Econometric Measures of Financial Risk in High Dimensions

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

doctor rerum politicarum
(Doktor der Wirtschaftswissenschaft)
eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
Humboldt-Universität zu Berlin

von
M.Sc. Shi Chen

Präsidentin Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Dr. Christian D. Schade

Gutachter:

1. Prof. Dr. Wolfgang Karl Härdle, Humboldt-Universität zu Berlin

2. Prof. Dr. Melanie Schienle, Karlsruher Institut für Technologie

List of Figures

List of Tables

Abstract

Abstract

Modern financial system is complex, dynamic, high-dimensional and often possibly non-stationary. All these factors pose great challenges in measuring the underlying financial risk, which is of top priority especially for market participants. High-dimensionality, which arises from the increasing variety of the financial products, is an important issue among econometricians. A standard approach dealing with high dimensionality is to select key variables and set small coefficient to zero, such as lasso. In financial market analysis, such sparsity assumption can help highlight the leading risk factors from the extremely large portfolio, which constitutes the robust measure for financial risk in the end. In this paper we use penalized techniques to estimate the econometric measures of financial risk in high dimensional, with both low-frequency and high-frequency data. With focus on financial market, we could contrcut the risk network of the whole system which allows for identification of individual-specific risk.

In Chapter 1, we conduct network analysis for limit order books (LOB) across stocks for a better understanding of market impact, and show how network for LOB can be constructed in the presence of microstructure noise and non-synchronous trading. This paper contributes to directed network estimation through penalized Vector autoregressive (VAR) approach. The connectedness table is directly derived from generalized impulse response function with attractive property of order invariance, by way of bootstrapping techniques. The directional connectedness "from" and "to" are associated with the forecast error variation for specific order book across various stocks when the arising shocks transmit from one stock to the others. To balance the sparsity and estimation accuracy, a moderately tuning parameter in penalized VAR is determined by Bayesian information criterion (BIC), and we apply ordinary least squares (OLS) post-Lasso for our estimator, which can be configured to reduce finite-sample bias and ensure better model performance. With large scale or high-dimensionality in this study, we introduce prescreening based on graphical lasso which filters out less relevant variables. Moreover we look for the short-horizon large portfolio allocation decisions, and provide novel empirical evidence using high frequency data trading in NASDAQ market.

Chapter 2 studies the German power derivative market, we could be able to identify the relevant risk drivers from the portfolio that are unknown to the power market investors. The selection of important market drivers via iterated-SIS algorithm enables us to investigate a ultra high-dimensional portfolio, since the the number of

parameters to estimate increases quadratically in the number of variables included in the context of VAR estimation. In our paper, the network of interest follows the Dieboldt-Yilmaz tradition. With the wide range of power derivative contracts trading in the German electricity market, we are able to identify, estimated the risk contribution of individual power contract, this helps us to have a better understanding of the German power market functioning and environment.

In Chapter 3, we aim to estimate the financial risk arising from bond market. As we all know, inflation expectation is acknowledged to be an important indicator for policy makers and financial investors. To capture a more accurate real-time estimate of inflation expectation on the basis of financial markets, we propose an arbitrage-free model across different countries in a multi-maturity term structure, where we first estimate inflation expectation by modelling the nominal and inflation-indexed bond yields jointly for each country. The Nelson-Siegel model is popular in fitting the term structure of government bond yields, the arbitrage-free model we proposed is the extension of the arbitrage-free dynamic Nelson-Siegel model proposed by Christensen et al. (2011). We discover that the extracted common trend for inflation expectation is an important driver for each country of interest. Moreover, the model will lead to an improved forecast in a benchmark level of inflation and will provide good implications for financial markets.

Zusammenfassung

Das moderne Finanzsystem ist komplex, dynamisch, hochdimensional und oftmals nicht stationär. All diese Faktoren stellen große Herausforderungen beim Messen des zugrundeliegenden Finanzrisikos dar, das speziell für Marktteilnehmer von oberster Priorität ist. Hochdimensionalität, die aus der ansteigenden Vielfalt an Finanzprodukten entsteht, ist ein wichtiges Thema für Ökonometriker. Ein Standardansatz, um mit hoher Dimensionalität umzugehen, ist es, Schlüsselvariablen auszuwählen und kleine Koeffizientenen auf null zu setzen, wie etwa Lasso. In der Finanzmarktanalyse kann eine solche geringe Annahme helfen, die führenden Risikofaktoren aus dem extrem großen Portfolio, das letztendlich das robuste Maß für finanzielles Risiko darstellt, hervorzuheben. In dieser Arbeit nutzen wir penalisierte Verfahren, um die ökonometrischen Maße für das finanzielle Risiko in hoher Dimension zu schätzen, sowohl mit nieder-, als auch hochfrequenten Daten. Mit Fokus auf dem Finanzmarkt, können wir das Risikonetzwerk des ganzen Systems konstruieren, das die Identifizierung individualspezifischen Risikos erlaubt.

In Kapitel 1 führen wir eine Netzwerkanalyse für Limit Order Books (LOB) über Aktien zum besseren Verständnis der Marktwirkung durch und zeigen, wie ein Netzwerk für LOB in Anwesenheit von Mikrostruktur-Noise und nicht synchronem Handel konstruiert werden kann. Diese Arbeit leistet einen Beitrag zur gerichteten Netzwerkschätzung durch einen penalisierten Vektorautoregressiven (VAR) Ansatz. Die Vernetzungstabelle ist direkt aus verallgemeinerten Impulse-Response Funktionen mit anziehender Eigenschaft von Order Invarianz, mittels Bootstrapping-Techniken. Die gerichteten Vernetzungen "vonünd ßußtehen im Zusammenhang mit der Vorhersagefehlervariation für spezifische Orderbücher über verschiedene Aktien, wenn die aufkommenden Schocks von Aktie zu Aktie transmittiert werden. Um Einfachheit und Schätzgenauigkeit zu balancieren, wird ein moderater Tuningparameter in penalisierten VAR durch das Bayessche Informationskriterium (BIC) bestimmt und Ordinary Least Squares (OLS) post-Lasso für unseren Schätzer angewandt, der so konfiguriert werden kann, dass endliche Stichprobenverzerrungen reduziert werden und eine bessere Modell-Performance gesichert wird. Mit großen Umfängen oder Hochdimensionalität in dieser Studie führen wir Prescreening basiert auf graphischem Lasso ein, das weniger relevante Variablen herausfiltert. Des Weiteren untersuchen wir kurzfristige große Portfolio Allokationsentscheidungen und geben neue empirische Evidenz, indem wir hochfrequente Handelsdaten im NASDAQ-Markt nutzen.

Kapitel 2 untersucht den deutschen Energiederivatemarkt. Wir haben die Möglichkeit die relevanten Risikotreiber aus dem Portfolio identifizieren, die den Energiemarktinvestoren unbekannt sind. Die Auswahl wichtiger Markttreiber mittels iterated-SIS Algorithmen ermöglicht uns ein ultra-hochdimensionales Portfolio zu untersuchen, da die Anzahl der Schätzparameter quadratisch mit der Anzahl an Variablen, die im Kontext der VAR-Schätzung inkludiert sind, ansteigt. In unserer Arbeit folgt das Interessennetzwerk der Dieboldt-Yilmaz Tradition. Mit dem brei-

ten Spektrum an Energiederivaten, die im deutschen Elektrizitätsmarkt gehandelt werden, ist es uns möglich, den Risikobeitrag eines einzelnen Energiederivats zu identifizieren. Dies hilft uns die Funktionsweise und Umwelt des deutschen Energie-markts besser zu verstehen.

In Kapitel 3 zielen wir darauf, das finanzielle Risiko ausgehend vom Anleihe-markt zu schätzen. Wie wir wissen, ist die Inflationserwartung ein wichtiger Indi-kator für politische Entscheidungsträger und Finanzinvestoren. Um eine akkuratere Echtzeitprognose der Inflationserwartung auf Basis der Finanzmärkte einzufangen, schlagen wir ein arbitragefreies Modell über verschiedene Länder in einer multi-Fälligkeitslaufzeitstruktur, in dem wir zuerst durch gemeinsame Modellierung der nominalen und inflations-indexierte Anleiherenditen die Inflationserwartungen für jedes Land schätzen. Das Nelson-Siegel Modell ist populär darin die Laufzeitstruk-tur der Renditen von Staatsanleihen zu fitten, das arbitragefreie Modell, das wir vorstellten, ist die Erweiterung des arbitragefreien dynamischen Nelson-Siegel Mo-dell, vorgeschlagen von Christensen (2011). Wir stellen fest, dass der extrahierte gemeinsame Trend für die Inflationserwartung ein wichtiger Treiber für jedes Land von Interesse ist. Des Weiteren wird das Modell zu einer verbesserten Prognose auf Benchmarkniveau der Inflation führen und gute Implikationen für Finanzmärkte be-reitstellen.

Schlüsselwörter :   Limit Order Book, hochdimensional, verallgemeinerte impulse-response, Hochfrequenz, Finanzrisiko, Markteinfluss, large VAR, makroökonomi-sches Risiko, Maße für Finanzrisiko

Contents

1 Measuring the Information Contect of Limit Order Books

1.1 Introduction

Advancements in trading technologies allow an extremely rapid placement of buy and sell orders. These rapid-fire trading algorithms can make decisions in milliseconds. The changing of high frequency (HF) limit order book (LOB) gives us more insights into the market behavior. In an LOB shown in Figure **??**, the order book contains a quantity of limit orders and the corresponding price at which you would issue a "buy" or "sell" limit order. When an investor places an order to purchase or sell a stock, there are two fundamental execution options: place the order "at market" or "at limit." The market ones are orders of purchase or sale at the best available quote. On the other hand the limit orders are not immediately executed since they are placed at a quote which is less favorable than the best quote, e.g. the second level bid/ask order. The schematic representation of an LOB reflects the local decisions and interactions between thousands of investors, it generates a high dimensional dynamic and complex system, which then results in the pricing of assets. Here we provide insights into this highly dynamic LOB by dimension reduction techniques in combination with generalized impulse response analysis. In order to do so the raw HF data has to be prepared.
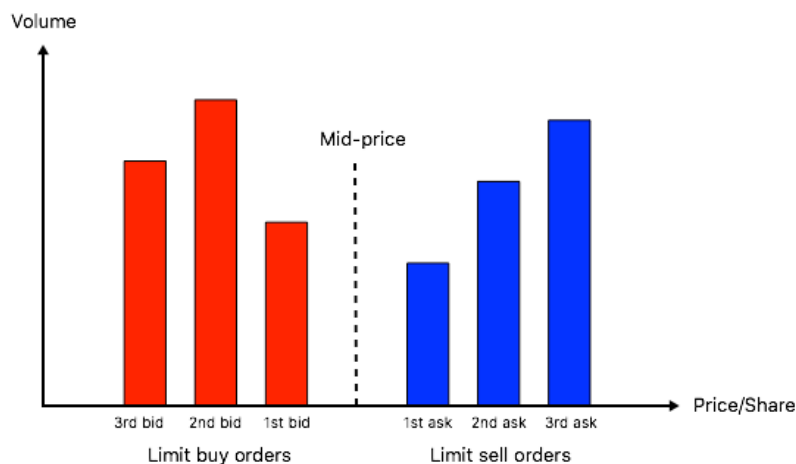
However empirical evidence on the actual market impact of limit order placements across stocks is virtually not existent, many questions of interest to regulators and traders are unsolved: i) How does the order flows interact with price dynamics? ii) How to measure the impact of an incoming limit order quantitatively? iii) Are the impacts on return responds to incoming ask and bid limit orders widely symmetric? iv) If not symmetric, how does the heterogeneous market impact caused by bid and ask order for various stocks affect the whole market? To address the arising questions, in this paper we provide a comprehensive study on the interaction between price and bid/ask order sizes in terms of vast network system. LOB provides a more complicated scenario that inspires us to construct a high-dimensional object using both price and several levels of depth of order sizes with historical order flow. This may help us to understand how information is impounded into price. The underlying assumption is that there exists a sparse representation of the data. The orders posted on the selected order levels that induce significant price impact would be treated as price drivers. In this way, investors' decision-making can be addressed by making trading price driven by order flows. Of particular interest is vast directed network analysis based on the constructed high-dimensional object. The motivation to construct a network of LOB stems from both the lack of theoretical setup, and the market impacts caused by size imbalance.

To do so the VAR model is without doubt one of the most useful tools that allows us to capture in a simple fashion their dynamic evolution. However it imposes challenges of high dimensionality when we incorporating a variety of time series, particularly where the vector observed at each time is high dimensional relative to the time period. Researchers have developed various penalized estimators to filter out less relevant variables, such as the Lasso estimator of Tibshirani (1996a), SCAD of Fan and Li (2001), adaptive Lasso of Zou (2006), elastic net estimator of Zou and Hastie (2005), Dantzig selector of Candes and Tao (2007). This paper focuses on building up the network connectivity, where the connectedness table is directly derived from generalized impulse response function. There has been a large literature discussing acquiring sparse VAR estimation by adding different penalty terms. For instance, Negahban and Wainwright (2011) imposed sparse dependence assumption on the transition matrix of VAR model and studied the theoretical properties. Kock and Callot (2015a) discussed theoretical properties of LASSO and adaptive Lasso in VAR model that may reveal the correct sparsity pattern asymptotically. Basu et al. (2015) investigated theoretical properties of Lasso-type estimators for high-dimensional Gaussian processes. Wu and Wu (2016) studied the systematic theory for high-dimensional linear models with correlated errors. The Lasso-type estimators penalize the regression coefficients with the model size via a shrinkage procedure. Belloni et al. (2012) and Belloni et al. (2013) studied the post-model selection estimator that apply OLS to the first-step penalized estimators to alleviate shrinkage bias.

Diebold and Yılmaz (2014) proposed connectedness measures built from generalized

forecast error variance decomposition (GFEVD) based on VAR systems, where the GFEVD is developed by Pesaran and Shin (1998) and Koop et al. (1996) with an intrinsic appeal to order-invariance. However the contributions of shares of forecast error variation in various locations do not add to unity, and it is restricted to Gaussian innovations. To solve this, we use the new GFEVD has been recently proposed by Lanne and Nyberg (2016). They proposed a simple modification that, the proportions of the relative contributions to the h-period impact of the shocks sum to unity. The LN-GFEVD is thus economic interpretable, and can be implemented to both Gaussian and non-Gaussian models. This paper contributes to network estimation through penalized VAR approach. To keep the sparsity structure of VAR estimation, we apply bootstrap-based method rather than moving-average (MA) transformation which is often done in fixed dimensional cases. Besides, bootstrapped GFEVD relies neither on the ordering of the variables nor on the distribution of the innovations. Hence, the new connectedness table is obtained by bootstrapping from the estimated innovations, using generalized impulse response function. The directed connectedness "from" and "to" are associated with the forecast error variation for specific order book across various stocks when the arising shocks transmit from one stock to the others.

We progress by focusing on the models that capture the dynamics of LOB and their influence over time. We aim to mimic the LOB trading mechanism to find out significant market impact. Our primary finding is that order imbalance general exists across stocks, bootstrapped market impacts can be quantified. With our methodology, we identify the significant market impact caused by the arrival of a large limit order, and several robust risk transmission channels and risk clusters. The financial institutions are connected more closely compared with the firms come from other industry. The estimation results of LN-network are also compared and studied.

The rest of the paper is organized as follows. Section 1.2 introduces NASDAQ LOB market and the non-synchronous LOB data, we then elaborate the data preparation process. In Section 1.3 we present the theoretical framework for high-dimensional VAR estimation, and discuss the connectedness estimator based on our setting. Section 1.4 reports the novel evidence on market impact using bootstrapped estimator. Section 1.5 presents the robust risk channels and risk clusters. The network study will be illustrated in Section 1.6. Section 1.7 concludes, while more technical details are relegated to the Appendix.

1.2 Description of the Market and Data Preparation

1.2.1 NASDAQ Limit Order Book Market

In this paper, our sample consists of intraday trading data for selected NASDAQ stocks for the sample period spanning 1st, June 2016 to 30th, July of 2016. These data come from the LOBSTER academic data, which is powered by NASDAQ's historical TotalView using very detailed event information. The industry breakdown of NASDAQ market is technology of 57.18%, consumer service of 21.27% and Health Care of 12.38% (as of 01.09.2016).

The sample is stratified by market capitalization and industry sector. We consider a sample portfolio with 9 assets listed in Table 1.1, together with their first three levels of limit orders, which attracts the majority of trading activity, therefore becoming our research interest.

Industry	Stock	Company	MktCap (billion $)
Technology	IBM	International Business Machines Corp.	171.72
	MSFT	Microsoft Corporation	499.35
	T	AT&T Inc.	257.53
Healthcare	JNJ	Johnson & Johnson	328.91
	PFE	Pfizer Inc.	206.69
	MRK	Merck & Co. Inc.	181.56
Finance	JPM	JP Morgan Chase & Co.	326.04
	WFC	Wells Fargo & Company	293.39
	C	Citigroup Inc.	168.06

Table 1.1. Sample data. MktCap is the market capitalization by Feb 25th, 2017.

We present the summary statistics of sample dataset in Table 1.2. The data is collected for the normal trading day, which runs from 9:30 a.m to 4 p.m ET. To avoid erratic effects during the market opening and closure, our sample period covers only the continuous trading periods between 9:45 and 16:00.

The basic structure of LOB is shown in Figure 1.1. The sample file has one time-stamped record for every order entered for each stock throughout the trading day. Trades are time stamped up to the nanosecond and signed to indicate whether they were initiated by a buyer or seller by the "Direction" ticker. The ticker of "Event Type" indicates the trading type, for example, 1: Submission of a new limit, 2: Cancellation (partial deletion of a limit order), 3: Deletion (total deletion of a limit order) etc. Another important feature of this dataset is that each quote has been associated with trading

	NumObs ($*10^3$)	AvgTrd ($*10^3$)	AvgAP1 (in \$)	AvgBP1 (in \$)	AvgAS1 (100 shrs)
IBM	118.25	5.82	153.07	153.04	1.92
MSFT	584.55	25.91	52.28	52.26	24.19
T	223.45	6.67	38.75	38.74	36.36
JNJ	172.77	8.17	113.99	113.98	4.11
PFE	427.51	12.49	34.83	34.82	41.96
MRK	188.84	5.82	56.70	56.68	7.43
MDT	173.96	10.30	82.45	82.43	4.22
JPM	414.35	11.49	65.48	65.46	9.47
WFC	275.29	10.91	50.90	50.89	18.02
BAC	308.53	7.55	14.87	14.86	359.94
C	472.90	12.19	46.82	46.81	14.19
	AvgBS1 (100 shrs)	AvgAS2 (100 shrs)	AvgBS2 (100 shrs)	AvgAS3 (100 shrs)	AvgBS3 (100 shrs)
IBM	2.17	1.95	2.26	2.09	2.26
MSFT	24.53	28.12	31.06	33.90	35.37
T	33.76	43.63	41.96	55.53	63.67
JNJ	3.62	5.86	4.44	7.74	4.90
PFE	42.29	48.07	48.09	50.94	55.68
MRK	7.36	14.34	11.30	24.20	13.87
MDT	3.99	5.12	4.94	6.08	5.79
JPM	9.45	13.10	11.82	17.41	15.09
WFC	17.01	20.68	17.72	23.58	19.05
BAC	341.64	456.38	489.76	373.91	345.64
C	12.97	18.58	16.48	22.23	19.60

Table 1.2. Summary statistics of selected stocks. $NumObs$ denotes the average number of observation. $AvgTrd$ is the average number of execution trades of a limit order. $AvgAP1$ gives the average ask price for the first order, and $AvgAS1$ represents the corresponding ask size.

information and limit order book. To be more specific, the k-th row in the "message" file (upper panel of Figure 1.1) describes the limit order event causing the change in the limit order book from line $k-1$ to line k in the "orderbook" file (lower panel).

Time (sec)	Event Type	Order ID	Size	Price	Direction
⋮	⋮	⋮	⋮	⋮	⋮
34713.685155243	1	206833312	100	118600	-1
34714.133632201	3	206833312	100	118600	-1
⋮	⋮	⋮	⋮	⋮	⋮

Ask Price 1	Ask Size 1	Bid Price 1	Bid Size 1	Ask Price 2	Ask Size 2	Bid Price 2	Bid Size 2	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1186600	9484	118500	8800	118700	22700	118400	14930	...
1186600	9384	118500	8800	118700	22700	118400	14930	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Figure 1.1. Structure of LOBSTER data

The main challenge in dealing with HFT data is the presence of microstructure noise arising from market frictions, where the noise-induced bias at very high sampling frequencies contaminates the observed price. Whereas infrequent sampling frequency leads to imprecise estimates, optimal sampling frequency is needed to acquire bias-variance tradeoff, see Bandi and Russell (2006), Aït-Sahalia et al. (2005), Bandi and Russell (2008). In particular, several main approaches to improve the RV estimator include the preaveraging estimator of Jacod et al. (2009), the realized kernel estimator of Barndorff-Nielsen et al. (2008), the two scales estimator of Zhang et al. (2005) and multiscale estimator of Zhang et al. (2006) and Zhang (2011). The aim of this paper is to give more insights to the covariance structure in limit order markets, in order to match the size of the limit orders with the corresponding bid/ask prices, we implement pre-averaging approach introduced in subsection 1.2.2. Subsection 1.2.3 summarizes the data preparation procedure for high-dimensional statistical setting.

1.2.2 Pre-averaging estimation

Suppose that we observe non-synchronous noisy data Y_t following,

$$Y_t = X_t + \varepsilon_t, \quad t \geq 0 \tag{1.1}$$

with efficient log price X_t is latent. The error term ε_t represents microstructure noise and is assumed to be independent and identically distributed with

$$\mathsf{E}(\varepsilon_t) = 0, \quad \mathsf{E}\left(\varepsilon_t^2\right) = \psi \tag{1.2}$$

The price process X_t follows a semi-martingale form, Delbaen and Schachermayer (1994),

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s \tag{1.3}$$

where $(a_s)_{s \geq 0}$ is a càdlàg drift process, $(\sigma_s)_{s \geq 0}$ is an adapted càdlàg volatility process, $(W_s)_{s \geq 0}$ is a Brownian motion. In addition, we assume X_t and ε_t are independent, i.e.

$$\mathsf{E}(\varepsilon_t \mid X) = 0 \tag{1.4}$$

If one can only observe Y_i^n at discrete times t, i indexes the time points with interval length Δ_n, the returns $\Delta_i^n Y$ is thus defined as,

$$Y_i^n = Y_{i\Delta_n}, \quad \Delta_i^n Y = Y_i^n - Y_{i-1}^n, \quad i = 1, \dots, n \tag{1.5}$$

A pre-averaging is conducted to alleviate microstructure noise and solve non-synchronicity, we follow the notations originally used by Jacod et al. (2009). The basic idea is to construct smoothing functions to diminish the impact of the noise induced by ε_t. Specifically, there is a sequence of integers denoted as k_n which satisfies,

$$\exists \theta > 0, \quad k_n \sqrt{\Delta_n} = \theta + o\left(\Delta_n^{\frac{1}{4}}\right) \tag{1.6}$$

and a continuous weight function $g : [0,1] \mapsto \mathbb{R}$. g is piecewise C^1 with a piecewise derivative g', $g(0) = g(1) = 0$, and $\int_0^1 g^2(s) ds > 0$. Furthermore, the following real-valued numbers and functions are associated with function g on $\mathbb{R}_+$,

$$\psi_1 = \int_0^1 \{g'(u)\}^2 du, \quad \psi_2 = \int_0^1 \{g(u)\}^2 du$$

$$\Phi_1(s) = \int_s^1 g'(u)g'(u-s) du, \quad \Phi_2(s) = \int_s^1 g(u)g(u-s) du$$

$$\Phi_{ij} = \int_0^1 \Phi_i(s)\Phi_j(s) du, \quad i,j = 1,2, \quad u \in [0,1] \tag{1.7}$$

Here we choose $g(x) = x \wedge (1-x)$, as in Podolskij et al. (2009), Christensen et al. (2010b)

and Hautsch and Podolskij (2013). Therefore we have

$$
\begin{aligned}
\psi_1 &= 1, \quad \psi_2 = \frac{1}{12}, \quad \Phi_{11} = \frac{1}{6} \\
\Phi_{12} &= \frac{1}{96}, \quad \Phi_{22} = \frac{151}{80640}
\end{aligned}
\tag{1.8}
$$

The pre-averaged returns $\overline{Y}_i^n$ associated with the weight function g are given as,

$$
\begin{aligned}
\overline{Y}_i^n &= \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right) \Delta_{i+j}^n Y \\
&= -\sum_{j=0}^{k_n-1} \left\{ g\left(\frac{j+1}{k_n}\right) - g\left(\frac{j}{k_n}\right) \right\} Y_{i+j}^n, \quad i = 0, \ldots, n - k_n + 1
\end{aligned}
\tag{1.9}
$$

The window size k_n defined in equation (1.6) is chosen of $\mathcal{O}\left(\sqrt{\frac{1}{\Delta_n}}\right)$, balance the noise $\overline{\varepsilon}_i^n = \mathcal{O}_p\left(\sqrt{\frac{1}{k_n}}\right)$ and the efficient price $\overline{X}_i^n = \mathcal{O}_p\left(\sqrt{k_n \Delta_n}\right)$.

1.2.3 Data Preparation for High-dimensional Statistical Setting

This section is firstly devoted to the construction of covariance estimates. It follows that the sizes of the book at which limit orders are submitted are driving the price. In contrast to a moderate interval for price to remove the microstructure noise, the interval for the size should be smaller enough to capture the large orders submitted by the market trader, and yet detect the price-size interconnection.

Inspired by these, we propose size intensity $\tilde{S}_{t_j}$ as

$$
\tilde{S}_{t_j} = S_{t_j}(t_{j+1} - t_j)
\tag{1.10}
$$

where t_j denotes the time stamp of jth LOB, and S_{t_j} is the corresponding tick size. By size intensity can be summed up over a given time interval and therefore, matched with returns over a moderate time interval. We can provide valuable information and access to a new angle on the covariance structure.

In the following we shall illustrate how to explicitly proceed data preparation subject to the given LOB. For ease of illustration, it can be divided into four steps,

1. Set equally-spaced time intervals starting at time T_0

$$
T_0 + k\Delta T, \quad k = 0, 1, 2, \ldots, K
$$

2. Define the price and size at time $T_0 + k\Delta T$ as

$$
\begin{aligned}
\tilde{P}_{T_0+k\Delta T} &= P_{t_m}, \; t_m = \max\{t_j; \; t_j \leq T_0 + k\Delta T\} \\
\tilde{S}_{T_0+k\Delta T} &= \sum_{T_0+(k-1)\Delta T \leq t_j \leq T_0+k\Delta T} S_{t_j}(t_{j+1} - t_j)
\end{aligned}
$$

3. The changes of the log values are

$$
\begin{aligned}
\Delta p_{T_0+k\Delta T} &= \log \tilde{P}_{T_0+k\Delta T} - \log \tilde{P}_{T_0+(k-1)\Delta T} \\
\Delta s_{T_0+k\Delta T} &= \log \tilde{S}_{T_0+k\Delta T} - \log \tilde{S}_{T_0+(k-1)\Delta T}
\end{aligned}
$$

4. Pre-averaging both $\Delta p_{T_0+k\Delta T}$ and $\Delta s_{T_0+k\Delta T}$ by

$$
\begin{aligned}
\Delta \tilde{p}_{T_0+k\Delta T} &= \sum_{j=0}^{J} g_j \Delta p_{T_0+j\Delta T} \\
\Delta \tilde{s}_{T_0+k\Delta T} &= \sum_{j=0}^{J} g_j \Delta s_{T_0+j\Delta T}
\end{aligned}
$$

where $g_j \geq 0$ and $\sum_{j=0}^{J} g_j = 1$.

Cleaning data in this way can help alleviate microstructure noise, match the price to the size in a moderate interval and solve the problem of non-synchronicity. In our setting based on a 3-level order book, we take the mid price $\Delta \tilde{p}_t$ on the first level, the bid and ask sizes $(\Delta \tilde{s}_t^a, \Delta \tilde{s}_t^b)$ on the first 3 levels, i.e.,

$$
y_t^{(n)\top} = [\Delta \tilde{p}_t^{(n)}, \quad \Delta \tilde{s}_t^{a1(n)}, \Delta \tilde{s}_t^{a2(n)}, \Delta \tilde{s}_t^{a3(n)}, \Delta \tilde{n}_t^{b1(n)}, \Delta \tilde{s}_t^{b2(n)}, \Delta \tilde{s}_t^{b3(n)}] \tag{1.11}
$$

where $\Delta \tilde{s}_t^{aj(n)}$ stands for the preaveraged jth level of ask order for stock n, whereas $\Delta \tilde{s}_t^{bj(n)}$ stands for the jth level of bid order for stock n. For each stock we divide the trading period into 1-minute intervals and pre-average both $\Delta \tilde{p}_t^{(n)}$, $\Delta \tilde{s}_t^{bj(n)}$ and $\Delta \tilde{s}_t^{aj(n)}$ to reduce microstructure noise over 15-min, yielding 375 observations per day.

Note that a critical assumption imposed to ensure the consistency of estimator is the observations are weakly dependence. Here we define the large vector $Y_t^\top$ to estimate as

$$
Y_t^\top = [y_t^{(1)\top}, y_t^{(2)\top}, \ldots, y_t^{(N)\top}] \tag{1.12}
$$

by stacking the vector X_t^T for different N stocks together.

1.3 Methodology

1.3.1 High-dimensional VAR estimation

Statistically, the large VAR model facilitates consistent estimation and better finite-sample performance. Economically, estimation results derived from sparse assumption help explain the economic intuition. By incorporating the lags terms in the penalized VAR model, we aim to show the "sluggished" price adjustments caused by limit orders.

The standard $\text{VAR}(p)$ model, Lütkepohl (2005) is,

$$
\begin{aligned}
Y_t &= A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + u_t \\
&= (A_1, A_2, \ldots, A_p) \left(Y_{t-1}^\top, Y_{t-2}^\top, \ldots, Y_{t-p}^\top \right)^\top + u_t
\end{aligned} \tag{1.13}
$$

where $Y_t = (y_{1t}, y_{2t}, \ldots, y_{Kt})^\top \in \mathbb{R}^K$ is a random vector, $t = 1, \ldots, T$. A_i are fixed $(K \times K)$ coefficient matrices. p is the lag and $u_t = (u_{1t}, u_{2t}, \ldots, u_{Kt})^\top \in \mathbb{R}^K$ is the i.i.d innovation process. In our LOB setting, dimension of $K = 7N$ with N is the number of stocks in the portfolio.

Assumption 1. Assume (2.1) satisfies that,

1. The roots of $|I_K - \sum_{j=1}^p A_j z^j| = 0$ lie outside unit circle.

2. u_t are i.i.d innovations with fourth moment condition.

3. $\|\Sigma_u\|_2 < \infty$ and $\sum_{j=1}^p \|A_j\|_2 < \infty$.

In practice, the coefficients $A_1, \ldots, A_p$ are unknown and has to be estimated from $\{Y_t\}_{t=1}^T$. The multiple time series data will be partitioned into sample and pre-sample values to facilitate the following thoughts. Define,

$$
\begin{aligned}
Y &= (Y_1, Y_2, \ldots, Y_T) & A &= (A_1, A_2, \ldots, A_p) \\
Z_t &= (y_t, y_{t+1}, \ldots, y_{t-p+1})^\top & Z &= (Z_0, Z_1, \ldots, Z_{T-1})
\end{aligned} \tag{1.14}
$$

Then equation (2.1) reads,

$$
Y = AZ + U \tag{1.15}
$$

with $U = (u_1, u_2, \ldots, u_T)$. The compact form (1.15) is equivalent to

$$
\mathbf{y} = (Z^\top \otimes I_K)\beta + \mathbf{u} = \mathbf{x}\beta + \mathbf{u} \tag{1.16}
$$

where the length of the parameter vector β is pK^2, the number of observations is KT. In practice, the ration $\frac{Kp}{T}$ could be large due to high dimensionality, which deteriorates the accuracy of final estimate. Worse still, if $Kp > T$, the coefficients to be estimated increases quadratically in terms of the number of lags p, therefore the model cannot identified with traditional methods such as OLS. Therefore we introduce variable selection techniques, such as LASSO, to estimate the model because only a subset of the pK^2 parameters might be non-zero. For multiple time series data, especially high dimensional time series, it is preferred to use elastic net approach rather than pure Lasso to remedy potentially strong correlation among regressors. Besides, under normal assumption of error term, the upper bound of estimated error is positively correlated in $\frac{\log(K^2 p)}{T}$, part of oracle inequality. The methodologies introduced in the proceeding paragraph are of great importance in the sense that the true underlying model has a sparse representation.

The penalized VAR estimates β by minimizing the objective function,

$$\arg\min_{\beta} \left(\|\mathbf{y} - \mathbf{x}\beta\|_2^2 + \alpha_{1,T}\|\beta\|_1 + \alpha_{2,T}\|\beta\|_2^2 \right) \tag{1.17}$$

which is equivalent to,

$$\arg\min_{A_1, A_2, \ldots, A_P} \sum_{t=1}^{T} \|Y_t - \sum_{j=1}^{P} A_j Y_{t-j}\|_2^2 + \alpha_{1,T} \sum_{j=1}^{P} \|vec(A_j)\|_1 + \alpha_{2,T} \sum_{j=1}^{P} \|vec(A_j)\|_2^2 \tag{1.18}$$

where A_j is the $(K \times K)$ coefficient matrices of interest. $\alpha_{1,T}$ and $\alpha_{2,T}$ are the penalty parameters. Note that the notation $\|M\|_p$ depends on whether M is a vector or a matrix. To avoid confusion, we use $vec(M)$ here to tranform the object within $\|\|_p$ into a vector.

We choose a sequence of decreasing positive numbers $\alpha_{1,T}$ and $\alpha_{2,T}$ to control the regularization. In the case of regularization parameter is large, setting it too high will throw away useful information, whereas the estimated graph is not sparse when the α_T is small. To balance the sparsity and estimation accuracy, we choose a moderately small tuning parameter using Bayesian information criterion (BIC). In addition, we apply OLS post-model selection estimator to the first-step penalized estimator (1.17) or (1.18) to reduce shrinkage bias and ensure better model model performance.

1.3.2 Network Construction

Following Pesaran and Shin (1998) and Lanne and Nyberg (2016), they assume shocks hitting only one equation at a time rather than all the shocks at time t. The generalized impulse response function (GIRF) is denoted as GI, when there is a shock δ_{jt} hits on

j-th equation of y_t at horizon l,

$$GI(l, \delta_{jt}, \omega_{t-1}) = \mathsf{E}(y_{t+l}|u_{jt} = \delta_{jt}, \omega_{t-1}) - \mathsf{E}(y_{t+l}|\omega_{t-1}) \qquad (1.19)$$

$\mathsf{E}(y_{t+1}|u_{jt} = \delta_{jt}, \omega_{t-1})$ represents the expectation conditional on the history ω_{t-1} and a fixed value of j-th shock on time t. ω_{t-1} consists of the information used to compute the conditional expectations based on equation (2.1).

The LN-GFEVD denoted as $\lambda_{ij,\omega_{t-1}}(h)$ is defined by j-th shock hitting i-th variable at time t,

$$\lambda_{ij,\omega_{t-1}}(h) = \frac{\sum_{l=0}^{h} GI(l, \delta_{jt}, \omega_{t-1})_i^2}{\sum_{j=1}^{K} \sum_{l=0}^{h} GI(l, \delta_{jt}, \omega_{t-1})_i^2}, \quad i, j = 1, \ldots, K \qquad (1.20)$$

where h is the horizon, ω_{t-1} refers to the history. Therefore $\lambda_{ij,\omega_{t-1}}(h) \in [0, 1]$, measuring the relative contribution of a shock δ_{jt} to the j-th equation in relation to the total impact of all K shocks on the i-th variable in y_t after h periods, and these contributions sum to unity.

To measure the persistent effect of a shock on the behaviour of a series, we aim to acquire the population connectedness table. Compared with the connectedness table in Diebold and Yılmaz (2014), this simple modification of GFEVD allows the proportions of the relative contributions to the h-period impact of the shocks sum to unity. The LN-GFEVD is thus economic interpretable, and can be implemented to both Gaussian and non-Gaussian models. Upon the penalized VAR estimation in equation (1.17) and (1.18), we acquire the sparsity structure that filters out less relevant variables. Instead of transforming into MA process, which is often done in fixed dimensional cases, we apply bootstrap-based method to produce Table 1.3. Besides, bootstrapped GFEVD relies neither on the ordering of the variables nor on the distribution of the innovations.

The details for computation steps can be found in Appendix A.1. In particular, the numerical techniques for conditional mean forecast from nonlinear models for more than one period ahead are implemented in this paper, we use bootstrap to calculate $GI(l, \delta_{jt}, \omega_{t-1})$, see more details in Terasvirta et al. (2010).

We then have the directional connectedness "from" and "to" associated with the forecast error variation $\lambda_{ij}^b(h)$ for a specific order book across various stock when the arising shocks transmit from one stock to the others. These two connectedness estimators can be obtained by adding up the row or column elements. Hence the pairwise directional

	x_1	x_2	$\dots$	x_K	From others
x_1	$\lambda_{11}^b(h)$	$\lambda_{12}^b(h)$	$\dots$	$\lambda_{1K}^b(h)$	$\sum_{j=1}^{K}\lambda_{1j}^b(h), j \neq 1$
x_2	$\lambda_{21}^b(h)$	$\lambda_{22}^b(h)$	$\dots$	$\lambda_{2K}^b(h)$	$\sum_{j=1}^{K}\lambda_{2j}^b(h), j \neq 2$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_K	$\lambda_{K1}^b(h)$	$\lambda_{K2}^b(h)$	$\dots$	$\lambda_{KK}^b(h)$	$\sum_{j=1}^{K}\lambda_{Kj}^b(h), j \neq K$
To others	$\sum_{i=1}^{K}\lambda_{i1}^b(h)$ $i \neq 1$	$\sum_{i=1}^{K}\lambda_{i2}^b(h)$ $i \neq 2$	$\dots$	$\sum_{i=1}^{K}\lambda_{iK}^b(h)$ $i \neq K$	$\frac{1}{K}\sum_{i=1,j=1}^{K}\lambda_{ij}^b(h)$ $i \neq j$

Table 1.3. Connectedness table of interest, estimated by bootstrap-based methods.

connectedness from j to i can be written as,

$$C_{i \leftarrow j}^{H} = \lambda_{ij}^b(h) \tag{1.21}$$

Furthermore, the total directional connectedness "from" $C_{i \leftarrow \cdot}$ (others to i) is defined as

$$C_{i \leftarrow \bullet} = \sum_{j=1}^{K}\lambda_{ij}^b(h), i \neq j \tag{1.22}$$

and the total directional connectedness "to" $C_{\cdot \leftarrow j}$ (j to others) is defined as

$$C_{\bullet \leftarrow j} = \sum_{i=1}^{K}\lambda_{ij}^b(h), i \neq j \tag{1.23}$$

The corresponding net total directional connectedness is given by

$$C_i = C_{to} - C_{from} = C_{\bullet \leftarrow i} - C_{i \leftarrow \bullet} \tag{1.24}$$

1.3.3 Prescreening - Graphical Model

In high dimensional setting, the denominator of LN-GFEVD connectedness measure might be unnecessarily large due to accumulated noise caused by the large amount of irrelevant variables. Therefore we introduce Graphical Lasso as prescreening step to reduce the number of variables summed up in the denominator. Graphical Lasso discussed in this subsection helps reveal the stocks, whose prices and sizes have significant correlation with others. This allows us to construct a reduced estimation framework for the network (graph) between prices and limit orders.

