Structural Instabilities in Factor Models for Time Series

Ragvir Sabharwal

London School of Economics (LSE)

24 February 2016

Outline

Generalised Dynamic Factor Models (GDFMs)

- Co-movements of (large) number of observed time series modelled in terms of small number of unobserved common factors.
- Literature on high-dimensional GDFMs focusses on case where parameters are stable; in particular, where factor loadings are time-invariant.

Outline

Generalised Dynamic Factor Models (GDFMs)

- Co-movements of (large) number of observed time series modelled in terms of small number of unobserved common factors.
- Literature on high-dimensional GDFMs focusses on case where parameters are stable; in particular, where factor loadings are time-invariant.

Outline of today's discussion

- Debate: do structural instabilities matter?

Outline

Generalised Dynamic Factor Models (GDFMs)

- Co-movements of (large) number of observed time series modelled in terms of small number of unobserved common factors.
- Literature on high-dimensional GDFMs focusses on case where parameters are stable; in particular, where factor loadings are time-invariant.

Outline of today's discussion

- Debate: do structural instabilities matter?
- Model: with/without structural breaks

Outline

Generalised Dynamic Factor Models (GDFMs)

- Co-movements of (large) number of observed time series modelled in terms of small number of unobserved common factors.
- Literature on high-dimensional GDFMs focusses on case where parameters are stable; in particular, where factor loadings are time-invariant.

Outline of today's discussion

- Debate: do structural instabilities matter?
- Model: with/without structural breaks
- Roadmap: what the key players are saying

Outline

Generalised Dynamic Factor Models (GDFMs)

- Co-movements of (large) number of observed time series modelled in terms of small number of unobserved common factors.
- Literature on high-dimensional GDFMs focusses on case where parameters are stable; in particular, where factor loadings are time-invariant.

Outline of today's discussion

- Debate: do structural instabilities matter?
- Model: with/without structural breaks
- Roadmap: what the key players are saying
- Extension: what I plan to do

I. What is the debate about?

Do structural instabilities matter?

NAY!

- Stock & Watson (JASA 2002)
- Bates, Plagborg-Moller, Stock & Watson (JoE 2013)

Claim:

Can consistently estimate the factor space despite structural instabilities

...provided certain conditions hold.

Do structural instabilities matter?

NAY!

- Stock & Watson (JASA 2002)
- Bates, Plagborg-Moller, Stock & Watson (JoE 2013)

Claim:

Can consistently estimate the factor space despite structural instabilities

...provided certain conditions hold.

AYE!

- Breitung & Eickmeier (JoE 2011)
- Corradi & Swanson (Forthcoming, JoE)
- Chen, Dolado & Gonzalo (JoE 2014)
- Han & Inoue ('metric Theory 2014)
- Baltagi, Kao & Wang (Manuscript, 2015)
- Yamamoto & Tanaka (JoE 2015)

Claim:

Stock & Watson conditions not realistic.

II. The standard model, with no breaks

Standard model: $X_t = \chi_t + e_t = \Lambda_t F_t + e_t$

For $t = 1, ..., T$

- Observed series: X_t; $dim(X_t) = N$
- Common component: χ_t; $dim(\chi_t) = N$
- Idiosyncratic error: e_t; $dim(e_t) = N$, with Σ_e not necessarily diagonal
- Common factors: F_t; $dim(F_t) = r$, where $r << N$
- Factor loadings: Λ_t; an $(N \times r)$ matrix, with Λ_0 fixed

Notation alert:

$$\begin{cases} \text{“N”: total number of observed variables (sometimes denoted by “p”)} \\ \text{“T”: total number of time periods (sometimes denoted by “n”)} \end{cases}$$

Standard model: $X_t = \chi_t + e_t = \Lambda_t F_t + e_t$

For $t = 1, ..., T$

- Observed series: X_t; $dim(X_t) = N$
- Common component: χ_t; $dim(\chi_t) = N$
- Idiosyncratic error: e_t; $dim(e_t) = N$, with Σ_e not necessarily diagonal
- Common factors: F_t; $dim(F_t) = r$, where $r << N$
- Factor loadings: Λ_t; an $(N \times r)$ matrix, with Λ_0 fixed

Notation alert:

$$\begin{cases} \text{“N”: total number of observed variables (sometimes denoted by “p”)} \\ \text{“T”: total number of time periods (sometimes denoted by “n”)} \end{cases}$$

Question:

How can we estimate this model? Can we use PCA?

