UNIVERSITY OF LIVERPOOL

CRISIS TRANSMITTING EFFECTS DETECTION AND EARLY WARNING SYSTEMS DEVELOPMENT FOR CHINA'S FINANCIAL MARKETS

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Philosophy

by

Peiwan Wang

Department of Mathematical Sciences

University of Liverpool

December 2021

DECLARATION

I hereby declare that the thesis entitled "CRISIS TRANSMITTING EFFECTS DETECTION AND EARLY WARNING SYSTEMS DEVELOPMENT FOR CHINA'S FINANCIAL MARKETS" submitted by me, for the award of the degree of *Doctor of Philosophy* to University of Liverpool is a bona fide and independent work carried out by me under my primary supervisor Dr. Lu Zong. I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university. I give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signature of the Candidate: 王珮瑜

Date: 2021.12.13

ABSTRACT

In the background of China's economic development mode being focused the worldwide attention, there is a growing trend to study the risk transmission pattern and the crisis forecasting mechanism for China's financial markets by domestic and global academics. The study progress, however, is observed to be affected by two gaping research problems: 1) few studies construct comparative contagion models and integrated crisis forecasting systems for China's financial markets and 2) current econometric models hired to the risk spreading effects detection and the financial crisis forecasts are yet deterministically investigated in terms of the effectiveness on China.

To fill the gaps, this research proposes two hybrid contagion models and prototypes the early warning systems with motivations of first analyzing the crisis linkages and transmission channels across domestic markets in hierarchical frameworks, and then predicting the market turbulence by integrating the crisis identifying techniques and time-dependent deep learning neuron networks. To accomplish our aims, the full project is progressed in phases by solving four technical challenges that portray two literature gaps of A) the crisis identification on the basis of price volatility state distinction, B) the decomposition for multivariate correlated patterns to infer the interdependence structure and risk spillover dynamics respectively, C) the real-time warning signals generation in comparison of between traditional and stylized predictive models and D) the contagion information fusion in the EWS frameworks to distinguish the leading indicators from between internal macroeconomic factors and external risk transmitters in statistical validation metrics.

The research mainly contributes to the comparative analysis on financial contagion effects detection and market turbulence prediction through the hybrid model innovations for CM and EWS development, and meanwhile brings practical significance to improve the risk management in investing activities and support the crisis prevention in policy-making. In addition, the model experimented results corroborate the China-characterized mode on risk transmissions and crisis warnings that 1) the stocks and real estate markets are verified to play the central role among risk transmitters, while the managed floating foreign exchange rate and the non-fully liberalized bond market are peripheral during the crisis; and 2) the all-round opening up policy increases the possibility of domestic security markets being exposed to external risk factors, especially relating to the cash flows, energy commodities and precious metals.

ACKNOWLEDGEMENT

Writing an elaborate thesis is the landmark for a PhD candidate to draw a good end to his/her doctoral career. For me, the significance of this thesis has been far more than meeting the requirements to get a degree since it reviews every bit of sweetness and bitterness in my four-year PhD journey.

My master advisor, Prof. Mark Beaumont, ever told me: "The most important quality for a PhD, in my experience, is endurance, not intelligence." Now, at the moment of approaching the end of my PhD progress, I finally understand his words. Endurance not only means trying to endure the pain of solitude but also learning how to enjoy it. Those deep late-night works with songs and stars, those missing meals because of debugging errors, those moonlight with literature reading and paper writing, and those 'faking' holidays with submitting deadlines, accompanying with thousand times of being frustrated and exhausted, sketch the full view of pursuit for truth and faith to be a qualified independent scholar.

It says that bitter learning is enough to usher in a harmonious life. I believe all the self-doubt and unspeakable loneliness will be internalized as the courage to move forward in the future challenge. Without experiencing the bitterness of life by striving upward, people tend to magnify the pessimism and hard to feel happiness. The depression once made me experience such an annoying state, becoming vulnerable and hypocritical. However, it allowed me to take a deeper look at myself, knowing what I love, what I am committed to, and how important it is to accept my shortcomings. I just hope what I have thought and understood along this way will continue to accompany me, blazing the brambles and thorns, braving the wind and waves, and hoisting the cloud sail to the raved sea.

Most importantly, countless people have supported my progress on this PhD, and I owe a great debt of gratitude.

Immense thanks to my primary supervisor, Dr. Lu Zong, who provided invaluable ideas to initiate fascinating projects and pertinent feedback on my analysis and article framing with great patience. I am sorry for my carelessness and recklessness once making trouble with her, and I sincerely hope she can accept my deep apology.

Great thanks to my supervisory team, Dr. Ahmet Goncu, Dr. David Liu and Dr. Yi Zhang, who actively participated in guiding, kind support and constructive feedback during my PhD, and all my IPAP interviewers, Dr. Jiajun Liu, Dr. Conghua Wen, Dr. Cihangir Kan, Dr. Fajin Wei and Dr. Feiyan Chen, who once generously shared their research experience with me in the annual progress review (APR).

Thanks to the pgsupport team of both the XJTLU and the UoL, especially

Ms. Yidan Qin, Ms. Jing Wen, and Dr. Nicola Pagani, for their support and help at each crucial stage of my PhD.

I am also indebted to my dear colleagues in Suzhou who helped me in my life and study, especially my best friends Dr. Yun Lu and Dr. Yurun Yang.

Special thanks to my lovely son, Ruiming Li. Whenever I felt depressed to lose confidence to hurdle the obstacles, he tightly cuddled me as if putting me on the strongest armor to resist all sorrows.

Deepest thanks to my parents, Mr. Xuewen Wang and Ms. Weihong Li, for their unconditional love give me the courage to not give up quickly in the face of setbacks.

Warm and heartfelt thanks go to all my families for their unfailing support and encouragement, remind me of the original aspiration to embark the doctoral adventures.

Last, it is reminiscent of my grandfathers in heaven. Though they were not able to witness their granddaughter persisting on the academic road, I hope they would not be disappointed in what they had expected of me - dash ahead to what the heart yearns, regardless of the pains.

The PhD research is funded by the XJTLU doctoral scholarship (Ref. RDF-16-01-11) and the Jiangsu Province Science and Technology Young Scholar Programme (No. BK20160391).

LIST OF PUBLICATIONS AND PROCEEDINGS

This thesis includes two journal publications, one conference paper, and two open-access proceedings. The published articles are included in chapters of $4-5$, and two proceedings are used in chapter of 6 and appendix of B respectively.

Peer-referred journal articles

1. Peiwan Wang and Lu Zong, (2019). *Contagion Effects and Risk Transmission Channels in the Housing, Stock, Interest Rate and Currency Markets: An Empirical Study in China and the U.S.*, The North American Journal of Economics and Finance, 54, 101113, ISSN 1062-9408. **This paper is referred to Chapter 4**.

2. Peiwan Wang, Lu Zong and Ye Ma, (2020). *An integrated early warning system for stock market turbulence*, Expert Systems with Applications, 135, 113463, ISSN 0957-4174. **This paper is referred to Chapter 5**.

Peer-refereed international conference papers

1. Peiwan Wang, Lu Zong, and Yurun Yang, (2020). *Predicting Chinese Bond Market Turbulences: Attention-BiLSTM Based Early Warning System*, In Proceedings of the 2020 2nd International Conference on Big Data Engineering (BDE 2020). Association for Computing Machinery, New York, NY, USA, 91–104. ISBN 978-1-4503-7722-5. **This proceeding is referred to Chapter 5**.

Open-access works

1. Peiwan Wang and Lu Zong, (2021). *Are Crises Predictable? A Review of the Early Warning Systems in Currency and Stock Markets*, arXiv:2010.10132v1. **Appendix B partially uses this work**.

2. Peiwan Wang, Yurun Yang, Lu Zong, (2021). *Attention Winged Deep Neural Early Warning System: How Does it Perform in Signaling Sovereign Crises for China?*, available at SSRN:`https://ssrn.com/abstract=3875942`. **This work is referred to Chapter 6**.

