IRTG 1792 Discussion Paper 2020-022

Tail Event Driven Factor Augmented Dynamic Model

Weining Wang [*] [*2]
Lining Yu [*]
Bingling Wang [*]

[*] Humboldt-Universität zu Berlin, Germany
[*2] University of London, UK

This research was supported by the Deutsche Forschungsgesellschaft through the International Research Training Group 1792 "High Dimensional Nonstationary Time Series".

http://irtg1792.hu-berlin.de
ISSN 2568-5619

International Research Training Group 1792

Tail Event Driven Factor Augmented Dynamic Model*

Weining Wang† Lining Yu‡ Bingling Wang §

Abstract

A factor augmented dynamic model for analysing tail behaviour of high dimensional time series is proposed. As a first step, the tail event driven latent factors are extracted. In the second step, a VAR (Vectorautoregression model) is carried out to analyse the interaction between these factors and the macroeconomic variables. Furthermore, this methodology also provides the possibility for central banks to examine the sensitivity between macroeconomic variables and financial shocks via impulse response analysis. Then the predictability of our estimator is illustrated. Finally, forecast error variance decomposition is carried out to investigate the network effect of these variables. The interesting findings are: firstly, GDP and Unemployment rate are very much sensitive to the shock of financial tail event driven factors, while these factors are more affected by inflation and short term interest rate. Secondly, financial tail event driven factors play important roles in the network constructed by the extracted factors and the macroeconomic variables. Thirdly, there is more connectedness during financial crisis than in the stable periods. Compared with median case, the network is more dense in lower quantile level.

Keywords: Quantile Regression, Expectile Regression, Dynamic Factor Model, Dynamic Network
JEL: C21, C51, G01, G18, G32, G38.

---

*The financial support from SFB649 is greatly acknowleged

†Professor at Ladislaus von Bortkiewicz Chair of Statistics, C.A.S.E. - Center for Applied Statistics and Econometrics, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. Email: wangwein@wiwi.hu-berlin.de. Department of Economics, City, University of London, Northampton Square, London EC1V 0HB, United Kingdom.

‡Research associate at Ladislaus von Bortkiewicz Chair of Statistics, C.A.S.E. - Center for Applied Statistics and Econometrics, IRTG 1792, Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany. Email: wsqiuning@hotmail.com.

§Humboldt-Universität zu Berlin, IRTG 1792, Dorotheenstr.1, 10117 Berlin, Germany.

1 Introduction

Quantile and expectile method are known as important tools for analysing tail behaviour, see Koenker and Bassett (1978) and Newey and Powell (1987). One very important application of quantile regression in finance is to calculate Value at Risk, for instance, Duffie and Pan (1997), Schaumburg (2012) and Xiao et al. (2015). Compared with quantile, the advantage of expectile is its coherency property. Gschöpf et al. (2015) introduced a flexible transformation framework, which provides a simple way to transfer from quantile to expectile, vice versa. However, when the dimensionality of the time series becomes too high, it will be difficult to analyse the tail dependency. One of the most effective solution for this is to use factor models.

Factor models are known to be important tools to reduce the dimensionality, see Cattell (1952), Fruchter (1954), McDonald (1985), Bai and Ng (2006) and Markus (2015). For instance, to analyse common behaviour in financial markets, the major factors could be extracted from the financial data. Moreover, the extracted factors are easily to connect with macro variables via a vector auto regression (VAR) model, which facilitates people to perform forecasting and study the interactions among them. This type of model is known as Factor-Augmented Vector Autoregressive (FAVAR) model. Favero and Marcellino (2005) used FAVAR model to analyse monetary policy in Europe. Bernanke et al. (2005) applied it to detect the policy shock on macro variables. Stock and Watson (2005) used this model to examine the underling shocks from macro economy. Moench (2008) forecasted the yield via FAVAR model.

However, the extracted factors of aforementioned literature are only provided in the median level, the tail event case has not been investigated. In the context of studying systemic risk, it is of interest to have tail risk driven factors isolated. A tail event driven factor augmented dynamic model is exactly a model to serve this purpose. In general, people use the Kalman filter and maximum likelihood to estimate the FAVAR model. But when the number of variables is large, this approach becomes computationally infeasible. Two alternative estimation methods are proposed by Bernanke et al. (2005): a single-step approach using Markov Chain Monte Carlo (MCMC) methods, and a two-step approach, that is principal components techniques are used to estimate the common factors first, then a VAR model could be constructed and estimated. Bernanke et al. (2005) found that the two-step approach outperformed the single-step one and was simpler to implement. In this paper instead of normal principal component, the principal component based on asymmetric norm is applied. Tran et al. (2019) developed an analogue of PCA in an asymmetric norm, these norms covered both quantiles and expectiles. However, they proposed a pure static setting for functional data, which is not suitable for the goal to study financial systems. In terms of systemic risk modelling their idea could be extended to a dynamic factor model context. Another paper related is De Rossi and Harvey (2009), who considered a state space quantile model in one dimensional setting. The approach was well developed but was too simple (one dimensional) for the needs of high dimensional data. Our approach can be considered as a multidimensional extension to their methodology. Namely, an expectile state space model with multivariate state variables is considered. The underlying state variables (factors) will be filtered by principal expectile component algorithm. Chen et al. (2005) proposed a quantile factor model, they applied an iterative procedure to estimate the common factors and the factor loadings at a given quantile. Chao et al. (2015) and Härdle et al. (2016) introduced quantile and expectile factor model combined with lasso technique. But their models do not have dynamic structure in terms of forecasting, their idea could be extended with dynamic framework.

It is very important to detect the dominant factors in the financial market, therefore, the tail event driven factor model will first be applied to the financial variables (e.g. stock returns). To make the systemic risk study more realistic, we would like to incorporate hedge funds data into our models and investigate further on the role of the hedge fund industry, and the impact of it on the overall financial stability of the system.

After the financial tail event driven factors are extracted, it is of interest to detect the interaction between these factors and macro economy. The macroeconomic variables applied in the Comprehensive Capital Analysis and Review (CCAR) for supervision stress testing are considered. We would like to examine the sensitivity between macro variables and the financial tail event driven factors. Furthermore, the connectedness of financial factors and macro variables are detected.

The contribution of this study is then three folds: First of all, the extracted factors are very important indicators of stock market risk levels. Secondly, the fitted dynamic model would serve as a good tool to detect the sensitivity between macroeconomic variables and financial factors. Finally, this methods could also detect the interaction between macroeconomic variables and financial market under different risk level.

This study is organised as follows: Section 2 introduce the setup of the tail event driven factor augmented dynamic model; Section 4 shows the empirical application by using financial data and macro economy variables; Section 6 concludes; Supplementary materials could be found in Appendix.

2 Model setup

In this section the structure of the factor augmented dynamic model is introduced, then the estimation method will be stated. Let X_{it} $(i = 1, \cdots, N, t = 1, \cdots, T.)$ be the response variable,

$$X_{it} = \sum_{l=1}^{L} \gamma_{il}^{\tau} f_{lt}^{\tau} + \varepsilon_{it}, \quad t = 1, \cdots, T, i = 1, \cdots, N, \tag{1}$$

γ_{il}^{τ}s are factor loadings, f_{lt}^{τ}s are the latent factors, and ε_{it}s are noise components. ε_{it}s are assumed to be i.i.d and $F^{-1}_{\varepsilon_{it}|f_{lt}^{\tau}}(\tau) = 0$. In the vector form, the model in (1) can be rewritten as

$$\underset{(N\times 1)}{X_t} = \underset{(N\times L)}{\Gamma^{\tau\top}} \underset{(L\times 1)}{F_t^{\tau}} + \underset{(N\times 1)}{\varepsilon_t}, \tag{2}$$

where $X_t = [X_{it}]_{i\leq N}$, $\Gamma^{\tau} = [\gamma_{il}^{\tau}]_{N\times L}^{\top} = [\gamma_i^{\tau}]_{i\leq N}$ and $F_t^{\tau} = [f_{lt}^{\tau}]_{l\leq L}^{\top}$ (a vector of f_{lt}^{τ}).
Then the augmented VAR model is assumed to be:

$$G_t^{\tau} = \alpha + B_1 G_{t-1}^{\tau} + B_2 G_{t-2}^{\tau} + \cdots + B_q G_{t-q}^{\tau} + \mathbf{u}_t, \tag{3}$$

where $G_t^{\tau} \overset{def}{=} [F_t^{\tau\top}, M_t^{\top}]^{\top}$, and M_t is a vector of macroeconomic variables, $\mathbf{u}_t$ is a K dimensional process with nonsingular covariance matrix Σ_u. The impulse response function analysis and forecast error variance decomposition will be carried out, see Appendix. The detailed analysis can be found in Lütkepohl (2005). We estimate factors and loadings according to an iterative approach.

2.1 Iteration between $\widehat{F}_t^{\tau}$ and $\widehat{\Gamma}^{\tau}$.

We start with initial value of $(R \times N)$ loading matrix $\widehat{\Gamma}^{\tau(0)}$ estimated from the iterative approach, then the estimated factors can be found using the following iterative procedure:

1. Given $\widehat{\Gamma}^{\tau(a)} = [\widehat{\gamma}_1^{\tau(a)}, \ldots, \widehat{\gamma}_L^{\tau(a)}]$, using quantile or expectile regression of X_t on $\widehat{\Gamma}^{\tau(a)}$ to estimate $\widehat{F}_t^{\tau(a+1)}$ for $t = 1, \ldots, T$.
2. Given $\widehat{F}^{\tau(a+1)} = [\widehat{F}_1^{\tau(a+1)}, \ldots, \widehat{F}_T^{\tau(a+1)}]$, using quantile or expectile regression of X_t on $\widehat{F}^{\tau(a+1)}$ to estimate $\widehat{\Gamma}^{\tau(a+1)}$ for $i = 1, \ldots, N$.
3. Repeat Steps 1 and 2 until $\widehat{\Gamma}^{\tau(R)}$ and $\widehat{\Gamma}^{\tau(R+1)}$ are close enough.

3 Simulation

We also perform a simulation to study the finite sample properties of our proposed algorithms. Assume we have two factors, F_{1t} follows AR(1) process with $F_{1t} = 0.5F_{1,t-1} + \epsilon_{1t}$, and F_{2t} follows AR(2) process with $F_{2t} = 0.3F_{2,t-1} + 0.2F_{2,t-2} + \epsilon_{2t}$. And the corresponding loadings follows normal distribution $\mathrm{N}(0, 9)$ and $\mathrm{N}(0, 4)$, the error term follows 1. standard Gaussian: iid $N(0, 0.01)$, 2. "fat tailed": iid $t(5)$. To identify the factors and loadings, we have fixed the sign of loadings and their corresponding factors. Then the simulated observations and the estimators are shown in Figure 26, Figure 27, Figure 28 and Figure 29.

We compare the simulated factors and loadings with the estimated factors and the loadings by using R^2 from the regression, see Table 13, Table 14, Table 15 and Table 16. The performance of the estimator are examined under three scenarios: (i) the sample size T goes larger while the dimension of time series N is fixed; (ii) the dimension of time serious N goes larger while the sample size T is fixed; (iii) the sample size and the dimension of time series goes larger simultaneously.

In $\tau = 0.5$ cases, under each of the three scenarios and two settings of error terms, the estimators of both the loadings and factors converge, meanwhile their explanatory power of the variation goes higher. In $\tau = 0.05$ or $\tau = 0.95$ cases, under the three scenarios and both assumption of the errors, none of the estimators converge. Besides, in each cases, the estimator of factor 1 and its loading can explain the variation very well while the estimator of factor 2 and its loading tends not to have a constantly good performance. In addition the estimators with $\tau = 0.5$ has a higher explanatory power on the variation than the ones in $\tau = 0.05$ or $\tau = 0.95$ cases.

4 Application

Now the methodology on financial data is illustrated. A sample of monthly log returns of 100 US financial companies in four sectors are taken: Depositories (25), Insurance (25), Broker-Dealers (25) and Hedge Funds (25). The time period considered is from February 30, 1991 to December 31, 2014 on a monthly frequency, $T = 287$. Since there are not enough Broker-Dealers available during this time period, several financial service firms in this group are included.

The data of Depositories, Insurance and Broker-Dealers come from Datastream. And the Hedge funds data are from Lipper Hedge Funds Database. We would like to include the data from the hedge funds sector to analyze their impact on the financial system. The motivation for this is the following. In 2014, the transaction volumes of global hedge funds are about 3.1 USD billions. Pure hedge fund assets were the third largest type of alternative investment (behind real estate and private equity). Although this suggests that the hedge fund industry should play a major active role when analyzing systemic risk in financial system, in the analysis of financial systems the shadow banking sector has been largely overlooked. Chan et al. (2005) quantify the potential impact of hedge funds on systemic risk by developing a number of new risk measures for hedge funds. Billio et al. (2012) apply several econometric measures of connectedness by using the monthly returns of hedge funds, banks, broker/dealers, and insurance companies. We also would like to investigate on the role of the hedge fund industry in our study.

4.1 Descriptive Statistics

Firstly, the difference between four industry groups is compared. Table 1 shows the descriptive statistics of the asset returns of four groups. The hedge funds group has highest mean and lowest standard deviation, which indicates that it is a highest return and lowest risk sector. From Figure 1 and Figure 2 it can be found that in 1998 there are lower mean of banks and broker-dealers, and higher standard deviation of hedge funds (see the last graph of Figure 2), the collapse of Long-Term Capital Management L.P. (LTCM) happened at that time, it had severe impact on the US financial market. During the financial crisis 2008 the mean of hedge funds is very stable, and the standard deviation is lowest than other groups. In this sense one could say that hedge funds applied a lot of strategies and played positive roles on reducing effects of the financial crisis.

4.2 Factor estimation and interpretation

Now the different τ level factors are extracted from the 100 selected US financial firms. A sequence of τ level from 0.05 to 0.95 is estimated, the explained variance for models with one to eight factors is shown in the left plot of Figure 3 and Table 2. Big proportion of variance of the data are explained at tail levels, e.g. a two factors model (orange line, i.e. the second line from bottom in the left panel of Figure 3), 70% for $\tau = 0.05$ and 71% for $\tau = 0.95$. But when τ approaches 0.50 the explained variance becomes smaller, as when $\tau = 0.50$ the value is 34%. This means that the model has more explanatory power at tail levels. The data in tail level are more risky and are more interesting in terms of risk controlling, $\tau = 0.05$ is therefore fixed in the following analysis. The right plot of Figure 3 is the scree plot for $\tau = 0.05$, while the first two factor explains 70%, the additional factors explain less variance of the data, two factors are selected in this analysis afterwards.

