Asymptotic Inference for Common Factor Models in the Presence of Jumps

Yohei Yamamoto

Hitotsubashi University

March 22, 2016
Australian National University

Introduction

- A short list of empirical applications of large dimensional common factor models with principal component estimation:
  - Stock returns: Lehmann and Modest (1988), Connor and Korajczyk (1988)
  - Government and corporate bonds: Litterman and Scheinkman (1991), Elton et al. (1995), Ang and Piazzesi (2003), Ludvigson and Ng (2009)
  - CDS spreads: Eichengreen et al. (2012), Longstaff et al. (2011)
  - Currency returns: Lustig et al. (2011), Engel et al. (2014)
  - Macroeconomic time series: Stock and Watson (2002)

Introduction

- Financial and macroeconomic time series data often have infrequent large jumps
- When the jumps are considered outliers, the inference on the underlying common factors is seriously contaminated
    - ARIMA models : Fox (1972), Box and Tiao (1975), Tsay (1986), Chen and Liu (1993)
    - Unit root & cointegration tests : Franses and Haldrup (1994), Vogelsang (1999), Perron and Rodríguez (2003)
    - Conditionally heteroskedastic models : Franses and Ghijsels (1999), Charles and Darne (2005)
- This paper also investigates the case where the common factors jump

Introduction

1. Model and estimation
2. Asymptotic results
3. Applications
    3.1 Jump-correction algorithm
    3.2 Factor jump tests
4. Empirical examples

Model

- Factor model

$$x_{it}^{*} = \lambda_i' F_t + u_{it}, \quad i = 1, ..., N \text{ and } t = 1, ..., T,$$

where $\lambda_i : r \times 1$, $F_t : r \times 1$. None of them are observed

- We observe

$$x_{it} = x_{it}^{*} + z_{it},$$

$$z_{it} = \eta_t^c \delta_{it}^c + \eta_{it} \delta_{it},$$

where $\eta_t^c \sim i.i.d.B(p_c/T)$, $\delta_{it}^c \sim i.i.d.N(0, \sigma_{NT}^2)$,
$\eta_{it} \sim i.i.d.B(p/T)$, $\delta_{it} \sim i.i.d.N(0, \sigma_{NT}^2)$.

- $\sigma_{NT} = k_{NT}\sigma$, where $\sigma \in (0, \infty)$ and $k_{NT} \geq 0$ is an arbitrary function of N and T

Estimation

- Factors and factor loadings are estimated by the principal component method

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

  where $\hat{F}'\hat{F}/T = I_r$
- The number of factors r is estimated by information criteria of Bai and Ng (2002)
- If x_{it}^* is used, we obtain as $N, T \to \infty$,

$$N^{1/2}(\hat{F}_t - H'F_t) \Rightarrow N(0, \Omega_{\hat{F}}) \text{ under } \sqrt{N}/T \to 0 \quad (1)$$

$$T^{1/2}(\hat{\lambda}_i - H^{-1}\lambda_i) \Rightarrow N(0, \Omega_{\hat{\Lambda}}) \text{ under } \sqrt{T}/N \to 0 \quad (2)$$

$$\hat{r} \xrightarrow{p} r \quad (3)$$

Asymptotic Results

Theorem (Upper bounds for the magnitudes of outliers)

Suppose Assumptions 1–7 hold.

- *(i-a) If* $\eta_t^c = 0$ *and* $k_{NT} < \sqrt{T}$ *, then (1) holds*
- *(i-b) If* $\eta_t^c = 1$ *and* $k_{NT} < \sqrt{N}$ *, then* $\left\|\hat{F}_t - H'F_t\right\| = o_p(1)$
- *(ii) If* $k_{NT} \leq T/\sqrt{N}$*, then (2) holds*
- *(iii) If* $k_{NT} \leq \max\left\{1, T^{1/4}N^{-1/4}\right\}$*, then (3) holds*

Jump-correction algorithm

- We propose a jump-correction algorithm series-by-series for x_{it} without accounting for the factor structure

- Suppose individual jump-free series follow a stationary ARMA model: $\theta_i(L)x_{it}^* = v_{it}$, v_{it} white noise

  Step 1. Compute $\tau_i(t) = \hat{v}_{it}^* / \hat{\sigma}_i$, where $\hat{v}_{it}^*$ is the residual from the individual jump-free model using x_{it}

- **Step 2.** If $\max_{1 \leq t \leq T} |\tau_i(t)| < \xi$, the ith series exhibits no (more) jumps. Assume that $\hat{x}^*_{it} = x_{it}$, and proceed with the $(i+1)$th series.

