{f}_{2,t} & \text{for} \quad t = \lfloor \pi T \rfloor + 1, \ldots, T. \end{cases}$$

Consider using subsample means of second moments process:

$$A_Z(\pi, \widehat{F}) = vech\left(\sqrt{T}\left(\frac{1}{\lfloor \pi T \rfloor}\sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t \widehat{f}_t^\intercal - \frac{1}{T - \lfloor \pi T \rfloor}\sum_{t=\lfloor \pi T + 1 \rfloor}^{T} \widehat{f}_t \widehat{f}_t^\intercal\right)\right). \tag{1.7}$$

Test Statistic for $\mathcal{H}_0 : \Sigma_F = Z\Sigma_F Z^\intercal$ vs $\mathcal{H}_1 : \Sigma_F \neq Z\Sigma_F Z^\intercal$

Recall combined series $\widehat{F} = \left[\widehat{f}_t, \ldots, \widehat{f}_T\right]^\intercal$:

$$\widehat{f}_t = \begin{cases} \tilde{f}_{1,t} & \text{for} \quad t = 1, \ldots, \lfloor \pi T \rfloor, \\ \tilde{Z}\tilde{f}_{2,t} & \text{for} \quad t = \lfloor \pi T \rfloor + 1, \ldots, T. \end{cases}$$

Consider using subsample means of second moments process:

$$A_Z(\pi, \widehat{F}) = vech\left(\sqrt{T}\left(\frac{1}{\lfloor \pi T \rfloor}\sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t\widehat{f}_t^\intercal - \frac{1}{T - \lfloor \pi T \rfloor}\sum_{t=\lfloor \pi T+1 \rfloor}^{T} \widehat{f}_t\widehat{f}_t^\intercal\right)\right). \qquad (1.7)$$

Define Wald test statistics for evidence of rotational change as:

$$\mathscr{W}_Z(\pi, \widehat{F}) = A_Z(\pi, \widehat{F})^\intercal \widehat{S}_Z(\pi, \widehat{F})^{-1} A_Z(\pi, \widehat{F}), \qquad (1.8)$$

where $\widehat{S}_Z(\pi, \widehat{F}) = \frac{1}{\pi}\widehat{\Omega}_{Z,(1)} + \frac{1}{1-\pi}\widehat{\Omega}_{Z,(2)}$, $\widehat{\Omega}_{Z,(1)}$, $\widehat{\Omega}_{Z,(2)}$ are estimates of long run variance.

Rotational Test - Null Distribution

Assumption 9. The Bartlett kernel of Newey and West (1987) is used.
Assumption 10. $\mathscr{W}_Z(\pi, FH_{0,1}) \Rightarrow Q_p(\pi)$, where
$Q_p(\pi) = [B_p(\pi) - \pi B_p(1)]^\intercal [B_p(\pi) - \pi B_p(1)]/(\pi(1-\pi))$, and $B_p(\cdot)$ is a $p = \frac{r(r+1)}{2}$ vector of independent Brownian motions on $[0, 1]$.

$Q_p(\pi) \sim \chi^2_p$ for fixed π implies the null distribution:

Rotational Test - Null Distribution

Assumption 9. The Bartlett kernel of Newey and West (1987) is used.
Assumption 10. $\mathscr{W}_Z(\pi, FH_{0,1}) \Rightarrow Q_p(\pi)$, where
$Q_p(\pi) = [B_p(\pi) - \pi B_p(1)]^\intercal [B_p(\pi) - \pi B_p(1)]/(\pi(1-\pi))$, and $B_p(\cdot)$ is a $p = \frac{r(r+1)}{2}$ vector of independent Brownian motions on $[0, 1]$.

$Q_p(\pi) \sim \chi^2_p$ for fixed π implies the null distribution:

Theorem 2. Under Assumptions 1 to 10, and if $\frac{\sqrt{T}}{N} \to 0$, then

$$\mathscr{W}_Z(\pi, \widehat{F}) \xrightarrow{d} \chi^2_{r(r+1)/2}.$$

Rotational Test - Consistency

Assumption 11. $Z\Sigma_F Z^\intercal \neq \Sigma_F$.
Assumption 12. $plim_{T\to\infty} \inf \left(vech(C)^\intercal \left[\max(b_{\lfloor \pi T \rfloor}, b_{T-\lfloor \pi T \rfloor}) \widehat{S}(F^* H_{0,1})^{-1} \right] vech(C) \right) > 0$,
where $C \equiv H_{0,1}^\intercal (\Sigma_F - Z\Sigma_F Z^\intercal) H_{0,1}$.