The left part of Figure 4 shows that most firms have positive loadings with factor 1 (F1). The exceptions are some hedge funds firms which have slightly negative loadings. The estimated F1 is presented, see the grey line on the right panel of Figure 4. It is found that the pattern of the first estimated factor is similar to average returns of all firms, see Figure 5. The correlation of F1 and average return sequence is 0.96, and both of them are stationary. F1 can be named as a Market Level Factor. As the hedge funds sector is not regulated by authorities, this could also be confirmed by the highest mean and lowest standard deviation in Table 1. Factor two's highest negative loadings are with AIG (American International Group), FNMA (Federal National Mortgage Association) and FDMC (Federal Home Mortgage Corporation), and are with slight positive loadings with other firms. One can track the historical event that AIG accepted the terms of the federal reserve board rescue package on September 16, 2008. This makes it to be the largest government bailout of a private company in US history. Federal Housing Finance Agency took over of FNMA and FDMC in September 2008. These firms were all severely affected by 2008 financial crisis, therefore factor 2 (F2) is called as a Crisis Sensitive Factor. This result can be confirmed by right panel of Figure 4, as the F2 (black line) is more volatile during the crisis time than in normal times. Until here the two factors are extracted from the tail level of the financial data, therefore they will be called as financial tail event driven factors in the following analysis.

4.3 Macroeconomic variables description

Stress testing became nowadays one of the most important supervision strategy after 2007-2008 financial crisis. In US, the Federal Reserve System has conducted CCAR since 2009, which is the annual stress test on the US banking system. Normally, the Federal Reserve releases the stress supervisory scenarios (baseline, adverse, and severely adverse), then the banks will access their minimum capital requirements under these three scenarios by using different methodologies. These macro variables are given and assumed to have important impact to financial firms. In this study, we would like to examine the sensitivity between macro variables and the financial tail event driven factors, the impulse response analysis facilitates us to reach this goal by using 8 macro economic variables involved in 2011 CCAR.

There are nine macroeconomic variables are: real GDP (GDP), Consumer Price Index (CPI), real disposable personal income (INC), unemployment rate (UNE), three-month Treasury bill rate (THR), 10-year Treasury bond rate (TEN), BBB corporate rate (BBB), Dow Jones S&P 500 Index (SP) and National House Price Index (HPI). Honwever, BBB corporate rate is not available before 1997, for consistency with the financial markets data, only eight macroeconomic variables (without BBB corporate rate) are adopted in this application. These eight variables are obtained from Federal Reserve Bank of St. Louis. The time horizon and frequency of these macroeconomic variables are the same as the previous dataset of financial institutions.

4.4 VAR model setup

Before a VAR estimation is performed, one have to ensure the stationarity of the data. The Augmented Dickey Fuller Test (ADF test) is conducted to test the stationarity of the time series. Table 3 shows that F1, F2, GDP, TEN are tested to be stationary. Then the first difference of HPI is taken, the CPI, INC are transformed as percent change, and SP is transformed into log returns, whose unit root behavior is suggested by the ADF test.

As for the order of this VAR model, we have kept it to be 15 based on different information criteria and autoregression tests, see Table 4 and Table 5. Figure 30 and Figure 31 in Appendix shows the autocorrelation function (acf) of the residuals, the acf of these variables are insignificant, which indicates the suitability of the VAR model. Stationary test is also conducted on the data to ensure the appropriateness of this model. Table 6 shows that the constructed VAR (15) model is stationary. As there is no roots in or on the complex circle, which is equivalent to the condition that all eigenvalues of reverse characteristic polynomial of the VAR(p) process have modulus less than 1 (see Lütkepohl (2005)).

4.5 Impulse Response Analysis

In this section an impulse response analysis is performed. As a structural VAR with correlated shocks is considered, the Choleski decomposition of the covariance matrices is applied in this study.

Firstly the impulse response from the two extracted financial tail event driven factors to macro economy is conducted. Left panel of Figure 6 shows the estimated impulse functions with bootstrap confidence intervals. For example, it is noticed that one standard deviation shock of F1 (market level factor) will have a significant positive impact on itself and F2 immediately. Also it will have a significant positive impact on GDP (in six period and from 14 to 17 period), SP and HPI immediately under 95% confidence bounds. It will decrease the CPI immediately and decrease UNE from 9th period on. From the magnitude of the response, it suggests that the positive shocks of the market level factor have large significant impact on GDP and unemployment rate.

As for the impulse effect of F2 (crisis sensitive factor) in right panel of Figure 6, it can be found that one standard deviation shock of crisis sensitive factor decreases the market level factor significantly in the first and fourth period under 95% confidence bounds. There is a significant positive impact on itself, it will decrease the GDP in the fourth period significantly. The employment will be increased from period 2 to period 11. There is no significant impact from F2 to other variables. It can be seen that GDP and unemployment rate are severely affected by crisis sensitive factor.

Next, the impulse response from macro economy to the financial market will be detected, which could be useful for central bank to check the sensitivity of the financial market shocked by the macroeconomic variables used in CCAR. One standard deviation shock of GDP (Left panel of Figure 7) will increase the F1 in the third period significantly, and decrease F2 from period 1 to period 3. It has immediate positive impact on itself, inflation will be decreased and personal income will be increased immediately, unemployment will be decreased from period 6 to period 20, which means that more people could find a job. However, house price will be increased significantly from period 2 on.

If there is one positive shock of CPI (Right panel of Figure 7), the market level factor F1 will be decreased significantly from 4th to 5th period. F2 will be increased in the third and tenth period significantly. GDP will be decreased from period 15 to period 18 significantly, there will be an immediate positive impact on CPI itself. The unemployment will be increased significantly from period 10 on.

One standard deviation shock of unemployment (Right panel of Figure 8) could increase F2 in the first period significantly, and increase itself immediately.

One standard positive shock of short term interest rate (Left panel of Figure 9) could increase market level factor in the third, seven and tenth period significantly, increase F2 from 3th to 4th periods, then decrease F2 from the 5th to 6th periods significantly, increase itself immediately and increases long term interest rate immediately until third period, and from period 10 to period 14 significantly, decrease house price from 16th period to 22th period significantly.

One standard positive shock of the long term interest rate (Right panel of Figure 9) could increase the short term interest rate at the beginning to 16th period and increase itself immediately. From here it can be concluded that the short term and long term interest rate reinforce each other.

Now the findings from impulse response analysis are summarized as follows: 1. GDP and unemployment rate are relative sensitive to the shocks from financial tail event driven factors. 2. Market level factor (F1) is more sensitive to the impact of GDP, CPI, THR, whereas crisis sensitive factor (F2) is more sensitive to CPI, UNE, THR. So the financial tail event driven factors are both sensitive to CPI and THR. 3. Shock of GDP decreases unemployment rate significantly, however, the unemployment rate does not have severe impact on GDP. 4. The short term and long term interest rate reinforce each other for long term.

4.6 Forecasting

Now the out-of-sample forecast result of this model is presented. Since the behavior of two financial tail event driven factors are more useful to indicate the possible risk levels in the financial market, the forecasting of them will be shown in the following, while the forecasting of macro variables is less interesting for us.

Figure 11 shows the 9 quarters forecasting of financial factors by using VAR model. Compared with higher values during 2008 financial crisis, the values in forecasting are not very high for both factors. The possible risk level in the financial market could be therefore predicted, which is very helpful in terms of risk controlling and supervision.

4.7 Network Analysis

In this section the network effects of financial factors and the macroeconomic variables are studied. Namely we would like to detect the connectedness between financial sector and macro economy. There is a lot of literature regarding network based connectedness. For instance, Hautsch et al. (2015) proposed realized systemic risk beta based on linear LASSO regression, which measures financial companies' contribution to systemic risk given network interdependence between firms' tail risk exposures. Härdle et al. (2016) identified Systemically Important Financial Institutions based on the network constructed by the components from Single-Index Model with LASSO technique. Diebold and Yilmaz (2014) apply generalized variance decomposition

(GVD) to measure the connectedness for financial firms. Since the VAR model is applied, the variance decomposition (VD) is the direct and simplest approach to build a network. However, using GVD (see Pesaran and Shin (1998)) or VD (Sims (1980)) is a trade-off. This is related to the issue about ordering of variables. While VD is sensitive to ordering, GVD is invariant to ordering. In Diebold and Yilmaz (2014), they detected the connectedness of the financial firms, each firm should be equally treated, GVD is good for them. But in this study, based on the result of Granger causality test from Table 7, it shows that two financial tail event driven factors Granger cause other variables, while some of the macroeconomic variables do not Granger cause others. So the financial factors should be placed first in the constructed VAR model, and then would be the macroeconomic variables, respectively. Another reason is that it is of interest to detect the network effect of macro economy changed by financial shocks. The ordering is an important feature in this analysis and can not be neglected, therefore VD is applied in this application.

Forecast error variance decomposition is carried out to construct the network, and the predictive horizon for variance decomposition is 12 months, as one year is normal economy circle, moreover, it can be found out there is no much difference of the estimation results between 6 months and 12 months. Window size n is chosen as 48 months, i.e. 4 years, since it could cover the longest duration of the recession in the history. Then the adjacency matrix in each time point is constructed and the diagonal element is set as zeros (since the impact from one variable to the others are more interesting than to itself), see Table 8. There are two levels connectedness, the aggregated level and the individual level. For the definition of them we reference Härdle et al. (2016). First of all, the direct connection from variable j to variable i is defined as $C^w_{i,j}$ for window w. The total direct connectedness matrix is simply the sum of each direct connection over all moving windows: $\sum_{w=1}^{T-n} C^w_{i,j}$. Where the column direction stands for the OUT-going (emit) link, i.e. how one variable affect the others. And the row connection means the IN-coming (receive) link, i.e. how one variable is affected by others. The total Incoming links is defined as $\sum_{j=1}^{K}\sum_{w=1}^{T-n} C^w_{i,j}$, while the total Outgoing links is defined as $\sum_{i=1}^{K}\sum_{w=1}^{T-n} C^w_{i,j}$. Table 9 shows the total IN and OUT links, it can be seen that the major impact emitter is market level factor with the value 314, the crisis sensitive factor and GDP follow it with 187 and 155 respectively. It indicates that the two financial tail event driven factors dominate the network.

Figure 12 shows the aggregated total direct connections, the highest direct connection is from F1 to SP with 129. The second direct connection is from

F1 to F2 with the value 80. The total connectedness is simply the sum of each adjacency matrix over time $\sum_{i=1}^{K}\sum_{j=1}^{K}\sum_{w=1}^{T-n} C_{i,j}^{w}$. Figure 13 shows the result of total connectedness: the connectedness in financial crisis period (from 2008 to 2010) is higher than the other periods. Incoming link ($\sum_{j=1}^{K} C_{i,j}^{w}$) and Outgoing link ($\sum_{i=1}^{K} C_{i,j}^{w}$) over time are shown in Figure 14 and Figure 15. Comparing these two graphs it can be found that the outgoing links is more volatile than incoming links. Very interesting phenomenon is that in the peak of the financial crisis the crisis sensitive factor dominates the whole network, which can be seen as a crisis indicator in this network.

Next the individual level connectedness (i.e. the direct connection) is detected. The direct connection among these variables is shown from Figure 16 to Figure 18. There are plots of direct connections on different dates: 11.28.1997 which is during the asian financial crisis, 30.11.2009 which is during global financial crisis, and 31.12.2015 which is during financial stable period. It is clear that during the crises the network is denser than the stable period.

5 $\tau = 0.5$ case

In $\tau = 0.5$ case the factor loadings are obtained, see left plot of Figure 19. The major difference of the first factor with $\tau = 0.05$ case can be found in Figure 20, in $\tau = 0.5$ case, the first factor have higher correlation (0.97) with the averaged mean, while in $\tau = 0.05$ case, the correlation of factor 1 and the averaged mean is 0.96. For the second factor, from the plot of the factor loadings in right plot of Figure 4 and right plot of Figure 19, it can be found that in $\tau = 0.05$ case the defaulted firms are more sensitive to factor two (AIG, FNMA, FDMC), while only AIG and FNMA are sensitive to factor two in $\tau = 0.5$ case. In Figure 21, we found that the second factors in the $\tau = 0.05$ and $\tau = 0.5$ cases is larger than the first factors, see also Table 12, especially the skewness and kurtosis of second factor in $\tau = 0.05$ case is much different from $\tau = 0.5$ case. After the comparison, it can be conclude that in median case, the first factor could be identified precisely, while in $\tau = 0.05$, i.e. high risk case, the second factor could be better identified to indicate the possible risk. In terms of systemic risk study, the second factor is more important to play a role of risk indicator, while the first factor which represents the average return is not so interesting for us.
In the VAR model, same order 15 is selected, see from Table 6 to Table 9. From the forecasting in Figure 22 it can be seen that the pattern of second factor is flatter than the predicted second factor in $\tau = 0.05$ case.
In the network analysis, we find out that total connectedness in $\tau = 0.05$

case is higher than in $\tau = 0.5$ case, see Table 9 and Table 11. Especially, for two financial factors, there is also clearly more connection between them in $\tau = 0.05$ case compared with $\tau = 0.5$, either total received impacts or total emit impacts, see Figure 12 and Figure 23. Also the macro variables emit more impact to the system in $\tau = 0.05$ case with the value 757 larger than 753 in $\tau = 0.5$ case. We also find out that macro variables play more important role in $\tau = 0.05$ case than in $\tau = 0.5$ case, while in $\tau = 0.05$ case the macro variables received the total impact from financial factors with value 406 less than in $\tau = 0.5$ case with value 408. They emit to the financial factors 81 impact in $\tau = 0.05$ case which is larger than 80 in $\tau = 0.5$ case. Also for the macro variables themselves they mutually received 1082 in $\tau = 0.05$ case impact larger than 1081 in $\tau = 0.5$ case, and emit 757 to macro system larger than 753 in $\tau = 0.5$ case. The total connection of macro variables is 676 in $\tau = 0.05$ case larger than 673 in $\tau = 0.5$ case.
We conclude that if the financial market is under higher risk then not only financial system itself will have more connectedness, but also the macro economic will be affected and will have more connectedness than in the financial stable case. Figure 12 and Figure 23 shows the aggregated total direct connections for $\tau = 0.5$ and $\tau = 0.05$ case. In Figure 12 i.e. $\tau = 0.05$ case, the highest direct connection is from F1 to SP with 129, it is similar as in $\tau = 0.5$ case (see Figure 23), while the direct connection from F1 to SP is 133, even higher than high risk case, this effect also confirms that the first factor in $\tau = 0.5$ case could be better identified as market level factor. The second direct connection $\tau = 0.05$ is from F1 to F2 with the value 80, whereas in $\tau = 0.5$ case, is from unemployment to house price index with the value 42. In Figure 25 we find out that the total connectedness in $\tau = 0.5$ (black line) and $\tau = 0.05$ (red line) case. The $\tau = 0.05$ case is always higher than $\tau = 0.5$ case. But they have the similar trends, the connectedness in financial crisis period (from 2008 to 2010) is higher than the other periods.

6 Conclusion

Tail event driven factor augmented dynamic model with financial market data in $\tau = 0.05$ case is applied in this study. In the first step, two tail event driven factors in the financial market are filtered by using PEC algorithm. In the second step a VAR model is constructed, includes extracted factors and macroeconomic variables. Then, the impulse response analysis is conducted and the sensitivity between the macro economy and the financial shocks are detected. After the forecast error variance decomposition, the network of these variables is built.