  If $\max_{1 \leq t \leq T} |\tau_i(t)| \geq \xi$,

  $$\hat{T}_i = \arg\max_{1 \leq t \leq T} |\tau_i(t)|$$

  is considered a possible jump location. Go to Step 3.

- **Step 3.** Estimate the realized jump magnitude with least squares estimation of coefficient ω_i in the regression

$$\hat{v}_{it}^* = \omega_i w_{it} + \epsilon_{it}, \quad \text{for } t = 1, ..., T,$$

where $w_{it} = 0$ for $t < \hat{T}_i$, $w_{it} = 1$ for $t = \hat{T}_i$, and $w_{it} = -\hat{\theta}_{il}$ for $t = \hat{T}_i + l$

Compute $\hat{x}_{it}^* = x_{it} - \hat{\omega}_i w_{it}$. Go back to Step 1 and use $\hat{x}_{it}^*$ as a new x_{it}.

Jump-correction algorithm

Theorem (Jump-free estimates)

Suppose that F, Λ and r are estimated using $\hat{x}^*_{it}$ and $\hat{\omega}_i - \omega_i = O_p(1)$ for every jump detected by the algorithm.

- *(i-a) If $\eta^c_t = 0$, then (1) holds*
- *(i-b) If $\eta^c_t = 1$, then $\left\|\hat{F}_t - H'F_t\right\| = o_p(1)$*
- *(ii) (2) holds under an additional condition $\sqrt{N}/T \to c$ $(0 \leq c < \infty)$*
- *(iii) (3) holds*

Monte Carlo simulation

- Setting

$$\begin{aligned} x_{it}^* &= \lambda_i' f_t + u_{it}, \\ x_{it} &= x_{it}^* + z_{it}, \\ z_{it} &= \eta_t^c \delta_{it}^c + \eta_{it} \delta_{it}, \end{aligned}$$

where

- $f_t \sim i.i.d.N(0, I_r),\ \lambda_i \sim i.i.d.N(0, I_r),\ u_{it} \sim i.i.d.N(0,1)$
  $\eta_t^c \sim i.i.d.B(p_c/T),\ \delta_{it}^c \sim i.i.d.N(0, \sigma^2)$
  $\eta_{it} \sim i.i.d.B(p/T),\ \delta_{it} \sim i.i.d.N(0, \sigma^2)$

Monte Carlo simulation

- Coverage ratios of the common component $\lambda_1 f_T$

(at the 90% level)

σ	N=20, T=500 no correction	N=20, T=500 correction	N=100, T=100 no correction	N=100, T=100 correction	N=500, T=20 no correction	N=500, T=20 correction
	pc=1, p=0					
0	0.88	0.88	0.89	0.89	0.87	0.86
5	0.87	0.87	0.88	0.89	0.51	0.86
10	0.84	0.87	0.58	0.89	0.42	0.86
50	0.58	0.88	0.50	0.90	0.42	0.86
100	0.58	0.88	0.50	0.90	0.43	0.85
	pc=0, p=1					
0	0.88	0.88	0.90	0.90	0.87	0.86
5	0.87	0.87	0.89	0.89	0.87	0.85
10	0.87	0.87	0.89	0.89	0.76	0.85
50	0.40	0.87	0.18	0.90	0.14	0.84
100	0.15	0.88	0.11	0.89	0.11	0.85

Monte Carlo simulation

- Average length of the common component $\lambda_1 f_T$

(at the 90% level)

	N=20, T=500		N=100, T=100		N=500, T=20	
σ	no correction	correction	no correction	correction	no correction	correction
	pc=1, p=0					
0	0.60	0.60	0.41	0.41	0.59	0.59
5	0.60	0.60	0.43	0.41	0.47	0.66
10	0.60	0.60	0.43	0.42	0.34	0.67
50	0.72	0.60	0.38	0.41	0.34	0.62
100	0.71	0.60	0.38	0.41	0.33	0.60
	pc=0, p=1					
0	0.60	0.60	0.41	0.41	0.59	0.59
5	0.61	0.60	0.43	0.42	0.77	0.66
10	0.60	0.59	0.46	0.40	0.91	0.67
50	0.54	0.61	0.29	0.42	0.19	0.64
100	0.36	0.60	0.16	0.41	0.13	0.62

