高次元データのモデリング理論と金融市場の実証分析

慶應義塾大学

安道知寛

Agenda

- Issues to be discussed

- Modeling high-dimensional panel data with applications

- Summary

Part 1: Issues

Issue 1

- Statistical independence?

Issue 2

- Simplicity

Issue 3

- Limited access to information

Issue 4

- Underlying structure

Part 2:
Modeling high-dimensional panel data

Details on theoretical results can be found in our papers

- Ando, T. and Bai, J. (2014a). Panel data models with grouped factor structures under unknown group membership. *Journal of Applied Econometrics*, forthcoming.

- Ando, T. and Bai, J. (2014b). Selecting the regularization parameters in high-dimensional panel data models: consistency and efficiency. *Econometric Reviews*, forthcoming.

- Ando, T. and Bai, J. (2014c). Asset pricing with general multifactor structure. *Journal of Financial Econometrics*, forthcoming.

- Ando, T. and Bai, J. (2014d). A simple new test for slope homogeneity in panel data models with interactive effects. Working Paper.

- Ando, T. and Bai, J. (2015). The subprime financial crisis and spillover effects on global financial markets. Working Paper.

- **Please see the references therein**

Some related econometric literatures

- **Dynamic exact factor model** (Geweke, 1977; Sargent and Sims, 1977)

- **Static approximate factor model** (Chamberlain and Rothschild, 1983)

- **Generalized dynamic factor model** (Forni et al., 2000; Forni and Lippi, 2001; Amengual and Watson, 2007; Hallin and Liska, 2007),

- **Bayesian factor model** (Aguilar and West, 2000; Lopes and West, 2004; Lopes et al., 2008; Ando, 2009; Bhattacharya and Dunson, 2011; Tsay and Ando, 2012).

- **Approximate factor model** (Chamberlain and Rothschild, 1983; Connor and Korajczyk, 1986, 1988; Jones, 2001; Bai and Ng, 2002; Bai, 2003, Korajczyk and Sadka, 2008);

Some related econometric literatures

- **Hierarchical factor model** (Moench, Ng, and Potter 2012; Diebold et al. 2008; Kose et al. 2008; Wang 2014 and Moench and Ng 2011).

- **Panel models with multifactor error structures** (Pesaran, 2006; Bai 2009)

- **Shrinkage estimation** (Tibshirani 1996, Zou 2006, Fan and Li 2001, Fan and Peng 2004, Caner 2009; Belloni, Chernozhukov and Hansen 2010; Caner and Zhang 2012; Bai and Liao 2013; Fan and Liao 2013; Liao 2013; Lu and Su 2013).

- **Panels with unknown group structures** (Bonhomme and Manresa 2014).

Asset pricing model

Excess returns

Error

$$y_{it} = \alpha_i + \beta_{i1} z_{it,1} + \cdots + \beta_{ip} z_{it,p} + u_{it}$$

Risk factor 1

Risk factor p

- Macroeconomic factors

- Fundamental factors

- Unobservable factors

World 20 major stock market

Source: http://www.mapsofworld.com/stock-exchange/

Ando, T. and Bai, J. (2014c).
Asset pricing with general multifactor structure.
Journal of Financial Econometrics, forthcoming.

Model description (1/4)

Excess returns

$$y_{it} = x'_{it}\boldsymbol{\beta}_i + f'_{c,t}\boldsymbol{\lambda}_{c,i} + f'_{gi,t}\boldsymbol{\lambda}_{gi,i} + \varepsilon_{i,t}$$

Common factors

Predictors   **Group specific factors**

Model description (2/4)

$$y_{it} = x'_{it}\boldsymbol{\beta}_i + f'_{c,t}\boldsymbol{\lambda}_{c,i} + f'_{g_i,t}\boldsymbol{\lambda}_{g_i,i} + \varepsilon_{i,t}$$

- S: the number of groups (known)
- Group membership (known)
- r: the number of common factors
- rj: the number of group-specific factors
- N: total number of assets
- Nj: the number of assets within the group j