It can be concluded: 1. Hedge funds played positive roles on reducing effects of the financial crisis; 2. GDP, Unemployment rate are relative sensitive to the shock of financial tail event driven factors, whereas the financial tail event driven factors are more sensitive to the impulse of inflation and short term interest rate; 3. Short term and long term interest rates reinforce each other; 4. During 2008 crisis, the total connection among these variables became higher than normal periods, the crisis sensitive factor plays the most important role which could be a risk indicator; 5. The financial tail event driven factors are major impact emitters in the constructed network for the overall period; 6. The outgoing impacts of these variables are more volatile than the received impacts over time. 7. After the comparison, the second factor could be better identified in $\tau = 0.05$ case than in $\tau = 0.5$ case. Moreover, the network is denser in $\tau = 0.05$ case, and also the connections among macro variables are denser in tail case than in median case. 8. If the financial market is under higher risk then not only financial system itself will have more connectedness, but also the macro economic will be affected and will have more connectedness than in the financial stable case.

7 Appendix

7.1 Tables

	Depositories	Insurance	Broker-Dealers	Hedge Funds
Mean	0.18	0.17	0.20	0.21
SD	2.35	2.09	2.75	1.29

Table 1: Descriptive Statistics of four industry groups.

τ level	1 Factor	2 Factors	3 Factors
0.05	0.66	0.70	0.73
0.10	0.54	0.59	0.63
0.15	0.47	0.52	0.56
0.20	0.41	0.47	0.52
0.25	0.37	0.42	0.48
0.30	0.33	0.39	0.45
0.35	0.31	0.37	0.43
0.40	0.29	0.35	0.41
0.45	0.29	0.35	0.40
0.50	0.28	0.34	0.40
0.55	0.29	0.35	0.40
0.60	0.30	0.36	0.41
0.65	0.32	0.37	0.43
0.70	0.34	0.40	0.45
0.75	0.38	0.43	0.48
0.80	0.42	0.47	0.52
0.85	0.48	0.52	0.57
0.90	0.56	0.60	0.63
0.95	0.67	0.71	0.74

Table 2: Explained τ- variance with different number of factors, where τ from 0.05 to 0.95.

	Original variables	p-values	Original variables	p-values
$\tau = 0.05$ case	F1	0.01	F2	0.01
$\tau = 0.50$ case	F1	0.01	F2	0.01
both cases	GDP	0.01	CPI	0.36
	INC	0.85	UNE	0.32
	THR	0.09	TEN	0.01
	SP	0.70	HPI	0.20

Table 3: ADF test statistics for all variables.

	Criteria	AIC	HQ	SC	FPE
Selected order	$\tau = 0.05$	20	4	2	7
	$\tau = 0.50$	20	4	1	7

Table 4: Model specification: different criteria for both $\tau = 0.05$ and $\tau = 0.5$.

	Four tests	PT (asymptotic)	PT (adjusted)	BG	ES
$\tau = 0.05$	Order 2	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-16}
	Order 4	1.77×10^{-11}	2.4×10^{-14}	2.2×10^{-16}	2.2×10^{-16}
	Order 7	8.4×10^{-08}	5.7×10^{-10}	5.9×10^{-15}	2.4×10^{-04}
	Order 15	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-16}	$1.8 \times 10^{-01}*$
	Order 20	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-03}
$\tau = 0.50$	Order 1	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-16}
	Order 4	2.3×10^{-10}	4.6×10^{-13}	2.2×10^{-16}	2.2×10^{-16}
	Order 7	5.7×10^{-07}	5.4×10^{-09}	2.2×10^{-16}	1.2×10^{-05}
	Order 15	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-16}	$4.5 \times 10^{-01}*$
	Order 20	2.2×10^{-16}	2.2×10^{-16}	2.2×10^{-16}	1.7×10^{-02}

Table 5: Model Selection: different tests for both $\tau = 0.05$ and $\tau = 0.5$.

	Minimum	1st Quantile	Median	Mean	3rd Quantile	Maximum
$\tau = 0.05$	0.16	0.89	0.94	0.90	0.96	0.99
$\tau = 0.50$	0.40	0.89	0.93	0.91	0.96	0.98

Table 6: Eigenvalues of the companion coefficient matrix for VAR models.

Cause	F1	F2	GDP	CPI	INC
p value	7.4×10^{-03}	3.8×10^{-05}	4.6×10^{-05}	2.6×10^{-01}	3.2×10^{-01}
Cause	UNE	THR	TEN	SP	HPI
p value	3.2×10^{-01}	8.7×10^{-04}	1.2×10^{-01}	2.2×10^{-01}	7.8×10^{-03}

Table 7: p vlaues of Granger causality test of each variable, while the effect variables are other variables except for the cause variable, where $\tau = 0.05$.

$$A^w = \begin{array}{c c} & \begin{array}{ccccc} V_1 & V_2 & V_3 & \cdots & V_K \end{array} \\ \begin{array}{c} V_1 \\ V_2 \\ V_3 \\ \vdots \\ V_K \end{array} & \begin{pmatrix} 0 & \widehat{C}^w_{12} & \widehat{C}^w_{13} & \cdots & \widehat{C}^w_{1K} \\ \widehat{C}^w_{21} & 0 & \widehat{C}^w_{23} & \cdots & \widehat{C}^w_{2K} \\ \widehat{C}^w_{31} & \widehat{C}^w_{32} & 0 & \cdots & \widehat{C}^w_{3K} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{C}^w_{K1} & \widehat{C}^w_{K2} & \widehat{C}^w_{K3} & \cdots & 0 \end{pmatrix} \end{array}$$

Table 8: $K \times K$ adjacency matrix for the constructed network at the wth window, V_k stands for the name of variables.

	F1	F2	GDP	CPI	INC	UNE	THR	TEN	SP	HPI	Total SUM
Total IN	56	120	118	80	67	136	175	162	194	150	1258
Total OUT	314	187	155	89	68	136	73	96	45	94	1258

Table 9: Total IN and OUT links over all periods by using forecast error variance decomposition, where $\tau = 0.05$, window size is 48 months, the predictive horizon for the underlying variance decomposition is 12 months, $T - n = 239$.

Cause	F1	F2	GDP	CPI	INC
p value	3.7×10^{-02}	3.4×10^{-04}	2.2×10^{-05}	2.7×10^{-01}	3.4×10^{-01}
Cause	UNE	THR	TEN	SP	HPI
p value	8.7×10^{-02}	1.3×10^{-05}	8.9×10^{-02}	5.2×10^{-02}	6.3×10^{-02}

Table 10: p vlaues of Granger causality test of each variable, while the effect variables are other variables except for the cause variable, where $\tau = 0.5$.

	F1	F2	GDP	CPI	INC	UNE	THR	TEN	SP	HPI	Total SUM
Total IN	54	74	113	78	65	135	174	169	198	148	1209
Total OUT	283	173	167	90	63	142	72	93	34	92	1209

Table 11: Total IN and OUT links over all periods by using forecast error variance decomposition, where $\tau = 0.5$, window size is 24 months, the predictive horizon for the underlying variance decomposition is 12 months, $T - n = 239$.

		Mean	Standard Deviation	Skewness	Kurtosis
F1	$\tau = 0.05$	0.05	0.48	-1.08	8.08
	$\tau = 0.5$	0.06	0.49	-1.26	7.85
F2	$\tau = 0.05$	0.02	0.22	3.92	45.95
	$\tau = 0.5$	0.01	0.23	1.38	25.53

Table 12: The moments for factors in both $\tau = 0.05$ and $\tau = 0.5$ cases.

Error term	τ	N	T	Factor 1	Factor 2	Loading 1	Loading 2	MSE
N(0,0.1)	0.05	100	100	0.97	0.77	0.93	0.51	3.00
		100	1000	0.98	0.82	0.98	0.43	2.91
		1000	100	0.97	0.78	0.93	0.51	2.94
		1000	1000	1.00	0.83	0.99	0.27	3.36
	0.5	100	100	0.97	0.94	0.97	0.95	1.03
		100	1000	0.98	0.98	0.99	0.97	0.75
		1000	100	0.99	0.95	0.98	0.98	0.94
		1000	1000	1.00	0.99	1.00	1.00	0.34
	0.95	100	100	0.96	0.77	0.92	0.51	2.97
		100	1000	0.98	0.82	0.98	0.45	2.95
		1000	100	0.98	0.78	0.94	0.49	3.05
		1000	1000	1.00	0.82	0.99	0.30	3.35
t(5)	0.05	100	100	0.96	0.75	0.92	0.53	3.77
		100	1000	0.98	0.81	0.98	0.50	3.58
		1000	100	0.98	0.80	0.93	0.54	3.67
		1000	1000	1.00	0.86	0.99	0.35	3.76
	0.5	100	100	0.97	0.94	0.97	0.95	2.11
		100	1000	0.99	0.98	0.99	0.97	1.64
		1000	100	0.99	0.95	0.98	0.97	1.91
		1000	1000	1.00	0.99	1.00	0.99	1.55
	0.95	100	100	0.96	0.73	0.92	0.48	3.78
		100	1000	0.98	0.80	0.98	0.48	3.68
		1000	100	0.98	0.78	0.93	0.50	3.70
		1000	1000	1.00	0.86	0.99	0.37	3.71

Table 13: The average R^2 of the regression of simulated factors (or loadings) on the estimated factors (or loadings), and the average mean squared error of simulated observations and the estimators in different error assumptions under different expectile levels in 500 simulations.

Error term	τ	N	T	Factor 1	Factor 2	Loading 1	Loading 2	MSE
N(0,0.1)	0.05	100	100	0.97	0.92	0.95	0.94	1.04
		100	1000	0.99	0.98	0.99	0.97	0.71
		1000	100	0.97	0.78	0.93	0.51	2.94
		1000	1000	1.00	0.83	0.99	0.27	3.36
	0.5	100	100	0.98	0.94	0.97	0.95	0.98
		100	1000	0.99	0.99	0.99	0.97	0.69
		1000	100	0.99	0.95	0.98	0.98	0.94
		1000	1000	1.00	0.99	1.00	1.00	0.34
	0.95	100	100	0.96	0.91	0.95	0.93	1.13
		100	1000	0.99	0.98	0.99	0.97	0.70
		1000	100	0.98	0.78	0.94	0.49	3.05
		1000	1000	1.00	0.82	0.99	0.30	3.35
t(5)	0.05	100	100	0.96	0.75	0.92	0.53	3.77
		100	1000	0.98	0.81	0.98	0.50	3.58
		1000	100	0.98	0.80	0.93	0.54	3.67
		1000	1000	1.00	0.86	0.99	0.35	3.76
	0.5	100	100	0.97	0.94	0.97	0.95	2.11
		100	1000	0.99	0.98	0.99	0.97	1.64
		1000	100	0.99	0.95	0.98	0.97	1.91
		1000	1000	1.00	0.99	1.00	0.99	1.55
	0.95	100	100	0.96	0.73	0.92	0.48	3.78
		100	1000	0.98	0.80	0.98	0.48	3.68
		1000	100	0.98	0.78	0.93	0.50	3.70
		1000	1000	1.00	0.86	0.99	0.37	3.71

Table 14: The average R^2 of the regression of simulated factors (or loadings) on the estimated factors (or loadings), and the average mean squared error of simulated observations and the estimators in different error assumptions under different expectile levels in 500 simulations, use PEC algorithm.

Error term	τ	N	T	MSE of Loadings	MSE of Factors	Total MSE
N(0,0.1)	0.05	100	100	0.20	0.10	2.97
		100	1000	0.22	0.08	3.12
		1000	100	0.19	0.11	3.23
		1000	1000	0.21	0.07	3.38
	0.5	100	100	0.06	0.07	1.12
		100	1000	0.04	0.03	0.68
		1000	100	0.03	0.05	0.87
		1000	1000	0.01	0.02	0.43
	0.95	100	100	0.21	0.13	3.57
		100	1000	0.20	0.08	2.93
		1000	100	0.19	0.11	2.90
		1000	1000	0.20	0.08	3.23
t(5)	0.05	100	100	0.20	0.11	3.72
		100	1000	0.18	0.08	3.61
		1000	100	0.18	0.10	3.70
		1000	1000	0.20	0.08	3.75
	0.5	100	100	0.07	0.06	1.87
		100	1000	0.05	0.04	2.00
		1000	100	0.05	0.06	2.03
		1000	1000	0.02	0.02	1.74
	0.95	100	100	0.17	0.11	3.54
		100	1000	0.20	0.09	3.83
		1000	100	0.18	0.11	3.61
		1000	1000	0.18	0.07	3.57

Table 15: The average MSE of simulated factors (or loadings) and the estimated factors (or loadings), and the average mean squared error of simulated observations and the estimators in different error assumptions under different expectile levels in 100 simulations.

Error term	τ	N	T	MSE of Loadings	MSE of Factors	Total MSE
N(0,0.1)	0.05	100	100	0.08	0.09	0.18
		100	1000	0.07	0.07	0.20
		1000	100	0.07	0.10	0.18
		1000	1000	0.03	0.04	0.18
	0.5	100	100	0.05	0.05	0.10
		100	1000	0.04	0.03	0.10
		1000	100	0.03	0.04	0.10
		1000	1000	0.01	0.01	0.10
	0.95	100	100	0.09	0.10	0.19
		100	1000	0.07	0.06	0.20
		1000	100	0.07	0.09	0.20
		1000	1000	0.06	0.05	0.29
t(5)	0.05	100	100	0.09	0.11	1.96
		100	1000	0.07	0.05	1.97
		1000	100	0.06	0.10	1.97
		1000	1000	0.04	0.04	1.97
	0.5	100	100	0.08	0.06	1.25
		100	1000	0.04	0.05	1.28
		1000	100	0.03	0.05	1.27
		1000	1000	0.01	0.02	1.29
	0.95	100	100	0.09	0.09	1.96
		100	1000	0.07	0.06	1.97
		1000	100	0.07	0.07	1.97
		1000	1000	0.04	0.05	1.99

Table 16: The average MSE of simulated factors (or loadings) and the estimated factors (or loadings), and the average mean squared error of simulated observations and the estimators in different error assumptions under different expectile levels in 100 simulations, using iteration.

7.2 Figures

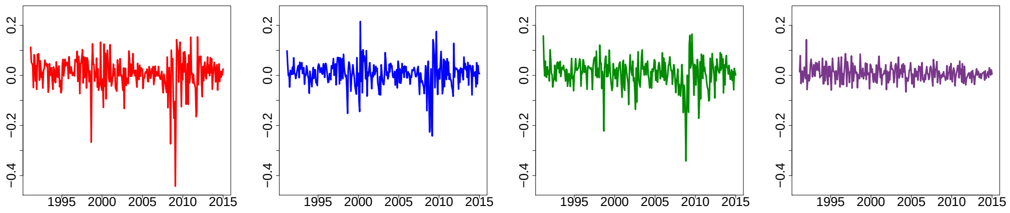

Figure 1: Average mean of each industry groups. From left to right: Depositories (red), Insurance (blue), Broker-Dealers (green), Hedge Funds (violet).