Theorem 3. Under Assumptions 1 to 9 and 12, and if Z satisfies Assumption 11, then

1. there exists some nonrandom matrix $C \neq 0$ such that

$$\frac{1}{\pi T} \sum_{t=1}^{\lfloor \pi T \rfloor} \widehat{f}_t \widehat{f}_t^\intercal - \frac{1}{T - \lfloor \pi T \rfloor} \sum_{t=\lfloor \pi T + 1 \rfloor}^{T} \widehat{f}_t \widehat{f}_t^\intercal \xrightarrow{p} C,$$

2. the test statistic $\mathscr{W}_Z(\pi, \widehat{F})$ is consistent under the alternative hypothesis that $\Sigma_F \neq Z\Sigma_F Z^\intercal$.

Orthogonal Shift Test $\mathcal{H}_0 : W = \mathbf{0}$ vs $\mathcal{H}_1 : W \neq \mathbf{0}$

Note that hypothesis is *infeasible* due to $N \to \infty$.

Orthogonal Shift Test $\mathcal{H}_0 : W = \mathbf{0}$ vs $\mathcal{H}_1 : W \neq \mathbf{0}$

Note that hypothesis is *infeasible* due to $N \to \infty$.
Strategy: consider individual and cross sectionally averaged Wald test statistics:

$$\mathscr{W}_{W,i} = (T)(\tilde{w}_i)^\intercal(\tilde{\Omega}_{W,i})^{-1}(\tilde{w}_i), \tag{2.1}$$

$$\mathscr{W}_{W} = (TN)\left(\frac{\sum_{i=1}^{N} \tilde{w}_i}{N}\right)^\intercal (\tilde{\Omega}_W)^{-1}\left(\frac{\sum_{i=1}^{N} \tilde{w}_i}{N}\right). \tag{2.2}$$

Orthogonal Shift Test $\mathcal{H}_0 : W = \mathbf{0}$ vs $\mathcal{H}_1 : W \neq \mathbf{0}$

Note that hypothesis is *infeasible* due to $N \to \infty$.
Strategy: consider individual and cross sectionally averaged Wald test statistics:

$$\mathscr{W}_{W,i} = (T)(\tilde{w}_i)^\intercal(\tilde{\Omega}_{W,i})^{-1}(\tilde{w}_i), \tag{2.1}$$

$$\mathscr{W}_W = (TN)\left(\frac{\sum_{i=1}^N \tilde{w}_i}{N}\right)^\intercal (\tilde{\Omega}_W)^{-1}\left(\frac{\sum_{i=1}^N \tilde{w}_i}{N}\right). \tag{2.2}$$

Assumption 13. Additional pooled moment conditions on errors
Assumption 14. Additional pooled moment conditions of Assumption F of Bai (2003)
Assumption 15. Additional pooled moment conditions for loadings
Assumption 16.

1. $\frac{1}{\sqrt{T}}\sum_{t=1}^T f_t e_{it} \xrightarrow{d} N(0, \Phi_{W,i})$, and $(T)^{-1}\sum_{t=1}^T f_t f_t^\intercal e_{it}^2 \xrightarrow{p} \Phi_{W,i}$
2. $\frac{1}{\sqrt{TN}}\sum_{t=1}^T \sum_{i=1}^N f_t e_{it} \xrightarrow{d} N(0, \Phi_W)$, and $(TN)^{-1}\sum_{t=1}^T \sum_{i=1}^N f_t f_t^\intercal e_{it}^2 \xrightarrow{p} \Phi_W$

Orthogonal Shift Test - Null Distribution & Power

Theorem 4. If $\frac{\sqrt{T}}{N} \to 0$, then:

1. Under Assumptions 1 to 9, and additionally Assumptions 13, 14 and 16, $\mathscr{W}_{W,i} \xrightarrow{d} \chi^2_r$ for each i, and
2. Under Assumptions 1 to 9, and additionally Assumptions 13 to 16, $\mathscr{W}_W \xrightarrow{d} \chi^2_r$.