Model description (3/4)

$$y_{it} = x'_{it}\boldsymbol{\beta}_i + f'_{c,t}\boldsymbol{\lambda}_{c,i} + f'_{g_i,t}\boldsymbol{\lambda}_{g_i,i} + \varepsilon_{i,t}$$

- **Common/group-specific factors**
  - assume orthogonality
- **Observable risk factors (x)**
  - allow correlations between the common factor and group-specific factors
  - but not perfectly correlated

Model description (4/4)

$$y_{it} = x'_{it}\boldsymbol{\beta}_i + f'_{c,t}\boldsymbol{\lambda}_{c,i} + f'_{g_i,t}\boldsymbol{\lambda}_{g_i,i} + \varepsilon_{i,t}$$

- **Security-specific returns (ε)**
  - zero mean
  - cross-sectional/serial dependence and heteroscedasticity.
  - Independent of the observable risk factors, the common factor and group-specific factors

Shrinkage estimation

$$\ell(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N, F_c, F_1, \ldots, F_G, \Lambda_c, \Lambda_1, \ldots, \Lambda_G | r, r_1, \ldots, r_G, \kappa)$$

$$= \sum_{i=1}^{N} \| \boldsymbol{y}_i - X_i \boldsymbol{\beta}_i - F_c \boldsymbol{\lambda}_{c,i} - F_{g_i} \boldsymbol{\lambda}_{g_i,i} \|^2 + T \sum_{i=1}^{N} p_{\kappa,\gamma}(|\boldsymbol{\beta}_i|)$$

subject to the constrains on the factors

- SCAD (Fan and Li (2001, JASA)) , Asymptotic principal components (Connor and Korajczyk (1986, JFE)) are jointly employed

Cp criterion

- We need to select
  - the regularization parameter
  - the number of common factors
 - the number of group-specific factors.

Cp = Fitness +Penalty on model complexity

China A-shares, B-shares

- A-shares
  - Traded in mainland currencies.
  - Generally only available for purchase by mainland citizens;
  - Used to be the market for mainlanders.
- B shares
  - Traded in foreign currencies.
  - Was only for foreign investors until February 2001

Empirical questions

- What type of observable risk factors explain the individual excess returns?

- Numbers of common and group-specific factors.

- Meanings of extracted factors.
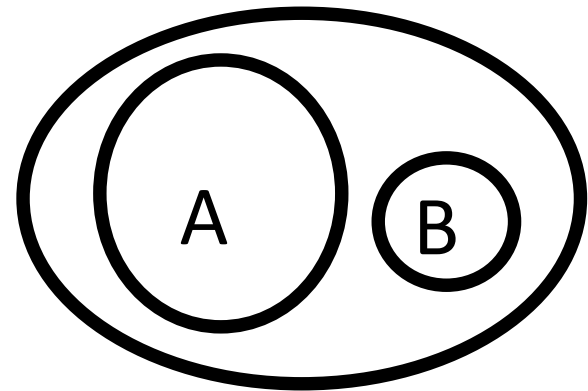
Data

- Monthly excess returns

- March 2002 to December 2012 period

- Systematically exclude stocks with missing returns data.

- 1,039 A-share firms and 102 B-share firms.

Observable risk factors

- China macroeconomic variables
  - (M2, CPI, Macroeconomic climate index)
- Exchange rates
  - (Dollar, EURO, UK, HK, Yen)
- Commodity price index
  - (Metals, Aluminum, Copper, Oil, Gold, Gas, Nickel, Silver)
- Major stock market index
  - (MSCI world, MSCI EU, MSCI China, SP500, HANG SENG, FTSE100, TOPIX)

Ando and Bai (2014c). An impact on observable factors (1/2)

Variables	A-shares	B-shares
China macroeconomic variables		
MACROECONOMIC CLIMATE INDEX (LEADING)	50.33	67.64
MONEY SUPPLY - M2	35.12	24.50
MACROECONOMIC CLIMATE INDEX (COINCIDENT)	4.52	0.98
MACROECONOMIC CLIMATE INDEX (LAGGING)	63.90	60.78
CONSUMER PRICE INDEX	0.00	0.98
Exchange rates		
CHINESE YUAN to US DOLLAR	18.76	32.35
CHINESE YUAN to YEN	16.55	28.43
CHINESE YUAN to EURO	4.13	1.96
CHINESE YUAN to UK POUND	35.70	73.52
CHINESE YUAN to HK DOLLAR	3.75	1.96