Can use PCA under certain conditions

Under certain assumptions (to be discussed), as $N \to \infty$, the following are equivalent:

- PCA - projection of X_t onto space spanned by the r first principal components;
- Factor Analysis - projection of X_t onto space spanned by the r factors F_t.

Can use PCA under certain conditions

Under certain assumptions (to be discussed), as $N \to \infty$, the following are equivalent:

- PCA - projection of X_t onto space spanned by the r first principal components;
- Factor Analysis - projection of X_t onto space spanned by the r factors F_t.

Emphasis:

- Critical that $N \to \infty$ for equivalence between PCA and Factor Analysis to hold.
- Critical that certain assumptions are not violated. We discuss these now...

Under what conditions can we use PCA estimators? (1 of 2)

Main assumptions; Bai & Ng (Econometrica 2002)

1. Pervasiveness of the factors:
   - $\Lambda_0'\Lambda_0/N \rightarrow D$ as $N \rightarrow \infty$ for some positive definite $(r \times r)$ matrix D.

- Factors affect (most or) all of the observed series, even when the number of series grows very large.

Under what conditions can we use PCA estimators? (1 of 2)

Main assumptions; Bai & Ng (Econometrica 2002)

1. Pervasiveness of the factors:
    - $\Lambda_0'\Lambda_0/N \rightarrow D$ as $N \rightarrow \infty$ for some positive definite $(r \times r)$ matrix D.

- Factors affect (most or) all of the observed series, even when the number of series grows very large.
- F_t can be assumed stationary for simplicity.

Under what conditions can we use PCA estimators? (1 of 2)

Main assumptions; Bai & Ng (Econometrica 2002)

1. Pervasiveness of the factors:
   - $\Lambda_0'\Lambda_0/N \rightarrow D$ as $N \rightarrow \infty$ for some positive definite $(r \times r)$ matrix D.

- Factors affect (most or) all of the observed series, even when the number of series grows very large.
- F_t can be assumed stationary for simplicity.

It is equivalent to assume that:

- Σ_χ has all r eigenvalues diverging as $N \rightarrow \infty$.

Under what conditions can we use PCA estimators? (2 of 2)

2. Mild cross-sectional correlation in idiosyncratic component:
   - $\frac{1}{N}\Sigma_{i,j=1}^{N}|Cov(e_{i,t}, e_{j,t})| \leq M$ for some $M \in (0, \infty)$.

- This assumption allows for the *generalised/approximate* factor structure; yet...

Under what conditions can we use PCA estimators? (2 of 2)

2. Mild cross-sectional correlation in idiosyncratic component:
   - $\frac{1}{N}\Sigma_{i,j=1}^{N}|Cov(e_{i,t}, e_{j,t})| \leq M$ for some $M \in (0, \infty)$.

- This assumption allows for the *generalised/approximate* factor structure; yet...
- it prevents combinations of idiosyncratic terms from having unusually large variation; otherwise it would be difficult to identify genuine factors.

Under what conditions can we use PCA estimators? (2 of 2)

2. Mild cross-sectional correlation in idiosyncratic component:
    - $\frac{1}{N}\Sigma_{i,j=1}^{N}|Cov(e_{i,t}, e_{j,t})| \leq M$ for some $M \in (0, \infty)$.

- This assumption allows for the *generalised/approximate* factor structure; yet...
- it prevents combinations of idiosyncratic terms from having unusually large variation; otherwise it would be difficult to identify genuine factors.
- We can assume that e_t is stationary. No restriction on temporal dependence.

Under what conditions can we use PCA estimators? (2 of 2)

2. Mild cross-sectional correlation in idiosyncratic component:
    - $\frac{1}{N}\Sigma_{i,j=1}^{N}|Cov(e_{i,t}, e_{j,t})| \leq M$ for some $M \in (0, \infty)$.

- This assumption allows for the *generalised/approximate* factor structure; yet...
- it prevents combinations of idiosyncratic terms from having unusually large variation; otherwise it would be difficult to identify genuine factors.
- We can assume that e_t is stationary. No restriction on temporal dependence.
- Heteroscedasticity in both dimensions is also allowed.

It is (almost) equivalent to assume that:

- Σ_e has all eigenvalues bounded for any N.

The assumptions on eigenvalues of Σ_χ and Σ_e imply...

- ...that the r^{th} largest eigenvalue of Σ_x diverges as $N \to \infty$; and
- the $(r+1)^{th}$ largest eigenvalue of Σ_x stays bounded for any N.