Orthogonal Shift Test - Null Distribution & Power

Theorem 4. If $\frac{\sqrt{T}}{N} \to 0$, then:

1. Under Assumptions 1 to 9, and additionally Assumptions 13, 14 and 16, $\mathscr{W}_{W,i} \xrightarrow{d} \chi^2_r$ for each i, and
2. Under Assumptions 1 to 9, and additionally Assumptions 13 to 16, $\mathscr{W}_W \xrightarrow{d} \chi^2_r$.

Assumption 17. There exists a constants $0 < \alpha \leq 0.5$ and $C > 0$ such that as $N, T \to \infty$,

$$Pr\left(\left\| \frac{T^{\alpha/2}}{\sqrt{N}} \sum_{i=1}^{N} w_i \right\| > C \right) \to 1. \tag{2.3}$$

Orthogonal Shift Test - Null Distribution & Power

Theorem 4. If $\frac{\sqrt{T}}{N} \to 0$, then:

1. Under Assumptions 1 to 9, and additionally Assumptions 13, 14 and 16, $\mathscr{W}_{W,i} \xrightarrow{d} \chi^2_r$ for each i, and
2. Under Assumptions 1 to 9, and additionally Assumptions 13 to 16, $\mathscr{W}_W \xrightarrow{d} \chi^2_r$.

Assumption 17. There exists a constants $0 < \alpha \leq 0.5$ and $C > 0$ such that as $N, T \to \infty$,

$$Pr\left(\left\| \frac{T^{\alpha/2}}{\sqrt{N}} \sum_{i=1}^{N} w_i \right\| > C \right) \to 1. \tag{2.3}$$

Theorem 5. Suppose that $\frac{\sqrt{T}}{N} \to 0$, and the alternative hypothesis $\mathcal{H}_1 : W \neq 0$ holds. Then:

1. under Assumptions 1 to 8, 13, 14 and 16, and if $w_i \neq 0$, then $\mathscr{W}_{W,i} \to \infty$ as $N, T \to \infty$
2. under Assumptions 1 to 8 and 13 to 17, $\mathscr{W}_W \to \infty$ if $\frac{\sqrt{N}}{T^{1-\alpha/2}} \to 0$ as $N, T \to \infty$.

Monte Carlo Study

1. Simulate $W = \Lambda_2 - (\Lambda_1^\intercal \Lambda_1)^{-1}\Lambda_1^\intercal \Lambda_2$, $\Lambda_1, \Lambda_2 \sim MVN(0_3, I_3)$
2. $Z = I$ for no break, or a lower triangular matrix with $[2.5, 1.5, 0.5]$ on the main diagonal and its lower triangular entries drawn from $N(0, 1)$, (Duan et al. (2022))
3. $\omega = 0$ for no break, otherwise controls "size" or orthogonal shift

$$x_{it} = \begin{cases} (\lambda_{1,i})^\intercal f_t + \sqrt{\theta} e_{it}, & t = 1, \dots, \lfloor \pi T \rfloor \\ (Z\lambda_{1,i} + \omega w_i)^\intercal f_t + \sqrt{\theta} e_{it}, & t = \lfloor \pi T \rfloor + 1, \dots, T. \end{cases} \tag{3.1}$$

AR(1) Factors and errors:

$$f_{k,t} = \rho f_{k,t-1} + \mu_{it}, \mu_{it} \sim i.i.d.N(0, 1-\rho^2), \tag{3.2}$$

$$e_{it} = \alpha e_{i,t-1} + v_{it}, \tag{3.3}$$

where $v_t = v_{1,t}, \dots, v_{N,t}^\intercal$ being i.i.d. $N(0, \Omega)$ for $t = 2, \dots, T$. Cross sectional dependence set by $\Omega_{ij} = \beta^{|i-j|}$, serial correlation controlled by α.

Table 1: Size of Rotation and Orthogonal Shift Tests, $N = 200, r = 3$, nominal 5%

T	ρ	α	β	Z Test Unadj.	Z Test Adj.	W Test Unadj.	W Test Adj.	W Individual
200	0.0	0.0	0.0	0.283	0.212	0.150	0.123	0.027
			0.3	0.052	0.033	0.064	0.033	0.005
		0.3	0.0	0.326	0.266	0.182	0.141	0.030
			0.3	0.147	0.101	0.092	0.055	0.016
500		0.0	0.0	0.133	0.087	0.003	0.001	0.011
			0.3	0.045	0.030	0.044	0.019	0.003
		0.3	0.0	0.139	0.088	0.003	0.000	0.012
			0.3	0.108	0.058	0.064	0.032	0.007
200	0.7	0.0	0.0	0.217	0.160	0.147	0.106	0.027
			0.3	0.219	0.150	0.061	0.038	0.017
		0.3	0.0	0.237	0.177	0.187	0.150	0.042
			0.3	0.215	0.154	0.087	0.062	0.029
500		0.0	0.0	0.158	0.085	0.005	0.000	0.010
			0.3	0.145	0.100	0.051	0.023	0.005
		0.3	0.0	0.155	0.090	0.009	0.002	0.014
			0.3	0.134	0.085	0.062	0.042	0.011