Monte Carlo simulation

- Number of factors ($r = 4$)

σ	N=50, T=200 no correction	N=50, T=200 correction	N=100, T=100 no correction	N=100, T=100 correction	N=200, T=50 no correction	N=200, T=50 correction
	pc=1, p=0					
0	4.00	4.00	4.00	4.00	4.00	4.03
5	5.01	4.00	5.02	4.30	5.06	5.08
10	5.00	4.00	4.98	4.11	5.01	4.95
50	4.99	4.00	4.97	4.00	5.01	4.02
100	5.01	4.00	5.01	4.00	4.98	4.02
	pc=0, p=1					
0	4.00	4.00	4.00	4.00	4.00	4.03
5	4.00	4.00	4.00	4.00	4.00	4.01
10	4.14	4.00	4.07	4.00	4.02	4.01
50	19.68	4.00	3.56	4.00	1.04	4.01
100	19.88	4.00	5.21	4.00	1.09	4.02

Factor jump tests

- Suppose common outliers are detected at $t = T^c$. They may be resulted from a jump in factors.

- We consider a test for the null hypothesis:

$$x_{i,T^c} = \lambda_i' F_{T^c} + \eta_{T^c}^c \delta_{i,T^c}^c + u_{i,T^c}, \tag{4}$$

against the alternative hypothesis

$$\begin{aligned} x_{i,T^c} &= \lambda_i'(F_{T^c} + J_{T^c}) + u_{i,T^c}, \\ &= \lambda_i' F_{T^c} + \lambda_i' J_{T^c} + u_{i,T^c}, \end{aligned} \tag{5}$$

where $J_{T^c} \sim (0, \sigma_{NT}^2 I_r)$

Factor jump tests

- We use a cross-sectional regression of the residuals on the jump-free factor loading estimates

$$\hat{u}_{iT^c} = \gamma_0 + \hat{\lambda}_i' \gamma_1 + \varepsilon_i, \quad i = 1, ..., N$$

- Implement an F test for $H_0 : \gamma_1 = 0$

$$F^J = \frac{(SSR_r - SSR_u)/r}{SSR_u/(N-r-1)}.$$

Monte Carlo simulation

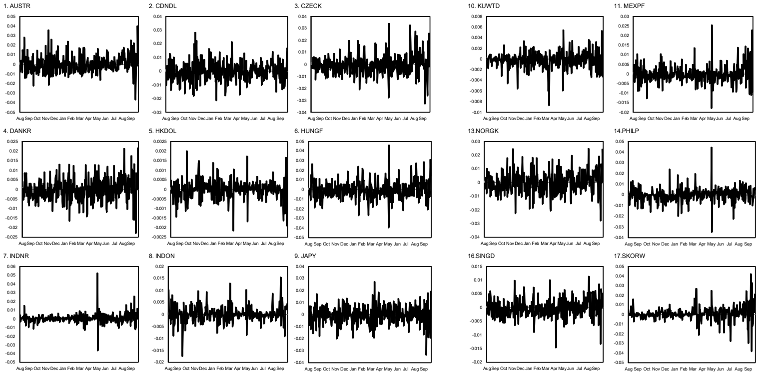
- Size and power of the factor jump test

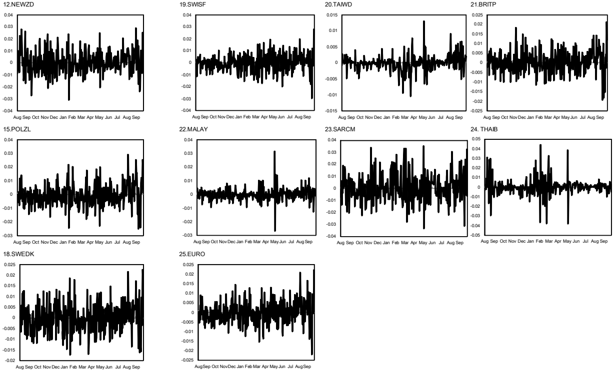
(at the 5% nominal level)

