Distribution of probabilities of Financial Risk Meter (FRM)

Master's Thesis submitted to

Prof. Dr. Wolfgang Karl Härdle

Dr. Rui Ren

School of Business and Economics

Ladislaus von Bortkiewicz Chair of Statistics

Humboldt-Universität zu Berlin

by

Ranqing Song

599676

in partial fulfillment of the requirements

for the degree of

Master of Science in Statistics

Frankfurt am Main, March 26, 2021

Acknowledgement

I want to express my sincere gratitude to my supervisors Prof. Dr. Wolfgang Karl Härdle and Dr. Rui Ren, for their inspiration and support during this thesis. Thank you for the advice and support, Michael Althof and Anna Shchekina. I would also like to thank all members of International Research Training Group 1792 “High Dimensional Nonstationary Time Series” for their help.

Abstract

This paper studies a systemic risk indicator, Financial Risk Meter (FRM), which is calculated based on quantile Lasso regression. The standard FRM index is the average of daily penalization parameters for all selected financial institutions. This paper extends the standard FRM to numerous novel FRM candidates that could capture systemic risk and predict the upcoming recession. FRM candidates are defined by using quantiles of penalization parameters derived from the distribution of financial institutions' returns. The co-movement of FRM candidates and commonly used systemic risk measures are checked with the correlation coefficient, the Kolmogorov-Smirnov test statistic and the Granger causality test. Furthermore, FRM candidates are able to predict the probability of economic recessions by applying binary regression models. Empirical experiments are implemented during two periods, namely the financial crisis of 2007 and the COVID-19 pandemic, in two major financial markets, the Americas and Europe stock markets. The results prove that FRM candidates are suitable systemic risk measures and recession predictors, since they can capture the increase of overall distress and market downturn, move similarly and even better than popular systemic risk measures for both Americas and Europe stock markets. Additionally, the recession probabilities estimated from FRM candidates are close to the actual recession indicators. In conclusion, FRM candidates can be regarded as systemic risk indicators in terms of feasibility and robustness.

Keywords: systemic risk, Financial Risk Meter (FRM), Quantile Regression, Lasso regularization, penalization parameters, recession prediction, financial crisis of 2007, COVID-19 pandemic.

Contents

List of Abbreviations

FRM	Financial Risk Meters
VIX	CBOE Volatility Index
SRISK	Systemic Risk Index
VSTOXX	Euro Stoxx 50 Volatility Index
GT	Google Trends of key word ”financial Risk
FT	Financial turbulence index
CISS	Composite Indicator of Systemic Stress
VaR	Value at Risk
SIM	Single Index Model
GACV	Genalized approximate cross-validation
BIC	Bayesian Information Criterion
AIC	Akaike Information Criterion
FPE	Prediction Error Criterion
HQ	Hannan-Quinn information criterion
SC	Schwarz criterion
PT	Portmanteau test
BG	Breusch-Godfrey LM test
ES	Edgerton-Shukur F test
NBER	National Bureau of Economic Research
CEPR	Center for Economic and Policy Research

List of Figures

List of Tables

1 Introduction

After the financial crisis of 2007, market participants have paid much attention to systemic risk. At the same time, governments and international organisation are calling for increased regulation of systemic risk. These works intended to shed light on understanding the systemic risk, how dangerous the systemic risk is in terms of the whole financial system, and how to measure systemic risk to eliminate the negative influences of systemic risk so that reasonable precautions may be taken before suffering catastrophic losses?

There are various definitions and explanations about systemic risk. Schwarcz (2008) defined systemic risk as "the risk that an economic shock such as market or financial institutional failure triggers either (X) the failure of a chain of markets or institutions or (Y) a chain of significant losses to financial institutions, resulting in increases in the cost of capital or decreases in its availability, often evidenced by substantial financial-market price volatility". Federal Reserve Governor Daniel Tarullo defined systemic risk as "financial institutions are systemically important if the failure of the firm to meet its obligations to creditors and customers would have significant adverse consequences for the financial system and the broader economy", which is cited by Brownlees et al. (2012). Both definitions indicate that considering the interdependence and interconnectedness of institutions in the financial system, the bankruptcies or critical financial intermediaries' failure will cause the spread of systemic distress along with capital-market linkages of institutions and even the whole economic system. For example, the financial crisis of 2007 was caused by the banking panic. As Yu et al. (2019) mentioned, after the bankruptcy of Lehman Brothers, a sequence of the international bank was bankrupted due to their interdependence of Lehman Brothers. As a result, the stability of entire financial system was threatened international, and the broad economy suffered from the great recession, see Hautsch et al. (2015) and Brownlees et al. (2012).

This paper is in the context of one systemic risk measure called Financial Risk Meter (FRM), which was proposed by Härdle et al. (2016), Zbonakova et al. (2016), Mihoci et al. (2020), Yu et al. (2017), and Yu et al. (2019). FRM is an augmented systemic risk measure that expresses the high-dimensional tail risk with a single accurate value indicator. The standard FRM is the average over series of the penalization parameters λ for all selected financial institutions. The penalization parameters in linear quantile Lasso regressions, which expresses financial institutions based on related institutions and macroeconomic factors, are subsequently estimated by the generalized approximate cross-validation criterion (GACV). The quantile in each linear quantile regression is set as 5% or 10%, which corresponds to tail risk. In this paper, despite the standard FRM based on all financial institutions, we want to extend the definition of standard FRM using quantiles of the distribution of penalization parameters. So that the newly defined FRM is related to the each financial institution's rank order in terms of penalization parameters, i.e. only a portion of financial institutions are included into defining FRM candidates. As data input for FRM calculation, the empirical penalization parameters across two time intervals with high systemic risk in the Americas and Europe are estimated. For both regions, the first period ranges from 03 April 2007 until 31 December 2009 and the second period is from 01 January 2019 until 31 December 2020.

Based on characteristics of the time series of penalization parameters, novel FRM candidates with Interquartile range (IQR) and quantiles at 50%, 60%, 70%, 80%, 90% of a series of penalization parameters for all financial are proposed. All novel FRM and the standard FRM for each market are called FRM candidates for the respective market from now on. This paper aims to compare all FRM candidates so as to select an optimal FRM that works well as a systemic risk measure and can be further applied in related fields like predicting the upcoming recession. The first procedure compares these FRM candidates with commonly used systemic risk measures based on testings for correlation, goodness-of-fit and Granger causality. Furthermore, these FRM candidates are applied to forecast the probability of observing a current recession since the high risk of systemic risks is in line with the recession periods by Hautsch et al. (2015). Each FRM candidate can capture the evolution of systemic risk, and forecast recession will be taken as a suitable systemic risk measure.

The striking result of the paper is that there is no significant difference among these FRM candidates. All FRM candidates are qualified and useful systemic risk measures for the Americas and Europe. They fluctuate similarly to other systemic risk measures and have the ability to predict the upcoming and the end of the recession. This finding proves that penalization parameters from our linear Lasso quantile regression generate a robust estimation of FRM. The properties of various FRM candidates calculated based on different subsamples of financial institutions are not differentiated significantly from the property of the standard FRM based on the whole sample. Moreover, the result of this paper implies that the standard FRM proposed by previous research is practical and meaningful because no novel FRM candidate has an overwhelming advantage over it.

The paper is structure as follows: after reviewing literature about systemic risk measures, especially the tail risk measure methodologies in chapter 2, the FRM framework is elaborated in chapter 3. The linear quantile lasso regression model and the estimation procedure of penalization parameters are presented in chapter 3.1 and 3.2. Chapter 3.3 describes the empirical application of the model based on data collected from the two periods mentioned above. According to the characteristics of penalization parameters which are interpreted in chapter 3.4, the mathematical representations and empirical distribution of novel FRM candidates are described in chapter 3.5. In the next two chapter, all comparison methods and results are presented. The comovement tests, i.e. correlation test and Kolmogorov-Smirnov test as well as Granger causality test of individual FRM candidates and individual commonly used systemic risk measures, are available in chapter 4. In chapter 5, recession prediction models and corresponding implied recession probabilities for all FRM candidates are reported. The last chapter 6 are conclusions and discussions.

All underlying codes have been written in the R software. They can be downloaded at www.quantlet.de, which is indicated in this paper with Q.

2 Literature Review

FRM is a systemic risk measure, which integrates joint tail risk comovement of critical financial institutions into one index. Before investigating how FRM works, systemic risk and common systemic risk measurements will be introduced. The methodology in the context of tail risk measures are explained successively: based on Value at Risk (VaR) in chapter 2.2, Adrian and Brunnermeier (2011) proposed a quantile regression based linear bivariate model see chapter 2.3. In chapter 2.4, Hautsch et al. (2015) extended the bivariate model into a high-dimensional linear model. Fan et al. (2018) achieved the balance between precision and dimension through applying a non-linear Single Index Model (SIM) to previous algorithms in chapter 2.5. These researches have motivated the idea of using penalization parameters as a systemic risk measure.

2.1 Systemic risk measures

As Schwarcz (2008) emphasized, systemic risk is not caused by normal market volatility. It cannot be avoided through diversification and therefore affects almost all market participants. Moreover, Schwarcz (2008) mentioned that because of the ongoing trend towards disintermediation or enabling companies to access the ultimate source of funds, systemic risk should increasingly be viewed by its impact on the whole financial markets. The cross-sectional interdependencies among financial institutions may reveal the overall risk level of the financial market. Therefore, financial institution-specific risk cannot be appropriately evaluated without considering potential risk spillover effect from other institutions by Hautsch et al. (2015). The systemic risk measure that aimed at analyzing, monitoring and controlling system risk should take the inevitable consequences followed by institutions being under distress or bankruptcies into account.

For major financial markets, there are already several widely used systemic risk measures. For example, Zbonakova et al. (2016), Yu et al. (2017) and Yu et al. (2019) indicated that CBOE Volatility Index (VIX), Systemic Risk Index (SRISK), financial turbulence index (FT), and Google Trends data (GT) are common systemic risk measures in Americas, and Mihoci et al. (2020) mentioned that Composite Indicator of Systemic Stress (CISS) and Euro Stoxx 50 Volatility Index (VSTOXX) are most widely used in European counties.

VIX, the expected volatility measure based on S&P500 index over the next 30 days, is published by the Chicago Board Options Exchange (CBOE). Corrado and Miller (2005) proved that VIX performed well to forecast the volatility in the American stock market. The VIX index is well-known as a "fear indicator" for stock markets according to Brownlees et al. (2012). Kritzman and Li (2010) introduced a mathematical measure of financial turbulence using Mahaland's distance. Kritzman and Li (2010) defined financial turbulence as a condition in which asset prices, conditioning on their historical patterns of behaviour, behave in an uncharacteristic fashion such as extreme price moves, decoupling of correlated asset and convergence of uncorrelated assets. The turbulence for a particular time period t will be high if the asset return either move further from the average of historical joint returns of assets or not follow the historical correlation structure. Corresponding mathematical representation is available in 4.1. SRISK proposed by Brownlees et al. (2012), Brownlees and Engle (2017) is to measure expected capital shortage across all financial institutions over a given time horizon. The capital shortfall of a firm depends on its degree of leverage, and the expected equity loss resulted from a crisis. Preis et al. (2013) argued that query data from Google Trends of the keyword related to finance reflected human interaction with the Internet in response to changes in the financial markets. It may offer a perspective on the behaviour of market participants in periods of large market movement like a financial crisis. In this paper, we apply the Google Trends volumes of the keyword "financial crisis" as a systemic risk measure.

In the Euro area, Hollo et al. (2012) introduced CISS, which is used to measure the current state of distress and instability in the financial system and to condense that state of instability into a single statistic. The CISS not only captures the real-time stress level in the entire financial system but also can be compared and studied empirically in the context of early warning signal models. VSTOXX measures implied variance across all options over a given time horizon. The option contracts on Euro Stoxx 50 are products of Eurex among the Exchange with the highest trading volumes, see Qontigo (2020).

2.2 Value at Risk (VaR)

As Adrian and Brunnermeier (2011) indicated, systemic risk measures capture the potency for spreading financial distress or risk across the institution by gauging this increase in tail comovement. The value at risk (VaR) has established its role in the tail risk measure field. It measures the level of financial risk within one institution by estimating its profits and losses in future periods. Under this condition, the definition of VaR by Jorion (2007) was reinterpreted by Adrian and Brunnermeier (2011) and Härdle et al. (2017), i.e. for the cutoff return of the institution j at target time period t such that there is a pre-specified probability q that the actual log return will be smaller. This definition could be mathematically represented as

$$\mathrm{P}(X_{j,t} \leq VaR^q_{j,t}) \stackrel{\text{def}}{=} q, \tag{2.1}$$

where $X_{j,t}$ is the log return of institution j at time t for which the $VaR^q_{j,t}$ is defined with $q \in (0,1)$. $VaR^q_{j,t}$ could be calculated with rectangular moving average (RMA) and Delta-Normal model see Härdle et al. (2017). If taking quantile q as 99%, $VaR^q_{j,t}$ is the minimal log return of institution j at time t that the probability of having a lower expected log return is 99 percent. Several methodologies regarding tail risk measure elaborated in next chapters are based on VaR.

2.3 Bivariate CoVaR

The risk that the stability of the whole financial system is threatened does not necessarily be reflected by the individual institution risk measure VaR. Hence, Adrian and Brunnermeier (2011) and Tobias and Brunnermeier (2016) proposed a systemic risk measure $CoVaR$, i.e. adding to existing tail risk measure the prefix "Co", which stands for the conditional comovement. $CoVaR$ is defined as the VaR of one institution j conditional on one particular institution i being in distress:

$$\mathbb{P}\{X_j \leq CoVaR^q_{j|i} | X_i = VaR^q\,(X_i)\} \stackrel{def}{=} q. \tag{2.2}$$

Once in the case where $j = system$, $CoVaR$ denotes the return of all financial institutions in one system conditional on the institution i being at its VaR level, i.e. $CoVaR_{system|i}$ reflect how the whole system was affected by the failure of the institution i. On the other hand, $CoVaR_{j|system}$ characterizes which institution is suffered from the highest risk when a financial crisis occurs. The difference between $CoVaR$ conditional on the distress of an institution and $CoVaR$ conditional on the normal state of the institution is called $\Delta CoVaR$:

$$\Delta CoVaR^q_{j|i} = CoVaR^q_{j|X^i=VaR^q_i} - CoVaR^q_{j|X^i=Var^{50}_i}. \tag{2.3}$$

The advantage of $\Delta CoVaR$ is that for two institutions that are equally risky according to VaR, the institution with the higher $\Delta CoVaR$ contributes more to the systemic risk than another institution. By Adrian and Brunnermeier (2011), the difference between the VaR of the financial system conditional on the distress of a specific institution i and the VaR of the financial system conditional on the media state of the institution is captured by $\Delta CoVaR^q_{system|i}$. Moreover, $\Delta CoVaR^q_{j|system}$, which is called "exposure $CoVaR$", measures institution j's increase in VaR in the case of under a financial crisis, i.e. the extent to which the systemic financial event influences an institution. In summary, taking i and j as different institutions and financial system makes $CoVaR$ and $\Delta CoVaR$ generally enough to measure the risk spillover from institution to institution across the whole financial network.

To capture time-varying $CoVaR_t$ and VaR_t, a vector of lagged macroeconomic variable M_{t-1} is considered as conditioning variables:

$$\mathbb{P}\{X_{j,t} \leq CoVaR^q_{j|i,t} | X_{i,t} = VaR^q\left(X_{i,t}\right), M_{t-1}\} \overset{def}{=} q. \tag{2.4}$$

$X_{j,t}$ is the growth rate of market-valued assets based on balance sheet data and market equity data, including leverage, size, and market-to-book. Under this setting, we use quantile regression to estimate $CoVaR$. Quantile regression introduced by Koenker and Bassett Jr (1978) aims to estimate the conditional quantile function, in which quantile of the conditional distribution of the response variables are expressed as functions of observed covariates, e.g. equation 2.4. In the financial risk world, it was observed that error distribution with longer tails than Gaussian distribution was common. Koenker and Bassett Jr (1978) mentioned that in these cases, weights on extreme observations should be put to modify the sample mean, i.e. giving different weights to positive and negative residuals. Quantiles yield minimizing the sum of asymmetrically weighted absolute residuals. Analogue to linear regression, the estimates of conditional quantile regression are obtained by minimizing the sum of residuals, which herein are asymmetrically weighted with the loss function. Because of this advantage, quantile regression is "gradually developing into a comprehensive strategy for completing the regression picture" according to Koenker and Hallock (2001). Following Tobias and Brunnermeier (2016), we focus on the $\Delta CoVaR^q_{system|i,t}$ from now on. Firstly, we estimate $X_{i,t}$ and $X_{system,t}$ based on following quantile regressions:

$$X_{i,t} = \alpha_i^q + \gamma_i^q M_{t-1} + \varepsilon_{i,t}^q, \tag{2.5a}$$

$$X_{system|i,t} = \alpha_{system|i}^q + \gamma_{system|i}^q M_{t-1} + \beta_{system|i}^q X_{i,t} + \varepsilon_{system|i,t}^q. \tag{2.5b}$$

Then the estimated values are further applied if $F_{\varepsilon_{i,t}}^{-1}(q|M_{t-1}) = 0$ and $F_{\varepsilon_{i,t}}^{-1}(q|M_{t-1}, X_{i,t}) = 0$ are assumed:

$$VaR_{i,t}^q = \widehat{\alpha}_i^q + \widehat{\gamma}_i^q M_{t-1}, \tag{2.6a}$$

$$CoVaR_{system|i,t}^q = \widehat{\alpha}_{system|i} + \widehat{\gamma}_{system|i}^q M_{t-1} + \widehat{\beta}_{system|i}^q VaR_{i,t}^q. \tag{2.6b}$$

Finally, $\Delta CoVaR_{system|i,t}^q$ for each institution is computed as:

$$\Delta CoVaR_i^q = CoVaR_{i,t}^q - CoVaR_{50,t}^q = \widehat{\beta}_{system|i}^q \left(VaR_{i,t}^q - VaR_{50,t}^q \right), \tag{2.7}$$

where $\beta_{system|i}$ reflects the degree of interconnectedness between institution i and the whole financial system, i.e. how the institution influences the rest of the financial system by Härdle et al. (2016).

2.4 High-dimensional CoVaR

Hautsch et al. (2015) argued that $CoVaR$ has several drawbacks. Based on the equation 2.6b, $CoVaR$ varies over time only through the channel of individual $VaRs$. Due to the multicollinearity, VaR is not modelled in terms of firm-specific variables. Thus, the variation in underlying macroeconomic factors reflect changes in firm's systemic relevance. Under this background, Hautsch et al. (2015) refined the $CoVaR$ algorithm by adding a set of lagged company-specific characteristics $C_{i,t-1}$ to the set of lagged macroeconomic factors M_{t-1}. The combined set of tail risk drivers, which is denoted as W_{t-1}, replaces M_{t-1} in the two-stage linear quantile regressions 2.5a and 2.5b as well as equations 2.6a and 2.6b.

However, shrinking the high-dimensional set of possible cross-linkages between all financial institutions to a feasible number of relevant risk connections is the major challenge faced by Hautsch et al. (2015). Osborne et al. (2000) mentioned that high-dimensional data with highly correlated covariates tends to have an extremely high variance of least-square coefficient estimates. The normally calculated estimator has poor forecast ability due to unstable combination weights see Bayer (2018), since it is challenging to balance the model flexibility and statistical precision. As Hautsch et al. (2015) argued, appropriate model selection techniques are not straightforward because the test of individual significance of single variables does not account for the collinearity between the covariates. On the other hand, a sequence of significance tests has too many possible variations to be checked. Hence, the penalization parameters like Lasso, ridge and elastic are generally applied to quantile regression estimator for regularization purpose, i.e. eliminating the multicollinearity among predictors. The L_1- norm penalization regression, known as the least absolute shrinkage and selection operation (Lasso) by Tibshirani (1996) could not only shrink the estimated coefficients towards zero but also cause several of them to be precisely zero when making λ sufficient large. By shrinking or setting some coefficients to zero, the variance is to be reduced more than the increase of bias. In terms of the mean squared error, the overall prediction accuracy and interpretation ability of the whole predictors are improved straightforwardly by Tibshirani (1996). Although the exact calculation of the bias-variance trade-off is still unexplored for the asymmetric loss function in quantile regression cases, intuitively, it should be reasonable that a decrease of the variance of the estimated error decreases the expected loss of the estimation error, see Bayer (2018). Since the subset selection property is not shared with other types of penalization like L_2-norm penalty and the parsimonious model puts more light on the relationship between the response and covariates, the L_1-norm penalty may perform better in where there are many noise variables, particularly with high dimensional data $p \geq n$. In other words, by Tibshirani (1996), the form of Lasso regression is standard for a high-dimensional conditional mean regression problem. Therefore, tail risk drivers are selected in a data-driven way by adopting Lasso to quantile regression, following Belloni et al. (2011). $\varepsilon_{i,t}^q$ is estimated in a quantile Lasso regression with a fixed individual penalization parameter λ, which is constant and determined in a completing data-driven way for each financial institution. If any component of Lasso-selected relevant drivers W_i^q is related to another institution j, then the estimated coefficient of this component marks the impact of institution j on institutions i.