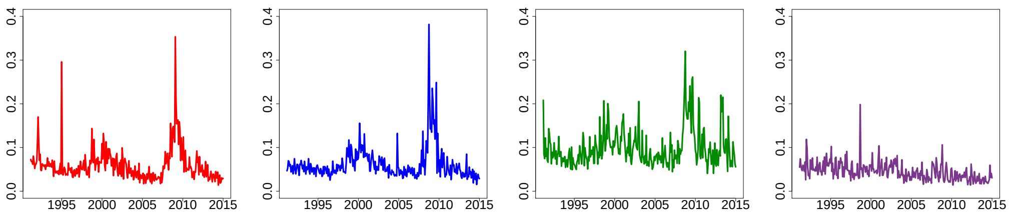

Figure 2: Average standard deviation of each industry groups. From left to right: Depositories (red), Insurance (blue), Broker-Dealers (green), Hedge Funds (violet).

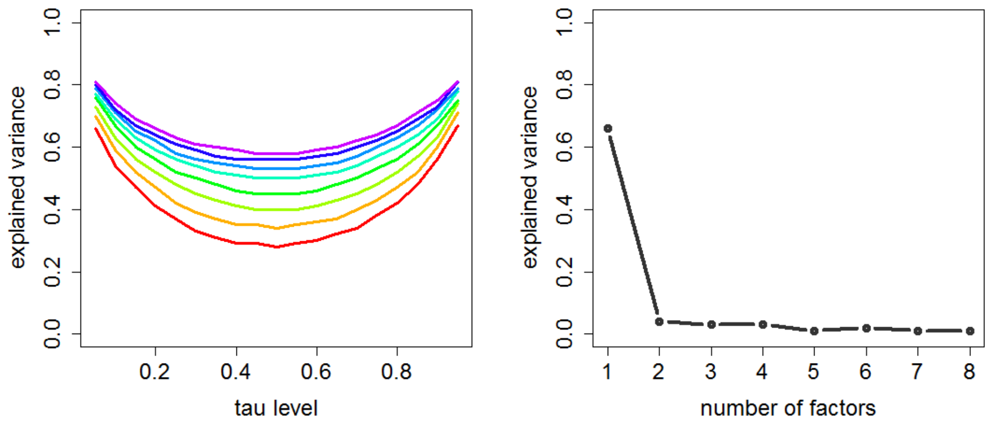

Figure 3: Left graph: x-axis represents the different τ levels, y-axis is the explained variance. Colors from red to violet (bottom to top) stand for the models with 1 to 8 factors. Right graph: the scree plot with $\tau = 0.05$.

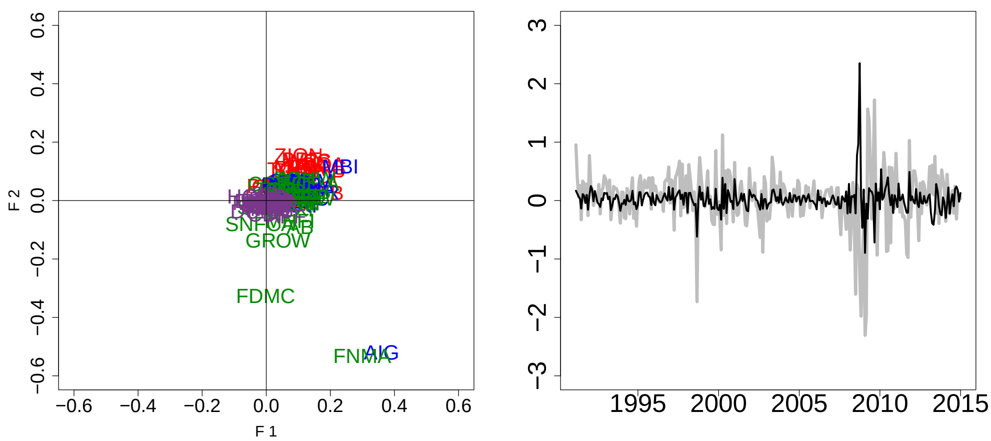

Figure 4: Left: The estimated factor loadings, Depositories (red), Insurance (blue), Broker-Dealers and Real Estates (green), Hedge Funds (violet). Right: The estimated factors with $\tau = 0.05$, F1 (thicker grey) and F2 (thinner black).

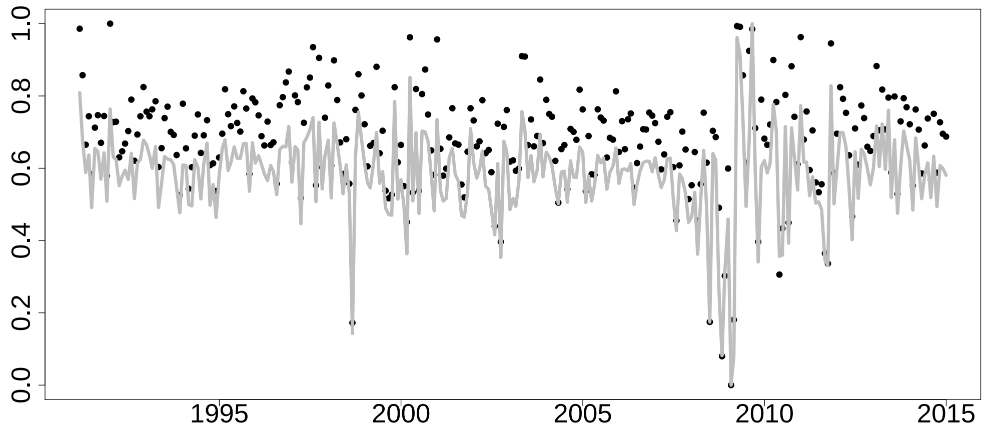

Figure 5: The estimated factor 1 (grey line) and average returns of all firms (black points) with $\tau = 0.05$. For comparison, both of them are scaled into zero and one.

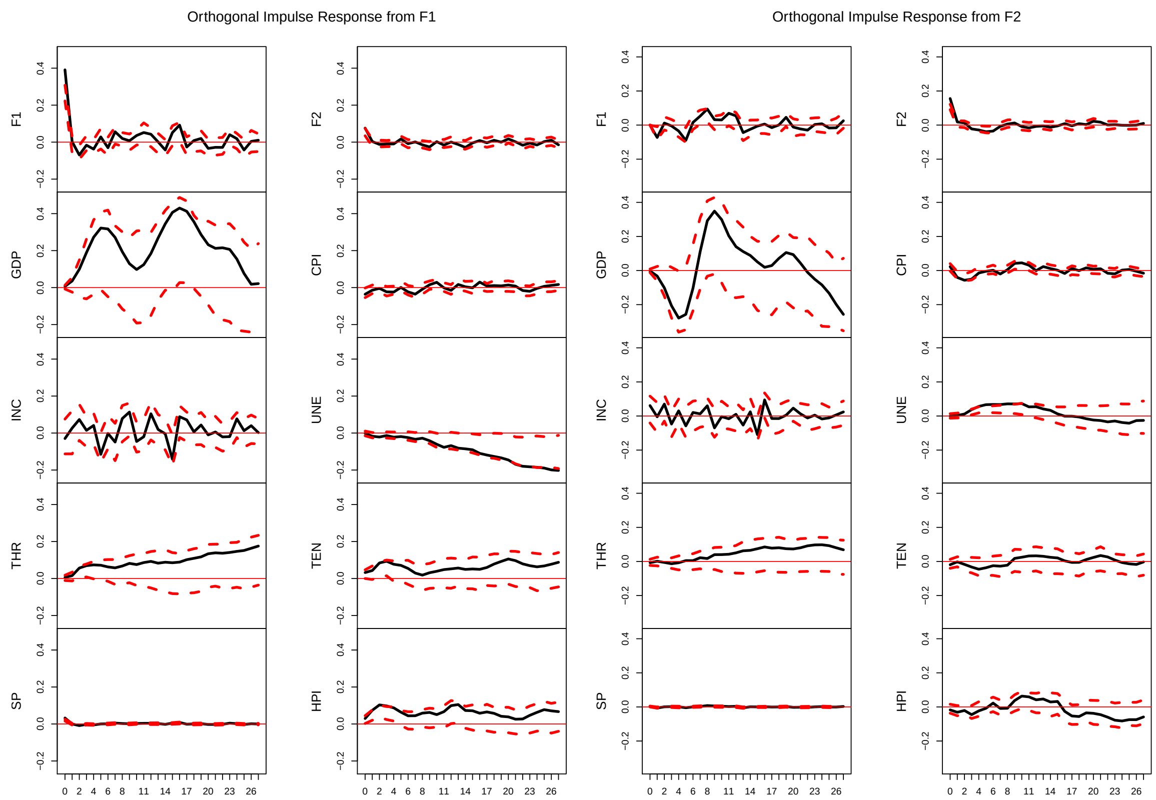

Figure 6: Impulse Response analysis of F1 and F2 for $\tau = 0.05$.

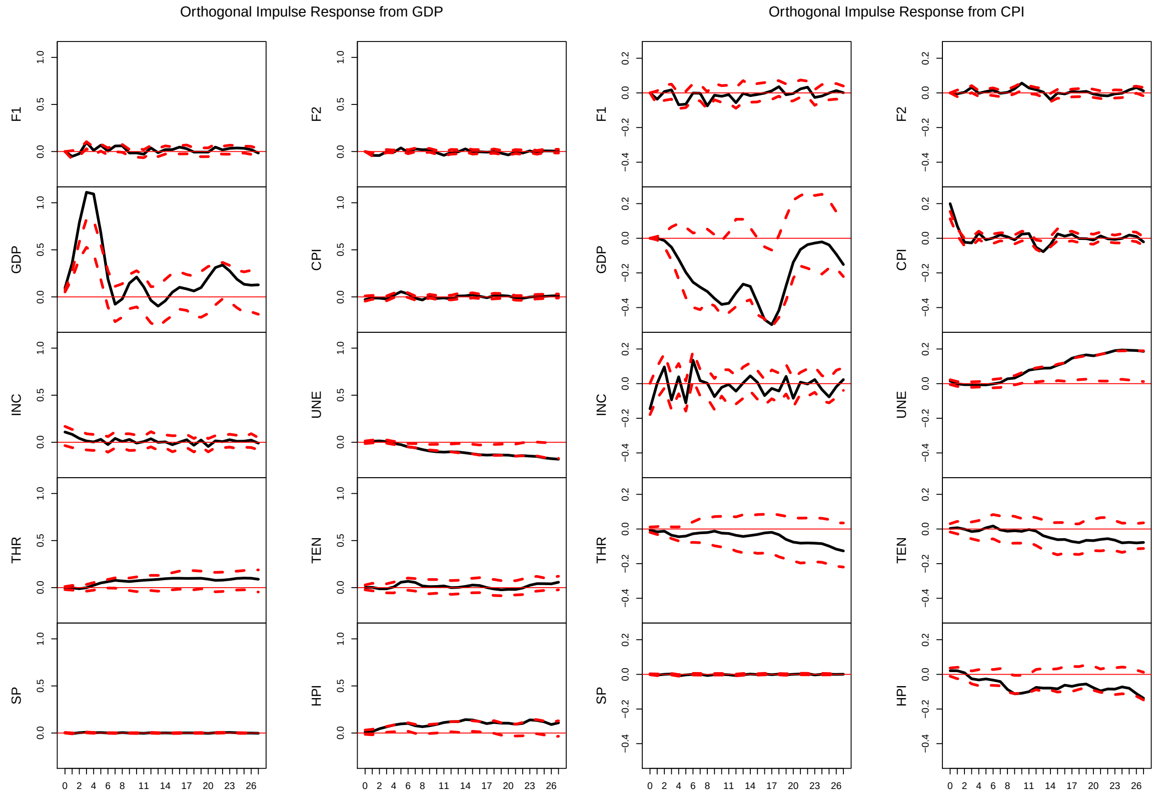

Figure 7: Impulse Response analysis of GDP and CPI for $\tau = 0.05$.

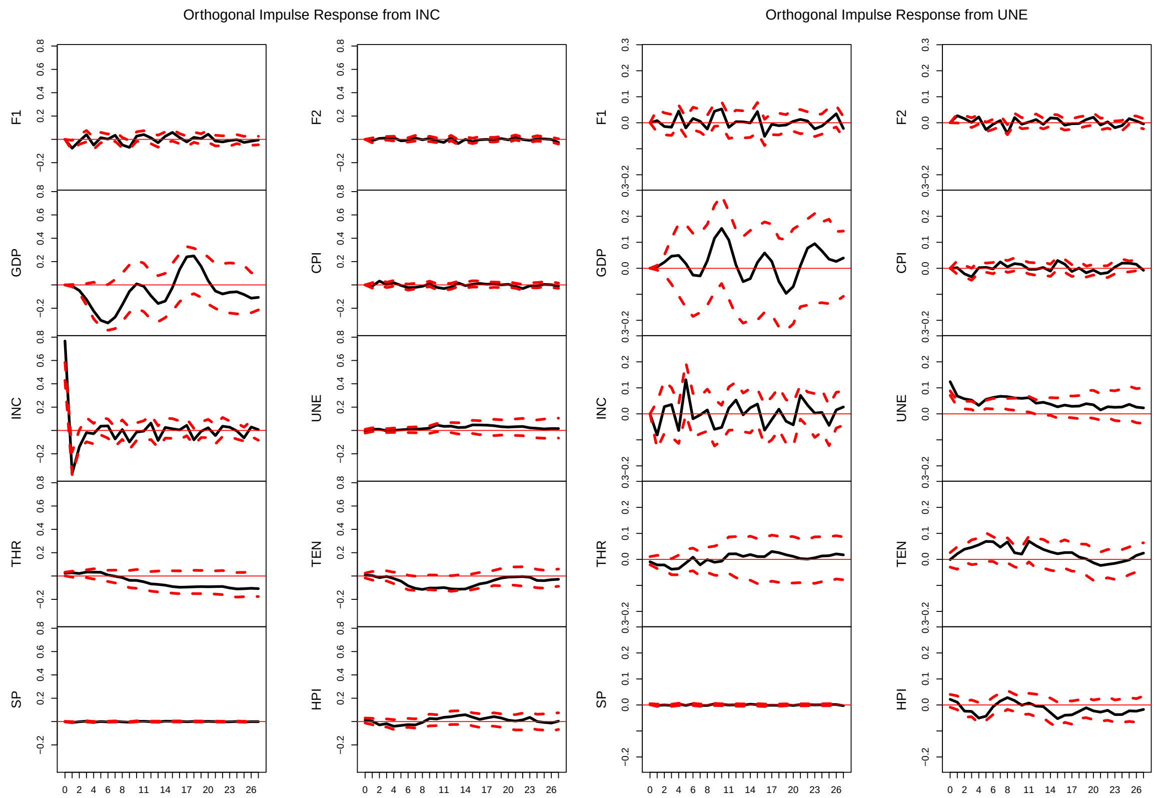

Figure 8: Impulse Response analysis of INC and UNE for $\tau = 0.05$.

