RECENT DEVELOPMENTS IN LARGE DIMENSIONAL FACTOR ANALYSIS

Serena Ng

June 2007
SCE Meeting, Montreal

Factor Model: for $i = 1, \ldots N$, $t = 1, \ldots T$,

$$x_{it} = \lambda_i' F_t + e_{it}$$

- F_t: vector of r common factors
- λ_i: vector of r factor loadings
- $c_{it} = \lambda_i' F_t$: the common component
- e_{it}: idiosyncratic component.

Factor Model: for $i = 1, \ldots N$, $t = 1, \ldots T$,

$$x_{it} = \lambda_i' F_t + e_{it}$$

- F_t: vector of r common factors
- λ_i: vector of r factor loadings
- $c_{it} = \lambda_i' F_t$: the common component
- e_{it}: idiosyncratic component.

• Key feature • N large, T large

Outline of the talk

- Applications of factor models

Outline of the talk

- Applications of factor models
- Sketch econometric framework

Outline of the talk

- Applications of factor models
- Sketch econometric framework
- Main Statistical results

Outline of the talk

- Applications of factor models
- Sketch econometric framework
- Main Statistical results
- Caveats, theory and practice

Outline of the talk

- Applications of factor models
- Sketch econometric framework
- Main Statistical results
- Caveats, theory and practice
- What next?

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;
- George Connor, George Korajczyk

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;
- George Connor, George Korajczyk
- Gary Chamerlain, Michael Rothschild

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;
- George Connor, George Korajczyk
- Gary Chamerlain, Michael Rothschild
- Jushan Bai, Jim Stock, Mark Watson;

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;
- George Connor, George Korajczyk
- Gary Chamerlain, Michael Rothschild
- Jushan Bai, Jim Stock, Mark Watson;
- Lucrezia Reichlin, Marco Lippi, Mario Forni, Marc Hallin;

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;
- George Connor, George Korajczyk
- Gary Chamerlain, Michael Rothschild
- Jushan Bai, Jim Stock, Mark Watson;
- Lucrezia Reichlin, Marco Lippi, Mario Forni, Marc Hallin;
- Domenico Giannone, Catherine Doz

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;
- George Connor, George Korajczyk
- Gary Chamerlain, Michael Rothschild
- Jushan Bai, Jim Stock, Mark Watson;
- Lucrezia Reichlin, Marco Lippi, Mario Forni, Marc Hallin;
- Domenico Giannone, Catherine Doz
- Jean Boivn, Marc Giannoni

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;
- George Connor, George Korajczyk
- Gary Chamerlain, Michael Rothschild
- Jushan Bai, Jim Stock, Mark Watson;
- Lucrezia Reichlin, Marco Lippi, Mario Forni, Marc Hallin;
- Domenico Giannone, Catherine Doz
- Jean Boivn, Marc Giannoni
- Fashid Vahid, Heather Robertson, Roger Moon, Benoit Perron;

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;
- George Connor, George Korajczyk
- Gary Chamerlain, Michael Rothschild
- Jushan Bai, Jim Stock, Mark Watson;
- Lucrezia Reichlin, Marco Lippi, Mario Forni, Marc Hallin;
- Domenico Giannone, Catherine Doz
- Jean Boivn, Marc Giannoni
- Fashid Vahid, Heather Robertson, Roger Moon, Benoit Perron;
- Alexei Ontaski, Matt Harding.

Contributors to the Literature

- T.W. Anderson, Lawley and Maxwell, ...;
- Chris Sims, John Geweke, Tom Sargent, Danny Quah;
- George Connor, George Korajczyk
- Gary Chamerlain, Michael Rothschild
- Jushan Bai, Jim Stock, Mark Watson;
- Lucrezia Reichlin, Marco Lippi, Mario Forni, Marc Hallin;
- Domenico Giannone, Catherine Doz
- Jean Boivn, Marc Giannoni
- Fashid Vahid, Heather Robertson, Roger Moon, Benoit Perron;
- Alexei Ontaski, Matt Harding.

Example 1. Arbitrage Pricing Theory (APT):

$$\begin{aligned} R_{it} &= a_i + b_i' F_t + e_{it} \\ E(e_{it}|F_t) &= 0 \\ E(e_{it}^2) &= \sigma_i^2 \leq \sigma^2 < \infty. \end{aligned}$$

Example 1. Arbitrage Pricing Theory (APT):

$$\begin{aligned} R_{it} &= a_i + b_i' F_t + e_{it} \\ E(e_{it}|F_t) &= 0 \\ E(e_{it}^2) &= \sigma_i^2 \leq \sigma^2 < \infty. \end{aligned}$$

- F_t: common (pervasive) factors in asset returns;
- e_{it} in large, well-diversified portfolios vanishes;
- e_{it} sufficiently uncorrelated across assets
- no single asset dominates wealth in competitive equilibrium.

What are the factors?

- observed
  - portfolios
  - macroeconomic variables
    - innovations in GDP, inflation, changes in bond yields

What are the factors?

- observed
  - portfolios
  - macroeconomic variables
    - innovations in GDP, inflation, changes in bond yields
- latent
  - estimation of F_t: N large, T small.

Example 2. Interest Rate Models:

$$
\begin{aligned}
r_t &= a_0 + b_0' F_t^M + b_1' F_{t-1}^M + \ldots b_p' F_{t-p}^M + e_t \\
&= a_0 + \beta' \vec{F}_t^M + e_t .
\end{aligned}
$$

Example 2. Interest Rate Models:

$$
\begin{aligned}
r_t &= a_0 + b_0' F_t^M + b_1' F_{t-1}^M + \ldots b_p' F_{t-p}^M + e_t \\
&= a_0 + \beta' \vec{F}_t^M + e_t.
\end{aligned}
$$

- Taylor rule:
  - F_t^M: current and macro variables orthogonal to e_t

Example 2. Interest Rate Models:

$$\begin{aligned} r_t &= a_0 + b_0' F_t^M + b_1' F_{t-1}^M + \ldots b_p' F_{t-p}^M + e_t \\ &= a_0 + \beta' \vec{F}_t^M + e_t. \end{aligned}$$

- Taylor rule:
  - F_t^M: current and macro variables orthogonal to e_t
- affine term structure models:
  - bond yields are linear in the underlying state variables

Example 3. Demand Systems: J goods, H households

- $E_h =$ total spending on J goods by household h;
- Marshallian demand: $X_{jh} = g_j(p, E_h)$
- budget share: $w_{jh} = X_{jh}/E_h$

Example 3. Demand Systems: J goods, H households

- $E_h =$ total spending on J goods by household h;
- Marshallian demand: $X_{jh} = g_j(p, E_h)$
- budget share: $w_{jh} = X_{jh}/E_h$
- the rank of a demand system holding prices fixed
  - the smallest integer r such that

$$w_j(E) = \lambda_{j1} G_1(E) + \ldots \lambda_{jr} G_r(E).$$

- $F_h = (G_1(E_h), \ldots G_r(E_h))'$ are r factors across goods

Example 4. Coincident index

$$\begin{aligned} y_{1t} &= \lambda_1 F_t + z_{1t} \\ y_{2t} &= \lambda_2 F_t + z_{2t} \\ y_{3t} &= \lambda_3 F_t + z_{3t} \\ y_{4t} &= \lambda_4 F_t + z_{4t} \\ F_t &= \phi F_{t-1} + v_t \\ z_{it} &= \alpha_i z_{it-1} + e_{it}, \quad i = 1, \ldots 4. \end{aligned}$$

$N = 4$, T large.