The assumptions on eigenvalues of Σ_χ and Σ_e imply...

- ...that the r^{th} largest eigenvalue of Σ_x diverges as $N \to \infty$; and
- the $(r+1)^{th}$ largest eigenvalue of Σ_x stays bounded for any N.

In other words:

PCA may be used, alongside any number of existing criteria for estimating r, in order to consistently recover the space spanned by the r factors.

The assumptions on eigenvalues of Σ_χ and Σ_e imply...

- ...that the r^{th} largest eigenvalue of Σ_x diverges as $N \to \infty$; and
- the $(r+1)^{th}$ largest eigenvalue of Σ_x stays bounded for any N.

In other words:

PCA may be used, alongside any number of existing criteria for estimating r, in order to consistently recover the space spanned by the r factors.

As $N, T \to \infty$

$$||\hat{F}_t - H' F_t|| = O_p\left(max\left\{ \frac{1}{\sqrt{N}}, \frac{1}{\sqrt{T}} \right\} \right)$$

The rate of convergence is determined by the smaller of N or T, and thus depends on the panel structure.

III. The extended model, with breaks in factor loadings

What is the problem?

- A FM with breaks in the loadings can be re-parameterized as another FM with constant loadings but a larger set of factors.

What is the problem?

- A FM with breaks in the loadings can be re-parameterized as another FM with constant loadings but a larger set of factors.
- Consider, for instance, a one-factor model with a structural break:

$$X_{it} = \begin{cases} \lambda_{1i} f_t + e_{it}, t \leq \kappa \\ \lambda_{2i} f_t + e_{it}, t > \kappa \end{cases}$$

What is the problem?

- A FM with breaks in the loadings can be re-parameterized as another FM with constant loadings but a larger set of factors.
- Consider, for instance, a one-factor model with a structural break:

$$X_{it} = \begin{cases} \lambda_{1i} f_t + e_{it}, \, t \leq \kappa \\ \lambda_{2i} f_t + e_{it}, \, t > \kappa \end{cases}$$

- Define $g_{1t} = \begin{cases} f_t, \, t \leq \kappa \\ 0, \, t > \kappa \end{cases}$ and $g_{2t} = f_t - g_{1t}$.

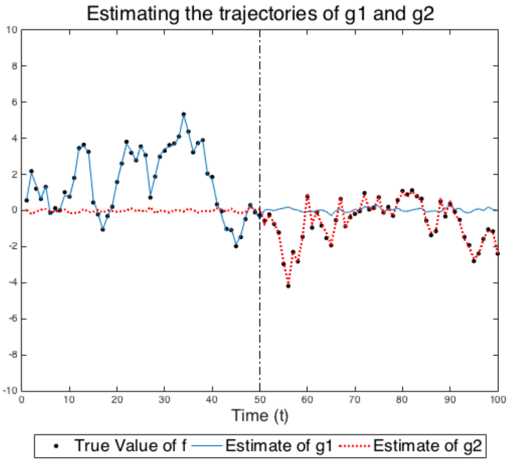
What is the problem?

- A FM with breaks in the loadings can be re-parameterized as another FM with constant loadings but a larger set of factors.
- Consider, for instance, a one-factor model with a structural break:

$$X_{it} = \begin{cases} \lambda_{1i} f_t + e_{it}, t \leq \kappa \\ \lambda_{2i} f_t + e_{it}, t > \kappa \end{cases}$$

- Define $g_{1t} = \begin{cases} f_t, t \leq \kappa \\ 0, t > \kappa \end{cases}$ and $g_{2t} = f_t - g_{1t}$.
- Then, we have $X_{it} = \lambda_{1i} g_{1t} + \lambda_{2i} g_{2t} + e_{it}$, an equivalent stable model.

Graphical illustration with one-factor model



Potential problems

- Existing criteria to determine the number of factors, r, will fail.

$$\begin{cases} \text{Bai \& Ng (Econometrica 2002)} \\ \text{Onatski (RES 2010)} \\ \text{Alessi, Barigozzi \& Capasso (Stat. \& Prob. Letters 2010)} \\ \text{Ahn \& Horenstein (Econometrica 2013)} \end{cases}$$

- Existing methods to estimate the space spanned by the factors will fail.

$$\begin{cases} \text{PCA} \\ \text{Kalman Filter} \end{cases}$$

- Implications:

$$\begin{cases} \text{Pseudo-factors - not necessarily interpretable} \\ \text{Forecasting - performance may deteriorate} \\ \text{Structural analysis - ``impulse responses'' not recoverable} \end{cases}$$

IV. (Moderate) instabilities do not matter

Model with instability: $\begin{cases} X_t = \Lambda_t F_t + e_t \\ \Lambda_t = \Lambda_0 + \xi_t \end{cases}$

- ξ_t: a random process of dimension $N \times r$

For example, a single abrupt break at a common date κ:
Specifically, let $r = 1$, and $\Delta \in \mathbb{R}^N$ be a shift parameter.