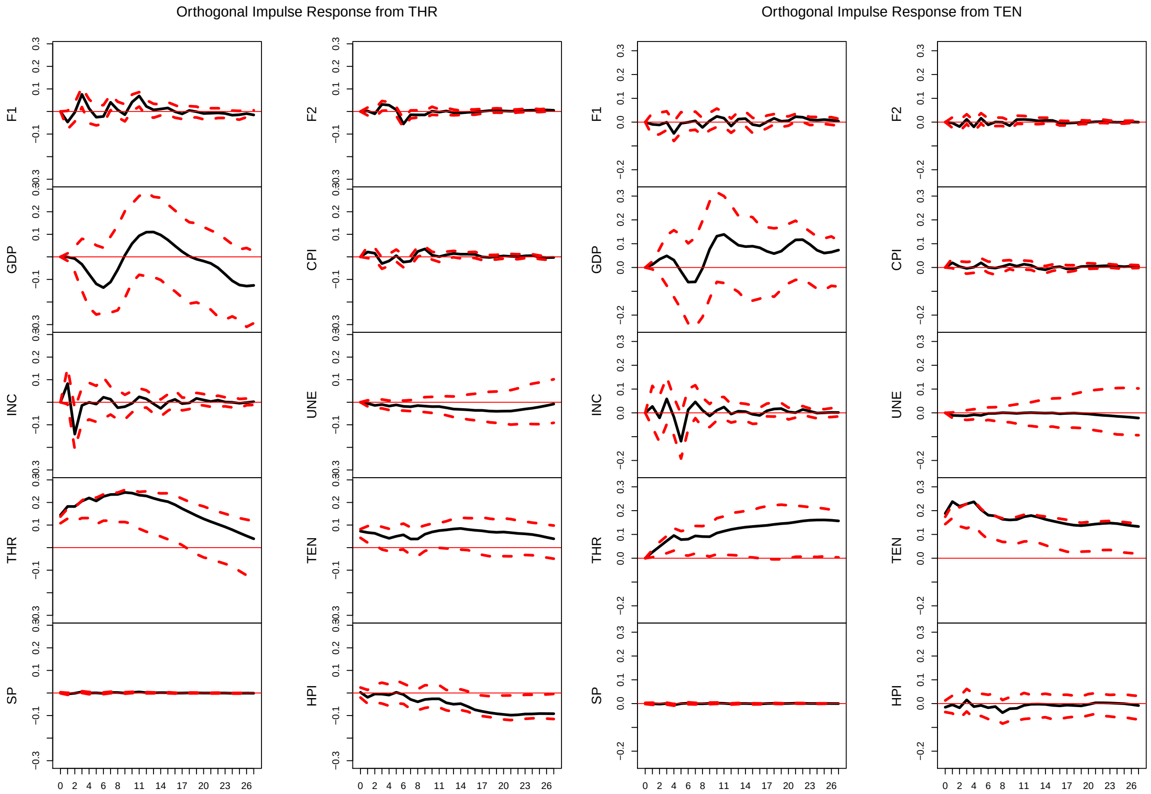

Figure 9: Impulse Response analysis of THR and TEN for $\tau = 0.05$.

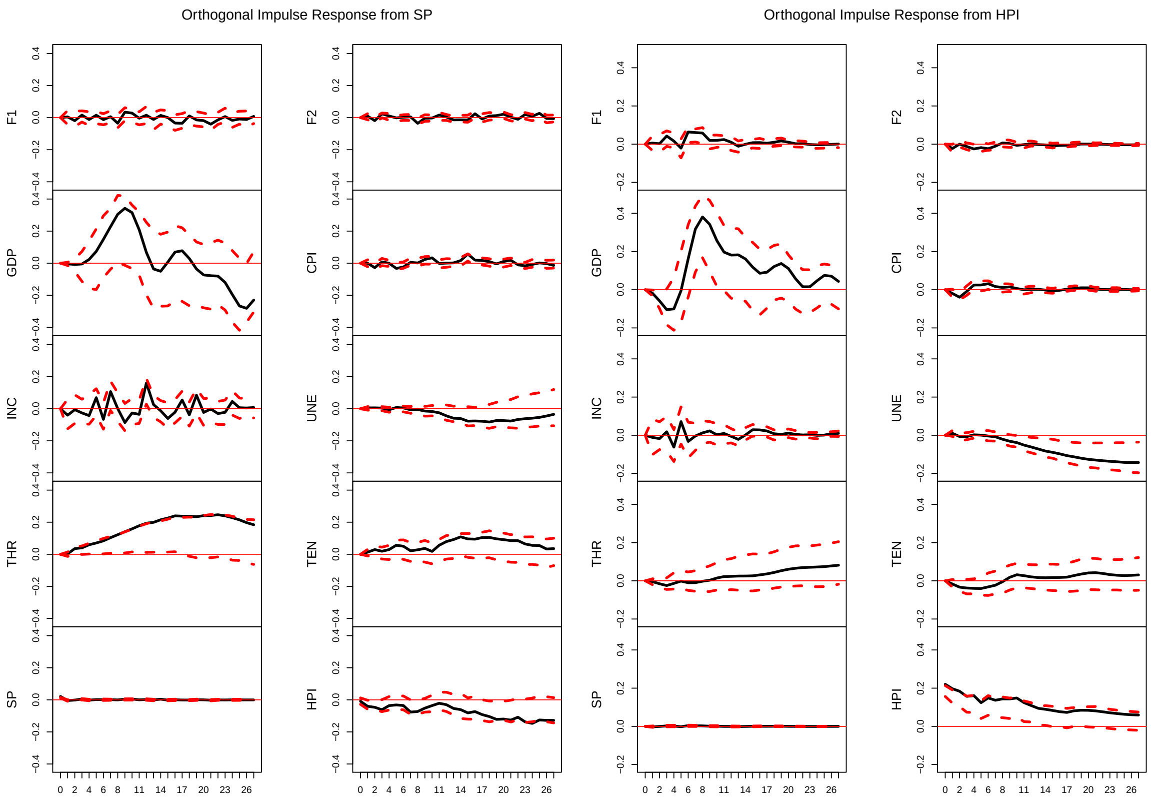

Figure 10: Impulse Response analysis of SP and HPI for $\tau = 0.05$.

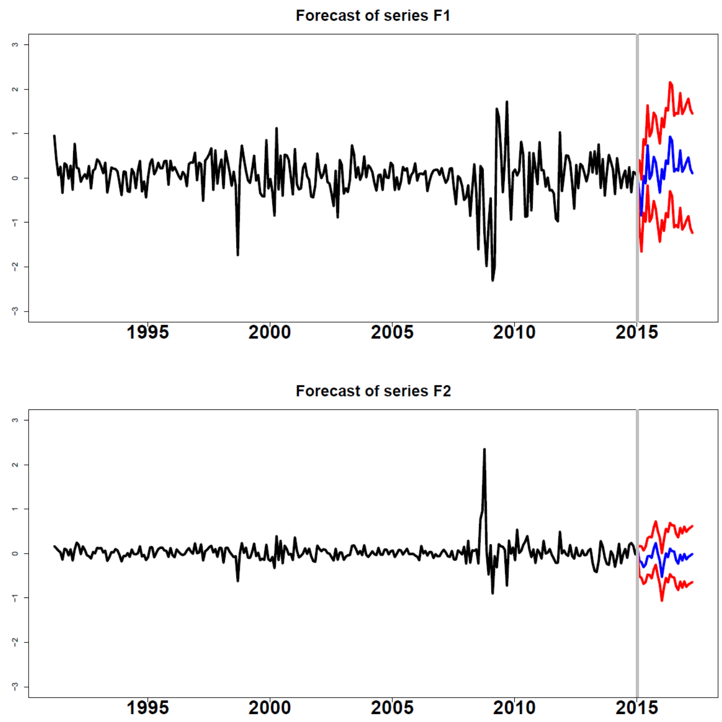

Figure 11: Forecasting of two financial factors with $\tau = 0.05$ in 27 months (9 quarters from 31-01-2015 to 31-03-2017).

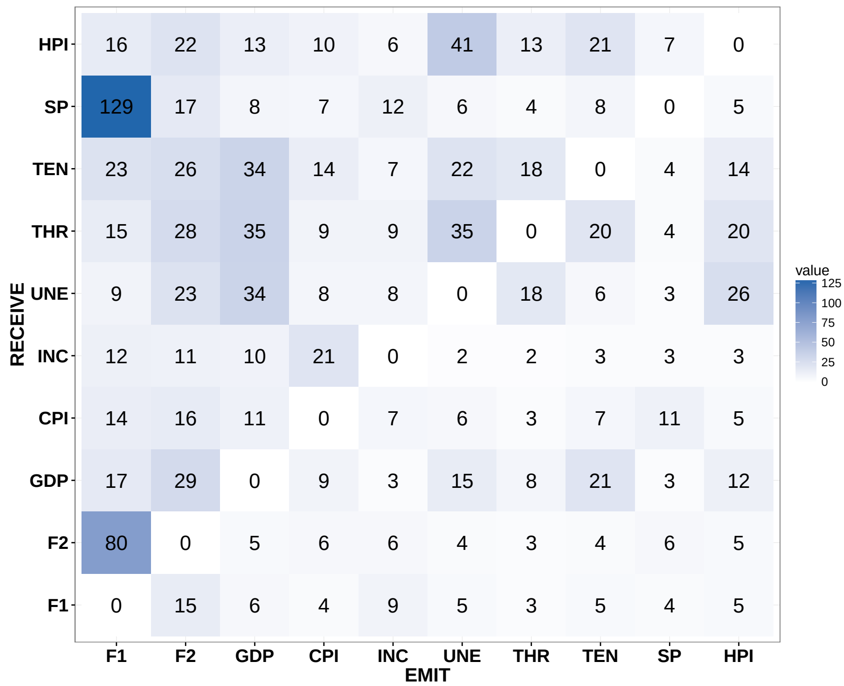

Figure 12: Stacked total direct connectedness over all periods by using forecast error variance decomposition, window size n is 48 months, the predictive horizon for the underlying variance decomposition is 12 months, $T - n = 239$

Figure 13: Total connectedness by using forecast error variance decomposition, window size is 48 months, the predictive horizon for the underlying variance decomposition is 12 months.

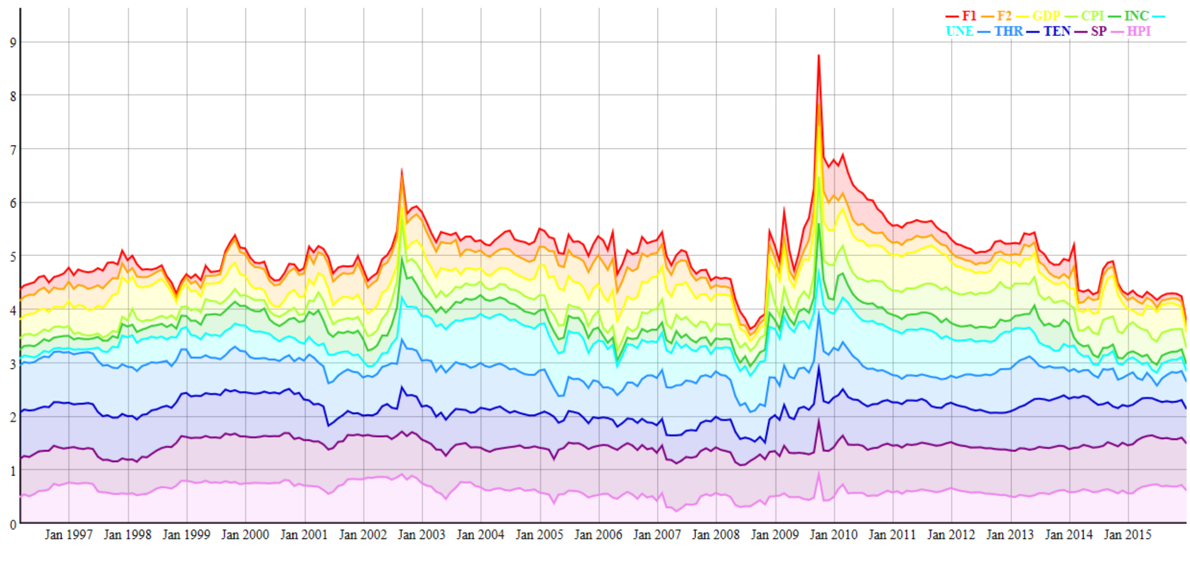

Figure 14: Total Incoming connectedness over all periods by using forecast error variance decomposition, window size n is 48 months, the predictive horizon for the underlying variance decomposition is 12 months, $T - n = 239$.

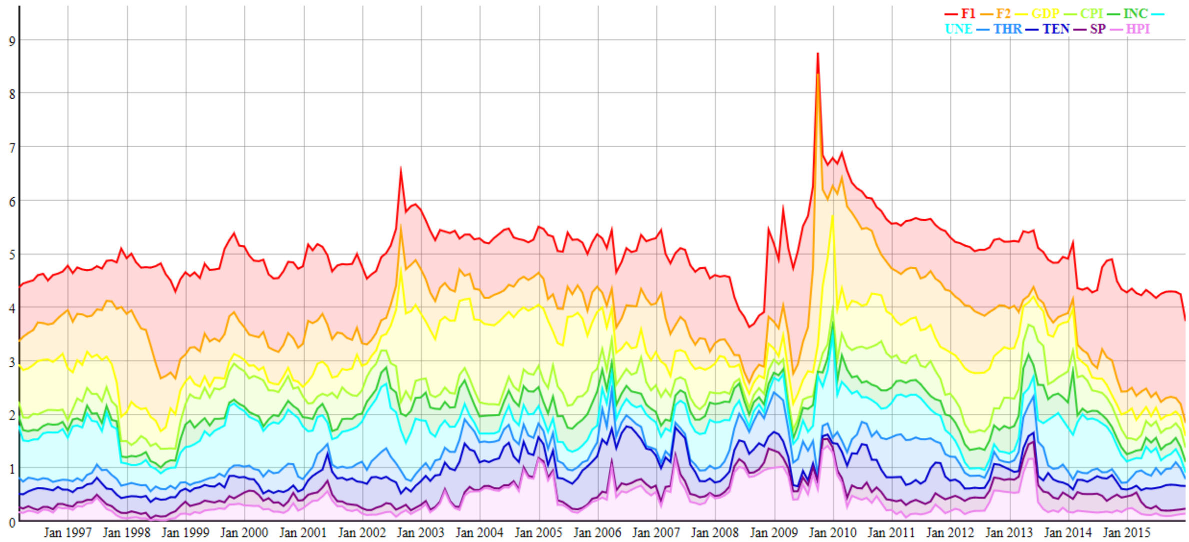

Figure 15: Total Outgoing connectedness over all periods by using forecast error variance decomposition, window size n is 48 months, the predictive horizon for the underlying variance decomposition is 12 months, $T - n = 239$

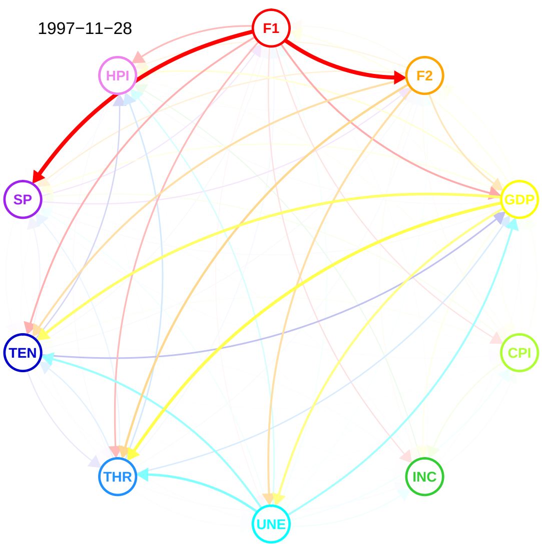

Figure 16: Network by using forecast error variance decomposition on 30.11.2009, window size is 48 months, the predictive horizon for the underlying variance decomposition is 12 months.