Example 6. Forecasting

$$y_{t+1} = a'X_t + \beta' W_t + \epsilon_{t+1}.$$

Example 6. Forecasting

$$y_{t+1} = a' X_t + \beta' W_t + \epsilon_{t+1}.$$

- X_t: N observed variables.
- W_t: observed variables

Example 6. Forecasting

$$y_{t+1} = a' X_t + \beta' W_t + \epsilon_{t+1}.$$

- X_t: N observed variables.
- W_t: observed variables
- N large: inefficient
- Assume X_{it} have common sources of variation F_t.

Diffusion Index Forecasting:

$$y_{t+1} = \alpha' F_t + \beta' W_t + \epsilon_{t+1}.$$

Diffusion Index Forecasting:

$$y_{t+1} = \alpha' F_t + \beta' W_t + \epsilon_{t+1}.$$

eg: Fed reacts to state of the economy.

$$r_t = \alpha' F_t + \epsilon_t.$$

Example 6. VAR: m variables

$$y_{t+1} = \sum_{k=0}^{p} \alpha_{11}(k) y_{t-k} + e_{1t+1}.$$

Example 6. VAR: m variables

$$y_{t+1} = \sum_{k=0}^{p} \alpha_{11}(k) y_{t-k} + e_{1t+1}.$$

FAVAR: m variables + r factors

$$\begin{aligned} y_{t+1} &= \sum_{k=0}^{p} a_{11}(k) y_{t-k} + \sum_{k=0}^{p} a_{12}(k) F_{t-k} + e_{1t+1} \\ F_{t+1} &= \sum_{k=0}^{p} a_{21}(k) y_{t-k} + \sum_{k=0}^{p} a_{22}(k) F_{t-k} + e_{2t+1}. \end{aligned}$$

Main Econometric issues

- the factors F_t are not observed;

Main Econometric issues

- the factors F_t are not observed;
- inference when F_t has to be estimated

Main Econometric issues

- the factors F_t are not observed;
- inference when F_t has to be estimated
- the number of factors r is unknown

Main Econometric issues

- the factors F_t are not observed;
- inference when F_t has to be estimated
- the number of factors r is unknown
- N and T both large

Statistical Factor Models

$$x_{it} = \lambda_i' F_t + e_{it}.$$

Covariance Structure with $\Sigma_F = I_r$.

$$\Sigma_x = \Lambda\Lambda' + \Omega$$

Statistical Factor Models

$$x_{it} = \lambda_i' F_t + e_{it}.$$

Covariance Structure with $\Sigma_F = I_r$.

$$\Sigma_x = \Lambda\Lambda' + \Omega$$

- Strict factor model: Ω diagonal

Statistical Factor Models

$$x_{it} = \lambda_i' F_t + e_{it}.$$

Covariance Structure with $\Sigma_F = I_r$.

$$\Sigma_x = \Lambda\Lambda' + \Omega$$

- Strict factor model: Ω diagonal
- Classical factor model:
  - (i) Ω diagonal
  - (ii) F_t and e_t serially uncorrelated

Statistical Factor Models

$$x_{it} = \lambda_i' F_t + e_{it}.$$

Covariance Structure with $\Sigma_F = I_r$.

$$\Sigma_x = \Lambda\Lambda' + \Omega$$

- Strict factor model: Ω diagonal
- Classical factor model:
  (i) Ω diagonal
  (ii) F_t and e_t serially uncorrelated
- Anderson and Rubin: assume
  (i) e_{it} is iid over t,
  (ii) normality,
  (iii) N fixed $T \to \infty$.

Large dimensional factor models:

- approximate factor structure:
  - e_{it} can be 'weakly' cross-sectionally and serially correlated

Large dimensional factor models:

- approximate factor structure:
  - e_{it} can be 'weakly' cross-sectionally and serially correlated
- Ω need not be diagonal

Large dimensional factor models:

- approximate factor structure:
  - e_{it} can be 'weakly' cross-sectionally and serially correlated
- Ω need not be diagonal
- N and T are large

Large dimensional factor models:

- approximate factor structure:
  - e_{it} can be 'weakly' cross-sectionally and serially correlated
- Ω need not be diagonal
- N and T are large
- distribution assumptions not imposed on e_{it}

Static vs. Dynamic Factors

- dynamic factor model

$$x_{it} = \lambda_{i1} f_{1t} + \lambda_{i2} f_{1t-1} + e_{it}.$$

Static vs. Dynamic Factors

- dynamic factor model

$$x_{it} = \lambda_{i1} f_{1t} + \lambda_{i2} f_{1t-1} + e_{it}.$$

Put $F_{1t} = f_{1t}$, $F_{2t} = f_{1t-1}$

- static factor model

$$x_{it} = \lambda_{i1} F_{1t} + \lambda_{i2} F_{2t} + e_{it}.$$

Static vs. Dynamic Factors

- dynamic factor model

$$x_{it} = \lambda_{i1} f_{1t} + \lambda_{i2} f_{1t-1} + e_{it}.$$

Put $F_{1t} = f_{1t}$, $F_{2t} = f_{1t-1}$

- static factor model

$$x_{it} = \lambda_{i1} F_{1t} + \lambda_{i2} F_{2t} + e_{it}.$$

- q dynamic factors and s lags give $r = q(s+1)$ static factors.

Properties of a model with r factors:

- the r largest eigenvalues of Σ_x diverge as N increases;
- the $r+1$ eigenvalue is bounded.
- example

$$x_{it} = F_t + e_{it}, \quad e_{it} \sim iid(0,1).$$

  - $eig_1^x = N + 1$,
  - $eig_i^x = 1,\ i = 2, \ldots N$.
- the population principal components converge to the population factors as N increases.
- will need the sample principal components to converge to the population principal components.

Principal Components (PC) estimator

$$(\tilde{F}, \tilde{\Lambda}) = \min_{\Lambda, F} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i' F_t)^2.$$

Principal Components (PC) estimator

$$(\tilde{F}, \tilde{\Lambda}) = \min_{\Lambda, F} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i' F_t)^2.$$

- $\tilde{F}$: r eigenvectors (times $\sqrt{T}$) associated with the r largest eigenvalues of the matrix $XX'/(TN)$.
- $\tilde{\Lambda} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_N)' = X'\tilde{F}/T$
- $\tilde{e} = X - \tilde{F}\tilde{\Lambda}'$.