Ando and Bai (2014c).
An impact on observable factors (2/2)

Variables	A-shares	B-shares
Commodity price index		
S&P GSCI Industrial Metals Spot	7.12	22.54
S&P GSCI Aluminum Spot	3.17	34.31
S&P GSCI Copper Spot	17.22	0.00
S&P GSCI Crude Oil Spot	5.10	28.43
S&P GSCI Gold Spot	18.09	9.80
S&P GSCI Natural Gas Spot	0.76	0.00
S&P GSCI Nickel Spot	19.05	24.50
S&P GSCI Silver Spot	14.91	13.72
Major stock market indexes		
S&P 500 INDEX	10.10	14.70
MSCI WORLD INDEX	0.76	0.00
FTSE 100 INDEX	24.63	53.92
MSCI EUROPE INDEX	1.05	0.00
TOPIX INDEX	2.79	7.84
HANG SENG INDEX	11.06	0.98
MSCI CHINA INDEX	47.06	93.13

The number of factors

- Common factors:
  - 2
- A-share group-specific factors:
  - 1
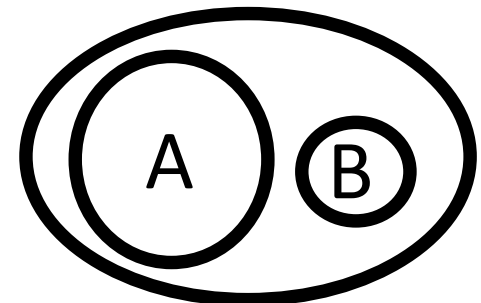- B-share group-specific factors:
  - 3
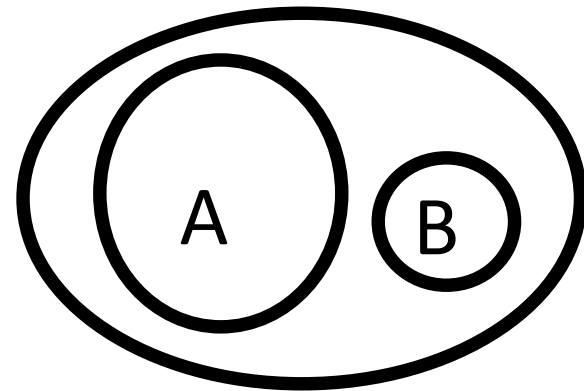
Interpretation of the factors

- Consumer confidence index in China (CCI)
- CBOE volatility index (VIX)
- Market excess returns of A-shares (ER–A)
- Market excess returns of B-shares (ER–B)
- Book-to-market ratio (HML)
- Market capitalization (SMB)

Ando and Bai (2014c) Common factors

		CCI	VIX	ER–A	ER–B	HML	SMB
Common factors							
	First	−0.070	−0.017	0.579	0.425	−0.066	0.443
	SD	0.077	0.116	0.084	0.099	0.085	0.083
	t-value	−0.915	−0.151	6.873	4.277	−0.783	5.306
	Second	−0.107	−0.021	0.084	−0.039	0.046	−0.172
	SD	0.079	0.081	0.075	0.091	0.069	0.076
	t-value	−1.355	−0.262	1.109	−0.434	0.670	−2.246

Ando and Bai (2014c)
Group-specific factors

		CCI	VIX	ER–A	ER–B	HML	SMB
Group A							
	First	0.026	0.005	0.067	0.018	0.056	0.005
	SD	0.073	0.077	0.096	0.081	0.063	0.098
	t-value	0.356	0.068	0.698	0.222	0.892	0.052
Group B							
	First	0.073	−0.058	−0.079	−0.357	−0.094	−0.239
	SD	0.081	0.059	0.101	0.11	0.074	0.095
	t-value	0.903	−0.988	−0.779	−3.235	−1.261	−2.507
	Second	−0.126	−0.097	−0.019	0.065	−0.077	0.057
	SD	0.100	0.085	0.092	0.098	0.079	0.085
	t-value	−1.26	−1.133	−0.204	0.658	−0.978	0.678
	Third	0.043	0.062	−0.095	−0.024	−0.167	−0.029
	SD	0.084	0.060	0.098	0.114	0.085	0.113
	t-value	0.513	1.027	−0.972	−0.214	−1.958	−0.264

Robustness check

- Some companies issue "twin" A and B shares.
  - Contained 50 "twin" A and B shares.