Here the thresholding technique is implemented for consistent high dimensional covari-

ance estimation by bounding the small eigenvalues from zero. At first the thresholded covariance matrix $\hat{\Sigma}_0$ is defined by,

$$\hat{\Sigma}_0 = \{S_{ij}\mathbf{I}_{|S_{ij}|>u_0}\}; \qquad S = \frac{1}{T}\sum_{t=1}^{T}\bar{X}_t\bar{X}_t^\top$$

where matrix S is the sample covariance matrix. $\bar{X}_t^\top$ is the large vector of interest in equation (1.12). However, in finite sample, some of its eigenvalues could be close to or smaller than zero, thus we introduce a positive-definitization-version $\hat{\Sigma}_n$,

$$\hat{\Sigma}_n = \sum_{j=1}^{m}(\hat{\eta}_j \vee \lambda)\hat{q}_j\hat{q}_j^\top; \qquad \hat{\Sigma}_0 = \sum_{j=1}^{m}\hat{\eta}_j\hat{q}_j\hat{q}_j^\top \tag{1.25}$$

where $\lambda \in \Lambda$ is the sequence to control the regularization.

In graphical model, the edges $\hat{\mathcal{E}}$ are identified by non-zero partial correlations we introduced in the previous section. Thus far we have the basic setup for the undirected connectedness measure.

The details for the graphical model are summerized in Appendix A.2. In particular, the partial correlation between two nodes of $X^{(j)}$ and $X^{(k)}$ given $X^{(\mathcal{V}\setminus\{j,k\})}$ is defined as,

$$\rho_{jk|\mathcal{V}\setminus\{j,k\}} = -\frac{\Sigma_{jk}^{-1}}{\sqrt{\Sigma_{jj}^{-1}\Sigma_{kk}^{-1}}} \tag{1.26}$$

where Σ^{-1} is the sparse concentration matrix estimated by graphical Lasso via the penalized concentration matrix given by,

$$\hat{\Sigma}_n^{-1}(\lambda) = \underset{\Psi \succ 0}{\arg\min}\{tr(\Psi\hat{\Sigma}_n) - \log\det(\Psi) + \lambda|\Psi|_1\} \tag{1.27}$$

where $\hat{\Sigma}_n$ is the thresholded estimator for covariance matrix. $|\Psi|_1$ denotes the sum of the absolute values of Ψ, and λ is a tuning parameter controlling the amount of l_1 shrinkage.

Consistency of the estimator is proved by Chen et al. (2013). Therefore the estimated graph model is,

$$\hat{E}(\lambda) = \{(j,k) \in \mathcal{V} \times \mathcal{V}; \ \hat{\Sigma}_{jk}^{-1}(\lambda) \neq 0\} \tag{1.28}$$

In high dimensional setting, the p-dimensional covariance matrix Σ is assumed to satisfy,

$$\mathcal{G}_q(M) = \left\{ \Sigma | \max_{j \leq p} \sigma_{jj} \leq 1; \quad \max_{1 \leq k \leq p} \sum_{j=1}^{p} |\sigma_{jk}|^q \leq \tilde{M} \right\} \tag{1.29}$$

where $0 \leq q < 1$ to ensure Σ is sparse. $\tilde{M}$ is constant or $C_0(p)$. By setting the small off-diagonal elements to zero, we obtain the sparse covariance matrix $\hat{\Sigma}_n$, and therefore, sparse precision matrix $\hat{\Sigma}_n^{-1}$ by graphical lasso based on eqution (1.27).

1.4 Bootstrapped Market Impact

In this section we aim to characterize the dynamic links between uncertainty, price and size (trading volume) across stocks in a high dimensional penalized VAR system.

When a large market order to buy a stock arrives, the limit orders with the lowest ask prices will automatically execute, this causes a temporary market impact. We will answer the following questions proposed in the very beginning, i) How does the order flows interact with price dynamics? ii) How to measure the impact of an incoming limit order quantitatively? iii) Are the impacts on return responds to incoming ask and bid limit orders widely symmetric? iv) If not symmetric, how does the heterogeneous market impact caused by bid and ask order for various stocks affect the whole market?

1.4.1 Price and Order Flows

To check whether shocks of order flows affect price dynamics and whether the impacts identified by our model are temporary or robust over time, we resort to generalized impulse response analysis based on bootstrapped estimation of GI.

Here we follow the VAR literature and interpret the shock as the market shock. In particular, the market impacts of the order flows to price factors are quantified by equation (1.19), when the shock δ_{jt} is treated as one of the size factors ($\Delta \tilde{s}_t^{a1(i)}$, $\Delta \tilde{s}_t^{a2(i)}$, $\Delta \tilde{s}_t^{a3(i)}$, $\Delta \tilde{s}_t^{b1(i)}$, $\Delta \tilde{s}_t^{b2(i)}$, $\Delta \tilde{s}_t^{b3(i)}$) hitting on the jth equation for stock i. The response of price factor $\Delta \tilde{p}_t^{(i)}$ is given by,

$$GI(l, \delta_{jt}, \omega_{t-1}) = \mathsf{E}(y_{t+l} | u_{jt} = \delta_{jt}, \omega_{t-1}) - \mathsf{E}(y_{t+l} | \omega_{t-1})$$

where $\mathsf{E}(y_{t+1} | u_{jt} = \delta_{jt}, \omega_{t-1})$ represents the expectation conditional on the history ω_{t-1} and a fixed value of j-th shock on time t. ω_{t-1} consists of the information used to compute the conditional expectations based on equation (2.1).
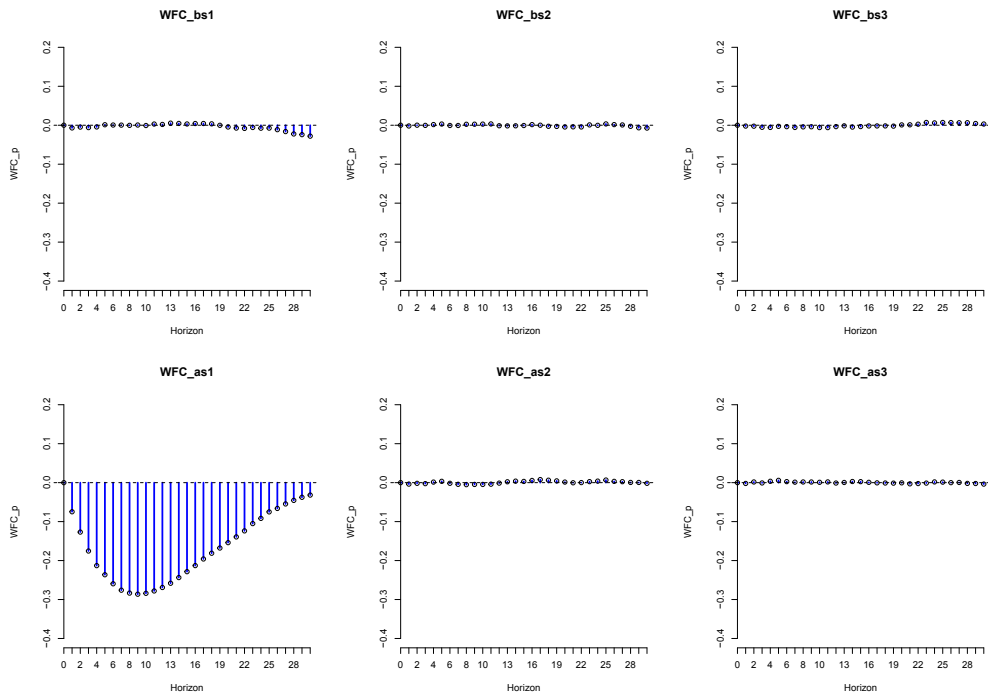
Figure 1.2. The bootstrapped market impact of WFC (Wells Fargo) on 25th of July, 2016.

hfhd_marketimpact

With a moderate sparse structure selected by BIC, we could successfully identify the existence of significant market impact. Figure 1.2 depicts the estimated market impacts of limit orders for a single trading day for WFC (Wells Fargo). We observe a negative correlation between the magnitude of WFC's first-order ask size factor and its price factor. For orders posted deeper in the book, the market impacts are not as significant as the first-level. It is normal for financial market in the sense that the investors will start marking down their bid price when there is a wave of sell orders coming into the order book. As expected, the price (average of bid and ask quotes) factor tends to decrease significantly after the arrival of a large ask limit order. The impact can last for almost 32 minutes before the price shifts back, this gives the HF investors enough time of reaction to arbitrage opportunities.
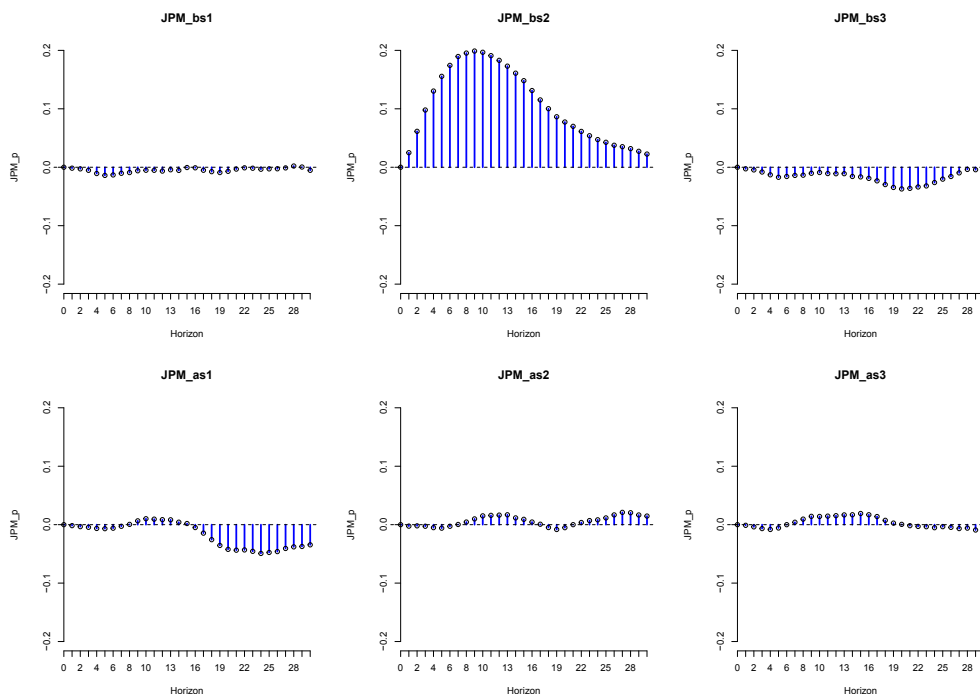
Figure 1.3. The bootstrapped market impact of C (Citigroup) on 1st of June, 2016.

hfhd_marketimpact

Figure 1.3 and 1.4 show the market impacts from second-level bid factor for Citigroup on two separate trading days. This implies the positive pile-on effect where larger bid order may further perpetuating a price increase, the orders may not necessarily set at the current market price of the stock (i.e. they are not market orders, they are limit orders). The estimated market impact lasts for almost 20 minutes, the price goes down
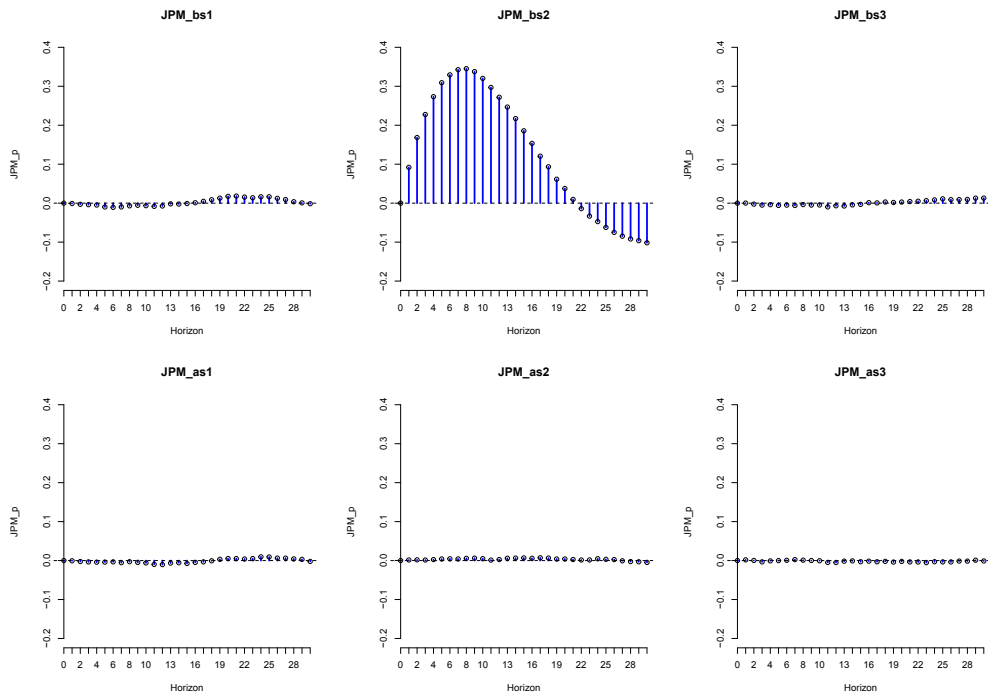
Figure 1.4. The bootstrapped market impact of C (Citigroup) on 20th of June, 2016.

 hfhd__marketimpact

after 10 minutes because the market investors may sell trades picking up the posted volume or by cancellations on the bid side.

1.4.2 Order Imbalance

As discussed above, the bid/ask sizes of the limit order book will be affected when market investors submit limit orders to buy/sell. Hence the order imbalance, that caused by difference between the bid and ask sizes, would allow us to further understand the direction of the upcoming price changes. To quantify the investors' trading intention, we calculate the relative weights for different levels of depth of limit order with different orders of lag. In this way we will investigate if there is a link between net size (ask trading volume minus bid trading volume) and price changes, based on our model setup and undirected network estimation.

Figure 1.5 shows the net market impact of size factors to the price factor for three continuous trading days for JPM (JP Morgan). Generally, when there is a strong sell pressure induced by huge sell volume queued on the ask side, all three levels of net ask size factors have negative impacts affect the price factor in a negative pattern and the deeper orders posted in the book, the weaker impacts they have.

According to Figure 1.6, the history information shows which stock is stable in the price-size relations and thus robust for statistical arbitrage, see Hautsch and Huang (2012). All these three financial institutions exhibit robust impacts from net sizes, with WFC performs best among them. Based on our methodology, the 'gap' between the impacts from ask and bid size factors is quantified and could serve as a strong signal for market investors.

1.5 Transmission Channels of Risk

Our model has implied that in an LOB market, the huge sell/buy volume queued on the ask/bid side could induce strong sell/buy pressure on the market and therefore changing the price. In this section, our aim is to gain some insights into the details of the price formation by estimating the cross-sectional dynamics of returns (price factor).

1.5.1 Risk Channel & Risk Cluster

To check the shock persistence and asymmetric effects of shocks derived from the generalized impulse response function, we first visualize the cross-sectional impact of returns
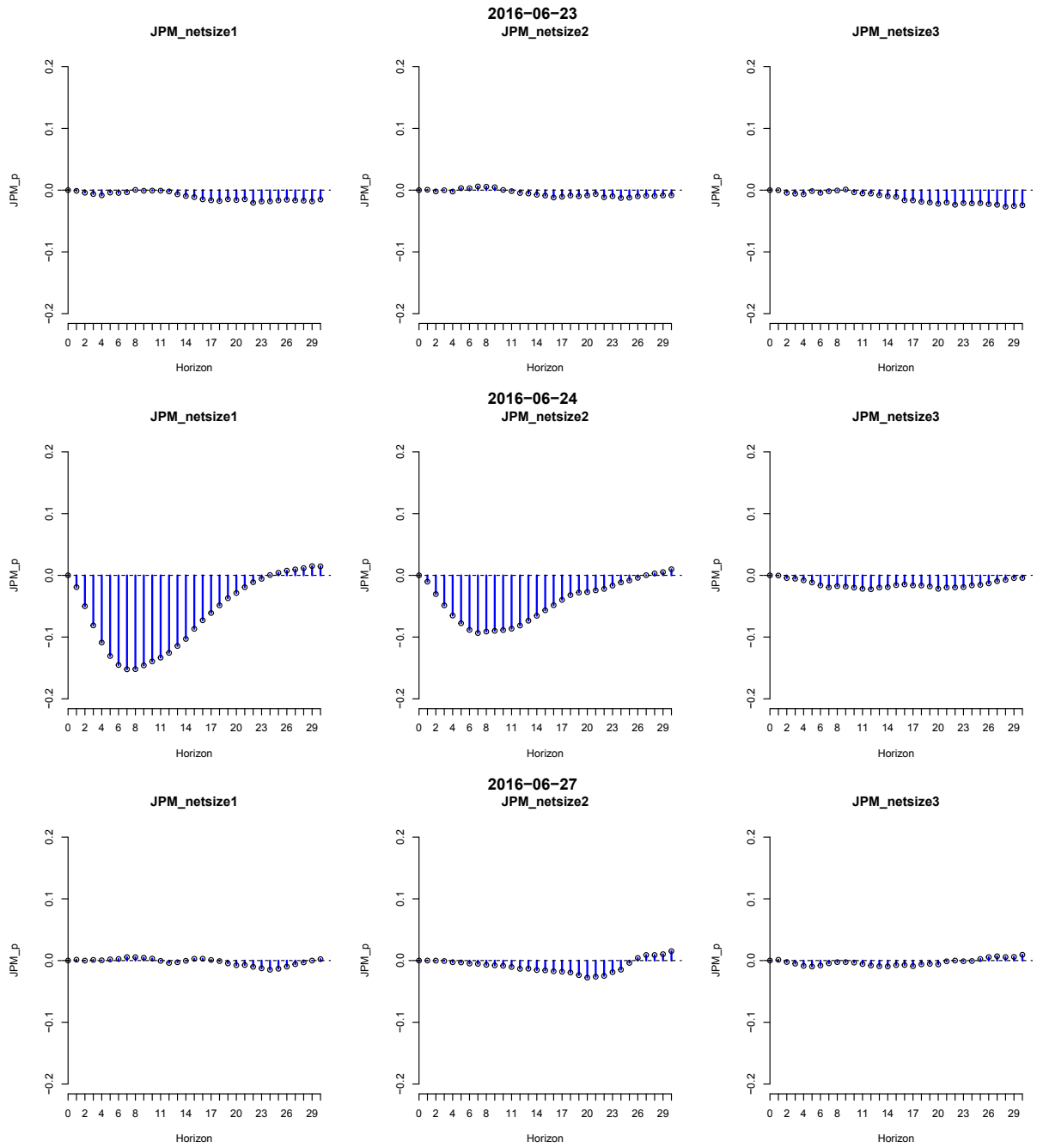
Figure 1.5. The bootstrapped market impacts from net size factors for JPM from 23rd of June to 27th of June.

hfhd_marketimpact_diff
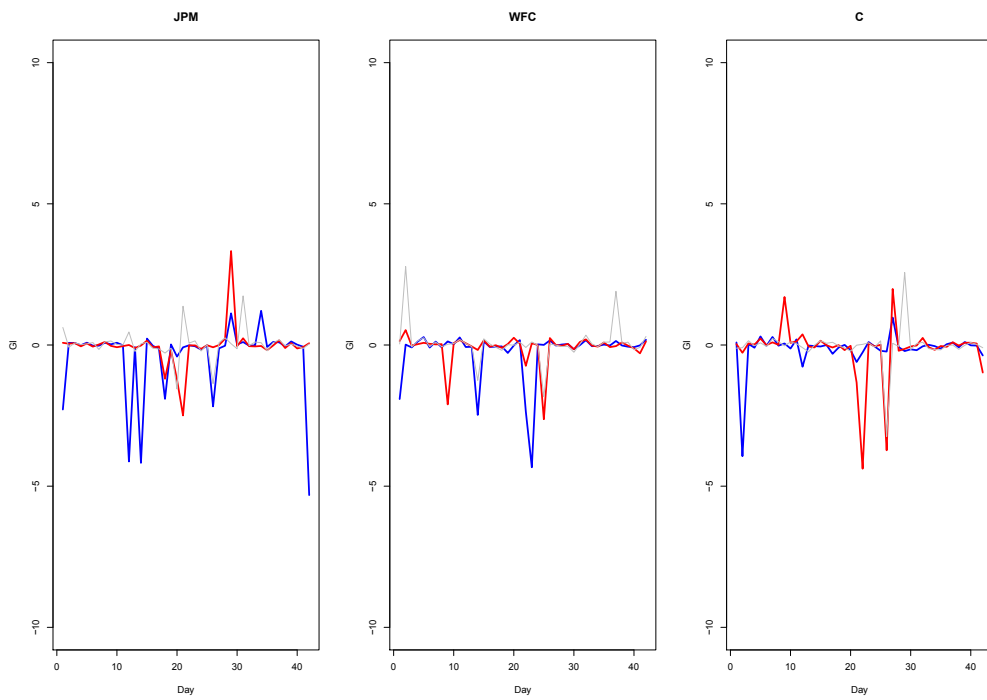
Figure 1.6. The bootstrapped net market impact from 1st level net size (blue), 2nd level net size(red) and 3rd level net size (grey), for three financial institutions over 42 days, from 1st, June of 2016 to 29th, July of 2016.

hfhd_marketimpact_diff

in Figure 1.7 and 1.8. The impulse response estimates for the companies from the same industry sector to the whole system are presented for financial institutions and Health-care companies respectively.

The estimation results indicate that the transmission of shocks within the companies from the same industry sector. In particular, we observe the risk transmission channels of JPM -> C -> WFC for finance institutions, and the risk cluster PFE -> MRK -> JNJ -> PFE. With respect to the response of the financial companies, the positive impact dies out very slowly.

To sum up, impulse responses show that risk transmission is more likely to happen within the same type of companies. Our evidence suggests that the stocks within the same industry are a more potent channel of risk sharing. The bootstrapped GIRF allows us to provide novel evidence on robust risk channels and risk clusters, with focuses on the NASDAQ market to examine the way risk arises endogenously and how shocks propagate within one industry sector.

1.5.2 Robustness

To confirm and quantify the robustness of all the above statements, we summarize the significant pairwise impacts in a time-varying pattern in Figures 1.9, 1.10 and 1.11. They plot the interactions between two companies, separately for each industry sector. For instance, the upper panel of Figure 1.9 consists of vertical parallel bars that show the impacts of "T-> MSFT" (dark gray bars, impact from left to right as shown in subfigure title) and "MSFT -> T" (light gray bars, impact from right to left). Moreover, we observe some clusters of significant pairwise impacts, e.g., the "MSFT->IBM" lasting from 09th trading day to 16th trading day (i.e., 13th to 22nd of June, 2016), "MSFT->T"(light grey) from 27th to 34th trading day (i.e., 8th to 19th of July, 2016), the "WFC -> JPM" from 24th to 30th trading day (i.e., 05th to 13th of July, 2016) and the "JPM ->C" from 12th to 18th trading day (16th to 24th of June, 2016). In general, the clusters of the significant pairwise impacts indicates the dependence structure of cross-sectional impacts over time, the lag is normally at least 6 trading days.

For the Technology companies in Figure 1.9, we find that the pairwise companies does not exhibit a particular strong pattern on their mutual interactions. For both MSFT and IBM, they have strong effects on T. IBM, on the other hand, effects MSFT as well. These findings suggest that when the risk(shock) of IBM arises, the shocks will transmit to T via MSFT. From 9th of June to 34th of June 2016, the impact of MSFT->IBM went through two smaller cycles, during which it moved within the 12-20 range.
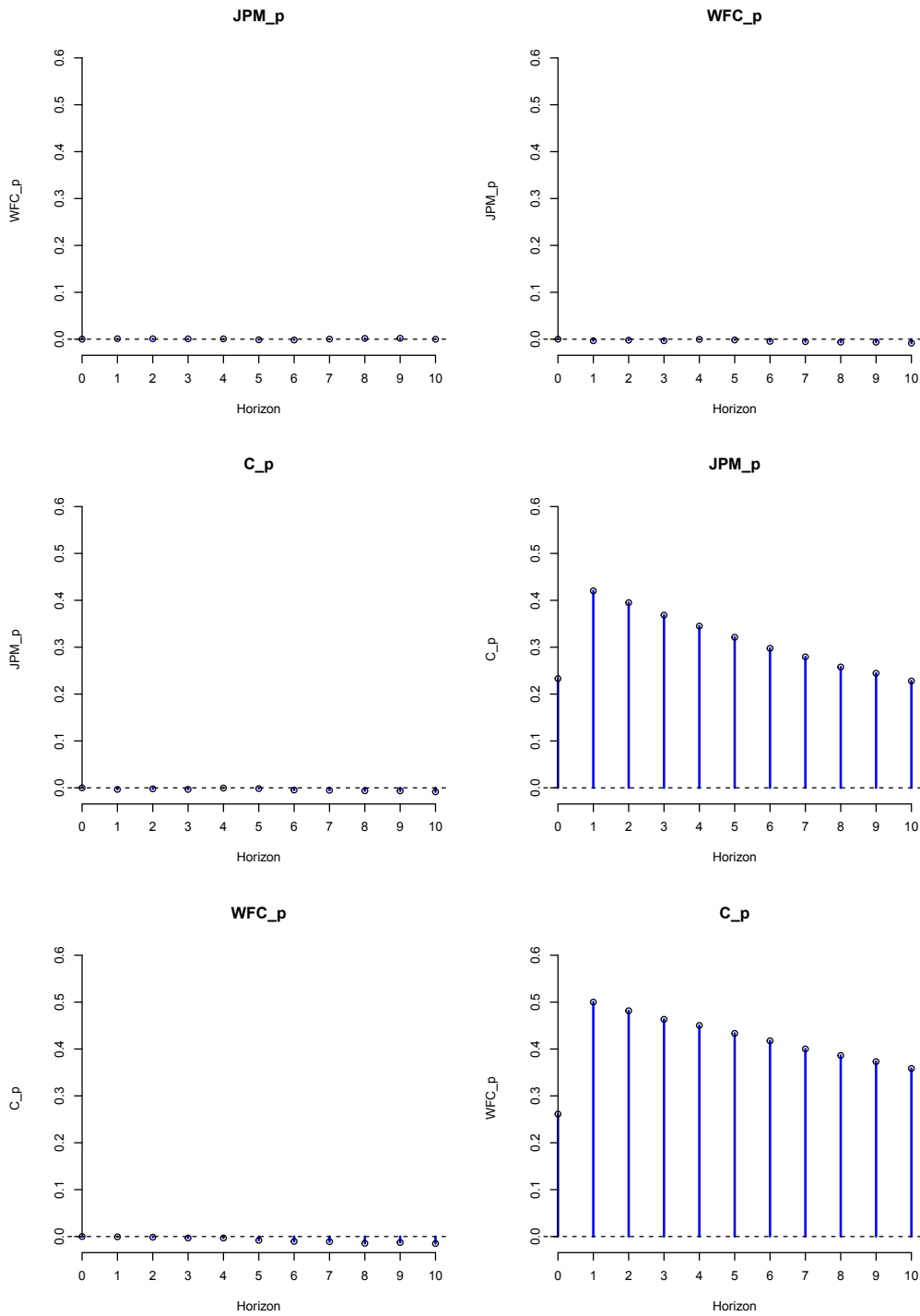
Figure 1.7. The transmit of for three financial institutions on 28th, June of 2016.

hfhd_risktran
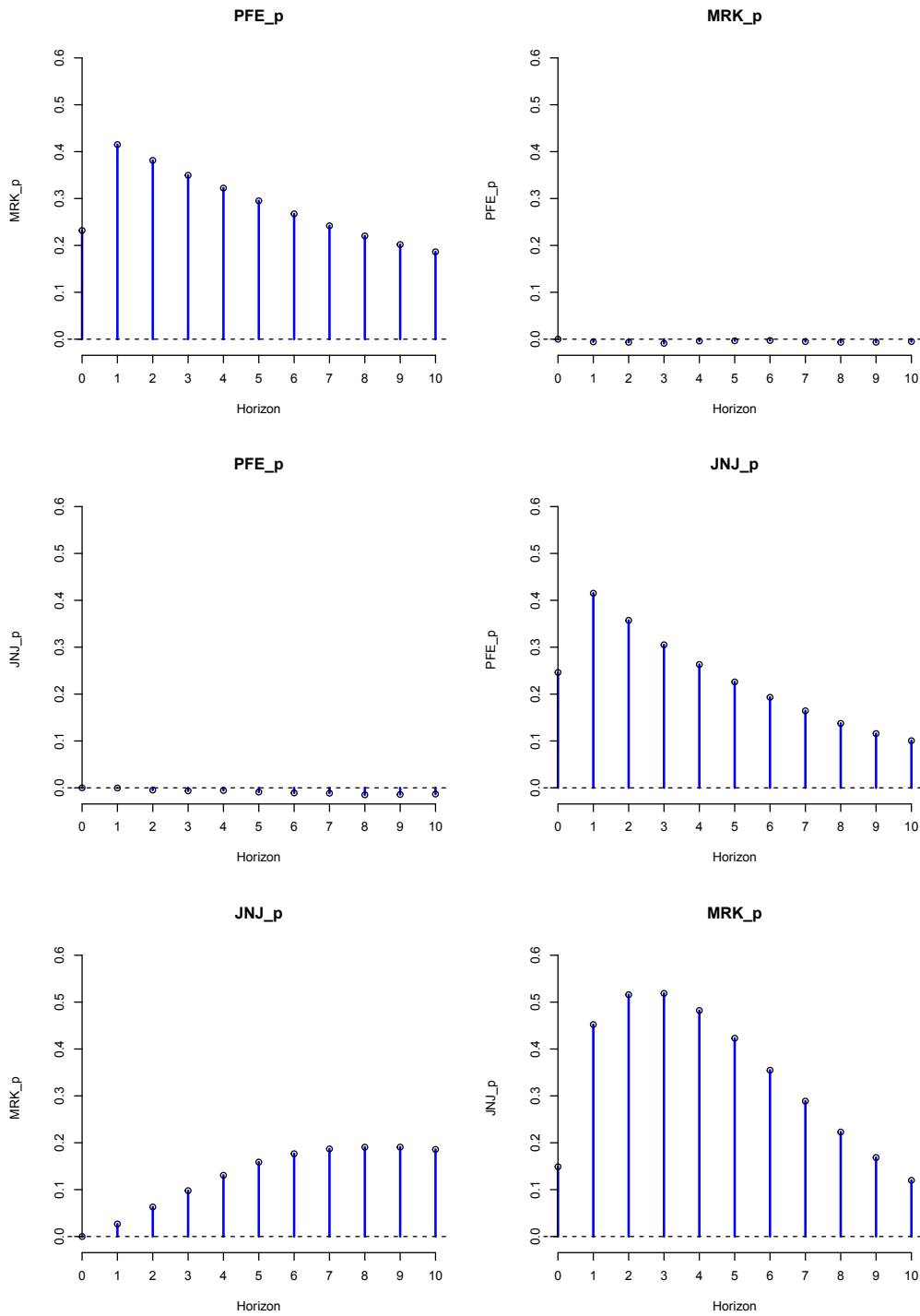
Figure 1.8. The transmit of for three financial institutions on 28th, June of 2016.

hfhd_risktran

Figures 1.10 and 1.11, which show the cross-sectional direct and indirect impacts for financial institutions and healthcare companies, reinforces the analysis. The impacts for financial institutions are stronger than the heathcare, which suggests that financial sector might be a useful leading indicator of risk, and the previous empirical literature lends support to this conjecture. Furthermore, we find some robust risk transmission channels based on the estimation, e.g. "MRK->JNJ->PFE", "MRK->PFE" and "JPM->C->WFC". The general message is similar, when looking at the the at the bank level. We identify the JPM, WFC are strongly correlated, which contains information as sources of systemic risk.
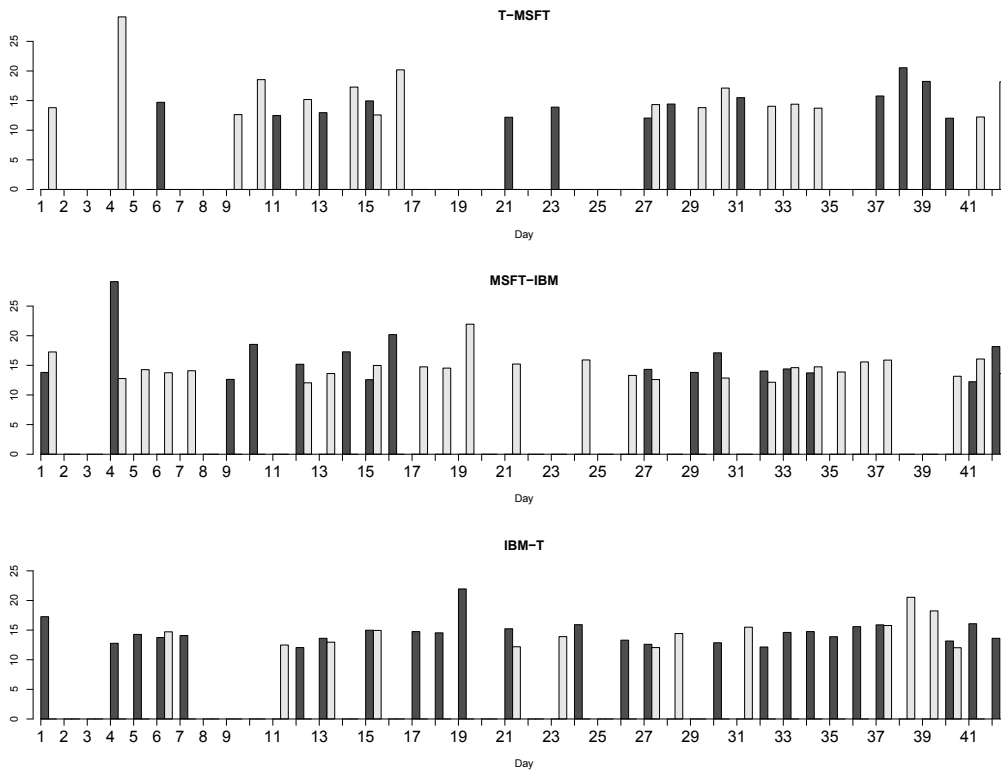
Figure 1.9. The time-varying pairwise impacts for three technology companies.

hfhd_rob

1.6 Network Analysis

The network in this study consists of the stocks with limit order books which are neglected in current literature. Specifically, we study the complex system in terms of not
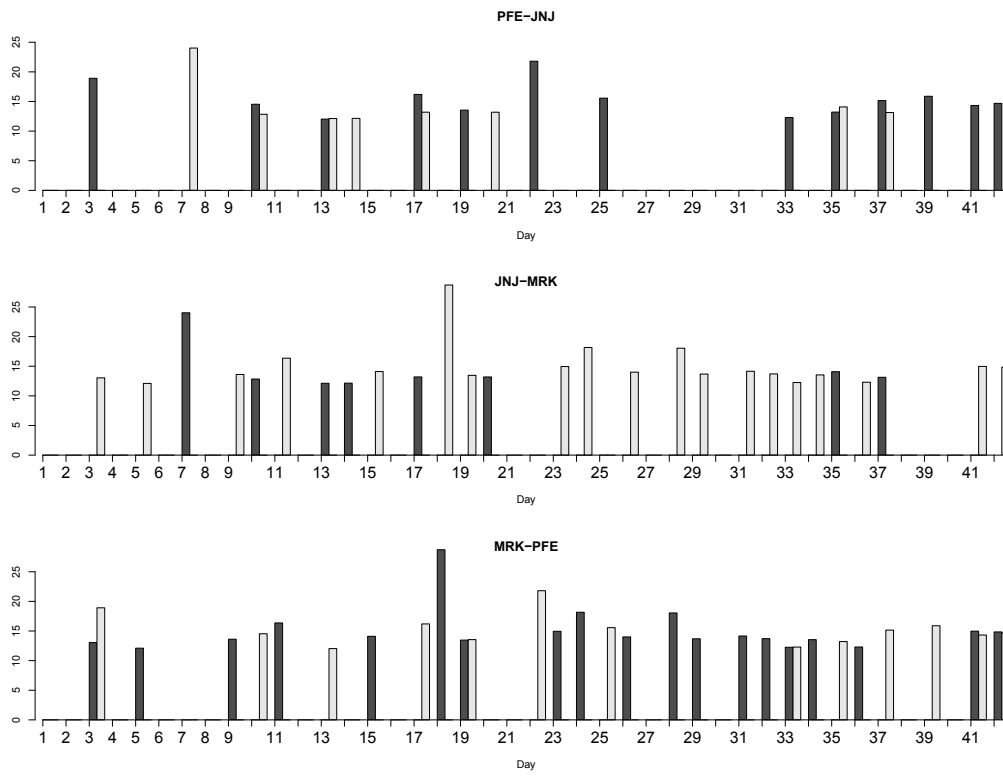
Figure 1.10. The time-varying pairwise impacts for three healthcare companies.

hfhd_rob
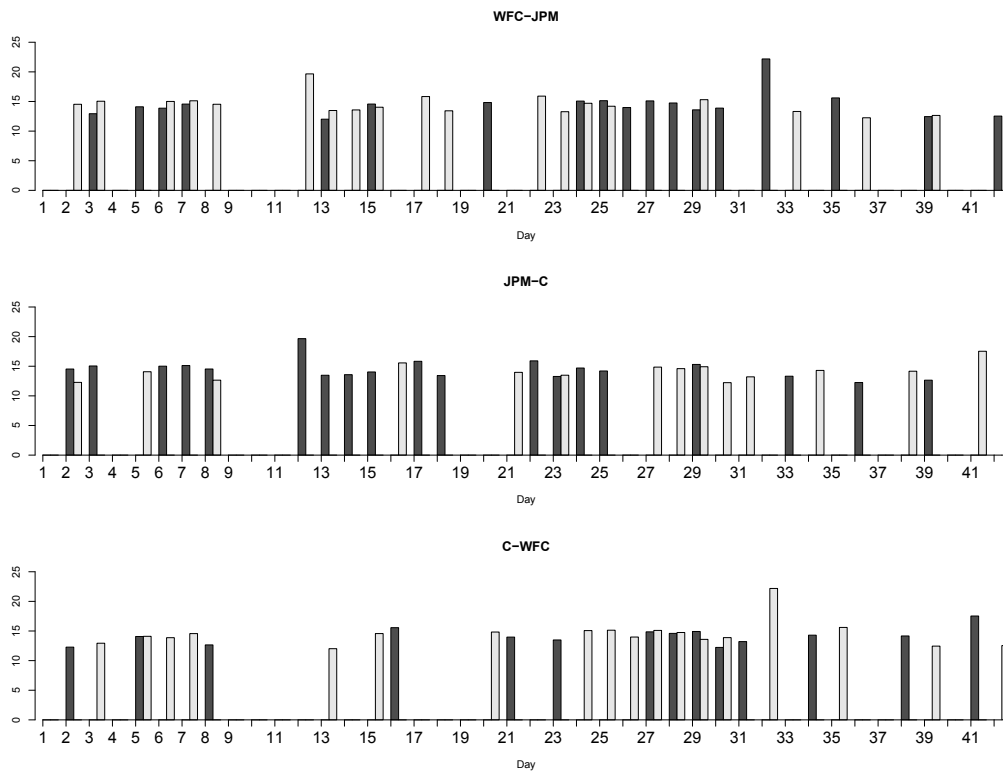
Figure 1.11. The time-varying pairwise impacts for three financial companies.

hfhd_rob
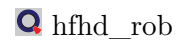
only the dynamic links across stocks but also the market impacts of limit orders across stocks.