2.5 Single-Index-based CoVaR

Fan et al. (2013) and Fan et al. (2018) proposed a nonlinear Single Index Model (SIM) combining with CoVaR by Adrian and Brunnermeier (2011), Tobias and Brunnermeier (2016) and Hautsch et al. (2015) to investigate possible non-linearities in tail interconnectedness based on several facts: Firstly, nonlinearity may occur by employing the methodologies introduced by Adrian and Brunnermeier (2011) and Hautsch et al. (2015) because of the complexity of financial system, see Härdle et al. (2016). Secondly, the selected factors are difficult to be interpreted, and need to be summarized to be an index according to Fan et al. (2018). SIM performs efficiently with variable selection in the case of high-dimensional covariates since the index yields interpretability and low dimension simultaneously, i.e. SIM is a data driven technique that combines dimension reduction, variable selection and generalized tail event by Fan et al. (2018). The model motivated by Adrian and Brunnermeier (2011) and Hautsch et al. (2015) are defined based on SIM quantile variable selection technique:

$$X_{i,t} = g\left(R^\top \beta^q_{i|R_{i,t}}\right) + \varepsilon_{i,t}, \tag{2.8}$$

where $g(\cdot)$ is an unknown smooth link function, $\widehat{R}_{i,t} \stackrel{\text{def}}{=} [W_{i,t-1}, X_{-i,t}]$. $X_{-i,t}$ is a vector of log returns for all institutions except the institution i. If taking $F^{-1}_{\varepsilon_i|\widehat{R}}(p|R) = 0$, then analogue to Equation 2.6b

$$\widehat{CoVaR}^q_{i|\widehat{R}_{i,t}} = \widehat{g}\left(\widehat{R}_{i,t}\widehat{\beta}_{i|\widehat{R}_{i,t}}\right), \tag{2.9}$$

where $\widehat{R}_{i,t} \stackrel{\text{def}}{=} \left[W_{t-1}, \widehat{VaR}^q_{-i,t}\right]$. $\widehat{VaR}^q_{i,t}$ is estimated $VaRs$ from equation 2.6a for all institution except i-th institution. In the equation 2.9, not only the influence of financial institutions except for i are included, but also non-linearity reflected in the shape of link function $g(\cdot)$ is incorporated. The minimal average contrast approach (MACE) with Lasso penalization is adopted to estimated the shape of smooth link function $g(\cdot)$ and β, more details see Fan et al. (2018).

Fan et al. (2018) applied $CaViaR$ test for backtesting the estimations of VaR, CoVaR and SIM-based CoVaR. This comparison is conducted based on the data of 200 financial institutions collected from 2006 until 2015. SIM-based $CoVaR$ performed better than $CoVaR$ by Adrian and Brunnermeier (2011) and VaR in overall period from 2006 to 2015. Both SIM-based $CoVaR$ and VaR performed well during the crisis period.

While estimating SIM-based $CoVaR$, it was found that the series of time-varying penalization parameter has a striking pattern: higher values correspond to financial crisis period, and lower values tend to be generated during stable periods. This finding has led to the idea to use the Lasso penalization parameter λ as a systemic risk measure by Yu et al. (2019). Above mentioned methodologies by Hautsch et al. (2015) and Fan et al. (2018) provide one series of penalization parameters for individual financial institution. Yu et al. (2019), Yu et al. (2017) and Mihoci et al. (2020) would like to generalize the penalization parameters for all financial institutions to reflect the whole financial system, i.e. the overall behaviour of λ then works as a systemic risk measure. Härdle et al. (2016) concluded that both linear quantile regression and SIM are valid in terms of backtesting. The analysis of the entire financial system is based on empirical data for a enormous number of financial institutions within a long time interval.

Although SIM method performs better for a single institution, generating λ series during 100 institutions for more than 300 trading days with SIM is already not realistic. In this paper, linear quantile Lasso regression is applied to construct FRM framework considering its overwhelming advantage of being time-saving. More details about FRM methodology will be elaborated in the next chapter 3.

3 FRM Framework

As mentioned in chapter 2, the penalization parameters estimated in linear quantile Lasso regression is expected to capture the overall behaviour of a financial system. The linear quantile Lasso regression, which is proposed based on the methodologies in chapter 2 will be explained in detail in chapter 3.1. Approaches to estimating and optimizing penalization parameters will also be discussed in chapter 3.2. In this paper, we would like to define novel FRM candidates based on the distribution of estimated penalization parameters. The distribution of daily series of penalization parameters is available in chapter 3.3. The probabilities of above-mentioned distribution are used to define novel FRM candidates see chapter 3.5.

3.1 Linear Quantile Lasso Regression

Following Härdle et al. (2016), Yu et al. (2019), Mihoci et al. (2020), FRM is proposed in the context of methodologies elaborated in chapter 2. Let the stock price of a financial institution j at time t be P_t such that the log return of this stock at day t is calculated as $X_{j,t} = \log(\frac{P_t}{P_{t-1}})$. $X^s_{j,t}$ is a J-dimensional vector of log returns for all financial institutions on a specified trading day $t \in \{2, \ldots, T\}$ in the moving window s. s denotes the index of moving window, $s \in \{2, \ldots, (T-(n-1))\}$. T is the total number of time series observations and n is the length of window size. Then the linear quantile Lasso regression is analogue to the equation 2.5a and 2.8:

$$X^s_{j,t} = \alpha^s_{j,t} + A^{s\top}_{j,t}\beta^s_j + \varepsilon^s_{j,t} \tag{3.1}$$

with $A^s_{j,t} \stackrel{def}{=} \begin{pmatrix} X^s_{-j,t} \\ M^s_{t-1} \end{pmatrix}$. M^s_{t-1} the m-dimensional vector of macroeconomic variables includes log returns of macroeconomic factors at the corresponding trading day $t-1$ within the moving window s. $X^s_{-j,t}$ is the $(J-1)$-dimensional vector of log returns of all other institutions expect for institution j. β^s_j is a $p = (J+m-1)$ dimensional vector reflecting time-varying interdependence between the financial institution j and both financial institution $1, \ldots, j-1, j+1, \ldots, J$ and the m-dimensional macroeconomic factors in the moving window s. $\alpha^s_{j,t}$ is a constant intercept for each moving window.

Following Li and Zhu (2008) and Belloni et al. (2011), the regularized model with penalization parameter $\lambda \geq 0$:

$$\min_{\alpha^s_j, \beta^s_j} \{\frac{1}{n} \sum_{t=s}^{s+n-1} \rho_\tau(X^s_{j,t} - \alpha^s_{j,t} - A^{s\top}_{j,t}\beta^s_j) + \lambda^s_j \|\beta^s_j\|_1\} \tag{3.2}$$

is referred to as the quantile Lasso regression for the $100\tau\%$ quantile function. Here $\tau \in (0,1)$ represented the tail risk level, which is normally defined as 5% or 10%. The necessaries and advantages of using Lasso in the high-dimensional model have been explained in chapter 2.4. $\|\beta_j^s\|_1$ is the sum of all β_j within one moving window s. The loss or equivalently the check function ρ_τ could be rewritten as the asymmetric absolute deviation function by Belloni et al. (2011):

$$\rho_\tau(u) = |u|^c |\tau - \mathbf{I}_{\{u<0\}}|, \tag{3.3}$$

with $c = 1, 2$. $c = 1$ corresponds to quantile regression which is hereinafter applied and $c = 2$ corresponds to expectile regression, see Ren et al. (2021). Taking $u = X_{j,t}^s - \alpha_j^s - A_{j,t}^{s\top}$, then the first part of equation 3.2 is rewritten as:

$$\rho_\tau(X_{j,t}^s - \alpha_j^s - A_{j,t}^{s\top}\beta_j^s) = \begin{cases} \tau \cdot \left(X_{j,t}^s - \alpha_j^s - A_{j,t}^{s\top}\right) & \text{if } X_{j,t}^s - \alpha_j^s - A_{j,t}^{s\top}\beta_j^s > 0, \\ -(1-\tau)\cdot(X_{j,t}^s - \alpha_j^s - A_{j,t}^{s\top}\beta_j^s) & \text{otherwise.} \end{cases} \tag{3.4}$$

As Belloni et al. (2011) argued, the use of penalization parameters may restore the consistency of high-dimensional quantile regression. As Bayer (2018) argues, the optimal quantile estimation consequently minimizes the expected loss of the error and penalization part because of the consistent check function.

3.2 penalization parameter λ

The choice of regularization parameter λ is critical since it balances the quantile loss and the penalty in equation 3.4 proposed by Li and Zhu (2008). Since the L1-norm loss function and penalization parameter are non-differentiable, numerical estimation and optimization is not a trivial problem. Zbonakova et al. (2016) mentioned that appropriate penalization parameters λ could be determined in a completely data-driven way. The research by Osborne et al. (2000) has led light to derive a formula for estimating penalization parameterλ in linear Lasso regression, see also Zbonakova et al. (2016). If λ is treated as a fixed value in the objective function of the penalized regression within each moving window s:

$$f(\beta,\lambda) = \frac{1}{n}\sum_{t=s}^{s+n-1}\left(X_{j,t}^s - \alpha_j^s - A_{j,t}^{s\top}\beta\right)^2 + \lambda\sum_{j=1}^{p}|\beta_j| \tag{3.5}$$

where X_t is a vector of log returns over a moving window and A is a design matrix. Then $f(\beta,\lambda)$ is a convex function in β. Moreover, as $\beta \to \infty$, then $f(\beta,\lambda) \to \infty$. Therefore, there is at least one minimum of the function $f(\cdot,\lambda)$ and $\widehat{\beta}(\lambda)$ minimizes $f(\beta,\lambda)$ if and only if the null-vector is an element of the sub-differential $\partial_\beta f(\beta,\lambda)$. From the formulation of penalized regression model and its objective function 3.5, Osborne (1985) and Mihoci et al. (2020) have defined the sub-differential as:

$$\partial_\beta f(\beta, \lambda) = -A^\top ((X - \alpha) - A\beta) + \lambda v(\beta) \tag{3.6}$$

where $v(\beta) = (v_1(\beta), \ldots, v_p(\beta))^\top$ is given as $v_i(\beta) = 1$ if $\beta_j > 0$, $v_i(\beta) = -1$ if $\beta_j < 0$, and $v_i(\beta) \in [-1, 1]$ if $\beta_j = 0$. Then $\widehat{\beta}$ minimizes $f(\cdot, \lambda)$ if the condition is to be satisfied:

$$0 = -A^\top \left((X - \alpha) - A\widehat{\beta}(\lambda)\right) + \lambda v\left(\widehat{\beta}(\lambda)\right). \tag{3.7}$$

Mihoci et al. (2020) and Zbonakova et al. (2016) emphasized that λ is firstly selected and then $\widehat{\beta}(\lambda)$ which minimizes the function 3.5 is searched. Hence, the estimation of a parameter vector β is denoted as a function of λ. The definition of $v(\beta)$ implies that $v(\beta) \cdot \beta = \|\beta\|_1$, $\|\cdot\|$ denotes L1-norm with p elements. Then the equation 3.7 is rewritten as:

$$\lambda = \frac{\left((X - \alpha) - A\widehat{\beta}(\lambda)\right)^\top A\widehat{\beta}(\lambda)}{\|\widehat{\beta}(\lambda)\|_1}. \tag{3.8}$$

Zbonakova et al. (2016), Zboňáková et al. (2019) and Mihoci et al. (2020) indicated that the form of λ characterizes three main effects which influence the values of the penalization parameter λ:

1. The variance or magnitude of the residuals $\left((X - \alpha) - A\widehat{\beta}(\lambda)\right)$. As the variance of residuals increases, so does the associated λ leading to an increase of sparsity of $\widehat{\beta}(\lambda)$. An increase in the variance of residuals indicates a drop in the signal, so the ration of the data is raised. If the size of the residual decreases, the value of λ decreases.

2. The absolute size of model coefficients $\widehat{\beta}$ of the model. This effect could translated into the effect of the number of nonzero covariates, which is so-called active set of the model $q = \|\beta\|_0 = \sum_{j=1}^{p} \mathbf{I}(\beta_j \neq 0)$. It has to be explained that a small number of λ is required in order to recover regression coefficients with a large L1-norm. The contributions of these effects are discussed and empirically tested by Zboňáková et al. (2019) in the L1-regularized linear model using time-varying data. If the size of the active set increases, the value of λ will decrease.

3. The correlation within design matrix A the matrix $A^\top A$. The change of λ is no longer linear dependent on the change of $A^\top A$. As the correlation of covariates increases, the higher value from the numerator of λ makes it higher. However, the active set of model coefficients decreases once we reach the peak of multicollinearity, which results in a decrease of λ.

According to the work by Tibshirani (1996) and Li and Zhu (2008), the equation 3.5 could be rewritten as a constrained optimization problem:

$$\min_{\alpha_j,\beta_j} \frac{1}{n} \sum_{t=s}^{s+n-1} \rho_\tau(X_{j,t}^s - \alpha_{j,t}^s - A_{j,t}^{s\top} \beta_j^s), \tag{3.9a}$$

$$\text{subject to } |\beta_j^s\|_1 \leq u \tag{3.9b}$$

for $u > 0$. Equations 3.9a and 3.9b can be rewritten as a Lagrangian primal function, details see Li and Zhu (2008). Through setting the derivatives of the Lagrangian primal function 3.9b to zero and the keeping all Lagrangian multipliers non-negative, Li and Zhu (2008) derived the size of penalization parameter λ in quantile Lasso regressions as follows:

$$\lambda = \frac{\theta^\top A \widehat{\beta}(\lambda)}{\|\widehat{\beta}(\lambda)\|_1} \tag{3.10}$$

where $\theta^\top = (\theta_s, \cdots, \theta_{s+n-1})$ satisfies the following relationships:

$$\theta t = \begin{cases} \tau & \text{if } X_{j,t}^s - \alpha_j^s - A_{j,t}^{s\top} \beta_j^s > 0, \\ (1-\tau) & \text{if } X_{j,t}^s - \alpha_j^s - A_{j,t}^{s\top} \beta_j^s < 0, \\ \in (-(1-\tau), \tau) & \text{if } X_{j,t}^s - \alpha_j^s - A_{j,t}^{s\top} \beta_j^s = 0. \end{cases} \tag{3.11}$$

Equation 3.10 for λ in the quantile Lasso regression revealed that λ depends on the size of residual, the active set of model coefficients which is influenced by the covariance matrix of the design matrix. By Härdle et al. (2017), considering the definition of the design matrix in 3.1, equations 3.8 and 3.10 could be interpreted that λ depends not only on the volatility but also the connectedness of financial institutions.

In theory, every solution of the Lasso optimization problem holds the equation 3.8 and 3.10 since after λ being chosen, the model is post-fitted according to the given λ. The standard approaches for choosing the regularization parameter λ include information criteria and cross-validation. The latter methods have three forms, namely k-folds, leave-one-out and generalized cross-validation (Koenker et al., 1994). Among former approaches, Schwarz information criterion (SIC) by Koenker et al. (1994), also known as Bayesian Information Criterion (BIC) is the most widely used criterion for selecting λ. Cross-validation is aimed to minimize prediction error on each grid of the penalization parameter λ. Although, as Leng et al. (2006) argued, this method of choosing penalization parameter based on predicting accuracy is generally not consistent when variable selection is required. The generalized approximate cross-validation(GACV) by Yuan (2006) is still efficient in terms of model error.

Following criteria are conducted in previous research about FRM:

$$SIC(\lambda_j^s) = \ln(\frac{1}{n}\sum_{t=s}^{s+(n-1)} \rho_\tau(X_{j,t}^s - \alpha_{j,t}^s - A_{j,t}^{s\top}\beta_j^s) + \frac{\ln n}{2n}df, \tag{3.12}$$

$$GACV(\lambda_j^s) = \frac{\sum_{t=s}^{s+(n-1)} \rho_\tau(X_{j,t}^s - \alpha_{j,t}^s - A_{j,t}^{s\top}\beta_j^s(\lambda))}{n - df}, \tag{3.13}$$

where df is a measure of the effective dimensionality of the fitted model. Li and Zhu (2008) proved that the number of interpolated observations $X_{j,t}$, or equivalently, the number of non-zero coefficients of the fitted model is an estimate of the df. Because GACV also works for $p > n$, in where the size of covariates including the number of institutions and macroeconomic factors could be larger than the moving window size. Moreover, in terms of statistical efficiency, GACV outperforms SIC argued by Yuan (2006). Therefore, GACV is implemented in this paper as the criterion to optimize our penalization parameter λ in our quantile Lasso model.

3.3 Empirical penalization parameters $\widehat{\lambda}_j$

In this chapter, we apply data from the real stock markets to quantile Lasso regression model introduced in chapter 3.1, and penalization parameter estimation and optimization algorithm described in chapter 3.2, then we obtain empirically estimated penalization parameters λ of individual financial institutions in the stock market during a particular time period respectively. Based on the characteristics of these estimated penalization parameters, we will define other FRM candidates to capture systemic risk.

This paper focuses on two major stock markets, the Americas and Europe, because both stock markets attract large financial institutions to be published there. The performance of publicly traded financial institutions in both stock markets reflects whether the financial system works properly. Moreover, Zbonakova et al. (2016), Yu et al. (2017), Yu et al. (2019) and Mihoci et al. (2020) have proved that penalization parameters λ fluctuate similar along with the evolution of systemic risk. Furthermore, we are interested in the performance of systemic risk measures under high systemic risk and we want to reduce the computation cost, so we focus on the application around the financial crisis of 2007 and the current COVID-19 pandemic. The purpose of choosing two stock markets and two time intervals is to double-check and compare the performance of FRM candidates under financial distress that occurred for different reasons in two regions.

Analogue to the procedure introduced by Mihoci et al. (2020), we need to first collect the relevant data about individual financial institutions and the overall market on each trading day for both stock markets. In order to prevent the possible survivorship bias, a list of financial institutions that have been an active constituent of the aimed stock market indices are selected. In the Americas, all (395) financial institutions from the US's S&P 1500 Composite Index and the Canadian TSX Toronto Composite Index have been considered. At the same time, all (187) financial institutions from S&P 600 Europe have been selected among 17 European countries: Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland and the United Kingdom. For each component of the above mentioned financial institution lists, daily closing prices P_t and the market capitalization at the closing price are downloaded from the Bloomberg dataset. Moreover, the daily value of six macroeconomic factors for the Americas and seven macroeconomic factors for Europe are also downloaded from the Bloomberg dataset. Detailed information on these factors is available in Table 3.1. All macroeconomic factors are chosen in line with Adrian and Brunnermeier (2011) in order to capture the common explored macroeconomic risks in both markets. In this setup, data from 02 January 2007 until 31 December 2009 is collected to represent the stock market change around the financial crisis of 2007, while the stock market during COVID-pandemic is reflected with the data from 03 October 2020 and 31 December 2020. Since Equation 3.1 takes the log-returns of stock prices and macroeconomic factors into account, for the first moving window $s = 2$, $A^2_{j,2}$ includes the log-returns of individual financial institutions $X^2_{j,2}$ that is obtained from the logarithm of difference between the stock prices on 04 January 2007 and those on 03 January 2007, and the log-returns of macroeconomic factors are calculated as the difference between values on 03 January 2007 and on 02 January 2007. $T = 782$ indicates the last series of log returns within the first time interval. Analogously, 587 series of log returns are collected in the Americas and Europe around the COVID-19 pandemic.

Despite that FRM@Americas and FRM@Europe candidates are separately computed based on their data sources, the basic settings of linear quantile Lasso regression models are identical for both markets. In each stock market, we selected the largest 100 (J=100) financial institutions from the financial institutions list on every trading day based on daily market capitalization. The largest 100 financial institutions for each market vary over time so that we capture the crucial market participants precisely. We determine J as 100 according to the following reasons mentioned by Yu et al. (2017). Firstly, the market capitalization of the largest 100 institutions covers more than 85% of total market capitalization in the US stock markets. Secondly, if we use 200 financial institutions, the smaller companies in this set change

FRM@Americas	FRM@Europe
S&P 500 Index Returns	S&P Europe 600 Index Returns
CBOE Volatility Index(VIX) Returns	Euro Stoxx 50 Volatility Index Returns
REIT Index Returns	MSCI Europe REIT Index Returns
3 months Treasury Constant Maturity Rate	1 year German Treasury Constant Maturity
3 months Treasury Constant Maturity Rate Differences to 10 year Treasury Constant Maturity Rate Spread Differences	German Treasury Constant Maturity Rate 1 to 10 years Slope Spread Differences
Moodys Seasoned Baa Corp Bond Yield Spread to 10 year Treasury Constant Maturity Rate Spread Differences	Barclays Bloomberg EuroAgg Corporate Yield Spread to 10 year German Treasury Constant Maturity Rate Spread Differences
	10 year Italy Treasury to 10 year German Treasury Constant Maturity Rate Spread Differences

Table 3.1: Macroeconomic risk factors for FRM@Americas and FMR@Europe.

regularly due to bankruptcies or other reasons. The financial institutions list for Europe is smaller than 200. Moreover, the penalization parameters estimated based on the largest 100 or 200 institutions are very similar. Then each linear quantile Lasso regression for Americas has 105 covariates and each regression for Europe has 106 covariates. The moving window size is determined as 63 days (n=63), which represents three months since only trading days are considered as stated by Härdle et al. (2017). Yu et al. (2017) found that deciding on the optimal window size is a trade-off. On the one hand, the moving window size smaller than 50 may be imprecise. On the other hand, the larger moving window size leads to more lagged data. Relying on the estimate of standard FRM with $n = 63$ and $n = 126$, Yu et al. (2017) concluded that standard FRM with $n = 63$ leads standard FRM with $n = 126$ at least 22 trading days. Therefore, the moving window size is set to be $n = 63$. We take tail risk level τ as 5%, following Mihoci et al. (2020).

With reference to the data input and parameter settings, I will elaborate the process of estimating penalization parameters step by step. Within each moving window, the algorithm designed by Li and Zhu (2008) is used to estimate penalization parameter λ in the aforementioned linear quantile Lasso regression model. The estimation result at 64th trading day is calculated based on log returns within the first moving window $s = 2$, i.e. the log-returns of macroeconomic factors from trading day $s - 1 = 1$ until trading day $s + (n - 2) = 63$ and the log-returns of financial institutions from trading day $s = 2$ until trading day $s + (n - 1 = 64)$ are considered. Based upon the τ at 5%, among 63 quantile regressions within each moving window for each institution, the first 5% of them and the rest 95% are differently treated in accordance with the check function 3.4. Optimized penalization parameter on each trading day for each financial institution $\widehat{\lambda}_j^s$ has the minimal GACV value among all $GACV\ (\lambda_j)$ after 25 iterations. Next, we repeat this estimation and optimization process for J institutions. Because on each trading day t within a given window s, we take the log return of 100 largest financial institutions $X_{1,t}^s, \ldots, X_{100,t}^s$ on each market as the dependent variable $X_{j,t}^s$ in equation 3.1 sequentially. In other words, we get $J = 100$ linear regressions on a given trading day within a given moving window. As a result, we obtain a vector $\widehat{\lambda}_j^s \in \{\widehat{\lambda}_1^s, \ldots, \widehat{\lambda}_{100}^s\}$ within the moving window s. Then after rolling the moving window along the pre-defined time interval, we get $T - 63$ series of penalization parameters λ from $T - 63$ moving windows. To simplify the notation hereinafter, we take the set of penalization parameters estimated from s-th moving window as the penalization parameters on the $s + (63 - 1)$-th trading day. Thus, we attain estimated penalization parameters λ for Americas and Europe from 02 April 2007 until 31 December 2009 (719 trading days) and from 01 January 2019 until 31 December 2020 (524 trading days). In order to find out reasonable quantiles to define novel FRM, the characteristics of empirically estimated penalization parameters $\widehat{\lambda}_{j,t}$ are illustrated in the next chapter.