Table 2: Power of Z and W Tests, $r = 3$, $N = 200$, $\alpha = \beta = 0.3$

Break Type	T	ω	ρ	Z Test Unadj.	Z Test Adj.	W Test Unadj.	W Test Adj.	W Test Individual	HI	BKW	$\tilde{r}$
$W \neq 0$	200	1	0.0	0.136	0.129	0.860	0.821	0.849	1.000	1.000	5.928
			0.7	0.244	0.233	0.916	0.896	0.908	1.000	1.000	6.000
	500		0.0	0.079	0.076	0.950	0.939	0.947	1.000	1.000	6.000
			0.7	0.146	0.144	0.968	0.965	0.968	1.000	1.000	6.000
$Z \neq I$	200	0	0.0	1.000	1.000	0.100	0.100	0.026	1.000	1.000	3.000
			0.7	1.000	1.000	0.106	0.106	0.035	1.000	1.000	3.000
	500		0.0	1.000	1.000	0.094	0.094	0.009	1.000	1.000	3.000
			0.7	1.000	1.000	0.096	0.096	0.012	1.000	1.000	3.000
$W \neq 0$ and $Z \neq I$	200	1	0.0	1.000	1.000	0.804	0.803	0.765	1.000	1.000	4.206
			0.7	1.000	1.000	0.867	0.867	0.846	1.000	1.000	5.047
	500		0.0	1.000	1.000	0.919	0.919	0.901	1.000	1.000	4.772
			0.7	1.000	1.000	0.946	0.946	0.938	1.000	1.000	5.511

Note:

Entries denote the rejection rates across different simulated break types; a break type of W denotes a break in the factor loadings, Z a break in the factor variance, and W and Z denoting a break in both. HI denotes Han and Inoue (2015)'s test, and BKW denotes Baltagi et al (2021)'s test. The scalar ω denotes the "size" of the break in the loadings.

Empirical Study

FRED-QD Dataset 1959Q3 - 2019Q4, McCracken and Ng (2020)

Great Moderation (1984Q1)

- Documented decrease in variance of *all* series
- Considered *a priori* by Stock and Watson (2009); dated by Baltagi et al. (2021), Breitung and Eickmeier (2011), and Chen et al. (2014)

Great Recession (2008Q3)

- Baltagi et al. (2021), Duan et al. (2022), and Ma and Su (2018), and others

Consider 2-6 factors

Empirical Joint Test Results

Table 3: Joint Test Results

	Z Test p values		W Test p values			
$\tilde{r}$	Unadjusted	Adjusted	Unadjusted	Adjusted	Han and Inoue (2015)	Baltagi et al (2021)
Great Moderation (1984 Q1), 1959 Q3 - 2008 Q3 Sample						
2	0.001	0.001	0.000	0.000	0.633	0.097
3	0.000	0.000	0.000	0.000	0.000	0.004
4	0.008	0.007	0.000	0.000	0.000	0.001
5	0.000	0.000	0.001	0.001	0.000	0.001
6	0.000	0.000	0.000	0.000	0.000	0.002
Great Recession (2008 Q3), 1984 Q2 - 2019 Q4 Sample						
2	0.000	0.000	0.004	0.004	0.183	0.012
3	0.000	0.000	0.000	0.000	0.019	0.006
4	0.000	0.000	0.000	0.000	0.000	0.005
5	0.000	0.000	0.000	0.000	0.000	0.014
6	0.000	0.000	0.000	0.000	0.000	0.074