TABLE OF CONTENTS

List of Figures

List of Tables

List of Abbreviations

ANN Artificial Neural Networks
ARCH Autoregressive Conditional Heteroskedasticity
AUC Area Under Curve
Bi-LSTM Bidirectional Long-short term memory networks
Bi-SWARCH Bivariate Markovian Regime Switching ARCH
BN Bivariate Normal Distribution
BPNN Backpropagation Neural Networks
CM Contagion Model
DCC Dynamic Correlation Coefficient
DL Deep Learning
DNN Deep Neural Networks
EVT Extrem Value Theory
EWS Early Warning System
GARCH Generalized Autoregressive Conditional Heteroskedasticity
GPD Generalized Pareto Distribution
LSTM Long-Short Term Memory
MDA Mean Decrease Accuracy
MDG Mean Decrease Gini
ML Machine Learning
MLE Maximum Likelihood Estimation
NN Neural Networks
POT Peak-Over-Threshold
RCM Regime Classification Measure
RF Random Forests
RMSE Root Mean-squared Error
RNN Recurrent Neural Networks
ROC Receiver Operating Characteristic Curve
RP Research Problem
RV Realized Variance

SVM Support Vector Machine

SVR Support Vector Regression

SWARCH Markovian Regime Switching ARCH

Chapter 1

Introduction

1.1 Background information

Numerous literature studies have focused on China's economy as its meteoric rise to the world stage since the milestone point of either 1978, the start of reform and opening-up, or 2001, the formal join the WTO. A volume of financial models on the basis of experience in other countries have been referred to as they notice the similarity between the progress of China adopting market-oriented reforms and the paradigmatic framework from other free markets, especially for risk transmissions during financial shocks. The most important paradigm in the U.S.. Retrospecting the U.S. subprime mortgage crisis in 2007, its spreading effects triggered the nationwide collapse of domestic economic major sectors of housing to banks and led to worldwide recession in the West and some Asian countries in the aftermath. The domestic contagion started from successive bankruptcies. By the end of 2008, the U.S. stock market drastically glided as the Dow Jones index dropped from 14279 points to 7800 points with a total loss of 45%, and the U.S. dollar index fell by 12% in the same year. The Federal Reserve (FED) then cut the effective fund rate to $0\%-0.25\%$ and held such a 'zero' interest rate level for more than six years to spur economic recovery. The most noxious international transmission is subsequent sovereign debt crises burst in European economies since 2009, and even a decade later, the recovery of the global financial system is still undergoing due to the vast damage.

Being one of the greatest emerging markets and the second-largest economy in the globe, China has attracted increasing attention from domestic and over-

seas investors, and the main concern is thus naturally placed on whether the crisis-episode risk spillover across principal sectors in China will copy the U.S. mode towards a free market in a discursive style or will make distinctive economic models to shed the global shocks by virtue of solving problems in quintessential "China-style" ways considering to its historical and politic particularity. Such an argument is seemingly unsettled. The former opinion is supported by referring to the cultural similarity across Asian countries (Nagayasu, 2001; Ghosh and Basurto, 2006) that have been frequently observed in the financial system's stability during crises. For instance, in the 1997 Asian crisis, the currency of Thailand collapsed in January 1998 with a drop of over 100%. Within one month, the banking interest rate bubbled to the top before the nationwide financial markets started to avalanche. Soon after that, Thailand SET stock index hit the bottom points of 207 with a maximum draw-down of 76% at the start of September 1998, and the local housing price also dramatically fell by 20% in such unfolding and lasting impact in the final quarter of 1999. The latter one stems from the researches (Shalendra and Sharma, 2010; Huang and Ali, 2012) that realize China's special operation mode by imposing fiscal policy to regulate the financial disputes as to the crisis burst. The practical evidence has been provided by Yuan and Qiang (2010), who analyses the risk transmission path across principal Chinese markets of stocks, bonds, and the gold, and surprisingly find that the market contagion between the stock and the gold exists but becomes insignificant during the stock crisis, and the China's bond market performs as a safe haven in the extreme stock market collapse.

In the pros and cons of the argument, the econometric model for detecting contagious effects in China must be developed and applied, and better to be comparative to either of advanced and emerging countries. That is the first task of this thesis explores to specify the transmission channels across a broad range of assets in China and make appropriate comparisons to the free-market paradigm on the basis of contagion model (CM) development. As for which assets will be included, we opt for the real-estate, equity (of stocks), bond, and currency, which four are vigorously growing and taking the majority of transactions in China's trading markets.

The follow-up work for the contagion models' development across Chinese financial markets is to design a forecasting system that absorbs and analyzes both contagious and macro-economic information to generate in-time warnings to assist market participants and policymakers in decision-making. In the ideal case, the perfect warning signals output will contribute to hedge investment risks and guide authorities to block the systematic crackings. Such forecasting mechanism has a formal name of Early Warning System (EWS) and covers not only individual financial markets (Eichengreen et al., 1995; Schimmelpfennig et al., 2003; Gennaioli et al., 2014; Babecký et al., 2014) but the entire economic system (Oh et al., 2006; Samitas et al., 2020) as well. In our study, China's early warning system construction will be first discussed for two essential investment markets of stocks and bonds to 1) display the idiosyncratic building process for each investment market and 2) explain their corresponding results to propose individually targeted advice for different types of investors. Further exploration on merging contagious factors into studying the possibility of sovereign crises appearing in China will be implemented for guiding power-oriented policy.

Recollecting the turmoil days of Chinese stock markets, the Composite index of Shanghai Stock Exchange (SSEC) experienced one of its greatest falls at the end of 2007. The crashed bubbles in mid-2015 led to another extreme turbulence that triggered the instability in domestic financial surroundings. As the burst and lasting effect of stock market crises is recognized as the cause of critical society stress and results in increasing financial loads of the government, a systematic model that monitors the economic scenarios of financial markets and generates early warning signals for potential extreme risks is in heavy demand. Meanwhile, considering the Chinese bond market has been over 98 trillion yuan (14.5 trillion dollars) in 2019 and overtaken Japan to be the world's second-largest bond market, in such booming market shares and complex international surroundings, strengthening the institutional norms and reinforcing the risk control to maintain a steady and neutral situation for the bond market thus becomes demanding. Recently, the sudden rise of default events, especially for the high credit rating corporate bonds, for example, 120 bonds issued by 46 companies defaulted with a total scale of 111.217 billion yuan in the year of 2018, also make investors urge

to find an effective early warning system for the rapidly expanding but crisply toppling China's bond market.

Besides, the concern from international investors on the Chinese sovereign bonds is not totally immune to tension during the outbreak on COVID-19, even though both the S & P's global rating and Fitch rating have given 'A' for China issued euro sovereign bonds and the 5-year euro bond with a negative yield rate of -0.152% sold out in a short order time (Zheng, 2018). The concern is sourced from two worries. First, the outstanding government debt load has exceeded 10 trillion yuan at the end of 2010, which appeals that the exaggerated expansion of local governments' debt may lead to moral risks to sovereign default especially considering the poor opacity of data on local debts. On the other hand, China shows strong ties between countries' banking industries and governments sector (Deev and Hodula, 2016; Chen, 2017), specifically the state-owned banks can directly participate in large governmental projects, which implies the banking fragility may result in the deterioration of state funds and further raise the risk transmitting to sovereign default. Thus, the early warning system for predicting Chinese sovereign crises is required to suggest appropriate political instructions to maintain the low level of systematic risks in national debts and credits.

1.2 Research topics

In response to the argument on CM exploration and the rising concern on EWS construction for China, two mentioned issues in Section 1.1, we first do a survey on current studies of analyzing the contagious relationship across financial markets and structuring the financial crisis forecasting system, and then, in a nutshell, find two main types of research problems (RP).

- **RP1:** *From the knowledge and experience level, there is limited literature being directly referred as to develop the CM and EWS for China.*
- **RP2:** *In the technique perspective, the methodologies being used to model the econometric connectedness and to predict forwarding output based on historical data are manifold abundantly exploited for developed and emerging countries but uncertainly determined for China.*

1.2.1 Current questions

The mentioned two research problems can be explored as following four challenge questions according to our literature survey.

1. **How to identify the crisis episodes for specific markets?**
   To be the most important basis of the CM and EWS researches, identifying the crisis episodes (or defining the crisis variable in the EWS development) should be delicately addressed. In the previous studies for advanced and emerging countries, the CM often recklessly disregarded the univariate analysis on each asset for being the origin of crises, instead of straightforward model the co-movement dynamics by either decomposing the common and idiosyncratic components (i.e. latent factor models) or capturing the conditional heteroskedasticity of time-varying variances and co-variances (i.e. correlation-based models). The EWS studies, though, greatly promote crisis identification by quantifying the composite index based on several essential asset price dynamics and forming the crisis indicator function, scantly discuss the cutoff level for determining the significance of being abnormally deviated from being crisis observations.

2. **How to exploit appropriate technique fragments to formulate the contagious linkages as well as trace the risk transmission channels across different domestic markets?**
   In a massive volume of hired methodologies, the contagious effect is recognized as the joint tail behavior for multiple assets. The complex correlation is usually pairwise decomposed and investigated in a static way, rarely performed in a dynamic way to display the time-varying contagious effect dynamics, not to mention the inappropriateness of directly modeling the high dimensional joint distribution in simple structural breaks. Furthermore, previous works of CM studies have paid overwhelming efforts to explain the variation of correlated structure across different nations or continents to serve for the risk control during global financial crashes but rarely mine the insights to crisis-period risk transmitting across domestic financial sectors in the same country to drive the diversification of portfolio allocations.