3.4 Distribution of penalization parameter $\widehat{\lambda}_j$

The $\widehat{\lambda}_{j,t}$ estimated based on stock returns and macroeconomic factors over the last 63 trading days reveals the interconnections among large financial institutions in the corresponding financial system will be visualized and investigated in this chapter. To understand the risk level of an individual institution that is mirrored by the estimated penalization parameter $\widehat{\lambda}_j$ within the predetermined time frame, we have produced a daily boxplot to visualize the descriptive statistics of these estimated penalization parameters on each trading day. After sorting the financial institutions in ascending order by their daily $\widehat{\lambda}_j$ values, the lower bound and upper bound of daily boxplot in Figure 3.1 show the penalization parameter value of the 25-th financial institution and the 75-th financial institution on corresponding trading day. Red and blue lines in Figure 3.1 represent the maximal and averaged daily estimated penalization parameters $\widehat{\lambda}_j$ respectively.

From left plots in 3.1, we observe that $\widehat{\lambda}$ in the Americas and in Europe are generally stable from April 2007 until the second quarter of 2008. $\widehat{\lambda}$ increases during the third quarter of 2008, which corresponds to the bankruptcy of Lehman Brothers and the start of an international bank crisis. The peak at the end of 2008 indicates the great financial crisis and great recession. Our penalization parameters captured the increase of stress level before the upcoming of a worldwide crisis comparing with the timeline of the financial crisis of 2007. For both markets, estimated penalization parameters $\widehat{\lambda}$ recovered to a stable pattern after the curve around March 2009. The decrease of stress is the response to government interactions, and the rebound could be a seasonal downturn in stock markets. The crisis decayed quickly during the second quarter of 2009, when the global financial markets were recovering. The second plot in both figures is the results for the COVID-19 pandemic. It shows that $\widehat{\lambda}$ in Europe and the Americas kept at a low level and stable until the beginning of 2020. $\widehat{\lambda}$ increases suddenly in March 2020 because of the spread of Coronavirus. The high-stress level of $\widehat{\lambda}$ across the second quarter of 2020 is the result of worldwide preventive measures, including lockdowns, travel restrictions, faculty and store closures. From the above analysis, we may conclude that the estimated penalization parameters could capture periods with high systemic risk correctly and efficiently. Comparing $\widehat{\lambda}$ of 2007-2008 with $\widehat{\lambda}$ of 2019-2020, we found that the $\widehat{\lambda}$ reacted with delay after the outbreak of COVID-19 in the Americas and Europe, and the high-stress level has no seasonal fluctuations. This finding implies that the penalization parameter may capture high systemic risk much better and earlier, for which trigger event is relative to financial institutions. Comparing the results for Americas and in Europe at the same time period, $\widehat{\lambda}$ in Americas tends to be higher than $\widehat{\lambda}$ in Europe.

In both markets, we observe that several maxima were extraordinary high. Moreover, many financial institutions were detected as outliers on specific trading days since they had high daily penalization parameters, which are identified with $\circ$. Higher $\widehat{\lambda}_{j,t}$ indicates the institution j has a high-risk level as companies at the origin of the crisis. Therefore, the institutions having $\widehat{\lambda}_{j,t}$ that is picked out with black circles are expected to have high "co-stress" or even to be seen as major contributors of systemic stress by Mihoci et al. (2020). In the context of standard FRM, it may be sensitive to these financial institutions with high "co-stress". The outstanding maxima may lead to an overestimated standard FRM. This motivates us to define novel FRM with quantiles so that only partition of the penalization parameters will be

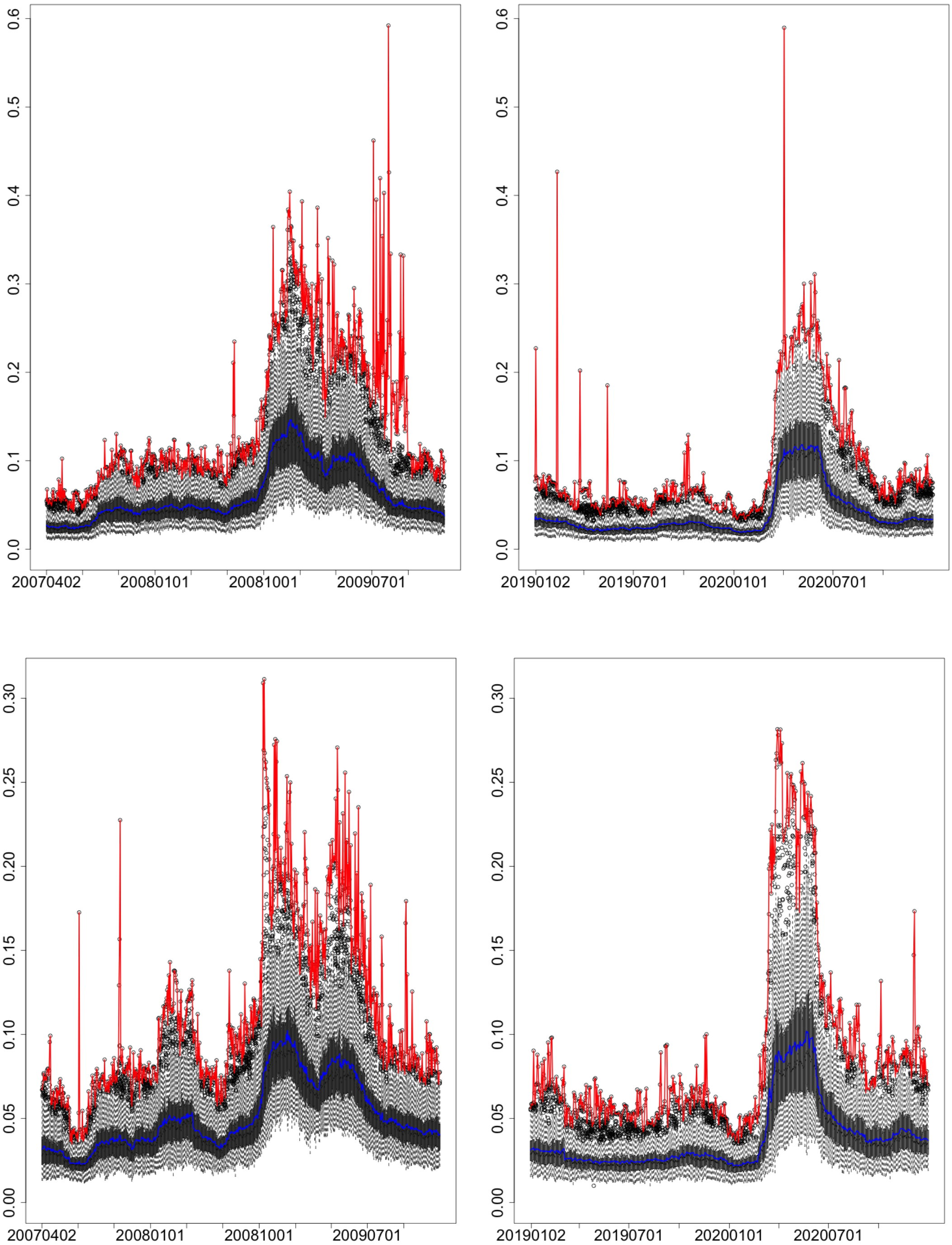

Figure 3.1: Boxplot, average and maxima of estimated daily penalization parameters from April 2007 to December 2009 and from January 2019 to December 2020 in the Americas (top) and Europe (bottom).

FRMQLambdaboxplot

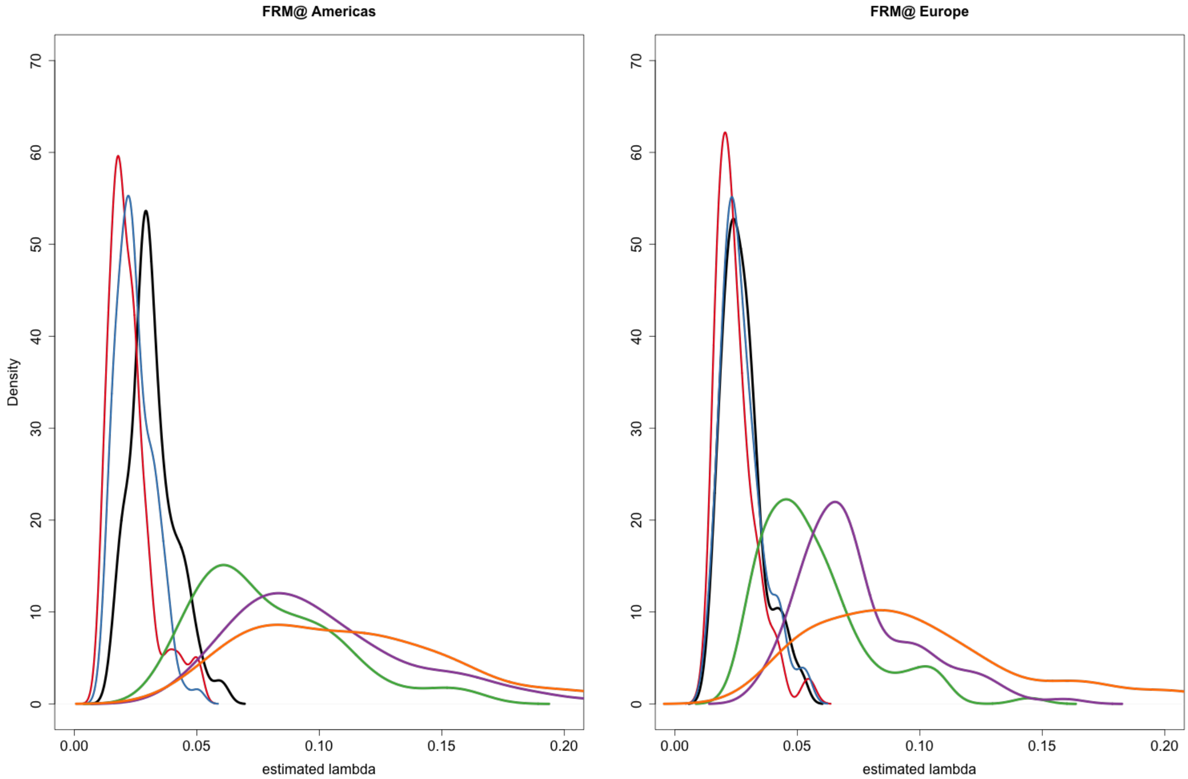

Figure 3.2: Kernel density estimation examples of $\widehat{\lambda}_{j,t}$ for 100 financial institutions on 20070801, 20190603, 20191203, 20081003, 20090202, 2020602.

FRMQlambdaDistr

employed as a systemic risk measure. During the period with higher financial distress, for both financial markets, the distances between upper and lower bounds of boxplots are larger. This finding is consistent with the estimated distribution of daily penalization parameters series in Figure 3.2. Figure 3.2 shows kernel density estimations of daily penalization parameters on six trading days in the Americas and in Europe. 20081003, 20090202 and 20200602 in Figure 3.2 are selected from periods under high stress, while the other three lines are examples from stable periods. During trading days under high risk, the daily series of $\widehat{\lambda}_{j,t}$ have a flatter and broader distributions than the series of $\widehat{\lambda}_{j,t}$ estimated from stable periods. On the basis of this fact, we would like to take the variability of daily penalization parameters $\widehat{\lambda}_{j,t}$ as a systemic risk measure. From Figure 3.2 we discover that daily penalization parameters $\widehat{\lambda}_{j,t}$ tend to be distributed as Weibull distribution. Although the exact distributions of $\widehat{\lambda}_{j,t}$ series are still unknown, we may conclude that they are right-skewed distributions. Consequently, if a fixed proportion of financial institutions should be considered as a systemic risk measure, the size of this proportion should be relatively large. For example, if we take the penalization parameters of the best 70 financial institutions into account rather than that of the best 30 financial institutions, the differentiation of penalization parameter value due to the change of general systemic risk will be larger. Thus, it is reasonable that in further analysis, we will focus on financial institutions with relative high risks. The definition of novel FRM premised on the characteristics of daily penalization parameters for financial institutions will be demonstrated in the next chapter.

3.5 Definition and visualization of Financial Risk Meter (FRM)

Yu et al. (2019), Yu et al. (2017) and Mihoci et al. (2020) defined standard FRM as

$$FRM^s_{mean} = J^{-1} \sum_{j=1}^{J} \widehat{\lambda}^s_j \tag{3.14}$$

in the specified moving window s for j companies. Analogue to the notation of penalization parameters, FRM^s_{mean} is further denoted as $FRM_{mean,t}$, which denotes FRM on trading day t with $t = s + (63 - 1)$. Standard FRM is induced by risk levels of all financial institutions represented with values of penalization parameters. The maxima and outliers of daily penalization parameters may result in an unrobust FRM. Therefore, in this paper, we extend the previous analysis by defining the following novel FRM candidates based on p-th quantile of F for $0 < p < 1$:

$$F^{-1}(p) = \inf\{\widehat{\lambda}_t : F\left(\widehat{\lambda}_t\right) \geq p\}, \tag{3.15}$$

where $\widehat{\lambda}_t$ is a vector of $\widehat{\lambda}_j$ for all financial institutions on trading day t. F is the empirical cumulative distribution function of $\widehat{\lambda}_t$ such that $F(FRM_{p-,t}) \leq p \leq F(FRM_{p,t})$. $FRM_{p,t}$ delivers the level of estimated penalization parameter collected from best p institutions. Better institutions herein corresponds to lower values of penalization parameters, which signify that these financial institutions are less like to contribute to a growing systemic risk. As argued in chapter 3.4, we have found that the increased value of $\widehat{\lambda}_t$ for each one percent in left tail of daily $\widehat{\lambda}_t$ distribution is smaller than that in right tail because of the positive skewness of density distributions in Figure 3.2. In other words, taking a large p may result in a significant difference of $F^{-1}(p)$, equivalently to $FRM_{p,t}$. Therefore, we are interested in p larger than 50%. We define FRM_{q50}, FRM_{q60}, FRM_{q70}, FRM_{q80}, FRM_{q90} to reflect the interdependent stress level among 50, 60, 70, 80, 90 top financial institutions. In other words, these novel FRM candidates discover the distress level referring to 50, 40, 30, 20 and 10 most risky financial institutions. The interquartile range, which is calculated as the distance between first and third quartile of $\widehat{\lambda}$, increases under high distress in Figure 3.1 and Figure 3.2. Then we derive $FRM_{IQR} = \frac{1}{2}\left(F^{-1}(75\%) - F^{-1}(25\%)\right)$ to represent systemic risk with the volatility level of penalization parameters. To simplify the notation, we denote standard FRM and novel FRM for each trading day as $FRM_{p,t}$ with the index of FRM measures $q \in \{mean, q50, q60, q70, q80, q90, IQR\}$. All novel FRM definitions and standard FRM for Americas and Europe are taken as FRM candidates for the optimal FRM@Americas and FRM@Europe.

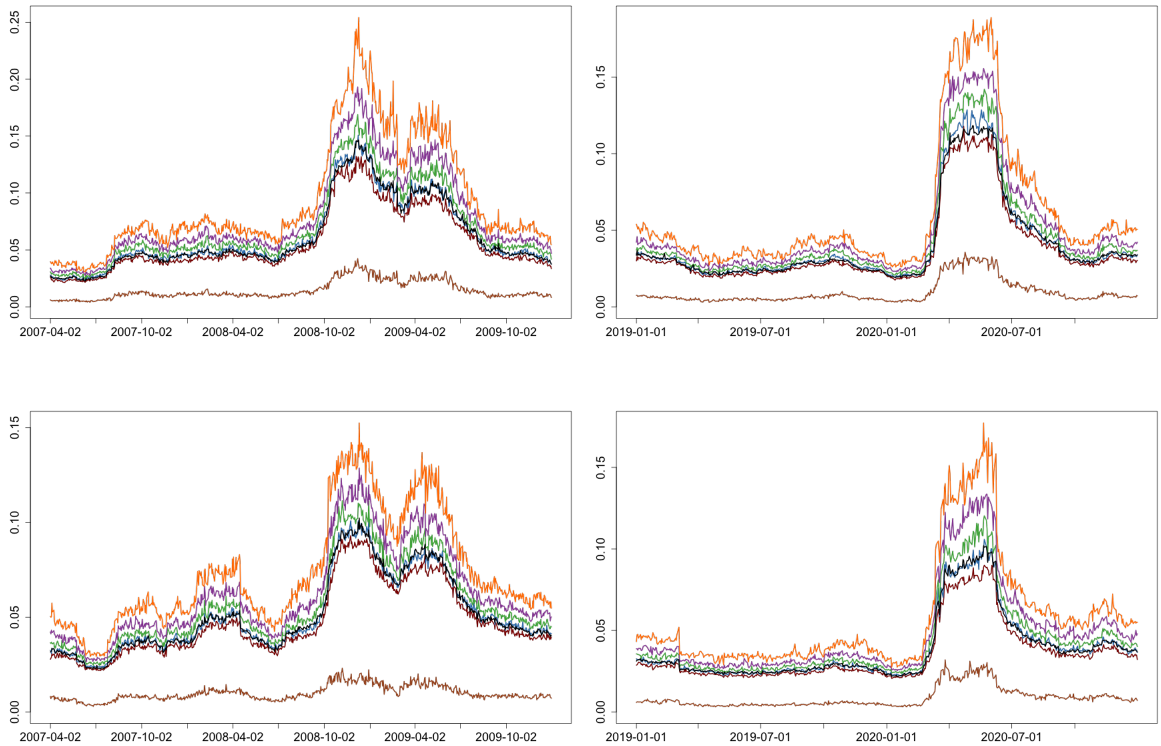

Figure 3.3: Time series of standard FRM, FRM_{q50}, FRM_{q60}, FRM_{q70}, FRM_{q80}, FRM_{q90}, FRM_{IQR} for Americas (top) and Europe (bottom) from April 2007 to December 2009 and from January 2019 to December 2020.

FRMQdefineFRMs

Figure 3.3 visualizes evolution of $FRM_{p,t}$@Americas and $FRM_{p,t}$@Europe candidates. In both markets, the overall fluctuations of FRM candidates are identical to the evolution of penalization parameters in chapter 3.4. FRM@Americas and FRM@Europe candidates hold identical characteristics during two time intervals, so the following interpretations are appropriate to both markets during both periods. It is observed that novel FRM candidates shift almost parallel upwards along with the increase of their underlying probabilities, especially during stable periods. Standard FRM lies above FRM_{q50} and is close to FRM_{q60}. The change of FRM_{q80} and FRM_{q90} is much more considerable than the change of other FRM candidates when large variations came, such as around October 2008 and around March 2020. Moreover, the distance between two the higher quantiles is also larger than the distance between the two lower quantiles. For example, FRM_{q80} and FRM_{q90} is larger than the distance between FRM_{q50} and FRM_{q60}. This difference is enlarged under high stress. However, it is impossible to detect whether FRM candidates based on higher quantiles could forecast the upcoming crisis much earlier than other candidates from the visualization plots. Despite the lower values, the trend and peak of FRM_{IQR} are similar to that of standard FRM. In summary, novel FRM candidates generally perform similarly to the standard FRM. Above mentioned findings could be proved by the kernel density estimations of all FRM candidates in Figure 3.4. In both markets during both periods, all density distributions of FRM candidates are bimodal. The first peak stands for the FRM value that occurred most frequently within a stable period, and the second peak corresponds to the mode of FRM value during periods under high distress. FRM defined in different ways have the same mode during stable periods. However, FRM derived with higher quantiles tend to deliver a higher value under distress.

After comparing the same candidates in two markets, we may conclude that FRM@Americas are likely to be higher than FRM@Europe, especially when the whole financial system was under stress. In other words, financial institutions based in the Americas may suffer more losses during the financial crisis of 2007 and the COVID-19 pandemic than financial institutions based in Europe. We cannot distinguish the density distributions of FRM candidates during the financial crisis between the density distributions of FRM candidates during the COVID-19 pandemic easily. It means although the COVID-19 pandemic is not due to regular activities in financial markets, it made an enormous influence on the financial system. All FRM candidates are able to capture the volatility of systemic risk regardless of its cause.

In this chapter, we recognize that it is hard to distinguish FRM candidates based on visualization plots. Hence, in the next chapter 4.1, we will implement correlation, goodness-of-fit and causality tests between individual FRM candidates and individual commonly used systemic risk measures to compare the capabilities of FRM candidates in terms of measuring systemic risk in American and European countries.