- The effect of foreign denominated currencies.
  - Express the B-share returns in Chinese yuan

Ando, T. and Bai, J. (2014a).
Panel data models with grouped factor structures
under unknown group membership.
Journal of Applied Econometrics, forthcoming.

Issue

- Do we know real group membership?

Model description: Ando and Bai (2014a).

Group specific factors

$$y_{it} = x'_{it}\beta_{g_i} + f'_{g_i,t}\lambda_{g_i,i} + \varepsilon_{i,t},$$

Group specific coefficients

- Further unknowns
  - Number of groups
  - Group membership indicator
- Estimation by the shrinkage methods
- Cp type model selection

Theoretical results

- Consistency of membership, $\hat{\boldsymbol{\beta}}_i$, $\{\hat{F}_g,\ g=1,...,G\}$

- Asymptotic normality of $\hat{\boldsymbol{\beta}}_i$

- Variable selection consistency of $\hat{\boldsymbol{\beta}}_i$

- Cp criterion for model selection

The number of groups

- China stock market

- We found 6 groups in our sample

Ando and Bai (2014a) Scatter matrices of the estimated group membership against nominal classification

Classification	G1	G2	G3	G4	G5	G6
1 Location of stock exchanges						
Shanghai stock exchange	179	67	132	77	105	81
Shenzhen stock exchange	125	29	94	64	95	93
2 Types of share						
A-shares	211	95	224	141	196	172
B-shares	93	1	2	0	4	2
3 Category based on Industry						
Chemicals, Construction, Manufacturing	76	15	70	36	53	49
Food, Beverages, Personal Goods	40	14	24	21	25	13
Gas, Metals, Mining, Oil	42	16	16	17	17	26
Banks, Financial Services, Real Estate	30	6	25	15	23	17
Retails	29	18	26	19	19	21
Utilities	17	8	16	6	19	9
Pharmaceuticals, Health	24	6	21	10	16	12
Information Technology	27	8	21	9	19	11
Others	11	4	4	5	7	13

Ando, T. and Bai, J. (2015).
The subprime financial crisis and spillover effects on global financial markets, Working Paper

Model and its construction: Ando and Bai (2015)

$$y_{i,t} = \mathbf{x}'_{i,t}\boldsymbol{\beta}_i + \mathbf{f}'_{g_i,t}\boldsymbol{\lambda}_{g_i,i} + \varepsilon_{i,t,}$$

- Further unknowns
  - Number of groups
  - Group membership indicator
- Estimation by the shrinkage methods
- Cp type model selection

Data description (1/2)

- Financial industry
  - Banks
  - Life Insurance
  - Nonlife Insurance
  - Financial Services
  - Real Estate Investment and Services
- 109 international stock exchanges
  - US (NYSE, NASDAQ, Non-NASDAQ OTC)
  - Asia (TSE, Shanghai, Shenzen, Hong Kong,…)
  - Europe (Frankfurt, London, Luxembourg,…)
  - Emerging (BSE, Thailand,…)

Data description (2/2)

- Periods
  - May 1 2006 to December 31, 2006
  - May 1 2007 to December 31, 2007
  - February 1 2008 to August 31, 2008
  - September 1 2008 to March 31, 2009
  - May 1 2009 to December 31, 2009

- Systematically exclude stocks with missing
- More than 6,000 stocks in each of the periods.

Observable risk factors

- Averages of lagged returns of each market

- More than 50 stocks among approx. 6000 stocks.

- 31 predictors, including NYSE, TSE, Shanghai, etc.

Period 1

Period 4

Number of groups

- Period 1: S = 11 groups.

- Period 2: S = 13 groups.

- Period 3: S = 15 groups.