3. **How to make the predictive mechanisms in EWS development more practically comparative to produce in-time warning signals?**
   The progress of developing the EWS generally embeds essential components of the crisis classifier and the predictive model. The former one will be settled as the first question can be solved. The latter one includes three competitively accredited models of classic parametric regression, non-parametric signaling approach, and stylized machine learning techniques, where they have been richly learnt for predicting banking and currency crashes based on low-frequency data to instruct the policy scheming for regulators, yet systematically compared in the EWS horse race to examine their practical significance for directing investment on specific financial assets based on high-frequency data for practitioners. Besides, the sophisticated frame of machine learning networks is always criticized by its complexity and ambiguity in explaining the relationship between input factors and the output (especially when compared to the parametric regression, which model allows to infer the marginal effects and visualize the impacting significance for each factor), in spite of considerably boosting the predicting accuracy.

4. **How to integrate the contagion information into the crisis forecasting mechanism and investigate whether they share equivalent contributing degrees with macro-economic factors in producing effective warning signals?**
   Some researchers have realized the importance of contagion information in developing the crisis forecasting system, for example, Dawood et al. (2017) pioneers that the contagion factors should be regarded as an essential part to predict the sovereign debt crisis. However, few studies dig deep into either proposing the workable ways to transfer the contagion information into time-varying dynamics as other input factors or evaluating their significance of contributing degree as compare to empirical macro-economic inputs.

The four questions are logically connected according to the order in which we find the gaps as learning the current literature.

1.2.2 Overview of jumping-offs

To jump over these obstacles, we then intuitively appoint solvable start-ups.

First, to solve *Question 1*, the usual way is either to collect experts' opinions or to define the crisis indicator function based on quantized/composite price index. However, the former one is too expensive for individual academics and practitioners to afford, especially for China, whose financial sector yet sets up a public database with transparency and free access, the quantified econometric techniques are more preferred to identify the crisis observations. In practice, the latter method is handier, especially being combined with appropriate statistical models. For example, the crisis samples are usually described as extreme events, which can be fitted in Extreme Value Theory (EVT) frameworks (McNeil and Frey, 2000; Lestano, 2007; Samuel, 2008). The alternative way is based on the price index dynamics volatility jump, which has been proven more efficient in distinguishing between the turmoil and tranquil states for specific financial markets (Hamilton, 1989; A and B, 1994; Abiad, 2003; Christiansen et al., 2012). It seems the question can be perfectly solved, but in formulating the binary crisis indicator variable for the EWS development, new troubles relating to the arbitrary thresholding cutoff and the fixed volatility levels are accused of lacking flexibility, which thus inspires us to innovate schemes that allow for variation on cutoffs and volatility regimes.

Second, *Question 2* mentions dual requirements to construct the CM for China: 1) technically, the complex joint tail behavior cannot be simply described but hierarchically decomposed to model the financial series that is full of the non-normality and 2) economically, the analysis cannot be dull for the single country but lively as compared to one of the other paradigm countries. The first requirement can be solved by the hybrid contagious models (Jondeau and Rockinger, 2006; Rodriguez, 2007; Aloui and Ben Aïssa, 2016; BenSaïda, 2018), that provide the feasible way to estimate the joint tail behavior layer by layer, which not only decomposes the dependent relationships and infer the transmitting channels in structure but saves computational cost as well. For the second one, the domestic markets in the U.S. will be opted as the paradigm of developed countries to make a comparison for discussing whether the risk transmission shows any difference

between advanced and developing countries[1].

Third, corresponding to *Question 3*'s requirement on producing a time-efficient prediction, the deep neural networks of the LSTM, which specialize in processing the time-dependent sequence data, will be first included in the horse race of stylized predictive models based EWSs (Nag and Mitra, 1999; Ohtsu, 2007; Sevim et al., 2014; Samitas et al., 2020). Meanwhile, to tease out the input factors' importance to the final prediction, like a traditional predictive model of logit regression-based EWS estimating parameters to gain the understanding of factors' marginal effects (Eichengreen et al., 1995; Bussiere and Fratzscher, 2006; Dawood et al., 2017; Filippopoulou et al., 2020), the attention mechanism (the technique that has been widely adopted in the natural language processing) is first combined with machine learning-based EWS models to estimate the contributing degree for each factor. In addition, to calculate the statistical metrics as previous studies done to evaluate the EWS performance, we directly apply the produced early warning signals to the investment portfolio and use the back-testing and reality check to investigate the practical guiding role of the EWS in investment risk control.

Last, as mentioned in *Question 4*, the fused EWS requires input factors in the form of time series, thus to quantify the contagion information as time-varying dynamics, we resort to both of DCC-GARCH model (that estimates the correlated relationship between two assets) and SWARCH model (that infers the crisis originator's probability of being unstable) to innovate the contagious intensity index given the specific crisis transmitting direction from the contagious source to the risk receiver. The further requirement is to validate the drawn weights' significance deviating from zero and the stability of referred non-zero contributors performing on between in-sample and out-of-sample sets. It inspires us to implement the hypothesis testing given appropriate assumption, similar to the p-value inference process in logit regression, on attention drawn weights. Here, we do not use the Gaussian distribution assumption that most traditional parametric regression models impose on factors, but first investigate their empirical distri-

[1]China has become the paradigm of developing countries because of its rapid economic growth and strong comprehensive prosperity in recent decades.

butions to make sure the rationality of implementing the hypothesis testing[2].

1.3 Objectives

According to current research problems (proposed in Section 1.2), the overall aim of our study is thus to develop effective CM for detecting the co-movement dynamics and risk transmitting channels across principal markets and to construct robust EWS for predicting turmoils or crises for a specific market, to finally instruct investors and governors to prevent themselves being exposed in unexpected damage from market turbulence oriented crises.

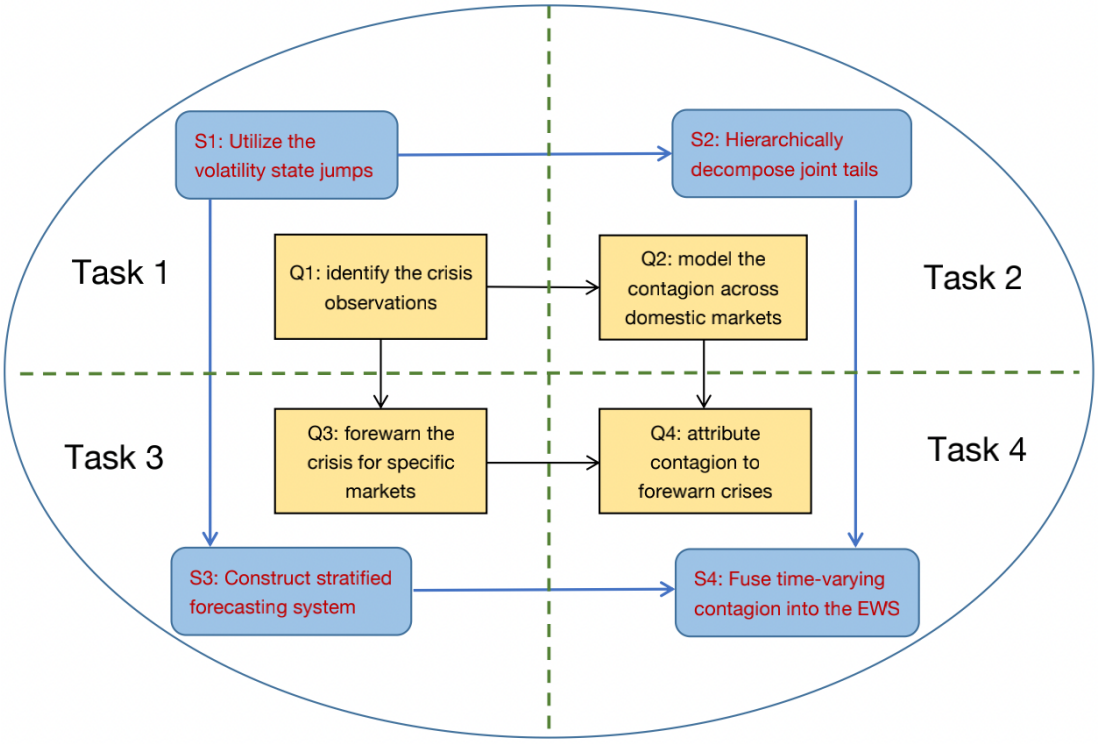


Figure 1.1: A diagram that illustrates the specific progress of structuring the thesis.