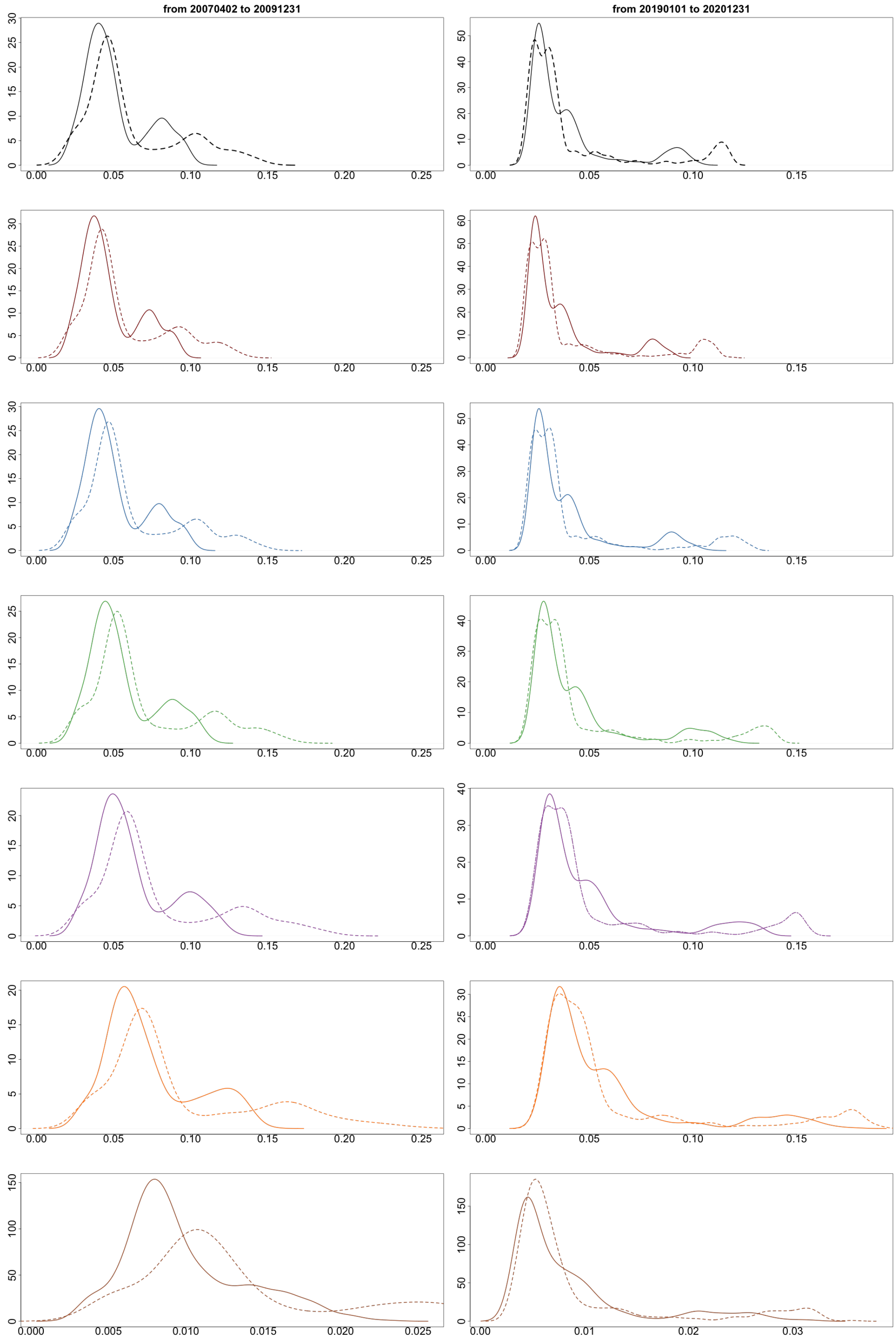

Figure 3.4: Kernel density estimation of standard FRM, FRM_{q50}, FRM_{q60}, FRM_{q70}, FRM_{q80}, FRM_{q90}, FRM_{IQR} in the Americas and Europe.

FRMQFRMdistr

4 Comparison of FRM and other systemic risk measures

Yu et al. (2017), and Yu et al. (2019) have proved the mutual correlations and causalities between standard FRM@Americas and other systemic risk measures like VIX, $SRISK$, Google Trends and Financial Turbulence. Moreover, Mihoci et al. (2020) illustrated the similarity between standard FRM@Europe and systemic risk measures for Europe, namely CISS and VSTOXX. In this chapter, we will extend the correlation and causality tests firstly to standard FRM@Europe, and secondly to other novel FRM candidates for both markets. Another extension is that we will try to test the goodness-of-fit between the distribution of FRM candidates and the distribution of each systemic risk measures.

4.1 Data source of other systemic risk measures

Historical time series of estimated penalisation parameters $\widehat{\lambda}$ and FRM candidates have been illustrated in chapters 3.3 and 3.5. The overall trend seems to be reasonable according to the results of previous research and our common sense. However, we still need statistical methodologies to validate that FRM candidates are qualified systemic risk measures. Moreover, FRM candidates are not distinctly differentiated from each other. Thus, several systemic risk measures are introduced to validate the co-movement and co-stress of FRM candidates in term of working as a systemic risk indicator in American and European countries. For Americas, as mentioned in Chapter 2.1, we collected daily VIX index from Bloomberg database, weekly Google Trends (GT) data in the United States and Canada with the keyword "financial risk" with R package by Massicotte et al. (2016), and monthly SRISK data in the United States and Canada from Engle et al. (2020). Cubic interpolation is applied to transform weekly GT series and monthly SRISK records into daily time series, see Yu et al. (2017). For Europe, daily CISS are downloaded from European Central Bank Statistical Data Warehouse and daily VSTOXX are also from the Bloomberg database. For both markets, we calculated a Financial turbulence index (FT) according to the algorithm introduced by Kritzman and Li (2010) and Zbonakova et al. (2016) based on the same data that is applied to calculate FRM candidates. Kritzman and Li (2010) proposed a turbulence index measure the financial turbulence statistically:

$$d_t = (X_t - \mu)\,\Sigma^{-1}\,(X_t - \mu)^\top\,, \tag{4.1}$$

where d_t is the turbulence for a given time t, X_t is a series of historical stock returns for time t. μ is sample average historical returns, and Σ is a covariance matrix of historical returns.

We take stock returns of a list of financial institutions as X_t. The financial institutions list is a joint list of the largest 100 financial institutions within the whole time frame for simplicity. t is specified as the same time frame in which FRM candidates are calculated. Following Yu et al. (2017), all series are standardized to interval zero and one for comparison purposes. VIX and VSTOXX, as the most commonly used systemic risk measures in the Americas and Europe, are macroeconomic factors that are considered to calculate FRM@Americas and FRM@Europe, respectively. However, as mentioned in chapter 2.4, several macroeconomic factors may be removed from our initially structured linear quantile Lasso regression model 3.1 so that they will not be included in the estimation of penalization parameter $\lambda_{j,t}$. Through checking the selected variables in quantile Lasso regression models on at a daily frequency. On average, we find that on each trading day, the VIX is removed from approximately 45 models among 100 models for American financial institutions, and the VSTOXX is removed from about 46 models among 100 models for European financial institutions. Furthermore, after averaging the estimated coefficients of VIX and VSTOXX for all models among the entire time frame, we obtained an estimated coefficient of VIX at -0.004 and an estimated coefficient of VSTOXX at -0.006. In the context of equation 3.8, such slight coefficients cannot influence the estimation of the penalization parameter significantly, not to mention the value of FRM candidates. Therefore, VIX and VSTOXX are still reasonable to be considered as comparable systemic risk measures. VIX and VSTOXX are considered as the standard systemic risk measure in the Americas and Europe, respectively. So we will focus more on them in further analysis. Since the measures mentioned above are separately defined and evaluated in their way, they will be scaled or normalized for further comparison purpose as follows:

$$\text{normalizedFRM}_{\text{p,t}} = \left(FRM_{p,t} - \min\left(FRM_{p,t}\right)\right) / \left(\max\left(FRM_{p,t}\right) - \min\left(FRM_{p,t}\right)\right) \quad (4.2)$$

The first two plots in Figure 4.1 visualize the comovement of $FRM@Americas$ candidates and other systemic risk measures in the Americas during the financial crisis of 2007. The first finding is that all $FRM@Americas$ candidates and systemic risk measures for Americas except SRISK reached at peak right when the financial crisis became systemically essential. Standard FRM captured the increased level of systemic risk much earlier than VIX, GT and FT. Although SRISK increased earliest among all systemic risk measures before the crisis occurred, it cannot indicate the peak of crisis exactly. Two summits of SRISK were either earlier or later than the maximal values of other measures. SRISK showed a progressive growth in the systemic risk while VIX, GT, FT and $FRM@Americas$ candidates experienced a dramatic enlargement of systemic risk level until the whole financial systemic was threatened. Not only

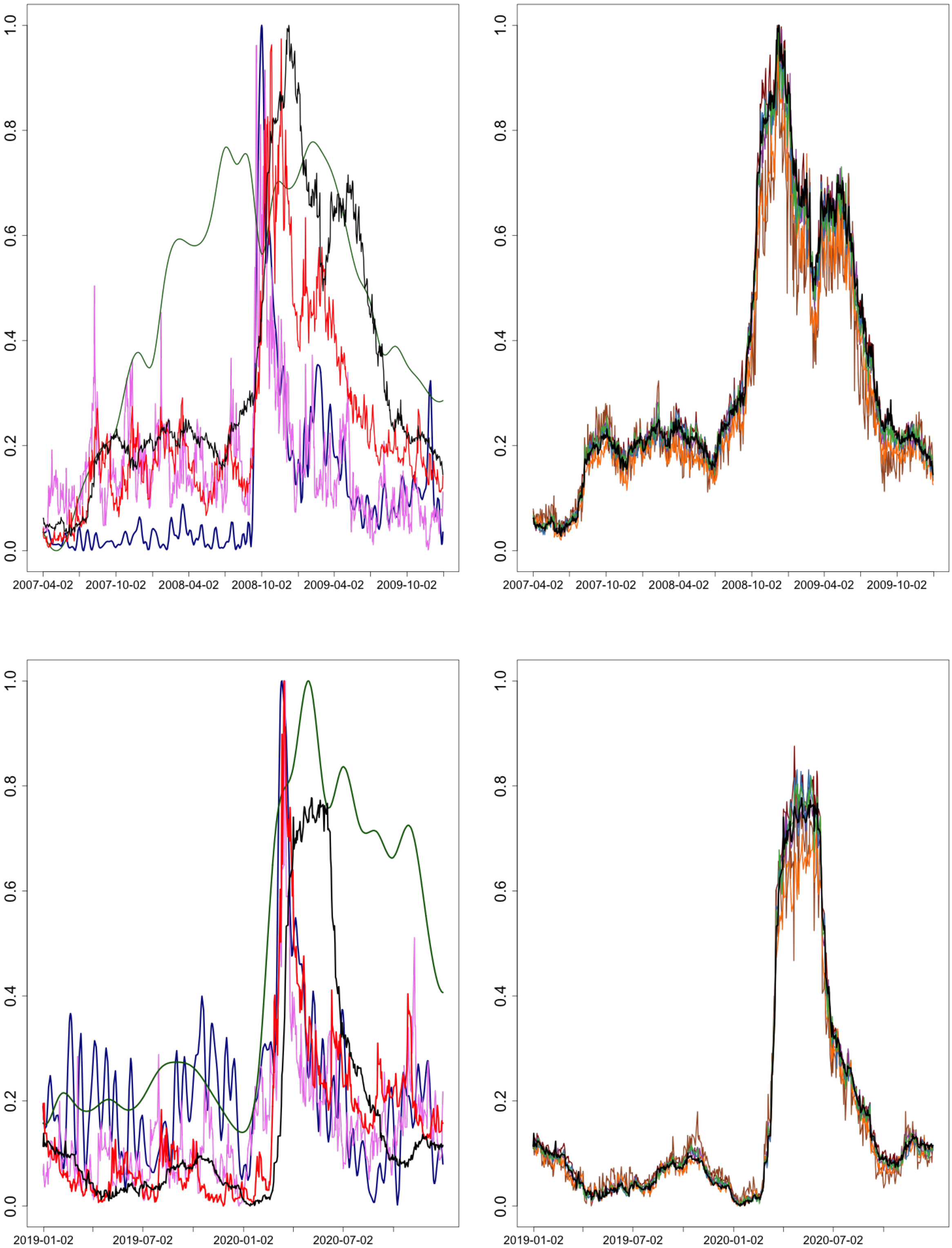

Figure 4.1: Normalized systemic risk measures VIX, SRISK, GT, FT, and standard FRM, FRM_{q50}, FRM_{q60}, FRM_{q70}, FRM_{q80}, FRM_{q90}, FRM_{IQR} for Americas during the financial crisis of 2007 and the COVID-19 pandemic.

$FRM@Americas$, but also VIX, GT, and FT had a synchronized bounce of systemic risk level at the beginning of 2009. However, in this period, $FRM@Americas$ may overestimate the undertaken distress since its value was higher than the value of the other three measures. In other words, there was a longer delay for standard $FRM@Americas$ to reflect the recovery from the financial crisis 2007. This could be seen from the increased deviation of standard $FRM@Americas$ and other systemic risk measures in the mid of 2009. As for the comparison of $FRM@Americas$ candidates, we find that they generally have a synchronized fluctuation pattern as the standard $FRM@Americas$. These normalized standard $FRM@Americas$, $FRM_{q50}@Americas$, $FRM_{q60}@Americas$, $FRM_{q70}@Americas$ and FRM_{q80} are close to each other is consistent with the Figure 3.4, in where estimated density distributions of respective standard $FRM@Americas$ are almost identical. It means the risk that the top 20 risky financial institutions suffered was almost the same as the risk for top 30, 40 or 50 risky financial institutions. In this sense, we may conclude that even for relatively safe financial institutions that lie around the median among all the largest financial institutions, the influence of the financial crisis was still considerable. If we look into the slight difference among standard $FRM@Americas$ candidates, we may observe that the smaller quantile factor is used to calculate standard $FRM@Americas$ the higher its respective normalized FRM_{q50} is obtained. Considering the definition of normalization and the minimal value of all $FRM@Americas$ in Figure 3.4 being similar. The higher normalized $FRM@Americas$ values compared with values from other FRM types result from either a higher $FRM@Americas$ values at the same time or a smaller range of $FRM@Americas$ series within the given time frame. A striking finding is that the normalized standard $FRM@Americas$ tends to maintain a high level among all $FRM@Americas$ candidates during the period under distress. During the stable period, the normalized standard $FRM@Americas$ lie in the middle among all $FRM@Americas$ candidates. The higher normalized standard $FRM@Americas$ under distress proves our assumption that several standard $FRM@Americas$ during the crisis may be influenced by large outliers of daily λ_j. In contrast, the normalized $FRM_{IQR}@Americas$ tends to be the highest FRM candidates during a stable period and the lowest under distress. We also find that normalized $FRM_{q50}@Americas$, $FRM_{q60}@Americas$, $FRM_{q70}@Americas$ and FRM_{q80} series are closer to each other during the stable period, in comparison to the series during the crisis. This reveals that several $FRM@Americas$ candidates fluctuate more strongly than others, which is consistent with the larger deviation of $FRM@Americas$ series in Figure 3.3. The broader range of $FRM_{q90}@Americas$ leads to a smaller normalized series, so even if its absolute value is higher than other candidates, its normalized series is lower than other others. This implied that $FRM_{q90}@Americas$ cannot reflect the increased systemic risk as quickly and precisely as others since it always undertake a relatively high level of systemic risk. However, the lower normalized $FRM_{q90}@Americas$ does not imply a lower systemic risk for top 10 risky financial institutions because the high range is due to the high maximum of $FRM_{q90}@Americas$ during the crisis, which is the consequence of a high systemic risk of the most risky financial institution.

The other two plot in Figure 4.1 is the comparison results of above mentioned measures for the Americas during the COVID-19 pandemic. It is seen that $FRM@Americas$ candidates captured the high systemic risk later than other measures. The reason for this delay is that this COVID-19 pandemic was not relevant to activities in the American stock market. Thus, the reactions of the stock market to this spread of COVID-19 has postponed. Moreover, the influences of the pandemic on the whole financial system or economics were not real-time. $FRM@Americas$ candidates remains at a high level despite the sharp decrease of VIX, GT and FT from March 2020 to May 2020. After the dramatic decay in June, FRM candidates reached a comparable level as VIX, GT, FT. Like what occurred in the first research period, SRISK reacted earliest at the beginning of the pandemic and maintained the highest level when other systemic risk measures decayed. Among $FRM@Americas$ candidates, analogue to the finding from the first period, normalized $FRM_{q50}@Americas$, $FRM_{q60}@Americas$ and $FRM_{q80}@Americas$ were very close to standard $FRM@Americas$. The normalized $FRM_{q80}@Americas$ series was significantly higher during the pandemic than the normalized $FRM_{q90}@Americas$. We then conclude that the top eleven to 20 risky financial institutions were suffered the largest during this pandemic. From $FRM_{IQR}@Americas$ time series, we see that the deviation between risky and safe financial institutions during the period under distress is smaller than that before the pandemic. Nevertheless, the fluctuation range of $FRM_{IQR}@Americas$ is more extensive than the fluctuation range during the financial crisis of 2007. Another finding from the comparison between the two distress periods is that the systemic risk caused by the pandemic was not as considerable as that triggered by the financial crisis of 2007.

Figure 4.2 validates that $FRM@Europe$ candidates have also predicted the financial crisis of 2007 earlier than VSTOXX and FT. Like the SRISK for Americas, CISS for Europe was higher than $FRM@Europe$ and VSTOXX before and after the financial crisis. Similar to the performance of $FRM@Americas$ candidates, $FRM_{q90}@Europe$ and $FRM_{IQR}@Europe$ cannot

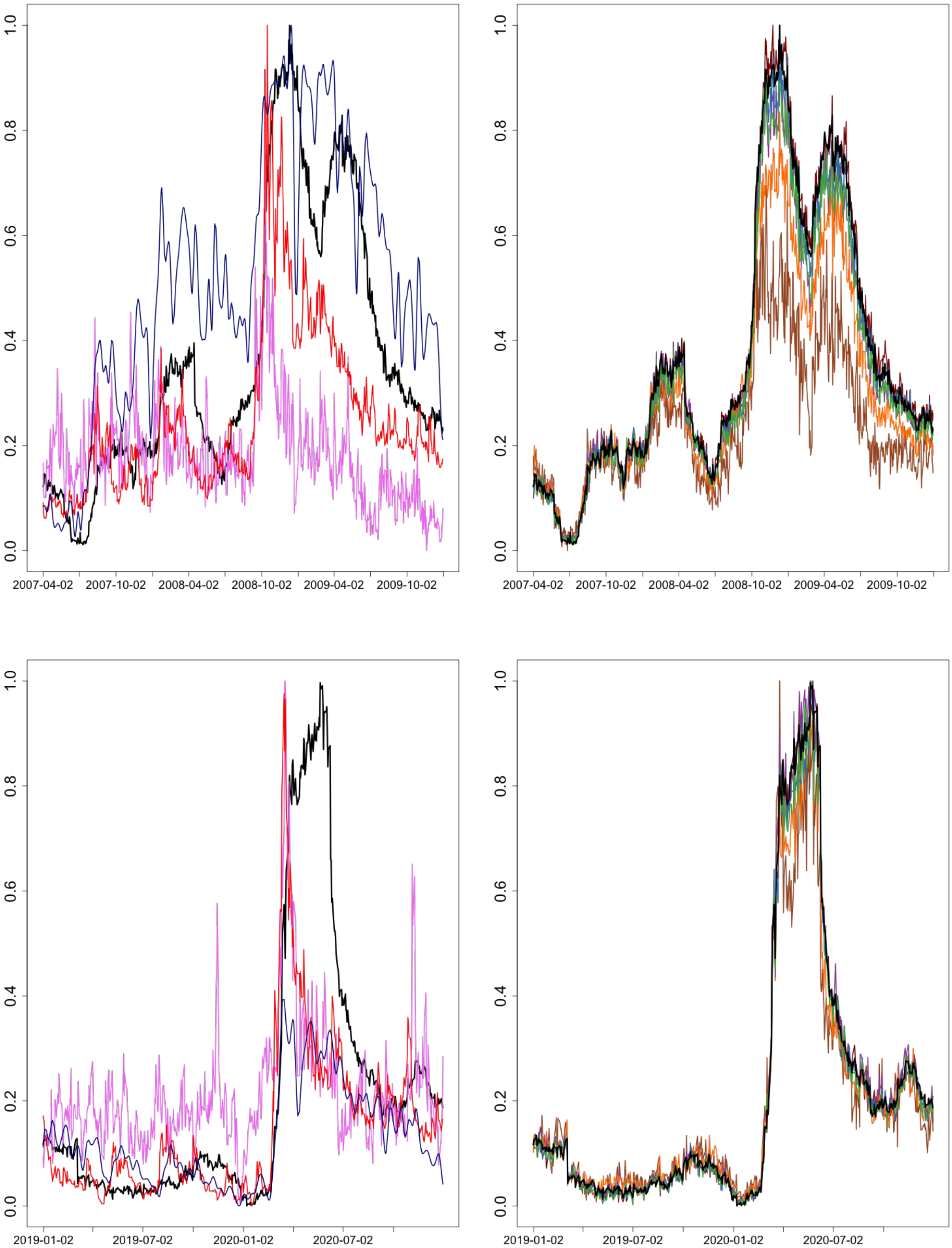

Figure 4.2: Normalized systemic risk measures VSTOXX, CISS, FT, and standard FRM, FRM_{q50}, FRM_{q60}, FRM_{q70}, FRM_{q80}, FRM_{q90}, FRM_{IQR} for Europe during the financial crisis of 2007 and the COVID-19 pandemic

reach such a high normalized systemic risk level as standard $FRM@Europe$, $FRM_{q50}@Europe$, $FRM_{q60}@Europe$, $FRM_{q70}@Europe$ and $FRM_{q80}@Europe$. The overall characteristics of normalized $FRM@Europe$ candidates are similar to normalized $FRM@Americas$ candidates. However, during the financial crisis of 2007, levels of $FRM_{q90}@Europe$ and $FRM_{IQR}@Europe$ were smaller than the respective $FRM@Americas$ candidates. In contrast, normalized $FRM_{q90}@Europe$ and $FRM_{IQR}@Europe$ during the COVID-19 pandemic were higher than both during the financial crisis of 2007. $FRM_{IQR}@Europe$ had a dramatic increase at the beginning of the COVID-19 pandemic. It proved that the top 10 financial institutions might suffer more because they were already under high distress before this pandemic and were not well prepared for this unexpected outbreak of virus. The $FRM_{IQR}@Europe$ series indicates that the difference between the systemic risk levels of the top 25 risky and safe financial institutions was large during the pandemic. This implies that risk top 25 safe financial institutions were not suffered as much has during the financial crisis after taking the considerable $FRM_{q70}@Europe$ into account. In general, the impact of the COVID-19 pandemic in European countries was the same as the impact of the financial crisis of 2007.

From the above visualizations of normalized FRM candidates in the Americas and Europe, we have observed their similarity. However, it is impossible to compare FRM candidates directly from the plot. Therefore, in the next chapters, several statistical tests will be conducted to compare FRM candidates with these commonly used systemic risk measures numerically.