Principal Components (PC) estimator

$$(\tilde{F}, \tilde{\Lambda}) = \min_{\Lambda, F} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i' F_t)^2.$$

- $\tilde{F}$: r eigenvectors (times $\sqrt{T}$) associated with the r largest eigenvalues of the matrix $XX'/(TN)$.
- $\tilde{\Lambda} = (\tilde{\lambda}_1, \ldots, \tilde{\lambda}_N)' = X'\tilde{F}/T$
- $\tilde{e} = X - \tilde{F}\tilde{\Lambda}'$.

The space spanned by the factors can be consistently estimated by $\tilde{F}$ when N and T are both large

Intuition: let $r = 1$, $\lambda_i = 1$ and $\sigma_i^2 = \sigma^2$ for all i.

$$x_{it} = \lambda_i F_t + e_{it}$$

Intuition: let $r = 1$, $\lambda_i = 1$ and $\sigma_i^2 = \sigma^2$ for all i.

$$x_{it} = \lambda_i F_t + e_{it}$$

- given x_{it}, we cannot separately identify F_t and e_{it}

Intuition: let $r = 1$, $\lambda_i = 1$ and $\sigma_i^2 = \sigma^2$ for all i.

$$x_{it} = \lambda_i F_t + e_{it}$$

- given x_{it}, we cannot separately identify F_t and e_{it}
- with a large N: $\bar{x}_t$ is $\sqrt{N}$ consistent for F_t

Intuition: let $r = 1$, $\lambda_i = 1$ and $\sigma_i^2 = \sigma^2$ for all i.

$$x_{it} = \lambda_i F_t + e_{it}$$

- given x_{it}, we cannot separately identify F_t and e_{it}
- with a large N: $\bar{x}_t$ is $\sqrt{N}$ consistent for F_t
    - $\frac{1}{N}\sum_{i=1}^{N} x_{it} = F_t + \frac{1}{N}\sum_{i=1}^{N} e_{it}$
    - $\text{var}(\frac{1}{N}\sum_{i=1}^{N} e_{it}) \to 0$ as $N \to \infty$

Intuition: let $r = 1$, $\lambda_i = 1$ and $\sigma_i^2 = \sigma^2$ for all i.

$$x_{it} = \lambda_i F_t + e_{it}$$

- given x_{it}, we cannot separately identify F_t and e_{it}
- with a large N: $\bar{x}_t$ is $\sqrt{N}$ consistent for F_t
  - $\frac{1}{N}\sum_{i=1}^{N} x_{it} = F_t + \frac{1}{N}\sum_{i=1}^{N} e_{it}$
  - $\mathrm{var}(\frac{1}{N}\sum_{i=1}^{N} e_{it}) \to 0$ as $N \to \infty$
- with a large T:
  - regressing each x_i on $\tilde{F}_t$ gives $\sqrt{T}$ consistent estimates of λ_i.

Intuition: let $r = 1$, $\lambda_i = 1$ and $\sigma_i^2 = \sigma^2$ for all i.

$$x_{it} = \lambda_i F_t + e_{it}$$

- given x_{it}, we cannot separately identify F_t and e_{it}
- with a large N: $\bar{x}_t$ is $\sqrt{N}$ consistent for F_t
  - $\frac{1}{N}\sum_{i=1}^N x_{it} = F_t + \frac{1}{N}\sum_{i=1}^N e_{it}$
  - $\text{var}(\frac{1}{N}\sum_{i=1}^N e_{it}) \to 0$ as $N \to \infty$
- with a large T:
  - regressing each x_i on $\tilde{F}_t$ gives $\sqrt{T}$ consistent estimates of λ_i.
- precision of factor estimates depends on both N and T.

Intuition: let $r = 1$, $\lambda_i = 1$ and $\sigma_i^2 = \sigma^2$ for all i.

$$x_{it} = \lambda_i F_t + e_{it}$$

- given x_{it}, we cannot separately identify F_t and e_{it}
- with a large N: $\bar{x}_t$ is $\sqrt{N}$ consistent for F_t
  - $\frac{1}{N}\sum_{i=1}^N x_{it} = F_t + \frac{1}{N}\sum_{i=1}^N e_{it}$
  - $\text{var}(\frac{1}{N}\sum_{i=1}^N e_{it}) \to 0$ as $N \to \infty$
- with a large T:
  - regressing each x_i on $\tilde{F}_t$ gives $\sqrt{T}$ consistent estimates of λ_i.
- precision of factor estimates depends on both N and T.
- method of PC weights X_{it} appropriately to yield $\tilde{F}$ when $r > 1$, and/or there is heterogeneity in λ_i, σ_i^2.

Assumptions

- F(0) – moments
- LFE – independence
- L – identification
- E – weak correlation
- IE – homoskedsaticity

Assumption F(0)
$E\|F_t^0\|^4 \leq M$ and $\frac{1}{T}\sum_{t=1}^{T} F_t F_t' \xrightarrow{p} \Sigma_F > 0$, is a $r \times r$ non-random matrix.

Assumption LFE
$\{\lambda_i\}$, $\{F_t\}$, and $\{e_{it}\}$ are three mutually independent groups. Dependence within each group is allowed.

Assumption L λ_i^0 is either deterministic such that $\|\lambda_i^0\| \leq M$, or it is stochastic such that $E\|\lambda_i^0\|^4 \leq M$. In either case, $N^{-1}\Lambda^{0\prime}\Lambda^0 \xrightarrow{p} \Sigma_\Lambda > 0$, a $r \times r$ non-random matrix, as $N \to \infty$.

Assumption E:

b.i $E(e_{it}) = 0$, $E|e_{it}|^8 \leq M$.

b.ii $E(e_{it}e_{js}) = \sigma_{ij,ts}$

- $\frac{1}{NT}\sum_{i,j,t,s=1} |\sigma_{ij,ts}| \leq M$
- $|\sigma_{ij,ts}| \leq \bar{\sigma}_{ij}$ for all (t,s) and $\frac{1}{N}\sum_{i,j=1}^{N} \bar{\sigma}_{ij} \leq M$;
- $|\sigma_{ij,ts}| \leq \tau_{ts}$ for all (i,j) and $\frac{1}{T}\sum_{t,s=1}^{T} \tau_{ts} \leq M$.

b.iii For every (t,s), $E|N^{-1/2}\sum_{i=1}^{N}\left[e_{is}e_{it}-E(e_{is}e_{it})\right]|^4 \leq M$.

b.iv For each t, $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\lambda_i e_{it}\xrightarrow{d}N(0,\Gamma_t)$, as $N\to\infty$ where

$$\Gamma_t=\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}E(\lambda_i\lambda_j' e_{it}e_{jt}).$$

b.v For each i, $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}F_t e_{it}\xrightarrow{d}N(0,\Phi_i)$ as $T\to\infty$ where

$$\Phi_i=\lim_{T\to\infty}T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}E(F_t^0 F_s^{0\prime}e_{is}e_{it}).$$

Assumption IE for all T and N and for all $t\leq T$, $i\leq N$, $\sum_{s=1}^{T}|\tau_{st}|\leq M$, and $\sum_{i=1}^{N}|\sigma_{ij}|\leq M$.