1.6.1 Static Full-sample Connectedness

In accordance with the discussion in section 1.3.2, we randomly select one day and depict the sparse full sample directional connectedness Table 1.3 in Figure 1.12. Figure 1.13 shows the connectedness between different stocks. The price factor and size factors that belong to the same company appear in the same color, the width of edges between two nodes represents the connectedness.
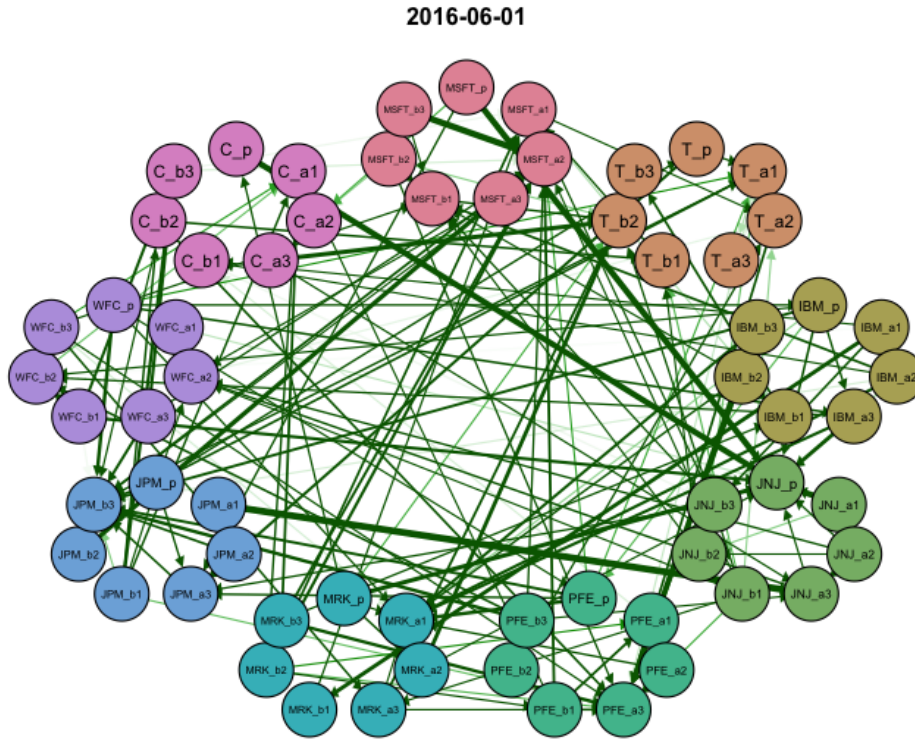
Figure 1.12. The full sample network plot on 1st of June, 2016.

Table 1.4 summarizes LN-GFEVD full sample connectedness Table 1.3, we report the directional connectedness estimators "from" and "to" associated with the forecast error variation $\lambda_{ij}^b(h)$ for a specific order book across various stock when the arising
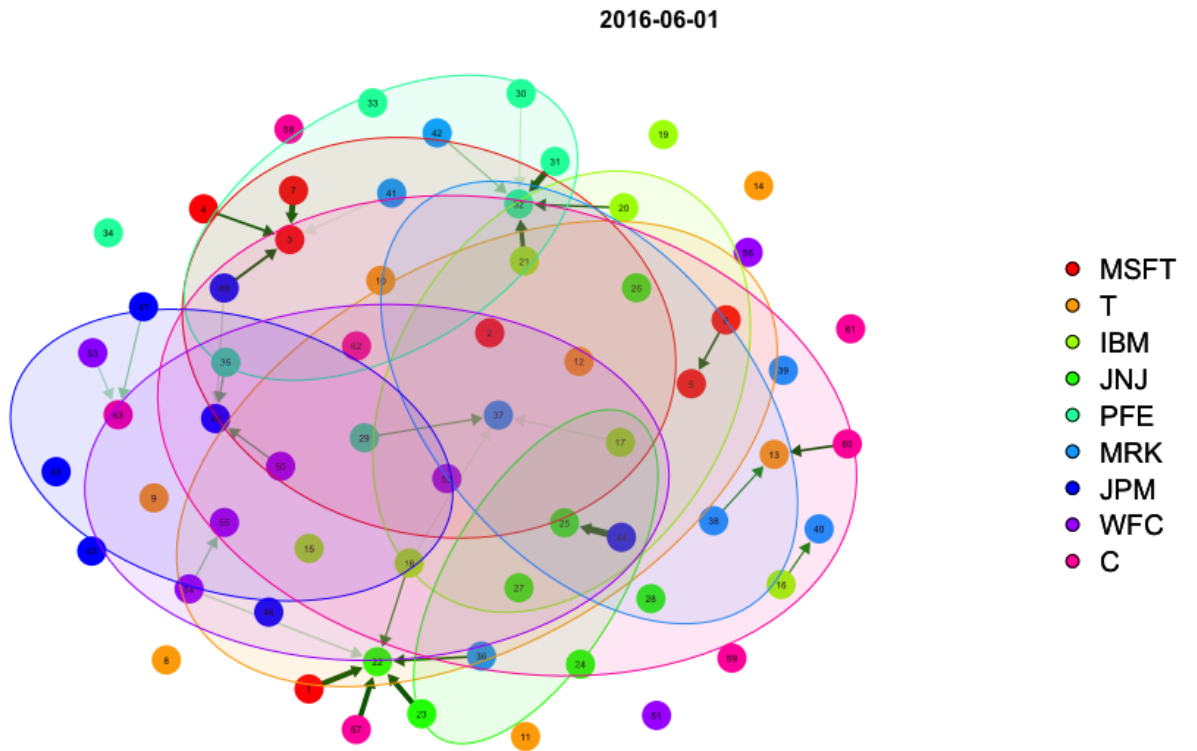
Figure 1.13. The full sample network plot on 1st of June, 2016.

shocks transmit from one stock to the others. As discussed before, the element in the connectedness table measures the total impact of all K shocks on the i-th variable, and these contributions sum to unity, which suggests the "from" estimator is unity. The strong positive and negative impacts are marked in red and blue respectively. In general, the price factors are more sensitive to the shocks than the size factors.

	MSFT			T			IBM		
	from	to	net	from	to	net	from	to	net
price	1.00	0.76	-0.24	1.00	0.72	-0.28	1.00	0.88	-0.12
1st ask	1.00	0.98	-0.02	1.00	1.28	0.28	1.00	1.20	0.20
1st bid	1.00	0.78	-0.22	1.00	0.58	-0.42	1.00	1.09	0.09
2nd ask	1.00	1.26	0.26	1.00	0.82	-0.18	1.00	1.52	0.52
2nd bid	1.00	0.90	-0.10	1.00	0.99	-0.01	1.00	1.07	0.07
3rd ask	1.00	0.79	-0.21	1.00	0.88	-0.12	1.00	0.92	-0.08
3rd bid	1.00	0.95	-0.05	1.00	0.72	-0.28	1.00	1.15	0.15
	JNJ			PFE			MRK		
	from	to	net	from	to	net	from	to	net
price	1.00	0.46	-0.54	1.00	1.14	0.14	1.00	0.94	-0.06
1st ask	1.00	1.17	0.17	1.00	1.50	0.50	1.00	1.15	0.15
1st bid	1.00	0.77	-0.23	1.00	1.19	0.19	1.00	1.14	0.14
2nd ask	1.00	0.95	-0.05	1.00	0.93	-0.07	1.00	1.66	0.66
2nd bid	1.00	1.13	0.13	1.00	0.89	-0.11	1.00	1.35	0.35
3rd ask	1.00	0.94	-0.06	1.00	0.62	-0.38	1.00	1.24	0.24
3rd bid	1.00	0.97	-0.03	1.00	0.94	-0.06	1.00	0.74	-0.26
	JPM			WFC			C		
	from	to	net	from	to	net	from	to	net
price	1.00	0.48	-0.52	1.00	1.39	0.39	1.00	0.88	-0.12
1st ask	1.00	1.53	0.53	1.00	1.17	0.17	1.00	1.18	0.18
1st bid	1.00	1.10	0.10	1.00	0.90	-0.10	1.00	0.90	-0.10
2nd ask	1.00	1.05	0.05	1.00	1.11	0.11	1.00	0.86	-0.14
2nd bid	1.00	1.18	0.18	1.00	0.91	-0.09	1.00	0.85	-0.15
3rd ask	1.00	0.87	-0.13	1.00	0.72	-0.28	1.00	0.93	-0.07
3rd bid	1.00	0.76	-0.24	1.00	1.16	0.16	1.00	1.02	0.02

Table 1.4. Summary of the full connectedness table on 1st of June, 2016, based on h=30 (30 minutes/steps ahead forecasts).

Figure 1.14 shows the sparse network graph of six continuous trading days, from 21st of June to 28th of June, 2016. It is interesting that the estimated graph on 27th of June (first trading day after Brexit on 24th of June) is more sparse than the other trading days. This is, perhaps because idiosyncratic shocks have always hit individual

companies and these shocks have been transmitted to more other companies. Hence the averaged shock transmitted to others is smaller and therefore generating a more sparse graph. These results, on the other hand, suggest the companies are expected to be less interconnected during important event because the companies tend to react more independently to important event than usual.

1.6.2 Time-varying Connectedness

With our analysis, the network we estimate can be exhibited in a time-varying pattern. In this way, we identify several connections which remain quite stable over time.

Figures 1.15, 1.16, 1.17 and 1.18 establish the time-varying impacts for the limit order book of MSFT. In total, each factor of MSFT limit order book can be effect by 63 kinds of shocks, here we select top two risk contributors. The time varying network for the other 8 companies are plotted in Appendix A.3.

1.7 Conclusion

In this study, we conduct network analysis for limit order books (LOB) across stocks for a better understanding of market impact, and show how network for LOB can be constructed in the presence of microstructure noise and non-synchronous trading. This paper contributes to directed network estimation through penalized Vector autoregressive (VAR) approach. The connectedness table is directly derived from generalized impulse response function with attractive property of order invariance, by way of bootstrapping techniques. We progress by focusing on the models that capture the dynamics of LOB and their influence over time. We aim to mimic the LOB trading mechanism to find out significant market impact. Our primary finding is that order imbalance general exists across stocks, bootstrapped market impacts can be quantified. With our methodology, we identify the significant market impact caused by the arrival of a large limit order, and several robust risk transmission channels and risk clusters. The financial institutions are connected more closely compared with the firms come from other industry. The estimation results of LN-network are also compared and studied.
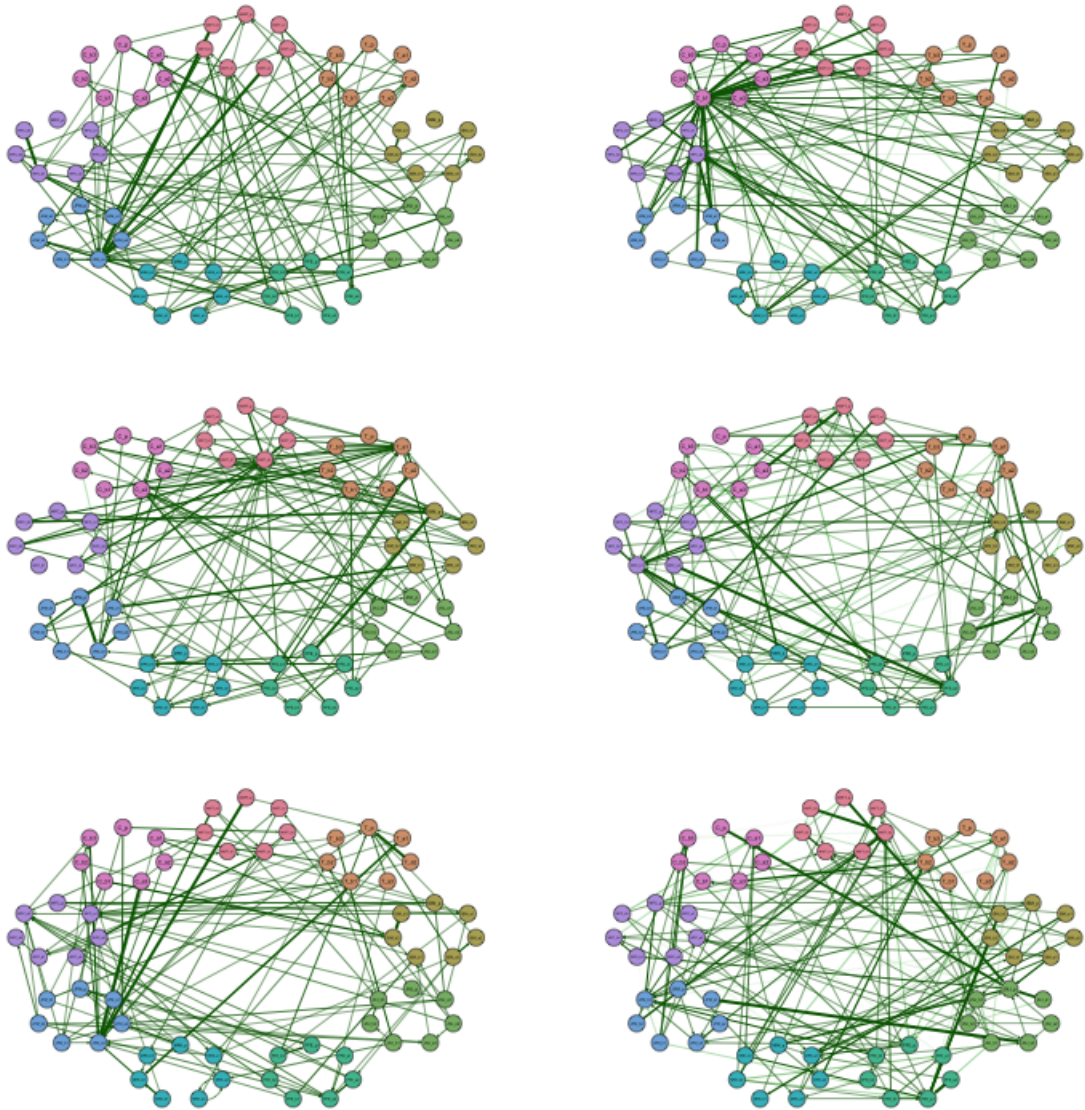
Figure 1.14. The full sample network graph of six continuous trading days, from 21st of June to 28th of June, 2016. 25th and 26th of June are Saturday and Sunday.

MSFT_p
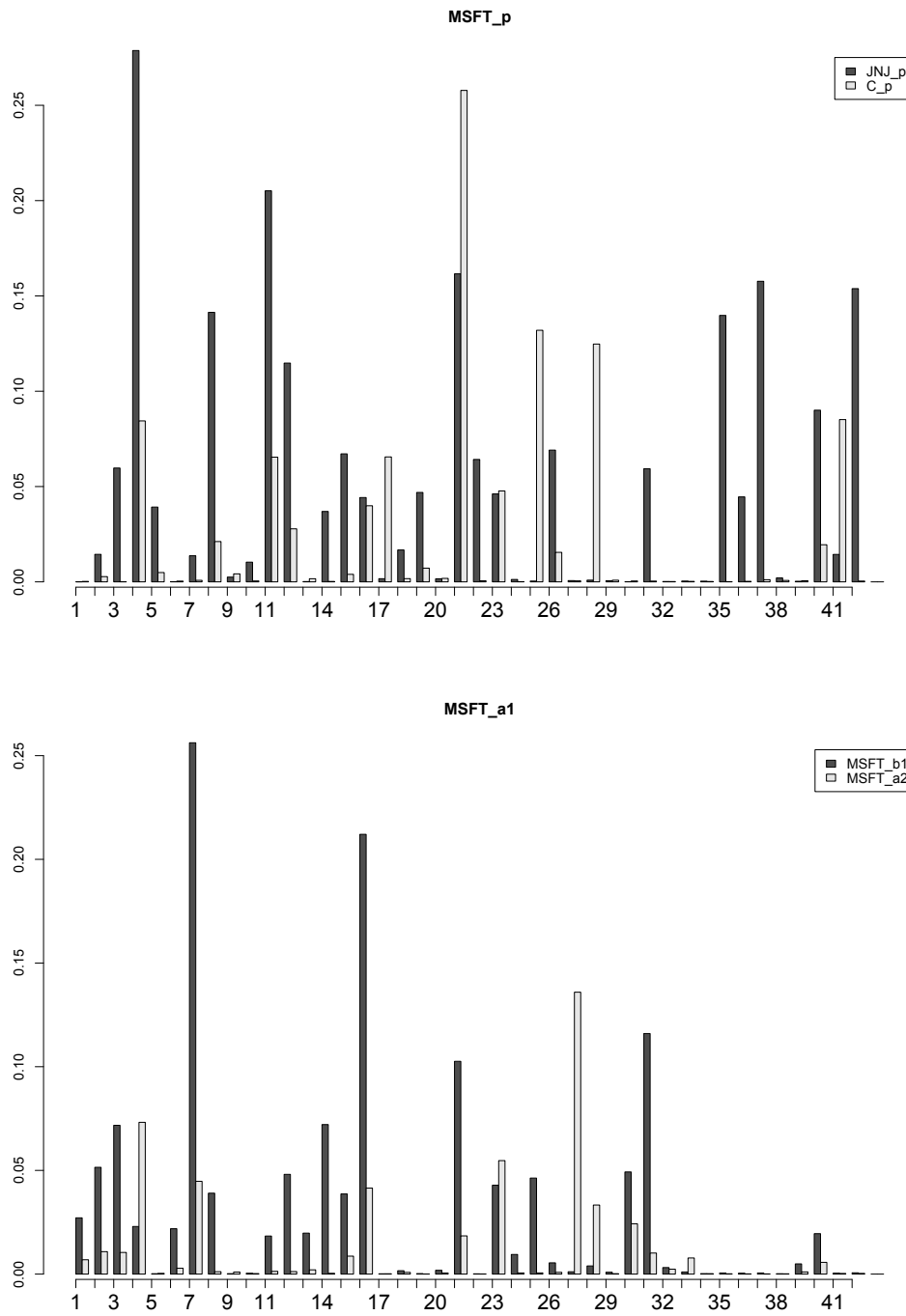
MSFT_a1

Figure 1.15. The time-varying network for price factor of MSFT, and 1st level ask size factor of MSFT.

hfhd_tvnet
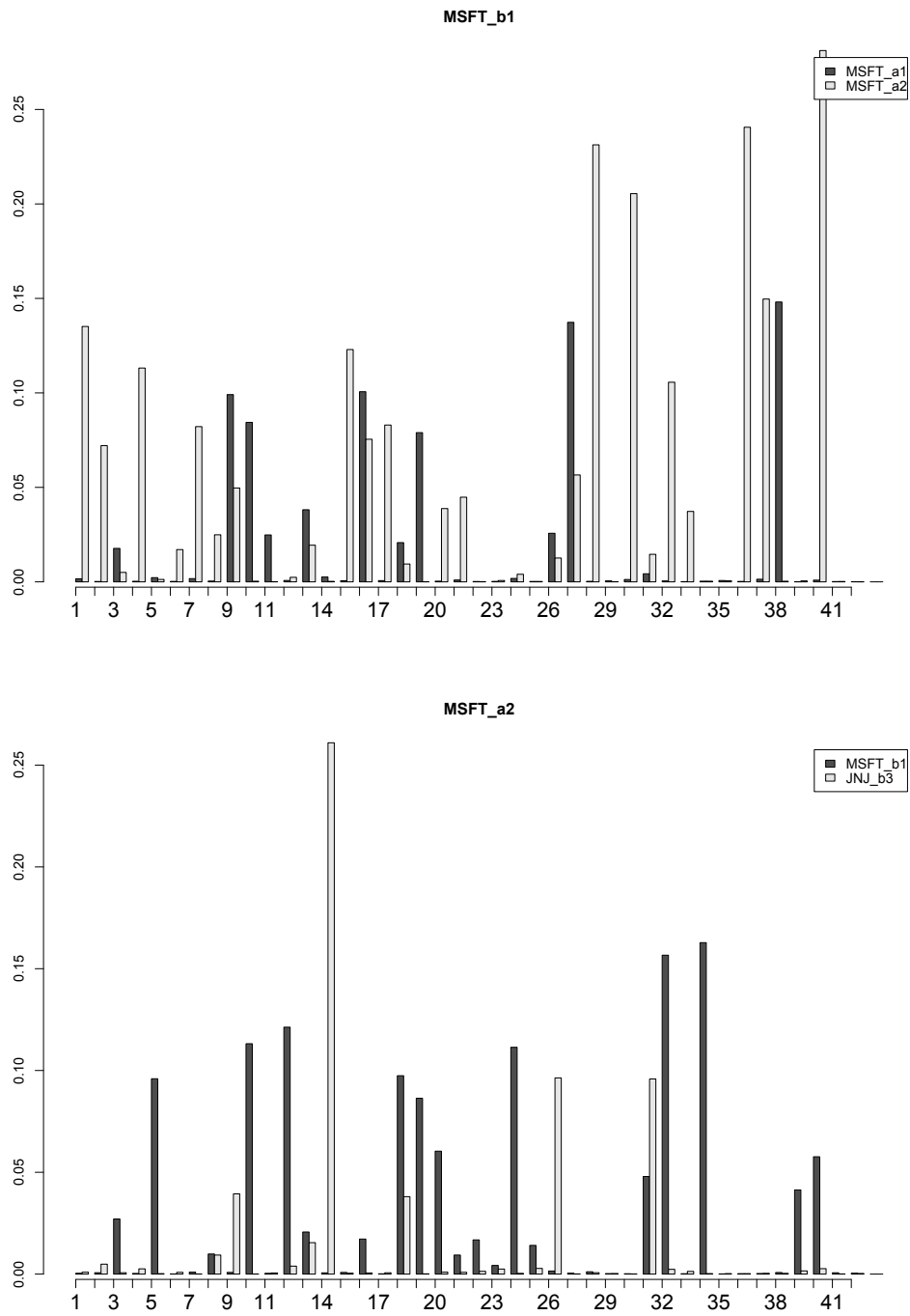
Figure 1.16. The time-varying network for 1st level bid size factor of MSFT, 2nd level ask size factor of MSFT.
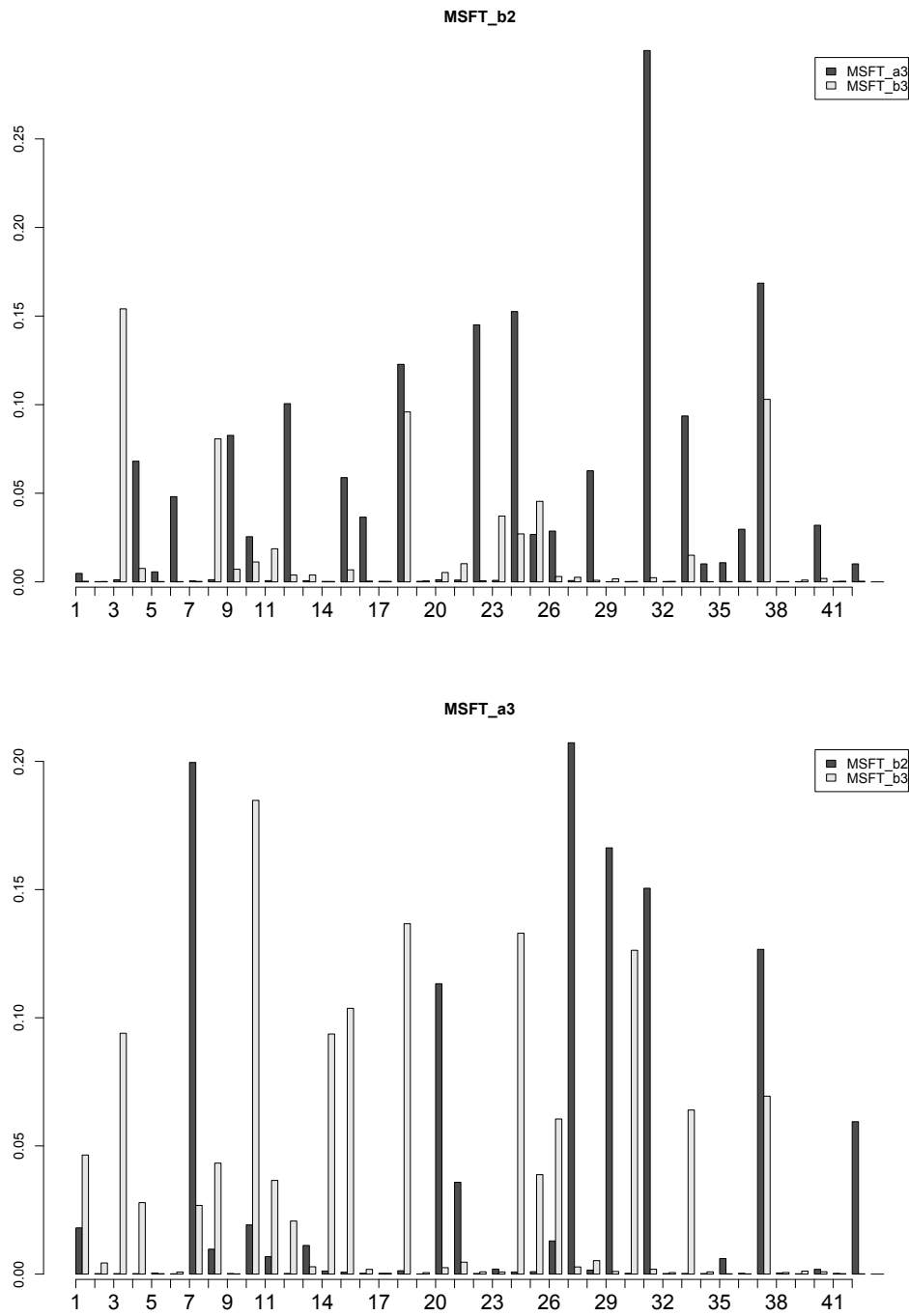
Q hfhd__tvnet

Figure 1.17. The time-varying network for 2nd level bid size factor of MSFT, 3rd level ask size factor of MSFT
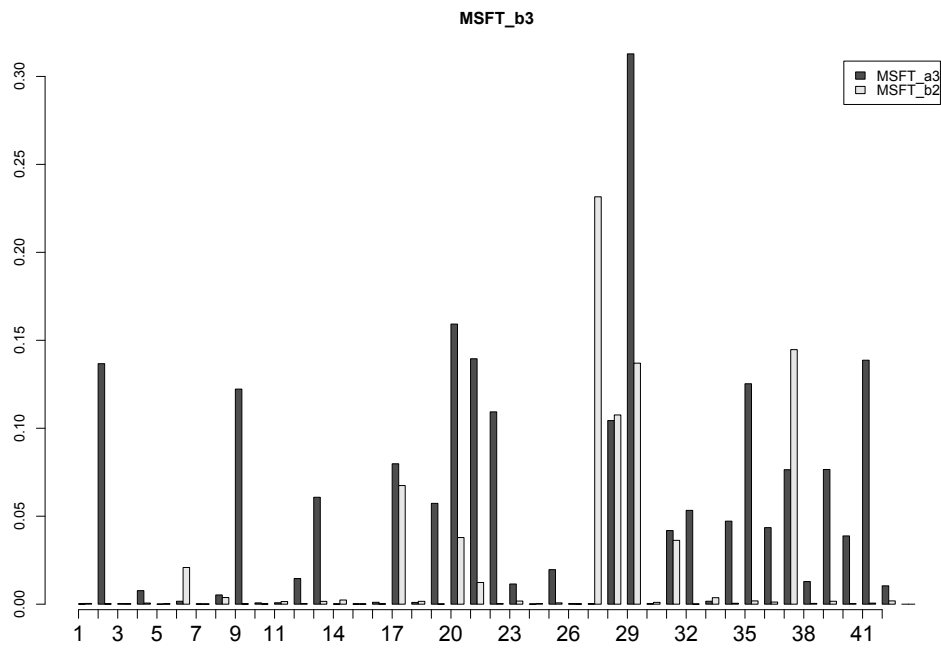
hfhd_tvnet

Figure 1.18. The time-varying network for 3rd level bid size factor of MSFT.

hfhd_tvnet

2 Sparsity Analysis of Energy Price Forecasting

2.1 Introduction

Affordable and reliable energy supply is essential for our industrial growth. Achieving these in times of growing demand, raw materials shortage and climate change poses great challenges for us all. It is essential for us to keep ahead in future power supply. Germany's electricity supply is undergoing radical change. At present, conventional energy sources generate approximately 74% of Germany's electricity. However, the ongoing expansion of renewable energy and the phase-out of nuclear energy for power generation will change the composition of the electricity mix, which in return, will generate the pricing signal that affects the electricity trading. Therefore a study on the electricity derivative market, especially in the forward market can be a way to hedge against the risk appeared in the electricity market. Based on these insights, energy companies may decide to invest both electricity spot and derivatives markets for portfolio risk management. However the number of relevant factors may be huge, subset selection methods could be a useful tool to select important power contracts to capture the energy market risk.

High-dimensional statistical problems arise from diverse fields of scientific research and technological development, including energy market. The traditional idea of best subset selection methods is computationally too expensive for many modern statistical applications. Variable selection techniques have been successfully developed in recent years and they play a pivotal role in contemporary statistical learning and scientific discoveries. Researchers have proposed various penalized estimators, for example, least absolute selection and shrinkage operator (lasso) of Tibshirani (1996b) is famous for its simultaneous model selection and estimation. Besides, it has sound statistical properties and is efficient to implement. In recent years, lasso has been extended to high-dimensional case, one typical work is Bickel et al. (2009). There are many other popular methods contribute to the literature, such as smoothly clipped absolute deviation (SCAD) estimator of Fan and Li (2001), adaptive Lasso of Zou (2006), elastic net estimator of Zou and Hastie (2005), Dantzig selector of Candes and Tao (2007). In a ultra high-dimensional case where the dimensionality of the model is allowed to grow exponentially in the sample size, it is helpful to begin with screening to delete some significantly irrelevant variables

from the model. Fan and Lv (2008) introduce a method called sure independence screening (SIS hence-after) for this goal. Even when the regularity conditions may not hold, Fan et al. (2009) extend the iterated-SIS method to work by iteratively performing feature selection to recruit a small number of features.

The German power derivative market is an interconnected world, with a wide range of electricity derivative contracts. In this paper we build up a ultra high-dimensional network in which nodes represent power contracts and links represent the magnitude of connectedness. To better understand the interaction between power contracts, the iterated-SIS methods combined with penalized estimators are applied to estimate the sparse web of connections. Our network of interest is constructed in the context of time series based on vector autoregressions (VAR), the iterated-SIS methods are of much use when building VAR models since the number of parameters to estimate increases quadratically in the number of variables included. In addition, asymptotic properties of lasso for high-dimensional time series have been considered by Loh and Wainwright (2011), Wu and Wu (2016). Kock and Callot (2015b) establish the high-dimensional VAR estimation with focus on lasso and adaptive lasso. Basu et al. (2015) investigate the theoretical properties of regularized estimates in sparse high-dimensional time series models when the data are generated from a multivariate stationary Gaussian process.

To quantify the associations between individual power contract and energy exchange market, the network we constructed is obtained from the forecast error variance decomposition (FEVD) based on VAR estimates in the framework of Koop et al. (1996) and Pesaran and Shin (1998). This kind of connectedness measure is proposed by Diebold and Yılmaz (2014) for conceptualizing and empirically measuring weighted, directed network at a variety of levels. However one may confront the dimensionality problem while nevertheless remaining squarely in the Diebold-Yilmaz connectedness traditon, Demirer et al. (2017) use lasso method to select, shrink and estimate the high-dimensional network. Related empirical work are with more focus on financial banking contexts, for examples see Acharya et al. (2012), Acharya et al. (2017), Giglio et al. (2016), Hautsch et al. (2014). While estimates of the network yield the qualitative links between power contracts, individual impact from specific contract can be estimate and speculate accordingly. Hence the risk contribution from the market component can be identified clearly, this will help us to learn more about the German power market functioning and environment.

The rest of the paper proceeds as follows. Section 2 introduces the structure of electricity market in Germany. In section 3, we discuss the econometric methodology used in the variable selection. The technical details are in the Appendix. The empirical results are shown in Section 4. Finally section 5 concludes.

2.2 Overview of Energy Market

2.2.1 German Electricity Derivative Market

The German electricity market is Europe's largest, with annual power consumption of around 530 TWh and a generation capacity of 184 GW. As a net energy exporter, the export capacity of Germany is expected to continue to grow as planned interconnections expand cross-border transmission capacity with several neighbouring countries. Germany has significant interconnection capacity with neighbouring EU member states as well. It is interconnected with Austria, Switzerland, the Czech Republic, Denmark, France, Luxembourg, the Netherlands, Poland, and Sweden. To maintain stable and reliable supply of electricity, the so-called Transmission system operators (TSOs) keep control power available. Primary control, secondary control, and tertiary control reserve are procured by the respective TSOs within a non-discriminatory control power market in accordance with the requirements of the Federal Cartel Office. Demand for control energy is created when the sum of power generated varies from the actual load caused by unforeseeable weather fluctuations in the case of renewable energies.

Electricity is traded on the exchange and over the counter. Standardised products are bought and sold in a transparent process on the exchange, which, for Germany, is the European Energy Exchange EEX in Leipzig, the European Energy Exchange EPEX SPOT in Paris and the Energy Exchange Austria (EXAA) in Vienna. The European Energy Exchange (EEX) is the leading energy exchange in Europe. It develops, operates and connects secure, liquid and transparent markets for energy and commodity products. Contracts on power, coal and CO2 emission allowances as well as freight and agricultural products are traded or registered for clearing on EEX. EPEX SPOT, Powernext, Cleartrade Exchange (CLTX) and Gaspoint Nordic are also members of EEX Group. The German wholesale electricity market is broadly made up of three elements, a forward market, a day-ahead market and an intra-day market. These submarkets generate the pricing signal which electricity production and consumption align to. The objective of this paper is to analyse the interaction of different future contracts traded in the forward market, whether forward market is influenced by market power of spot prices traded in EPEX market.

Electricity providers and electricity purchasers submit their bids in their national day-ahead market zones. The exchange price on the day-ahead market is determined jointly for coupled markets. Electricity providers and electricity purchasers submit their bids in their national day-ahead market zones. In an iterative process, the demand for electricity in the market zone is served by the lowest price offers of electricity from all the market areas until the capacity of the connections between the market zones (cross-border interconnectors) is fully utilised. As long as the cross-border interconnectors

have sufficient capacity, this process aligns the prices in the coupled market areas. On account of market coupling, the national power demand is covered by the international offers with lowest prices. The upshot is that on the whole less capacity is required to meet the demand. As shown in Figure 2.1, Phelix Future, as the product traded in Germany, is a financial derivatives contract settling against the average power spot market prices of future delivery periods for the German/Austrian market area.
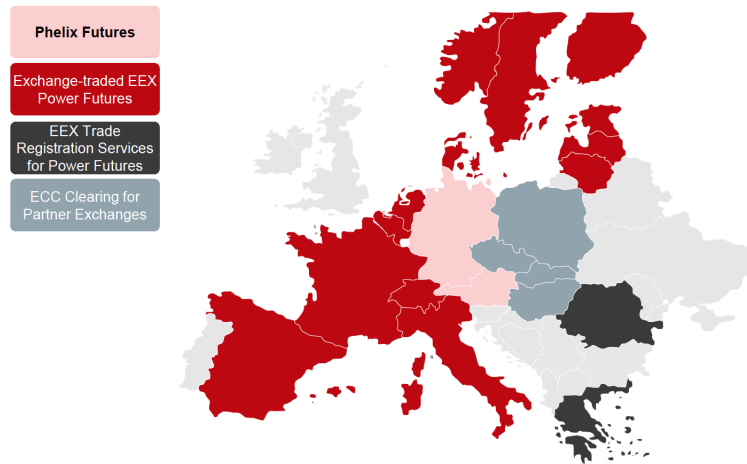
Figure 2.1. The distribution of European power derivatives in EEX market. Source: EEX website

2.2.2 Phelix Futures

Electricity supply deliveries in the forward market can be negotiated up to seven years in advance, but for liquidity reasons typically only look out three years, and in fact most futures trading focuses one year ahead. The Phelix Future is a financial derivatives contract referring to the average power spot market prices of future delivery periods of the German/Austrian market area.

As the most liquid contract and benchmark for European power trading, the underlying of these future contracts is the Physical Electricity Index determined daily by EPEX Spot Exchange for base and peak load profiles. To be more specific, the Phelix Base contract is average price of the hours 1 to 24 for electricity traded on spot market, while the Phelix Peak is the average price of the hours 9 to 20 for electricity traded on spot market. EEX offers continuous trading and trade registration of financially fulfilled Phelix Futures, with Day/Weekend Futures, Week Futures, Month Futures, Quarter Futures and Year Futures available.

In addition, the Phelix market is also successfully connected to other European power markets. The products of Location Spread enables members to trade price differences between markets, thus enabling participants to benefit from improved liquidity and tighter spreads, for instance, Phelix / French Power, Italian / Phelix Power, Phelix / Nordic Power and Phelix / Swiss Power. For the empirical work of this paper, we use the Phelix Future data to find price drivers and important variables in the big system we construct. The decision-making mechanism of energy companies will also be explored.

2.3 Theoretical Framework

2.3.1 Model Description

We have access to spot prices, trading prices of different future contracts. An interesting question is the how all these prices interact with each other? Which variables are crucial for the whole system? To answer this question, we are going to build a Vector Autoregressive Model. However, due to the large number of variables in the system, some sparsity assumption must be imposed for the sake of an accurate estimate. Sparsity will also help find out the most important variables. Large dimension comes from two parts:

1. varieties of power derivative products;

2. large lag in VAR model to avoid the correlation of error terms.

The VAR(p) model (VAR model of order p) is constructed according to Lütkepohl (2005),

$$
\begin{aligned}
y_t &= \nu + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + u_t \\
&= \nu + (A_1, A_2, \ldots, A_p) \left(y_{t-1}^\top, y_{t-2}^\top, \ldots, y_{t-p}^\top \right)^\top + u_t
\end{aligned}
\tag{2.1}
$$

where $y_t = (y_{1t}, y_{2t}, \ldots, y_{Kt})^\top$ is a $(K \times 1)$ random vector consisting K prices we have at time t, t from 1 to T. A_i are fixed $(K \times K)$ coefficient matrices. ν is a $(K \times 1)$ vector of intercept terms, p is lag and $u_t = (u_{1t}, u_{2t}, \ldots, u_{Kt})^\top$ is a K-dimensional innovation process.

The coefficients $\nu, A_1, \ldots, A_p$ are assumed to be unknown in the following. The time series data $y_1, y_2, \ldots, y_T$ of the y variable is available and will be used to estimate the coefficients. The multiple time series data will be partitioning into sample and presample

values to facilitate the following analysis. Define

$$
\begin{aligned}
Y &= (y_1, y_2, \ldots, y_T) \\
B &= (\nu, A_1, A_2, \ldots, A_p) \\
Z_t &= (1, y_t, y_{t-p+1})^\top \\
Z &= (Z_0, Z_1, \ldots, Z_{T-1})
\end{aligned}
\tag{2.2}
$$

Hence for multivariate case, the model described in equation (2.1) can also be rewritten as

$$
Y = BZ + U \tag{2.3}
$$

where $U = (u_1, u_2, \ldots, u_T)$. The compact form (2.3) is equivalent to

$$
vec(Y) = (Z^\top \otimes I_K)vec(B) + vec(U) \tag{2.4}
$$

If the vector of intercept terms ν is assumed to be zero, we can thus conclude that the total dimension of the model to be estimated is pK^2 and the total number of observations is KT.

The ration $\frac{Kp}{T}$ could be large due to the reasons mentioned earlier, which deteriorates the accuracy of final estimate. Worse still, if $Kp > T$, the model is not identified with traditional method. Therefore, we use variable selection technique, such as lasso, to estimate the model. Besides, under normal assumption of error term, the upper bound of error in estimation is positively correlated in $\frac{\log(K^2 p)}{T}$, part of oracle inequlity. the estimation results can be further developed by adding one more step of sure independence screening (SIS hence-after) before variable selection step. Another advantage of SIS is that it could mitigate the problem caused by multicollinearity, which is common in time series setting. The methodologies introduced in the proceeding paragraph are of great importance in the sense that the true underlying model has a sparse representation.

2.3.2 Penalized Least Square and Variable Selection

Variable selection is an important tool for the linear regression analysis. A popular method is the lasso estimator of Tibshirani (1996b), which can be viewed to simultaneously perform model selection and parameter estimation. Related literature includes bridge regression studied by Frank and Friedman (1993) and Fu (1998), the least angle regression of Efron et al. (2004) and adaptive lasso proposed by Zou (2006). Another remarkable example is a smoothly clipped absolute deviation (SCAD) penalty for variable selection proposed by Fan and Li (2001), they proved its oracle properties.

Let us start with consider model estimation and variable selection in a linear regression model,

$$y = X\beta + \varepsilon \tag{2.5}$$

where $y = (y_1, \ldots, y_n)^\top$ is an $n \times 1$ response vector, $X = (x_1, \ldots, x_p)$ is an $n \times p$ matrix with $x_j = (x_{1j}, \ldots, x_{nj})^\top$, $j = 1, \ldots, p$. $\hat{\beta}$ denotes the coefficient estimator produced by the fitting procedure. $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^\top$ is an $n \times 1$ vector of iid random errors.