4.2 Comovement of FRM candidates and other systemic risk measures

Firstly, to check the relationship of FRM candidates and above mentioned systemic risk measures, we will perform a correlation test following Yu et al. (2017). Secondly, we extend the Kolmogorov-Smirnov(KS) statistic to compare the goodness-of-fit between the empirical series of one FRM candidate and the empirical series of one systemic risk measure.

According to Lee Rodgers and Nicewander (1988), the Pearson's correlation coefficient is still most commonly used to measure correlation. The correlation coefficient developed by Pearson in 1895 is as follows:

$$r = \frac{\sum_{t=1}^{T}\left(FRM_{p,t} - \overline{FRM_{p,t}}\right)\left(S_{i,t} - \overline{S_{i,t}}\right)}{\sqrt{\sum_{t=1}^{T}\left(FRM_{p,t} - \overline{FRM_{p,t}}\right)^2 \sum_{t=1}^{T}\left(S_{i,t} - \overline{S_{i,t}}\right)^2}}, \tag{4.3}$$

where $\overline{FRM_{p,t}} = \frac{1}{t}\sum_{t=1}^{T} FRM_{p,t}$ is the average of individual FRM candidates over two periods and $\overline{S_{i,t}} = \frac{1}{t}\sum_{t=1}^{T} S_{i,t}$ is the average of individual series of commonly used systemic

	VIX	SRISK	FT	GT
FRM_{mean}	0.81	0.67	0.37	0.32
FRM_{q50}	0.81	0.67	0.39	0.34
FRM_{q60}	0.81	0.67	0.38	0.34
FRM_{q70}	0.81	0.67	0.39	0.33
FRM_{q80}	0.81	0.67	0.38	0.33
FRM_{q90}	0.80	0.66	0.37	0.32
FRM_{IQR}	0.79	0.66	0.39	0.33

Table 4.1: Correlation of $FRM@Americas$ candidates and other systemic risk measures in Americas.

FRMQcorrelationAM

risk measures over the corresponding periods. The i in $S_{i,t}$ ranges from 1 to 4 for Americans and varies from 1 to 3 for Europe, representing other systemic risk measures.

Tables 4.1 and 4.2 show that FRM candidates are strongly correlated with other systemic risk measures. The standard FRM, FRM_{q50}, FRM_{q60}, FRM_{q70} and FRM_{q80} for Americas and Europe are slightly more correlated with VIX and VSTOXX than FRM_{q90} and FRM_{IQR}. $FRM@Americas$ and $FRM@Europe$ are almost same correlated to VIX, VSTOXX and CISS, and the correlation is positive and extremenly strong. Although in Figure 4.1, the SRISK is largely different from VIX and $FRM@Americas$ in terms of the general trend and the performance under high distress. The correlations of SRISK and individual $FRM@Americas$ candidates are strong. In contrast, FT and GT fluctuated similarly to FRM candidates, but the statistical correlations of FT and FRM candidates, as well as that of GT and FRM candidates, are smaller due to the high volatility of FT and GT.

In summary, theses FRM candidates are not significantly distinct in terms of the correlation test. After applying Pearson correlation testing, the null hypothesis that the correlation of FRM candidates and one other systemic risk indicators are equal to 0 should be rejected. For $FRM@Americas$ candidates, it is difficult to distinguish their performance since the correlation coefficients are almost the same. Among $FRM@Europe$ candidates, except for $FRM_{IQR}@Europe$, the performance of other $FRM@Europe$ candidates are similar.

Furthermore, the Kolmogorov-Smirnov (KS) test is used to test the goodness-of-fit between the empirical distributions of one FRM candidate and one series of other systemic risk measures.

	VSTOXX	CISS	FT
FRM_{mean}	0.81	0.75	0.39
FRM_{q50}	0.81	0.76	0.28
FRM_{q60}	0.81	0.75	0.29
FRM_{q70}	0.81	0.73	0.31
FRM_{q80}	0.81	0.72	0.33
FRM_{q90}	0.80	0.70	0.34
FRM_{IQR}	0.76	0.57	0.43

Table 4.2: Correlation of $FRM@Europe$ candidates and systemic risk measures in Europe. FRMQcorrelationEU

Following Simard et al. (2011), the KS two-sided test statistic is defined as

$$D_t = \sup_{FRM_p, S_i} |F_t(FRM_p), G_t(S_i)|, \tag{4.4}$$

where F_t and G_t are empirical cumulative distribution functions of FRM candidates and other systemic risk measures. This KS statistic quantifies a distance between the empirical distribution of one FRM candidate and the empirical distribution of one comparable systemic risk measure. The KS statistics in Table 4.3 indicate that VIX fits $FRM@Americas$ candidates at most in comparison of other systemic risk measures. Moreover, $FRM_{q90}@Americas$ and $FRM_{IQR}@Americas$ fit the empirical distribution of VIX, GT and FT better than other FRM candidates. Unlike the results of correlation statistics, the lower KS statistics of GT and FT show that their distributions are more close to the distributions of $FRM@Americas$ candidates than SRISK. However, for all $FRM@Americas$ candidates, the KS test with the null hypothesis that one $FRM@Americas$ candidate fit one systemic risk measure should be rejected, i.e. no FRM candidates could fit any above mentioned systemic risk measure well.

Figure 4.4 presents the KS statistics of $FRM@Europe$ and systemic risk measures in European countries. Analogue to the results of $FRM@Americas$, the standard systemic risk measure in Europe VSTOXX is closest to $FRM@Europe$ candidates. The KS test with the null hypothesis that $FRM_{IQR}@Europe$ fit the series of VSTOXX cannot be rejected, i.e. the distribution of $FRM_{IQR}@Europe$ fits the VSTOXX distribution well. Moreover, $FRM_{q50}@Europe$ fits the empirical distribution of CISS best while $FRM_{IQR}@Europe$ fits the empirical distribution of FT best. $FRM_{IQR}@Europe$ fits the worst in comparison to other $FRM@Europe$ candidates. This finding is consistent with the lowest correlation coefficients

	VIX	SRISK	GT	FT
FRM_{mean}	0.14	0.40	0.31	0.31
FRM_{q50}	0.14	0.39	0.33	0.33
FRM_{q60}	0.16	0.41	0.31	0.31
FRM_{q70}	0.14	0.41	0.32	0.32
FRM_{q80}	0.15	0.41	0.31	0.31
FRM_{q90}	0.10	0.44	0.24	0.24
FRM_{IQR}	0.10	0.42	0.25	0.25

Table 4.3: KS test statistics D_t of $FRM@Americas$ candidates and other systemic risk measures in Americas

FRMQcorrelationAM

of $FRM@Europe$ candidates and CISS in Table 4.2.

Although most FRM candidates cannot pass the KS test, we can use KS test statistics to compare their performance in terms of goodness-of-fit. As a result, we find that $FRM_{q90}@Americas$ and $FRM_{IQR}@Americas$ fit the systemic risk measures in Americas best, and $FRM_{IQR}@Europe$ fits the systemic risk measures in Europe optimally. However, their advantages over other FRM candidates in the KS test are not significant. Moreover, the FRM_{IQR} is not as correlated as other candidates with VIX, VSTOXX and CISS. Hence, we cannot conclude that the comovement of FRM_{IQR} and other systemic risk measures is stronger than the comovements of other FRM candidates and other systemic risk measures.

4.3 Causality of FRM candidates and other systemic risk measures

Yu et al. (2017) introduced a Granger causality test to further analyze the relationship between standard $FRM@Americas$ and other systemic risk measures. This methodology is extended hereafter to check the Granger causality of all FRM candidates for Americas and Europe and systemic risk measures.

The test procedure is presented in Figure 4.3. After testing the stationarity of data in Americas using Augmented Dickey-Fuller (ADF) test, we find only for GT and FT series, we cannot reject the null hypothesis that there is a unit root in the time series because its p-value is smaller than 0.05. In other word, the GT and FT series are stationary while all FRM candidates, VIX and SRISK series are non-stationary. Hence, we do not need to consider the

	VSTOXX	CISS	FT
FRM_{mean}	0.18	0.19	0.25
FRM_{q50}	0.19	0.18	0.26
FRM_{q60}	0.17	0.19	0.23
FRM_{q70}	0.15	0.21	0.22
FRM_{q80}	0.18	0.19	0.25
FRM_{q90}	0.11	0.24	0.18
FRM_{IQR}	0.04	0.28	0.14

Table 4.4: KS test statistics D_t of $FRM@Europe$ candidates and other systemic risk measures in Europe.

FRMQcorrelationEU

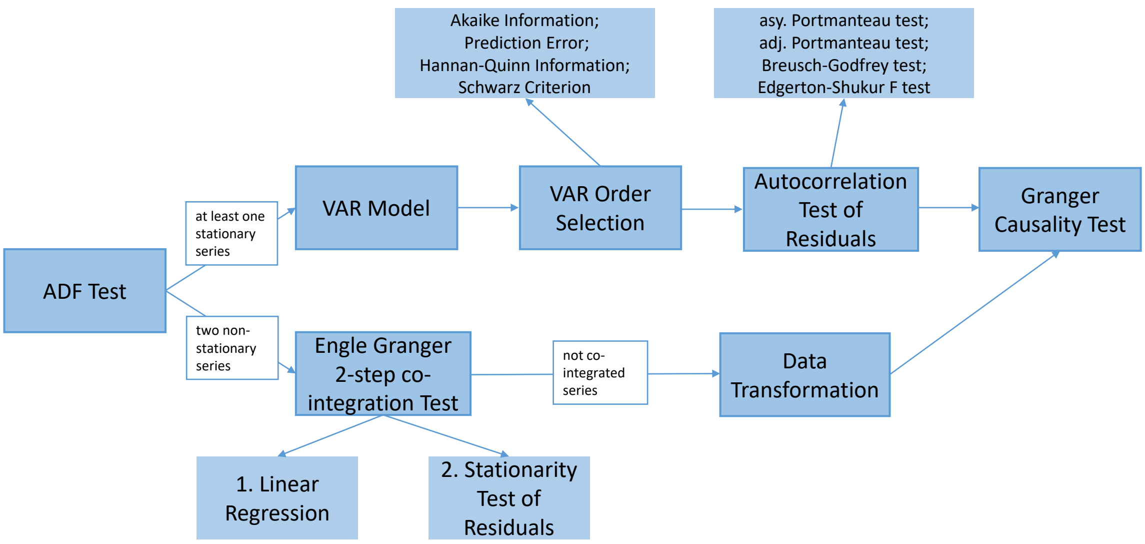

Figure 4.3: Granger causality test procedure.

co-integration problem for the Granger causality test of FRM candidates and GT or FT. The Vector Autoregression model (VAR) proposed by Lütkepohl (2005), see Equation 4.5, is fair to test series that are not co-integrated.

$$SYS_t = \upsilon + B_1 \times SYS_{t-1} + B_2 \times SYS_{t-2} + \ldots + B_p \times SYS_{t-p} + \zeta_t \tag{4.5}$$

$SYS_t \overset{def}{=} (FRM_t, S_t)$ is split into two subvectors: FRM_t represents one specific FRM candidate and S_t indicats one of stationary systemic risk measures. Then the VAR(P) model is rewritten as:

$$SYS_t = \begin{pmatrix} FRM_t \\ S_t \end{pmatrix} = \begin{pmatrix} \upsilon_1 \\ \upsilon_1 \end{pmatrix} + \begin{pmatrix} B_{11,1} & B_{12,1} \\ B_{21,1} & B_{22,1} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ B_{2,t-1} \end{pmatrix} + \ldots \tag{4.6}$$

$$+ \begin{pmatrix} B_{11,P} & B_{12,P} \\ B_{21,P} & B_{22,P} \end{pmatrix} \begin{pmatrix} y_{1,t-P} \\ B_{2,t-P} \end{pmatrix} + \begin{pmatrix} \zeta_{1t} \\ \zeta_{2t} \end{pmatrix}. \tag{4.7}$$

The null hypothesis of Granger causality test is that the subvector FRM_t does not Granger-cause the subvector S_t, i.e. $B_{21,i} = 0$ for $i = 1, 2, \ldots, P$.

Four criteria: the Akaike Information Criterion (AIC), the Prediction Error Criterion (FPE), Hannan-Quinn information criterion (HQ) and Schwarz criterion (SC) are used to propose the smallest possible VAR order so that $B_p \neq 0$ and $B_i = 0$ for $i > p$ by Lütkepohl (2005). Next, we check the autocorrelation of the residuals to determine the optimal order P among these proposed orders using four tests: the asymptotic Portmanteau test (PT(asymptotic)), the adjusted Portmanteau test (PT(adjusted)), the Breusch-Godfrey LM test (BG) and the Edgerton-Shukur F test (ES). The null hypothesis of these four tests is that there is no first-order autocorrelation among residuals. We check the proposed orders in ascending order until the order having at least one p-value larger than 0.05 occurs. If all orders proposed by four criteria lead to the rejection of all tests, we will try other orders subsequently. If no tests are passed after trying all orders up to 20, the order whose p-value of autocorrelation test is most close to the critical value 0.05 is selected. Table A.1 indicates that there is always the first-order autocorrelation among residuals of FRM candidates and GT series based on four tests. The VAR(P) model is further used to test the Granger causality with the order P listed in Table A.2. Since the t ranges from 1 to T and the two subvector FRM_t, S_t both have a dimension (1×1). The test statistic follows F(P, $2 \cdot (T-P) \cdot 2(2P+2)$) distribution. Two is the number of dimensions for SYS_t. (T-P) and 2(2P+2) are the sample size and the total number of the parameters mentioned above. Taking 5% as the critical value, the Table 4.5 shows that standard $FRM@Americas$, $FRM_{q80}@Americas$ and $FRM_{IQR}@Americas$

Granger cause GT while GT Granger causes all $FRM@Americas$ candidates. The p-value of $FRM_{q70}@Americas$ Granger causes GT at is at 9% shows there is a Granger causality between $FRM_{q70}@Americas$ and GT if the critical value is set as 10%.

FRM	p-values (FRM Granger causes GT)	p-values (GT Granger causes FRM)
FRM_{mean}	0.03	4.88×10^{-08}
FRM_{q50}	0.80	2.48×10^{-09}
FRM_{q60}	0.62	2.32×10^{-08}
FRM_{q70}	0.09	4.14×10^{-07}
FRM_{q80}	2.35×10^{-03}	1.91×10^{-07}
FRM_{q90}	0.36	5.67×10^{-09}
FRM_{IQR}	8.95×10^{-03}	2.51×10^{-03}

Table 4.5: p-values of Granger causality test of $FRM@Americas$ candidates and GT.

FRMQcausalityAM

The above procedure is repeated to test the Granger causality between $FRM@Americas$ candidates and FT. Table A.2 shows the optimal order P that is selected by four autocorrelation tests. The autocorrelation test shows that all model has the first order autocorrelated residuals. The p-values of the Granger causality test in the context of VAR(P) model are available in Table 4.6. $FRM_{q50}@Americas$, $FRM_{q60}@Americas$, $FRM_{q70}@Americas$ Granger cause FT, while FT Granger cause all $FRM@Americas$ candidates. The p-value of the hypothesis test that standard $FRM@Americas$ Granger causes FT is 0.07, i.e. standard $FRM@Americas$ can almost Granger cause FT.

From Granger (1988) we know that there must be at least one way Granger causality caused by one series if two series are co-integrated. For co-integrated series, the aforementioned VAR(P) testing cannot detect the underlying causality. Following Yu et al. (2017), we perform the Engle Granger 2-step test for co-integrated FRM candidates and VIX, SRISK series. The linear regression of one FRM candidate on VIX or SRISK is carried out in the first step. Secondly, the stationary of residuals of the above linear regression is tested to check whether this FRM candidate can Granger cause VIX or SRISK. For stationary residuals, there is a co-integration between this FRM and VIX or SRISK. For two co-integrated series, there is always at least one way Granger causality. If we regression VIX or SRISK on this FRM candidate in the first step, we can test whether VIX or SRISK are co-integrated, i.e. we can

FRM	p-values (FRM Granger causes FT)	p-values (FT Granger causes FRM)
FRM_{mean}	0.07	2.67×10^{-15}
FRM_{q50}	5.78×10^{-04}	$< 2.20 \times 10^{-16}$
FRM_{q60}	2.76×10^{-03}	8.88×10^{-16}
FRM_{q70}	5.33×10^{-03}	$< 2.20 \times 10^{-16}$
FRM_{q80}	0.44	1.34×10^{-10}
FRM_{q90}	0.37	1.58×10^{-05}
FRM_{IQR}	0.32	1.50×10^{-11}

Table 4.6: p-values of Granger causality test of $FRM@Americas$ candidates and FT.

FRMQcausalityAM

test the existence of Granger causality in another direction in the second step. The Granger causality test between two non-stationary series is conducted by checking the co-integration in both directions.

After applying this method to all $FRM@Americas$ candidates and VIX series, we obtain summary table 4.7. Based on this table, we conclude that there are mutual causalities between all FRM candidates and VIX since there are bilateral co-integration between FRM candidates and VIX. The co-integration is found because the absolute values of test-statistics for all $FRM@Americas$ candidates and VIX are larger than the absolute value of critical value at 5%.

The Engle Granger 2-step test results for co-integration of FRM candidates and SRISK are summarized in Table 4.8. We conclude that $FRM_{q90}@Americas$ and SRISK as well as SRISK and $FRM_{q90}@Americas$ are co-integrated, i.e. they have a mutual Granger causality. Analogously, $FRM_{IQR}@Americas$ and SRISK as well as SRISK and $FRM_{IQR}@Americas$ are co-integrated, i.e. they have a mutual Granger causality. Other FRM candidates are not co-integrated with SRISK. However, Ghosh (2002) argued that no co-integration does not imply that the Granger causality does not exist. Then, we do the Granger causality test on transformed data. The transformed $\Delta SRISK$ and ΔFRM_p with $p = \{mean, q50, q60, q70, q80\}$, where the Δ is the first difference order as well as the growth rate of the SRISK and FRM, is denoted as DSRISK and $DFRM_p$. They can be modelled as an unrestricted VAR(P) since they are not co-integrated according to Ghosh (2002). Moreover, Yule and Granger argued that the original non-stationary data might lead to untrusted estimation. Table A.3 lists the

FRM	Test-statistic (FRM Granger causes VIX)	Test-statistic (VIX Granger causes FRM)	Critical value at 5%
FRM_{mean}	-5.16	-4.27	-1.95
FRM_{q50}	-5.35	-4.54	-1.95
FRM_{q60}	-5.49	-4.72	-1.95
FRM_{q70}	-5.30	-4.50	-1.95
FRM_{q80}	-5.22	-4.41	-1.95
FRM_{q90}	-5.33	-4.68	-1.95
FRM_{IQR}	-6.04	-5.94	-1.95

Table 4.7: Results of Engle Granger 2-step test of $FRM@Americas$ candidates and VIX. FRMQcausalityAM

optimal order for transformed series, which are further applied to the VAR(P) model. The p-values in the Table A.4 show that standard $DFRM@Americas$, $DFRM_{q50}@Americas$, $DFRM_{q60}@Americas$, $DFRM_{q70}@Americas$ and $DFRM_{q80}@Americas$ do not Granger-cause $DSRISK$, since the null hypothesis that these transformed FRM do not Granger-cause the transformed SRISK cannot be rejected and the null hypothesis in the reverse direction cannot be rejected. The finding that is inconsistent with the results in Yu et al. (2017) may due to either the limited sample size in this analysis or the large deviation of FRM and SRISK during the COVID-19 pandemic.

Analogously, the causality of individual $FRM@Europe$ candidates and systemic risk measures in Europe is checked following the the same process. Except for FT, all FRM candidates and the systemic risk measures in Europe are non-stationary. The causality test procedure for $FRM@Europe$ candidates and FT for Europe is the same as the procedure for each $FRM@Americas$ candidates and FT for Americas. For each VAR(P) model composed of one $FRM@Europe$ candidate, the optimal VAR order in the Table A.5 is selected among four orders proposed by four criteria since at least one autocorrelation test is passed. The Table 4.9 shows mutual Granger causality between each $FRM@Europe$ candidate and FT for Europe. Since the p-values of $FRM@Europe$ Granger causes FT for Europe and the p-values of FT for Europe Granger causes $FRM@Europe$ are smaller than the critical value 5%.

Furthermore, the Engle Granger 2-step tests are applied to $FRM@Europe$ candidates and VSTOXX as well as $FRM@Europe$ candidates and CISS. The results of co-integration test are

FRM	Test-statistic (FRM Granger causes SRISK)	Test-statistic (SRISK Granger causes FRM)	Critical value at 5%
FRM_{mean}	-1.70	-1.46	-1.95
FRM_{q50}	-1.76	-1.67	-1.95
FRM_{q60}	-1.82	-1.81	-1.95
FRM_{q70}	-1.76	-2.71	-1.95
FRM_{q80}	-1.77	-1.74	-1.95
FRM_{q90}	-2.02	-2.29	-1.95
FRM_{IQR}	-2.75	-3.73	-1.95

Table 4.8: Results of Engle Granger 2-step test of $FRM@Americas$ candidates and SRISK.
FRMQcausalityAM

available in Table 4.10 and 4.11. We conclude that all FRM candidates are co-integrated with VSTOXX and CISS. Moreover, VSTOXX and CISS are co-integrated with all FRM candidates. Thus, we conclude that there are mutual Granger causality between FRM candidates and VSTOXX as well as CISS.

As a summary, the result of the Granger causality test of FRM candidates and systemic risk measures is as follows:

1. All $FRM@Americas$ candidates mutually Granger-cause VIX;
2. standard $FRM@Americas$, $FRM_{q80}@Americas$ and $FRM_{IQR}@Americas$ mutually Granger-cause GT, and GT Granger-cause all $FRM@Americas$ candidates;
3. $FRM_{q50}@Americas$ $FRM_{q60}@Americas$ $FRM_{q70}@Americas$ mutually Granger-cause FT, and FT Granger-cause all $FRM@Americas$ candidates;
4. $FRM_{q90}@Americas$ and $FRM_{IQR}@Americas$ mutually Granger-cause SRISK;
5. All $FRM@Europe$ candidates mutually Granger-cause VSTOXX, CISS and FT.