- Period 4: S = 17 groups.

- Period 5: S = 14 groups.

- The number of group-specific factors varies over the groups

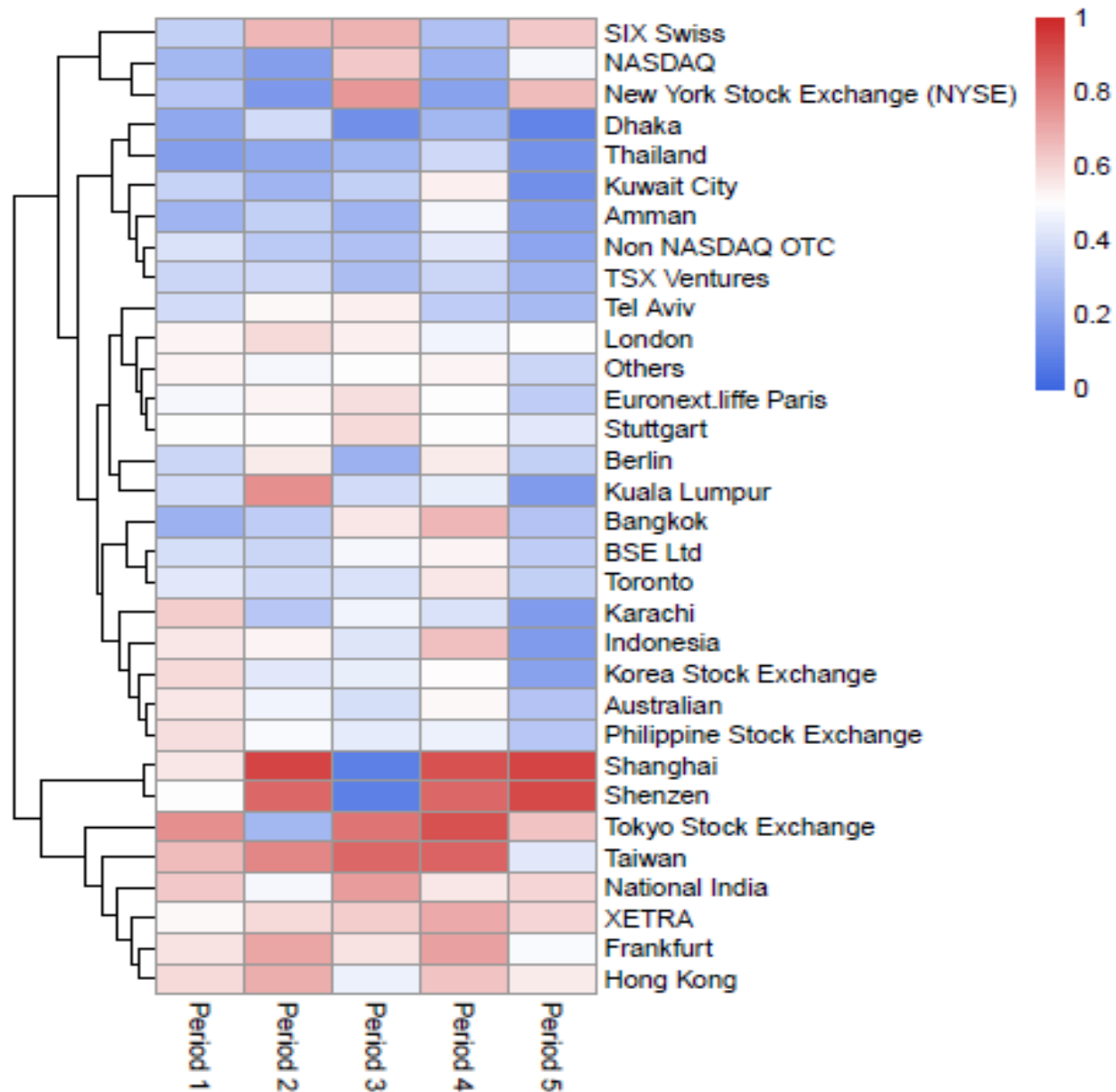# Period 2: Market vs grouping

Period 4: Market vs grouping

Factor interpretation

- Thomson Routers Global Index

- Global, U.S., EU.