To accomplish this aim, we implement four major tasks. As the diagram in Figure 1.1 shows, these tasks are logically linked with each other in the sequence of achieving the following objectives embedded on both economic backgrounds for China and technical research problems for model construction mentioned in

[2]The Gaussian assumption is usually imposed on the parameter hypothesis testing in logit regression, but for attention drawn weights, the scheme will be more delicate as not any theoretical deduce has been appeared to prove such assumption's rationality on machine learning mechanism.

Section 1.1 and 1.2, respectively.

Objective 1. ***To classify the crisis observations for specific markets.***

- Use the volatility of price index dynamics to identify different levels from the tranquil to turmoils, and then samples locating in most high volatility state are recognized as crisis observations.

Objective 2. ***To make comparative contagion models for both of the US and China to investigate the connectedness across principal markets of stocks, bonds, real-estate and currency.***

- Visualize the co-movement dynamics and transmitting channels.
- Decompose the multivariate joint tail behavior in hierarchical model frameworks.

Objective 3. ***To construct early warning systems for China's stock and bond markets, respectively.***

- Discuss the impact of cutoff levels' variation (detailed in Chapter 5.2 by introducing the dynamic thresholding scheme) and volatility state number's change (discussed in Chapter 5.3 by applying the RCM in Markov switching model frameworks) on identifying the crisis and non-crisis samples.
- Produce in-time warning signals by hiring time-dependent deep neural networks in comparison with other stylized machine learning models and classical econometric models based EWS's output in statistical metrics.
- Retrieve the predicted results in portfolio construction for evaluating in practice.

Objective 4. ***To construct the contagion information fused early warning system for China's sovereign bond market.***

- Quantify the contagion information in time-varying series to embody the risk transmission intensity given the contagious source being recognized as the crisis originator (being in turmoil state).

- Distinguish the contributing degree between macro-economic factors and contagious factors, and judge whether they perform equal importance on predicting crises.
- Validate the attention mechanism drawn contributing degrees by adopting hypothesis tests given appropriate statistic variables for each input feature.

1.4 Main research workflow

This section will introduce the workflow of exploring the full PhD project during the research progress.

1.4.1 Literature basis for projects

At the start of exploring the projects, the literature survey should be well collected, filtrated, and sorted corresponding to solve current problems. Thus, three main clusters of references will be reviewed: (1) crisis identification methods applied in CM and EWS development, (2) contagious effects detection models in terms of decomposing the joint tails behavior across different markets, and (3) early warning system models developed for predicting financial crises for both of systematic economy and specific markets. The analysis for each cluster's pros and cons will be explored in Chapter 2.

1.4.2 Methodology basis for projects

To attain the objectives being inspired by jumping-off points, methodological preliminaries relating to filling the research gaps will be emphatically studied in terms of (1) crisis identification based on SWARCH model specified volatility levels, (2) contagious models that hierarchically analyze the correlated relationship across markets in Copula structured inference, Bi-SWARCH depicted co-movement periods and EVT estimated heavy tails, and (3) time-dependent deep learning model of LSTM based predictive model in EWS development on high-frequency data. The specific method formulation process will be detailed in Chapter 3.

1.4.3 Development of two hybrid CMs

Based on the second objective and proposed jumping-off points, two hybrid contagion models are developed and applied parallel to the log returns of housing and stock prices, the short-term interest rates, and the off-shore exchange rates in China and the U.S.. The attempt is to gain a more holistic view of the risk transmission channels across a broader range of assets, including the real estate, equity, bond, and currency in the emerging and developed markets.

The development process will be arranged as follows. First, the crisis and contagion episodes of the financial markets will be detected by univariate SWARCH and bivariate SWARCH, respectively. Second, by adopting the Extreme Value Theory (EVT) and regular vine (R-vine) copulas, we make an attempt to explore the joint behavior of heavy tails as well as the risk transmission channels in both Chinese and American financial systems during crises. Last, implications obtained from the empirical results will be discussed from a practical perspective by comparing the asset contagion behaviors in China and the U.S.. Chapter 4 will be further referred to diagram the procedure of implementing two CMs.

1.4.4 Development of EWS

To achieve the goal of constructing EWS for China, two essential markets for stocks and bonds that share the greatest trading volume in China will be differentiated in Chapter 5.

- **For stocks**, we first attempt to develop a robust crisis classifier based on SWARCH combined dynamic thresholding technique to precisely identify stock market turbulence on a daily basis, and then hire the LSTM network as the predictive model to produce alarming signals. In the comparison of CMAX based crisis classifier and other machine learning based predictor, the proposed integrated EWS algorithm will be examined by not only statistical metrics on out-of-sample set but cross-validation and back-testing as well. The reality check for data snooping will also be applied to retrieve the EWS performance in practice.

- **For bonds**, the ambiguity of defining the crisis variable will be first re-

solved by adopting SWARCH and clustering the observations into selecting appropriate count of volatility states in regime classification measure (RCM). Then, to weigh the economic factors in perceiving the forewarned turmoils and recognize the key economic factors to be inspected for making a timely macro-control policy, the attention combined bidirectional LSTM is adopted to predict warning signals and fully verify in multi-dimensional evaluating measurements.

1.4.5 Development of contagion fused EWS

To terminate the debate on existence of sovereign crises in China, a contagion fused early warning system is constructed by 1) hiring the RCM-SWARCH based quantified tools to relieve the brunt of chronological database deficiency to define the sovereign crisis, 2) quantifying the external contagion effects as the input features for EWS, and 3) instead of synthesizing a composite indicator index by thresholding leading factors, using the attention mechanism to simplify the feature selection between the leading indicators for machine learning prediction, as well as 4) designing three types of hypothesis tests to gain the inferred leading factors' credibility. Different from the EWS for specific markets to guide the practical application in investment, the contagion fused EWS is more apt to gain the economic understanding to grasp the key vulnerability channels in policy control. Its implementing procedure is diagrammed and detailed in Chapter 6.

1.5 Outline of the thesis

The thesis arranges seven chapters in total to outline the full projects.

Chapter 1 Introduction − briefly describes the backgrounds, the current research gaps, the intuitive jumping-off points, the research objectives, the overview of main research methodologies, the guide to read through the thesis and the study limitations.

Chapter 2 Literature review − introduces the knowledge basis, review the previous studies of CM and EWS development, and make comments on literature gaps.

Chapter 3 Methodological preliminaries − clarifies the main research methodologies adopted in our CM and EWS development, including volatility-based crisis identification, models that decompose the joint tails in terms of correlated pattern and co-movement dynamics, and deep neural networks based predictive models for EWS.

Chapter 4 Hybrid contagion models construction − specifies the components and structure of two hybrid CMs and empirically analyze their performance on domestic assets for China and the U.S.

Chapter 5 Integrated early warning systems development − introduces the conceptual prototype of the integrated EWS, including the technical improvement on both crisis classifier and predictive models, evaluates the developed EWS performance on specific markets of stocks and bonds in China.

Chapter 6 Contagion fused early warning system − introduces linking the contagion with the crisis prediction, illustrates the steps of constructing the contagion fused EWS, and experiments on predicting sovereign crises for China.

Chapter 7 Conclusions and discussions − underlines the study key contributions, practical implications, current limitations and further explorations.

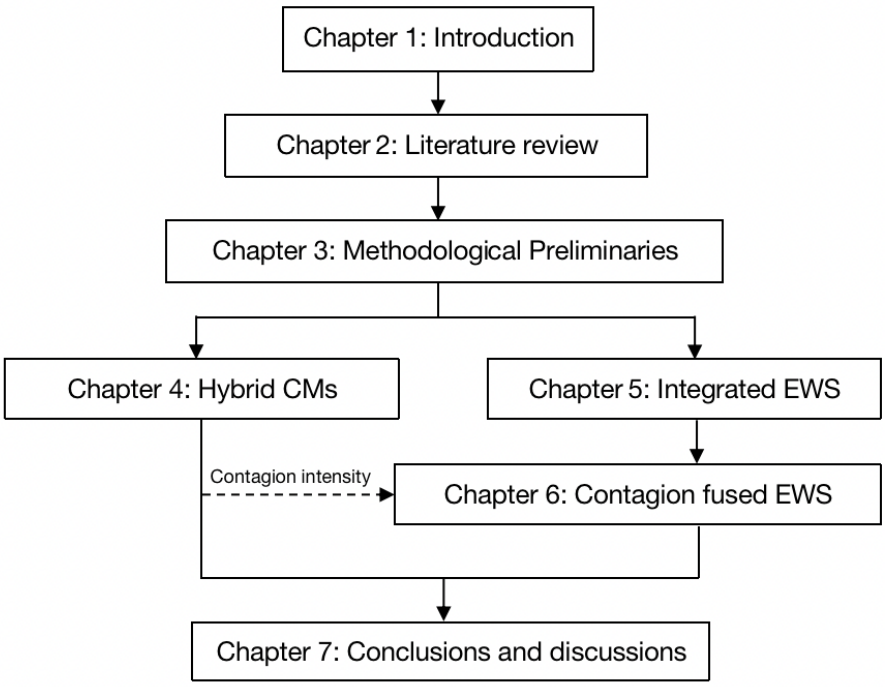


Figure 1.2: A flowchart of the thesis.