FRM	p-values (FRM Granger causes FT)	p-values (FT Granger causes FRM)
FRM_{mean}	2.54×10^{-04}	2.06×10^{-12}
FRM_{q50}	3.31×10^{-06}	$3,64 \times 10^{-08}$
FRM_{q60}	4.03×10^{-06}	4.75×10^{-11}
FRM_{q70}	3.37×10^{-07}	2.99×10^{-06}
FRM_{q80}	4.91×10^{-05}	3.92×10^{-07}
FRM_{q90}	8.52×10^{-03}	5.54×10^{-08}
FRM_{IQR}	1.01×10^{-10}	9.00×10^{-07}

Table 4.9: p-values of Granger causality test of $FRM@Europe$ candidates and FT.

FRMQcausalityEU

FRM	Test-statistic (FRM Granger causes VSTOXX)	Test-statistic (VSTOXX Granger causes FRM)	Critical value at 5%
FRM_{mean}	-6.02	-5.02	-1.95
FRM_{q50}	-6.14	-5.18	-1.95
FRM_{q60}	-6.03	-5.09	-1.95
FRM_{q70}	-6.07	-5.15	-1.95
FRM_{q80}	-5.96	-5.00	-1.95
FRM_{q90}	-6.03	-5.25	-1.95
FRM_{IQR}	-2.82	-6.15	-1.95

Table 4.10: Results of Engle Granger 2-step co-intergration test of $FRM@Europe$ candidates and VSTOXX.

FRMQcausalityEU

FRM	Test-statistic (FRM Granger causes CISS)	Test-statistic (CISS Granger causes FRM)	Critical value at 5%
FRM_{mean}	-2.81	-2.40	-1.95
FRM_{q50}	-2.91	-2.63	-1.95
FRM_{q60}	-2.69	-2.55	-1.95
FRM_{q70}	-2.66	-2.59	-1.95
FRM_{q80}	-2.46	-2.47	-1.95
FRM_{q90}	-2.63	-2.92	-1.95
FRM_{IQR}	-2.82	-4.27	-1.95

Table 4.11: Results of Engle Granger 2-step test of $FRM@Europe$ candidates and CISS.

FRMQcausalityEU

5 FRM as recession predictors

Comparing with other systemic risk measures in terms of comovement and causality cannot distinguish FRM candidates. In this chapter, we would like to transform FRM candidates into recession predictors. The major reason is that the occur of severe systemic risk commonly coincident with a great economic recession. Secondly, forecasting economy activity or the likelihood of recession based on indicators like the interest rate, stock price indices and the term structure of yield curve has already gained success. More details see Estrella and Hardouvelis (1991), Estrella and Mishkin (1996), Chauvet (1998) and Chauvet and Potter (2005). Although the yield curve is defined specifically as the spread between the interest rates on the ten-year treasury note and the three-month treasury bill by Estrella and Mishkin (1996), several factors that may be included in the estimation of penalization parameters and the calculation of FRM candidates are similar to the yield curve. As mentioned in Table 3.1, the three months Treasury Constant Maturity Rate to ten years Treasury Constant Maturity Rate Spread Differences and the Moodys Seasoned Baa Corp Bond Yield Spread to ten years Treasury Constant Maturity Rate Spread Differences may be considered into the estimation of FRM@Americas candidates. Analogously, the German Treasury Constant Maturity Rate one to ten years Slope Spread Differences and the Barclays Bloomberg EuroAgg Corporate Yield Spread to ten year German Treasury Constant Maturity Rate Spread Differences may be taken as covariates to estimate FRM@Europe candidates. These above mentioned factors can be decomposed into expected real interest rate and expected inflation, each of which may be associated with the expectation of future monetary police and future economic growth as stated by Estrella and Mishkin (1996). Besides, stock prices of essential financial institutions are determined by expectations of future dividend streams, which may also reflect future economic and financial state development. FRM candidates for Americas and Europe may be taken as a measure to reflect the expectation of future growth since the their calculation contains a large proportion of market capitalization and the evolution of stock prices of these critical financial institutions. Thirdly, by (Mihoci et al., 2020) applied standard FRM to predict recession probability. They found that both standard $FRM@Americas$ and $FRM@Europe$ are able to predict the recessions derived from the financial crisis of 2007 and the Euro Area debt crisis of 2011. This outcome motivates us to compare predefined FRM candidates based upon their forecast abilities. The optimal FRM candidate in one financial market is identified as the one that could forecast the recession probability more precise than the standard FRM, i.e. could capture the recession earlier than the standard FRM.

5.1 Recession prediction models

In this chapter, we will introduce several regression models to predict the recession probabilities. All possible models will be elaborated and be empirically compared in terms of statistical inference and prediction results in chapters 5.2 and 5.3.

The binary regression models are utilized to examine whether these FRM candidates could be taken as predictors for recession. Because Estrella and Hardouvelis (1991), Estrella and Mishkin (1996) and Chauvet and Potter (2005) argued that models predicting a binary indicator of recession or expansion are more successful and stable over time than basic linear regression. Here we propose recession models based on monthly FRM candidates and recession indicators in the Americas and in Europe. The probability of observing recession at month t $\pi_{p,t}$ is estimated based on FRM candidates obtained k months ago:

$$\pi_{p,t} = \mathrm{P}\left(Y_t = 1\right) = h\left(\eta_p\right) = h\left(\omega_0 + \sum_{k=1}^{6} \omega_k \cdot FRM_{p,u-k}\right), \tag{5.1}$$

where the response function h is a strictly monotonically increasing cumulative distribution function on the real line as claimed by Fahrmeir et al. (2007). $h\left(\eta_p\right)$ is ensured being bounded within $[0, 1]$, and the Equation 5.1 can always be expressed as $\eta_p = g\left(\pi_p\right)$ with the link function $g = h^{-1}$. Mihoci et al. (2020) used logistic regression, while Estrella and Hardouvelis (1991), Chauvet and Potter (2005) used probit model to predict the upcoming recession. In this analysis, both logit and probit models, which known as most widely used binary regression models, will be applied. The logistic model is constructed with the logistic response function $h(\eta_p) = \frac{\exp(\eta_p)}{1+\exp(\eta_p)}$, or equivalently, the logit link function $g(\pi_p) = \log\left(\frac{\pi_p}{1-\pi_p}\right) = \eta_p$. In other words, logistic regression model yields a linear relationship between the logarithmic odds and predictors lagged FRM indices. The response function for probit model is the standard normal cumulative distribution function Φ, i.e. $\pi_p = \Phi(\eta_p)$. In chapter 4.2, the visualization of FRM candidates and the correlation coefficients indicate that FRM candidates are highly correlated other each other. Considering the multicollinearity problem that may be resulted from these high correlated FRM candidates, only one FRM candidate will be included in each regression. FRM@Americas and FRM@Europe could not be used in mixed form because we are interested in the analysis of separate regions.

In the USA, the peak and trough dates reported by the National Bureau of Economic Research (NBER) are most widely utilized in the recession analysis. In European countries, the recession indicator reported by Center for Economic and Policy Research (CEPR) is commonly used. The series of NBER monthly recession indicators for the United States are downloaded from FederalReserveBankofSt.Louis (2020) while CEPR monthly recession indicators for the Europe are collected from EuroAreaBusinessCycleNetwork (2020). All observations are binary, with a value at one standing for recession periods and a value at zero representing the expansion periods for both datasets. All recession indicators are collected and updated until January 15th 2021. In the Americas, from January 2008 to June 2009 and from March 2020 until December 2020 are recession periods. In Europe, the recession periods contain the time intervals from April 2008 to June 2009 and from January 2020 to June 2020 since the CEPR has not reported indicators after the second quarter of 2020. Figure 5.1 visualizes the distributions of standard FRM in both markets during recessions and expansions. These plots support the idea that a high value of FRM or high systemic risk level is associated with the recession, regardless of the individual characteristics of the systemic risk and the regions. The standard $FRM@Amercias$ and $FRM@Europe$ during expansion were concentrated lower than 0.05. However, during the recession periods, the standard $FRM@Americas$ and $FRM@Europe$ had another peak around 0.08, and the proportion of FRM, which was higher than 0.05, was significantly large. From the comparison of the recession due to the financial crisis of 2007 and the recession due to the COVID-19 pandemic, we conclude that the later one may not be as severe as the former one, but the later one may last for a longer time and the recession is still on-going. From the comparison between the two regions, we find out that countries in the Americas suffered more from systemic risks or recessions.

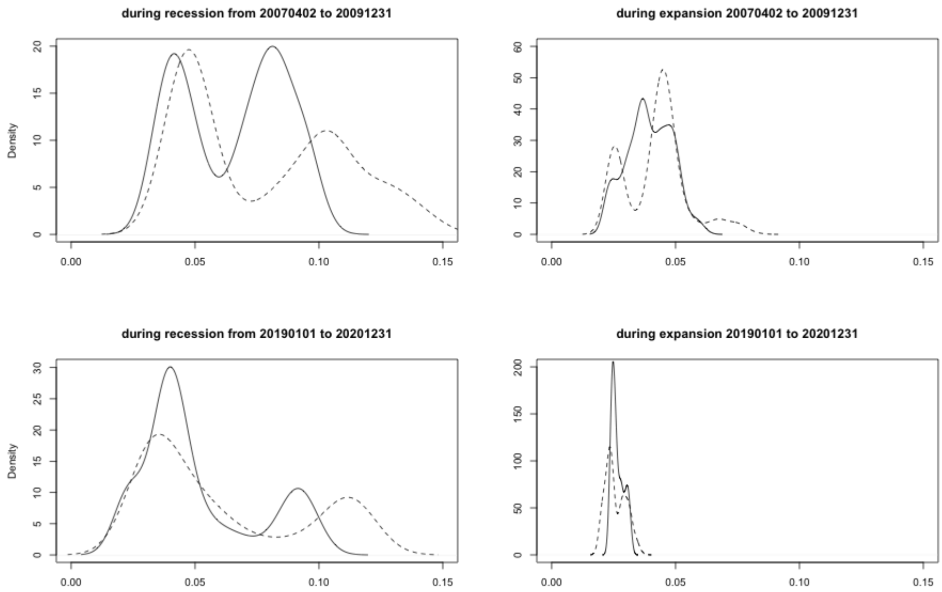

Figure 5.1: Standard FRM@Americas (- - -) and FRM@Europe (——) during recessions and expansions

FRMQRecDensity

Regarding the FRM candidates data, we obtain $FRM_{p,u}$ with u indicating the index of the month after averaging daily FRMs in a month. Lagged monthly FRM candidates $FRM_{p,u-k}$ correspond to FRM_p with $p \in \{mean, 50, 60, 70, 80, 90, IQR\}$ at $k = (1, \ldots, 6)$ months ago. For example, the predictor obtained one month ago is denoted as $FRM_{p,u-1}$. Most research has forecast the upcoming recession at the quarterly frequency rather than monthly. However, Mihoci et al. (2020) argued that the capability of standard FRM to predict the upcoming recession in two quarters is already low, so k quoted monthly will be more suitable for FRM candidates. Hence, six lagged sequences are generated for each FRM candidate as possible predictors to forecast the upcoming recession. In other words, historical FRM data in the last months may be applied to predict the probability of observing the recession in the current month. After joining FRM candidates with available recession indicators, we attain empirical data for the Americas and Europe over 33 months around the financial crisis of 2007. Moreover, data for the Americas of 24 months and Europe of 18 months are employed in the analysis of the COVID-19 pandemic.

We have also tried to replace monthly average FRM candidates with daily FRM candidates to have a larger size of observations for each model. The FRM value at k months ago is replaced with daily FRM obtained at $21 \times k$ trading days ago because, the average number of trading days in each month is 21 days. Then $FRM_{p,u-k}$ will be denoted as $FRM_{p,d-21\times k}$ with d indicating the index of trading days. In this sense, each recession prediction model is used to predict the recession for today. Thus the size of observations will be transformed from the number of months to the number of trading days so that the effect of FRM candidates are more precisely estimated. On the other hand, the confidence interval is getting narrower, and it is more likely to reject the null hypothesis that the $\omega_k = 0$, so FRM candidates are more likely to be statistically significant. However, the daily FRM data as predictors will reflect the daily evolution of the recession indicator directly. Unlike the stock price, macroeconomic factors and FRM candidates, the recession indicators are not so volatile and tend to remain unchanged at least for several months. In this regard, using monthly data is helpful to smooth the fluctuation and evolution of the time series of daily FRM data so that the underlying relationship of FRM candidates and recession indicators will be extracted.

The Equation 5.1 including six FRM predictors from one to six months ago, is the saturated model for each FRM candidate. On the basis of the saturated model, we will select the most essential lagged FRM candidates to predict the upcoming recession. Besides AIC, BIC, deviance, the pseudo coefficient of determination R^2 are used to measure the goodness-of-fit of possible models. In this analysis, Estrella pseudo R^2 developed by Estrella and Mishkin (1998) is applied, which is widely used to compare recession prediction models, see Bernard and Gerlach (1998) and Nyberg (2010). This measure is calculated as $1 - \left(\frac{\log L_u}{\log L_c}\right)^{\left(\frac{2}{n}\right)\log L_c}$ with $\log L_u$ denoting the unrestricted maximum value of log-likelihood and $\log L_c$ denoting the maximum value under the constraint that all coefficients are zero except the intercept. Similar to R^2 in linear case, the pseudo R^2 ensures that the values 0 and 1 correspond to "not fit" and "perfect fit". The intermediate values have similar interpretations, namely the proportion of variance in the recession indicators that could be forecasted from one FRM candidate. The procedure of model selections and the estimation results will be explained and interpreted in the following two chapters 5.2 and 5.3. Based on the selected recession prediction model for each FRM candidate, we check the in-sample results over the entire time frame. The in-sample results are called implied recession probabilities in this paper, which reports the probability of recession attributed to the quantified systemic risk measures at 5% level of each FRM candidate. These in-sample results should be plausible and robust in general since empirical data applied in this paper consists of two recessions periods in two large markets. Out-of-sample validation is performed as an addition once the in-sample results seem to be implausible or problematic.

5.2 Uniperiod recession Prediction Models

In this chapter, the most commonly used standard static recession prediction model proposed by Estrella and Hardouvelis (1991) and Estrella and Mishkin (1996) is applied. It means that only one predictor, the FRM candidate with the lag at k, is kept in Equation 5.2:

$$\pi_{p,t} = \mathrm{P}\left(Y_t = 1\right) = h\left(\eta_p\right) = h\left(\omega_0 + \omega_{FRM_{p,k}} \cdot FRM_{p,u-k}\right).$$

After comparing the AIC, BIC, deviances, pseudo R^2 and the implied recession probabilities, we find there is almost no difference between probit and logistic regression. The probit regression results will be illustrated further since some predictors are statistically insignificant in logistic regression but significant in probit regression.

For each FRM candidate, the forecast horizon k varies from one to six. Tables A.6 and A.8 are the summaries of six uniperiod models for seven $FRM@Americas$ and $FRM@Europe$ candidates respectively. The first lines for each forecast horizon include the estimated coefficients $\widehat{\omega}_k$. Each second lines contain the standard deviation of the corresponding coefficient. Values in each third lines represent the pseudo R^2. We could see that as the forecast horizon increases, the effect of all FRM candidates on predicting the upcoming recession decreases. At the same time, the statistical significance of all FRM candidates decays along with the increase of forecast horizon. FRM candidates obtained five or six months ago are not statistically significant to predict the recession. For both markets, the prediction models with the shortest forecast horizon performed best in terms of statistical significance and the coefficient of determination pseudo R^2. Comparing recession probabilities implied from individual $FRM@Americas$ candidates with $k = 1$ in Figure 5.2 with recession probabilities implied from corresponding individual candidates with $k = 2, 3, 4$ in Figure B.1, we observe that the uniperiod model with a predictor lagged at higher-order has a longer delay in predicting a recession. Figure B.3 shows the same conclusion for $FRM@Europe$ candidates. Hence, we conclude that the FRM candidates who lagged at one month perform best in forecasting the probability of having a current recession. All FRM candidates lagged at one month are statistically significant at 0.1% level. The pseudo R^2 around 40% means that the 40% variance of true recession indicators could be explained with one FRM obtained one month ago. We find that $FRM_{q70,u-1}$ is the best predictor to predict recession because of its highest pseudo R^2 value. The values of pseudo R^2 among all FRM candidates are directly comparable since these seven models have the same size of predictors and degree of freedom.

FRM/FRM_{IQR}*5	π_{mean}	π_{q50}	π_{q60}	π_{q70}	π_{q80}	π_{q90}	π_{IQR}
0.00	0.05	0.05	0.05	0.05	0.05	0.06	0.07
0.01	0.10	0.10	0.09	0.09	0.09	0.09	0.11
0.02	0.17	0.18	0.16	0.15	0.14	0.13	0.17
0.03	0.27	0.29	0.26	0.23	0.20	0.18	0.25
0.04	0.39	0.43	0.38	0.33	0.29	0.24	0.34
0.05	**0.52**	**0.58**	**0.51**	0.44	0.38	0.31	0.44
0.06	0.65	0.72	0.64	**0.57**	0.48	0.39	**0.54**
0.07	0.77	0.83	0.76	0.68	**0.59**	0.48	0.65
0.08	0.86	0.91	0.85	0.78	0.69	**0.57**	0.74
0.09	0.92	0.96	0.92	0.86	0.77	0.65	0.82
0.10	0.96	0.98	0.96	0.92	0.84	0.73	0.88
0.11	0.98	0.99	0.98	0.96	0.90	0.79	0.93
0.12	0.99	1.00	0.99	0.98	0.94	0.85	0.96
0.13	1.00	1.00	1.00	0.99	0.96	0.89	0.98
0.14	1.00	1.00	1.00	1.00	0.98	0.93	0.99
0.15	1.00	1.00	1.00	1.00	0.99	0.95	0.99
0.16	1.00	1.00	1.00	1.00	1.00	0.97	1.00

Table 5.1: Convert FRM@Americas to Recession probabilities

FRMQRecOne

Tables A.7 and A.9 compare the performances of FRM candidates with the lag at one month $FRM_{p,t-1}$ for Americas and Europe respectively. The results of criteria AIC, BIC, Log-likelihood, deviance and Area under the ROC curve (AUC) show that the performances of FRM candidates are similar. If we classify our data into two classes based on the implied recession probabilities, receiver operating characteristics (ROC) is a technique to compare the actual and predicted classification according to Fawcett (2006). Figure B.2 presents the confusion matrix based on true and predicted class as well as the relevant metrics used to construct ROC space, in which false positive rate is plotted on the x-axis, and the true positive rate is plotted on the y-axis. The area under the curve in ROC space is called AUC, which reduces ROC performance to a single scalar value representing the prediction

FRM/FRM_{IQR}*5	π_{mean}	π_{q50}	π_{q60}	π_{q70}	π_{q80}	π_{q90}	π_{IQR}
0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02
0.01	0.05	0.05	0.05	0.04	0.04	0.03	0.05
0.02	0.11	0.12	0.11	0.09	0.08	0.06	0.12
0.03	0.21	0.24	0.20	0.17	0.14	0.11	0.22
0.04	0.35	0.40	0.34	0.29	0.24	0.18	0.36
0.05	**0.51**	**0.58**	**0.51**	0.44	0.36	0.27	**0.53**
0.06	0.68	0.74	0.67	**0.59**	0.49	0.37	0.69
0.07	0.81	0.87	0.81	0.73	**0.63**	0.49	0.82
0.08	0.91	0.94	0.90	0.84	0.75	**0.61**	0.91
0.09	0.96	0.98	0.96	0.92	0.85	0.72	0.96
0.10	0.98	0.99	0.98	0.96	0.92	0.81	0.99
0.11	1.00	1.00	0.99	0.99	0.96	0.88	1.00
0.12	1.00	1.00	1.00	0.99	0.98	0.93	1.00
0.13	1.00	1.00	1.00	1.00	0.99	0.96	1.00
0.14	1.00	1.00	1.00	1.00	1.00	0.98	1.00
0.15	1.00	1.00	1.00	1.00	1.00	0.99	1.00
0.16	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 5.2: Convert FRM@Europe to Recession probabilities

FRMQRecOne

performance. AUC will always range from zero to one. Normally, a model with a higher AUC value corresponds to better average performance. Among seven FRM candidates for Americas, we find that $FRM_{q70}@Americas$ performed best in terms of pseudo R^2, AIC, BIC, Log-likelihood, deviance, and standard $FRM@Americas$ performed best in terms of AUC. Among seven FRM candidates for Europe, the $FRM_{q90}@Europe$ performed best in terms of R^2, AIC, BIC, Log-likelihood, Deviance, and $FRM_{q80}@Europe$ performed best in terms of AUC.

Based on the coefficients estimated from the uniperiod model with $k = 1$, the current probability of observing the recession can be predicted with the values of historical FRM candidates, respectively. Table 5.1 and 5.2 are conversion tables from the value of FRM candidates to the probability of recession for the Americas and Europe, respectively. From both tables, we conclude that the FRM candidate constructed with the higher quantile needs a higher value to imply the occurrence of recession. Moreover, for each FRM candidate, the increment of the probability of recession resulted from the increase of FRM value at 0.01 are not equivalent. For example, if the $FRM_{q50}@Americas$ increases from 0.04 to 0.05, the increment of the π_{q50} is 15%. However, if the $FRM_{q50}@Americas$ increases from 0.09 to 0.10, the increment of the π_{q50} is 2%. The bold numbers in tables 5.1 and 5.2 are the first predicted probability that could be classified as a recession for each FRM candidates if the threshold is assumed to be 50%.