The least square estimate is obtained via minimizing $\|y - X\beta\|^2$, where the ordinary least squares (OLS) gives nonzero estimates $\omega = X^\top y$ to all coefficients. Normally best-subset selection are implemented to select significant variables, but the traditional idea of best subset selection methods is computationally too expensive for many statistical applications. Therefore the penalized least square with a penalty term that is separable with respect to the estimated parameter $\hat{\beta}$ is considered here. In this paper we consider two popular estimators, lasso and SCAD.

The lasso is a regularization technique for simultaneous estimation and variable selection. Its estimate is defined as,

$$
\begin{aligned}
\hat{\beta}_{LASSO} &= \ \arg\min_{\beta} \|y - X\beta\|^2 + \lambda \sum_{j=1}^{p} |\beta_j| \\
&= \ \arg\min_{\beta} \|y - \sum_{j=1}^{p} x_j \beta_j\|^2 + \lambda \sum_{j=1}^{p} |\beta_j|
\end{aligned}
\tag{2.6}
$$

where λ is a tuning parameter. The second term in equation (2.3.2) is known as the ℓ_1-penalty. The idea behind lasso is the coefficients shrinks toward 0 as λ increase. When λ is sufficiently large, some of the estimated coefficients are exactly zero. The estimation accuracy comes from the trade-off between estimation variance and the bias.

To sum up, lasso is the penalized least square estimates with the ℓ_1 penalty in the general least squares and likelihood settings. Furthermore, the ℓ_2 penalty results in a ridge regression and ℓ_p penalty will lead to a bridge regression. In the setting of bridge regression, the penalty term is of ℓ_p norm.

We proceed to a brief introduction of the SCAD method. In the present context, the

SCAD estimator is given by,

$$\hat{\beta}_{SCAD} = \begin{cases} \text{sgn}(\omega)(|\omega| - \lambda)_+ & \text{when } |\omega| \leq 2\lambda \\ \dfrac{\{(a-1)\omega - \text{sgn}(\omega)a\lambda\}}{a-2} & \text{when } 2\lambda < |\omega| \leq a\lambda \\ \omega & \text{when } |\omega| > a\lambda \end{cases} \tag{2.7}$$

where $a > 2$ is an additional tuning parameter. The continuous differentiable penalty function for SCAD estimator is defined by,

$$p_\lambda'(\beta) = \lambda \left\{ I(\beta \leq \lambda) + \frac{(a\lambda - \beta)_+}{(a-1)\lambda} I(\beta > \lambda) \right\} \quad \text{for } a > 2 \text{ and } \beta > 0 \tag{2.8}$$

Both estimators are members of this penalized likelihood family. LASSO has better performance when the noise to signal ratio is large, but this approach creates bias. SCAD can generate variable selection results without generating excess biases.

2.3.3 Iterated-SIS Estimation

Fan and Lv (2008) proposed a SIS method to select important variables in ultra high-dimensional linear models. The proposed two-stage procedure can perform better than other methods in the sense of statistical learning problems. The SIS method is based on the concept of sure screening, is defined as the correlation learning which filters out the features that have weak correlation with the response. By sure screening, all the important variables survive after variable screening with probability tending to 1.

Fan et al. (2009) improve iterated-SIS to a general pseudo-likelihood framework by allowing feature deletion in the iterative process. Fan et al. (2010) further extend the SIS model and consider an independent learning by ranking the maximum marginal likelihood estimator or maximum marginal likelihood itself for generalised linear models. Here we combine the VAR(p) model and SIS algorithm to find out the key elements in a big system. The basic idea of SIS is introduced in the following.

Let $\omega = (\omega_1, \omega_2, \ldots, \omega_p)^\top$ be a p-vector that is obtained by componentwise regression, i.e.,

$$\omega = X^\top y \tag{2.9}$$

where y is n vector of response and X is a $n \times p$ data matrix. ω is a vector of marginal correlations of predictors with the response of predictors with the response variable, rescaled by the standard deviation of the response.

When there are more predictors than observation, LS (least square) estimator is noisy,

that's why ridge regression is considered. Let $\omega^\lambda = (\omega_1^\lambda, \ldots, \omega_p^\lambda)^\top$ be a $p-$vector obtained by ridge regression, i.e.,

$$\omega^\lambda = (X^\top X + \lambda I_p)^{-1} X^\top y \tag{2.10}$$

where $\lambda > 0$ is a regularization parameter. Obviously, when $\lambda \to 0$, $\omega^\lambda \to \hat{\beta}_{LS}$ and $\lambda \to \infty$, $\lambda\omega^\lambda \to \omega$. The componentwise regression is a specific case of ridge regerssion with $\lambda = \infty$.

The iterated-SIS algorithem applied for estimating the $VAR(p)$ model is,

1. Apply SIS for initial screening, reduce the dimensionality to a relative large scale d;

2. Apply a lower demensional model selection method (such as lasso, SCAD) to the sets of variables selected by SIS;

3. Apply SIS to the variables selected in the previous step;

4. Repeat step 2 and 3 until the set of selected variables do not decrease.

2.3.4 Connectedness Measure

The interactions between the variables, i.e., the directional connectedness measure $\theta_{ij}(q)$ is calculated by the generalized impulse response analysis using the sparse estimation of VAR(p) models. With interated-SIS algorithm to estiamte the sparse VARs structure, we can acquire its moving average (MA) transformation,

$$y_t = \sum_{i=0}^{\infty} B_i u_{t-i} \tag{2.11}$$

The coefficient matrices B_i obey $B_i = \sum_{j=1}^{iy} B_{i-j} A_j$, with $B_0 = I_K$ and $A_j = 0$ for $j > p$. $A_j, j = 1, 2, \ldots, p$ is the coefficient matrices of VAR(p) model.

Denoting the GFEVD by $\theta_{ij}(q)$,

$$\theta_{ij}(q) = \frac{\sigma_{jj}^{-1} \sum_{q=0}^{Q-1} \left(e_i^\top \hat{B}_q \Sigma e_j \right)^2}{\sum_{q=0}^{Q-1} \left(e_i^\top \hat{B}_q \Sigma \hat{B}_q^\top e_i \right)} \tag{2.12}$$

where q is the lag order, e_i is an $pK^2 \times 1$ selection vector with unity as its ith element and zeros elsewhere. $\Sigma = \mathrm{E}\left(u_t u_t^\top \right)$, is the covariance matrix of the non-orthogonalized VAR(p) in equation (2.1). σ_{jj} is the corresponding jth diagonal element of Σ. The

matrics $\hat{B}_l$ are the estimated coefficient matrices of equation (2.11).

To measure the persistent effect of a shock on the behavior of a series, we aim to acquire the population connectedness table 2.1, according to Diebold and Yılmaz (2014).

	x_1	x_2	...	x_n	From others
x_1	$\theta_{11}(q)$	$\theta_{12}(q)$	...	$\theta_{1n}(q)$	$\sum_{j=1}^{n} \theta_{1j}(q), j \neq 1$
x_2	$\theta_{21}(q)$	$\theta_{22}(q)$	...	$\theta_{2n}(q)$	$\sum_{j=1}^{n} \theta_{2j}(q), j \neq 2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	$\theta_{n1}(q)$	$\theta_{n2}(q)$	...	$\theta_{nn}(q)$	$\sum_{j=1}^{n} \theta_{nj}(q), j \neq n$
To others	$\sum_{i=1}^{n} \theta_{i1}(q), i \neq 1$	$\sum_{i=1}^{n} \theta_{i2}(q), i \neq 2$	...	$\sum_{i=1}^{n} \theta_{in}(q), i \neq n$	$\frac{1}{n} \sum_{i=1,j=1}^{n} \theta_{ij}(q), i \neq j$

Table 2.1. Connectedness table of interest.

The rightmost column gives the "from" effect of total connectedness, and the bottom row gives the "to" effect. In particular, the directional connectedness "from" and "to" associated with the forecast error variation θ_{ij} for specific power contract when the arising shocks transmit from one stock to the others. These two connectedness estimators can be obtained by adding up the row or column elements, the pairwise directional connectedness from j to i is given by,

$$C_{i \leftarrow j}^{Q} = \theta_{ij}(q) \tag{2.13}$$

The total directional connectedness "from" $C_{i \leftarrow \cdot}$ (others to i), "to" $C_{\cdot \leftarrow j}$ (j to others) and the corresponding net connectedness are defined as

$$C_{i \leftarrow \bullet} = \sum_{j=1}^{n} \theta_{ij}, i \neq j$$

$$C_{\bullet \leftarrow j} = \sum_{i=1}^{n} \theta_{ij}, i \neq j$$

$$C_i = C_{to} - C_{from} = C_{\bullet \leftarrow i} - C_{i \leftarrow \bullet} \tag{2.14}$$

2.4 Empirical Study

2.4.1 Data

As introduced in Section 2.2, EEX offers continuous trading data of Phelix Futures. The available load profiles are base, peak and off-peak. The available products with different maturities have five kinds, Day/Weekend Futures, Week Futures, Month Futures, Quarter Futures and Year Futures. Nevertheless the products of Day/Weekend Futures and Week Futures only have the off-peak load data, for all other contracts base and

peak only. Here we recall the underlying of the Phelix Futures data, the Phelix Base contract is average price of the hours 1 to 24 for electricity traded on spot market, while the Phelix Peak is the average price of the hours 9 to 20 for electricity traded on spot market. Therefore we involve the products of spot prices as well. The contracts of spot prices are diversified in Hours from 00-01h up to 23-24h, and in Blocks of Base Monthly, off-peak 01-08, off-peak 21-24, Peak Monthly. The dataset we constructed is provided by Bloomberg, we have 90 kinds of contracts in total. The time span is from 30.09.2010 to 31.07.2015. All the contracts are listed on Table 2.2 with detailed information in Table 2.3

Kind	Contract Types			
GI	GI1.Comdty	GI2.Comdty	GI3.Comdty	GI4.Comdty
	GI5.Comdty	GI6.Comdty	GI7.Comdty	
GT	GT1.Comdty	GT2.Comdty	GT3.Comdty	GT4.Comdty
	GT5.Comdty	GT6.Comdty	GT7.Comdty	
HP	HP1.Comdty	HP2.Comdty	HP3.Comdty	HP4.Comdty
	HP5.Comdty	HP6.Comdty		
GJ	GJ1.Comdty	GJ2.Comdty	GJ3.Comdty	GJ4.Comdty
	GJ5.Comdty	GJ6.Comdty	GJ7.Comdty	
HI	HI1.Comdty	HI2.Comdty	HI3.Comdty	HI4.Comdty
	HI5.Comdty	HI6.Comdty	HI7.Comdty	
NE	NE1.Comdty	NE2.Comdty	NE3.Comdty	NE4.Comdty
	NE5.Comdty	NE6.Comdty		
POA	POA1.Comdty	POA2.Comdty	POA3.Comdty	POA4.Comdty
	POA5.Comdty	POA6.Comdty	POA7.Comdty	
PDA	PDA1.Comdty	PDA2.Comdty	PDA3.Comdty	PDA4.Comdty
	PDA5.Comdty	PDA6.Comdty	PDA7.Comdty	
PBA	PBA1.Comdty	PBA2.Comdty	PBA3.Comdty	PBA4.Comdty
	PBA5.Comdty	PBA6.Comdty		
LPXBHR	LPXBHR01.Index	LPXBHR02.Index	LPXBHR03.Index	LPXBHR04.Index
	LPXBHR05.Index	LPXBHR06.Index	LPXBHR07.Index	LPXBHR08.Index
	LPXBHR09.Index	LPXBHR10.Index	LPXBHR11.Index	LPXBHR12.Index
	LPXBHR13.Index	LPXBHR14.Index	LPXBHR15.Index	LPXBHR16.Index
	LPXBHR17.Index	LPXBHR18.Index	LPXBHR19.Index	LPXBHR20.Index
	LPXBHR21.Index	LPXBHR22.Index	LPXBHR23.Index	LPXBHR24.Index
LPXBHxx	LPXBHBMI.Index	LPXBHOP1.Index	LPXBHOP2.Index	LPXBHPMI.Index
	LPXBHRBS.Index	LPXBHRPK.Index		

Table 2.2. The contracts we use for estimation.

Name	Detailed Info
GI1.Comdty - GI7.Comdty	Phelix Base Month Option and the respective next six delivery months
GT1.Comdty - GT7.Comdty	Phelix Base Quarter Option and the respective next six delivery quarters
HP1.Comdty - HP6.Comdty	Phelix Base Year Option and the respective next five delivery years
GJ1.Comdty - GJ7.Comdty	Phelix Peak Month Future and the respective next six delivery months
HI1.Comdty - HI7.Comdty	Phelix Peak Quarter Future, and the respective next six delivery quarters
NE1.Comdty - NE6.Comdty	Phelix Peak Year Future and the respective next five delivery years
POA1.Comdty - POA7.Comdty	Phelix Off-Peak Month Future and the respective next six delivery months
PDA1.Comdty - PDA7.Comdty	Phelix Off-Peak Quarter Future and the respective next six delivery quarters
PBA1.Comdty - PBA6.Comdty	Phelix Off-Peak Year Future and the respective next five delivery years
LPXBHR01.Index - LPXBHR24.Index	EEX Day-ahead Spot Market with Bid Type from 00-01 to 23-24h, e.g. LPXBHR14.Index is EEX Day-ahead Spot price based on bid hours from 13 -14.
LPXBHRxx.Index	EEX Day-ahead Spot Market with different Bid Types: LPXBHB.Index is Base Monthly 00-14h; LPXBHOP1.Index is Off Peak1 01-08h; LPXBHOP2.Index is Off Peak2 21-24h; LPXBHP.Index is Peak Monthly 08 - 20h; LPXBHRB.Index is Baseload; LPXBHRP.Index is Peakload.

Table 2.3. The detailed information of the selected contracts, summersized from the file "Products 2016" provided by European Energy Exchange AG.

To remove the redundant variable, we apply screening technique to select variables using the Phelix Futures consisting of different contracts and over different maturities. To implement the VAR model, first order difference of the data in Figure 2.2 is needed to transform non-stationary data to stationary time series. The contour plot of the constructed dataset are depicted in Figure 2.3.

In the market of Phelix Futures, final settlement at negative prices is also possible. There are some missing values after transforming the original data to stationary time series by first order difference. To deal with the missing data in dataset, some quick fixes such as mean-substitution may be fine in some cases. While such simple approaches usually introduce bias into the data, for instance, applying mean substitution leaves the mean unchanged (which is desirable) but decreases variance, which may be undesirable. In our paper, we impute missing values with plausible values drawn from a distribution using an approach proposed by Van Buuren and Oudshoorn (2000).
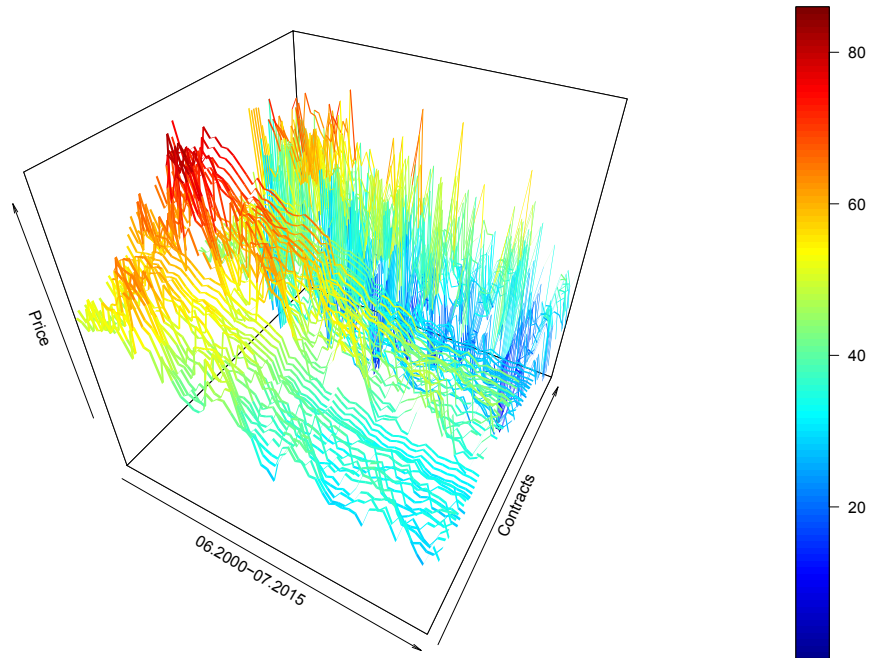
Figure 2.2. Ribbon plot of prices over 90 contracts
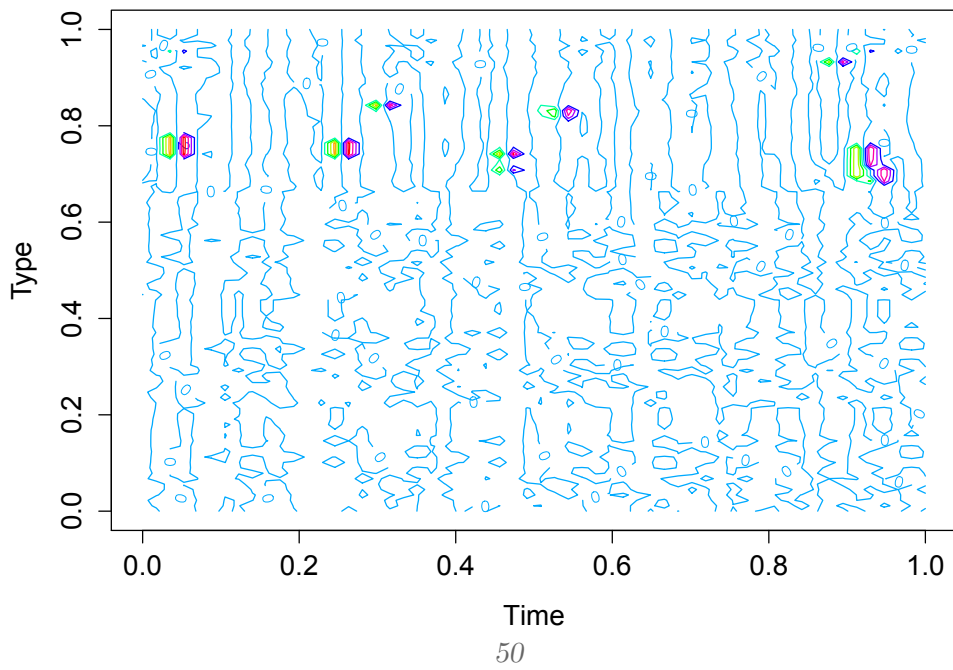
Figure 2.3. Contour plot of log return

Figure 2.4. Overview of dataset
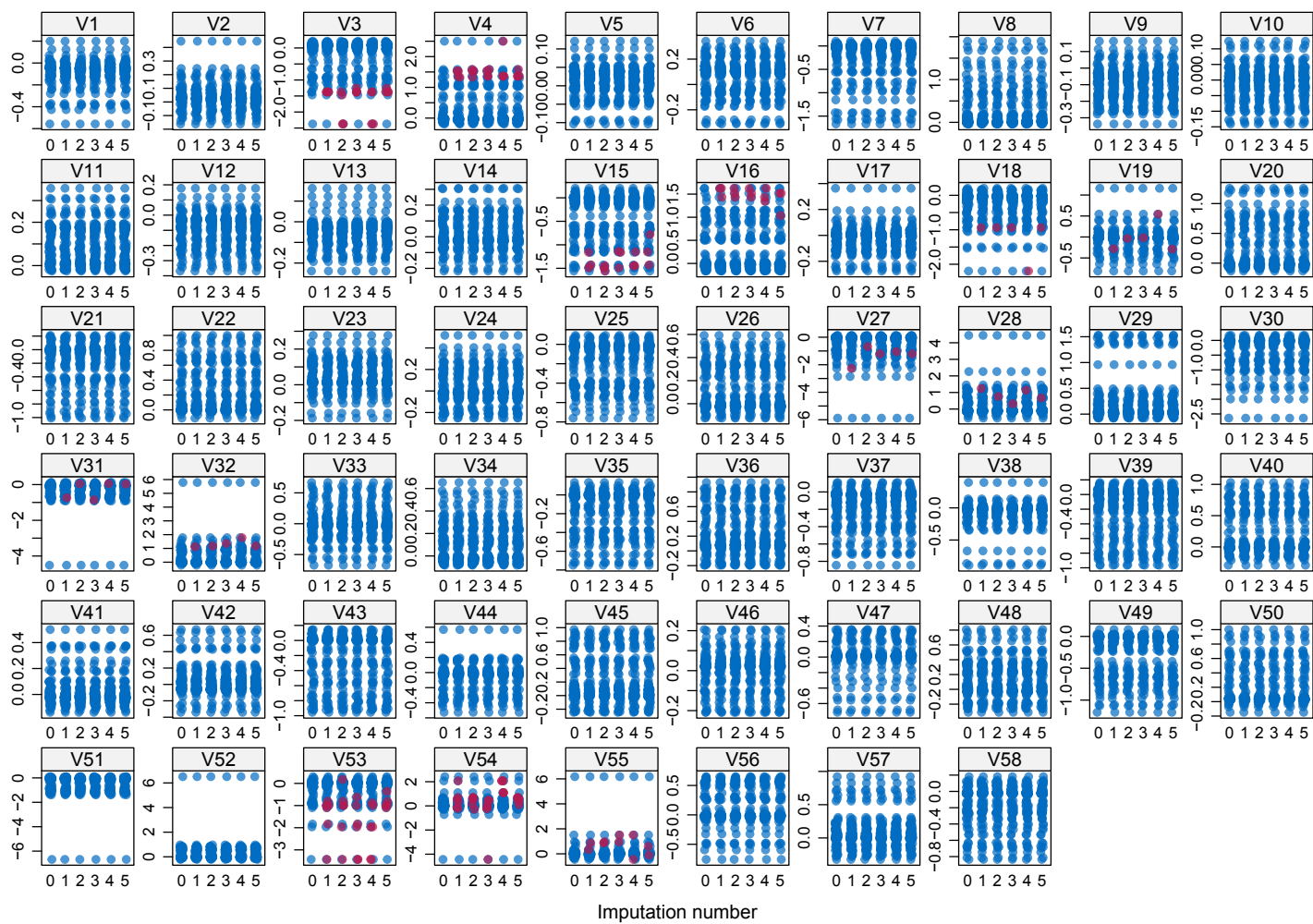
Figure 2.5. Pattern of imputed data.

The patterns of missing data for the original dataset and imputation dataset are compared in the Figure 2.5. The distributions of the variables are shown as individual points, the imputed data for each imputed dataset is showed in magenta while the density of the observed data is showed in blue. The distributions are expected to be similar based on the assumption. We can observe that the shape of the magenta points (imputed) matches the shape of the blue ones (observed). The matching shape tells us that the imputed values are indeed plausible values. With the imputed dataset of interest, we proceed to the estimation results derived from the iterated-SIS methodology.

2.4.2 Variable selection

With iterated-SIS, we can show the consistency of variable selection. Besides, with only a small set of relevant variables in the final model after iterated-SIS with lasso, the final estimate can be shown to be normally distributed.

Each curve corresponds to a variable. It shows the path of its coefficient against the $l1$-norm of the whole coefficient vector as λ varies. The estimation results are plotted in Figure 2.6, while the iterated-SIS-SCAD output is shown in Figure 2.7.

To estimate the accuracy of forecasting behaviour concerning the model selected, the observations y_t is partitioning into sample and presample values. In this paper we select the presample as 30.09.2010 - 28.11.2014, and the rest is treated as the sample from 31.12.2014 to 07.12.2015. The variable selected by the algorithm is,

$$
\begin{aligned}
Y_t &= 1.7730 + 0.4982Y_{t-92} + 0.3252Y_{t-104} + u_t \\
&= 1.7730 + [0.4982, B_0.3252][Y'_{t-92}, Y'_{t-104}]' + u_t
\end{aligned}
\tag{2.15}
$$

Once the VAR model has been estimated, we are interested in causal inference, forecasting and diagnosing the empirical model's dynamic behavior, i.e., impulse response functions (IRF) and forecast error variance decomposition. The impulse response analysis is based upon the Wold moving average representation of a VAR(p) process. It is used to investigate the dynamic interactions between the endogenous variables. The (i, j)th coefficients of the matrice

The tuning parameter used to adjust the penalty of glmnet estimation is Bayesian information criterion (BIC). As far as we acquired the estimated VAR model, the impulse response technique can be implemented to find the variables that accounts for most variation of dataset. For the continuous-time system, the impulse response analysis quantifies the reaction of every single variable in the model on an exogenous shock to the model. We assume all effects of omitted variables are assumed to be in the innovations. If important variables are omitted from the system,this may lead to major distortions in
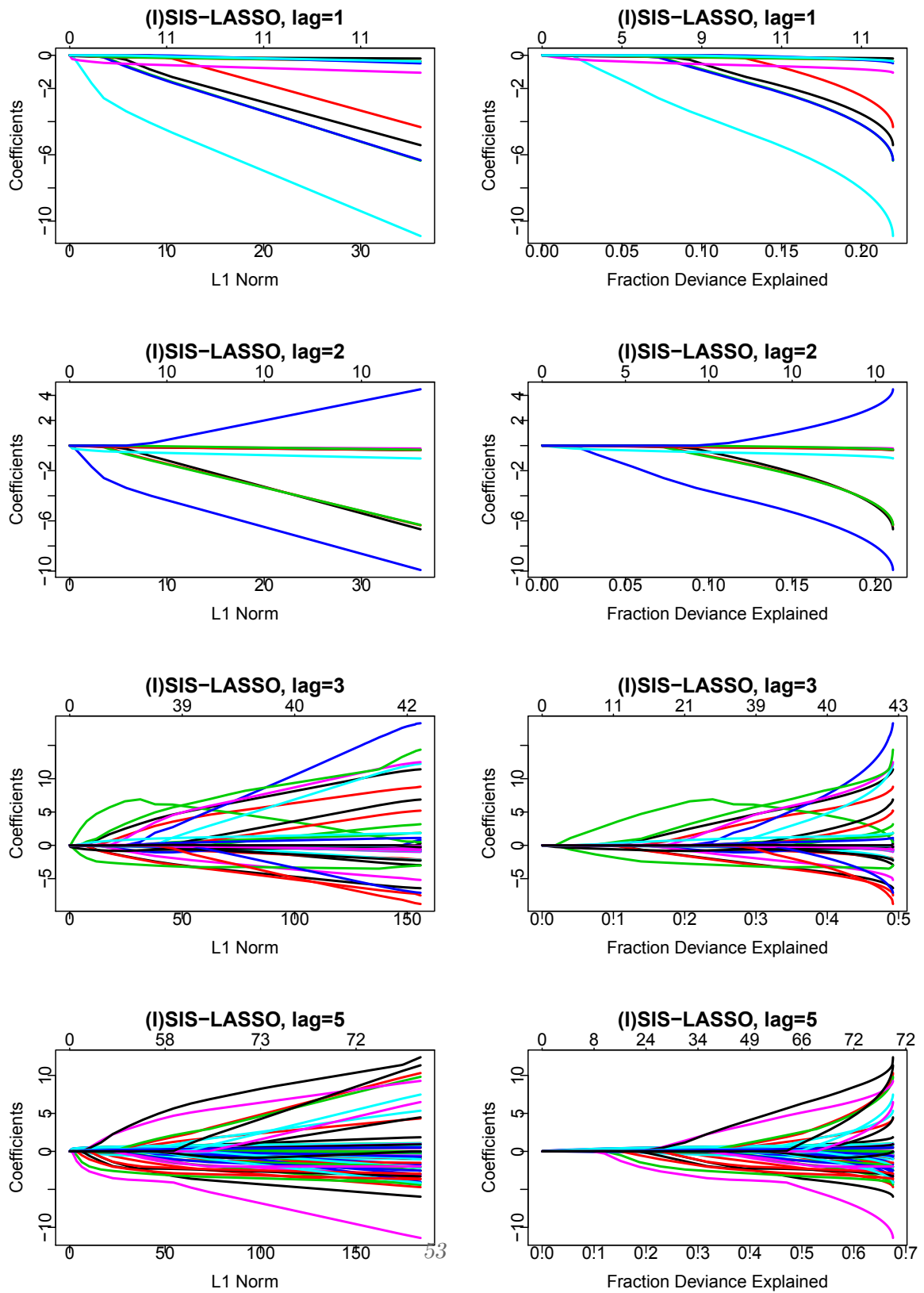
Figure 2.6. iterated-SIS-LASSO estimation results.
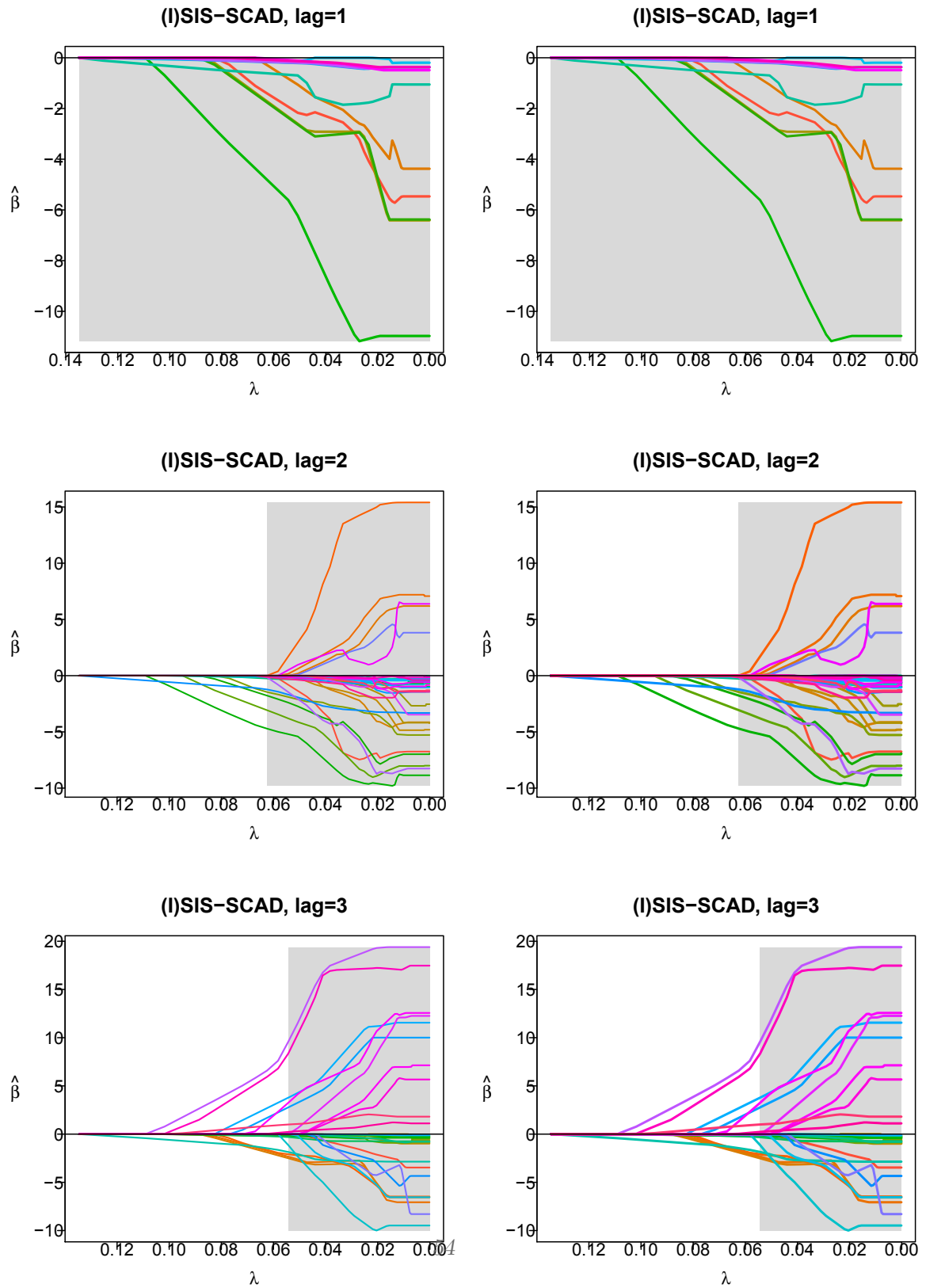
Figure 2.7. iterated-SIS-SCAD estimation results.

the impulse responses and makes them worthless for structural interpretations. Here we construct a systerm y_t with VMA representation, . . . The impulse response function for VAR model here is the orthogonal Impulse-Response function. The results derived by IR analysis are as follows,. . .

The tuning parameter used to adjust the penalty of glmnet estimation is Bayesian information criterion (BIC). As far as we acquired the estimated VAR model, the impulse response technique can be implemented to find the variables that accounts for most variation of dataset. For the continuous-time system, the impulse response analysis quantifies the reaction of every single variable in the model on an exogenous shock to the model. We assume all effects of omitted variables are assumed to be in the innovations. If important variables are omitted from the system, this may lead to major distortions in the impulse responses and makes them worthless for structural interpretations.

2.4.3 Lag Length Selection

We use different lags for estimating the VAR(p)model. The lag length for the VAR(p) model may be determined using model selection criteria. General approach is as follows,

- Fit the VAR(p) models with different lags $p = 0, \ldots, p_{max}$,

- Choose the value of p which minimizes some model selection criteria.

Model selection criteria for VAR(p) can be written as,

$$IC(p) = \log|\hat{H}(p)| + \varphi(K,p)c_T \tag{2.16}$$

where $\varphi(K,p)$ is a penalty function. c_T is a sequence indexed by the sample size T. The residual covariance matrix without a degrees of freedom correction is defined as,

$$\hat{H}(p) = \frac{1}{T}\sum_{t=1}^{T} u_t^\top u_t \tag{2.17}$$

Rewrite equation 2.16 with different penalty functions, the three most common information criteria are the Akaike (AIC), Schwarz-Bayesian (BIC) and Hannan-Quinn (HQ),

$$AIC = \log|\hat{H}(p)| + \frac{2}{T}pK^2 \tag{2.18}$$

$$HQ = \log|\hat{H}(p)| + \frac{2\log\log T}{T}pK^2 \tag{2.19}$$

$$BIC = \log|\hat{H}(p)| + \frac{\log T}{T}pK^2 \tag{2.20}$$

The model selection results are shown in Table 2.4.

Model	AIC	HQ(C)	BIC
iterated-SIS-lasso, $p = 1$	4.5686	4.7249	5.7864
iterated-SIS-lasso, $p = 2$	4.5006	4.6426	5.6076
iterated-SIS-lasso, $p = 3$	7.7034	8.3143	12.4637
iterated-SIS-lasso, $p = 5$	7.0839	8.1209	15.1652
iterated-SIS-SCAD, $p = 1$	4.5714	4.7277	5.7892
iterated-SIS-SCAD, $p = 2$	6.1043	6.1043	9.5782
iterated-SIS-SCAD, $p = 3$	7.2559	7.6820	10.5770

Table 2.4. The three most common information criteria: the Akaike (AIC), Schwarz-Bayesian (BIC) and Hannan-Quinn (HQ) are compared.

Lag	iterated-SIS-lasso	iterated-SIS-SCAD
$p = 1$	0.0697	0.0697
$p = 2$	0.0670	0.0701
$p = 3$	1.9598	0.1413
$p = 5$	0.1397	-

Table 2.5. MSE of out-of-sample forecasting during 31.12.2014 - 31.07.2015

Recall equation **??**, the VAR(p) model,

$$
\begin{aligned}
y_t &= \nu + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + u_t \\
&= \nu + (A_1, A_2, \ldots, A_p) \left(y_{t-1}^\top, y_{t-2}^\top, \ldots, y_{t-p}^\top \right)^\top + u_t
\end{aligned}
$$

We select the in-sample dataset as 30.09.2010-28.11.2014, the out-of-sample dataset used to measure model performance is from 31.12.2014 to 31.07.2015. We roll each model through the out-of-sample data set one observation at a time while each time forecasting the target variable one month ahead. By rolling window. The mean squared errors (MSE) for different models are calculated and reported in Table 2.5. VAR(2) with iterated-SIS-lasso technique performs best.

2.5 Network Analysis

We estimate VAR models using the iterated-SIS algorithm as described in previous section. Then we compute variance decompositions and corresponding connectedness measures at horizon H = 10, using the estimated VAR parameters.

2.5.1 Full-sample connectedness
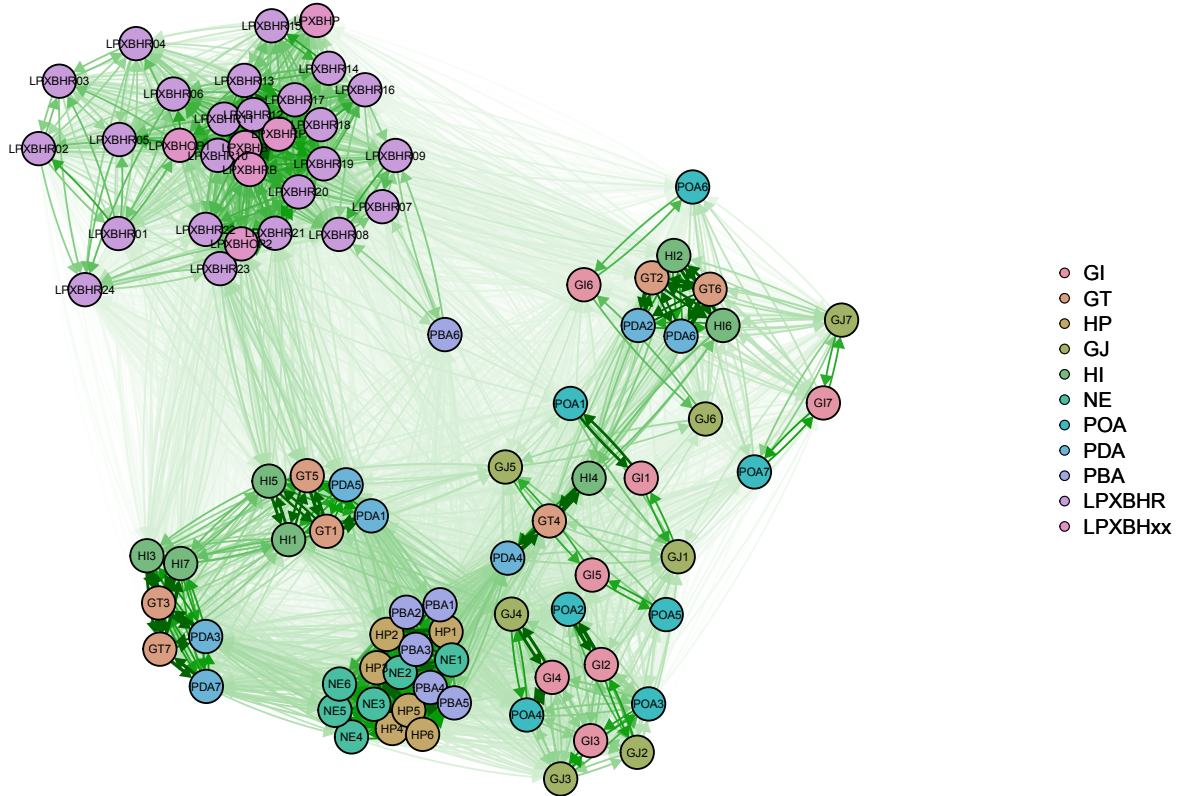
Figure 2.8. The graph for full-sample energy market network, across 11 different kinds with in total 90 contracts.

The graph of our full-sample energy market network is depicted in Figure 2.8. We observe the cluster phenomena in this graph, which motivates us to study the connectedness between contracts within and across 11 different kinds of energy contracts. In general, the contracts that belong to the same type tend to appear inside the same cluster. We find out several pairs of strong connections between different types of contracts, for example, the upper-left area reveals that the LPXBHR-type and LPXBHxx-type are massively connected. In addition, a cluster consisting of HP-type (Phelix Base Year Future), NE-type (Phelix Peak Year Future) and PBA-type (Phelix Off-Peak Year Future) indicates the closer relationship among these contracts, this implies the year futures are

closer to each other while the week future and quarter future remain distinct.