1.6 Study scope and limitations

This research aims to erect a compact detecting and predicting system for China's financial markets. The projects in this thesis focus on the research methodologies to improve the system performance on either empirical analysis or practical instruction. The linkage between the CM and EWS, as Figure 1.2 shows, is dashed with the contagion intensity dynamics, not directly connected with each other. The further project will focus on bridging the limitation of converting the contagion model structurally decomposed correlation into the EWS required time-series dynamics input.

In the inefficient markets, the risk transmission across financial markets is meanwhile dynamically changed by investors' behavior against to the potential crisis forewarning danger and their confidence in the authorities ability of preventing the crisis. Such real-time updated contagious effects are required to fed the EWS to boost the prediction time-efficiency. Although our current research stage does not cover this project, the methodological basis presented in this study can be applied to develop the investing behavior uncertainty driven contagion backflow EWS in further explorations.

Chapter 2

Literature review

In this Chapter, we introduce the literature surveys on fundamental knowledge basis for defining financial crisis and contagion (see Section 2.1), modeling financial contagions (see Section 2.2) and then review the studies for financial crisis early warning systems (see Section 2.3). The crisis identification will not be solely discussed in one chapter but separately explored for the CM and (fused-)EWS projects, since it though performs as the joint research basis, serves for different goals in subsequent CM and EWS developments: in CM, the crisis identification focuses on detecting the abnormal behaviors of market asset price fluctuations to infer whether they possibly produce spillover effects, while in EWS, it is the essential basis to form the crisis classifier, which will further influence on the predicting effectiveness of the entire early warning system. In the following sections, the arguments that provide supportive evidence to our research aims will be highlighted in bold font.

2.1 Fundamental knowledge basis

2.1.1 Financial crisis

Before going through the literature content, the elementary definition basis for the financial crisis should be first clarified. At the end of 20^{th} century, the IMF[1] classified financial crises into four types, namely, currency crisis, banking crisis, systemic crisis, and debt crisis. However, with the rise of emerging markets and the increasing market openness of Asian countries, the economic map pattern of

[1]According to the IMF published *World Economic Outlook* in May 1998.

the world has changed, which makes the definition of the financial crisis more diversified.

At present, the financial crisis is no longer a spontaneous occurrence of a single financial institution or sector but either comes from the transmission of financial crises burst in other regions or relies on the downturn of economic conditions derived from abnormal changes in several markets. For example, the global economic shock in 2008 is generally considered to be caused by the Subprime crisis in the United States.

By far, the most accepted financial crisis definition by most scholars and academics is "all or most of the financial indicators including the short-term interest rates, monetary assets, and securities' price, real estate and landing price, business bankruptcies and the fails of financial institutions perform a sharp, short and super cycle deterioration in one country or several countries and regions ". Regardless of the defining diversity, the common feature is attached among all types, that is, the crisis being recognized as the sharp fall in asset prices marked financial turmoils. Thus, the author believes that, in addition to the four classified crises, the sharp decline or the abrupt jump for the market price that has been paid extra attention by policymakers and investors should be included in the crisis studies.

2.1.2 Financial contagion

Financial contagion is one of the inevitable consequences of the financial crisis. As exampled in Chapter 1, Section 1.1, the transmission process of the 2008 subprime mortgage crisis after it broke out in the United States is witnessed like toppled dominoes, that the linkage effect among financial markets makes them fall one after another. Besides the observed phenomenon, how to precisely define the financial contagion?

According to the summary work by Pericoli and Sbracia (2003), there are four definitions: 1) the crisis in one country increases the possibility of crisis appearing in another country, which is usually cited in empirical analysis for currency crisis; 2) the volatility of asset prices in the financial markets of the countries being in crises spills over the financial markets of other countries; 3) as

the crisis occurs in a regional or a group of markets, the market price occurs "idiosyncratic co-movement"; 4) the "volatile contagion", which is explained as the market structure of crisis countries changes brought the change of crisis transmission channel. Based on the different definitions, experts have constructed various models to simulate the crisis transmitting process.[2] This study considers the second defined contagion to construct the risk transmission model for China since its definition relies on the asset price performed abnormal fluctuations during the crisis periods, which keeps consistent with the aforementioned study basis on market assets' volatility defined financial crisis.

2.2 Modeling contagious effects

Contagion effects emerge as a result of risk spillover in the financial crisis. Despite the definition of contagion remaining diverse, the majority of existing literature chooses to focus on either A) the increased probability of crises or B) the volatility jump of asset returns in one country given the occurring crisis in a different country (Dungey et al., 2005; Pericoli and Sbracia, 2003). In particular, Dungey and Martin (2001) investigate the contagion between currency and equity markets for a panel of countries in the East Asian crisis of 1997-98 using the latent factor structure that decomposes the asset return by a factor model of the common and the idiosyncratic components. The model is further extended considering the high volatility nature of asset returns during crises by including a $GARCH$ component that captures the heteroskedasticity as well as the time-varying variances and covariance, and applied to study the east Asian currency crisis (Dungey and Martin, 2004) and stock market crash in Europe, South-East Asia, and Latin America (Bekaert et al., 2005). The other group of correlation/covariance-based models is built by Forbes and Rigobon (2002) to capture stock price co-movements during the 1997 East Asian crises, the 1994 Mexican peso collapse, and the 1987 US stock market crash.

[2]In the following section of 2.2, the proposed contagion models will be specifically reviewed.

2.2.1 Volatility based CM models

One of the essential drawbacks of both the latent factor models and the correlation- or covariance-based models is that neither provides an empirical solution to the classification of the crisis and non-crisis periods, especially for small samples. To address this issue, the Markovian Switching ARCH (SWARCH) model (Hamilton, 1994) provides a synchronized solution to estimate contagion effect together with classifying crisis and non-crisis periods. According to Hamilton (1989), the SWARCH model suggests a probabilistic mechanism for dating the tranquil and turmoil periods by calculating the filtering and smooth probabilities of high and low volatility states of financial returns. This mechanism thus transfers the volatility dynamics into a visualized mapping of the two regimes. Hamilton and Gang (1996) extend the model to the bi-variate case, wherewith the initial purpose of studying the relationship between the business cycle and the stock market. The so-called Bi-SWARCH is then applied to detect contagion effects across paired stock/equity markets (Ramchand and Susmel, 1998; Edwards and Susmel, 2001), based on the argument that it captures the cross-market state-varying covariance and correlations according to the volatility regimes.

2.2.2 Decomposing tail dependence in CM

To exploit the risk transmission channel across different markets, hybrid contagion models integrating (vine) copulas are widely used to decompose the tail dependence of asset returns into (tree-structured) conditional bi-variate dependence. The flexibility that copulas offer to investigate the joint tail behavior makes it tempting in many studies that intend to discuss financial contagion matters, such as Sriboonchitta et al. (2014) (which adopts the copula-based GJR-GARCH model to analyze the financial risk and co-movement of stock markets across three Southeast Asian countries of Indonesia, Philippine and Thailand), Aloui and Ben Aïssa (2016) (which applies the vine copula pair-wisely and finds significant evidence of symmetric linkages across oil, stock and exchange rate markets in the U.S.), and Fink et al. (2016) (which hires Markov-switching R-vine models to investigate the existence of different

global dependence regimes). Rodriguez (2007) applies the SWARCH model with copulas to stock indices from five East Asian countries during the Asian Crisis and from four Latin-American countries during the Mexican Crisis, aiming to analyze the joint tail behavior of extreme returns in the characterization of non-linearity and asymptotic dependence during the contagion episodes. Moreover, a regime-switching vine copula approach is adopted by BenSaïda (2018) to study the contagion effect in European sovereign debt markets, whereas BenMim and BenSaïda (2019) employ a regular vine copula approach to model the dependence dynamics between major American and European stock markets by distinguishing the effects during crisis and tranquility periods.