Furthermore, we will check the in-sample performance of the above mentioned uniperiod recession prediction models. The left plots in Figures 5.2 prove that FRM candidates for the Americas and Europe were able to predict the recession before it was announced officially, since the implied recession probabilities were around or higher than 50%, and the trend of further increase also clear. As the financial system recovers from the great recession, the value of FRM candidates decreased dramatically. The implied recession probabilities for the Americas and Europe during the COVID-19 pandemic and corresponding recession indicators were visualized in the right plots in Figure 5.2. We observe that FRM candidates cannot predict this recession that was not rooted in financial activities. The fluctuations of stock prices, macroeconomic indices on the financial markets were delayed information of the COVID-19 outbreak. It is worth mentioning that FRM candidates adjusted themselves quickly to fit the recession resulted from the coronavirus. All FRM candidates for the Americas and Europe grew considerable within March, so that implied recession probability reached 100% in April. NEPR announced that the recession came around March, then $FRM@Americas$ candidates had a delay at one month. However, since CEPR reported that European countries went into recession in January 2020, the delay of $FRM@Europe$ was longer. Regarding the comparison of FRM candidates, the conclusion from visualization is consistent with that from chapter 4, i.e. it is hard to distinguish the performance of candidates for both markets.

5.3 Multiperiod Recession Prediction Models

In this chapter, we will describe multiperiod regression models, which is differentiated from uniperiod models by including more lags of each FRM candidate into one regression model. Based on this saturated model, predictors have been backwards selected using AIC stepwise method. Each model's AIC value takes the goodness of fit, which is assessed by the likelihood function, and a parameter to penalize the increased number of predictors into account. The penalty aims to eliminate the overfitting effect so that the model is not only suitable for our

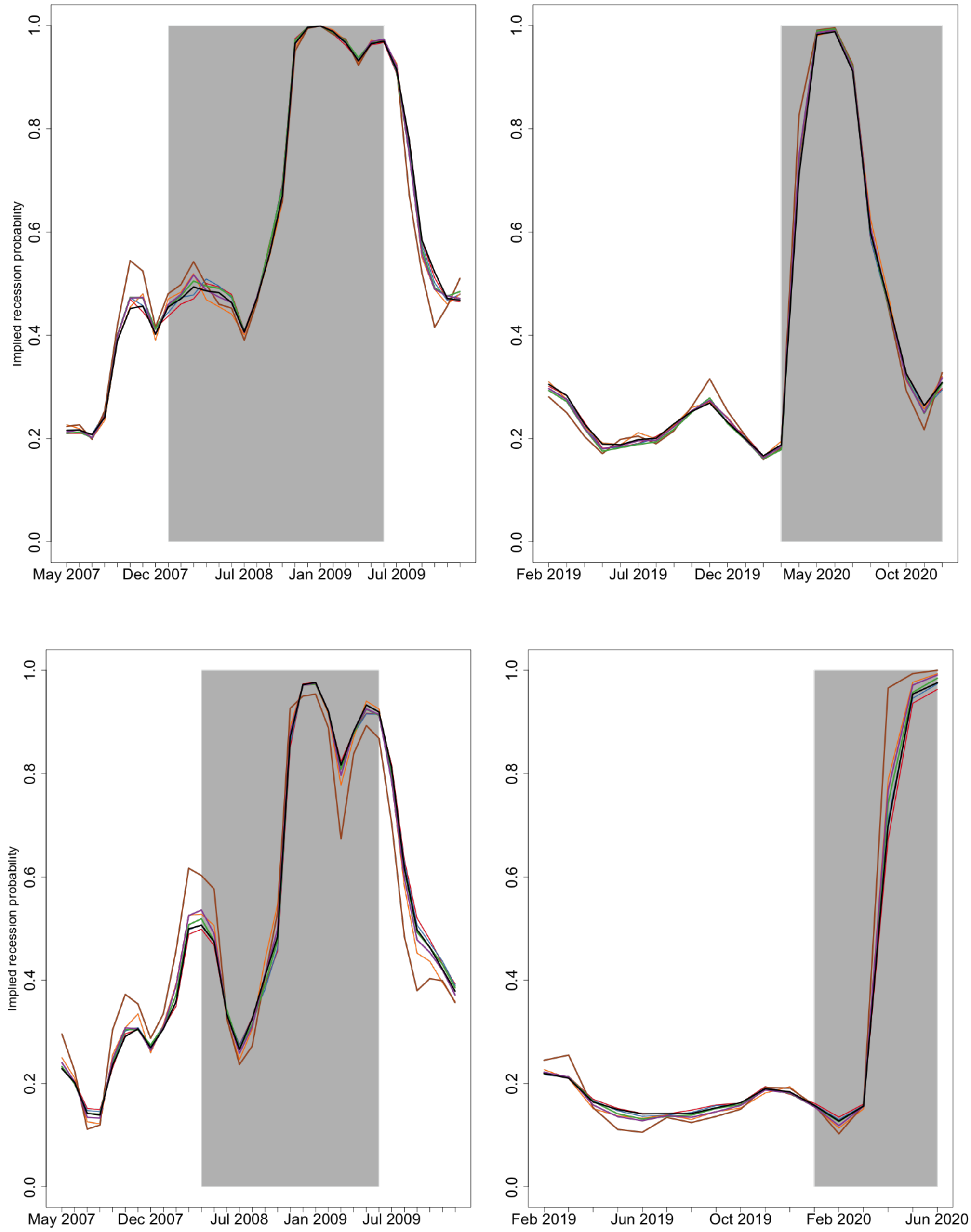

Figure 5.2: Recession probabilities implied from uniperiod model with standard FRM, FRM_{q50}, FRM_{q60}, FRM_{q70}, FRM_{q80}, FRM_{q90}, FRM_{IQR} and recession indicators for Americas (top) and Europe (bottom)

FRMQRecOne

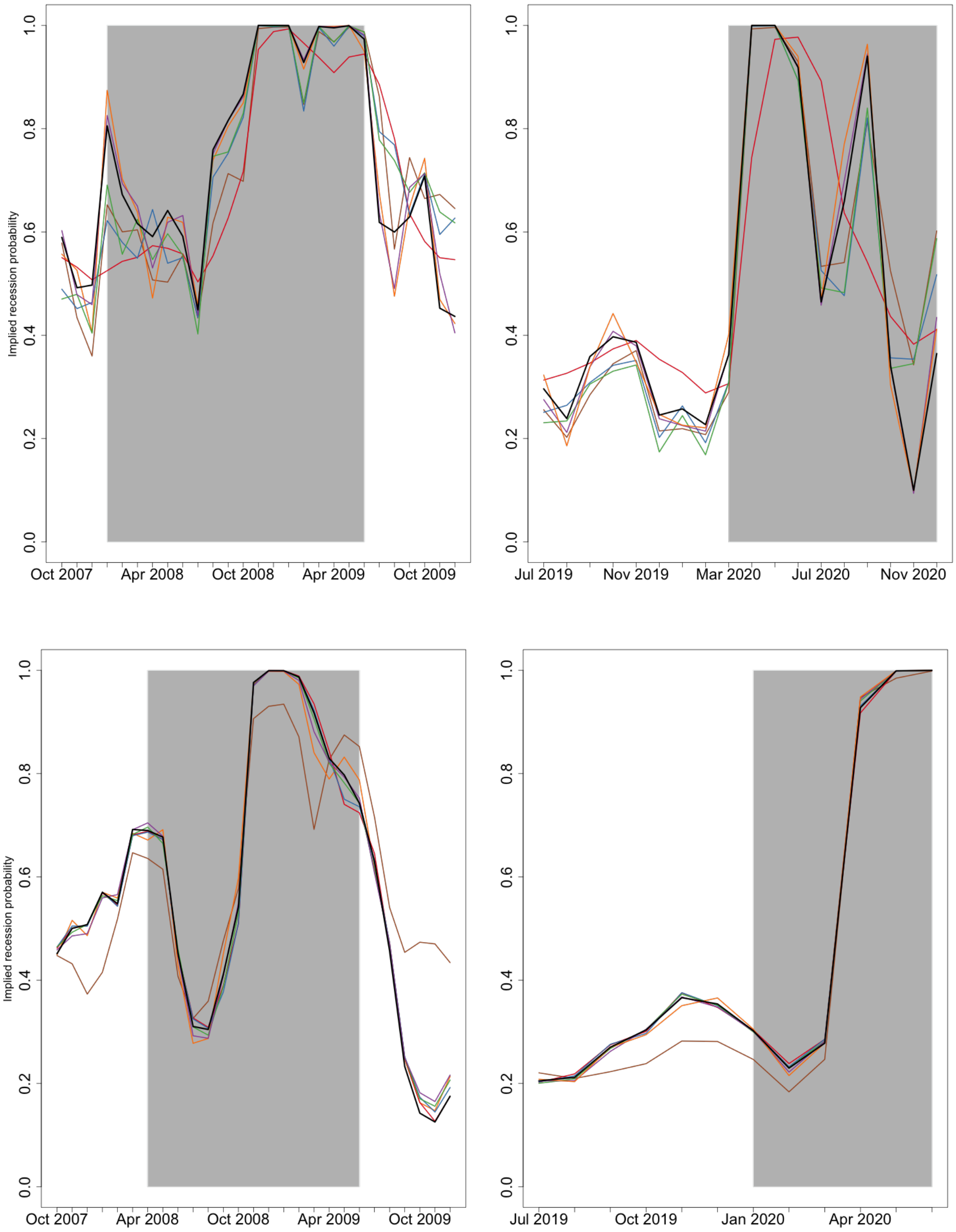

Figure 5.3: Recession probabilities implied from multiperiod model with standard FRM, FRM_{q50}, FRM_{q60}, FRM_{q70}, FRM_{q80}, FRM_{q90}, FRM_{IQR} and recession indicators for Americas (top) and Europe (bottom)

FRMQRecMulti

particular empirical data. For each FRM candidate, there is a model with the lowest AIC value, in which the most efficient predictors are selected so that the complexity of multiperiod recession models is decreased. Several predictors are removed once their corresponding coefficients shrink to zero. Analogue to the results of uniperiod models, only the results from probit regression models will be elaborated. Tables A.10 and A.11 report the optimal multiperiod model with the lowest AIC value among all models for each FRM@Americas and FRM@Europe candidate. Each column of both tables indicates the summary of the optimal regression model for individual FRM candidates. Both tables list multiple selected predictors used to construct the multiperiod regression models for individual FRM candidates. For Americans, only the latest predictor is statistically significant among models of other FRM candidates expect for $FRM_{q80}@Americas$ and $FRM_{q90}@Americas$. Multiperiod models perform relatively better than uniperiod models considering statistical criteria for model comparison such as AIC, BIC,and deviance. Including more predictors leads to a higher level of pseudo R^2 for each FRM candidate. The amount of pseudo R^2 increases from about 40% to around 63%, even to 67%. However, according to the definition of pseudo R^2, it is not feasible to compare the performances between the uniperiod model and the multiperiod. For $FRM_{q50}@Americas$, the optimal selected model contains only the predictor with $k = 1$, which is identical to the uniperiod model. However, the smaller observation number, which is resulted from removing the first 6-months data, contributes to the rise of pseudo R^2 by approximately 18%. Therefore, the higher pseudo R^2 cannot imply the better performance of multiperiod models. Nevertheless, except for $FRM_{q50}@Europe$, AUC values calculated from multiperiod models are relatively lower than those calculated from uniperiod models. It indicates that multiperiod models cannot predict the upcoming recession as accurate as uniperiod models, although more historical information is taken into account. We observe that each positive predictors tend to be followed by a negative predictor and vice versa.

Figure 5.3 represents that the overall trend of recession probabilities implied from different FRM candidates are identical. The tendencies of recession probabilities implied from multiperiod models and uniperiod models are also synchronized. All FRM@Americas and FRM@Europe candidates are capable of predicting the upcoming recession before the financial crisis of 2007. Comparing with the results from uniperiod recession models, the implied recession probabilities estimated from multiperiod models experienced a higher level and more considerable increase before the financial crisis of 2007. When that recession was announced, the implied recession probability for the Americas was more than 80% while that for Europe was about 70%, which was significantly larger than the values implied from uniperiod models. Moreover, $FRM@Europe$ candidates were able to capture the economic recovery before the beginning of an expansion period was reported. $FRM@Americas$ candidates predicted an increased probability of observing recession from February 2020, which corresponds to the recession caused by the COVID-19 pandemics and was not captured by uniperiod models. The FRM@Europe implied recession probability also increased from February. However, the CEPR recession indicator defines that there was already a recession in Europe in January 2020. Although the recession probabilities implied from multiperiod models based on FRM@Americas candidates fit the true recession indicator well, the high level of implied recession probability before and after the financial crisis of 2007 in the first plot of Figure 5.3 motivates us to check whether these multiperiod regression models may overestimate the probability of observing a recession. Before and after the financial crisis of 2007 May, the high level was due either to the high stress or to the model misspecification. Hence we need to check the prediction results during stable periods to ensure that using too many predictors does not deteriorate the out-of-sample performance of recession models. We conducted a validation by applying the standard FRM series for Americas produced by Yu et al. (2017) from 2010 until 2016 to multiperiod models listed in A.10. Figure 5.4 illustrates a generally low and stable estimated probability of recession during an expansion period. This empirical validation proved that these multiperiod probit regressions are plausible in both in-sample and out-of-sample tests. Comparing individual FRM candidates based on their results of multiperiod models, we conclude that $FRM_{q80}@Americas$ and $FRM_{q50}@Europe$ performed best among all FRM candidates for the Americas and Europe. However, their advantages over other candidates are also slight.

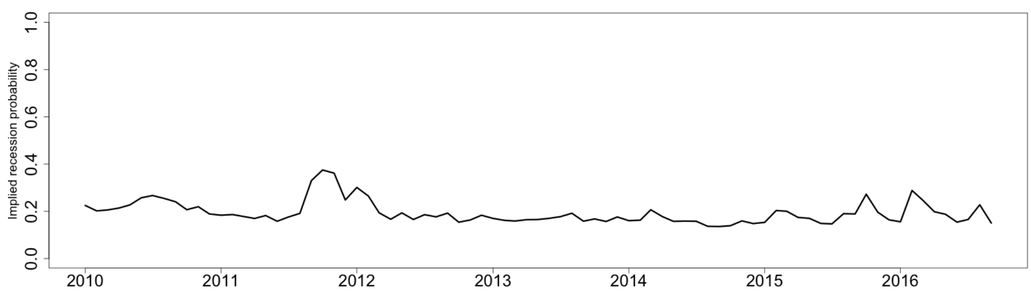

Figure 5.4: Recession probability implied from out-of-sample standard FRM.

FRMQRecMulti

After empirically applying historical FRM candidates as predictors to predict the upcoming or even current recession, we find that the choice regarding the type of binary regression makes a slight difference on the prediction accuracy. Besides performance of probit regression models that are interpreted above, logistic regression models convey similar results. Moreover, we observe that implied recession probabilities of uniperiod and multiperiod regression models prove the predictability of all FRM@Americas and FRM@Europe candidates. Nevertheless, considering the statistical criteria and the prediction accuracy, FRM candidates within one market are not significantly differentiated. Hence, we conclude that all FRM candidates, regardless of the Americas or Europe, have similar prediction ability to forecast the recession.

6 Conclusions

Following the linear quantile lasso regression model introduced by Yu et al. (2017), Yu et al. (2019) and Mihoci et al. (2020), we have estimated penalization parameters at daily frequency among 100 largest financial institutions of Americas and Europe, respectively. Besides standard FRM, we have defined six novel FRM candidates based on various quantiles, namely, probabilities at 50%, 60%, 70%, 80%, 90% and interquartile range(IQR). These quantiles are selected based on the characteristic of series of penalization parameters. Six FRM candidates reflect the systemic risk level of financial institutions in the middle, the worst 40, 30, 20, 10 percent of financial institutions in Americans and European financial markets, and dispersion of systemic risk of both markets. In order to prove their functionality and feasibility as systemic risk measures, these FRM candidates are statistically checked by taking other commonly used systemic risk measures as benchmarks. We have extended the correlation test and Granger causality test proposed by Yu et al. (2017) to newly defined $FRM@Americas$ candidates and all $FRM@Europe$ candidates. Each FRM candidates are highly positively correlated with other systemic risk measures. The correlation coefficients for different systemic risk measures are distinct, but the coefficients of various FRM candidates and a given systemic risk measure have a slight difference. The Granger causality test result indicates that several FRM candidates cannot mutually Granger cause several commonly used systemic risk measures. However, if we focus on the most prevalent systemic risk measure in the respective market, VIX in the Americas and VSTOXX in Europe, all FRM candidates can pass the Granger causality test, i.e. they have mutual Granger causality with other systemic risk measures. Kolmogorov-Smirnov test is introduced to compare the empirical distributions of individual FRM candidates and the distribution of each other systemic risk measures. Although only $FRM_{IQR}@Europe$ and VSTOXX can pass this test, the amount of Kolmogorov-Smirnov test statistics convey that the goodness-of-fit between the distribution of one specific systemic risk measure and individual FRM candidates is similar. Motivated by Mihoci et al. (2020), historical FRM candidates are further applied to predict the upcoming recession. Histories of FRM candidates at daily frequency are averaged into monthly time series to predict the probability of observing recession at the current month. Uniperiod and multiperiod recession models are used to explore the relationship between FRM candidates and recession indicators. After comparing possible recession models based on statistical criteria and implied recession probabilities, we found that the latest FRM data play essential roles in predicting the upcoming recession. Although the predictability of individual FRM candidates cannot be improved significantly by considering FRM data obtained more than one month ago, both uniperiod and multiperiod models can be used to predict the recession based on individual historical FRM candidates. However, the predictability and implied recession probabilities of each FRM candidate are similar.

From above empirical results, we conclude that all FRM candidates are feasible to be taken as a systemic risk measure and applied as a predictor of recession. Moreover, it is implied that penalization parameters in linear quantile Lasso regressions are robust since different portions of penalization parameters contribute to stable various FRM candidates. This analysis also offers the evidence that previous research based on standard FRM is reasonable, see Zbonakova et al. (2016), Yu et al. (2017), Yu et al. (2019), Mihoci et al. (2020) and Ren et al. (2020). Moreover, their analysis may further be applied to FRM candidates proposed in this paper. We have observed that subsamples of crucial financial institutions offer also useful FRM candidates. Hence, FRM candidates that are calculated based on quantiles could be used if we are interested in a specific group of financial institutions such as top 10 risky financial institutions.

There are several limitations to this study. Firstly, two time intervals with each at around two years are employed in this paper. Such a short time frame of empirical date limits the accuracy of comparison tests and recession models. For example, Granger causality tests that are inconsistent with the results in Yu et al. (2017) may be revised with larger sample size. Moreover, longer time intervals may include more recessions, which may result in more accurate recession models. We have found that the performance of FRM candidates based on two recession periods is better than that based solely on data collected from the COVID-19 pandemic. Secondly, simple recession prediction models employed in this paper could be further improved, such as using dynamic prediction models.

Therefore, the first suggestion for future research to employ a more extended research period if possible. Secondly, more applications in the context of FRM candidates, such as those proposed by Mihoci et al. (2020), could be applied to compare the performance of FRM candidates. Last but not least, the synchronized performance of FRM@Americas and FRM@Europe induces that the estimation of a cross-country FRM might be meaningful.

References

Adrian, T. and M. K. Brunnermeier (2011): "CoVaR," Tech. rep., National Bureau of Economic Research.

Bayer, S. (2018): "Combining Value-at-Risk forecasts using penalized quantile regressions," *Econometrics and statistics*, 8, 56–77.

Belloni, A., V. Chernozhukov, et al. (2011): "l1-penalized quantile regression in high-dimensional sparse models," *The Annals of Statistics*, 39, 82–130.

Bernard, H. and S. Gerlach (1998): "Does the term structure predict recessions? The international evidence," *International Journal of Finance & Economics*, 3, 195–215.

Brownlees, C. and R. F. Engle (2017): "SRISK: A conditional capital shortfall measure of systemic risk," *The Review of Financial Studies*, 30, 48–79.

Brownlees, C. T., R. Engle, et al. (2012): "Volatility, correlation and tails for systemic risk measurement," *Available at SSRN*, 1611229.

Chauvet, M. (1998): "An econometric characterization of business cycle dynamics with factor structure and regime switching," *International economic review*, 969–996.

Chauvet, M. and S. Potter (2005): "Forecasting recessions using the yield curve," *Journal of forecasting*, 24, 77–103.

Corrado, C. J. and T. W. Miller, Jr (2005): "The forecast quality of CBOE implied volatility indexes," *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, 25, 339–373.

Engle, R., R. Capellini, B. Reis, and G. DeNard (2020): "The Volatility Laboratory of the NYU Stern Volatility and Risk Institute," `http://vlab.stern.nyu.edu/`, [Online; accessed 11-December-2020].

Estrella, A. and G. A. Hardouvelis (1991): "The term structure as a predictor of real economic activity," *The journal of Finance*, 46, 555–576.

Estrella, A. and F. S. Mishkin (1996): "The yield curve as a predictor of US recessions," *Current issues in economics and finance*, 2.

——— (1998): "Predicting US recessions: Financial variables as leading indicators," *Review of Economics and Statistics*, 80, 45–61.

EuroAreaBusinessCycleNetwork (2020): "CEPR based Recession Indicators for Euro Area Business Cycle," `https://eabcn.org/dc/recession-indicators/`, [Online; accessed 20-November-2020].

Fahrmeir, L., T. Kneib, S. Lang, and B. Marx (2007): *Regression*, Springer.

Fan, Y., W. K. Härdle, W. Wang, and L. Zhu (2013): "Composite quantile regression for the single-index model," .

——— (2018): "Single-index-based CoVaR with very high-dimensional covariates," *Journal of Business & Economic Statistics*, 36, 212–226.

Fawcett, T. (2006): "An introduction to ROC analysis," *Pattern recognition letters*, 27, 861–874.

FederalReserveBankofSt.Louis (2020): "NBER based Recession Indicators for the United States from the Period following the Peak through the Trough," `https://fred.stlouisfed.org/series/USREC/`, [Online; accessed 20-November-2020].

Ghosh, S. (2002): "Electricity consumption and economic growth in India," *Energy policy*, 30, 125–129.

Granger, C. (1988): "Some recent development in a concept of causality," *Journal of Econometrics*, 39, 199 – 211.

Härdle, W. K., C. Y.-H. Chen, and L. Overbeck (2017): *Applied quantitative finance*, Springer.

Härdle, W. K., W. Wang, and L. Yu (2016): "Tenet: Tail-event driven network risk," *Journal of Econometrics*, 192, 499–513.

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

Hollo, D., M. Kremer, and M. Lo Duca (2012): "CISS-a composite indicator of systemic stress in the financial system," .