σ	N=20 T=500	N=50 T=200	N=100 T=100	N=200 T=50	N=500 T=20
Size					
5	0.07	0.01	0.03	0.03	0.65
10	0.06	0.08	0.12	0.20	0.56
50	0.05	0.07	0.07	0.07	0.08
100	0.05	0.07	0.07	0.06	0.06
Power					
5	0.58	0.34	0.35	0.39	0.56
10	0.76	0.62	0.65	0.65	0.76
50	0.95	0.92	0.93	0.92	0.95
100	0.98	0.96	0.96	0.96	0.98

Empirical illustration (1)

- Daily returns of 25 currencies against US dollar (from Aug 1, 2007 to Sep 30, 2008)




12.NEWZD
19.SWISF
20.TAIWD
21.BRITP
15.POLZL
22.MALAY
23.SARCM
24. THAIB
18.SWEDK
25.EURO

Empirical illustration (1)

- The number of jumps

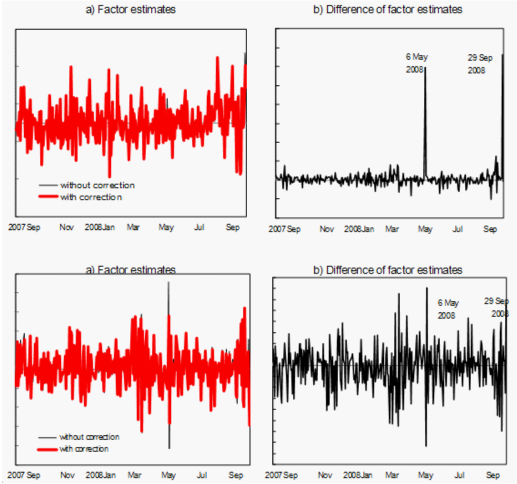
			# of jumps	common jump dates	
				May 06, 2008	Sep. 29, 2008
1	Australian Dollar	AUSTR	1		X
2	Canadian Dollar	CDNDL	0		
3	Czech Republic Koruna	CZECK	0		
4	Danish Krone	DANKR	0		
5	Hong Kong Dollar	HKDOL	7	X	X
6	Hungrian Forint	HUNGF	1		
7	Indian Rupee	INDNR	4	X	
8	Indonesian Rupiah	INDON	11	X	
9	Japanese Yen	JAPYN	0		
10	Kuwaiti Dinar	KUWTD	6	X	
11	Mexican Peso	MEXPF	3	X	X
12	New Zealand Dollar	NEWZD	0		
13	Norwegian Krone	NORGK	0		
14	Philippines Peso	PHILP	3	X	
15	Polish Zloty	POLZL	0		
16	Singaporean Dollar	SINGD	1		
17	South Korean Won	SKORW	21	X	X
18	Swedish Krona	SWEDK	0		
19	Swiss Franc	SWISF	0		
20	UK Pound	BRITP	11		
21	Malaysian Ringgit	MALAY	0		
22	Taiwan Dollar	TAIWD	2	X	
23	South African Rand	SARCM	0		
24	Thai Baht	THAIB	20	X	
25	Euro	EURO	0		

Notes : 1. "# of jumps" indicates how many jumps are detected by the proposed method between 1 Aug 2008 and 30 Sep. 2008.

2. The common jumps dates are those on which more than 3 currencies have a jump. These currencies have a mark "X".

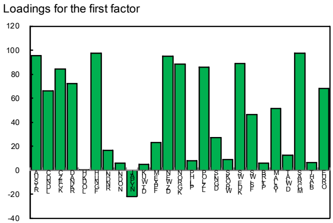
Empirical illustration (1)

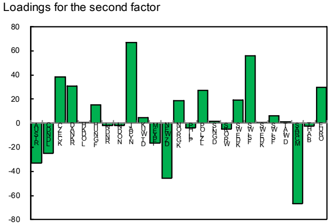
- Common factors



Empirical illustration (1)

- Jump-free factor loadings





Empirical illustration (1)