	GI	GT	HP	GJ	HI	NE	POA	PDA	PBA	LPXBHR	LPXBHxx	From
GI	1.75	0.46	0.86	1.46	0.44	0.71	1.48	0.48	0.80	0.15	0.15	8.74
GT	0.46	2.33	0.98	0.58	2.41	0.93	0.42	2.03	0.84	0.30	0.41	11.70
HP	0.73	0.84	5.27	0.63	0.51	4.78	0.70	1.21	4.19	0.06	0.03	18.95
GJ	1.46	0.58	0.73	1.80	0.63	0.64	1.09	0.51	0.65	0.13	0.13	8.36
HI	0.44	2.41	0.60	0.63	2.65	0.63	0.37	1.92	0.49	0.34	0.48	10.96
NE	0.61	0.79	4.78	0.55	0.54	5.09	0.55	1.05	3.34	0.08	0.12	17.52
POA	1.48	0.42	0.81	1.09	0.37	0.64	1.60	0.48	0.79	0.19	0.18	8.04
PDA	0.48	2.03	1.42	0.51	1.92	1.23	0.48	1.99	1.27	0.26	0.32	11.91
PBA	0.68	0.72	4.19	0.56	0.42	3.34	0.67	1.09	3.88	0.20	0.12	15.88
LPXBHR	0.80	1.24	0.50	0.67	1.35	0.50	0.90	1.12	1.01	7.86	9.81	25.79
LPXBHxx	0.13	0.35	0.03	0.11	0.41	0.12	0.16	0.28	0.13	2.70	3.86	8.28
To	9.03	12.18	20.17	8.59	11.65	18.62	8.41	12.16	17.38	12.28	15.63	146.12
Net	0.29	0.48	1.22	0.23	0.70	1.10	0.37	0.25	1.50	-13.50	7.35	

Table 2.6. Population connectedness table for 11 kinds of contracts.
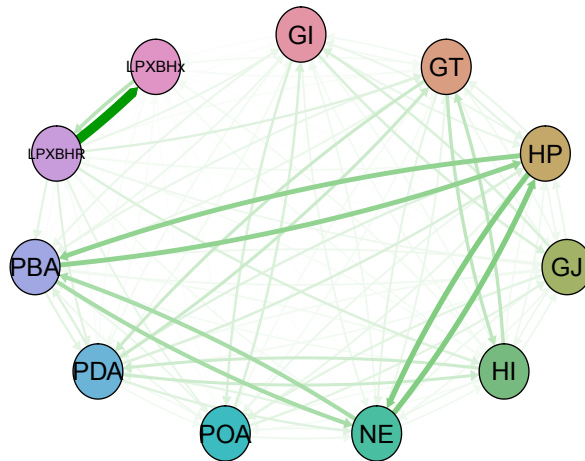
Figure 2.9. The graph for network across 11 different contract types.

Table 2.6 summerizes the full-sample connectedness of German power market across 11 different power contract types, with the own effects equal to the diagonal elements. We observe the contracts that have significant impacts are the "From" impact of "LPXBHR"-type and "To" impact of "HP"-type. The strongest link is the impact of LPXBHxx-type on the LPXBHR, however the inverse impact does not exist. Furthermore, HP-type contracts have stronger links from and to the other contract types. We can also conclude that the total impacts are mainly distributed among three types of contracts, i.e., HP-type, NE-type and PBA-type. The main risk of the whole market is mainly caused

by LPXBHxx-type, HP-type and NE-type. This is potentially interesting because, although HP-type, NE-type are important for the whole market as shown in Figure 2.9, their net connectedness are negligible, with 4.52% and 4.08% of the total market power contracts.

2.5.2 Determining Significant Market Component

In terms of magnitude for individual power contract reported in Table 2.6, the net directional connectedness from others is distributed rather tightly, in total 77.21% of "LPXBHR"-type and "LPXBHxx"-type.

We start with directional connectedness across 24 contracts of "LPXBHR"-type in Table **??**. Some blocks of high connectedness are detected, especially for the trading hours ranging from 9-13h and 16-19h. Table 2.8 provides the "from", "to" and "net" effects for 24 contracts in descending order of importance. Our finding clearly shows that, the impact from day-ahead spot power contracts that bidding between 9am and 13am are more relevant to the stability of the German power market.

The pairwise directional impacts between "LPXBHR"-type and "LPXBHxx"-type are plotted in Figure 2.10, we find a risk cluster of "LPXBHR10", "LPXBHR11", "LPXBHR11", "LPXBHR12", "LPXBHRP", "LPXBHRB" and "LPXBHB", the graph exhibits strong mutual links between some of the spot contracts. The "LPXBHB" (Base hours 00:00 - 24:00) has significant impacts on the spot contracts from hours 09 to 13, while the impacts from "LPXBHP" (Peak Hours 08:00 - 20:00) is negilitable. In addition, both "LPXBHRB" (Baseload) and "LPXBHRP" (Peakload) exhibits strong interconnectedness with the spot contracts from hours 09 to13. However only the "LPXBHRP" affects the spot prices from hours 16h to18h. We can infer that the Base spot contract "LPXBHB" is the largest risk contributor due to the strong linkage to the other spot contracts.

The network graph in Figure 2.11 illustrates the directional connectedness between "HP"-type and "NE"-type power contracts. Compared with Figure 2.8, we can see that, the links between these two types are obviously very strong. "NE" contracts are the Year Futures with maturities up to six years, the underlying of these contracts is the average price of hours 9 to 20 for electricity traded on the spot market. "HP" contracts are the European style options on the Phelix Base Future provided by EEX, the underlying of Phelix Base is the average price of the hours 1 to 24 for electricity traded on the spot market. It is calculated for all calendar days of the year as the simple average of the Auction prices for the hours 1 to 24 in the market area Germany / Austria. This figure shows the similar connectedness pattern between "HP1" and "NE1" contracts.

	LPXBHR01	LPXBHR02	LPXBHR03	LPXBHR04	LPXBHR05	LPXBHR06	LPXBHR07	LPXBHR08
LPXBHR01	1.00	0.71	0.52	0.35	0.53	0.43	0.13	0.14
LPXBHR02	0.42	0.59	0.47	0.36	0.31	0.18	0.06	0.07
LPXBHR03	0.27	0.41	0.54	0.42	0.36	0.23	0.12	0.12
LPXBHR04	0.19	0.32	0.41	0.53	0.38	0.25	0.21	0.15
LPXBHR05	0.25	0.24	0.30	0.33	0.47	0.37	0.19	0.15
LPXBHR06	0.18	0.12	0.18	0.18	0.34	0.41	0.22	0.23
LPXBHR07	0.10	0.08	0.13	0.20	0.22	0.29	0.45	0.40
LPXBHR08	0.10	0.08	0.13	0.14	0.18	0.30	0.45	0.54
LPXBHR09	0.11	0.08	0.14	0.15	0.18	0.35	0.58	0.70
LPXBHR10	0.25	0.23	0.23	0.33	0.49	0.66	0.46	0.49
LPXBHR11	0.27	0.28	0.26	0.35	0.50	0.61	0.42	0.46
LPXBHR12	0.26	0.29	0.23	0.30	0.42	0.48	0.28	0.31
LPXBHR13	0.23	0.30	0.24	0.30	0.39	0.45	0.21	0.27
LPXBHR14	0.16	0.15	0.17	0.23	0.26	0.30	0.13	0.13
LPXBHR15	0.22	0.18	0.17	0.19	0.22	0.27	0.14	0.13
LPXBHR16	0.07	0.06	0.10	0.13	0.19	0.34	0.36	0.41
LPXBHR17	0.13	0.13	0.18	0.22	0.35	0.51	0.29	0.36
LPXBHR18	0.13	0.11	0.15	0.20	0.32	0.47	0.24	0.27
LPXBHR19	0.15	0.11	0.12	0.24	0.32	0.45	0.33	0.30
LPXBHR20	0.10	0.07	0.08	0.17	0.24	0.38	0.32	0.28
LPXBHR21	0.06	0.07	0.05	0.12	0.17	0.28	0.27	0.25
LPXBHR22	0.11	0.13	0.11	0.12	0.19	0.31	0.23	0.26
LPXBHR23	0.10	0.12	0.06	0.07	0.11	0.18	0.12	0.14
LPXBHR24	0.00	0.00	0.01	0.01	0.01	0.00	0.00	0.00

	LPXBHR09	LPXBHR10	LPXBHR11	LPXBHR12	LPXBHR13	LPXBHR14	LPXBHR15	LPXBHR16
LPXBHR01	0.16	0.25	0.27	0.26	0.23	0.13	0.08	0.03
LPXBHR02	0.07	0.15	0.19	0.20	0.21	0.07	0.04	0.03
LPXBHR03	0.10	0.13	0.14	0.13	0.14	0.07	0.05	0.05
LPXBHR04	0.12	0.21	0.23	0.20	0.20	0.11	0.10	0.15
LPXBHR05	0.12	0.24	0.25	0.22	0.21	0.11	0.08	0.11
LPXBHR06	0.20	0.29	0.27	0.23	0.22	0.13	0.12	0.15
LPXBHR07	0.38	0.27	0.26	0.20	0.16	0.09	0.09	0.16
LPXBHR08	0.53	0.28	0.27	0.19	0.17	0.08	0.08	0.09
LPXBHR09	0.73	0.33	0.31	0.22	0.19	0.10	0.11	0.09
LPXBHR10	0.45	0.98	0.94	0.82	0.70	0.43	0.37	0.43
LPXBHR11	0.43	0.94	0.98	0.92	0.81	0.44	0.37	0.43
LPXBHR12	0.31	0.84	0.94	1.00	0.90	0.46	0.36	0.44
LPXBHR13	0.27	0.71	0.83	0.90	1.00	0.62	0.51	0.43
LPXBHR14	0.13	0.39	0.41	0.41	0.54	0.83	0.77	0.27
LPXBHR15	0.14	0.31	0.32	0.30	0.40	0.67	0.71	0.27
LPXBHR16	0.41	0.39	0.38	0.34	0.32	0.22	0.26	0.59
LPXBHR17	0.33	0.68	0.72	0.72	0.78	0.50	0.51	0.65
LPXBHR18	0.24	0.66	0.66	0.65	0.70	0.48	0.50	0.71
LPXBHR19	0.27	0.63	0.63	0.60	0.63	0.48	0.47	0.62
LPXBHR20	0.26	0.64	0.61	0.58	0.52	0.34	0.32	0.49
LPXBHR21	0.24	0.56	0.54	0.53	0.44	0.25	0.22	0.35
LPXBHR22	0.24	0.51	0.47	0.42	0.35	0.17	0.14	0.32
LPXBHR23	0.14	0.40	0.41	0.44	0.39	0.22	0.17	0.27
LPXBHR24	0.00	0.03	0.04	0.07	0.07	0.02	0.01	0.03

	LPXBHR17	LPXBHR18	LPXBHR19	LPXBHR20	LPXBHR21	LPXBHR22	LPXBHR23	LPXBHR24
LPXBHR01	0.13	0.13	0.15	0.10	0.06	0.11	0.10	0.00
LPXBHR02	0.08	0.07	0.08	0.07	0.14	0.17	0.22	0.38
LPXBHR03	0.09	0.08	0.07	0.06	0.11	0.12	0.15	0.42
LPXBHR04	0.16	0.16	0.18	0.15	0.17	0.16	0.18	0.36
LPXBHR05	0.17	0.15	0.16	0.14	0.19	0.18	0.22	0.50
LPXBHR06	0.22	0.20	0.21	0.19	0.25	0.24	0.27	0.55
LPXBHR07	0.20	0.19	0.23	0.22	0.20	0.18	0.16	0.17
LPXBHR08	0.20	0.16	0.18	0.18	0.21	0.20	0.17	0.20
LPXBHR09	0.24	0.17	0.20	0.19	0.17	0.18	0.10	0.00
LPXBHR10	0.67	0.65	0.61	0.63	0.55	0.50	0.39	0.04
LPXBHR11	0.71	0.65	0.62	0.60	0.53	0.46	0.40	0.05
LPXBHR12	0.72	0.65	0.61	0.58	0.53	0.42	0.43	0.08
LPXBHR13	0.78	0.70	0.63	0.52	0.44	0.35	0.39	0.08
LPXBHR14	0.42	0.41	0.41	0.30	0.22	0.16	0.21	0.07
LPXBHR15	0.38	0.37	0.36	0.25	0.18	0.13	0.15	0.05
LPXBHR16	0.46	0.46	0.42	0.36	0.28	0.26	0.20	0.02
LPXBHR17	1.00	0.93	0.74	0.62	0.46	0.36	0.27	0.02
LPXBHR18	0.93	1.00	0.83	0.68	0.46	0.38	0.30	0.02
LPXBHR19	0.74	0.83	1.00	0.81	0.59	0.45	0.46	0.05
LPXBHR20	0.62	0.68	0.81	1.00	0.82	0.64	0.56	0.08
LPXBHR21	0.46	0.46	0.59	0.82	1.00	0.82	0.71	0.27
LPXBHR22	0.36	0.38	0.45	0.64	0.82	1.00	0.81	0.23
LPXBHR23	0.27	0.30	0.46	0.56	0.71	0.81	1.00	0.38
LPXBHR24	0.03	0.03	0.06	0.08	0.22	0.19	0.31	0.82

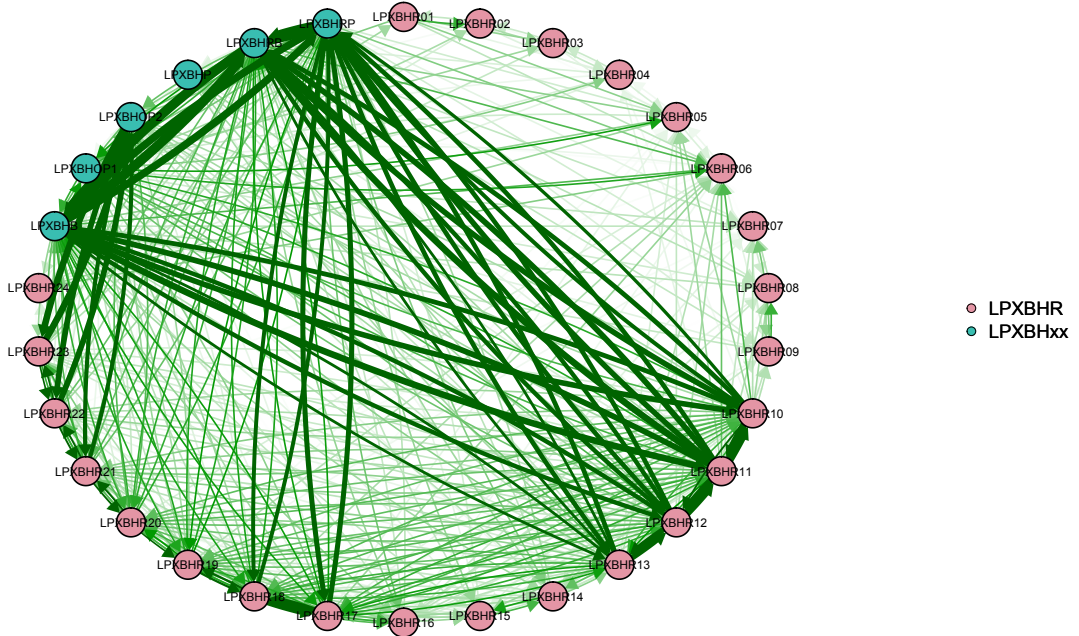**Table 2.7.** Population connectedness table for LPXBHR contracts.

Figure 2.10. The network graph for "LPXBHR"-type and "LPXBHxx"-type power contracts.

Figure 2.11. The network graph for "HP" and "NE"-type power contracts.

Power Contract	From	To	Net
LPXBHR11.Index	12.48	11.08	-1.40
LPXBHR10.Index	12.27	10.83	-1.45
LPXBHR12.Index	11.83	10.55	-1.27
LPXBHR13.Index	11.55	10.28	-1.27
LPXBHR17.Index	11.46	10.05	-1.41
LPXBHR19.Index	11.27	10.05	-1.23
LPXBHR18.Index	11.06	9.79	-1.27
LPXBHR20.Index	10.61	9.73	-0.88
LPXBHR21.Index	9.51	9.31	-0.19
LPXBHR22.Index	8.77	8.51	-0.26
LPXBHR23.Index	7.80	8.48	0.68
LPXBHR14.Index	7.48	8.16	0.68
LPXBHR16.Index	7.02	7.16	0.14
LPXBHR15.Index	6.51	7.15	0.64
LPXBHR01.Index	5.99	6.99	0.99
LPXBHR09.Index	5.64	6.54	0.90
LPXBHR06.Index	5.60	6.42	0.83
LPXBHR05.Index	5.35	6.30	0.94
LPXBHR04.Index	5.28	6.19	0.91
LPXBHR08.Index	5.13	5.65	0.52
LPXBHR07.Index	5.04	4.98	-0.06
LPXBHR02.Index	4.66	4.87	0.21
LPXBHR03.Index	4.36	4.85	0.49
LPXBHR24.Index	2.07	4.84	2.77

Table 2.8. Summary of "From", "To" and "Net" effects across "LPXBHR" contracts bidding from 0h to 24h.

2.6 Conclusion

By comparing our empirical findings, we could be able to identify the relevant risk drivers from the portfolio that are unknown to the power market investors. The selection of important market drivers via iterated-SIS algorithm enables us to investigate a ultra high-dimensional portfolio, since the the number of parameters to estimate increases quadratically in the number of variables included in the context of VAR estimation. In our paper, the network of interest follows the Dieboldt-Yilmaz tradition. With the wide range of power derivative contracts trading in the German electricity market, we are able to identify, estimated the risk contribution of individual power contract, this helps us to have a better understanding of the German power market functioning and environment.

3 Inflation Co-movement across Countries in Multi-maturity Term Structure: An Arbitrage-Free Approach

3.1 Introduction

Today most economists favour a low and steady rate of inflation because it facilitates real wage adjustments in the presence of downward nominal wage rigidity. Hence one of the major objectives of modern monetary policy is to bring inflation expectation under control, which is considered to be the first step in controlling inflation. Meanwhile, hedging the risk around the inflation forecast becomes more attractive in financial markets, as many investors rely on the stability and predictability of future inflation levels. Moreover, price stability is of immense importance to sustain social welfare, job opportunities and economic upturn. The objective of price stability refers to the general level of prices in the economy which implies avoiding both prolonged inflation and deflation. Inflation expectation that is involved in a contemporary macroeconomic framework anticipates future economic trends, will further affect monetary decisions. Since there is large demand on having reasonable estimates of inflation expectation levels, a large amount of literature has focused on analysing the government conventional and inflation-indexed bonds, which can implicitly provide a vast amount of information about the expectations of nominal and real interest rates obtained from the market. Such estimates are known to be an important complement to the estimates provided from the survey data. Despite the fact that inflation indexed bonds have been more frequently and widely issued in recent times, one would still have great difficulties in integrating the market information from multiple countries to get individual level estimates of the inflation expectation. The major problems lie in the relative short period of data availability and the existence of a lot of missing values. While the existing literature's focus is mainly on specific country, we would like to consider an estimation framework that allows us to analyse the co-movement of inflation expectation for multiple countries, and also provide the country specific estimates of inflation expectation(IE) and the inflation risk premium (IRP).

The starting point of our research is to analyse the break-even inflation rate (BEIR), which is known to be the difference between the yield on a nominal fixed-rate bond and the real yield on an inflation-linked bond of similar maturity and credit quality. The
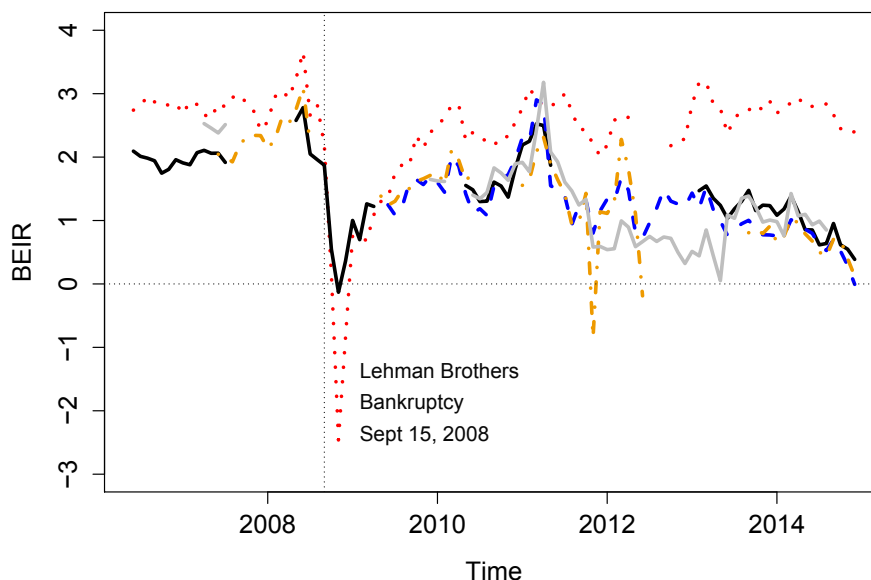
Figure 3.1. BEIR for five industrialized European countries - U.K.(red dotted line), Germany(blue dashed line), France(black line), Italy(orange dot-dashed line) and Sweden(grey line).

Q MTS_BEIR

BEIR can generally indicate how the inflation expectations are priced into the market. However they are not a perfect measure for IEs, as they may also encompass inflation risk premium, liquidity premium and "technical" market factors. We show the BEIR for five European countries - U.K., Germany, France, Italy and Sweden in Figure 3.1, which exhibits some degree of co-movement. This facilitates the following study in a multiple country framework. It is known that the euro-zone annual inflation rate was recorded at -0.2 percent in December of 2014 which matches, but are slightly higher than the overall BEIR shown in Figure 3.1. A fall in consumer prices first only appears since September 2009 due to a drop in energy costs. This motivates us to extract a joint time-varying structure of IEs estimated from individual (country-specific) BEIR.

The modelling of BEIR requires a model for the joint dynamics of the nominal and the real yields. For instance, Härdle and Majer (2014) investigated the yield curves using a Dynamic Semiparametric Factor Model (DSFM). To adopt a real time approach to help access the term structure of nominal and inflation-linked yields, in this study we consider a three-factor term structure model originally from Nelson and Siegel (1987). The attractiveness of factor models of the Nelson-Siegel type is due to its convenient

linear functions and good empirical performance. Diebold and Li (2006) extend the original Nelson-Siegel model to a dynamic environment. Theoretically, the Nelson-Siegel (NS) model does not ensure the absence of arbitrage opportunities, as shown by Bjork and Christensen (1999). Christensen et al. (2011) further develop the NS model to an AFNS model by imposing the arbitrage-free hypothesis, which reflects most of the real activities of financial markets. With arbitrage-free pricing, financial institutions apply arbitrage conditions to prices that are observable in financial markets in order to determine other prices that are not. The standard approaches for pricing forwards, swaps are all derived from such arbitrage arguments for both complete and incomplete markets. In our paper, we will use an AFNS model for the dynamics of the nominal and the real yield respectively, and combine the two models later on.

Based on the joint dynamics of the nominal and the real yields, a sizable amount of literature has analysed how to isolate IE and IRP from BEIR. Earlier work mainly focuses on U.K. data because the U.K. was one of the first developed economies to issue inflation-indexed bonds for institutional investors. Since the 1981 launch of the original U.K. index-linked gilts, various developments have occured in the international markets. Barr and Campbell (1997) estimated market expectations of real interest rates and inflation from observed prices of U.K. government nominal and inflation-linked bonds. Joyce et al. (2010) developed an affine term structure model to decompose forward rates to obtain IRP. Notably, Christensen et al. (2010a) used an affine arbitrage-free model of the term structure to decompose BEIR that captures the pricing of both nominal and inflation-indexed securities. A four-factor joint AFNS model was achieved by combining the AFNS models for nominal and inflation-linked yields, which proved efficient for fitting and forecasting analysis. Unlike Christensen et al. (2010a), we align the four factor models over different maturities to make the factors consistent over maturities.

With the AFNS model for the joint dynamics on hand, we proceed with our European country analysis. Most of the existing literature mentions little about the story of multiple countries. Diebold et al. (2008) are the first to consider a global multiple country model for nominal yield curves. There are a few European central bank reports, which focus on household and expert inflation expectation and the anchoring of inflation expectations in the two currency areas before and during the 2008 crisis.

Here we would like to look into five industrialised European countries by constructing a joint model of country-specific IEs. We construct an AFNS model in multi-maturity term structure for modelling nominal and inflation-indexed bonds simultaneously, we also propose a joint model of IE dynamics over European countries, which discovers the extracted common trend for IE is an important driver for each country of interest. Then we conduct an analysis to explore the estimated common factor by decomposing the variation into parts driven by common effect variation and macroeconomic effect variation.

The rest of the paper proceeds as follows. Section 2 estimates the joint AFNS model in multi-maturity term structure for estimating yields on nominal and inflation-linked

bonds and also covers the decomposition method of BEIR. In section 3, we discuss the econometric methodology used in the joint modelling of IE dynamics. The technical details are in the Appendix. The empirical results are shown in Section 4. Finally section 5 concludes.

3.2 Preliminary Analysis

In this section, we introduce the methodology to obtain the model-implied BEIR. Sub-section 3.2.1 briefly introduces the Nelson-Siegel model, and sub-section 3.2.2 constructs the joint AFNS structure for modelling nominal and inflation-indexed bonds. Sub-section 3.2.3 introduces the joint AFNS model across countries in a multi-maturity term structure. In the last sub-section 3.2.4, we describe the decomposition method of BEIR.

3.2.1 A factor model representation

The classic Nelson-Siegel (NS) yield curve model for fitting to static yield curves with simple functional form,

$$y(\tau) = \beta_0 + \beta_1 \left(\frac{1 - e^{-\lambda \tau}}{\lambda \tau} \right) + \beta_2 \left(\frac{1 - e^{-\lambda \tau}}{\lambda \tau} - e^{-\lambda \tau} \right) \tag{3.1}$$

where $y(\tau)$ is a zero-coupon yield with τ months to maturity, and $\beta_0, \beta_1, \beta_2$ and λ are parameters. This model is popular because it is simple and tractable. For a fixed value of parameter λ the remaining three βs can be estimated by the OLS method. Maturity τ determines the decay speed of parameters.

The aforementioned dynamic version of Nelson-Siegel (DNS) model enables institutional investors and policy makers to understand the evolution of the bond market over time, the DNS model can be written as,

$$y_t(\tau) = L_t + S_t \left(\frac{1 - e^{-\lambda \tau}}{\lambda \tau} \right) + C_t \left(\frac{1 - e^{-\lambda \tau}}{\lambda \tau} - e^{-\lambda \tau} \right) \tag{3.2}$$

where $y_t(\tau)$ denotes the continuously zero-coupon yield of maturity τ at time t. the time-varying parameters are defined as level L_t, slope S_t and curvature C_t. Such choice of the latent factors is motivated by principal component analysis, which gives us three principal components corresponding to the latent factors. For instance, the most variation of yields is accounted for by the first principal component - level factor L_t.

By incorporating the theoretical restriction of arbitrage-free, the AFNS model bridges the best of the Nelson-Siegel model and the AF model. Thus, the AFNS model consists of two equations by taking the structure of the DNS model and the real-world dynamics

(under P-measure) equation derived from the AF model respectively,

$$y_t(\tau) = X_t^1 + X_t^2 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + X_t^3 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) - \frac{A(\tau)}{\tau}$$

$$dX_t = K^P(\theta^P - X_t)dt + \Sigma dW_t^P$$

(3.3)

where $X_t^\top = (X_t^1, X_t^2, X_t^3)$ is a vector of latent factors, $\dfrac{A(\tau)}{\tau}$ is an unavoidable yield-adjustment term and only depends on maturity. K^P and θ^P correspond to drifts and dynamics terms, and both are allowed to vary freely. Σ is identified as a diagonal volatility matrix.

3.2.2 A joint factor model

The AFNS structure is a useful representation for term structure research. Christensen et al. (2010a) employed and conducted a separate AFNS model estimation of nominal and inflation-linked Treasury bonds respectively. Here we construct an extended AFNS structure for modelling nominal and inflation-indexed bonds simultaneously without exploring the estimated correlation of separate AFNS models.

The separate AFNS model of nominal and inflation-indexed type for a specific country i can be written as,

$$y_{it}^N(\tau) = L_{it}^N + S_{it}^N \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + C_{it}^N \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) - \frac{A_i^N(\tau)}{\tau}$$

$$y_{it}^R(\tau) = L_{it}^R + S_{it}^R \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + C_{it}^R \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) - \frac{A_i^R(\tau)}{\tau}$$

To explore the relationship between nominal and inflation-indexed bond yields within a country, we need combine two types and model them jointly.

To work with a simplified version of the yield curve, we assume the correlation between the latent factors of nominal and inflation-indexed bonds as follows,

$$S_{it}^R = \alpha_i^S S_{it}^N$$

$$C_{it}^R = \alpha_i^C C_{it}^N$$

(3.4)

The assumption will be justified by the performance of the joint model illustrated in

sub-section 3.4.2. The yield curve of the joint AFNS model is:

$$
\begin{pmatrix} y_{it}^N(\tau) \\ y_{it}^R(\tau) \end{pmatrix} = \begin{pmatrix} 1 & \dfrac{1-e^{-\lambda_i\tau}}{\lambda_i\tau} & \dfrac{1-e^{-\lambda_i\tau}}{\lambda_i\tau} - e^{-\lambda_i\tau} & 0 \\ 0 & \alpha_i^S\dfrac{1-e^{-\lambda_i\tau}}{\lambda_i\tau} & \alpha_i^C(\dfrac{1-e^{-\lambda_i\tau}}{\lambda_i\tau} - e^{-\lambda_i\tau}) & 1 \end{pmatrix} \begin{pmatrix} L_{it}^N \\ S_{it}^N \\ C_{it}^N \\ L_{it}^R \end{pmatrix}
$$
$$
+ \begin{pmatrix} \varepsilon_{it}^N(\tau) \\ \varepsilon_{it}^R(\tau) \end{pmatrix} - \begin{pmatrix} \dfrac{A_i^N(\tau)}{\tau} \\ \dfrac{A_i^R(\tau)}{\tau} \end{pmatrix} \tag{3.5}
$$

where y_{it}^N and y_{it}^R represent the nominal and inflation-linked yields for country i at time t. The real-world dynamics (under P-measure) equation takes the form of,

$$
dX_t = K^P(\theta^P - X_t)dt + \Sigma dW_t^P
$$

where the state variable $X_{it}^\top = \left(L_{it}^N, S_{it}^N, C_{it}^N, L_{it}^R\right)$ evolves dynamically.

3.2.3 Multiple Yield Curve Modelling

Diebold et al. (2008) extend the DNS model to a global version by modelling a potentially large set of country yield curves in a framework that allows for both global and country-specific factors. The model proposed here employs the joint AFNS model introduced in sub-section 3.2.2 and we further extend it to a multiple-maturity case.

For a specific country i, we first assume the state variable $X_{it}^\top$ introduced in sub-section 3.2.2 is a common state variable for the yield curves across different maturities. The multiple yield curve model may very well lead to efficient estimation. As in the following analysis examined, the small size of model residual represented in sub-section 3.4.4 accounts for the overall good fit of the model. More specifically, the joint AFNS

yield curve in multi-maturity term structure is,

$$
\begin{pmatrix} y_{it}^{N}(\tau_1) \\ y_{it}^{R}(\tau_1) \\ y_{it}^{N}(\tau_2) \\ y_{it}^{R}(\tau_2) \\ \vdots \\ y_{it}^{N}(\tau_n) \\ y_{it}^{R}(\tau_n) \end{pmatrix} =
\begin{pmatrix}
1 & \dfrac{1-e^{-\lambda_i\tau_1}}{\lambda_i\tau_1} & \dfrac{1-e^{-\lambda_i\tau_1}}{\lambda_i\tau_1}-e^{-\lambda_i\tau_1} & 0 \\
0 & \alpha_i^S\dfrac{1-e^{-\lambda_i\tau_1}}{\lambda_i\tau_1} & \alpha_i^C(\dfrac{1-e^{-\lambda_i\tau_1}}{\lambda_i\tau_1}-e^{-\lambda_i\tau_1}) & 1 \\
1 & \dfrac{1-e^{-\lambda_i\tau_2}}{\lambda_i\tau_2} & \dfrac{1-e^{-\lambda_i\tau_2}}{\lambda_i\tau_2}-e^{-\lambda_i\tau_2} & 0 \\
0 & \alpha_i^S\dfrac{1-e^{-\lambda_i\tau_2}}{\lambda_i\tau_2} & \alpha_i^C(\dfrac{1-e^{-\lambda_i\tau_2}}{\lambda_i\tau_2}-e^{-\lambda_i\tau_2}) & 1 \\
\vdots & \vdots & \vdots & \vdots \\
1 & \dfrac{1-e^{-\lambda_i\tau_n}}{\lambda_i\tau_n} & \dfrac{1-e^{-\lambda_i\tau_n}}{\lambda_i\tau_n}-e^{-\lambda_i\tau_n} & 0 \\
0 & \alpha_i^S\dfrac{1-e^{-\lambda_i\tau_n}}{\lambda_i\tau_n} & \alpha_i^C(\dfrac{1-e^{-\lambda_i\tau_n}}{\lambda_i\tau_n}-e^{-\lambda_i\tau_n}) & 1
\end{pmatrix}
\begin{pmatrix} L_{it}^{N} \\ S_{it}^{N} \\ C_{it}^{N} \\ L_{it}^{R} \end{pmatrix}
$$

$$
+ \begin{pmatrix} \varepsilon_{it}^{N}(\tau_1) \\ \varepsilon_{it}^{R}(\tau_1) \\ \varepsilon_{it}^{N}(\tau_2) \\ \varepsilon_{it}^{R}(\tau_2) \\ \vdots \\ \varepsilon_{it}^{N}(\tau_n) \\ \varepsilon_{it}^{R}(\tau_n) \end{pmatrix}
- \begin{pmatrix} \dfrac{A_i^N(\tau_1)}{\tau_1} \\ \dfrac{A_i^R(\tau_1)}{\tau_1} \\ \dfrac{A_i^N(\tau_2)}{\tau_2} \\ \dfrac{A_i^R(\tau_2)}{\tau_2} \\ \vdots \\ \dfrac{A_i^N(\tau_n)}{\tau_n} \\ \dfrac{A_i^R(\tau_n)}{\tau_n} \end{pmatrix}
\tag{3.6}
$$

where $y_{it}^{N}(\tau_n)$ and $y_{it}^{R}(\tau_n)$ represent the nominal and inflation-linked yields for country i at time t with maturity τ_n. The real-world dynamics equation is in the same form as before,

$$dX_t^P = K^P(\theta^P - X_t)dt + \Sigma dW_t^P$$

where state variables $X_{it}^{\top} = \left(L_{it}^N, S_{it}^N, C_{it}^N, L_{it}^R\right)$ evolves dynamically.

The methodology used to obtain the estimates of yield curves is the Kalman filtering technique. The technical details are in Appendix A.4.

3.2.4 BEIR decomposition

In order to find a more appropriate measure of expected inflation, it is necessary to understand the components of the bond yields intuitively and economically, for both nominal and inflation-linked types. A sizable amount of literature has adopted a parameterized approach for modelling the term structure of interest rates to estimate the IE and risk premia using data from both nominal and indexed bonds. Adrian and Wu (2009), Campbell and Viceira (2009), Pflueger and Viceira (2011) decomposed the yield of an inflation-linked bond into current expectation of a future real interest rate and a real interest rate premium. The yield on a nominal bond can be decomposed into parts of the yield on a real bond, expectations of future inflation and IRP. Therefore the spread between both yields, the BEIR, reflects the level of IE and IRP.

In the environment of an arbitrage-free model, there are no opportunities for investors to make risk-free profits, the bonds can be priced by basic pricing equations according to Cochrane (2005),

$$P_t = \mathsf{E}_t \left\{ \beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1} \right\} \tag{3.7}$$

where the price is denoted by P_t, the value c_t at time t has a payoff x_{t+1}, β is the discount factor. We break up the basic consumption-based pricing equation and get the stochastic discount factor (SDF) M_{t+1} at time $t+1$,

$$M_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)} \tag{3.8}$$

Then the prices of the zero-coupon bonds that pay one unit measured by the consumption basket at time t with maturity τ are formed as follows,

$$\begin{aligned} P_t^N(\tau) &= \mathsf{E}_t \left(M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N \right) \\ P_t^R(\tau) &= \mathsf{E}_t \left(M_{t+1}^R M_{t+2}^R \cdots M_{t+\tau}^R \right) \end{aligned} \tag{3.9}$$

where the nominal and the real (for inflation-linked bond) SDFs at time t are denoted by M_t^N and M_t^R. P_t^N and P_t^R represent the prices of nominal and real bonds respectively. The price of the consumption basket, which is known as the overall price level Q_t has the following link with SDFs given the assumption of no arbitrage,

$$\frac{M_t^N}{M_t^R} = \frac{Q_{t-1}}{Q_t} \tag{3.10}$$

Converting the price into the yield by the equation of $y_t(\tau) = -\frac{1}{\tau} \log P_t(\tau)$,

$$
\begin{aligned}
y_t^N(\tau) &= -\frac{1}{\tau} \log \mathsf{E}_t \left(M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N \right) \\
&= -\frac{1}{\tau} \mathsf{E}_t \left(\log M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N \right) - \frac{1}{2\tau} \mathsf{Var}_t \left(\log M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N \right) \\
y_t^R(\tau) &= -\frac{1}{\tau} \log \mathsf{E}_t \left(M_{t+1}^R M_{t+2}^R \cdots M_{t+\tau}^R \right) \\
&= -\frac{1}{\tau} \mathsf{E}_t \left(\log M_{t+1}^R M_{t+2}^R \cdots M_{t+\tau}^R \right) - \frac{1}{2\tau} \mathsf{Var}_t \left(\log M_{t+1}^R M_{t+2}^R \cdots M_{t+\tau}^R \right)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
y_t^N(\tau) - y_t^R(\tau) &= -\frac{1}{\tau} \mathsf{E}_t \left(\log \frac{M_{t+1}^N}{M_{t+1}^R} \frac{M_{t+2}^N}{M_{t+2}^R} \cdots \frac{M_{t+\tau}^N}{M_{t+\tau}^R} \right) + \frac{1}{2\tau} \mathsf{Var}_t \left(\log \frac{M_{t+1}^N}{M_{t+1}^R} \frac{M_{t+2}^N}{M_{t+2}^R} \cdots \frac{M_{t+\tau}^N}{M_{t+\tau}^R} \right) \\
&\quad - \frac{1}{\tau} \mathsf{Cov}_t \left(\log \frac{M_{t+1}^N}{M_{t+1}^R} \frac{M_{t+2}^N}{M_{t+2}^R} \cdots \frac{M_{t+\tau}^N}{M_{t+\tau}^R}, \log M_{t+1}^R M_{t+2}^R \cdots M_{t+\tau}^R \right)
\end{aligned}
$$

Given the log inflation is $\pi_{t+1} = \log \frac{Q_{t+1}}{Q_t}$ and the relationship between SDFs according to equation (10), the BEIR can be defined as,

$$
\begin{aligned}
y_t^N(\tau) - y_t^R(\tau) &= \frac{1}{\tau} \mathsf{E}_t \left(\log \pi_{t+1} \pi_{t+2} \cdots \pi_{t+\tau} \right) - \frac{1}{2\tau} \mathsf{Var}_t \left(\log \pi_{t+1} \pi_{t+2} \cdots \pi_{t+\tau} \right) \\
&\quad + \frac{1}{\tau} \mathsf{Cov}_t \left(\log \pi_{t+1} \pi_{t+2} \cdots \pi_{t+\tau}, \log M_{t+1}^R M_{t+2}^R \cdots M_{t+\tau}^R \right)
\end{aligned}
\tag{3.11}
$$

that is,

$$
BEIR_t(\tau) = y_t^N(\tau) - y_t^R(\tau) = \pi_t(\tau) + \eta_t(\tau) + \phi_t(\tau)
\tag{3.12}
$$

where $\pi_t(\tau)$ is the IE, $\eta_t(\tau)$ is the corresponding convexity effect and $\phi_t(\tau)$ is IRP. To link the $BEIR_t(\tau)$ with the estimated state variable mentioned in sub-section 3.2.3, we assume that the P-dynamics equations of the SDFs are,

$$
\begin{aligned}
\frac{dM_t^N}{M_t^N} &= -(r_t^N - r_{t-1}^N) dt - (\Gamma_t^N - \Gamma_{t-1}^N) dW_t^P \\
\frac{dM_t^R}{M_t^R} &= -(r_t^R - r_{t-1}^R) dt - (\Gamma_t^R - \Gamma_{t-1}^R) dW_t^P
\end{aligned}
\tag{3.13}
$$

where r_t is the stochastic risk-free rate and Γ_t represents the corresponding risk premium; their dynamics can be connected to the underling state variable $X_t^\top$ in equation (6),

details are given in Appendix A.5. Hence the dynamics of the overall price level is,

$$
\begin{aligned}
d\log\left(\frac{Q_{t-1}}{Q_t}\right) &= -(r_t^N - r_t^R)dt + (r_{t-1}^N - r_{t-1}^R)dt \\
d\log(Q_t) &= (r_t^N - r_t^R)dt
\end{aligned}
\tag{3.14}
$$

The IE is given by,

$$
\pi_t(\tau) = -\frac{1}{\tau}\log \mathsf{E}_t^P\left[\exp\left\{-\int_t^{t+\tau}(r_s^N - r_s^R)ds\right\}\right]
\tag{3.15}
$$

which can be solved by a system of ODEs with a Runge-Kutta method, see Appendix A.5. The convexity effect can be written as,

$$
\eta_t(\tau) = -\frac{1}{\tau}\mathsf{E}_t^P\left[\log\exp\left\{-\int_t^{t+\tau}(r_s^N - r_s^R)ds\right\}\right]
\tag{3.16}
$$

Then the IRP can be easily calculated out by equation (12).