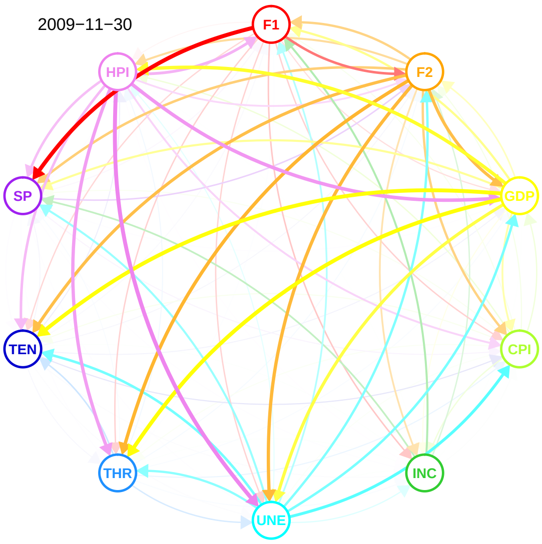

Figure 17: Network by using forecast error variance decomposition on 30.11.2009, window size is 48 months, the predictive horizon for the underlying variance decomposition is 12 months.

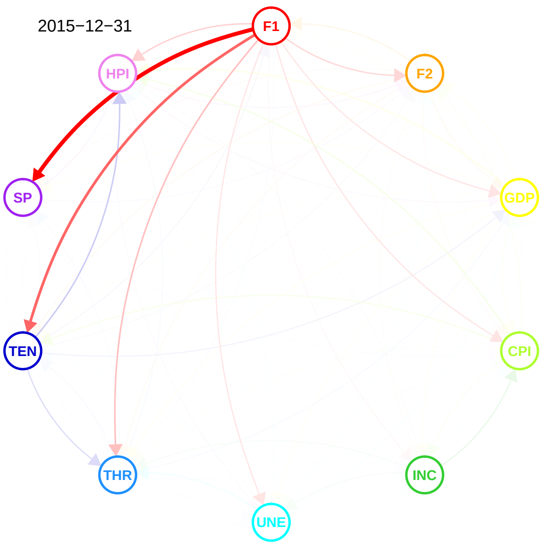

Figure 18: Network by using forecast error variance decomposition on 31.12.2015, window size is 48 months, the predictive horizon for the underlying variance decomposition is 12 months.

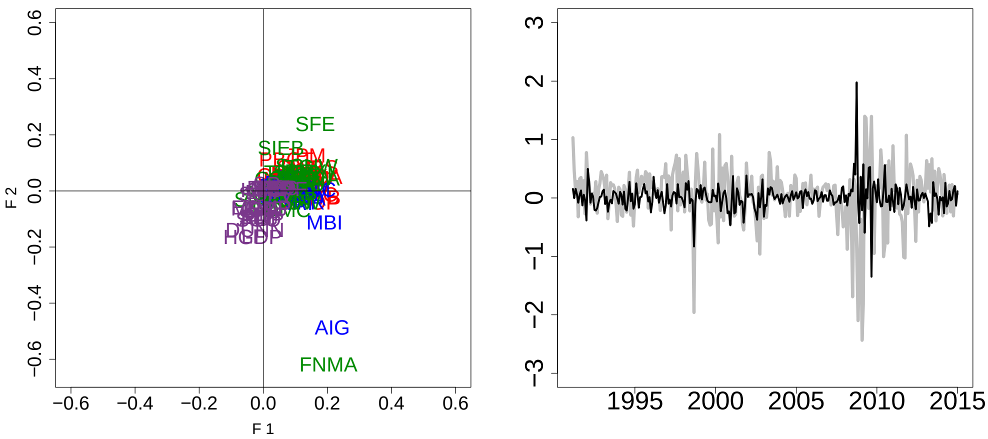

Figure 19: Left: The estimated factor loadings, Depositories (red), Insurance (blue), Broker-Dealers and Real Estates (green), Hedge Funds (violet). Right: The estimated factors with $\tau = 0.5$, F1 (thicker grey) and F2 (thinner black).

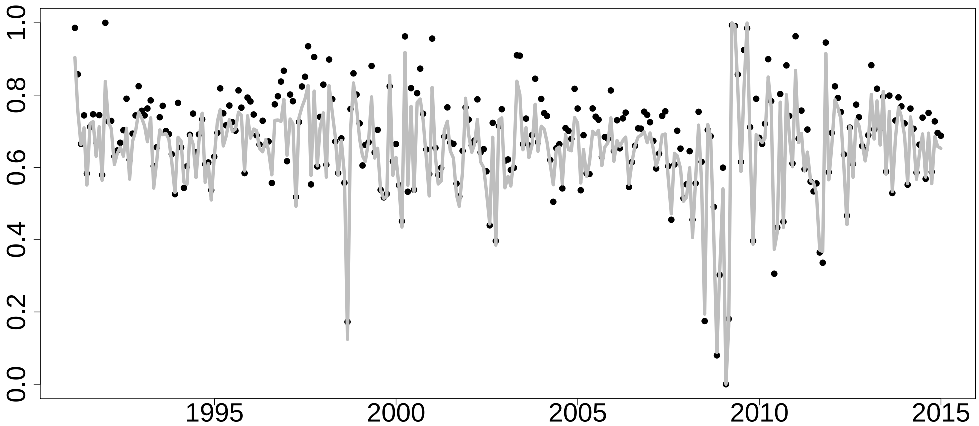

Figure 20: The estimated factor 1 (grey line) and average returns of all firms (black points) with $\tau = 0.5$. For comparison, both of them are scaled into zero and one.

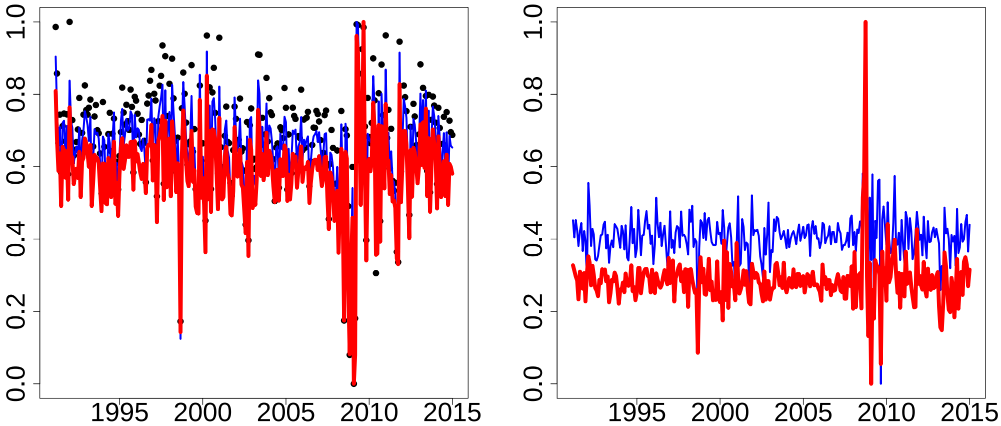

Figure 21: Left: the estimated factor 1 with $\tau = 0.05$ (red line) and $\tau = 0.5$ (blue line). Right: the estimated factor 2 with $\tau = 0.05$ (red line) with $\tau = 0.5$ (blue line).

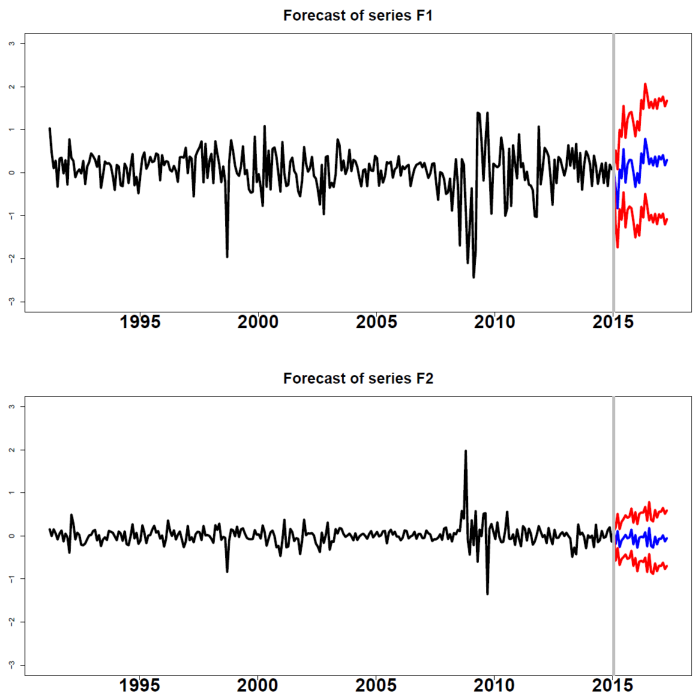

Figure 22: Forecasting of two financial factors with $\tau = 0.5$ in 27 months (9 quarters from 31-01-2015 to 31-03-2017).

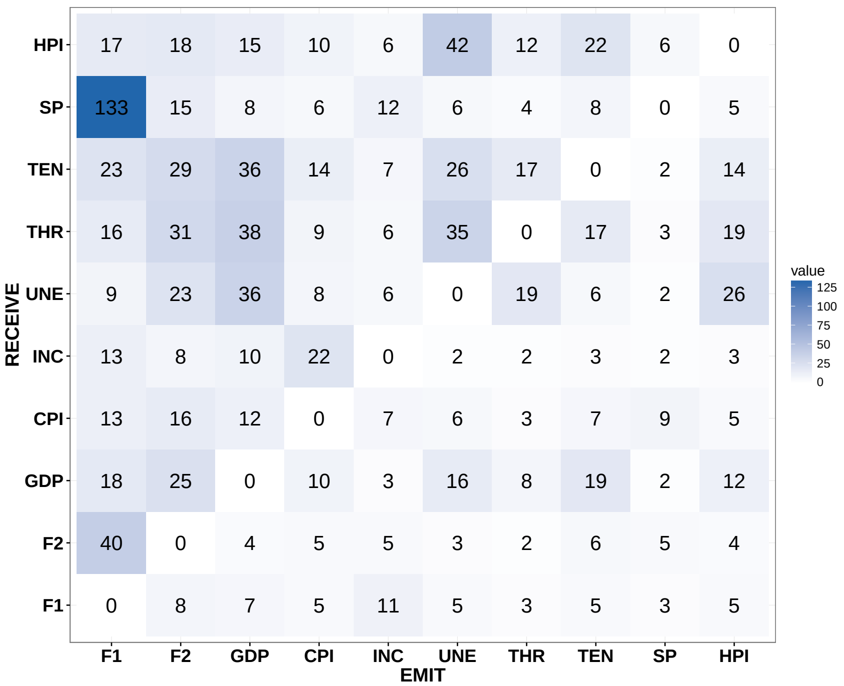

Figure 23: Stacked total direct connectedness over all periods by using forecast error variance decomposition with $\tau = 0.05$ and $\tau = 0.5$, window size is 48 months, the predictive horizon for the underlying variance decomposition is 12 months, $T - n = 239$

Figure 24: Total connectedness by using forecast error variance decomposition for $\tau = 0.05$ (red line) and $\tau = 0.5$ (black line), window size is 48 months, the predictive horizon for the underlying variance decomposition is 12 months.

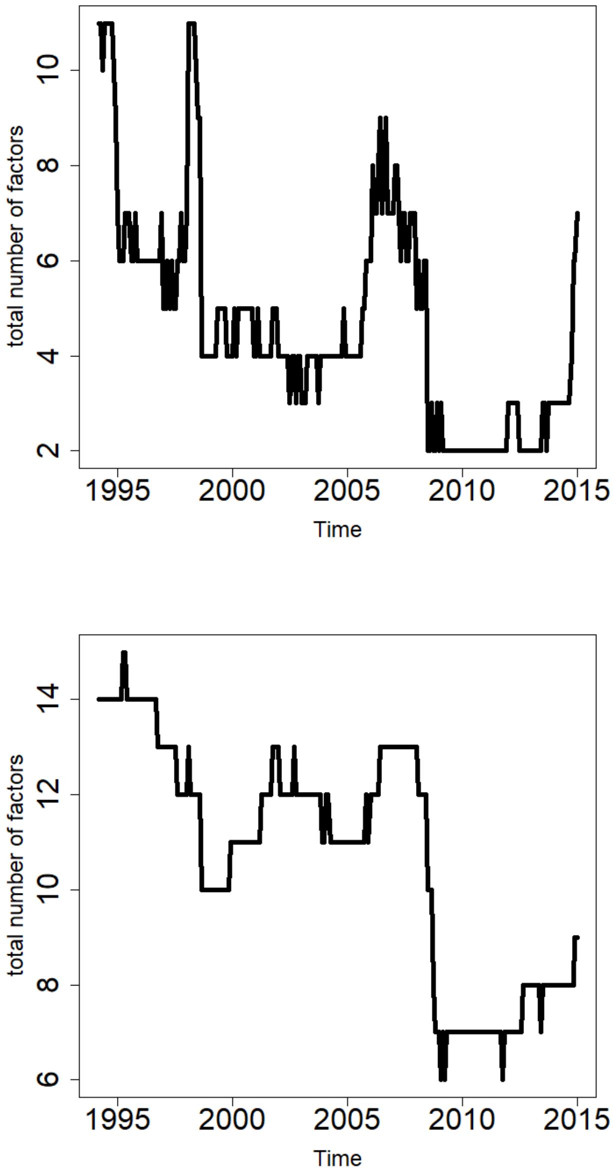

Figure 25: Number of factors over time, window size is 36, $\tau = 0.05$.

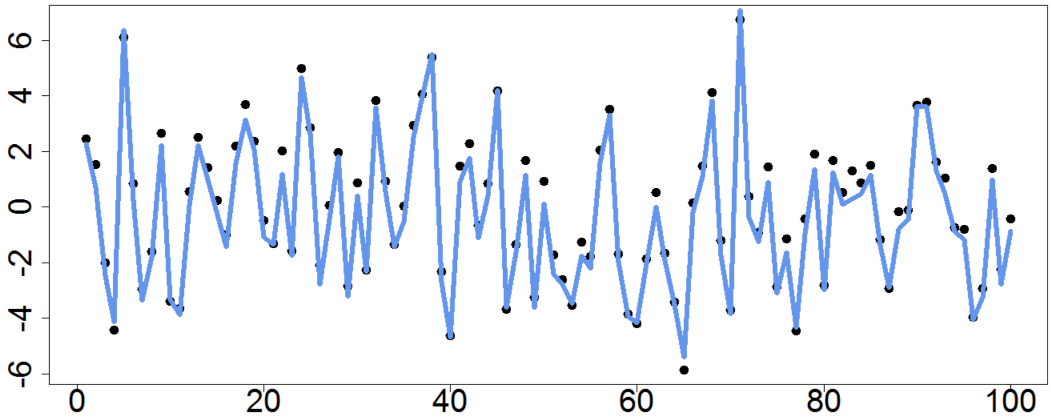

Figure 26: The simulated observations (black) and the estimators (blue) by using expectile regression, $\tau = 0.05$, $t = 1$, with iid $N(0, 0.01)$ error.