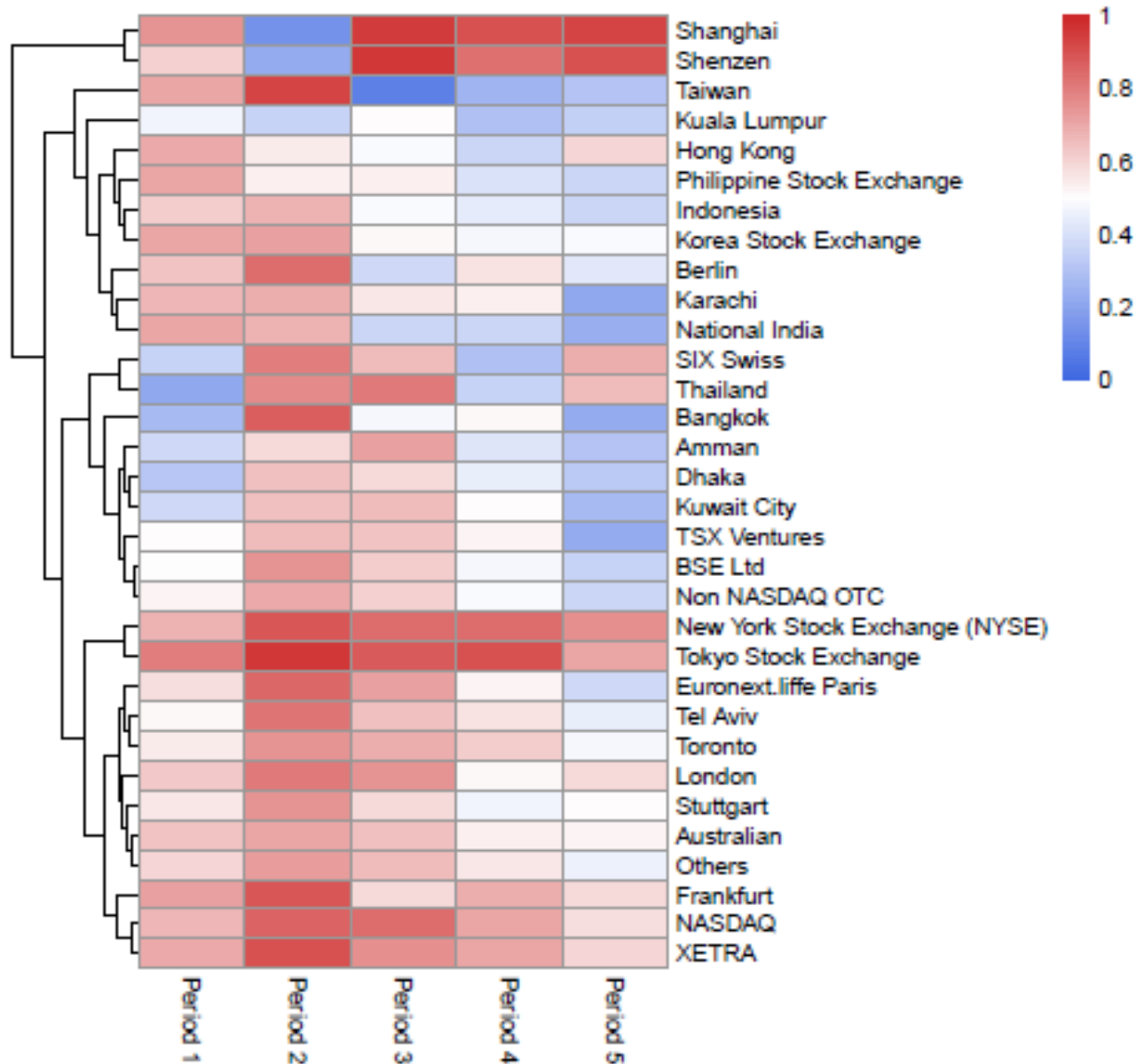- The equity market performance of Financial, Banks, Insurance, Real Estate, and REIT.