3.3 Econometric Modelling of Inflation Expectation

Diebold et al. (2008) extended the dynamic Nelson-Siegel (DNS) model proposed by Diebold and Li (2006) to a global version by modelling a potentially large set of country yield curves in a framework that allows for both global and country-specific factors. As far as we have obtained the country-specific estimates of IE, we can tell a story of multiple countries in this section. We aim to investigate the country-specific idiosyncratic factors to load on a common time-varying factor and country-specific factors. The dynamics of an extracted common trend is also evaluated.

The model without a macroeconomic factor is structured as follows, the idiosyncratic factors $\hat{\pi}_{it}^e$ for each country i at time t, to load on a common time-varying latent factor Π_t,

$$
\hat{\pi}_{it}^e = m_i + n_i\Pi_t + \mu_{it}
\tag{3.17}
$$

The dynamics of common factor,

$$
\Pi_t = p + q\Pi_{t-1} + \nu_t
\tag{3.18}
$$

where m, n, p and q are unknown parameters. The errors μ_{it} and ν_{it} are assumed to be i.i.d white noise.

Since there is a dynamic interaction between macro-economy and the yield curve as evidenced by Diebold et al. (2006), and in Figure 3.1 we can observe that the decrease of the BEIR appears around 2012 due to the European sovereign debt crisis. A straightforward

extension of the joint modelling equation (16) is adding a proxy of the macroeconomic factor - default risk factor. The model with a macroeconomic factor is,

$$\hat{\pi}_{it}^e = m_i + n_i \Pi_t + l_i d_{it} + \mu_{it} \tag{3.19}$$

where d_{it} is the default risk factor varying over time and m, n, p, l and q are unknown parameters. The noises μ_{it} and ν_{it} are assumed to be i.i.d. The dynamics of the common factor is the same form as in (18).

3.4 Empirical Results

We now turn to the analysis of the results obtained using the model proposed. Sub-section 3.4.1 describes the data, we then discuss the fitting performance of the multiple yield curve modelling, setting up a brief discussion on the estimation results of the preliminary analysis in sub-section 3.4.2. Sub-section 3.4.3 establishes the estimated IE for each country, and sub-section 3.4.4 discusses the common trend extracted from the joint modelling of country-specific IEs. The results of the cross-sectional forecast is shown in sub-section 3.4.5.

3.4.1 Data

Monthly nominal and inflation-linked yield data of zero-coupon government bonds used for model estimation are taken from Bloomberg and Datastream. The research databases are supported by the Research Data Center (RDC) from the Collaborative Research Center 649, Humboldt Universität zu Berlin. The time series for a specific country are estimated using data from the same source. We consider five samples from the industrialized European countries - United Kingdom (U.K.), France, Germany, Italy and Sweden, all member states of European Union (EU). It should be noted that, even though two of the selected five European countries are outside the euro-zone - U.K. and Sweden, they have their own currencies therefore independent central banks and monetary policy, the inflation co-movement can be observed across the selected countries in sub-section 3.4.4. This also motivates our analysis to extract a joint time-varying structure of country-specific IEs.

The lack of short-maturity inflation-linked bonds of the sample countries indicates that inflation-linked yield at short-maturity tends to be less reliable and accessible, so the shortest maturities we could get access for each country are limited. We therefore selected three maturities for each country to ensure that enough observations are available. The sample period involves the subprime crisis in 2008 and is sightly different for each country due to the integrity of the data. The surfaces of the yield data are plotted in Figure 3.2. The blank areas in the Figure are for missing values and not zeros. Summary statistics are depicted in Table 3.1.
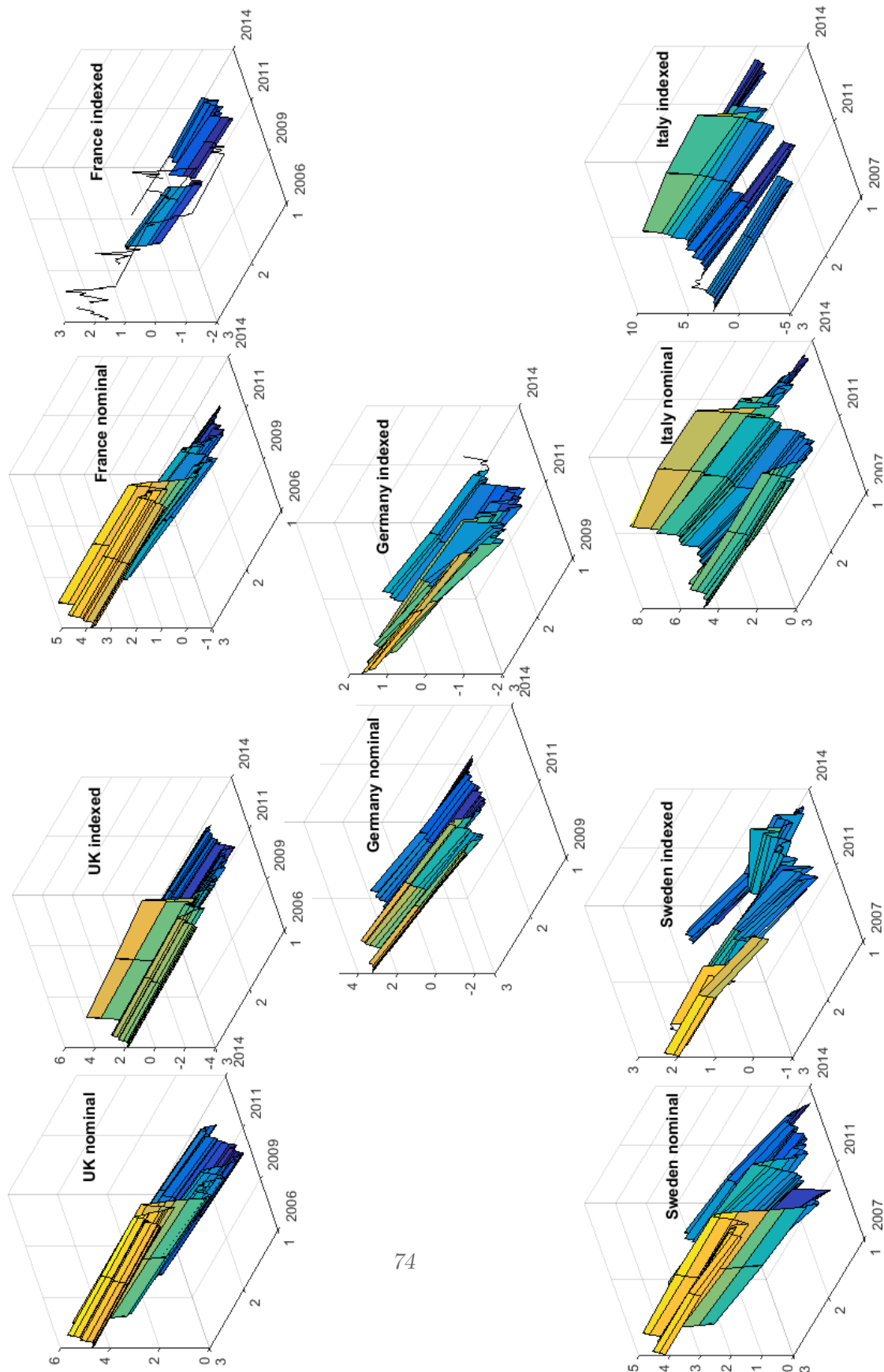
Figure 3.2. Term structures of nominal and inflation-linked bond yields across five European countries.

	Period	Maturity	Type	Min	Mean	Max	SD
UK	30.06.2006	3	nominal	0.16	2.20	5.69	1.73
			inflation-indexed	-2.87	-0.14	5.35	1.81
	—	4	nominal	0.35	2.44	5.62	1.59
			inflation-indexed	-2.62	-0.04	4.74	1.64
	31.12.2014	5	nominal	0.57	2.66	5.56	1.48
			inflation-indexed	-2.37	0.11	4.27	1.49
Sweden	30.04.2007	3	nominal	0.18	1.80	4.67	1.27
			inflation-indexed	-0.71	0.15	1.83	0.54
	—	5	nominal	0.58	2.33	4.71	1.11
			inflation-indexed	-0.84	0.51	2.33	0.79
	29.08.2014	10	nominal	0.88	2.59	4.61	1.03
			inflation-indexed	-0.30	0.98	2 29	0.64
France	30.06.2006	3	nominal	-0.02	1.86	4.74	1.46
			inflation-indexed	-1.19	0.43	2.75	1.29
	—	5	nominal	0.06	2.10	4.80	1.37
			inflation-indexed	-1.29	-0.40	1.06	0.60
	31.12.2014	10	nominal	0.18	2.34	4.80	1.28
			inflation-indexed	-1.09	1.03	2.66	1.03
Germany	30.06.2009	5	nominal	-0.07	0.90	2.38	0.77
			inflation-indexed	-1.39	-0.35	1.00	0.54
	—	7	nominal	0.05	1.37	2.85	0.86
			inflation-indexed	-1.16	0.05	1.36	0.65
	31.12.2014	10	nominal	0.39	1.94	3.29	0.84
			inflation-indexed	-0.53	0.33	1.67	0.67
Italy	30.06.2007	3	nominal	0.55	2.94	7.37	1.33
			inflation-indexed	-0.34	1.51	8.21	1.46
	—	5	nominal	0.95	3.53	7.54	1.20
			inflation-indexed	0.20	2.00	7.84	1.29
	31.12.2014	10	nominal	1.89	4.45	7.11	0.93
			inflation-indexed	1.02	2.77	6.72	1.06

Table 3.1. Descriptive statistics of the monthly bond yields data. SD is standard deviation.

3.4.2 Estimates of Yield Curve Modelling

We assess the performance of the previously discussed multiple yield curve model by conducting Kalman filtering, whose recursion is a set of equations allowing for an estimator to be updated once a new observation y_t is available; more technical details can be found in Appendix A.4. The objective is to see if the yields of nominal and inflation-linked bonds are suited to the model proposed in sub-section 3.2.3. Figure 3.3 shows the model residuals over different maturities for all five European countries.

While the model residuals have jumps for short periods, it is not entirely surprising as the jumps can be identified as the occurrence of extreme events. The outliers observed in the sub-figure of Italy happened to be the financial default crisis of Italy in 2012. We can also observe a jump around September 2008 for the U.K. and Sweden due to the well-known subprime crisis. Basically the overall small size of the model residuals indicates the good fit of the joint multiple yield curve model. The summary statistics of the model fit is represented in Table 3.2. Again, the value of the mean and RMSE of the model residuals are smaller due to the outliers in Figure 3.3.

The country-specific state variables are plotted in Figure 3.4, with four underlying latent factors $L_{it}^N, S_{it}^N, C_{it}^N, L_{it}^R$ presented. We observe that the level factors L_{it}^N, L_{it}^R are significantly positive, which in turn verifies the previously discussed choice of latent factors in sub-section 3.2.1. That is, the choice of the latent factors is motivated by principal component analysis, which gives us three principal components corresponding to the latent factors L_{it}, S_{it}, C_{it}. For instance, the most variation of yields is accounted for by the first principal component from principal component analysis - level factor L_t.

3.4.3 IE

We started by fitting the multiple yield curve modelling, where the model residuals are used to indicate the efficiency of the four-factor AFNS model over different maturities. We then conducted the decomposition of BEIR, as already described in sub-section 3.2.4, into parts of IE, the convexity effect and IRP to facilitate the following analysis. Figure 3.5 compares the estimated three-year and five-year forecasts of IE for each European country.

Figure 3.5 observes a decrease of the expected inflation for the U.K., which is also seemingly present in the other countries. To illustrates the similar trend among the five European countries, we present the country-specific three-year IE in Figure 3.6 to facilitate the following study of the similarity and difference among these five countries. Because the model-implied inflation expectation is on the three-year basis, the difference existing between the realized inflation level is understandable. We still find that the estimated IEs using the multiple AFNS model show similar trends as the realized levels. For instance, the realized inflation level of Sweden has two fluctuations in magnitude around the second half of the years 2008 and 2011, which is consistent with our finding.

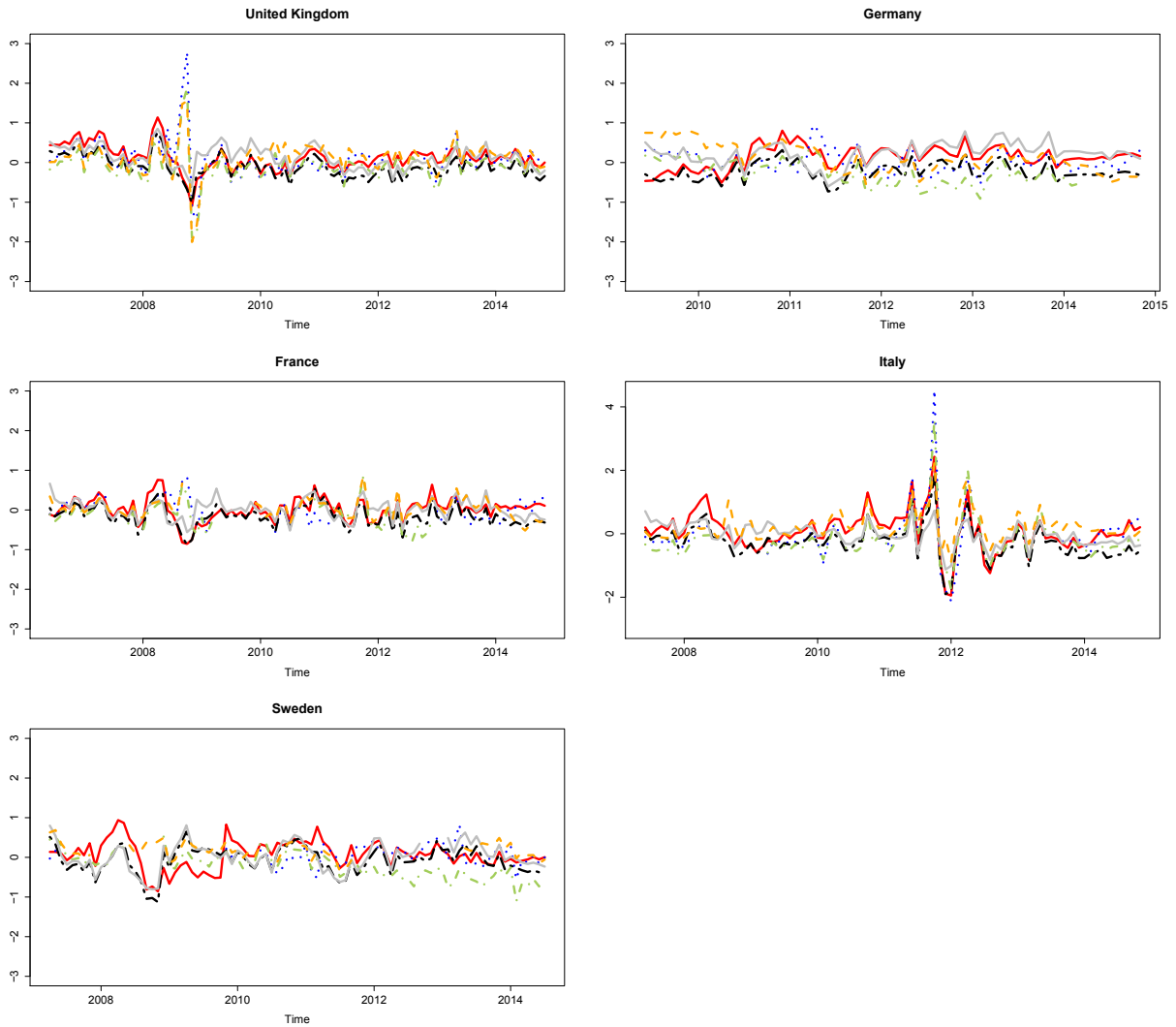**Figure 3.3.** The model residuals of multiple yield curve modelling over different maturities ($\tau_1 < \tau_2 < \tau_3$). The nominal type with τ_1 is the red line and the real type is the blue dotted line. The nominal and real types with τ_2 are the black long-dashed and green dot-dashed lines. For maturity of τ_3, the nominal type is grey and real type is an orange dashed line.

MTS_multi_modelres

	Maturity	Type	Mean	RMSE
UK	3	nominal	0.12	0.10
		inflation-indexed	0.10	0.22
	4	nominal	-0.10	0.08
		inflation-indexed	-0.09	0.18
	5	nominal	0.13	0.08
		inflation-indexed	0.12	0.17
Sweden	3	nominal	0.06	0.12
		inflation-indexed	0.06	0.04
	5	nominal	-0.07	0.11
		inflation-indexed	-0.20	0.51
	10	nominal	0.02	0.12
		inflation-indexed	0.18	0.03
France	3	nominal	0.01	0.08
		inflation-indexed	0.01	0.05
	5	nominal	-0.15	0.07
		inflation-indexed	-0.12	-0.06
	10	nominal	0.02	0.06
		inflation-indexed	0.04	0.05
Germany	5	nominal	0.14	0.08
		inflation-indexed	0.02	0.08
	7	nominal	-0.22	0.05
		inflation-indexed	-0.25	0.09
	10	nominal	0.23	0.08
		inflation-indexed	0.12	0.14
Italy	3	nominal	0.07	0.40
		inflation-indexed	0.01	0.50
	5	nominal	-0.25	0.27
		inflation-indexed	-0.18	0.40
	10	nominal	-0.02	0.13
		inflation-indexed	0.24	0.22

Table 3.2. Summary statistics of the model fit using the multiple yield curve model. RMSE is a root mean square error.
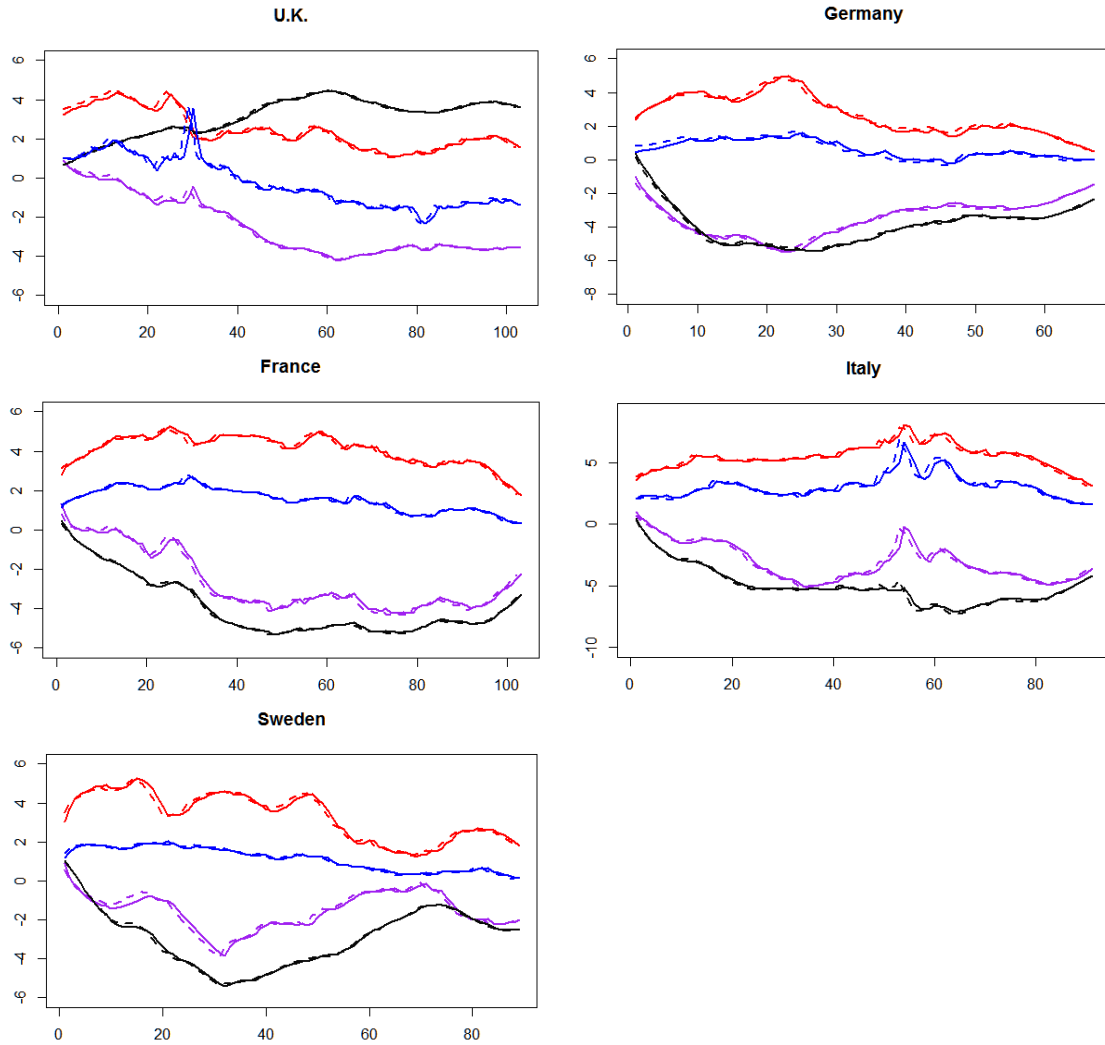
Figure 3.4. The estimated four latent factors of state variable $X_t = (L_{it}^N, S_{it}^N, C_{it}^N, L_{it}^R)$ for each European country - the nominal level factor L_{it}^N (red), the real level factor L_{it}^R (blue), the nominal slope factor S_{it}^N (purple) and the nominal curvature factor C_{it}^N (black). The preidicted state variables are presented as line type and the filtered state variables are dashed.

MTS_afns_uk, MTS_afns_de, MTS_afns_fr, MTS_afns_it, MTS_afns_sw
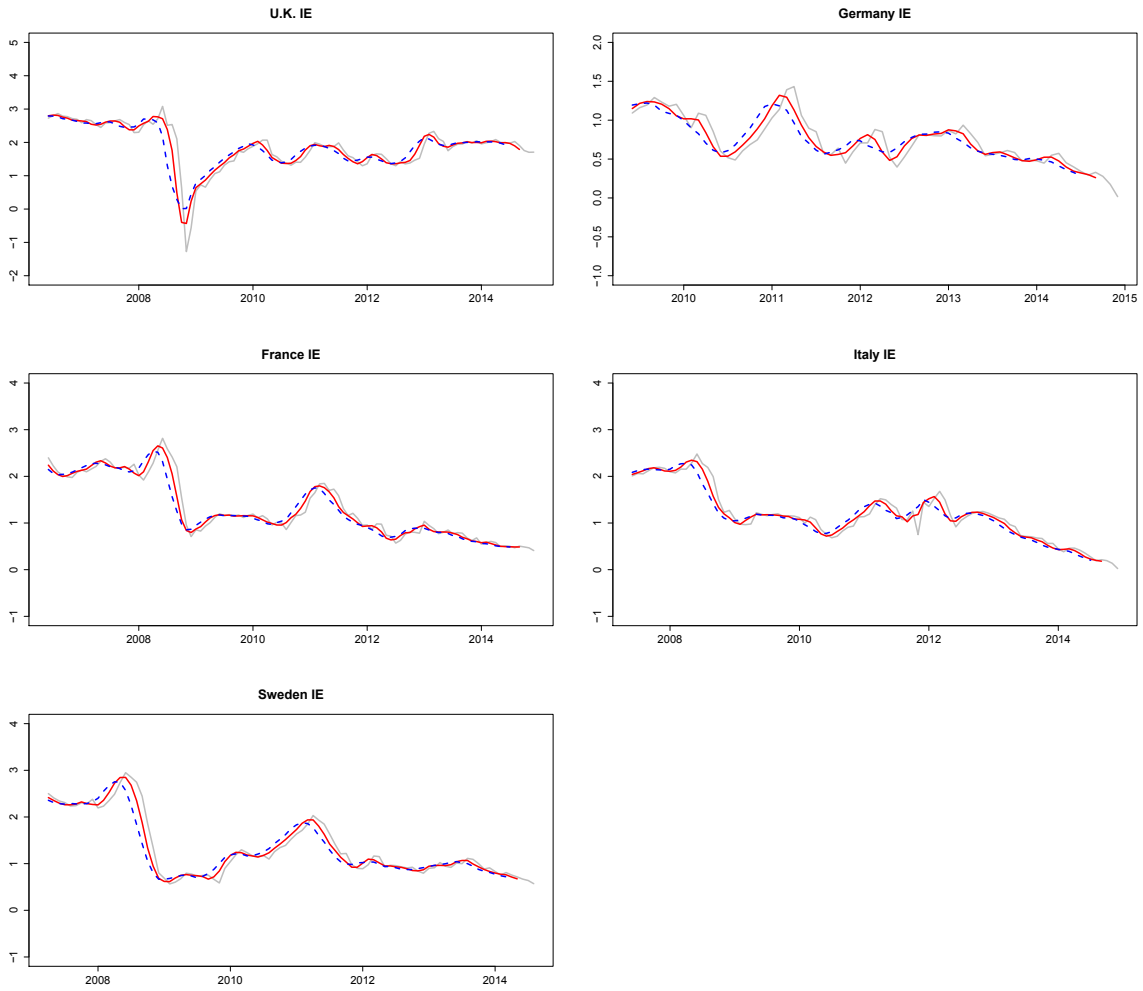
Figure 3.5. The model-implied IE for each European country. The 3-year IE is the red line and the 5-year IE is dashed blue.

MTS_expinf
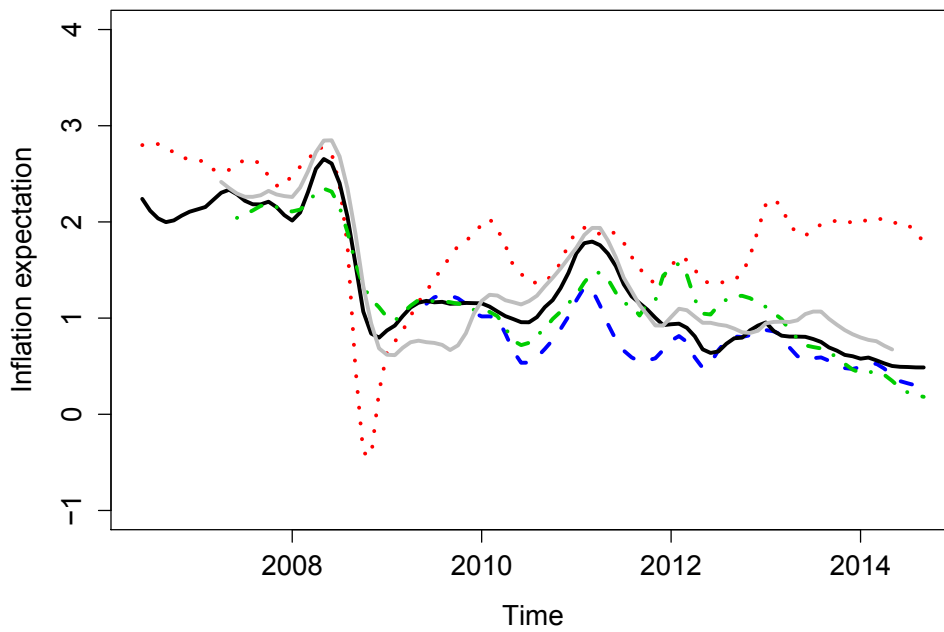
Figure 3.6. Model-implied inflation expectation for different countries - U.K.(red dotted line), Germany(blue dashed line), France(black line), Italy(green dot-dashed line) and Sweden(grey line)

MTS_expinf

3.4.4 Common Inflation Factor

Based on the methodology proposed in section 3.3, we can built the relationship among idiosyncratic countries to find out the similarities and differences for the model-implied IEs. The common inflation factor is extracted from the joint time-varying structure of IE dynamics and depicted in Figure 3.7. To be more specific, the estimated parameters
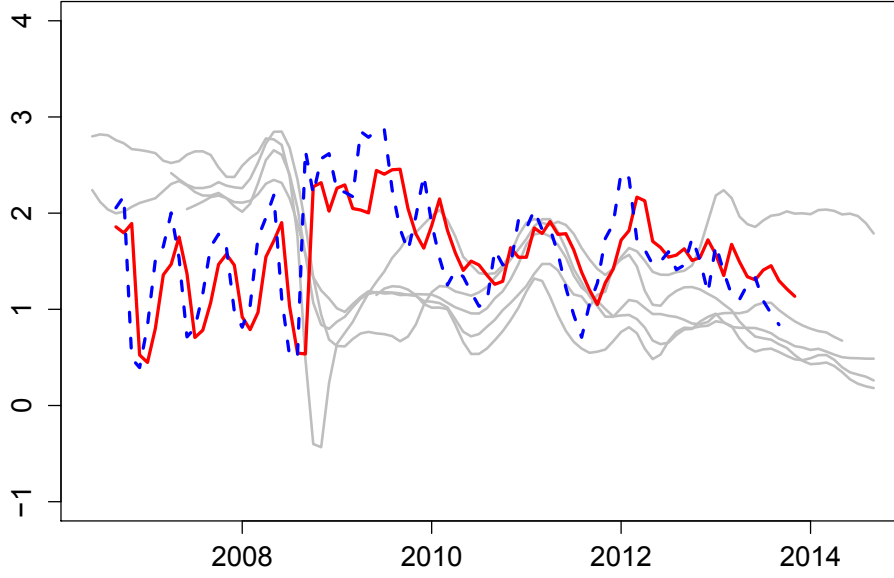
Figure 3.7. Common inflation factor in the red. The grey lines are the country-specific IEs. The predicted Π_t is the red line and the filtered Π_t is the blue dashed line.

MTS_comexpinf

for the joint modelling of IE dynamics are presented in the Table 3.3.

We conduct a variance decomposition to split the variation in model-implied IE into parts driven by the estimated common factor and the corresponding idiosyncratic factor. The variance equation is listed,

$$\mathsf{Var}(\pi_{it}^e) = \beta_i^2 \mathsf{Var}(\Pi_t) + \mathsf{Var}(\mu_{it}) \tag{3.20}$$

The variations explained by the common inflation factor are shown in the Table 3.4. The common inflation factor explains the least variation in the U.K., which can be

Country-specific equations		
UK	$\pi_{1t}^e(\tau) =$	$0.166 + 0.576\Pi_t$
France	$\pi_{2t}^e(\tau) =$	$-0.022 + 0.665\Pi_t$
Italy	$\pi_{3t}^e(\tau) =$	$-0.347 + 0.822\Pi_t$
Sweden	$\pi_{4t}^e(\tau) =$	$-0.057 + 0.665\Pi_t$
Germany	$\pi_{5t}^e(\tau) =$	$0.008 + 0.644\Pi_t$
Common Effect equation		
	$\Pi_t =$	$0.588 + 0.651\Pi_{t-1}$

Table 3.3. Estimates for the dynamics of IE.

	U.K.	France	Italy	Sweden	Germany
Common effect	24.91	30.66	40.32	30.65	29.32
Country-specific effect	69.34	50.69	69.35	58.50	70.68

Table 3.4. Variations explained in percentage

explained by the U.K. being outside the euro-zone and having its own currency, therefore being independent of European central bank and monetary policy. Even though Sweden is also outside the euro-zone, we find out it has closer relationships with other European countries compared with the U.K.. The international interaction among countries can also be observed through the estimation results.

To illustrate the efficiency of joint modelling of country-specific IEs, the model residuals are reported in Figure 3.8. The small size of the model residuals represents the overall good fit of the joint modelling of IE dynamics. However the model residuals are relatively high, around 2012, due to the European sovereign debt crisis including Italy's default. To eradicate this, we try to incorporate one more macroeconomic factor- default risk proxy to improve the model performance. By applying the method using the equations (18) and (19) proposed in section 3.3, we assess the joint model of IE dynamics by checking the model residuals in Figure 3.10 and the estimation results listed in Table 3.6.

The data we implement is the three-year CDS of Italy, the extracted common inflation factor derived from the joint model of IE dynamics with default proxy is presented in Figure 3.9, and successfully captures the decrease of IE caused by the subprime crisis. The estimated parameters for the joint modelling of inflation dynamics with macroeconomic factors - default proxy are presented in Table 3.5.

We also conduct a variance decomposition to split the variation in model-implied IE into parts driven by the estimated common factor Π_t and the default proxy d_t. The variations explained by the common inflation factor and default factor are reported in Table 3.6. The default factor explains the most variation in Italy and least variation
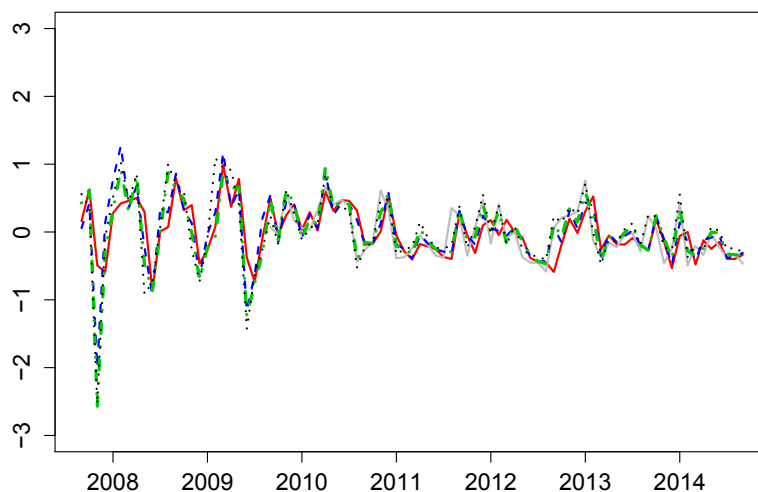
Figure 3.8. Model residual for modelling of inflation expectation dynamics over different countries - U.K.(red line), Germany(grey line), France(blue dashed line), Italy(black dotted) and Sweden(green dot-dashed).

<div align="right">⬛ MTS_comexpinf</div>

Country-specific equations		
UK	$\pi_{1t}^e(\tau) =$	$-0.358d_t + 0.798\Pi_t$
France	$\pi_{2t}^e(\tau) =$	$0.085d_t + 0.714\Pi_t$
Italy	$\pi_{3t}^e(\tau) =$	$1.078d_t + 0.531\Pi_t$
Sweden	$\pi_{4t}^e(\tau) =$	$-0.621d_t + 0.805\Pi_t$
Germany	$\pi_{5t}^e(\tau) =$	$0.045d_t + 0.700\Pi_t$
Common Effect equation		
	$\Pi_t =$	$0.382 + 0.976\Pi_{t-1}$

Table 3.5. Estimates for the dynamics of IE.
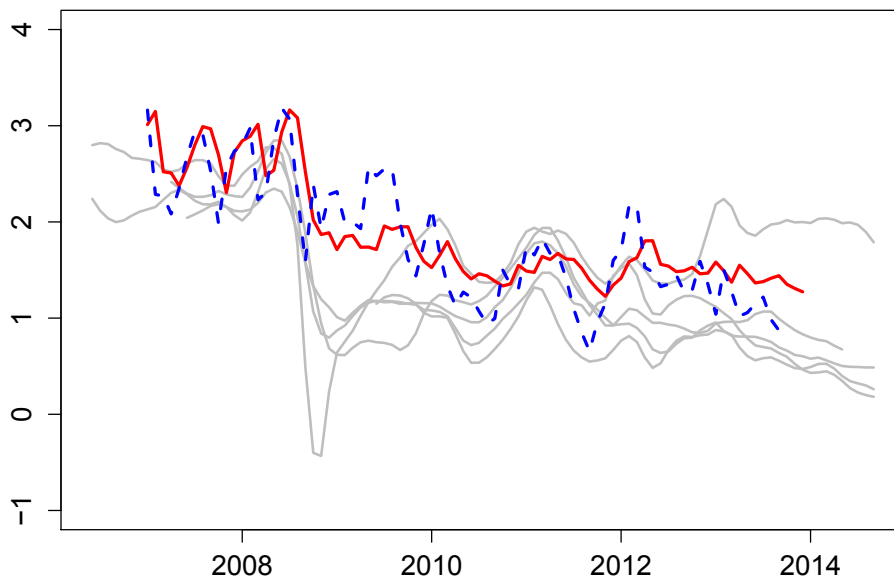
Figure 3.9. Common inflation factor with default proxy. The predicted estimation of common factor Π_t is the red line and the filtered Π_t is the blue dashed line.

MTS_comexpinf_cds

	U.K.	France	Italy	Sweden	Germany
Common effect	36.08	33.59	11.54	31.87	32.84
Country-specific effect	56.66	65.88	40.92	49.17	67.02
Default risk effect	7.26	0.53	47.55	18.96	0.14

Table 3.6. Variations explained in percentage

in Germany, which can be explained by the stability of the economy because Germany is generally considered to be the benchmark in the European financial system. It is generally known that the European sovereign debt crisis including Italy's default happened around 2012, the 47.55% variation explained by the default factor for Italy is understandable.

The model residuals are presented in Figure 3.10 to illustrate the efficiency of the joint model of IE dynamics with a default factor. Even though the model residuals with a default factor remain unchanged compared with Figure 3.8, we discover that the extracted common trend for IE is an important driver for each country of interest. Table 3.6 reports that the variation explained by the common inflation factor accounts for more than 30% of almost all the sample countries.

3.4.5 Forecast

Having obtained the estimation from the joint model of IE dynamics, we continue in this section by forecasting the common inflation factor. Figure 3.11 displays a forecast (in blue) containing 30 observations, that is, a two and a half year prediction. The predicted linear model we use involves trend and seasonality components. The confidence intervals are presented graphically and are shown at confidence levels of 80% and 95%.

Figure 3.12 clearly shows the difference among different measures of inflation. The real-time approach to measure IE proposed previously performs better than the other measures. A similar co-movement is seemingly present between the realized inflation level and the three-year IE estimated derived from our model. The 1 year and 2 year SPF (Survey Professional Forecast) data plotted in Figure 3.12 vary slightly over time therefore contains limited information of financial markets.