We define:

$$\xi_t = \begin{cases} 0 \text{ for } t = 1, ..., \kappa \\ \Delta \text{ for } t = \kappa + 1, ..., T \end{cases}$$

Model with instability: $\begin{cases} X_t = \Lambda_t F_t + e_t \\ \Lambda_t = \Lambda_0 + \xi_t \end{cases}$

- ξ_t: a random process of dimension $N \times r$

For example, a single abrupt break at a common date κ:
Specifically, let $r = 1$, and $\Delta \in \mathbb{R}^N$ be a shift parameter.

We define:

$$\xi_t = \begin{cases} 0 \text{ for } t = 1, ..., \kappa \\ \Delta \text{ for } t = \kappa + 1, ..., T \end{cases}$$

Question:

- Under what (extra) conditions is PCA estimator still consistent...
- ...with the same convergence rate as for stable model?

What additional assumptions do we need?

Bates et al. (JoE 2013):

- Factor loading innovations (breaks)

  1. Limited dependence between breaks and factors themselves; OR
  2. Full independence between breaks and factors, but limited dependence between breaks across series; OR
  3. Perfectly correlated breaks across series, but limited number of series undergoing breaks.

- Specifically, assume $|\Delta_i|$ for $i = 1, ..., N$ is bounded; and
- at most $O(N^{1/2})$ series undergo a break.

Averaging away the instabilities

For “moderate” sized breaks, we can treat the instabilities as simply another additive error term which satisfies the usual assumptions.

$$X_t = \Lambda_t F_t + e_t = (\Lambda_0 + \xi_t) F_t + e_t = \Lambda_0 F_t + e_t + \xi_t F_t$$

Assume a one-factor model with Λ_0 known. Consider $\hat{F}_t = ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}' X_t$.

$$\begin{aligned} \hat{F}_t - F_t &= ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}' e_t + ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}' \xi_t F_t \\ &\approx D^{-1}\left(({\Lambda_0}' e_t / N) + ({\Lambda_0}' \xi_t F_t / N) \right) \end{aligned}$$

Averaging away the instabilities

For "moderate" sized breaks, we can treat the instabilities as simply another additive error term which satisfies the usual assumptions.

$$X_t = \Lambda_t F_t + e_t = (\Lambda_0 + \xi_t) F_t + e_t = \Lambda_0 F_t + e_t + \xi_t F_t$$

Assume a one-factor model with Λ_0 known. Consider $\hat{F}_t = (\Lambda_0{}'\Lambda_0)^{-1}\Lambda_0{}'X_t$.

$$\begin{aligned} \hat{F}_t - F_t &= (\Lambda_0{}'\Lambda_0)^{-1}\Lambda_0{}'e_t + (\Lambda_0{}'\Lambda_0)^{-1}\Lambda_0{}'\xi_t F_t \\ &\approx D^{-1}\Big((\Lambda_0{}'e_t/N) + (\Lambda_0{}'\xi_t F_t/N)\Big) \end{aligned}$$

- Under mild cross-sectional dependence, the first term is $O_p(1/\sqrt{N})$.

Averaging away the instabilities

For "moderate" sized breaks, we can treat the instabilities as simply another additive error term which satisfies the usual assumptions.

$$X_t = \Lambda_t F_t + e_t = (\Lambda_0 + \xi_t) F_t + e_t = \Lambda_0 F_t + e_t + \xi_t F_t$$

Assume a one-factor model with Λ_0 known. Consider $\hat{F}_t = ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}'X_t$.

$$\begin{aligned}\hat{F}_t - F_t &= ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}'e_t + ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}'\xi_t F_t \\ &\approx D^{-1}\Big(({\Lambda_0}'e_t/N) + ({\Lambda_0}'\xi_t F_t/N)\Big)\end{aligned}$$

- Under mild cross-sectional dependence, the first term is $O_p(1/\sqrt{N})$.
- F_t is $O_p(1)$, so the order of the second term depends on the order of ${\Lambda_0}'\xi_t/N$.