Exposure to U.S. factors
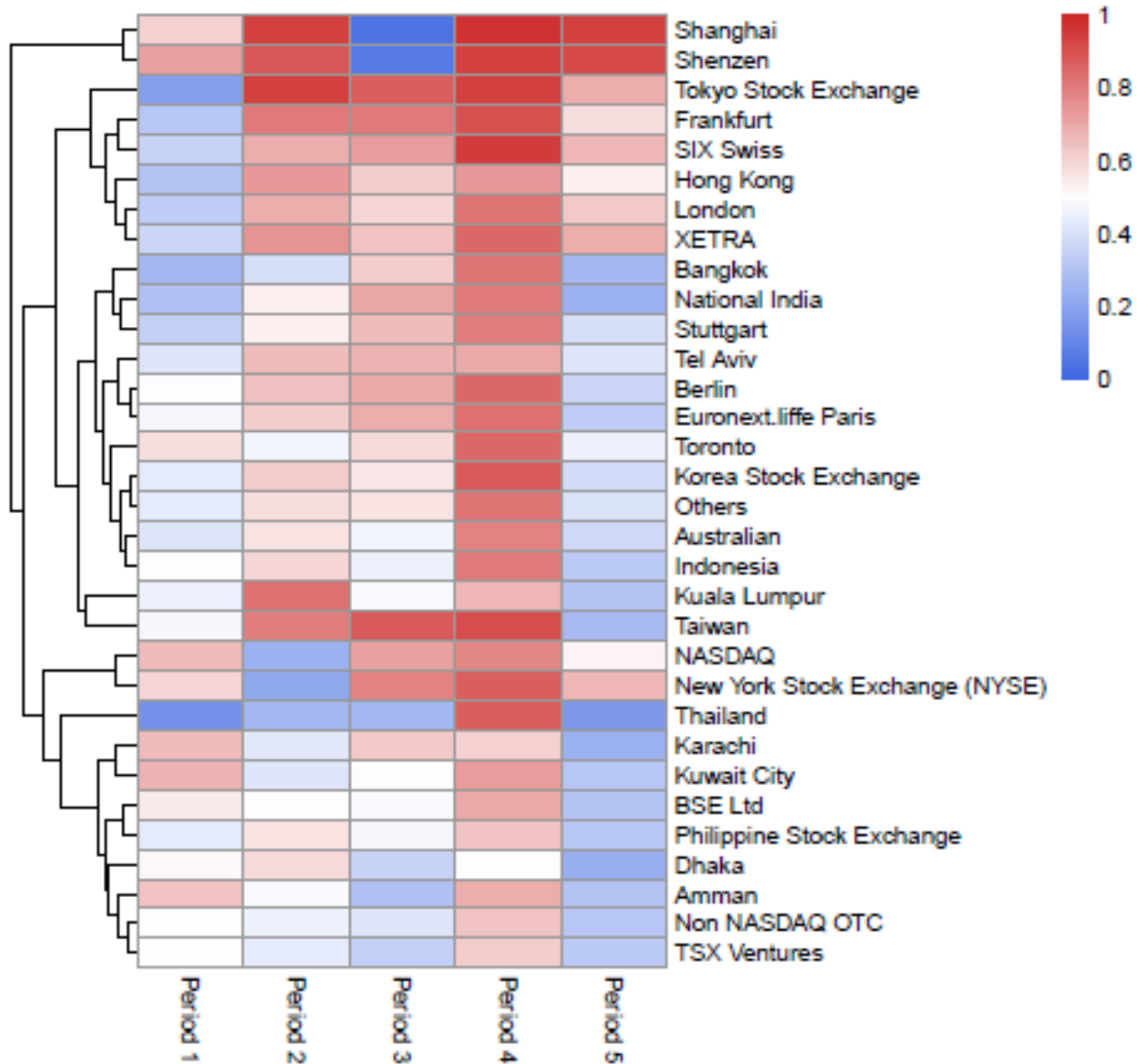
Exposure to global factors

Exposure to EU factors

Robustness check

- Different sub-period specification

- Analysis of daily returns instead of two day rolling average.

Summary

- Assumptions

- Reality, and business practice

- Recent advancement on econometric methods

- Applications