Factor Heteroskedasticity

Table 4: Estimated ratio of the factor variances

r	$tr(\tilde{Z}\tilde{Z}^\intercal)/tr(I_r)$	95% Bootstrap Confidence Interval
Great Moderation (1984 Q1), 1959 Q3 - 2008 Q3 Sample		
2	0.255	[0.185, 0.269]
3	0.294	[0.189, 0.306]
4	0.347	[0.239, 0.353]
5	0.306	[0.23, 0.324]
6	0.289	[0.223, 0.301]
Great Recession (2008 Q3), 1984 Q2 - 2019 Q4 Sample		
2	0.893	[0.888, 1.089]
3	1.303	[0.762, 1.49]
4	1.208	[0.957, 1.375]
5	1.097	[0.913, 1.158]
6	1.030	[0.894, 1.112]

Note:
The table presents estimates of the ratio of the total factor variance pre and post-break, or $tr(\Sigma_F)/tr(Z\Sigma_F Z^\intercal)$.

Which Series had breaks in their loadings?

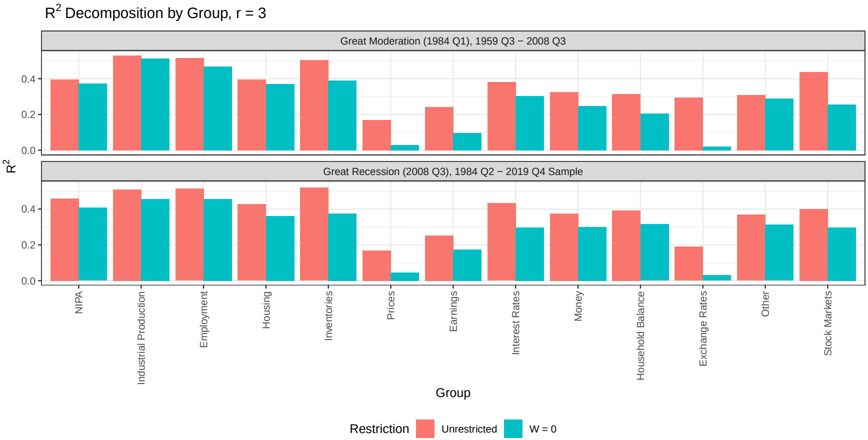

Figure 1: R^2 Statistics for unrestricted and restricted common component ($W = 0$) for Great Moderation Subsample, and Global Financial Crisis Subsample, for $r = 3$.

Great Moderation: Re-interpretation

$$X_t = \Lambda f_t + e_t,$$
$$f_t = \sum_{j=1}^{p} \Phi_j f_{t-j} + \eta_t, \quad \eta_t \sim (0, \Sigma_\eta).$$

"Good Luck" or "Good Policy"?

- Good Luck: smaller fortuitous shocks hitting economy $\Leftrightarrow$ break in Σ_η
- Good Policy: parameters not related to η, i.e. Φ_j or Λ

Great Moderation: Re-interpretation

$$X_t = \Lambda f_t + e_t,$$
$$f_t = \sum_{j=1}^{p} \Phi_j f_{t-j} + \eta_t, \quad \eta_t \sim (0, \Sigma_\eta).$$

"Good Luck" or "Good Policy"?

- Good Luck: smaller fortuitous shocks hitting economy $\Leftrightarrow$ break in Σ_η
- Good Policy: parameters not related to η, i.e. Φ_j or Λ

Nuanced Interpretation

Caution: break in Σ_F could also be from Φ_j.
Nonetheless, highlights importance of allowing/modeling breaks in factor variance

Conclusion

- Establish a new *projection*-based equivalent representation theorem to decompose any break into a rotational change (factor variance), and shift (loadings)

Conclusion

- Establish a new *projection*-based equivalent representation theorem to decompose any break into a rotational change (factor variance), and shift (loadings)
- Propose two separate tests: 1) evidence of rotational change and 2) evidence of shifts

Conclusion

- Establish a new *projection*-based equivalent representation theorem to decompose any break into a rotational change (factor variance), and shift (loadings)
- Propose two separate tests: 1) evidence of rotational change and 2) evidence of shifts
- Monte Carlo shows good size and power properties, and inability of existing tests to *differentiate* between these breaks

Conclusion

- Establish a new *projection*-based equivalent representation theorem to decompose any break into a rotational change (factor variance), and shift (loadings)
- Propose two separate tests: 1) evidence of rotational change and 2) evidence of shifts
- Monte Carlo shows good size and power properties, and inability of existing tests to *differentiate* between these breaks
- Evidence of both breaks on data
- Suggest more nuanced interpretation of Great Moderation - most variation explained by breaks in factor variance