Result A0.1:

Let $C_{NT}^2 = \min[N, T]$, H is a $r \times r$ matrix of rank r

a Under F0 + L+E:

$$C_{NT}^2 \left(\frac{1}{T} \sum_{t=1}^{T} \left\| \tilde{F}_t - H' F_t^0 \right\| \right) = O_p(1).$$

Result A0.1:

Let $C_{NT}^2 = \min[N, T]$, H is a $r \times r$ matrix of rank r

a Under F0 + L+E:

$$C_{NT}^2 \left(\frac{1}{T} \sum_{t=1}^{T} \left\| \tilde{F}_t - H' F_t^0 \right\| \right) = O_p(1).$$

- under F0+L+ E+LFE,

$$\max_{1 \leq t \leq T} \left\| \tilde{F}_t - H' F_t^0 \right\| = O_p(T^{-1/2}) + O_P((T/N)^{1/2}).$$

Result A0.1:

Let $C_{NT}^2 = \min[N, T]$, H is a $r \times r$ matrix of rank r

a Under F0 + L+E:

$$C_{NT}^2 \left(\frac{1}{T} \sum_{t=1}^{T} \left\| \tilde{F}_t - H' F_t^0 \right\| \right) = O_p(1).$$

- under F0+L+ E+LFE,

$$\max_{1 \le t \le T} \left\| \tilde{F}_t - H' F_t^0 \right\| = O_p(T^{-1/2}) + O_P((T/N)^{1/2}).$$

- if in addition $\sum_{s=1}^{T} \tau_{s,t} \le M$ for all t and T, then for each t,

$$C_{NT}^2 \left\| \tilde{F}_t - H^{k\prime} F_t^0 \right\|^2 = O_p(1).$$

Result A0.2: $\tilde{F}_t$ and $\tilde{\lambda}_i$:

a if $\sqrt{N}/T \to 0$, then for each t,

$$\sqrt{N}(\tilde{F}_t - H'F_t^0) \xrightarrow{d} N(0, Avar(\tilde{F}_t)).$$

Result A0.2: $\tilde{F}_t$ and $\tilde{\lambda}_i$:

a if $\sqrt{N}/T \to 0$, then for each t,

$$\sqrt{N}(\tilde{F}_t - H'F_t^0) \xrightarrow{d} N(0, Avar(\tilde{F}_t)).$$

If $\liminf \sqrt{N}/T > c \geq 0$, then $T(\tilde{F}_t - H'F_t^0) = O_p(1)$.

Result A0.2: $\tilde{F}_t$ and $\tilde{\lambda}_i$:

a if $\sqrt{N}/T \to 0$, then for each t,

$$\sqrt{N}(\tilde{F}_t - H'F_t^0) \xrightarrow{d} N(0, Avar(\tilde{F}_t)).$$

If $\liminf \sqrt{N}/T > c \geq 0$, then $T(\tilde{F}_t - H'F_t^0) = O_p(1)$.

b if $\sqrt{T}/N \to 0$, then for each i,

$$\sqrt{T}(\tilde{\lambda}_i - H^{-1}\lambda_i^0) \xrightarrow{d} N(0, (Avar(\tilde{\lambda}_i).$$

Result A0.2: $\tilde{F}_t$ and $\tilde{\lambda}_i$:

a if $\sqrt{N}/T \to 0$, then for each t,

$$\sqrt{N}(\tilde{F}_t - H'F_t^0)\xrightarrow{d}N(0, Avar(\tilde{F}_t)).$$

If $\liminf \sqrt{N}/T > c \geq 0$, then $T(\tilde{F}_t - H'F_t^0) = O_p(1)$.

b if $\sqrt{T}/N \to 0$, then for each i,

$$\sqrt{T}(\tilde{\lambda}_i - H^{-1}\lambda_i^0)\xrightarrow{d}N(0, (Avar(\tilde{\lambda}_i).$$

If $\liminf \sqrt{T}/N > c > 0$, then $N(\tilde{\lambda}_i H^{-1} - \lambda_i^0) = O_p(1)$.

Result A0.3: Common Component

Let $A_{it} = \lambda_i^{0\prime}\Sigma_\Lambda^{-1}\Gamma_t\Sigma_\Lambda^{-1}\lambda_i^0$, $B_{it} = F_t^{0\prime}\Sigma_F^{-1}\Phi_i F_t^0$.

a Under Assumption F(0), E, LFE, and IE,

$$(N^{-1}A_{it} + T^{-1}B_{it})^{-1/2}(\tilde{C}_{it} - C_{it}^0)\xrightarrow{d}N(0,1)$$

without restrictions on T/N or N/T.

b if $N/T \to 0$, then $\sqrt{N}(\tilde{C}_{it} - C_{it})\xrightarrow{d}N(0, A_{it})$;

c if $T/N \to 0$, then $\sqrt{T}(\tilde{C}_{it} - C_{it})\xrightarrow{d}N(0, B_{it})$

Result B: An Estimate of $Avar(\tilde{F}_t)$ is

$$\widehat{Avar}(\tilde{F}_t) = \tilde{V}^{-1}\tilde{\Gamma}_t\tilde{V}^{-1}.$$

- $\tilde{V}$ is the diagonal matrix of eigenvalues of $(NT)^{-1}XX'$.
- To estimate the $r \times r$ matrix Γ, let $\tilde{e}_{it} = x_{it} - \tilde{\lambda}_i'\tilde{F}_t$:

Result B: An Estimate of $Avar(\tilde{F}_t)$ is

$$\widehat{Avar}(\tilde{F}_t) = \tilde{V}^{-1}\tilde{\Gamma}_t\tilde{V}^{-1}.$$

- $\tilde{V}$ is the diagonal matrix of eigenvalues of $(NT)^{-1}XX'$.
- To estimate the $r \times r$ matrix Γ, let $\tilde{e}_{it} = x_{it} - \tilde{\lambda}_i'\tilde{F}_t$:

B1 heterogeneous panel: let

$$\tilde{\Gamma}_t = \frac{1}{N}\sum_{i=1}^{N}\tilde{e}_{it}^2\tilde{\lambda}_i\tilde{\lambda}_i'.$$

Result B: An Estimate of $Avar(\tilde{F}_t)$ is

$$\widehat{Avar}(\tilde{F}_t) = \tilde{V}^{-1}\tilde{\Gamma}_t\tilde{V}^{-1}.$$

- $\tilde{V}$ is the diagonal matrix of eigenvalues of $(NT)^{-1}XX'$.
- To estimate the $r \times r$ matrix Γ, let $\tilde{e}_{it} = x_{it} - \tilde{\lambda}_i'\tilde{F}_t$:

B1 heterogeneous panel: let

$$\tilde{\Gamma}_t = \frac{1}{N}\sum_{i=1}^{N}\tilde{e}_{it}^2\tilde{\lambda}_i\tilde{\lambda}_i'.$$

B2 homogeneous panel: let $\tilde{\Gamma}_t = \tilde{\sigma}_e^2\frac{1}{N}\sum_{i=1}^{N}\tilde{\lambda}_i\tilde{\lambda}_i'$.