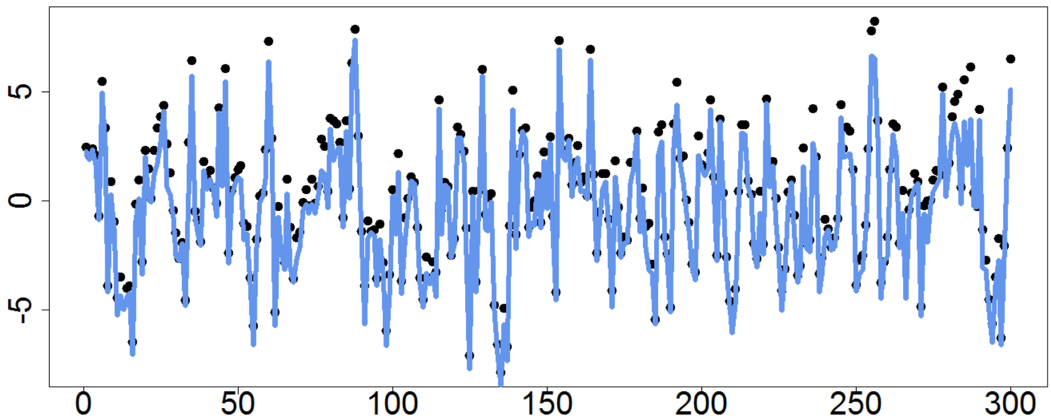

Figure 27: The simulated observations (black) and the estimators (blue) by using expectile regression, $\tau = 0.05$, $i = 1$, with iid $N(0, 0.01)$ error.

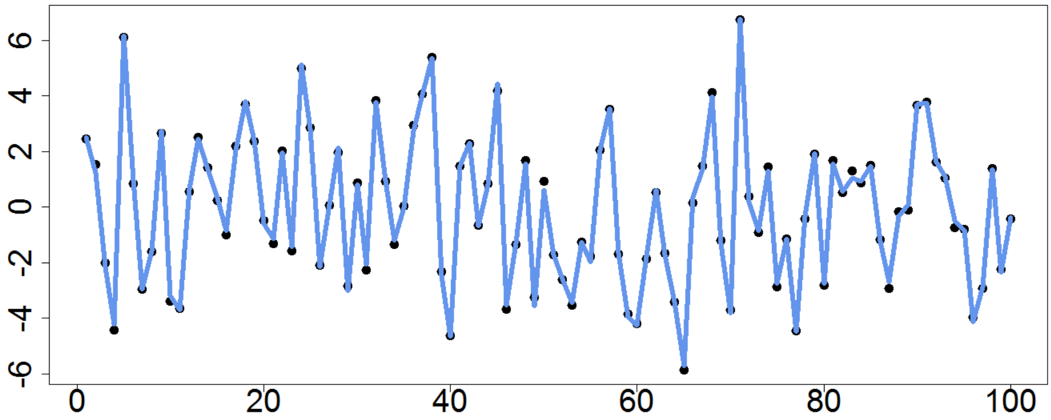

Figure 28: The simulated observations (black) and the estimators (blue) by using expectile regression, $\tau = 0.5$, $t = 1$, with iid $N(0, 0.01)$ error.

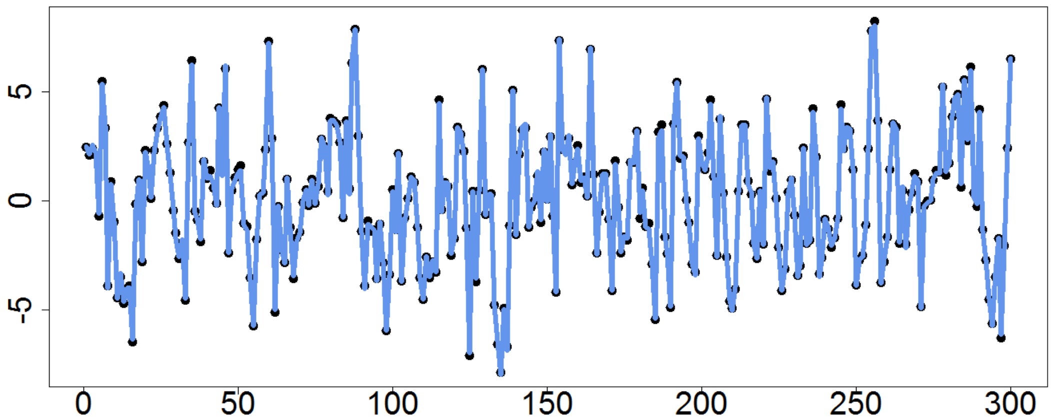

Figure 29: The simulated observations (black) and the estimators (blue) by using expectile regression, $\tau = 0.5$, $i = 1$, with iid $N(0, 0.01)$ error.

7.3 VAR model

For the VAR model in (3), Lütkepohl (2005) is referenced. $G_t \overset{def}{=} (g_{1t}, ..., g_{Kt})^\top$ is a $(K \times 1)$ random vector , $\alpha \overset{def}{=} (\alpha_1, \cdots, \alpha_K)^\top$, $\mathbf{u}_t \overset{def}{=} (u_{1t}, \cdots, u_{Kt})^\top$, and the coefficient matrix

$$B_i \overset{def}{=} \begin{pmatrix} \beta_{11,i} & \cdots & \beta_{1K,i} \\ \vdots & \ddots & \vdots \\ \beta_{K1,i} & \cdots & \beta_{KK,i} \end{pmatrix} \tag{4}$$

The moving average (MA) representation of VAR model can be used to conduct the impulse response analyisis, which could be derive as follows:

$$\begin{aligned} G_t &= \alpha + B_1 G_{t-1} + B_2 G_{t-2} + \cdots + B_q G_{t-q} + \mathbf{u}_t \\ &= \alpha + (B_1 L + \ldots + B_q L^q) G_t + \mathbf{u}_t \end{aligned}$$

$$B(L) G_t = \alpha + \mathbf{u}_t \tag{5}$$

with $B(L) = I_K - B_1 - \ldots - B_q L^q$. Let $\Phi(L) = \sum_{i=0}^{\infty} \Phi_i L^i$, s.t. $\Phi(L)B(L) = I_K$. Premultiplying (5) by $\Phi(L)$:

$$\begin{aligned} G_t &= \Phi(L)\alpha + \Phi(L)\mathbf{u}_t \\ &= \Big(\sum_{i=0}^{\infty} \Phi_i\Big)\alpha + \sum_{i=0}^{\infty} \Phi_i \mathbf{u}_{t-i} \\ &= \mu + \sum_{i=0}^{\infty} \Phi_i \mathbf{u}_{t-i} \end{aligned}$$

where $\Phi_0 = I_K$ Then impulse response function is:

$$\begin{aligned} G_{t+n} &= \sum_{i=0}^{\infty} \Phi_i \mathbf{u}_{t+n-i} \\ (\Phi_n)_{jk} &= \frac{\partial g_{t+n}}{\partial u_{kt}} \end{aligned}$$

where the element $(\Phi_n)_{jk}$ represents the effect of unit shock from variable k to variable j at period n. Note that the mean term μ is dropped here because it is of no interest in the impulse response analysis.

Until here the variance covariance of $\mathbf{u}_t$ might still be correlated. To obtain the orthogonal innovation, the Choleskey variance decomposition is applied, the positive definite symmetric matrix Σ_u can be decomposed as $\Sigma_u = PP^\top$, where P is a lower triangular nonsingular matrix with positive diagonal elements:

$$\begin{aligned} G_t &= \mu + \sum_{i=0}^{\infty} \Phi_i \mathbf{u}_{t-i} \\ &= \mu + \sum_{i=0}^{\infty} \Phi_i P P^{-1} \mathbf{u}_{t-i} \\ &= \mu + \sum_{i=0}^{\infty} \Theta_i \omega_{t-i} \end{aligned}$$

where $\Theta_i = \Phi_i P$, and $\omega_t = P^{-1}\mathbf{u}_t$, note that $\omega_t = (\omega_{1t}, \dots, \omega_{Kt})$ are uncorrelated and have unit variance $\Sigma_\omega = P^{-1}\Sigma_u(P^{-1})^\top = I_K$.

The MA representation of H step predictor $G_t(H)$ is as follows:

$$G_t(H) = \mu + \sum_{i=H}^{\infty} \Phi_i \mathbf{u}_{t+H-i} = \mu + \sum_{i=0}^{\infty} \Phi_{i+H} \mathbf{u}_{t-i}$$

The components from Forecast Error Variance Decomposition are applied to construct the network:

$$\begin{aligned} C_{jk,H} &= \sum_{i=0}^{H-1} (e_j^\top \Theta_i e_k)^2 / MSE_j(H) \\ &= \sum_{i=0}^{H-1} (e_j^\top \Theta_i e_k)^2 / \sum_{i=1}^{H-1} e_j^\top \Phi_i \Sigma_u \Phi_i^\top e_j \end{aligned}$$

where e_k is the ith element of I_K. $C_{jk,H}$ represents the amount of forecast error variance of variable j accounted for by exogenous shocks to variable k. In application, if H is fixed, I write C_{jk} for simplicity, and $\widehat{C}_{jk}$ represents the estimator of it. More details can be found in Lütkepohl (2005).

7.4 ACF plots

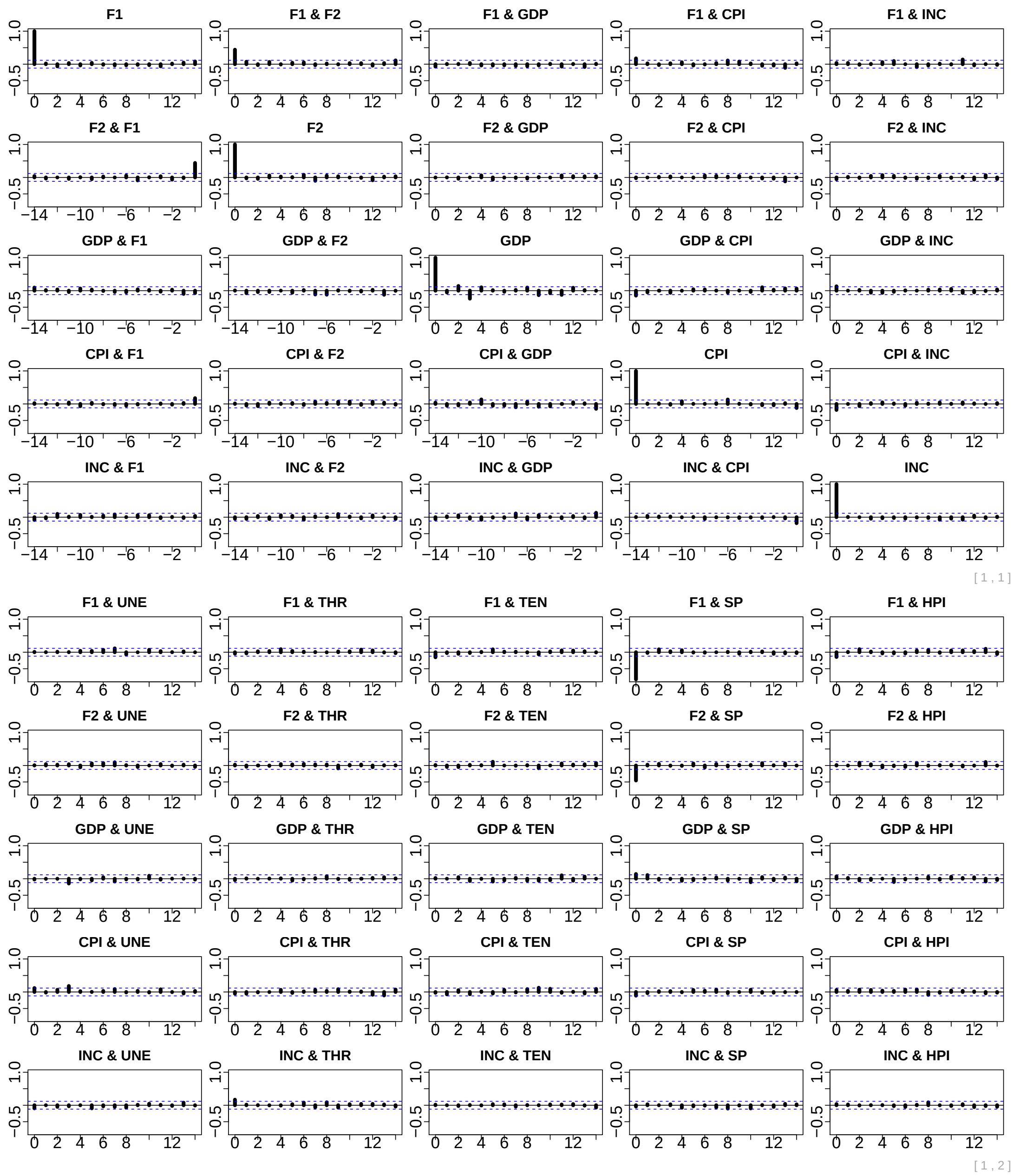

Figure 30: ACF of the constructed VAR(15) model, part 1.