Jorion, P. (2007): *Value at risk: the new benchmark for managing financial risk*, The McGraw-Hill Companies, Inc.

KOENKER, R. AND G. BASSETT JR (1978): "Regression quantiles," *Econometrica: journal of the Econometric Society*, 33–50.

KOENKER, R. AND K. F. HALLOCK (2001): "Quantile regression," *Journal of economic perspectives*, 15, 143–156.

KOENKER, R., P. NG, AND S. PORTNOY (1994): "Quantile smoothing splines," *Biometrika*, 81, 673–680.

KRITZMAN, M. AND Y. LI (2010): "Skulls, financial turbulence, and risk management," *Financial Analysts Journal*, 66, 30–41.

LEE RODGERS, J. AND W. A. NICEWANDER (1988): "Thirteen ways to look at the correlation coefficient," *The American Statistician*, 42, 59–66.

LENG, C., Y. LIN, AND G. WAHBA (2006): "A note on the lasso and related procedures in model selection," *Statistica Sinica*, 1273–1284.

LI, Y. AND J. ZHU (2008): "L1-norm quantile regression," *Journal of Computational and Graphical Statistics*, 17, 163–185.

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

MASSICOTTE, P., D. EDDELBUETTEL, AND M. P. MASSICOTTE (2016): "Perform and Display Google Trends Queries - gtrendsR," [R pacakge].

MIHOCI, A., M. ALTHOF, C. Y.-H. CHEN, AND W. K. HÄRDLE (2020): "FRM financial risk meter," in *The Econometrics of Networks*, Emerald Publishing Limited.

NYBERG, H. (2010): "Dynamic probit models and financial variables in recession forecasting," *Journal of Forecasting*, 29, 215–230.

OSBORNE, M. R. (1985): *Finite algorithms in optimization and data analysis*, John Wiley & Sons, Inc.

OSBORNE, M. R., B. PRESNELL, AND B. A. TURLACH (2000): "On the lasso and its dual," *Journal of Computational and Graphical statistics*, 9, 319–337.

PREIS, T., H. S. MOAT, AND H. E. STANLEY (2013): "Quantifying trading behavior in financial markets using Google Trends," *Scientific reports*, 3, 1684.

Qontigo (2020): "STOXX Strategy index guide," https://www.stoxx.com/documents/stoxxnet/Documents/Indices/Common/Indexguide/stoxx_strategy_guide.pdf, [Online; accessed 12-Jan-2020].

Ren, R., M. Althof, and W. K. Härdle (2020): "Tail Risk Network Effects in the Cryptocurrency Market during the COVID-19 Crisis," *Available at SSRN 3753421*.

Ren, R., M.-J. Lu, Y. Li, and W. K. Härdle (2021): "Financial Risk Meter Based on Expectiles," *Available at SSRN*.

Schwarcz, S. L. (2008): "Systemic risk," *Geo. LJ*, 97, 193.

Simard, R., P. L'Ecuyer, et al. (2011): "Computing the two-sided Kolmogorov-Smirnov distribution," *Journal of Statistical Software*, 39, 1–18.

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

Tobias, A. and M. K. Brunnermeier (2016): "CoVaR," *The American Economic Review*, 106, 1705.

Yu, L., W. K. Härdle, L. Borke, and T. Benschop (2017): "FRM: a Financial Risk Meter based on penalizing tail events occurrence," .

——— (2019): "An AI approach to measuring financial risk," *The Singapore Economic Review*, 1–21.

Yuan, M. (2006): "GACV for quantile smoothing splines," *Computational statistics & data analysis*, 50, 813–829.

Zbonakova, L., W. K. Härdle, and W. Wang (2016): "Time varying quantile lasso," .

Zboňáková, L., R. P. Monti, and W. K. Härdle (2019): "Towards the interpretation of time-varying regularization parameters in streaming penalized regression models," *Pattern Recognition Letters*, 125, 542–548.

A Tables

Model	Order	PT(asymptotic)	PT(adjusted)	BG	ES
FRM_{mean} and GT	20	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.05	0.07
FRM_{q50} and GT	14	3.75×10^{-03}	3.15×10^{-03}	0.29	0.32
FRM_{q60} and GT	15	0.07	0.07	0.43	0.47
FRM_{q70} and GT	20	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.04	0.06
FRM_{q80} and GT	14	0.05	0.05	0.52	0.55
FRM_{q90} and GT	20	$< 2.20 \times 10^{-16}$	$< 2.20 \times 10^{-16}$	0.86	0.88
FRM_{IQR} and GT	16	1.49×10^{-03}	1.24×10^{-03}	0.14	0.16

Table A.1: p-values of model selection test of $FRM@Americas$ candidates and GT.

FRMQcausalityAM

Model	Order	PT(asymptotic)	PT(adjusted)	BG	ES
FRM_{mean} and FT	15	0.01	9.94×10^{-03}	0.53	0.56
FRM_{q50} and FT	15	0.05	0.05	0.04	0.05
FRM_{q60} and FT	11	2.43×10^{-03}	2.00×10^{-03}	0.15	0.17
FRM_{q70} and FT	17	1.14×10^{-03}	9.57×10^{-04}	0.26	0.30
FRM_{q80} and FT	12	0.22	0.21	0.75	0.77
FRM_{q90} and FT	10	0.03	0.03	0.27	0.29
FRM_{IQR} and FT	17	8.24×10^{-03}	7.23×10^{-03}	0.30	0.34

Table A.2: p-values of model selection test of $FRM@Americas$ candidates and FT.

FRMQcausalityAM

Model	Order	PT(asymptotic)	PT(adjusted)	BG	ES
$DFRM_{mean}$	13	0.05	0.04	0.04	0.05
$DFRM_{q50}$	14	0.12	0.11	0.17	0.19
$DFRM_{q60}$	14	0.90	0.89	0.75	0.77
$DFRM_{q70}$	14	0.05	0.04	0.11	0.13
$DFRM_{q80}$	14	0.65	0.63	0.86	0.87

Table A.3: p-values of model selection test of $DFRM@Americas$ candidates and $DSRISK$.

FRMQcausalityAM

DFRM	p-values (DFRM Granger causes DSRISK)	p-values (DSRISK Granger causes DFRM)
$DFRM_{mean}$	0.1	0.51
$DFRM_{q50}$	0.99	0.68
$DFRM_{q60}$	0.99	0.75
$DFRM_{q70}$	0.99	0.77
$DFRM_{q80}$	1	0.85

Table A.4: p-values of Granger causality test of $DFRM@Americas$ candidates and $DSRISK$

FRMQcausalityAM

Model	Order	PT(asymptotic)	PT(adjusted)	BG	ES
FRM_{mean} and FT	5	0.02	0.02	0.10	0.11
FRM_{q50} and FT	12	0.16	0.14	0.53	0.56
FRM_{q60} and FT	8	0.06	0.05	0.15	0.16
FRM_{q70} and FT	11	0.02	0.01	0.46	0.49
FRM_{q80} and FT	13	0.07	0.07	0.14	0.15
FRM_{q90} and FT	9	0.01	0.01	0.21	0.22
FRM_{IQR} and FT	13	0.20	0.19	0.23	0.25

Table A.5: p-values of model selection test of $FRM@Europe$ candidates and FT.

FRMQcausalityEU

	$\omega_{FRM_{mean,k}}$	$\omega_{FRM_{q50,k}}$	$\omega_{FRM_{q60,k}}$	$\omega_{FRM_{q70,k}}$	$\omega_{FRM_{q80,k}}$	$\omega_{FRM_{q90,k}}$	$\omega_{FRM_{IQR,k}}$
k=1	33.80***	37.59***	33.89***	30.59***	26.35***	21.78***	133.68***
	(9.81)	(10.91)	(9.87)	(8.90)	(7.66)	(6.36)	(38.75)
	[39.64]	[39.85]	[40.03]	[**40.55**]	[40.22]	[39.86]	[40.30]
k=2	22.73**	25.57**	22.98**	20.47**	17.54**	14.63**	89.93**
	(7.33)	(8.18)	(7.35)	(6.54)	(5.64)	(4.73)	(28.79)
	[32.32]	[32.84]	[32.99]	[33.08]	[32.55]	[32.47]	[32.96]
k=3	16.70*	18.90**	17.05**	15.27**	12.99**	10.67*	68.76**
	(6.52)	(7.25)	(6.51)	(5.81)	(5.03)	(4.18)	(25.88)
	[30.19]	[30.71]	[30.90]	[31.08]	[30.52]	[30.21]	[31.52]
k=4	12.43*	14.18*	12.81*	11.48*	9.81*	8.03*	54.06*
	(6.14)	(6.82)	(6.12)	(5.45)	(4.74)	(3.94)	(24.51)
	[31.47]	[31.92]	[32.07]	[32.18]	[31.89]	[31.63]	[33.16]
k=5	7.63	9.06	8.14	7.37	6.14	4.84	34.95
	(5.83)	(6.48)	(5.80)	(5.17)	(4.50)	(3.73)	(23.09)
k=6	2.16	3.25	2.19	2.52	1.87	1.19	13.36
	(5.65)	(6.26)	(5.60)	(4.99)	(4.34)	(3.60)	(22.14)

***:significant at the 0.1% level; **: significant at the 1% level; *: significant at the 5% level.

Table A.6: Estimated coefficients and deviance of the one period forecast models for FRM@Americas candidates.

FRMQRecOne

	FRM_{mean}	FRM_{q50}	FRM_{q60}	FRM_{q70}	FRM_{q80}	FRM_{q90}	FRM_{IQR}
AIC	59.59	59.46	**59.34**	59.00	59.22	59.45	59.16
BIC	63.61	63.47	63.35	**63.02**	63.23	63.46	63.18
LogLik	−27.80	−27.73	−27.67	**-27.50**	−27.61	−27.72	−27.58
Deviance	55.59	55.46	55.34	**55.00**	55.22	55.45	55.16
AUC	**84.52%**	79.93%	84.26%	83.99%	83.99%	83.86%	83.07%

Table A.7: Comparing the prediction ability of FRM@Americas candidates with k=1.

FRMQRecOne

	$\omega_{FRM_{mean,k}}$	$\omega_{FRM_{q50,k}}$	$\omega_{FRM_{q60,k}}$	$\omega_{FRM_{q70,k}}$	$\omega_{FRM_{q80,k}}$	$\omega_{FRM_{q90,k}}$	$\omega_{FRM_{IQR,k}}$
k=1	42.75***	45.78***	42.42***	39.00***	35.10***	30.02***	213.19***
	(11.66)	(12.53)	(11.58)	(10.64)	(9.60)	(8.23)	(60.19)
	[39.75]	[38.66]	[39.11]	[39.83]	[40.56]	[**41.10**]	[40.82]
k=2	31.19 **	33.09 **	30.94 **	28.81 **	26.15 **	22.26 **	171.76 **
	(10.27)	(11.11)	(10.27)	(9.41)	(8.43)	(7.16)	(53.63)
	[30.85]	[29.84]	[30.33]	[31.15]	[31.97]	[32.28]	[34.68]
k=3	20.09 *	21.42 *	20.28 *	18.97 *	17.51 *	14.60 *	125.23 *
	(9.63)	(10.45)	(9.69)	(8.86)	(7.95)	(6.71)	(50.31)
	[24.56]	[24.17]	[24.58]	[25.05]	[25.71]	[25.47]	[28.80]
k=4	11.08	11.69	11.14	10.52	9.82	8.23	75.45
	(9.20)	(9.98)	(9.27)	(8.48)	(7.60)	(6.45)	(49.62)
k=5	5.53	5.65	5.57	5.47	5.24	4.45	49.03
	(9.13)	(9.90)	(9.19)	(8.41)	(7.52)	(6.39)	(49.02)
k=6	-0.88	-1.20	-0.85	-0.53	-0.11	0.10	15.27
	(9.20)	(10.00)	(9.28)	(8.47)	(7.57)	(6.43)	(49.16)

***:significant at the 0.1% level; **: significant at the 1% level; *: significant at the 5% level.

Table A.8: Estimated coefficients and deviance of the one period forecast models for FRM@Europe candidates.

FRMQRecOne

	FRM_{mean}	FRM_{q50}	FRM_{q60}	FRM_{q70}	FRM_{q80}	FRM_{q90}	FRM_{IQR}
AIC	52.25	52.86	52.61	52.20	51.79	**51.48**	51.64
BIC	56.03	56.64	56.39	55.99	55.57	**55.26**	55.42
Log Likelihood	−24.12	−24.43	−24.31	−24.10	−23.89	**-23.74**	−23.82
Deviance	48.25	48.86	48.61	48.20	47.79	**47.48**	47.64
AUC	80.44%	79.93%	80.44%	80.44%	**80.78**%	80.61%	79.93%

Table A.9: Comparing prediction ability of FRM@Europe candidates k=1.

FRMQRecOne

	$\theta_{FRM_{mean,k}}$	$\theta_{FRM_{q50,k}}$	$\theta_{FRM_{q60,k}}$	$\theta_{FRM_{q70,k}}$	$\theta_{FRM_{q80,k}}$	$\theta_{FRM_{q90,k}}$	$\theta_{FRM_{IQR,k}}$
intercept	-1.43*	-1.10**	-1.54*	-1.64*	-1.49*	-1.45*	-1.36*
k=1	102.83*	28.83**	84.83*	84.14*	80.76*	68.92*	280.64*
	(46.69)	(10.24)	(41.89)	(39.47)	(35.34)	(30.14)	(129.24)
k=2	-84.71		-85.27	-87.30	-66.82*	-38.23*	-281.21
	(44.77)		(53.04)	(49.73)	(33.69)	(29.27)	(169.54)
k=3			36.72	37.86			-140.04
			(26.25)	(24.36)			(94.13)
k=4	35.75				28.93	26.90	
	(20.45)				(15.64)	(14.41)	
k=5							
k=6	-16.86				12.81	12.43	
	(11.49)				(8.72)	(7.75)	
R^2	66.21%	57.83%	62.97%	64.22%	67.11%	67.28%	63.01%
AIC	52.58	52.31	52.86	51.99	51.93	51.81	52.77
BIC	61.61	55.92	60.09	59.22	60.96	60.84	59.99
LogLik	−21.29	−24.15	−22.43	−21.99	−20.97	−20.90	−22.38
Deviance	42.58	48.31	44.86	43.99	41.93	41.81	44.77
AUC	83.19%	78.15%	80.25%	80.88%	83.82%	83.19%	78.57%

***:significant at the 0.1% level; **: significant at the 1% level; *: significant at the 5% level.

Table A.10: Estimated coefficients and deviance of the stepwise selected multiperiod models for each FRM@Americas candidates.

FRMQRecMulti

	$\theta_{FRM_{mean,k}}$	$\theta_{FRM_{q50,k}}$	$\theta_{FRM_{q60,k}}$	$\theta_{FRM_{q70,k}}$	$\theta_{FRM_{q80,k}}$	$\theta_{FRM_{q90,k}}$	$\theta_{FRM_{IQR,k}}$
k=1	59.14**	64.99**	58.22**	52.00**	45.08**	37.45**	174.14**
	(21.56)	(24.08)	(21.59)	(19.19)	(16.53)	(13.54)	(60.47)
k=6	-30.18	-33.84	-29.54	-25.62	-21.26	-17.26	
	(16.15)	(17.79)	(16.04)	(14.34)	(12.42)	(10.42)	
R^2	63.77%	63.25%	62.98%	63.04%	62.99%	63.12%	58.07%
AIC	44.97	45.28	45.44	45.40	45.43	45.36	46.31
BIC	49.96	50.27	50.43	50.39	50.42	50.35	49.64
LogLik	-19.48	-19.64	-19.72	-19.70	-19.72	-19.68	-21.16
Deviance	38.97	39.28	39.44	39.40	39.43	39.36	42.31
AUC	79.89%	80.16%	80.16%	79.63%	79.37%	80.42%	75.66%

***:significant at the 0.1% level; **: significant at the 1% level; *: significant at the 5% level.

Table A.11: Estimated coefficients and deviance of the stepwise selected multiperiod models for each FRM@Europe candidate.

FRMQRecMulti

B Figures

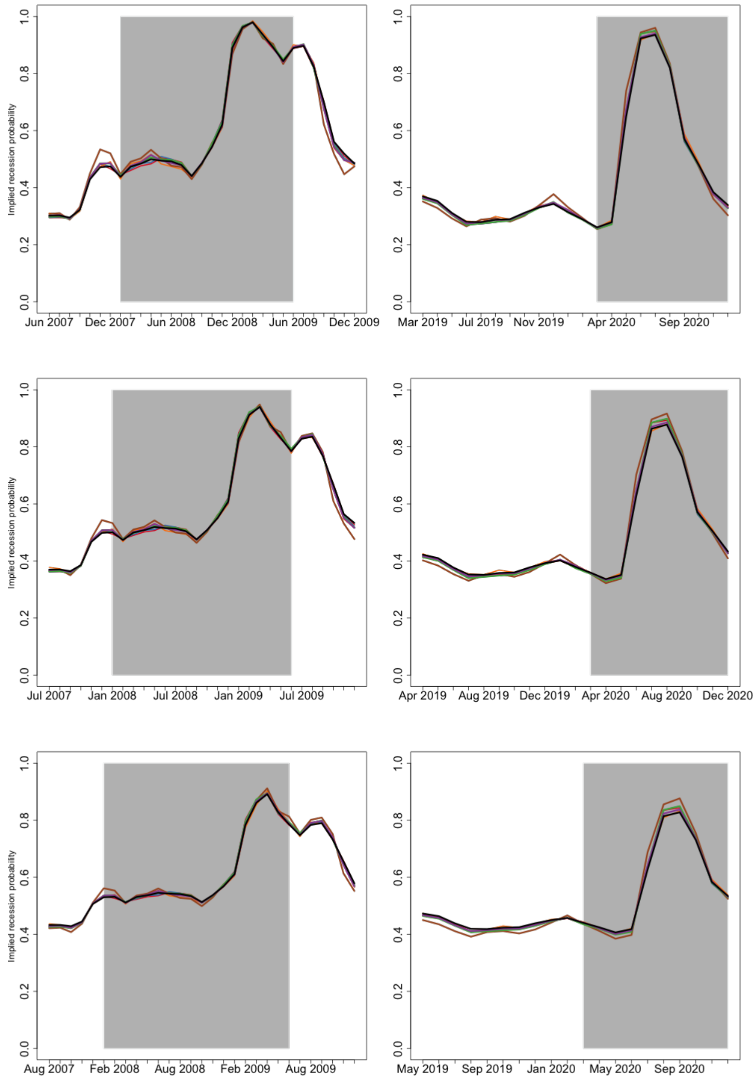

Figure B.1: FRM@Americas candidates implied recession probability with forecast horizons from two until four months from April 2007 until December 2009 and from January 2019 until December 2020

FRMQRecOne

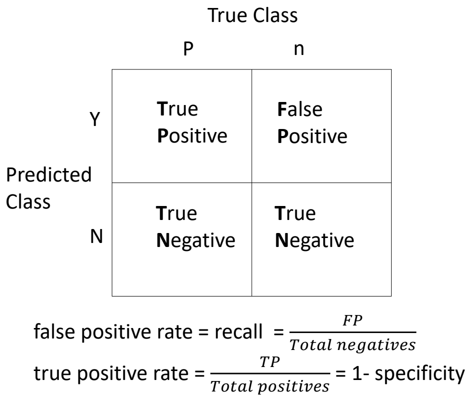

Figure B.2: Confusion matrix and corresponding performance metrics.

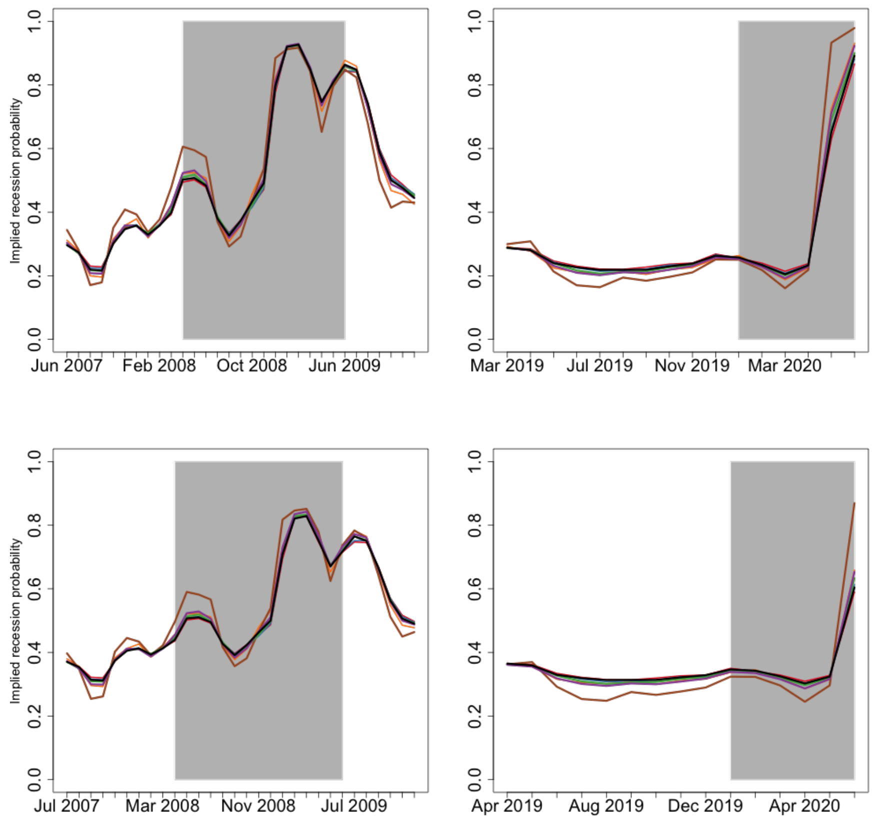

Figure B.3: FRM@Europe candidates implied recession probability with forecast horizons from two until three months from April 2007 until December 2009 and from January 2019 until October 2020.

FRMQRecOne

Declaration of Authorship

I hereby confirm that I have authored this Master's thesis independently and without use of others than the indicated sources. All passages which are literally or in general matter taken out of publications or other sources are marked as such.

Frankfurt am Main, March 26, 2021

Ranqing Song