Averaging away the instabilities

For "moderate" sized breaks, we can treat the instabilities as simply another additive error term which satisfies the usual assumptions.

$$X_t = \Lambda_t F_t + e_t = (\Lambda_0 + \xi_t) F_t + e_t = \Lambda_0 F_t + e_t + \xi_t F_t$$

Assume a one-factor model with Λ_0 known. Consider $\hat{F}_t = ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}'X_t$.

$$\begin{aligned}\hat{F}_t - F_t &= ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}'e_t + ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}'\xi_t F_t \\ &\approx D^{-1}\Big(({\Lambda_0}'e_t/N) + ({\Lambda_0}'\xi_t F_t/N)\Big)\end{aligned}$$

- Under mild cross-sectional dependence, the first term is $O_p(1/\sqrt{N})$.
- F_t is $O_p(1)$, so the order of the second term depends on the order of ${\Lambda_0}'\xi_t/N$.
- When $t > \kappa$, we have ${\Lambda_0}'\xi_t/N = (\lambda_{10}\Delta_1 + \ldots + \lambda_{N0}\Delta_N)/N$.

Averaging away the instabilities

For "moderate" sized breaks, we can treat the instabilities as simply another additive error term which satisfies the usual assumptions.

$$X_t = \Lambda_t F_t + e_t = (\Lambda_0 + \xi_t)F_t + e_t = \Lambda_0 F_t + e_t + \xi_t F_t$$

Assume a one-factor model with Λ_0 known. Consider $\hat{F}_t = ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}'X_t$.

$$\begin{aligned}\hat{F}_t - F_t &= ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}'e_t + ({\Lambda_0}'\Lambda_0)^{-1}{\Lambda_0}'\xi_t F_t \\ &\approx D^{-1}\Big(({\Lambda_0}'e_t/N) + ({\Lambda_0}'\xi_t F_t/N)\Big)\end{aligned}$$

- Under mild cross-sectional dependence, the first term is $O_p(1/\sqrt{N})$.
- F_t is $O_p(1)$, so the order of the second term depends on the order of ${\Lambda_0}'\xi_t/N$.
- When $t > \kappa$, we have ${\Lambda_0}'\xi_t/N = (\lambda_{10}\Delta_1 + ... + \lambda_{N0}\Delta_N)/N$.
- Thus, this term is $O_p(B/N)$ where B is the number of non-zero Δ_i terms in the summation.

Consistency under structural instability

As above, define B to be the number of series undergoing a break. Then:

$$\|\hat{F}_t - H' F_t\| = O_p\left(max \left\{ max \left\{ \frac{1}{\sqrt{N}}, \frac{1}{\sqrt{T}} \right\}, \left(\frac{B/\sqrt{N}}{\sqrt{N}} \right) \right\} \right)$$

- To summarise, the PCA estimator can accomodate a certain degree of temporal instability in the factor loadings.
- Under the right conditions, effects of instabilities can be eliminated asymptotically by averaging across series.

Are the Bates et al. (JoE 2013) conditions realistic?

AYE!

"The [simulation] results [calibrated to previous empirical work by Stock and Watson] confirm...

- Bates et al. (JoE 2013)

Are the Bates et al. (JoE 2013) conditions realistic?

AYE!

"The [simulation] results [calibrated to previous empirical work by Stock and Watson] confirm...

that the principal components estimator [is] robust to empirically relevant degrees of temporal instability in the factor loadings,...

- Bates et al. (JoE 2013)

Are the Bates et al. (JoE 2013) conditions realistic?

AYE!

> *"The [simulation] results [calibrated to previous empirical work by Stock and Watson] confirm...*
>
> *that the principal components estimator [is] robust to empirically relevant degrees of temporal instability in the factor loadings,...*
>
> *although the precise quantitative conclusions depend on the assumed type of structural instability"*

- Bates et al. (JoE 2013)

Are the Bates et al. (JoE 2013) conditions realistic?

AYE!

> *"The [simulation] results [calibrated to previous empirical work by Stock and Watson] confirm...*
>
> *that the principal components estimator [is] robust to empirically relevant degrees of temporal instability in the factor loadings,...*
>
> *although the precise quantitative conclusions depend on the assumed type of structural instability"*

- Bates et al. (JoE 2013)
- This is where the disagreement arises.

Are the Bates et al. (JoE 2013) conditions realistic?

NAY!