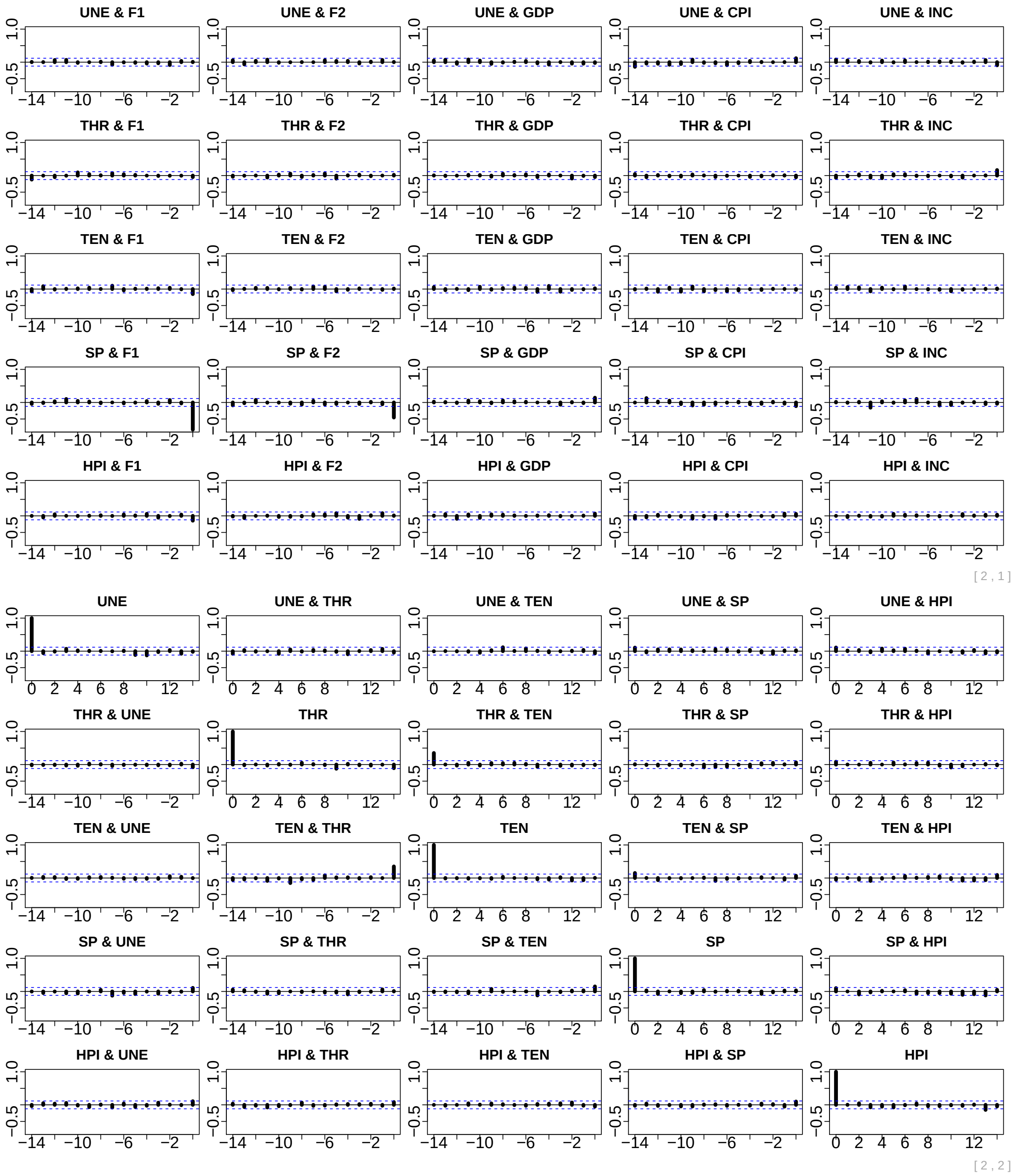

Figure 31: ACF of the constructed VAR(15) model, part 2.

7.5 Financial institutions

	Depositories (25)		Insurances (25)
WFC	Wells Fargo & Company	AIG	American International Group, Inc.
JPM	J P Morgan Chase & Co	TRV	The Travelers Companies, Inc.
BOA	Bank of America Corporation	MMC	Marsh & McLennan Companies, Inc.
C	Citigroup Inc.	AFL	Aflac Incorporated
USB	U.S. Bancorp	AON	Aon plc
BK	Bank Of New York Mellon Corporation	L	Loews Corporation
STT	State Street Corporation	PGR	Progressive Corporation (The)
BBT	BB&T Corporation	LNC	Lincoln National Corporation
STI	SunTrust Banks, Inc.	CNA	CNA Financial Corporation
FITB	Fifth Third Bancorp	MKL	Markel Corporation
MTB	M&T Bank Corporation	UNM	Unum Group
NTRS	Northern Trust Corporation	CINF	Cincinnati Financial Corporation
RF	Regions Financial Corporation	AJG	Arthur J. Gallagher & Co.
KEY	KeyCorp	Y	Alleghany Corporation
CMA	Comerica Incorporated	TMK	Torchmark
ZION	Zions Bancorporation	WRB	W.R.Berkley Corporation
SIVB	SVB Financial Group	AFG	American.FINL.GP.OHIO
CFR	Cullen/Frost Bankers, Inc.	BRO	Brown & Brown, Inc.
PBCT	People's United Financial, Inc.	WTM	White Mountain IN.GP.
CBSH	Commerce Bancshares, Inc.	ORI	Old Republic INTL.
SNV	Synovus Financial	ANAT	American National Insurance Company
BPOP	Popular	MCY	Mercury General
FHN	First Horizon National	RLI	RLI Corp.
WBS	Webster Financial	MBI	MBIA Inc.
TCB	TCF Financial	KMPR	Kemper Corporation

	Broker-Dealers (25)		**Hedge Funds (25)**
AXP	American Express Company	DUNN	Dunn Capital World Monetary & Agriculture Program
SCHW	The Charles Schwab Corporation	RED	Red Oak Fundamental Trading Program
BEN	Franklin Resources, Inc.	HGDP	Hawksbill Global Diversified Program
MHFI	McGraw Hill Financial, Inc.	ELP	Excelsior LP
TROW	T. Rowe Price Group, Inc.	SGO	Salem Global Opportunity LP
EFX	Equifax Inc.	CMF	Campbell Managed Futures
RJF	Raymond James Financial, Inc.	ATC	Abraham Trading Composite
SEIC	SEI Investments Company	LF	Legacy Futures LP
LM	Legg Mason, Inc.	SP	Simplon Partners LP
EV	Eaton Vance Corp.	SDP	Saxon Diversified Program
AB	AllianceBernstein Holding L.P.	FPLS	Falcon Point Long/Short
ITG	Investment Technology Group, Inc.	APF	Ardsley Partners Fund II LP
TRC	Tejon Ranch Co.	SCFV	SC Fundamental Value LP
SFE	Safeguard Scientifics, Inc.	CT	Campbell Trust
ATAX	America First Multifamily Investors, L.P.	KSCP	KS Capital Partners LP
OPY	Oppenheimer Holdings, Inc.	PG	Peconic Grenadier LP
CTO	Consolidated-Tomoka Land Co.	PP	PAW Partners LP
REXI	Resource America, Inc.	LMA	Longfellow Merger Arbitrage
SNFCA	Security National Financial Corporation	MRDP	Millburn Ridgefield Diversified Program
SIEB	Siebert Financial Corp.	CDP	Chesapeake Diversified Program
GROW	U.S. Global Investors, Inc.	EGMP	Eclipse Global Monetary Program
CSWC	Capital Southwest Corporation	RDP	Rabar Diversified Program
FNMA	Federal National Mortgage Association	MLP	Meteoric LP
FCE-A	Forest City Realty Trust, Inc	NP	Nestor Partners
FDMC	Federal Home Loan Mortgage Corporation	GPP	Gabelli Performance Partnership LP

Table 17: Financial companies with tickers classified by industry.

	Firms	RS1	RS2
1	WFC	0.53	0.20
2	JPM	0.43	0.16
3	BOA	0.60	0.13
4	C	0.69	0.07
5	USB	0.38	0.19
6	BK	0.43	0.05
7	STT	0.49	0.05
8	BBT	0.46	0.24
9	STI	0.48	0.18
10	FITB	0.49	0.08
11	MTB	0.40	0.25
12	NTRS	0.34	0.04
13	RF	0.41	0.09
14	KEY	0.37	0.11
15	CMA	0.52	0.20
16	ZION	0.38	0.16
17	SIVB	0.22	0.03
18	CFR	0.32	0.16
19	PBCT	0.11	0.08
20	CBSH	0.10	0.06
21	SNV	0.37	0.18
22	BPOP	0.41	0.08
23	FHN	0.33	0.10
24	WBS	0.45	0.20
25	TCB	0.30	0.20
26	AIG	0.42	0.18
27	TRV	0.28	0.04
28	MMC	0.32	0.06
29	AFL	0.43	0.07
30	AON	0.19	0.04
31	L	0.33	0.01
32	PGR	0.29	0.04
33	LNC	0.49	0.03
34	CNA	0.54	0.04
35	MKL	0.20	0.02
36	UNM	0.42	0.06
37	CINF	0.26	0.05
38	AJG	0.19	0.01
39	Y	0.17	0.06
40	TMK	0.51	0.07
41	WRB	0.15	0.03
42	AFG	0.38	0.11
43	BRO	0.16	0.05
44	WTM	0.20	0.02
45	ORI	0.30	0.05
46	ANAT	0.37	0.00
47	MCY	0.24	0.07
48	RLI	0.17	0.07
49	MBI	0.41	0.03
50	KMPR	0.45	0.01
51	AXP	0.51	0.03
52	SCHW	0.32	0.06

References

Bai, J. and Ng, S. (2006). Evaluating latent and observed factors in macroeconomics and finance. *Journal of Econometrics*, 131:507–537.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. *The Quarterly Journal of Economics*, 120(1):387–422.

Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, 104(3):535–559.

Cattell, R. B. (1952). *Factor analysis: an introduction and manual for the psychologist and social scientist.* New York: Harper.

Chan, N., Getmansky, M., Haas, S. M., and Lo, A. W. (2005). Systemic risk and hedge funds. Working Paper 11200, National Bureau of Economic Research.

Chao, S.-K., Härdle, W. K., and Yuan, M. (2015). Factorisable Sparse Tail Event Curves. *ArXiv e-prints*.

Chen, L., Dolado, J. J., and Gonzalo, J. (2005). Quantile factor models. *Working Paper*.

De Rossi, G. and Harvey, A. (2009). Quantiles, expectiles and splines. *Journal of Econometrics*, 152(2):179–185.

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

Duffie, D. and Pan, J. (1997). An overview of value at risk. *The Journal of Derivatives*, 4(3):7–49.

Favero, C. A. and Marcellino, M. (2005). Large datasets, small models and monetary policy in europe. *CLM Economia*, pages 249–269.

Fruchter, B. (1954). *Introduction to Factor Analysis.* Van Nostrand.

Gschöpf, P., Härdle, W. K., and Mihoci, A. (2015). TERES: Tail Event Risk Expectile based Shortfall. SFB 649 Discussion Paper 2015-047, Sonderforschungsbereich 649, Humboldt Universität zu Berlin, Germany.

Härdle, W., Wang, W., and Yu, L. (2016). TENET: Tail-Event driven NETwork risk. *Journal of Econometrics*, 192(2):499–513.

Härdle, W. K., Huang, C., and Chao, S.-K. (2016). Factorisable Sparse Tail Event Curves with Expectiles. SFB 649 Discussion Paper 2016-018, Sonderforschungsbereich 649, Humboldt Universität zu Berlin, Germany.

Hautsch, N., Schaumburg, J., and Schienle, M. (2015). Financial network systemic risk contributions. *Review of Finance*, 19(2):685–738.

Koenker, R. and Bassett, G. W. (1978). Regression quantiles. *Econometrica*, 46:33–50.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer.

Markus, P. (2015). Understanding systematic risk: A high-frequency approach.

McDonald, R. P. (1985). *Factor Analysis and Related Methods*. Hillsdale, NJ: Erlbaum.

Moench, E. (2008). Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented var approach. *Journal of Econometrics*, 146(1):26–43.

Newey, W. K. and Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica*, 55(4):819–847.

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

Schaumburg, J. (2012). Predicting extreme value at risk: Nonparametric quantile regression with refinements from extreme value theory. *Computational Statistics & Data Analysis*, 56(12):4081 – 4096.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1):1–48.

Stock, J. H. and Watson, M. W. (2005). Implications of Dynamic Factor Models for VAR Analysis. NBER Working Papers 11467, National Bureau of Economic Research, Inc.

Tran, N. M., Burdejová, P., Ospienko, M., and Härdle, W. K. (2019). Principal component analysis in an asymmetric norm. *Journal of Multivariate Analysis*, 171:1–21.

Xiao, Z., Guo, H., and Lam, M. S. (2015). *Quantile Regression and Value at Risk*, pages 1143–1167. Springer New York, New York, NY.

IRTG 1792 Discussion Paper Series 2020

For a complete list of Discussion Papers published, please visit http://irtg1792.hu-berlin.de.

001 "Estimation and Determinants of Chinese Banks' Total Factor Efficiency: A New Vision Based on Unbalanced Development of Chinese Banks and Their Overall Risk" by Shiyi Chen, Wolfgang K. Härdle, Li Wang, January 2020.

002 "Service Data Analytics and Business Intelligence" by Desheng Dang Wu, Wolfgang Karl Härdle, January 2020.

003 "Structured climate financing: valuation of CDOs on inhomogeneous asset pools" by Natalie Packham, February 2020.

004 "Factorisable Multitask Quantile Regression" by Shih-Kang Chao, Wolfgang K. Härdle, Ming Yuan, February 2020.

005 "Targeting Cutsomers Under Response-Dependent Costs" by Johannes Haupt, Stefan Lessmann, March 2020.

006 "Forex exchange rate forecasting using deep recurrent neural networks" by Alexander Jakob Dautel, Wolfgang Karl Härdle, Stefan Lessmann, Hsin-Vonn Seow, March 2020.

007 "Deep Learning application for fraud detection in financial statements" by Patricia Craja, Alisa Kim, Stefan Lessmann, May 2020.

008 "Simultaneous Inference of the Partially Linear Model with a Multivariate Unknown Function" by Kun Ho Kim, Shih-Kang Chao, Wolfgang K. Härdle, May 2020.

009 "CRIX an Index for cryptocurrencies" by Simon Trimborn, Wolfgang Karl Härdle, May 2020.

010 "Kernel Estimation: the Equivalent Spline Smoothing Method" by Wolfgang K. Härdle, Michael Nussbaum, May 2020.

011 "The Effect of Control Measures on COVID-19 Transmission and Work Resumption: International Evidence" by Lina Meng, Yinggang Zhou, Ruige Zhang, Zhen Ye, Senmao Xia, Giovanni Cerulli, Carter Casady, Wolfgang K. Härdle, May 2020.

012 "On Cointegration and Cryptocurrency Dynamics" by Georg Keilbar, Yanfen Zhang, May 2020.

013 "A Machine Learning Based Regulatory Risk Index for Cryptocurrencies" by Xinwen Ni, Wolfgang Karl Härdle, Taojun Xie, August 2020.

014 "Cross-Fitting and Averaging for Machine Learning Estimation of Heterogeneous Treatment Effects" by Daniel Jacob, August 2020.

015 "Tail-risk protection: Machine Learning meets modern Econometrics" by Bruno Spilak, Wolfgang Karl Härdle, October 2020.

016 "A data-driven P-spline smoother and the P-Spline-GARCH models" by Yuanhua Feng, Wolfgang Karl Härdle, October 2020.

017 "Using generalized estimating equations to estimate nonlinear models with spatial data" by Cuicui Lu, Weining Wang, Jeffrey M. Wooldridge, October 2020.

018 "A supreme test for periodic explosive GARCH" by Stefan Richter, Weining Wang, Wei Biao Wu, October 2020.

019 "Inference of breakpoints in high-dimensional time series" by Likai Chen, Weining Wang, Wei Biao Wu, October 2020.

IRTG 1792, Spandauer Strasse 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.

IRTG 1792 Discussion Paper Series 2020

For a complete list of Discussion Papers published, please visit http://irtg1792.hu-berlin.de.

020 "Long- and Short-Run Components of Factor Betas: Implications for Stock Pricing" by Hossein Asgharian, Charlotte Christiansen, Ai Jun Hou, Weining Wang, October 2020.

021 "Improved Estimation of Dynamic Models of Conditional Means and Variances" by Weining Wang, Jeffrey M. Wooldridge, Mengshan Xu, October 2020.

022 "Tail Event Driven Factor Augmented Dynamic Model" by Weining Wang, Lining Yu, Bingling Wang, October 2020.

IRTG 1792, Spandauer Strasse 1, D-10178 Berlin
http://irtg1792.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the IRTG 1792.