Result B: An Estimate of $Avar(\tilde{F}_t)$ is

$$\widehat{Avar}(\tilde{F}_t) = \tilde{V}^{-1}\tilde{\Gamma}_t\tilde{V}^{-1}.$$

- $\tilde{V}$ is the diagonal matrix of eigenvalues of $(NT)^{-1}XX'$.
- To estimate the $r \times r$ matrix Γ, let $\tilde{e}_{it} = x_{it} - \tilde{\lambda}_i'\tilde{F}_t$:

B1 heterogeneous panel: let

$$\tilde{\Gamma}_t = \frac{1}{N}\sum_{i=1}^{N} \tilde{e}_{it}^2\tilde{\lambda}_i\tilde{\lambda}_i'.$$

B2 homogeneous panel: let $\tilde{\Gamma}_t = \tilde{\sigma}_e^2 \frac{1}{N}\sum_{i=1}^{N} \tilde{\lambda}_i\tilde{\lambda}_i'$.

B3 cross-sectionally correlated panel: let

$$\tilde{\Gamma} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \tilde{\lambda}_i\tilde{\lambda}_j'\frac{1}{T}\sum_{t=1}^{T} \tilde{e}_{it}\tilde{e}_{jt}.$$

Theorem

Suppose Assumptions F(0), E and LFE hold,

- cross-sectionally uncorrelated panel: $\tilde{\Gamma}_t \xrightarrow{p} \Gamma_t$.
- cross-sectionally correlated panel: if $E(e_{it}e_{jt}) = \sigma_{ij}$ for all t so that $\Gamma_t = \Gamma$ not depending on t. If $\frac{n}{\min[N,T]} \to 0$.

$$\|\tilde{\Gamma} - H^{-1\prime}\Gamma H^{-1}\| \xrightarrow{p} 0.$$

Result C: Estimation of r :

Let

$$V(x,k,\hat{F}^k) = \min_{\Lambda}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \hat{\lambda}_i^{k\prime}\hat{F}_t^k)^2.$$

Result C: Estimation of r :

Let

$$V(x, k, \hat{F}^k) = \min_{\Lambda} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \hat{\lambda}_i^{k\prime} \hat{F}_t^k)^2.$$

Let $g(N, T)$ be a penalty function. Define

$$PCP(k) = V(x, k, \hat{F}^k) + k \hat{\sigma}^2_{kmax} g(N, T).$$

Result C: Estimation of r :

Let

$$V(x, k, \hat{F}^k) = \min_{\Lambda} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \hat{\lambda}_i^{k\prime} \hat{F}_t^k)^2.$$

Let $g(N, T)$ be a penalty function. Define

$$PCP(k) \;=\; V(x, k, \hat{F}^k) + k\hat{\sigma}^2_{kmax} g(N, T).$$

Let

$$\hat{k} = \text{argmin}_{0 \le k \le kmax} PCP(k).$$

Result C: Estimation of r :

Let

$$V(x, k, \hat{F}^k) = \min_{\Lambda} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \hat{\lambda}_i^{k\prime} \hat{F}_t^k)^2.$$

Let $g(N, T)$ be a penalty function. Define

$$PCP(k) \;=\; V(x, k, \hat{F}^k) + k \hat{\sigma}^2_{kmax} g(N, T).$$

Let

$$\hat{k} = \text{argmin}_{0 \le k \le kmax} PCP(k).$$

Under Assumptions F(0), L, E, and LFE, $\lim_{N,T \to \infty} \text{prob}(\hat{k} = r) = 1$ if

i $g(N, T) \to \infty$ and

ii $C_{NT}^2 g(N, T) \to 0$ as $N, T \to \infty$.

Result D: Estimation of q:

$$x_{it} = \lambda_i' F_t + \rho_i(L) x_{it-1} + e_{it}$$

Result D: Estimation of q:

$$x_{it} = \lambda_i' F_t + \rho_i(L) x_{it-1} + e_{it}$$

Suppose $F_t = A(L)^+ F_{t-1} + u_t$ and $u_t = R\epsilon_t$, R is $r \times q$. Then

$$x_{it} = \lambda_i' A^+(L) F_{t-1} + \rho_i(L) x_{it-1} + \lambda_i' R\epsilon_t + e_{it}.$$

Restricted Equation

Let $\hat{w}_{it}$ be the residuals from the restricted regression
Let

$$\hat{q} = \text{argmin}_k PCP(k),$$

where

$$PCP(k) = V(\hat{w}, k, \hat{F}^k) + k\hat{\sigma}^2_{kmax} g(N, T).$$

Then

$$\text{prob}(\hat{q} = q) \xrightarrow{p} 1.$$

Result E: Inference Issues with $\tilde{F}_t$

$$\begin{aligned} y_{t+h} &= \alpha' \tilde{F}_t + \beta' W_t + \epsilon_{t+h} \\ &= \tilde{z}'_{t+h} \delta + \epsilon_{t+h} \end{aligned}$$

Result E: Inference Issues with $\tilde{F}_t$

$$\begin{aligned} y_{t+h} &= \alpha'\tilde{F}_t + \beta' W_t + \epsilon_{t+h} \\ &= \tilde{z}'_{t+h}\delta + \epsilon_{t+h} \end{aligned}$$

If $\sqrt{T}/N \to 0$, then $\sqrt{T}(\hat{\delta} - \delta) \xrightarrow{d} N\Big(0, Avar(\delta)\Big)$.

Result E: Inference Issues with $\tilde{F}_t$

$$\begin{aligned} y_{t+h} &= \alpha' \tilde{F}_t + \beta' W_t + \epsilon_{t+h} \\ &= \tilde{z}'_{t+h}\delta + \epsilon_{t+h} \end{aligned}$$

If $\sqrt{T}/N \to 0$, then $\sqrt{T}(\hat{\delta} - \delta) \xrightarrow{d} N\Big(0, Avar(\delta)\Big)$.
A consistent estimator for $Avar(\hat{\delta})$ is

$$\widehat{Avar(\hat{\delta})} = \Big(\frac{1}{T}\sum_{t=1}^{T-h} \hat{z}_t \hat{z}_t'\Big)^{-1} \Big(\frac{1}{T}\sum_{t=1}^{T-h} \hat{\varepsilon}^2_{t+h} \hat{z}_t \hat{z}_t'\Big) \Big(\frac{1}{T}\sum_{t=1}^{T-h} \hat{z}_t \hat{z}_t'\Big)^{-1}.$$

Result E.2

Let δ_j be the parameters of the j-th equation of a FAVAR(p).
If $\sqrt{T}/N \to 0$,

$$\sqrt{T}(\hat{\delta}_j - \delta_j) \xrightarrow{d} N\Bigg(0, \text{plim } (\frac{1}{T}\sum_{t=1}^{T} \hat{z}_t\hat{z}_t')^{-1}\Bigg(\frac{1}{T}\sum_{t=1}^{T}(\hat{\epsilon}_{jt})^2\hat{z}_t\hat{z}_t'\Bigg)(\frac{1}{T}\sum_{t=1}^{T} \hat{z}_t\hat{z}_t'$$

Result F: IV estimation

Regression: $y_t = x_t'\beta + \epsilon_t$, $E(\epsilon_t x_t) \neq 0$.