> *"...in* *empirical applications* *parameters may change dramatically due to important economic events [...]."*

- Breitung & Eickmeier (JoE 2011)

Are the Bates et al. (JoE 2013) conditions realistic?

NAY!

"...in empirical applications parameters may change dramatically due to important economic events [...]."

- Breitung & Eickmeier (JoE 2011)

"...we find that a significant portion [around 80] of 132 U.S. macroeconomic time series have structural changes in their factor loadings. Although traditional principal components provide eight or more factors, there are significantly fewer nonspurious factors."

- Yamamoto (J o Bus. & Econ. Stats. 2016)

Are the Bates et al. (JoE 2013) conditions realistic?

NAY!

"...in empirical applications parameters may change dramatically due to important economic events [...]."

- Breitung & Eickmeier (JoE 2011)

"...we find that a significant portion [around 80] of 132 U.S. macroeconomic time series have structural changes in their factor loadings. Although traditional principal components provide eight or more factors, there are significantly fewer nonspurious factors."

- Yamamoto (J o Bus. & Econ. Stats. 2016)
- Burgeoning literature on this side of the debate!

IV. Accounting for instabilities

Specifically, let us consider Baltagi, Kao & Wang (Manuscript, 2015)

- Paper is recent and its representation is fairly general; models in many other papers correspond to special cases.
- This paper tackles estimation of an approximate static factor model with a single abrupt break.
- Allows for changes in the number of factors and/or partial changes in factor loadings.

Model description for X (in $T \times N$ matrix form)

$$X = \begin{bmatrix} F^0_{bef}\Lambda^{0'} + F^1_{bef}\Lambda^{1'} \\ F^0_{aft}\Lambda^{0'} + F^1_{aft}\Lambda^{2'} \end{bmatrix} + E$$

Notation:

- pre-change $\begin{cases} F^0_{bef} = [f^0_1, ..., f^0_\kappa]'; & \kappa \times (r - q) \\ F^1_{bef} = [f^1_1, ..., f^1_\kappa]'; & \kappa \times q \end{cases}$

Model description for X (in $T \times N$ matrix form)

$$X = \begin{bmatrix} F^0_{bef}\Lambda^{0\prime} + F^1_{bef}\Lambda^{1\prime} \\ F^0_{aft}\Lambda^{0\prime} + F^1_{aft}\Lambda^{2\prime} \end{bmatrix} + E$$

Notation:

- pre-change $\begin{cases} F^0_{bef} = [f^0_1, ..., f^0_\kappa]'; & \kappa \times (r-q) \\ F^1_{bef} = [f^1_1, ..., f^1_\kappa]'; & \kappa \times q \end{cases}$
- post-change $\begin{cases} F^0_{aft} = [f^0_{\kappa+1}, ..., f^0_T]'; & (T-\kappa) \times (r-q) \\ F^1_{aft} = [f^1_{\kappa+1}, ..., f^1_T]'; & (T-\kappa) \times q \end{cases}$

Model description for X (in $T \times N$ matrix form)

$$X = \begin{bmatrix} F^0_{bef}\Lambda^{0\prime} + F^1_{bef}\Lambda^{1\prime} \\ F^0_{aft}\Lambda^{0\prime} + F^1_{aft}\Lambda^{2\prime} \end{bmatrix} + E$$

Notation:

- pre-change $\begin{cases} F^0_{bef} = [f^0_1, \dots, f^0_\kappa]'; & \kappa \times (r-q) \\ F^1_{bef} = [f^1_1, \dots, f^1_\kappa]'; & \kappa \times q \end{cases}$
- post-change $\begin{cases} F^0_{aft} = [f^0_{\kappa+1}, \dots, f^0_T]'; & (T-\kappa) \times (r-q) \\ F^1_{aft} = [f^1_{\kappa+1}, \dots, f^1_T]'; & (T-\kappa) \times q \end{cases}$
- factor loadings $\begin{cases} \Lambda^0; & N \times (r-q) \\ \Lambda^1 \ \& \ \Lambda^2; & N \times q \end{cases}$

Model description for X (in $T \times N$ matrix form)

$$X = \begin{bmatrix} F^0_{bef}\Lambda^{0\prime} + F^1_{bef}\Lambda^{1\prime} \\ F^0_{aft}\Lambda^{0\prime} + F^1_{aft}\Lambda^{2\prime} \end{bmatrix} + E$$

Notation:

- pre-change $\begin{cases} F^0_{bef} = [f^0_1, \dots, f^0_\kappa]'; & \kappa \times (r-q) \\ F^1_{bef} = [f^1_1, \dots, f^1_\kappa]'; & \kappa \times q \end{cases}$
- post-change $\begin{cases} F^0_{aft} = [f^0_{\kappa+1}, \dots, f^0_T]'; & (T-\kappa) \times (r-q) \\ F^1_{aft} = [f^1_{\kappa+1}, \dots, f^1_T]'; & (T-\kappa) \times q \end{cases}$
- factor loadings $\begin{cases} \Lambda^0; & N \times (r-q) \\ \Lambda^1 \; \& \; \Lambda^2; & N \times q \end{cases}$
- define $\left\{ \Lambda_{bef} = \begin{bmatrix} \Lambda^0 & \Lambda^1 \end{bmatrix} \; \& \; \Lambda_{aft} = \begin{bmatrix} \Lambda^0 & \Lambda^2 \end{bmatrix}; \quad N \times r \right.$

Types of instability

$$X = \begin{bmatrix} F^0_{bef}\Lambda^{0'} + F^1_{bef}\Lambda^{1'} \\ F^0_{aft}\Lambda^{0'} + F^1_{aft}\Lambda^{2'} \end{bmatrix} + E$$

Factor loadings and/or number of factors may change:

- Structural change in the number of factors is incorporated as special case of structural change in factor loadings by allowing either Λ_{bef} or Λ_{aft} to have less than full column rank.
- $r = max\{r_{bef}, r_{aft}\}$
- Emerging factors: If $r_1 < r_2$, some columns in Λ_{bef} are zeros.
- Disappearing factors: If $r_1 > r_2$, some columns in Λ_{aft} are zeros.
- For now, assume both Λ_{bef} and Λ_{aft} have full column rank.

Equivalent model with stable loadings

$$X = G\Gamma' + E$$

Equivalent model with stable loadings

$$X = G\Gamma' + E$$

How do we define G and Γ?

$$G_{T\times(r+q)} = \begin{bmatrix} \begin{bmatrix} F^0_{bef} & F^1_{bef} \end{bmatrix} A' \\ \begin{bmatrix} F^0_{aft} & F^1_{aft} \end{bmatrix} B' \end{bmatrix}$$

where

$$A_{(r+q)\times r} = \begin{bmatrix} I_{(r-q)} & 0_{(r-q)\times q} \\ 0_{q\times(r-q)} & I_q \\ 0_{q\times(r-q)} & 0_q \end{bmatrix},$$

$$B_{(r+q)\times r} = \begin{bmatrix} I_{(r-q)} & 0_{(r-q)\times q} \\ 0_{q\times(r-q)} & 0_q \\ 0_{q\times(r-q)} & I_q \end{bmatrix},$$

and $\Gamma_{N\times(r+q)} = \begin{bmatrix} \Lambda^0 & \Lambda^1 & \Lambda^2 \end{bmatrix}$.

Suggested procedure

1. Obtain preliminary estimate, $\hat{r}$ of number of factors, r, ignoring structural change.
   - Use standard methods developed for stable models.
   - $\lim_{(N,T)\to\infty} P(\hat{r} = r + q) = 1$.

Suggested procedure

1. Obtain preliminary estimate, $\hat{r}$ of number of factors, r, ignoring structural change.
   - Use standard methods developed for stable models.
   - $\lim_{(N,T)\to\infty} P(\hat{r} = r + q) = 1$.
2. Given $\hat{r}$, estimate the space spanned by the pseudo-factors, G_t.
   - Use PCA.

Suggested procedure

1. Obtain preliminary estimate, $\hat{r}$ of number of factors, r, ignoring structural change.
   - Use standard methods developed for stable models.
   - $\lim_{(N,T)\to\infty} P(\hat{r} = r + q) = 1$.
2. Given $\hat{r}$, estimate the space spanned by the pseudo-factors, G_t.
   - Use PCA.
3. Estimate changepoint of factor loadings, κ.
   - Look for break in covariance matrix of pseudo-factors, Σ_G.
   - Note that $\Sigma_{G,bef} = A\Sigma_F A'$ and $\Sigma_{G,aft} = B\Sigma_F B'$.
   - Avoid potential infinite-dimensionality problem.