2.2.3 Modeling heavy tails in CM

Considering the heavy tail pattern of turmoils, the extreme value theory (EVT), an important approach to estimate the non-Gaussian tails of financial series, is widely applied in the study related to extreme risks. The Generalized Pareto Distribution (GPD), a semi-parametric approach that is widely adopted to estimate the heavy-tail of financial series, not only reduces the computational cost in parameter inferences also provides a more robust fitting, especially for small-sized datasets. McNeil and Frey (2000) propose a hybrid model of the GARCH model (that estimates the current volatility) and EVT (that describes the tail distribution of innovations from the GARCH model) to estimate tail-related risk measures. Their result shows that the combined model gives better performance than the methods that neglect the heavy tail. Allen et al. (2013) further hire the combined model, GARCH-EVT, to estimate the volatility indexes, FTSE100 and S&P500, for the U.K. and the U.S. stock market, and find the evidence to support hedging strategies in extreme market conditions (such as the Global Financial Crisis that commenced in 2007). Samuel (2008) develops a new conditional extreme value theory-based model that incorporates the SWARCH to forecast extreme risks in the U.S. and China's stock market and suggest that the EVT-SWARCH model outperforms the pure SWARCH and GARCH models in capturing the non-normality and providing accurate VaR forecasts in both in-sample and out-sample tests. Wang et al. (2021) use a dynamic mixture

copula-extreme value theory model (DMC-EVT) to study contagion channels as elucidating the complex and dynamic dependence between forex markets, and their results show that the DMC-EVT model outperforms the alternative copula models.

The works mentioned above mainly stem from the objectives related to enhancing the model accuracy of fitting and explaining contagion per se, which commonly serves for the purpose of risk control in crises. The contagion effect across financial markets of different regions, though have been emphasized by most of the researchers, **few studies mine insights into the crisis-time risk transmission among different financial sectors in a country**, which is deemed as a crucial factor that drives the diversification of domestic portfolios as well as the financial stability of a country. Besides, according to our survey, the majority of existing literature in this area concerns more about modeling the linkage between the real estate and other financial markets for developed countries, such as the U.S. (Chan et al., 2011; Hui and Chan, 2014; Hoesli and Reka, 2015; Caporin et al., 2019; Tiwari et al., 2019). However, **whether these risk transmission models inferred linkage results for developed countries are meanwhile comparatively applicable to China remains to be verified**.

2.3 Forecasting system for financial crises

2.3.1 Crisis predictability

Before constructing the forecasting system to predict a financial crisis, we should figure out whether the financial crisis is predictable. In fact, the predictability of the financial crisis is an open debating question and yet settled through transcenturies arguments among economists and practitioners. The core of controversies focuses on whether the crisis is triggered by unexpected exogenous factors or the market endogenous instability.

Malkiel and Fama (1970) propose the Efficient Market Hypothesis (EMH) theory in 70's of last century, which says the financial crises occur as long as external shocks (known as 'black swan' events) arise, given that the securities prices fully

reflect all available information in a perfect market. Malkiel and Fama (1970)'s theory shows strong assumption on human behaviors' rationality, however, market participants are hardly to be always rational and quickly respond to market information in practice. Arthur (1999) then relieves the strong assumption by pointing out the evolution process of the economy is dynamic, non-linear, and uncertain, which also inherently coincides to Schiller (2000)'s revolutionary proposal of the endogenous market imperfection being mainly induced by the irrational factors. Sornette (2009) formally advocates the argument that crises are the predictable 'dragon-king' not the unpredictable 'black swan' since the precursory symptoms such as substantial outliers can be observed in reality before the crisis abruptly bursts. It thus provides teeming evidence to acknowledge the crisis predictability with a positive attitude given a series of rigorous statistical tests on predicted results.

Even though perfectly grasping the exact time point of crises seems less attainable in practice, **both theory holders admit the fact that gaining the awareness on precursory turbulence signals for financial shocking is achievable**. Such admitted fact hence renders the essence of most arguments on fertilizing and flourishing the financial crisis forecasting system development to dilute the assertion that crises are unpredictable.

2.3.2 Financial early warning system

Financial early warning systems (EWSs) are designed to forecast crises via studying historical pre-turmoil patterns, thus allowing market participants to take early actions to hedge against vital risks. In practice, the target of early warning ranges from individual financial markets, such as the banking sector, the currency, and stock markets, to the entire economic system. The modeling of crises is then commonly formulated as a classification problem based on the identified crisis indicators.

2.3.3 Crisis identification in EWS

There are usually two ways being adopted to identify the crisis observations in the previous works - on the basis of either expert opinions or an indicator function

describing the market "crash".

The former approach of defining crises through the survey among experts is widely used in the early studies of EWS, especially those concerning banking and debt crises (Kaminsky and Reinhart, 1999; Kaminsky, 2006; Reinhart et al., 2011; Reinhart and Rogoff, 2013; Caprio and Klingebiel, 2002; Valencia and Laeven, 2008; Laeven and Valencia, 2010, 2012; Detragiache and Spilimbergo, 2001; Yeyati and Panizza, 2011). Despite that the expert-defined crises are considered to be reliable for long-term predictions (Oh et al., 2006), **this paradigm fails to offer an efficient modeling solution as the frequency of observation increases.**

The latter one, the indicator functions based on a pre-specified crisis threshold, are more frequently used to define currency or stock market crashes. Reinhart et al. (2011) define a currency crisis as the excessive exchange rate depreciation exceeds the threshold value of 15%. More often, the Financial Pressure Index (FPI) is used to measure the gross foreign exchange reserves of the Central Bank and the repo rate (Sevim et al., 2014). Currency crises are thus identified as the FPI raises more than 1.5 (Kibritcioglu et al., 1999), 2 (Eichengreen et al., 1995; Bussiere and Fratzscher, 2006), 2.5 (Edison, 2003) or, 3 (Kaminsky and Reinhart, 1999; Berg and Pattillo, 1999; Duan and Bajona, 2008)) standard deviations from the long-term mean. Such crisis indicator index compounding fixed thresholds method is also applied for identifying stock crises. Specifically, stock crashes are indicated by the CMAX index falling below its mean by 2 (Coudert and Gex, 2008), 2.5 (Li et al., 2015), or 3 (Fu et al., 2019) standard deviations.

In the crisis indicator function, two major drawbacks emerge in the practical aspect. First of all, despite that, the paradigm of handling crises as crashes captures the associated acute loss, **fails to consider the extreme risk that comes along with the volatility jump**. Moreover, the threshold selection criteria should be handled more delicately, taking into account **the trade-off between missing crises and false alarms resulted from over-/under-estimated thresholds** (Babecký et al., 2014).

Back to our aim - constructing the EWS for China, more disadvantages will appear as adopting two mentioned identification methods. On the one hand, **the crisis definition based on collecting experts' opinion is hard to achieve**

in China since 1) there is not an officially published database to be the open-source for evidencing China's specific market crises, and 2) the process of collecting experts' opinions to establish such database analyzing the instability of various China's markets for different periods is labor-intensive and time-consuming, as well as unlikely to be timely updated in high frequency. On the other hand, **the development degrees for markets are inconsistent in time horizon** (for example, the stock market performs greater maturity and openness than the bond), which leads the fluctuating amplitude of different market prices to present disparate characteristics (for example, the price dynamics fluctuation amplitude in the stock market are greater than that of more secured assets of bonds). Thus, the singular way[3] to cut the high quantile observations as the extreme crisis events cannot satisfy the personalized distribution patterns for each market.

We opt to define the target variable for crisis identification in the EWS frameworks by inspecting on the volatility since the index volatility has been proven efficient in both predicting the future state for specific financial markets (Harvey and Whaley, 1992; Fleming, 1998; Granger and Poon, 2003; Christiansen et al., 2012) and constructing the comprehensive crisis indicator (Kim et al., 2004a,b; Kim, 2013) to produce the warning signal for predicting country-specific financial crisis. Such works support the rationality of using volatility to define the target variable for the early warning system. As analyzed in Section 2.2, Hamilton and Susmel (1994) propose an appealing way to classify the volatility into different states and visualize the process of volatility state transition in Markov regime switching frameworks. The specific formulation process will be displayed in Chapter 3, Section 3.1. Note that the volatility based binary crisis indicator function cannot escape from the pain of cutoff selection either, and it moreover bears the extra problem of distinguishing volatility levels (i.e. regime count determination problem). We will explore the methodologies to relieve such pains relating to cutoff and volatility regime selections in Chapter 5.

[3]Based on the crisis binary function, the cutoff usually a fixed value by taking the mean of crisis indicator index with one or more than one standard deviations.

2.3.4 Predictive models in EWS

In terms of the predictive model, three types of methods are commonly applied to generate early warning signals for crisis predictions, namely the parametric logit-probit regression (Frankel and Rose, 1996; Eichengreen and Rose, 1998; Demirgüç-Kunt and Detragiache, 1998; Bussiere and Fratzscher, 2006; Beckmann, 2007), the non-parametric signaling approach (Kaminsky and Reinhart, 1999; Kaminsky, 1998; Berg and Pattillo, 1999; Davis and Karim, 2008) and machine learning-based models (Nag and Mitra, 1999; Kim et al., 2004a; Celik and Karatepe, 2007; Yu et al., 2010; Giovanis, 2012; Sevim et al., 2014).