Let z_{it} be a large panel of valid instruments and

$$\begin{aligned} x_t &= \psi' F_t + u_t \\ z_{it} &= \lambda_i' F_t + e_{it}. \end{aligned}$$

Result F: IV estimation

Regression: $y_t = x_t'\beta + \epsilon_t$, $E(\epsilon_t x_t) \neq 0$.

Let z_{it} be a large panel of valid instruments and

$$\begin{aligned} x_t &= \psi' F_t + u_t \\ z_{it} &= \lambda_i' F_t + e_{it}. \end{aligned}$$

F1: Let $g_t = \tilde{F}_t \varepsilon_t$. Then $\hat{\beta}_{FIV} = \beta^0 + o_p(1)$;

F2: If, in addition, $\frac{\sqrt{T}}{N} \to 0$ as $N, T \to \infty$,

$$\sqrt{T}(\hat{\beta}_{FIV} - \beta^0) \xrightarrow{d} N\left(0, Avar(\hat{\beta}_{FIV})\right)$$

where $Avar(\hat{\beta}_{FIV}) = \text{plim } (S_{\tilde{F}x}(\hat{S})^{-1} S_{\tilde{F}x}')^{-1}$.

Result F: IV estimation

Regression: $y_t = x_t'\beta + \epsilon_t$, $E(\epsilon_t x_t) \neq 0$.

Let z_{it} be a large panel of valid instruments and

$$\begin{aligned} x_t &= \psi' F_t + u_t \\ z_{it} &= \lambda_i' F_t + e_{it}. \end{aligned}$$

F1: Let $g_t = \tilde{F}_t \varepsilon_t$. Then $\hat{\beta}_{FIV} = \beta^0 + o_p(1)$;

F2: If, in addition, $\frac{\sqrt{T}}{N} \to 0$ as $N, T \to \infty$,

$$\sqrt{T}(\hat{\beta}_{FIV} - \beta^0) \xrightarrow{d} N\Big(0, Avar(\hat{\beta}_{FIV})\Big)$$

where $Avar(\hat{\beta}_{FIV}) = \text{plim } (S_{\tilde{F}x}(\hat{S})^{-1} S_{\tilde{F}x}')^{-1}$.

F3: Let $\hat{\beta}_{IV}$ be the estimator using z_2 observed instruments. Then

$$Avar(\hat{\beta}_{IV}) - Avar(\hat{\beta}_{FIV}) \geq 0.$$

Sketch of argument: Why need $\sqrt{T}/N \to 0$?

Sketch of argument: Why need $\sqrt{T}/N \to 0$?

$$\begin{aligned}\sqrt{T}\bar{g} &= T^{-1/2}\sum_{t=1}^{T}\tilde{F}_t\varepsilon_t \\ &= \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(\tilde{F}_t - HF_t^0)\varepsilon_t + HT^{-1/2}\sum_{t=1}^{T}F_t^0\varepsilon_t.\end{aligned}$$

- $\frac{1}{T}\epsilon'(\tilde{F} - HF) = O_p(\frac{1}{\min[N,T]})$

Sketch of argument: Why need $\sqrt{T}/N \to 0$?

$$
\begin{aligned}
\sqrt{T}\bar{g} &= T^{-1/2}\sum_{t=1}^{T}\tilde{F}_t\varepsilon_t \\
&= \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(\tilde{F}_t - HF_t^0)\varepsilon_t + HT^{-1/2}\sum_{t=1}^{T}F_t^0\varepsilon_t.
\end{aligned}
$$

- $\frac{1}{T}\epsilon'(\tilde{F} - HF) = O_p(\frac{1}{\min[N,T]})$
- $\frac{1}{\sqrt{T}}\epsilon'(\tilde{F} - HF) = O_p(\frac{\sqrt{T}}{\min[N,T]})$

Sketch of argument: Why need $\sqrt{T}/N \to 0$?

$$
\begin{aligned}
\sqrt{T}\bar{g} &= T^{-1/2}\sum_{t=1}^{T}\tilde{F}_t\varepsilon_t \\
&= \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(\tilde{F}_t - HF_t^0)\varepsilon_t + HT^{-1/2}\sum_{t=1}^{T}F_t^0\varepsilon_t.
\end{aligned}
$$

- $\frac{1}{T}\epsilon'(\tilde{F} - HF) = O_p(\frac{1}{\min[N,T]})$
- $\frac{1}{\sqrt{T}}\epsilon'(\tilde{F} - HF) = O_p(\frac{\sqrt{T}}{\min[N,T]})$
- if $\frac{\sqrt{T}}{N} \to 0$ as $N, T \to \infty$, $\tilde{F}$ can treated as though they were F in estimation

Sketch of argument: Why need $\sqrt{T}/N \to 0$?

$$
\begin{aligned}
\sqrt{T}\bar{g} &= T^{-1/2}\sum_{t=1}^{T}\tilde{F}_t\varepsilon_t \\
&= \sqrt{T}\frac{1}{T}\sum_{t=1}^{T}(\tilde{F}_t - HF_t^0)\varepsilon_t + HT^{-1/2}\sum_{t=1}^{T}F_t^0\varepsilon_t .
\end{aligned}
$$

- $\frac{1}{T}\epsilon'(\tilde{F} - HF) = O_p(\frac{1}{\min[N,T]})$
- $\frac{1}{\sqrt{T}}\epsilon'(\tilde{F} - HF) = O_p(\frac{\sqrt{T}}{\min[N,T]})$
- if $\frac{\sqrt{T}}{N} \to 0$ as $N, T \to \infty$, $\tilde{F}$ can treated as though they were F in estimation

Result G:
$G_t = (G_{1t}, \ldots G_{mt})$: a $m \times 1$ vector of observed proxies for the unobserved factors, F.

Result G:
$G_t = (G_{1t}, \ldots G_{mt})$: a $m \times 1$ vector of observed proxies for the unobserved factors, F.
Suppose

$$G_{jt} = \delta_j' F_t + \epsilon_{jt}.$$

Result G:

$G_t = (G_{1t}, \ldots G_{mt})$: a $m \times 1$ vector of observed proxies for the unobserved factors, F.