Estimating the changepoint, κ

- Split data into 2 subsamples for each "candidate" location of $\kappa \in [1, T-1]$.
- Pre-break subsample: $\hat{\Sigma}_{G,bef} = \sum_{t=1}^{\kappa} \hat{G}_t \hat{G}_t{}' / \kappa$
- Post-break subsample: $\hat{\Sigma}_{G,aft} = \sum_{t=\kappa+1}^{T} \hat{G}_t \hat{G}_t{}' / (T - \kappa)$

Estimating the changepoint, κ

- Split data into 2 subsamples for each "candidate" location of $\kappa \in [1, T-1]$.
- Pre-break subsample: $\hat{\Sigma}_{G,bef} = \sum_{t=1}^{\kappa} \hat{G}_t \hat{G}_t{}' / \kappa$
- Post-break subsample: $\hat{\Sigma}_{G,aft} = \sum_{t=\kappa+1}^{T} \hat{G}_t \hat{G}_t{}' / (T-\kappa)$
- Then, define sum of squared residuals:

$$\hat{S}(\kappa) = \sum_{t=1}^{\kappa} [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,bef})]' [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,bef})] + \sum_{t=\kappa+1}^{T} [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,aft})]' [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,aft})]$$

Estimating the changepoint, κ

- Split data into 2 subsamples for each "candidate" location of $\kappa \in [1, T-1]$.
- Pre-break subsample: $\hat{\Sigma}_{G,bef} = \sum_{t=1}^{\kappa} \hat{G}_t \hat{G}_t{}' / \kappa$
- Post-break subsample: $\hat{\Sigma}_{G,aft} = \sum_{t=\kappa+1}^{T} \hat{G}_t \hat{G}_t{}' / (T-\kappa)$
- Then, define sum of squared residuals:

$$\hat{S}(\kappa) = \sum_{t=1}^{\kappa} [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,bef})]'[vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,bef})] + \sum_{t=\kappa+1}^{T} [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,aft})]'[vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,aft})]$$

- Obtain LS estimator of changepoint: $\hat{\kappa} = \arg\min \hat{S}(\kappa)$.

Estimating the changepoint, κ

- Split data into 2 subsamples for each "candidate" location of $\kappa \in [1, T-1]$.
- Pre-break subsample: $\hat{\Sigma}_{G,bef} = \sum_{t=1}^{\kappa} \hat{G}_t \hat{G}_t{}' / \kappa$
- Post-break subsample: $\hat{\Sigma}_{G,aft} = \sum_{t=\kappa+1}^{T} \hat{G}_t \hat{G}_t{}' / (T-\kappa)$
- Then, define sum of squared residuals:

$$\hat{S}(\kappa) = \sum_{t=1}^{\kappa} [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,bef})]' [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,bef})] + \sum_{t=\kappa+1}^{T} [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,aft})]' [vec(\hat{G}_t \hat{G}_t{}' - \hat{\Sigma}_{G,aft})]$$

- Obtain LS estimator of changepoint: $\hat{\kappa} = \arg\min \hat{S}(\kappa)$.
- Estimator at most only "slightly misspecified": $\hat{\kappa} - \kappa = O_p(1)$:

Determining the number of factors

4. Identify the number of factors based on $\hat{\kappa}$.

   - Split the data into 2 subsamples based on $\hat{\kappa}$.
   - Obtain $\hat{r}_{bef}$ and $\hat{r}_{aft}$ using standard criteria.
   - $\hat{r}_{bef}$ and $\hat{r}_{aft}$ shown to be robust to "slightly misspecified" $\hat{\kappa}$:

Estimating the factor space

5. Estimate the factor space based on $\hat{\kappa}$ and $\hat{r}_{bef}, \hat{r}_{aft}$.
   - Use PCA on each of the subsamples.
   - Convergence rate is the same as in Bai & Ng (Econometrica 2002).

V. Sequential detection of structural instabilities in DFMs

Do structural instabilities matter?

NAY!

- Stock & Watson (JASA 2002)
- Bates, Plagborg-Moller, Stock & Watson (JoE 2013)

AYE!

- Breitung & Eickmeier (JoE 2011)
- Corradi & Swanson (Forthcoming, JoE)
- Chen, Dolado & Gonzalo (JoE 2014)
- Han & Inoue ('metric Theory 2014)
- Baltagi, Kao & Wang (Manuscript, 2015)
- Yamamoto & Tanaka (JoE 2015)
- Sabharwal (Journal of Optimism, 2016)

Thank you for your attention!

Contact

email: r.sabharwal2@lse.ac.uk
website: http://personal.lse.ac.uk/sabharw5/