3.5 Conclusion

This study attempts to provide an additional measure of IE on the basis of financial markets. We firstly construct an AFNS model in multi-maturity term structure of modelling nominal and inflation-indexed bonds simultaneously. The performance of this multiple yield curve modelling was assessed by conducting Kalman filtering, whose recursion was a set of equations allowing for an estimator to be updated once a new
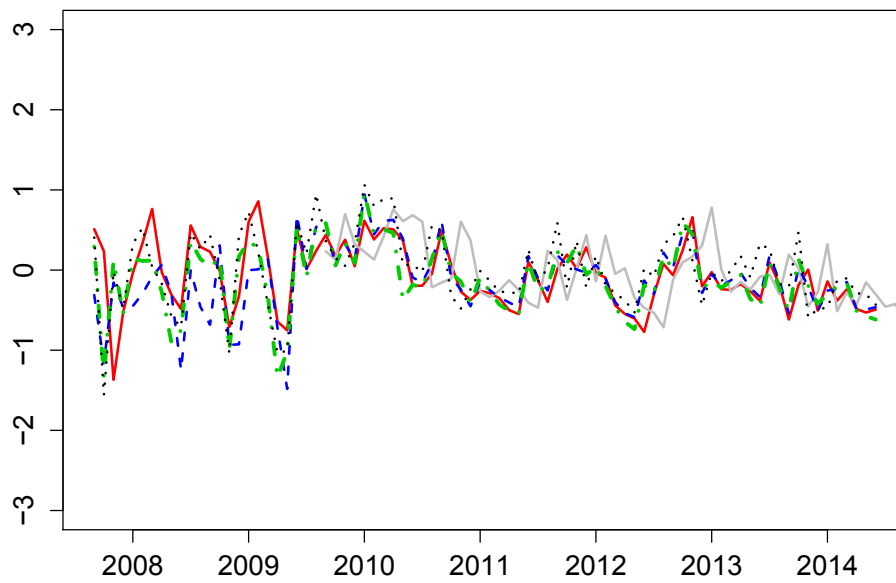
Figure 3.10. Model residual for modelling of inflation expectation dynamics with a default proxy factor over different countries - U.K.(red line), Germany(grey line), France(blue dashed line), Italy(black dotted) and Sweden(green dot-dashed).
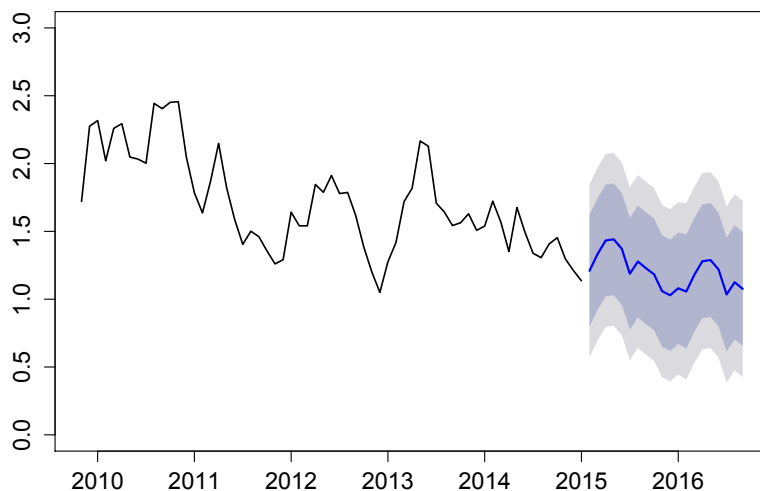
MTS_comexpinf_cds

Figure 3.11. The forecast of common inflation factor derived from the joint model of IE dynamics with default factor. The 80% and 95% confidence intervals are marked in the shaded area.

observation y_t is available. We then conducted the decomposition of model-implied BEIR into parts of IE, convexity effect and IRP to facilitate the modelling of joint structure of IE dynamics. The joint models of IE dynamics with, and without, macroeconomic factors indicated the extracted common inflation factor and was an important driver for each country of interest. Moreover, the model should lead to a better forecast in benchmark levels of inflation and give good implications for financial markets.
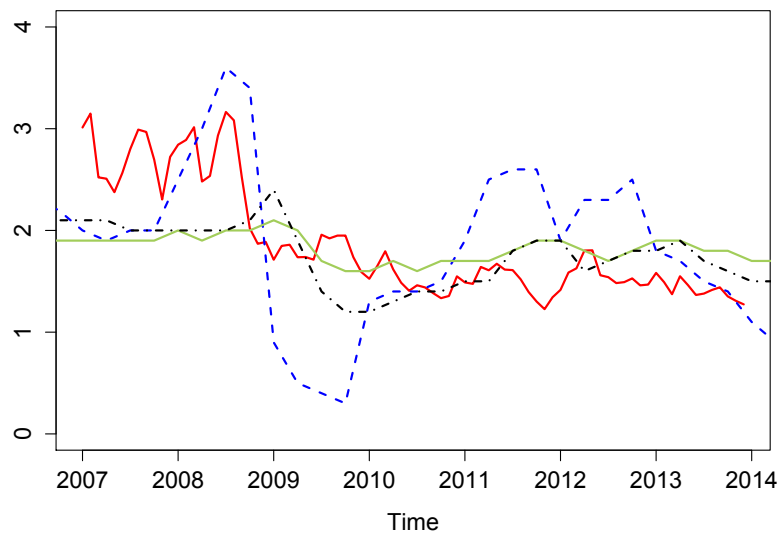
Figure 3.12. The comparison of different measures of inflation - the model-implied common inflation level (in red line), the observed inflation level (blue dashed line), the 1 year SPF forecast level of inflation (black dot-dashed) and the 2 year SPF forecast (in green).

A Appendix

A.1 Bootstrap-based multistep forecast methods

Here we describe the computation steps to yield bootstrapped GFEVD, more technical details can be found in Lanne and Nyberg (2016) and Terasvirta et al. (2010),

1. Pick a history ω_{t-1} which consists of all the information used to produce Y_t, and select a given forecast horizon h.

2. Randomly sample N_B vectors of shocks $(\delta_{1t}, \delta_{2t}, \dots, \delta_{Kt})^\top$ from the residuals of estimated model,

$$(\delta_{1t}, \delta_{2t}, \dots, \delta_{Kt})^\top \sim \hat{u}_t$$

$$\hat{u}_t = Y_t - \left(\hat{A}_1, \hat{A}_2, \dots, \hat{A}_p\right) \left(Y_{t-1}^\top, Y_{t-2}^\top, \dots, Y_{t-p}^\top\right)^\top = Y_t - g(Y_{t-1}) \qquad \text{(A.1)}$$

3. Compute conditional multistep forecast $\mathsf{E}(y_{t+l}|\omega_{t-1})$,

$$
\begin{aligned}
f_{t,0} &= g(Y_{t-1}) & \text{(A.2)} \\
f_{t,1} &= \mathsf{E}[Y_{t+1}|\omega_{t-1}] = \mathsf{E}[g(f_{t,0} + \hat{u}_t^\star)|\omega_{t-1}] \\
f_{t,2} &= \mathsf{E}[Y_{t+2}|\omega_{t-1}] = \mathsf{E}[g(f_{t,1} + \hat{u}_{t+1}^\star)|\omega_{t-1}] \\
&\dots
\end{aligned}
$$

4. Repeat steps 3 for all N_B vectors of estimated innovations with bootstrap methods, iterating on the estimated model,

$$
\begin{aligned}
fb_{t,1} &= \frac{1}{N_B}\sum_{j=1}^{N_B} g(f_{t,0} + \hat{u}_t^\star) & \text{(A.3)} \\
fb_{t,2} &= \frac{1}{N_B}\sum_{j=1}^{N_B} g(g(f_{t,0} + \hat{u}_t^\star) + \hat{u}_{t+1}^\star) \\
&\dots
\end{aligned}
$$

5. By the same logic, modify and repeat step 3 and 4 for all N_B vectors of shocks when the shock is given as $u_{jt} = \delta_{jt}$, this yields $\mathsf{E}(y_{t+l}|u_{jt} = \delta_{jt}, \omega_{t-1})$.

6. Plug in the GI function

$$GI(l, \delta_{jt}, \omega_{t-1}) = \mathsf{E}(y_{t+l}|u_{jt} = \delta_{jt}, \omega_{t-1}) - \mathsf{E}(y_{t+l}|\omega_{t-1}) \tag{A.4}$$

to obtain the relative contribution of a shock δ_{jt} to the i-th variable with horitzon h at time t,

$$\lambda_{ij,\omega_{t-1}}(h) = \frac{\sum_{l=0}^{h} GI(l, \delta_{jt}, \omega_{t-1})_i^2}{\sum_{j=1}^{K} \sum_{l=0}^{h} GI(l, \delta_{jt}, \omega_{t-1})_i^2}, \quad i, j = 1, \ldots, K \tag{A.5}$$

7. Repeat steps 2-6 for all histories.

8. Construct table 1.3 using averaged $\lambda_{ij,\omega_{t-1}}(h)$ generated from step 7.

We should note that if K is very large in empirical study, the denominator of equation (A.5) might be unnecessarily large due to accumulated noise caused by the large amount of irrelevant variables. Therefore it is one more step of prescreening is preferred to filter out less relevant variables.

A.2 Graphical models

Suppose a graph $\mathcal{G}$ consists of a set of vertices (nodes) $\mathcal{V}$, along with a set of edges $\mathcal{E}$ connecting pairs of $\mathcal{V}$. $\mathcal{E}$ is a subset of $\mathcal{V} \times \mathcal{V}$ consisting of ordered pairs of distinct vertices. The edges in a graph are parametrized by values or potentials that encode the strength of the conditional dependence between the random variables at the corresponding vertices.

Here we restrict our discussion to undirected graphical model where the edges have no direction. An edge is undirected if for two any nodes j and k, $(j, k) \in \mathcal{E}$ and $(k, j) \in \mathcal{E}$. The neighbors or adjacency set of a node j in the undirected graph $\mathcal{G}$ is denoted by

$$adj(\mathcal{G}, j) = \{k \in \mathcal{V}; (j, k) \in \mathcal{E} \quad \& \quad (k, j) \in \mathcal{E}\} \tag{A.6}$$

The pair $(\mathcal{G}, \mathcal{P})$ is called a graphical model with $\mathcal{P}$ as the probability distribution of the variables on the nodes

$$X = (X^{(1)}, X^{(2)}, \ldots, X^{(p)}) \sim \mathcal{P} \tag{A.7}$$

where $\mathcal{P}$ satisfies the pairwise Markov property w.r.t. $\mathcal{G}$ if for any pair of unconnected

vertices $(j, k) \notin \mathcal{E}$,

$$X^{(j)} \perp X^{(k)} | X^{(\mathcal{V} \backslash \{j,k\})} \tag{A.8}$$

A Gaussian Graphical Model (GGM) is defined if $\mathcal{P}$ is a normal distribution, i.e.

$$X = (X^{(1)}, X^{(2)}, \dots, X^{(p)}) \sim \mathcal{N}_p(0, \Sigma) \tag{A.9}$$

with mean $\mu = 0$ and positive definite $p \times p$ covariance matrix Σ. The edges in a GGM are given by the inverse of the covariance matrix due to

$$(j, k) \notin \mathcal{E} \iff X^{(j)} \perp X^{(k)} | X^{(\mathcal{V} \backslash \{j,k\})} \iff \Sigma_{jk}^{-1} = 0 \tag{A.10}$$

A.3 Time-varying network

Figure A.1 and A.2 show the time-varying network for T.

Figure A.3 and A.4 are of IBM.

Figure A.5 and A.6 are of JNJ.

Figure A.7 and A.8 are of PFE.

Figure A.9 and A.10 are of MRK.

Figure A.11 and A.12 are of JPM.

Figure A.13 and A.14 are of WFC.

Figure A.15 and A.16 are of Citigroup.

A.4 Estimation of multiple yield curve modelling

The analysis starts by introducing the yield-adjustment term proposed in the original AFNS model. Derived in an analytical form, the yield-adjustment term $\dfrac{A(\tau)}{\tau}$ with τ
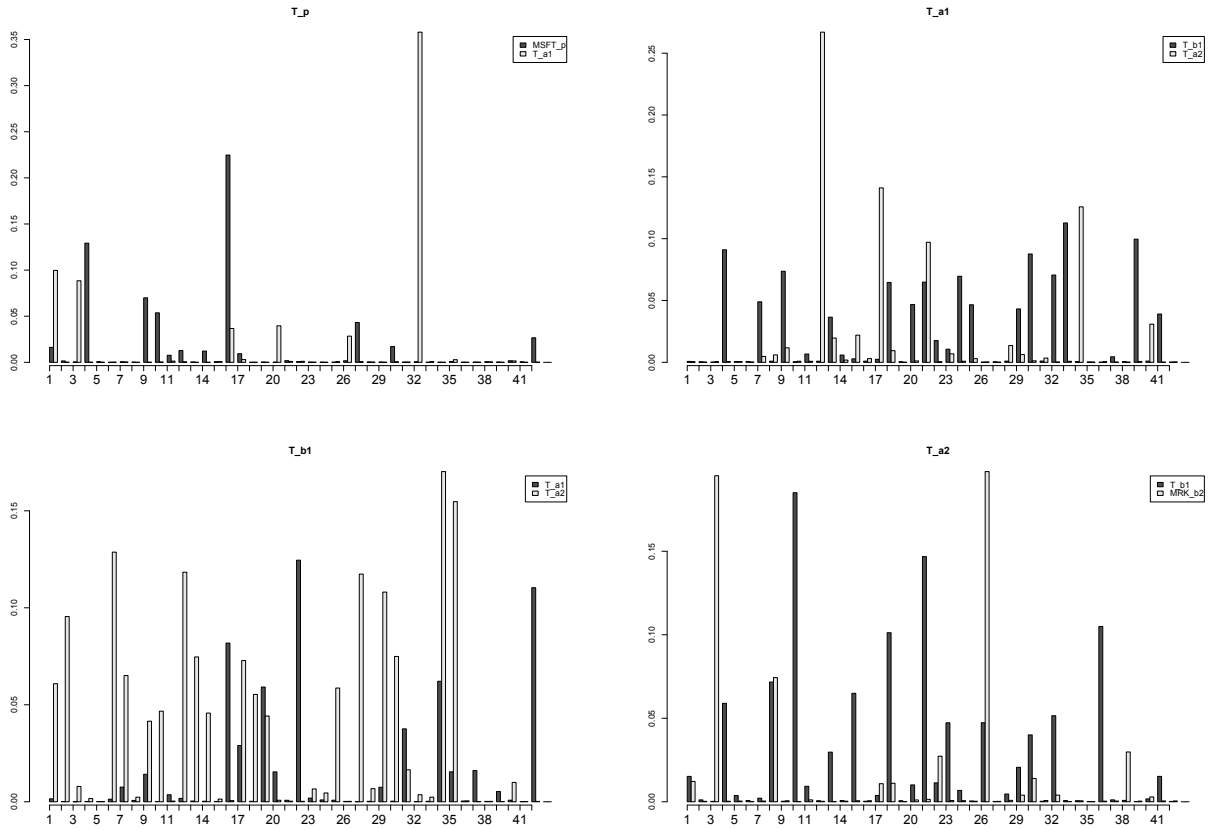
Figure A.1. The time-varying network of price factor, 1st level ask/bid size factor, 2nd level ask size factor of T.

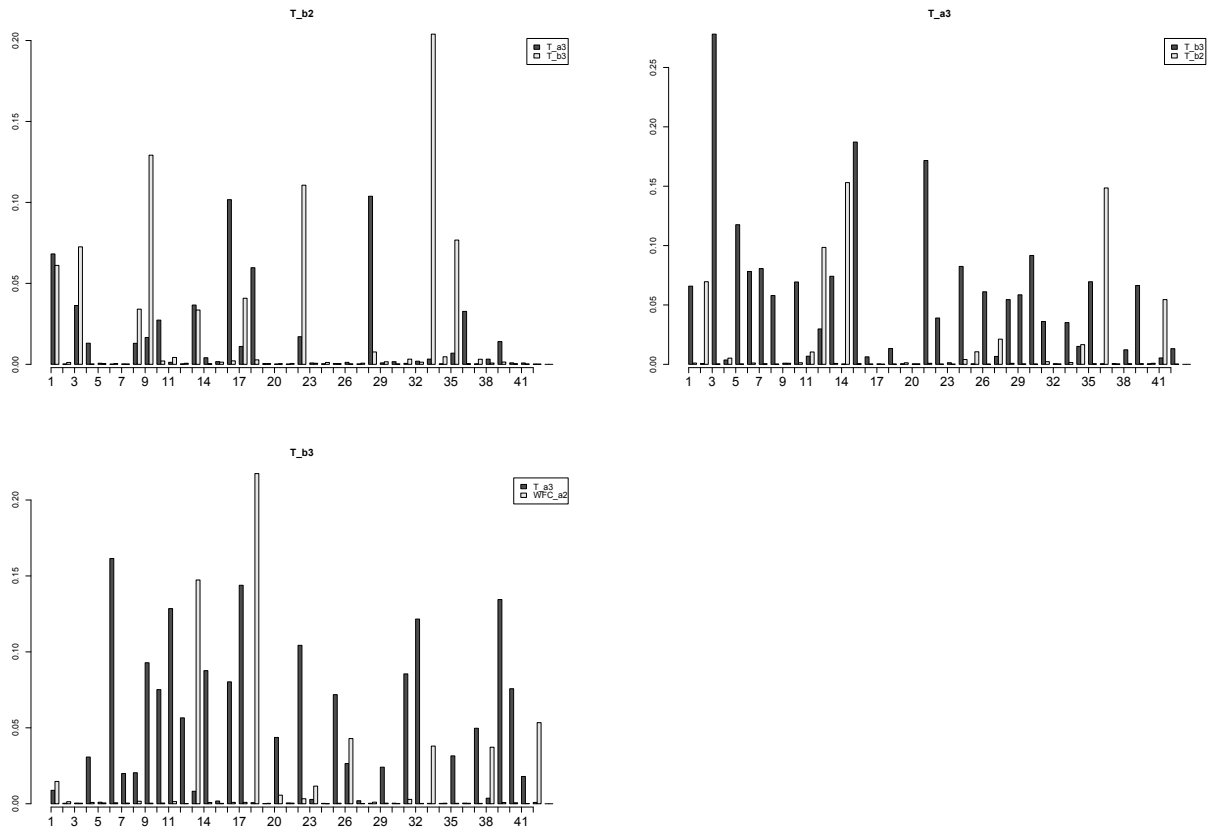**Figure A.2.** The time-varying network for 1st level ask/bid size factor, 2nd level ask/bid size factor and 3rd level ask/bid size factor of T.

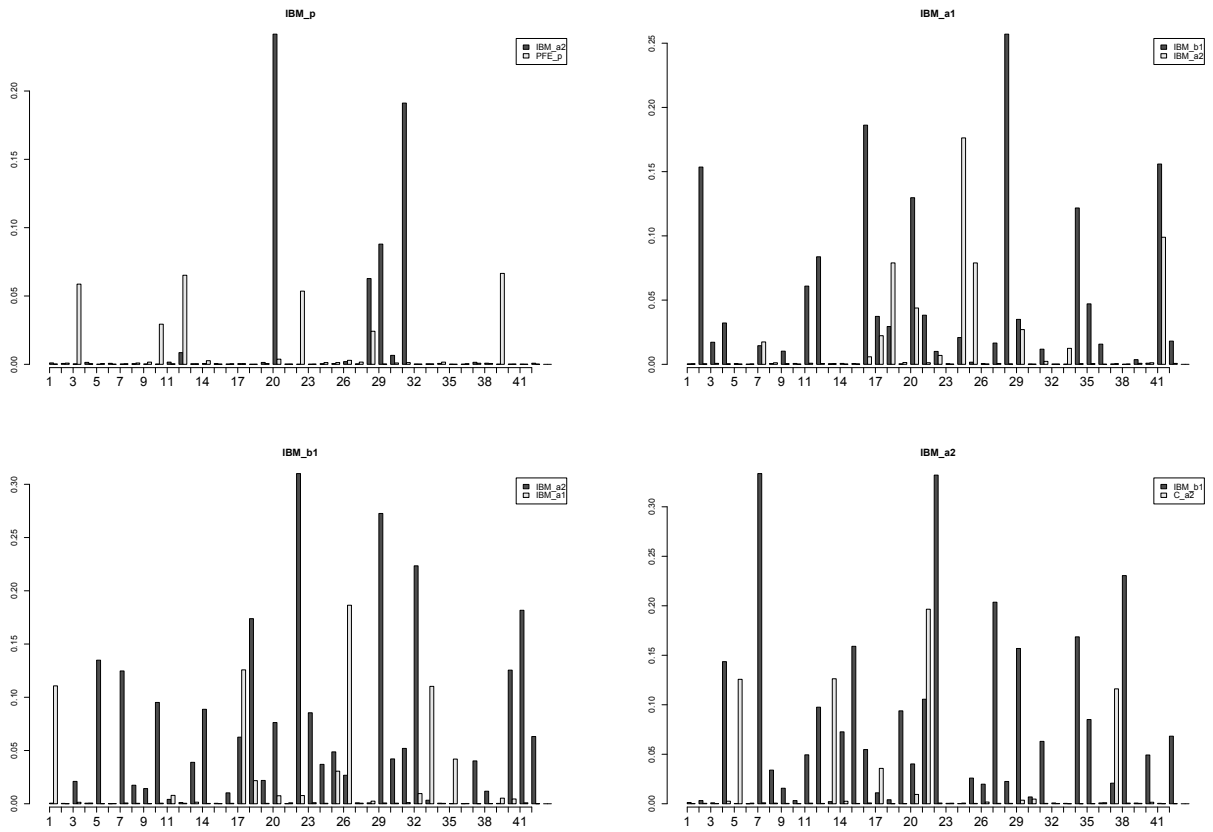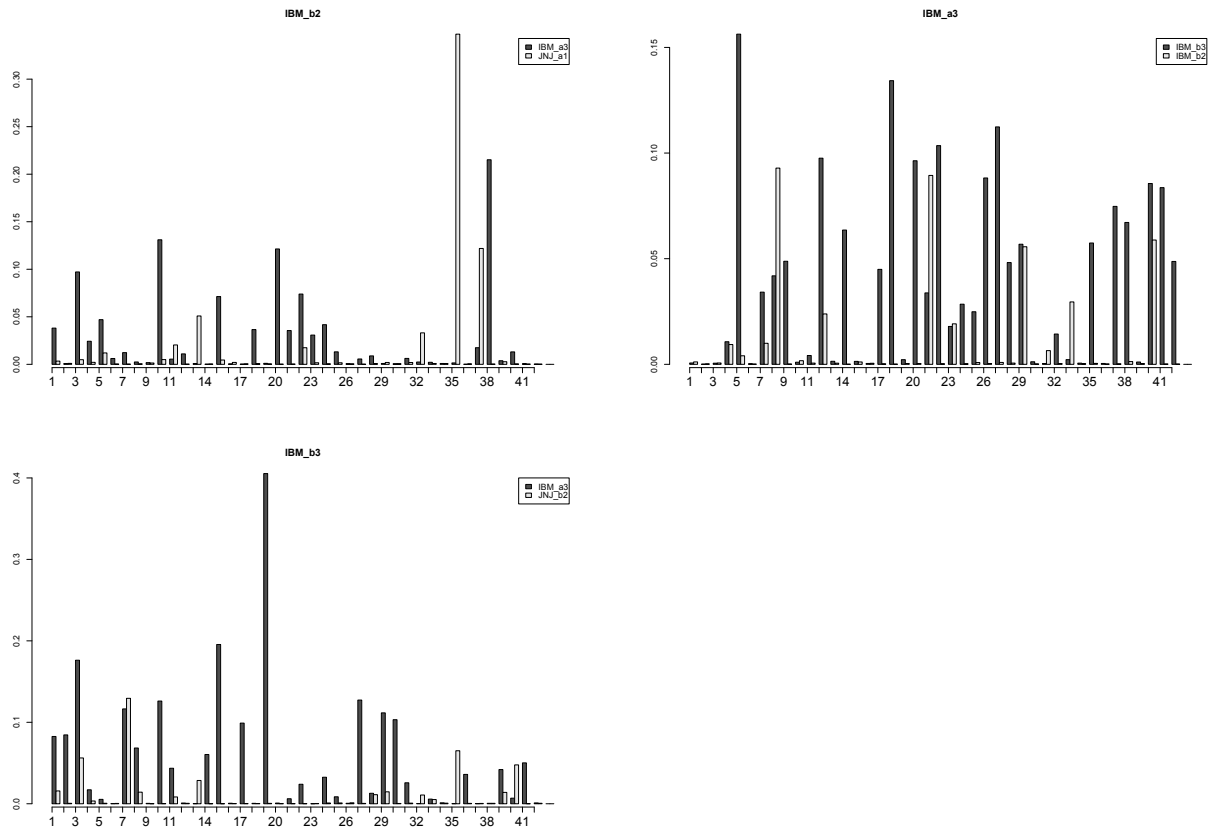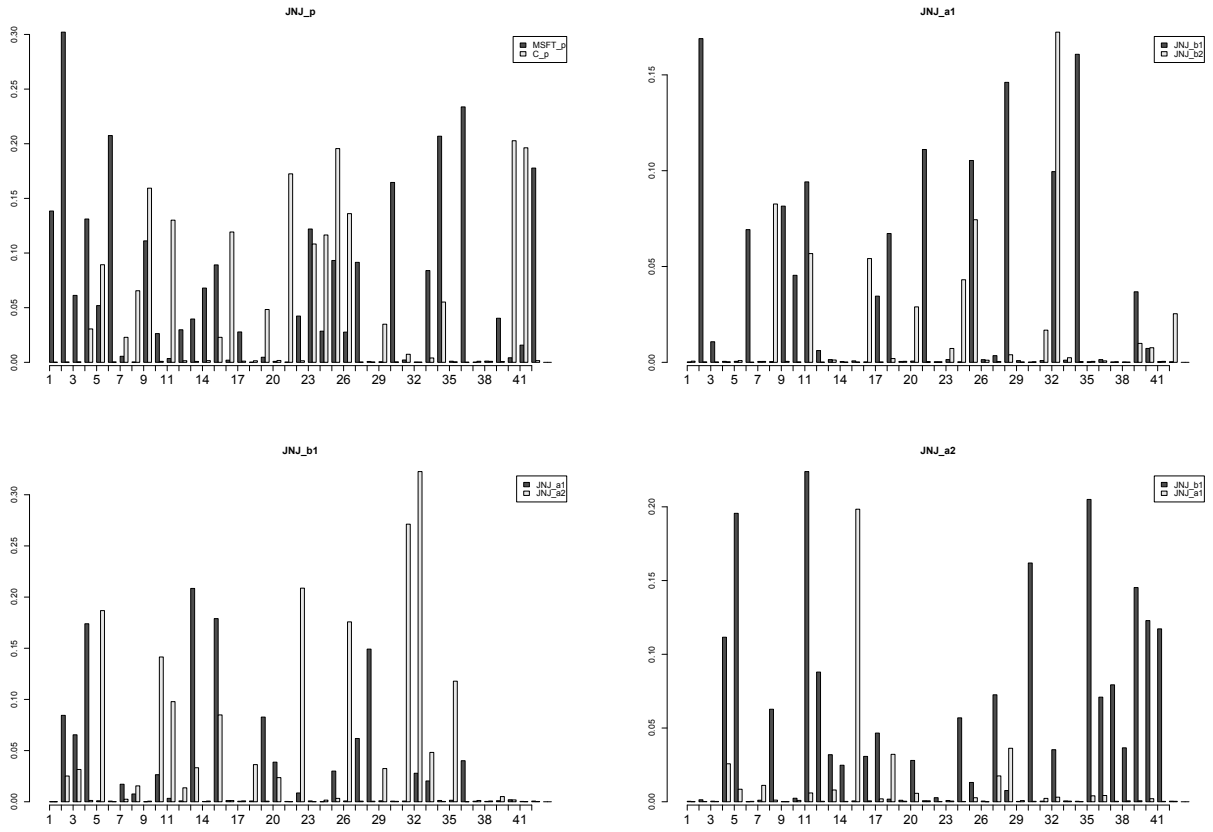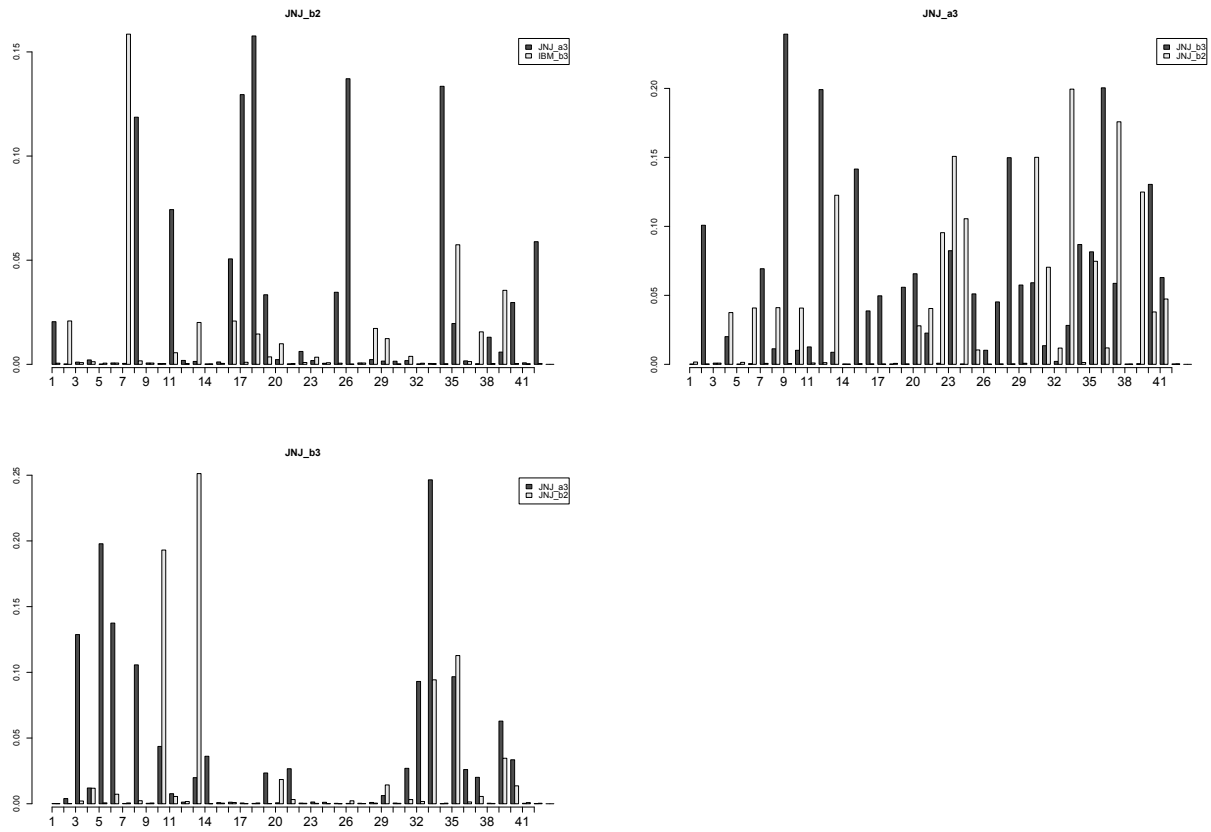**Figure A.3.** The time-varying network of price factor, 1st level ask/bid size factor, 2nd level ask size factor of IBM.

Figure A.4. The time-varying network for 1st level ask/bid size factor, 2nd level ask/bid size factor and 3rd level ask/bid size factor of IBM.

Figure A.5. The time-varying network of price factor, 1st level ask/bid size factor, 2nd level ask size factor of JNJ.

Figure A.6. The time-varying network for 1st level ask/bid size factor, 2nd level ask/bid size factor and 3rd level ask/bid size factor of JNJ.
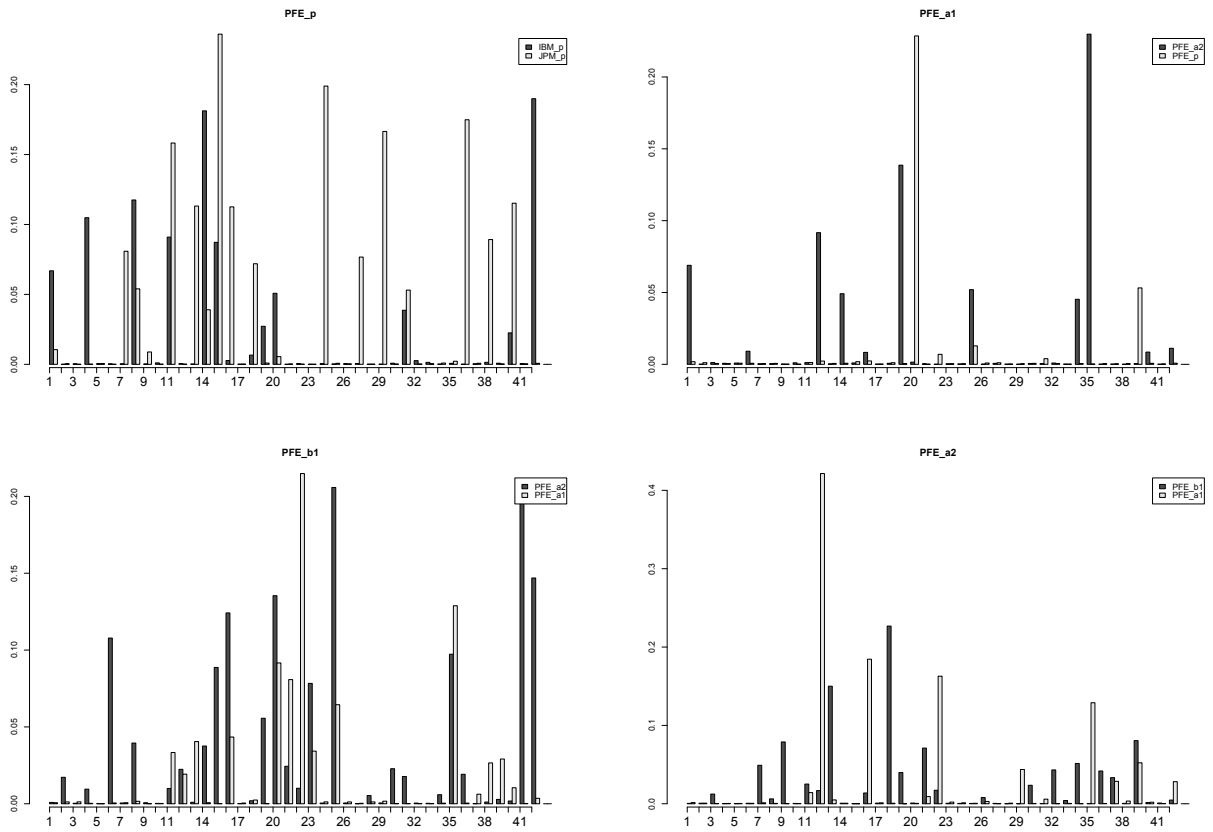
Figure A.7. The time-varying network of price factor, 1st level ask/bid size factor, 2nd level ask size factor of PFE.
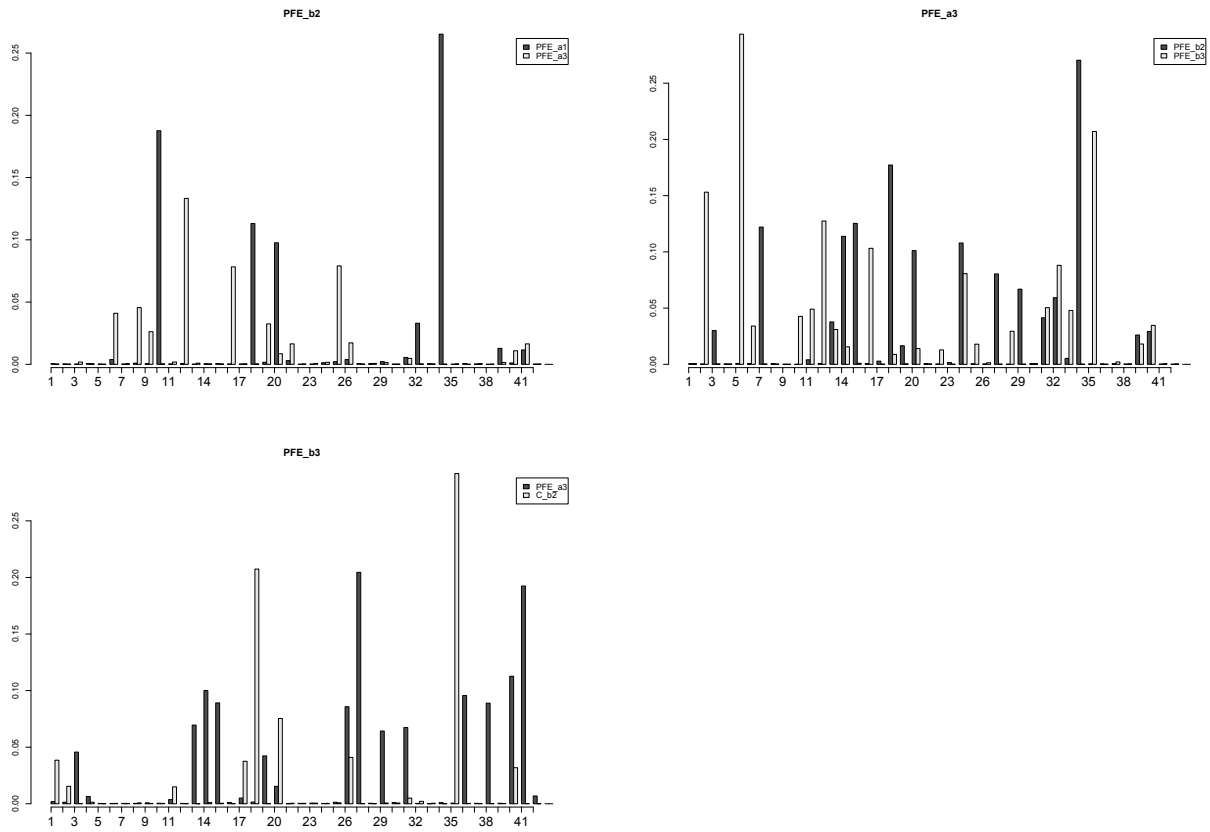
Figure A.8. The time-varying network for 1st level ask/bid size factor, 2nd level ask/bid size factor and 3rd level ask/bid size factor of PFE.
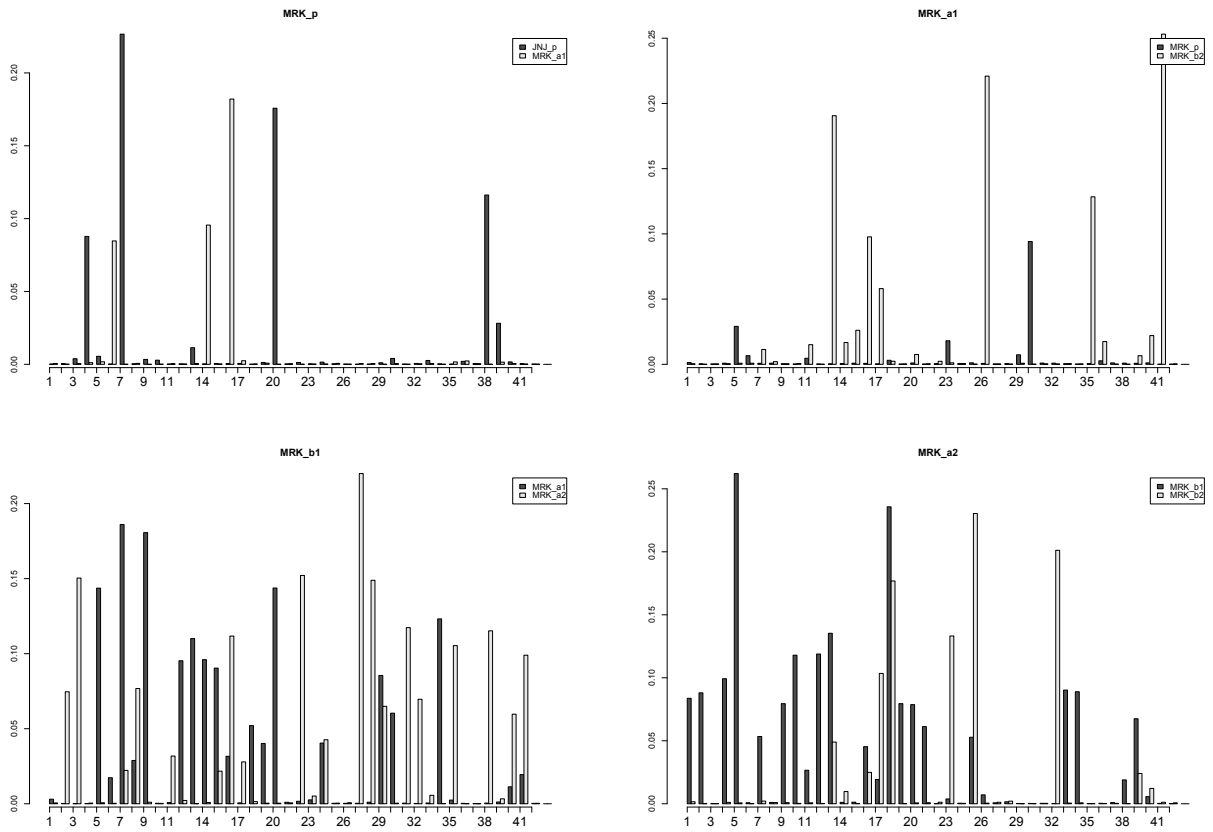
Figure A.9. The time-varying network of price factor, 1st level ask/bid size factor, 2nd level ask size factor of MRK.