Suppose

$$G_{jt} = \delta_j' F_t + \epsilon_{jt}.$$

Let

$$\hat{\epsilon}_{jt} = G_{jt} - \hat{G}_{jt}.$$

Result G:
$G_t = (G_{1t}, \ldots G_{mt})$: a $m \times 1$ vector of observed proxies for the unobserved factors, F.
Suppose

$$G_{jt} = \delta_j' F_t + \epsilon_{jt}.$$

Let

$$\hat{\epsilon}_{jt} = G_{jt} - \hat{G}_{jt}.$$

As $N, T \to \infty$

$$\frac{\hat{\epsilon}_{jt} - \epsilon_{jt}}{s_{jt}} \xrightarrow{d} N(0,1)$$

Result G:
$G_t = (G_{1t}, \dots G_{mt})$: a $m \times 1$ vector of observed proxies for the unobserved factors, F.
Suppose
$$G_{jt} = \delta_j' F_t + \epsilon_{jt}.$$
Let
$$\hat{\epsilon}_{jt} = G_{jt} - \hat{G}_{jt}.$$
As $N, T \to \infty$
$$\frac{\hat{\epsilon}_{jt} - \epsilon_{jt}}{s_{jt}} \xrightarrow{d} N(0,1)$$
$s_{jt}^2 = T^{-1} \tilde{F}_t' (T^{-1} \sum_{s=1}^{T} \tilde{F}_s \tilde{F}_s' \hat{\epsilon}_{js}^2)^{-1} \tilde{F}_t + N^{-1} Avar(\hat{G}_{jt}),$

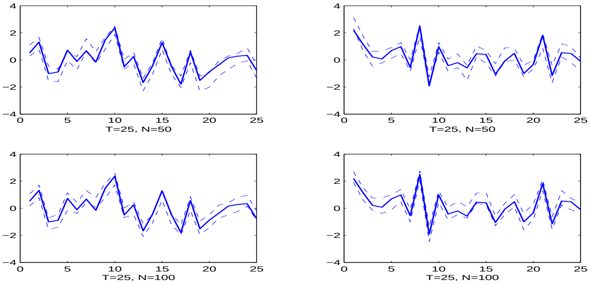
Let

$$NS(j) = \frac{\hat{\text{var}}(\hat{\epsilon}(j))}{\hat{\text{var}}(\hat{G}(j))}$$

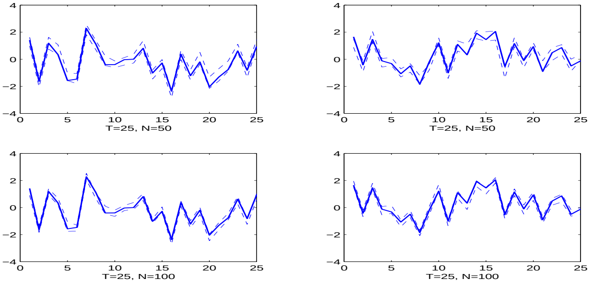
$$R^2(j) = \frac{\hat{\text{var}}(\hat{G}(j))}{\hat{\text{var}}(G(j))}.$$

Then $NS(j)$ should be close to zero and $R^2(j)$ should be close to one under the null hypothesis.

Figure 1: Measurement Errors and Their Confidence intervals



Dotted Lines: $\hat{\epsilon}_t \pm 1.96\hat{s}_{jt}$.



Solid Line: F_t

Dotted Lines: $\hat{G}_t \pm 1.96\hat{\text{var}}(\hat{G}_{jt})$.

Other applications:

- consistent estimation of the factors without knowing if the idiosyncratic errors are I(0) or I(1) (spurious regressions)

Other applications:

- consistent estimation of the factors without knowing if the idiosyncratic errors are I(0) or I(1) (spurious regressions)
- individual unit root tests

Other applications:

- consistent estimation of the factors without knowing if the idiosyncratic errors are I(0) or I(1) (spurious regressions)
- individual unit root tests
- panel unit root tests with cross-section dependence

Other applications:

- consistent estimation of the factors without knowing if the idiosyncratic errors are I(0) or I(1) (spurious regressions)
- individual unit root tests
- panel unit root tests with cross-section dependence
- panel cointegration analysis with cross-section dependence

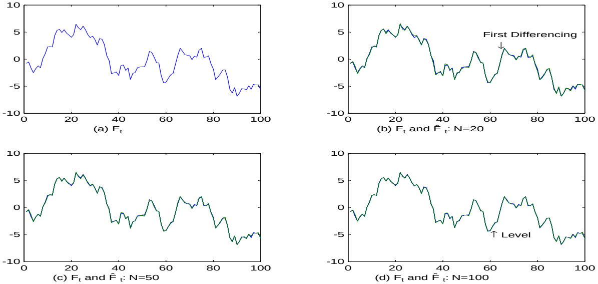
Key to all the results:

- the factor space can be consistently estimated by the method of principal components when N and T are both large.

Key to all the results:

- the factor space can be consistently estimated by the method of principal components when N and T are both large.
- 'ideal case': iid data, $\min[T, N] = 30$ yields precise estimates

Figure 2: True and Estimated F_t when e_{it} is I(0)



Solid Line: F_t

Practical issues

- is the principal components estimator efficient?
- are more data always better?
- weak factor structure?

When might the PC estimator be inefficient?

When might the PC estimator be inefficient?

- An unweighted objective function

$$V(k) = \min_{\Lambda,F} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i' F_t)^2.$$

When might the PC estimator be inefficient?

- An unweighted objective function

$$V(k) = \min_{\Lambda,F}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it} - \lambda_i' F_t)^2.$$

- ML estimation: $\hat{F}_t$ are the eigenvectors of $\tilde{\Omega}^{-1/2}\hat{\Sigma}_x\tilde{\Omega}^{-1/2}$.

When might the PC estimator be inefficient?

- An unweighted objective function

$$V(k) = \min_{\Lambda,F} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i' F_t)^2.$$

- ML estimation: $\hat{F}_t$ are the eigenvectors of $\tilde{\Omega}^{-1/2}\hat{\Sigma}_x\tilde{\Omega}^{-1/2}$.
- PC estimation of F_t: eigenvectors of $\hat{\Sigma}_x$.

When might the PC estimator be inefficient?

- An unweighted objective function

$$V(k) = \min_{\Lambda, F} (NT)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \lambda_i' F_t)^2.$$

- ML estimation: $\hat{F}_t$ are the eigenvectors of $\tilde{\Omega}^{-1/2} \hat{\Sigma}_x \tilde{\Omega}^{-1/2}$.
- PC estimation of F_t: eigenvectors of $\hat{\Sigma}_x$.
- When $\Omega \neq \omega I_n$, the PC will be less precise.

Implication: cross-section correlation and heteroskedasticity will affect the precision of the factor estimates.

How much data do we need?

How much data do we need?

- If the additional data are informative about the factor structure, more data always yield more efficient estimates.

How much data do we need?

- If the additional data are informative about the factor structure, more data always yield more efficient estimates.
- What if some of the data are 'noisy', or have a weak factor structure?

How much data do we need?

- If the additional data are informative about the factor structure, more data always yield more efficient estimates.
- What if some of the data are 'noisy', or have a weak factor structure?
- example (duplicated data) : $N = 2N_1$. Then $\text{var}(\tilde{F}_t) = O_p(N_1^{-1})$.

- the j-th eigenvalue of Σ_x measures the cumulative effect of the j factor on the cross-section units.

- the j-th eigenvalue of Σ_x measures the cumulative effect of the j factor on the cross-section units.
- strong factor asymptotics assumes that as N increases:
  - $eig_r^x / eig_{r+1}^x \rightarrow \infty$
  - eig_1^e is bounded

- the j-th eigenvalue of Σ_x measures the cumulative effect of the j factor on the cross-section units.
- strong factor asymptotics assumes that as N increases:
    - $eig^x_r / eig^x_{r+1} \to \infty$
    - eig^e_1 is bounded
- Implication:
    - eig^e_1 / eig^x_r (noise to signal ratio) should tend to zero

Why properties of eigenvalues are important?