According to the order in which they are proposed, we divide them into two generations. The first two types of models are mainly based on econometric and statistical approaches, and being started up in the late 1990's, we categorize them into the first-generation model. The second-generation model, namely stylized machine learning techniques, are initiated in the contemporary data science backgrounds considering the model specializes in predicting big size non-linear data. The impetus brought by these state-of-art models accelerates the progress of predicting financial crises in the data-driven situation.

Generation I: empirical models

The logit-probit regression is a widely applied statistical model that makes crisis predictions based on the leading factors among the pooled data of economic variables. The univariate probit regression is firstly adopted in Frankel and Rose (1996) to characterize currency crashes on the annual data of one hundred developing countries from 1971 to 1992. The multivariate-probit model proposed by Eichengreen and Rose (1998) cooperates the macro-economic factors to improve the forecasting accuracy on predicting banking crises for one hundred developing countries during 1975-1992. The multivariate-logit binomial model, which is applied in Demirgüç-Kunt and Detragiache (1998) to study the factors that led to the emergence of systemic banking crises from 1980 to 1994 for more than forty developed and developing countries. The extended works for the multivariate-logit model in Bussiere and Fratzscher (2006) proposes a multinomial early warning system to address the post-crisis bias problem that comes along with the

binomial discrete-dependent-variable model, and the proposed multinomial-logit model outperforms in predicting financial crises for emerging markets. As for application in specific markets, Coudert and Gex (2008) use logit and multi-logit models to predict stock and currency market crises and find the leading effect of risk aversion indicators in the stock market crisis warning. The significance of S&P 500 futures and options in predicting stock market crashes is also shown by Li et al. (2015) in the study based on a logit model. By combining the logit model and Ensemble Empirical Mode Decomposition, Fu et al. (2019) recently developed an EWS for daily stock crashes using daily stock market valuation and investor sentiment indicators and achieved good in-sample and test-set results.

The KLR model (Kaminsky and Reinhart, 1999) explores signals of increasing exposure to risks in the financial system by studying the deviation of variables with pre-specified thresholds. This method is firstly applied to predict currency crises for twenty countries in the period of 1970-1995. The KLR indicator is extended in Kaminsky (1998) by considering more variables associated with risks. Berg and Pattillo (1999) further include the correlations across variables and develop the composite KLR indicator in the emerging markets. Davis and Karim (2008) compare between the KLR and logit regression methods and argue that the multinomial logit model is more reliable for predicting global banking crises, and the KLR indicator approach is for constructing country specific early warning system. Peng and Bajona (2008) apply the KLR approach to study the probability of China suffering the currency crisis and successfully label two risky high periods of currency devaluation.

The vulnerability of crisis, however, cannot be convincingly revealed by the first generation of models cluster, as IMF published paper (Berg and Pattillo, 1999) summarized according to the exercised results on the 1997 Asian currency crisis, **their predicting power is barely satisfactory unless under plausible modifications**.

Generation II: stylized models

Machine learning techniques (Oh et al., 2006; Celik and Karatepe, 2007) are then locked to construct the second generation of early warning systems considering

its specialization in predicting big size non-linear data. Chamon et al. (2007) attempts the binary classification tree (BCT) to predict capital account crises for emerging markets and finds the model though is able to capture some signs of crisis, but the overall predicting effectiveness is unsatisfactory. Nag and Mitra (1999) construct the early warning system by artificial neural networks (ANNs) to forecast the currency crisis for developing countries. The ANNs model, in general, bring more robust and effective predictions than classic statistical methods in dealing with linear and non-linear time series. The back-propagation neural networks (BPNN) and feed-forward neural networks (FANN) are then applied to predict the financial and banking crises for Asian countries (Kim et al., 2004a; Celik and Karatepe, 2007; Yu et al., 2010), extend the ANNs model to multiple-scale and forecast the currency crisis by proposing an empirical decomposition learning paradigm. In addition, Giovanis (2012) develops an EWS based on the neuro-fuzzy inference to examine the U.S. recessions in the period 1950-2010.

The deep learning (DL) techniques revolute the machine learning (ML) model by empowering the algorithm in a rapid and accurate way. Its fast development has appealed to a lot of studies in solving the economic issues, especially for asset market predictions (Jiang, 2021). Chong et al. (2017) investigate the advantages and drawbacks of deep learning for stock market analysis and prediction using five-minute intra-day data from the Korean KOSPI stock market and prove the deep learning shows general effectiveness in predicting returns and risks. Long et al. (2019) suggest a multi-filters neural network (MFNN) model by integrating convolutional and recurrent neurons in a multi-filters network structure to predict stock movements on Chinese stock market index CSI 300, and such state-of-art model is proven to outperformed the ordinary machine learning and statistical models by at least 11.47% and 19.75% respectively.

Among all DL models, the recurrent neural networks (RNNs) (Jordan, 1997) stand out since it not only allows the time-dependency information injection from previous moments but endows the networks with 'memory' functions from the past neuron inferred content as well. Sezer et al. (2020) surveys the searchable studies of deep learning techniques for financial time series prediction and finds the RNN based DL models (LSTM and GRU included) are the most com-

mon among studies. As the special type of RNNs, the long-short term memory (LSTM) network in Hochreiter and Schmidhuber (1997) is further designed to learn both long- and short-term dependencies for sequential forecasting purposes. **This deep learning technique has been widely used for various predictions in financial engineering works**. Yan and Ouyang (2017) prove that the LSTMs have a better predicting performance on capturing complex features such as non-linearity, non-stationary, and sequence correlation of financial time series. More similar prediction works such as Fischer and Krauss (2018) and Liu (2019) validate the LSTM out-performance on forecasting the financial series.

The LSTMs networks can be further hierarchically combined with traditional time series models to predict the price index volatility. Kim and Won (2018) develop hybrid models as combining the LSTM with three different GARCH-type models to forecast the volatility of stock index KOSPI 200 and find the GEW-LSTM outperforms others in terms of comparing error measurements. Hu et al. (2020) propose a hybrid method that combines the LSTM and ANN to forecast the copper price volatility and comparing to the GARCH combined LSTM model. It generates better volatility forecasts even though the configuration of ANN combined model should be fine-tuned according to the measure of prediction errors. Zolfaghari and Gholami (2021) further extend the hybrid LSTM literature by combining the adaptive wavelet transform (AWT), LSTM and ARIMAX-GARCH family models, and AWT, LSTM and heterogeneous autoregressive model of realized volatility (HAR-RV) model to predict the U.S. stock price index and volatility respectively. Their results indicate the AWT-LSTM-ARMAX-FIEGARCH model performs more robust than benchmarks in the prediction of different size time horizons (from 1-day to 60-days ahead).

The dual-layer information processing mechanism embedded in deep neural networks is first proposed by Schuster and Paliwal (1997), who makes the information flow in RNNs frame **with both forward and backward directions as adding the extra backtracking layer to boost the 'memory' power in a more substantial way**. As the explored form of bidirectional RNN (Bi-RNN), BiLSTM is similarly structured and used in financial predictions. Ahmed et al. (2018) evaluate the BiLSTM and stacked LSTM by comparing with the bench-

mark LSTM networks on a public data set for stock market closing prices and conclude the BiLSTM is superior to the benchmark. Ren et al. (2020) establish a model based on the BIAS measured investors reaction-after-news index and the BiLSTM model to predict the short-term trend of stock prices using news text data and find the adopted model outperforms other models counterparts in terms of prediction accuracy.

Comparing to the logit-probit regression model and KLR indicator approach, the sole (deep) neural networks though greatly boom the forecasting power, barely make a comparable inference to illustrate the contribution degree of exogenous variables. To render the estimation on input variables importance available for neural networks but avoid the lengthy process of dropping out one feature variable for each time to draw the feature importance, **attention mechanism** (Bahdanau et al., 2014; Xu et al., 2015) from natural language processing (NLP) **is believed to evaluate the factors' contributing degree effectively for neural networks**. A few recently released publications have applied the attention model and validated its functionality in forecasting financial markets. Liu et al. (2018) adopt a two-level attention mechanism to quantify the importance of words and sentences in given news to predict stock movements. Hergott (2018) posted a blog of structuring an LSTM attention model to review the 2013 "Taper Tantrum" - the abrupt event that happened in the US bond market that the anticipation of a multi-trillion-dollar buyer disappearing brought a big selloff in bonds. Ouyang et al. (2021) investigate the attention mechanism based predictive model to study the systemic risk early warning of China and find the attention mechanism included LSTM deep neural network model is more accurate in all the cases. Thus, to make the proposed EWS model more comparable to KLR and parametric regression based in terms of weighing the exogenous variables, the attention mechanism will have coalesced with deep neural networks in our study.