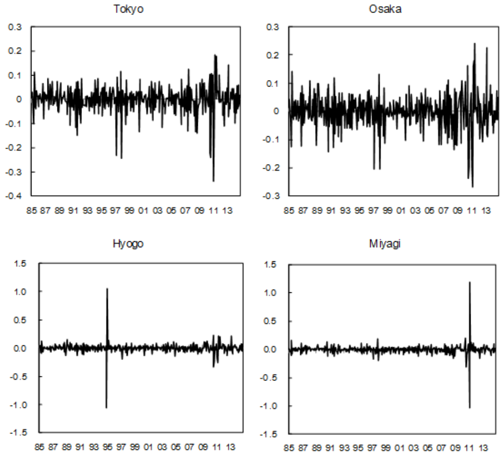
- Factor jump tests

	F	p-value	t (1st factor)	p-value	t (2nd factor)	p-value
2008/5/6	3.58**	(0.04)	-2.65**	(0.01)	-0.73	(0.47)
2008/9/29	2.90*	(0.07)	1.45	(0.16)	-1.71	(0.10)

Note: ** and * indicate significance at the 5% and 10% levels, respectively.

Empirical illustration (2)

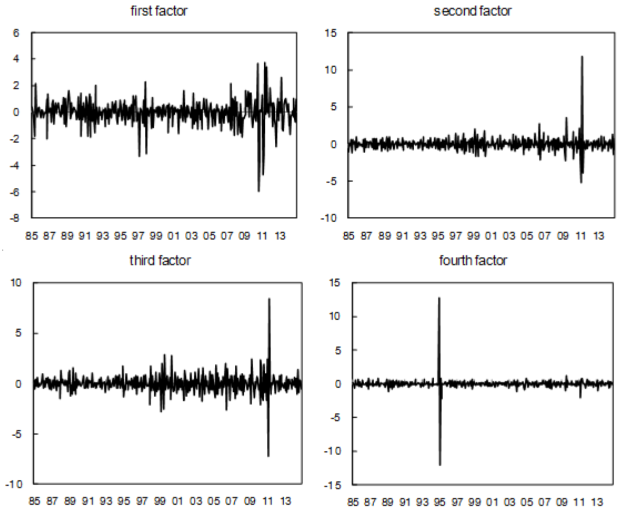
- Monthly growth of new car registration in Japanese 47 prefectures (from Jan. 1985 to Dec. 2014)



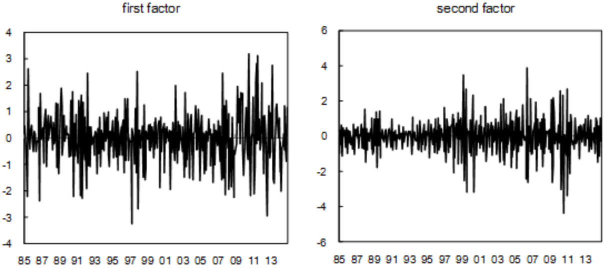
- Prefectures that have a jump in the two great earthquakes
  - Jan 1995: Great Hanshin earthquake
  - Mar 2011: Great East Japan earthquake

	# of pref.	Prefectures that have a jump
Jan 1995	1	Hyogo
Mar 2011	23	Hokkaido, Aomori, Iwate, Miyagi, Akita, Yamagata Fukushima, Ibaraki, Tochigi, Gunma, Saitama Chiba, Tokyo, Kanagawa, Yamanashi, Gifu, Nagano, Shizuoka, Aichi, Shimane, Okayama, Hiroshima, Fukuoka

- Factor estimates without jump correction
  - $\hat{r} = 4$



- Factor estimates with jump correction
    - $\hat{r} = 2$



- Factor jump tests

	F	p-value	t (1st factor)	p-value	t (2nd factor)	p-value
Jan 1995	0.04	(0.96)	0.06	(0.95)	0.17	(0.87)
Mar 2011	6.61***	(0.00)	2.21**	(0.03)	0.84	(0.40)

Note: ** and * indicate significance at the 5% and 10% levels, respectively.

Conclusion

- This paper has explored the effects of infrequent large jumps on recently developed large dimensional common factor models using principal components estimation
- Under a popular setting, we have derived the upper bounds of jump magnitudes. The results are used to propose 1) series-by-series jump correction algorithm and 2) factor jump tests
- Monte Carlo experiment and two empirical examples illustrate usefulness of these results