- if eig_1^e is bounded, the population principal components converge to the population factors as N increases

Why properties of eigenvalues are important?

- if eig_1^e is bounded, the population principal components converge to the population factors as N increases
- the sample principal components converge to the population principal components as T increases (irrespective of N)

Why properties of eigenvalues are important?

- if eig_1^e is bounded, the population principal components converge to the population factors as N increases
- the sample principal components converge to the population principal components as T increases (irrespective of N)

We assume $\frac{1}{N}\sum_i \sum_j |E(e_{it}e_{jt})| < M$.

- fact: $eig_1^e \leq \max_i \sum_j |E(e_{it}e_{jt})|$
- implication: eig_1^e can be bounded and yet $\max_i \sum_j |E(e_{it}ejt|$ can increase with N.
- we allow more cross-section correlation than if eig_1^e is bounded.

Weak instrument asymptotics :

Weak instrument asymptotics :
the least influential factor is comparable to the strongest idiosyncratic noise.

Weak instrument asymptotics :
the least influential factor is comparable to the strongest idiosyncratic noise.

- when the r-th eigenvalue is too small,

$$\tilde{F} = FQ + F^{\perp}$$

where Q is a random matrix with diagonal elements strictly smaller than unity, and $F^{\perp}$ is also random and orthogonal to F.

Weak instrument asymptotics :
the least influential factor is comparable to the strongest idiosyncratic noise.

- when the r-th eigenvalue is too small,

$$\tilde{F} = FQ + F^{\perp}$$

  where Q is a random matrix with diagonal elements strictly smaller than unity, and $F^{\perp}$ is also random and orthogonal to F.

Two indicators of precision of the factor estimates.

- eig^x_{r+1}/eig^x_r
- eig^e_1/eig^x_r

Simulations

For $i = 1, \ldots N$ and $t = 1, \ldots T$,

$$x_{it} = \lambda_i'(L) f_t + \sigma_i e_{it}$$

$\lambda_i(L) = \lambda_{i0} + \lambda_{i1} L + \ldots \lambda_i L^s$.

- σ_i^2 is set so that $R_i^2 \sim U[R_L^2, R_U^2]$,
- $R_U^2 = .8$.
- $\lambda_{ij} \sim N(0,1)$

$r = q(s+1)$ static factors;
$q = 1$:

$$\begin{aligned}(1-\rho_f L)f_t &= u_t, \quad u_t \sim N(0,1) \\ (1-\rho_e L)e_{it} &= \epsilon_{it}, \quad E(\epsilon_t \epsilon_t') = \Omega.\end{aligned}$$

Error variance matrix

- $\Omega = I_N$ (errors are cross-sectionally uncorrelated)
- cross-section correlation: $N_c \times N^2$ elements of Ω are non-zero.

Parameters of the simulations are

- (N, T)=(20,50), (50,100), (100,50), (100,100), (50,200), (100,200);
- $s = 0, 1$;
- ρ_f= 0, .4, .8 ;
- ρ_e=0, U(0, .5), or U(.4, .8)
- R^2_L=.1, .35, .6;
- N_c= 0, .15, .3;

- For a given s and sample size: 81 configurations

- For a given s and sample size: 81 configurations
- total of 486 configurations

- For a given s and sample size: 81 configurations
- total of 486 configurations
- 1000 replications each

- For a given s and sample size: 81 configurations
- total of 486 configurations
- 1000 replications each
- keep track of eigenvalues

- For a given s and sample size: 81 configurations
- total of 486 configurations
- 1000 replications each
- keep track of eigenvalues
  - eig_r^x: average of the r-th largest eigenvalue of the matrx $\Sigma_{xx} = x'x/(NT)$ over 1000 replications
  - eig_1^e: the largest eigenvalue of Ω.
  - $EIG_{A,B}(a, b)$: the ratio of the a-th largest eigenvalue of the covariance matrix of A to the b-th largest eigenvalue of the covariance matrix of B.

Let FIT$=R^2$ from a regression of $\tilde{F}_t$ on F_t and a constant.

Response surface analysis:
Regress FIT on

- C_{NT}^2, $C_{NT} = \min[\sqrt{N}, \sqrt{T}]$
- ratio of eigenvalues
- non-linear terms

Dependent variable: FIT

Regressor	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\beta}$	$t_{\hat{\beta}}$
	$r = 1$			
constant	0.974	21.244	1.000	66.855
C_{NT}^{-1}	0.158	0.238	0.219	1.066
C_{NT}^{-2}	-4.086	-1.819	-3.030	-4.250
$EIG_{x,x}(r+1,r)$			-0.116	-1.700
$EIG_{e,x}(1,1)$			0.025	7.906
$EIG_{x,x}(r+1,r)^2$			-0.952	-6.564
$EIG_{e,x}(1,1)^2$			-0.003	-10.694
$\bar{R}^2$	.246		.927	

Dependent variable: FIT

Regressor	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\beta}$	$t_{\hat{\beta}}$
	$r=2$			
constant	0.958	15.048	0.988	37.176
C_{NT}^{-1}	-0.257	-0.299	0.022	0.061
C_{NT}^{-2}	-3.196	-1.184	-1.499	-1.307
$EIG_{x,x}(r+1,r)$			0.286	7.681
$EIG_{e,x}(1,1)$			-0.019	-5.231
$EIG_{x,x}(r+1,r)^2$			-1.007	-19.892
$EIG_{e,x}(1,1)^2$			-0.000	-0.214
$\bar{R}^2$	.121		0.8454	

Future work

1. More efficient estimators in a large N and T environment
   - GLS type principal components estimator
   - QMLE
   - Dynamic bayesian analysis
2. (i, j, t) model

$$
\begin{aligned}
x_{ijt} &= \lambda_{ij} F_t + e_{ijt} \\
\lambda_{ij} &= \psi_i G_j + \epsilon_{ij}
\end{aligned}
$$

   - individual i in region j at time t
   - individual, regional, aggregate effects.

3. Identification of factors

- structural/confirmatory factor analysis
- time varying loadings: $\lambda_{it} = \lambda_{0i} + \lambda_{1i} t$

$$
\begin{aligned}
x_{it} &= \lambda_{it} F_t + e_{it} \\
&= \lambda_{0i} F_t + \lambda_{1i} F_t \cdot t + e_{it} \\
&= \lambda_{0i} F_{1t} + \lambda_{1i} F_{2t} + e_{it}.
\end{aligned}
$$

4. DSGE Models

- small number of common shocks
  - stochastic singularity
  - measurement error $\Rightarrow$ factor structure

4. DSGE Models

- small number of common shocks
  - stochastic singularity
  - measurement error $\Rightarrow$ factor structure
- identification and estimation
- Bayesian analysis in a large N and T setting

Conclusion:

- the factor model is a useful way of achieving dimension reduction
- factor estimates have good properties when N, T are large
- generated new theory and new applications

Thank You!