Figure A.10. The time-varying network for 1st level ask/bid size factor, 2nd level ask/bid size factor and 3rd level ask/bid size factor of MRK.

Figure A.11. The time-varying network of price factor, 1st level ask/bid size factor, 2nd level ask size factor of JPM.
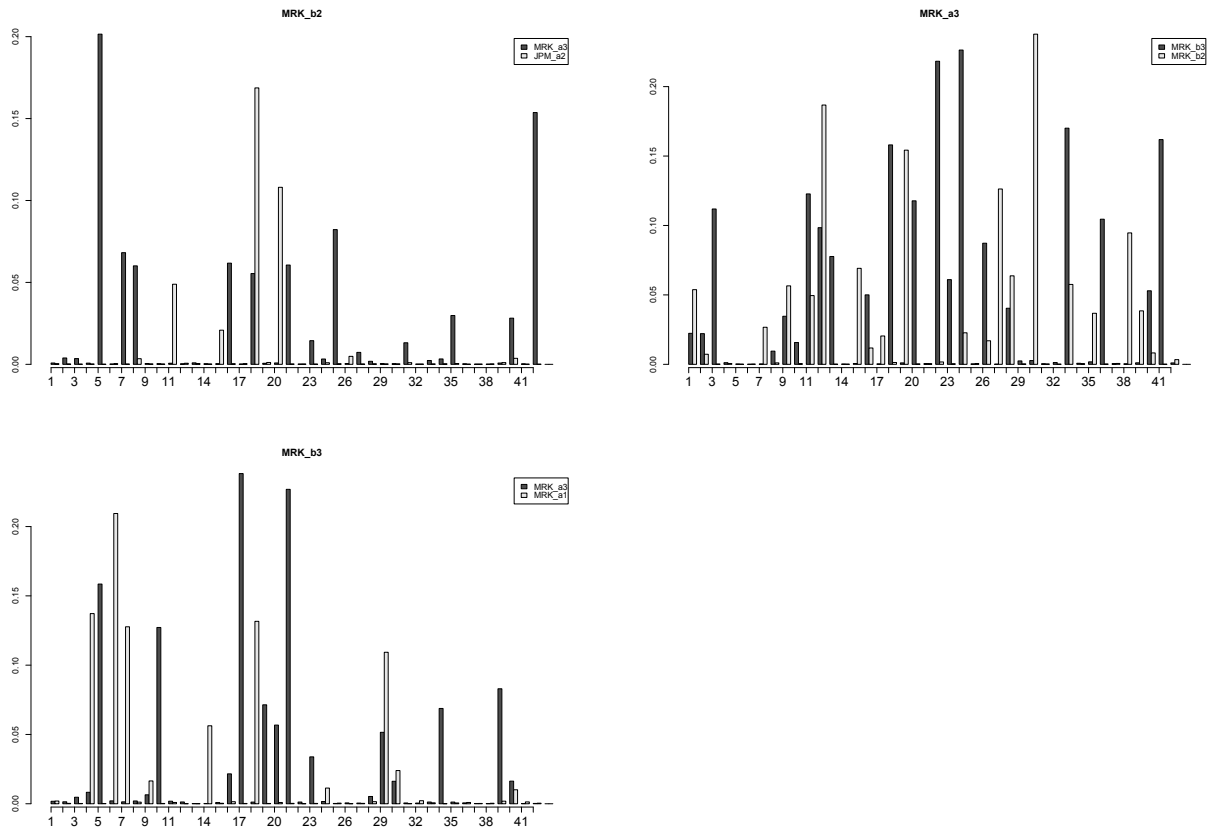
Figure A.12. The time-varying network for 1st level ask/bid size factor, 2nd level ask/bid size factor and 3rd level ask/bid size factor of JPM.

Figure A.13. The time-varying network of price factor, 1st level ask/bid size factor, 2nd level ask size factor of WFC.

Figure A.14. The time-varying network for 1st level ask/bid size factor, 2nd level ask/bid size factor and 3rd level ask/bid size factor of WFC.
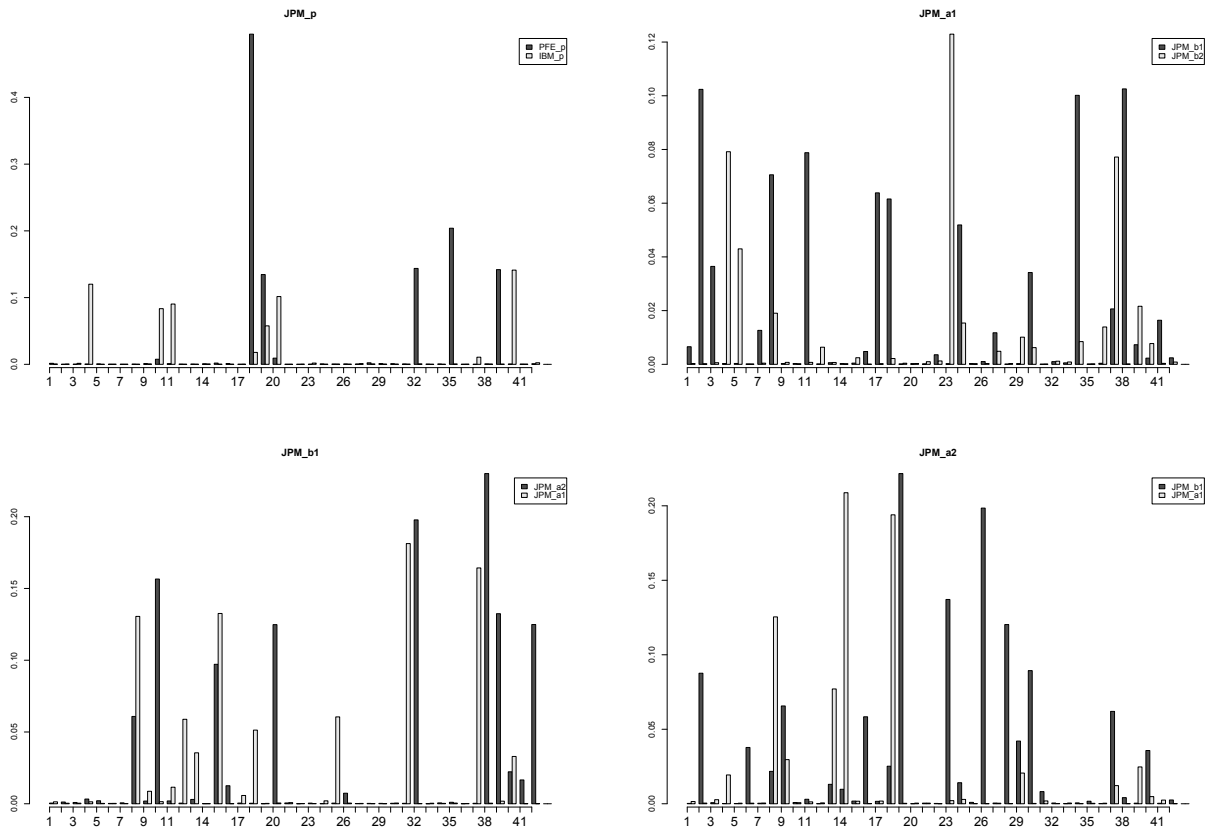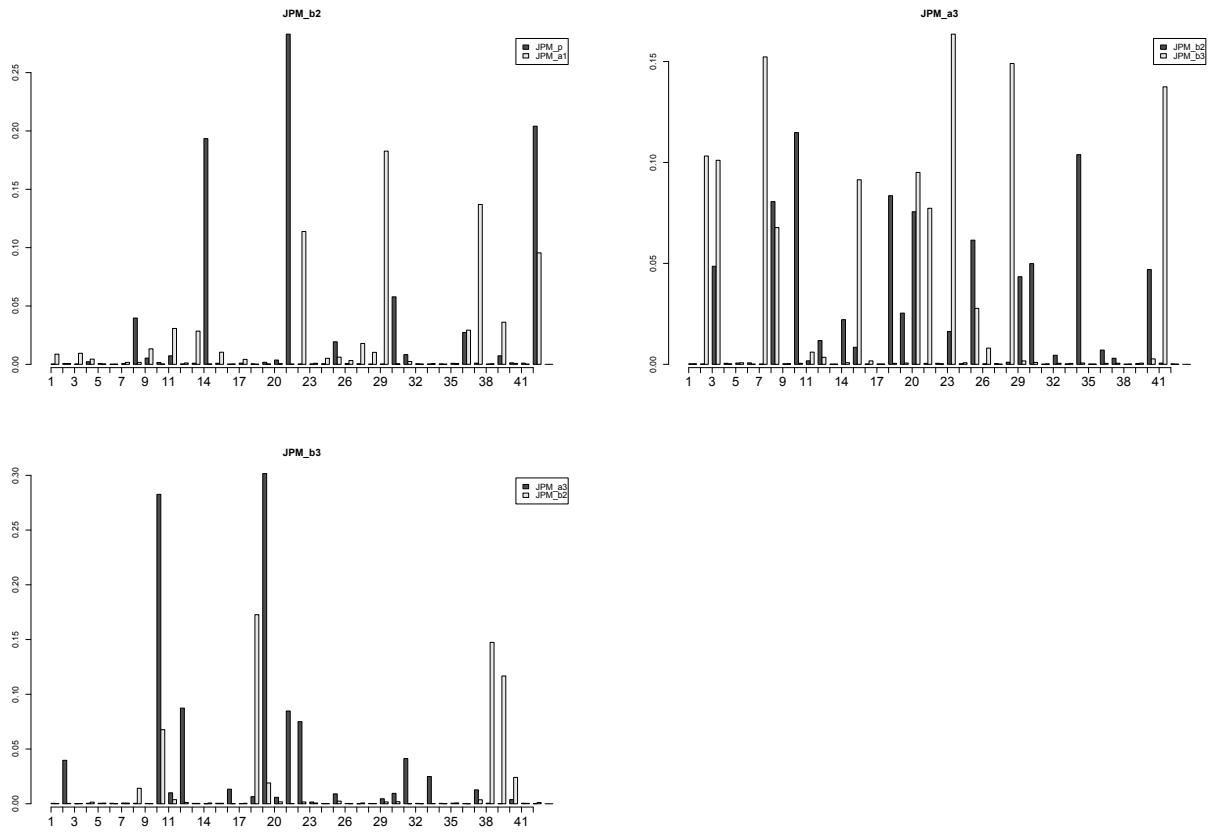
Figure A.15. The time-varying network of price factor, 1st level ask/bid size factor, 2nd level ask size factor of Citigroup.

Figure A.16. The time-varying network for 1st level ask/bid size factor, 2nd level ask/bid size factor and 3rd level ask/bid size factor of Citigroup.
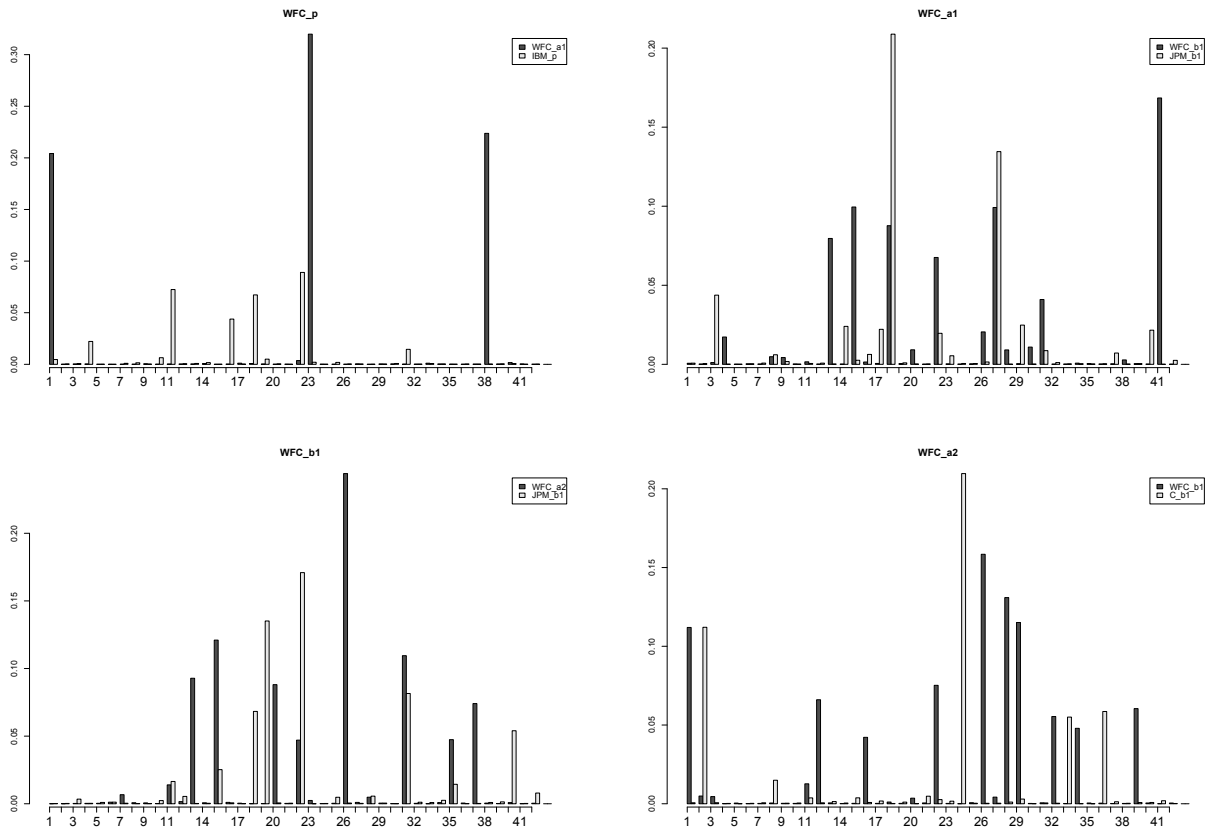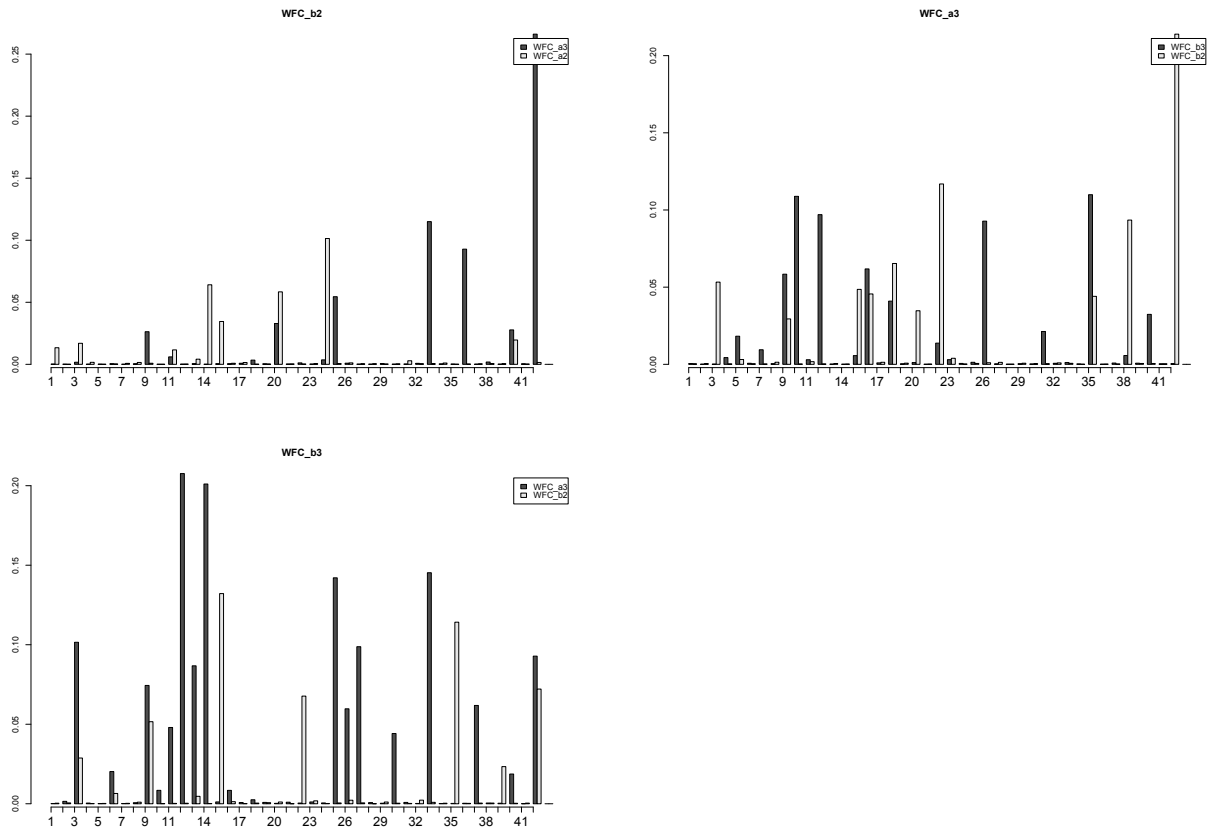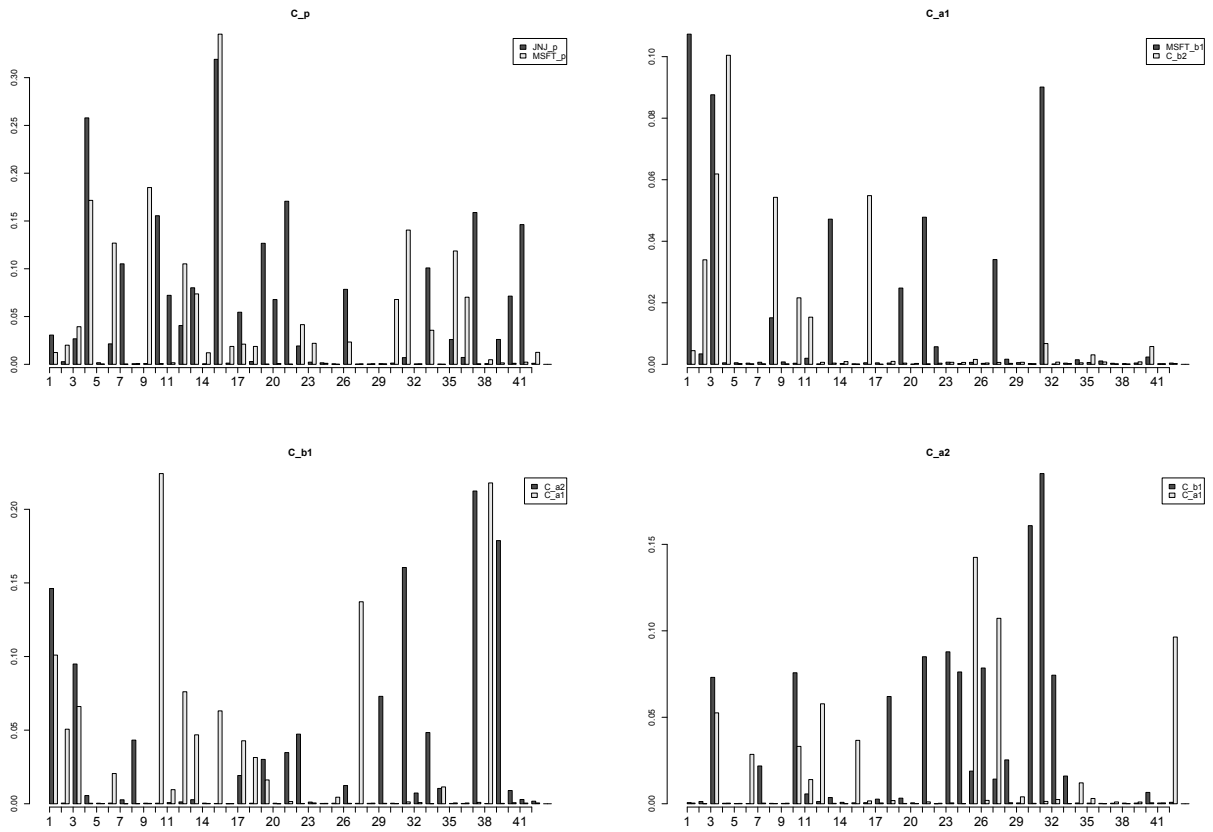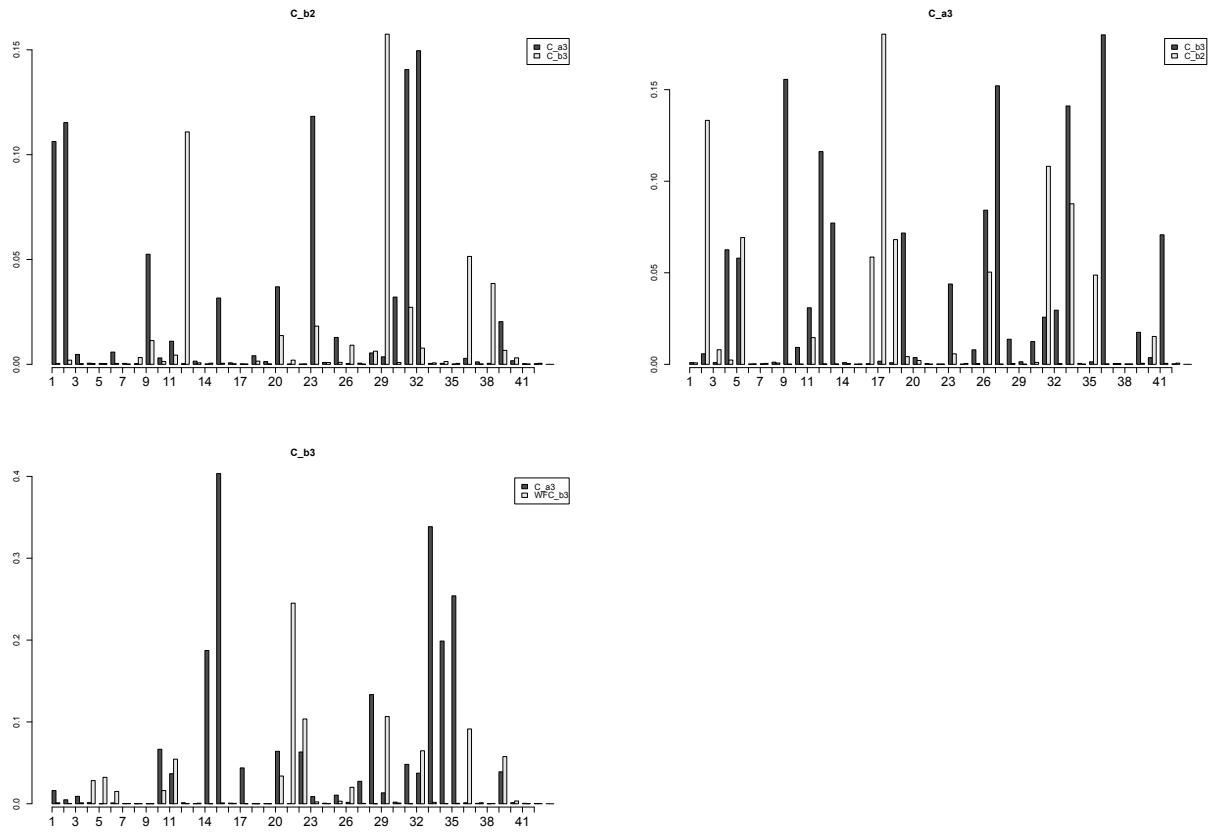
months to maturity can be written as,

$$
\begin{aligned}
\frac{A(\tau)}{\tau} &= \bar{A}\frac{\tau^2}{6} + \bar{B}\left\{\frac{1}{2\lambda^2} - \frac{1}{\lambda^3}\frac{1-\exp(-\lambda\tau)}{\tau} + \frac{1}{4\lambda^3}\frac{1-\exp(-2\lambda\tau)}{\tau}\right\} \\
&+ \bar{C}\left\{\frac{1}{2\lambda^2} + \frac{1}{\lambda^2}\exp(-\lambda\tau) - \frac{1}{4\lambda}\tau\exp(-2\lambda\tau) - \frac{3}{4\lambda^2}\exp(-2\lambda\tau)\right\} \\
&+ \bar{C}\left\{-\frac{2}{\lambda^3}\frac{1-\exp(-\lambda\tau)}{\tau} + \frac{5}{8\lambda^3}\frac{1-\exp(-2\lambda\tau)}{\tau}\right\} \\
&+ \bar{D}\left\{\frac{1}{2\lambda}\tau + \frac{1}{\lambda^2}\exp(-\lambda\tau) - \frac{1}{\lambda^3}\frac{1-\exp(-\lambda\tau)}{\tau}\right\} \\
&+ \bar{E}\left\{\frac{3}{\lambda^2}\exp(-\lambda\tau) + \frac{1}{2\lambda}\tau + \frac{1}{\lambda}\exp(-\lambda\tau) - \frac{3}{\lambda^3}\frac{1-\exp(-\lambda\tau)}{\tau}\right\} \\
&+ \bar{F}\left\{\frac{1}{\lambda^2} + \frac{1}{\lambda^2}\exp(-\lambda\tau) - \frac{1}{2\lambda^2}\exp(-2\lambda\tau) - \frac{3}{\lambda^3}\frac{1-\exp(-\lambda\tau)}{\tau} + \frac{3}{4\lambda^3}\frac{1-\exp(-2\lambda\tau)}{\tau}\right\}
\end{aligned}
$$
$$(A.11)$$

where the six terms $\bar{A}$, $\bar{B}$, $\bar{C}$, $\bar{D}$, $\bar{E}$ and $\bar{F}$ can be identified by the volatility matrix Σ defined in the dynamics equation under P-measure. The value of the adjustment term is constant in time t, but depends on time to maturity τ, coefficient λ that governs the mean reversion rate of slope and curvature factors, and the volatility parameters $\bar{A}$, $\bar{D}$ and $\bar{F}$.

The four latent factors defined in the state variable $X_{it}^{\top} = \left(L_{it}^N, S_{it}^N, C_{it}^N, L_{it}^R\right)$ evolve dynamically and hence we can identify their shocks accordingly,

$$
\begin{pmatrix} dL_{it}^N \\ dS_{it}^N \\ dC_{it}^N \\ dL_{it}^R \end{pmatrix} = \begin{pmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & \kappa_{14} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & \kappa_{24} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & \kappa_{34} \\ \kappa_{41} & \kappa_{42} & \kappa_{43} & \kappa_{44} \end{pmatrix} \begin{pmatrix} L_{it}^N \\ S_{it}^N \\ C_{it}^N \\ L_{it}^R \end{pmatrix} dt + \Sigma \begin{pmatrix} dW_t^{L^N} \\ dW_t^{S^N} \\ dW_t^{C^N} \\ dW_t^{L^R} \end{pmatrix} \quad (A.12)
$$

where $W_t^{L^N}$, $W_t^{S^N}$, $W_t^{C^N}$ and $W_t^{L^R}$ are independent Brownian motions.

We estimate the parameters in (A.2) using the Kalman filter technique. The Kalman filter recursion is a set of equations which allow for an estimator to be updated once a new observation y_t becomes available. It first forms an optimal predictor of the unobserved state variable vector given its previously estimated value. This prediction is obtained using the distribution of unobserved state variables, conditional on the previous estimated values. These estimates for unobserved state variables are then updated using the information provided by the observed variables.

By rewriting the yield equation (6) of the joint AFNS model in multi-maturity term

structure proposed in sub-section 3.2.3, we obtain the measurement equation as,

$$
\begin{pmatrix} y_{it}^N(\tau_1) \\ y_{it}^R(\tau_1) \\ \vdots \\ y_{it}^R(\tau_n) \end{pmatrix} \;=\; AX_{it} + \begin{pmatrix} \varepsilon_{it}^N(\tau_1) \\ \varepsilon_{it}^R(\tau_1) \\ \vdots \\ \varepsilon_{it}^R(\tau_n) \end{pmatrix} - \begin{pmatrix} \dfrac{A_i^N(\tau_1)}{\tau_1} \\ \dfrac{A_i^R(\tau_1)}{\tau_1} \\ \vdots \\ \dfrac{A_i^R(\tau_n)}{\tau_n} \end{pmatrix} \tag{A.13}
$$

The transition equation derived from Christensen et al. (2011) takes the form of,

$$
X_{i,t} = \left[I - expm\left(-K^P \Delta t\right) \right] \theta^P + expm\left(-K^p \Delta t\right) X_{i,t-1} + \eta_t \tag{A.14}
$$

where *expm* is a matrix exponential. The measurement and transition equations are assumed to have the error structure as,

$$
\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} \;=\; N \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & H \end{pmatrix} \right\}
$$

where Q has a special structure,

$$
Q = \int_0^{\Delta t} e^{-K^P s} \Sigma \Sigma^\top e^{-(K^P)^\top s} ds
$$

For estimation, the transition and measurement errors are assumed orthogonal to the initial state. The initial value of the filter is given by the unconditional mean and variance of the state variable $X_t^\top$ under the P measure,

$$
X_{\mathrm{o}} = \theta^P
$$

$$
\Sigma_{\mathrm{o}} = \int_0^\infty e^{-K^P s} \Sigma \Sigma^\top e^{-(K^P)^\top s} ds
$$

which can be calculated using the analytical solution provided in Fisher and Gilles (1996).

A.5 BEIR Decomposition

In the environment of an AF model, there are no opportunities to make risk-free profits. Based on the pricing equation from Cochrane (2005), the bond can be priced by the equation,

$$
P_t = \mathsf{E}_t \left\{ \beta \frac{u'(c_{t+1})}{u'(c_t)} x_{t+1} \right\} \tag{A.15}
$$

where the value of the bond c_t has a payoff x_{t+1}, β is the discount factor. We break up the basic consumption-based pricing equation (B.1) and get the stochastic discount factor (SDF) M_{t+1} at time $t+1$,

$$M_{t+1} = \beta \frac{u'(c_{t+1})}{u'(c_t)} \tag{A.16}$$

To estimate the expected value of inflation using the stochastic discount factor (SDF) M_t. Firstly we use the Taylor series to approximate the moments of the logarithm. Assuming that M_t, in a sense, significant from 0, so the yield for a nominal bond can be extended as follows,

$$\log\left(M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N\right) = \log\left\{\left(\mu_M + M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N - \mu_M\right)\right\} \tag{A.17}$$

where

$$\mu_M = \mathsf{E}_t\left(M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N\right) \tag{A.18}$$

Expand equation (B.3) using Taylor series and take the expectation on both sides,

$$\mathsf{E}_t\left(\log M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N\right) = \log \mu_M - \mathsf{Var}_t\left(\log M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N\right) \tag{A.19}$$

Therefore,

$$\begin{aligned} y_t^N(\tau) &= -\frac{1}{\tau}\log \mathsf{E}_t\left(M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N\right) \\ &= -\frac{1}{\tau}\mathsf{E}_t\left(\log M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N\right) - \frac{1}{2\tau}\mathsf{Var}_t\left(\log M_{t+1}^N M_{t+2}^N \cdots M_{t+\tau}^N\right) \end{aligned} \tag{A.20}$$

Similar solution could be obtained for the inflation-indexed bonds by the same logic.

To facilitate the calculation of equation (12), the instantaneous risk-free rate r_t and the risk premium Γ_t are given, more details can be found in Christensen et al. (2011) and Christensen et al. (2010a),

$$\begin{aligned} r_t &= \rho_0(t) + \rho_1(t)X_t \tag{A.21} \\ \Gamma_t &= \gamma_0 + \gamma_1 Y_t \tag{A.22} \end{aligned}$$

where $\rho_0(t)$, $\rho_1(t)$, γ_0 and γ_1 are bounded, continuous functions. X_t is the state variable and Y_t is the realized observations.

The estimation of the inflation expectation can be calculated by

$$\pi_t(\tau) = -\frac{1}{\tau}\log \mathsf{E}_t^P\left[\exp\left\{-\int_t^{t+\tau}(r_s^N - r_s^R)ds\right\}\right] \tag{A.23}$$

which are the solutions to a system of ordinary differential equations using the fourth-

order Runge Kutta method.

Bibliography

Acharya, V., Engle, R., and Richardson, M. (2012). Capital shortfall: A new approach to ranking and regulating systemic risks. *The American Economic Review*, 102(3):59–64.

Acharya, V. V., Pedersen, L. H., Philippon, T., and Richardson, M. (2017). Measuring systemic risk. *The Review of Financial Studies*, 30(1):2–47.

Adrian, T. and Wu, H. Z. (2009). The term structure of inflation expectations. *FRB of New York Staff Report*, (362).

Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial studies*, 18(2):351–416.

Bandi, F. M. and Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics*, 79(3):655–692.

Bandi, F. M. and Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. *The Review of Economic Studies*, 75(2):339–369.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536.

Barr, D. G. and Campbell, J. Y. (1997). Inflation, real interest rates, and the bond market: A study of uk nominal and index-linked government bond prices. *Journal of Monetary Economics*, 39(3):361–383.

Basu, S., Michailidis, G., et al. (2015). Regularized estimation in sparse high-dimensional time series models. *The Annals of Statistics*, 43(4):1535–1567.

Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6):2369–2429.

Belloni, A., Chernozhukov, V., et al. (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli*, 19(2):521–547.

Bickel, P. J., Ritov, Y., and Tsybakov, A. B. (2009). Simultaneous analysis of lasso and dantzig selector. *Ann. Statist.*, 37(4):1705–1732.

Bjork, T. and Christensen, B. J. (1999). Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9:323–348.

Campbell, J. and Viceira, L. (2009). Understanding inflation-indexed bond markets. Technical report, National Bureau of Economic Research.

Candes, E. and Tao, T. (2007). The dantzig selector: Statistical estimation when p is much larger than n. *The Annals of Statistics*, pages 2313–2351.

Chen, X., Xu, M., Wu, W. B., et al. (2013). Covariance and precision matrix estimation for high-dimensional time series. *The Annals of Statistics*, 41(6):2994–3021.

Christensen, J. H., Diebold, F. X., and Rudebusch, G. D. (2011). The affine arbitrage-free class of nelson–siegel term structure models. *Journal of Econometrics*, 164(1):4–20.

Christensen, J. H., Lopez, J. A., and Rudebusch, G. D. (2010a). Inflation expectations and risk premiums in an arbitrage-free model of nominal and real bond yields. *Journal of Money, Credit and Banking*, 42(1):143–178.

Christensen, K., Kinnebrock, S., and Podolskij, M. (2010b). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics*, 159(1):116–133.

Cochrane, J. H. (2005). *Asset pricing*, volume 1. Princeton University Press Princeton, NJ.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische annalen*, 300(1):463–520.

Demirer, M., Diebold, F. X., Liu, L., and Yılmaz, K. (2017). Estimating global bank network connectedness. Technical report, National Bureau of Economic Research.

Diebold, F. X. and Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2):337–364.

Diebold, F. X., Li, C., and Yue, V. Z. (2008). Global yield curve dynamics and interactions: a dynamic nelson–siegel approach. *Journal of Econometrics*, 146(2):351–363.

Diebold, F. X., Rudebusch, G. D., and Aruoba, S. B. (2006). The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics*, 131(1):309–338.

Diebold, F. X. and Yılmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134.

Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., et al. (2004). Least angle regression. *The Annals of statistics*, 32(2):407–499.

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association*, 96(456):1348–1360.

Fan, J. and Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(5):849–911.

Fan, J., Samworth, R., and Wu, Y. (2009). Ultrahigh dimensional feature selection: beyond the linear model. *Journal of Machine Learning Research*, 10(Sep):2013–2038.

Fan, J., Song, R., et al. (2010). Sure independence screening in generalized linear models with np-dimensionality. *The Annals of Statistics*, 38(6):3567–3604.

Fisher, M. and Gilles, C. (1996). Term premia in exponential-affine models of the term structure. *Manuscript, Board of Governors of the Federal Reserve System*.

Frank, L. E. and Friedman, J. H. (1993). A statistical view of some chemometrics regression tools. *Technometrics*, 35(2):109–135.

Fu, W. J. (1998). Penalized regressions: the bridge versus the lasso. *Journal of computational and graphical statistics*, 7(3):397–416.

Giglio, S., Kelly, B., and Pruitt, S. (2016). Systemic risk and the macroeconomy: An empirical evaluation. *Journal of Financial Economics*, 119(3):457–471.

Härdle, W. K. and Majer, P. (2014). Yield curve modeling and forecasting using semi-parametric factor dynamics. *The European Journal of Finance*, (ahead-of-print):1–21.

Hautsch, N. and Huang, R. (2012). The market impact of a limit order. *Journal of Economic Dynamics and Control*, 36(4):501–522.

Hautsch, N. and Podolskij, M. (2013). Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: theory, implementation, and empirical evidence. *Journal of Business & Economic Statistics*, 31(2):165–183.

Hautsch, N., Schaumburg, J., and Schienle, M. (2014). Financial network systemic risk contributions. *Review of Finance*, 19(2):685–738.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic processes and their applications*, 119(7):2249–2276.

Joyce, M. A., Lildholdt, P., and Sorensen, S. (2010). Extracting inflation expectations and inflation risk premia from the term structure: a joint model of the uk nominal and real yield curves. *Journal of Banking & Finance*, 34(2):281–294.

Kock, A. B. and Callot, L. (2015a). Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics*, 186(2):325–344.

Kock, A. B. and Callot, L. (2015b). Oracle inequalities for high dimensional vector autoregressions. *Journal of Econometrics*, 186(2):325–344.

Koop, G., Pesaran, M. H., and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1):119–147.

Lanne, M. and Nyberg, H. (2016). Generalized forecast error variance decomposition for linear and nonlinear multivariate models. *Oxford Bulletin of Economics and Statistics*, 78(4):595–603.

Loh, P.-L. and Wainwright, M. J. (2011). High-dimensional regression with noisy and missing data: Provable guarantees with non-convexity. In *Advances in Neural Information Processing Systems*, pages 2726–2734.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Negahban, S. and Wainwright, M. J. (2011). Estimation of (near) low-rank matrices with noise and high-dimensional scaling. *The Annals of Statistics*, pages 1069–1097.

Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of Business*, pages 473–489.

Pesaran, H. H. and Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters*, 58(1):17–29.

Pflueger, C. E. and Viceira, L. M. (2011). Inflation-indexed bonds and the expectations hypothesis. Technical report, National Bureau of Economic Research.

Podolskij, M., Vetter, M., et al. (2009). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3):634–658.

Terasvirta, T., Tjostheim, D., Granger, C. W., et al. (2010). Modelling nonlinear economic time series. *OUP Catalogue*.

Tibshirani, R. (1996a). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288.

Tibshirani, R. (1996b). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B*, 58(1):pp. 267–288.

Van Buuren, S. and Oudshoorn, C. (2000). Multivariate imputation by chained equations. *MICE V1. 0 user's manual. Leiden: TNO Preventie en Gezondheid.*

Wu, W. B. and Wu, Y. N. (2016). Performance bounds for parameter estimates of high-dimensional linear models with correlated errors. *Electronic Journal of Statistics*, 10(1):352–379.

Zhang, L. (2011). Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics*, 160(1):33–47.

Zhang, L. et al. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12(6):1019–1043.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101(476):1418–1429.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 28. Dezember 2017   Shi Chen