2.4 Sovereign crisis prediction

Different from previous EWS studies for predicting sovereign crisis on the macroeconomic level, we intend to renovate the issue in terms of 1) classifying the

volatility levels to define the sovereign crisis on a daily basis, 2) quantifying the contagion dynamics as the input factors of EWS and 3) distinguishing the leading indicators from between the contagious factors and the traditional monetary policy related indicators. To rationalize the innovated sovereign crisis definition and specify the process of fusing the quantized contagion information into the EWS, one intact section is used to review the literature for sovereign crisis definitions and the sovereign crisis determinants from contagious sources.

2.4.1 Define the sovereign crisis

In most previous studies, the sovereign crisis is virtually identical to be called the sovereign debt crisis (A et al., 2012; Gerali et al., 2017; Beltratti and Stulz, 2019), especially during the hardship for Euro-zone since 2010 (Mody and Sandri, 2012; Ahmad et al., 2013; A and B, 2020). However, **the specific rising and fading dates of the sovereign debt crisis are difficult to be uniformly clarified** in the experts' reports and academic studies (Ciarlone and Trebeschi, 2005; Fioramanti, 2008; Roman et al., 2016; Reusens and Croux, 2017; Afonso et al., 2018) **because the causes for sovereign crises stuffs the subjectivity and diversity** (Meier et al., 2021), which leads to further confusion to judge such works' credibility.

The alternative substitution for the sovereign crisis is the sovereign default (Fuertes and Kalotychou, 2007; Correa et al., 2014), though this name is accused of losing the categorized accuracy (Julianne Ams and Trebesch, 2018) among *technical, contractual and substantive* types of defaults. The IMF once published several working papers (Schimmelpfennig et al., 2003; Pescatori and Sy, 2007) to distinguish between defaults and debt crises, by clarifying **the debt crises as "events occurring when either a country defaults or its bond spreads are above a critical threshold.", however, in practice, either default events or critical thresholds**[4] **are difficult to be uniformly standardized** (Sy, 2004).

In the perspective of quantifying the crisis in market price dynamics, some

[4]The threshold determination is a cost work because of the complexity and variety among global countries and economies' development levels and fiscal policies.

market-based indicators such as CDS (credit default swaps) spreads and sovereign bond spreads, are recognized as the good proxies[5]. **Their deficiencies, however, are embodied in 1) being interfered by the credit-rating agencies' subjectivity, and 2) ignoring the seniority differentiation across bond creditors**. To remove the subjectivity interference, new indicators of CCA (Contingent Claims Approach) (Souto et al., 2007) and DtD (distance to default) (Singh et al., 2021) are recently proposed to sign the sovereign bond market vulnerability to default. **In practice, however, totally purifying the credit-judgment subjectivity in the quantifying process for sovereign crises by creating a new proxy is pricey and demanding**.

In 2011, the World Bank encouraged countries, especially emerging countries, to open up the local bank debt markets, which triggers the sovereign bond markets' instability being in charge of multiple types of risk sources, especially being exposed to diversified risks from large international financial tsunami transmission. Such risk exposure led instability can be fully reflected by volatility fluctuations, which makes **the methodology based on volatility clustering more appropriate to mitigate the pain of quantifying sovereign crisis, since it not only mirrors the market turbulence, the macroeconomic fundamentals fluctuations but credit-rating announcement changes as well** (Genberg and Sulstarova, 2008; Nowak et al., 2011; Daude, 2012; Ribeiro et al., 2017; Chatterjee and Eyigungor, 2019; Raimbourg and Salvadè, 2020). Some studies for sovereign bond markets have taken advantages of such method, for example, Keddad and Schalck (2020) use Markov switching time-varying model to analyze the sovereign risk transmission across domestic banks and BenSaïda (2018) combine vine copulas with Markov regime switching model to study the contagious effects across advanced countries' sovereign bond markets.

In the Markovian frameworks, the time-varying probabilities for estimating the market price dynamically changed volatile states will be opted for quantifying the sovereign crisis by inspecting the volatility regime switching process of sovereign bond observations.

[5]Lina et al. (2013) has proven bond spreads are better proxy than CDS in affine model frameworks.

2.4.2 Determinants for sovereign crisis

The determinants for the sovereign crisis can be categorized into two groups.

The first category is the macro-economic factors, which successfully lock researchers' attention even earlier than the EU countries bursting sovereign debt crisis in 2010, **because macro-fundamentals' dynamics change often underlies antecedent clues for financial shocks**. Schimmelpfennig et al. (2003) ever identified the macroeconomic variables relating to solvency and liquidity factors to be the leading indicators for sovereign debt crises by applying both logit and binary recursive tree model frameworks to 47 economies over the period of 1970-2002. Fuertes and Kalotychou (2006) find factors of trade to GDP, external debt to GDP, official debt to total debt, IMF credit to exports, and credit to the private sector over GDP robust to foresee the sovereign default risks by analyzing 96 emerging/developing economies over the year of 1983–2002. As a chain reaction of the 08's global financial crash, the EU debt crisis made such studies more flourished. Besides the macroeconomic indicators relating to debit and credit level to GDP, the suspicious shifted inflation level and accumulated local governments' hidden domestic debts also precede the sovereign crises for both emerging and advanced countries (Reinhart et al., 2011; Dawood et al., 2017; Ghulam and Derber, 2018; Rho and Saenz, 2021). Meanwhile, the inflation procyclical dynamics reveal the transmitting channel from political uncertainty to sovereign default risks (A and B, 2014; Kraussl et al., 2016; Ghulam and Derber, 2018) as it has been proven sensitive to monetary policy shifts (Dongho, 2017).

The second cluster is the contagious factors or the risk transmitting factors. Considering the jumbled causes for contagion impacting on the sovereign bond, we therefore further group the contagious factors into two sub-branches of *geographical economies* and *trading markets*[6] to discuss their respective roles as follows.

1. *Geographical factors mainly cover two types: the neighborhood community (e.g. the EU) and major economies in the globe (e.g. the US).*

[6]The tight correlation between sovereign bond and banking system is also being buzzed (Alter and Beyer, 2014; Georgoutsos and Moratis, 2017), but the contagious direction is too intricate to be determined since the sovereign bond abnormal fluctuation is sometimes recognized as leading to the banking sector's uncertainty, especially in the stressful finance episodes (Yu, 2017; Keddad and Schalck, 2020).

The crisis spillover effects across continental Europe have been proven by plentiful studies. Arghyrou and Kontonikas (2012) proves the sovereign debt crisis can be divided into two subs, and the early stage is mainly driven by Greece while the latter one is multiple sourced contagion with prominent performers of GIPS[7] countries. Fernández-Rodríguez et al. (2015) further show that the core EU countries are the triggers in spillover effects while the peripheral countries are dominant transmitters during the crisis, whose conclusion corresponds to Kalbaska and Gatkowski (2012)'s claim that PIIGS[8] members have lower capacity to trigger contagion than core EU countries. To date the crisis phases and test contagious effects over Germany for other Euro-zone countries, Cronin et al. (2016) develops MS-VAR models on daily 10-year sovereign bond spreads database, and the results show though the risk transmission is mutual across PIIGS and the core group of EU, the market co-movements are more often due to inter-dependency rather than contagion. From the most recent study of Hamill et al. (2021), the global financial and European sovereign debt crises are implied not only altering the sentiment of the Euro-zone countries but deteriorating and even exacerbating the contentedness across Euro-sovereign bond markets as well. In addition, the role of developed countries in cross-regional contagion is worth being further explored since they conventionally most reign over the discourse power in global economies. For instance, Kim et al. (2015) shows that news from the three major economies of the U.S., the EU, and China has significant spillover effects on other national sovereign bond markets by investigating the asymmetric news effects on the pricing and volatility of sovereign CDS spreads. Nitschka and Thomas (2018) and Zhang et al. (2020) also reinforce such conclusion by studying the impact of the U.S. monetary policy on other countries' government bond returns and analyzing the sovereign contagious effects across between China and other countries, respectively.

2. *Trading markets primarily refer to the vibrant financial markets with strong*

[7]GIPS is the abbreviation for Greece, Ireland, Portugal, and Spain
[8]PIIGS is the abbreviation for Portugal, Italy, Ireland, Greece, and Spain

liquidity, such as stocks, foreign exchange rates, commodities, and oil.

Most recent studies reveal the correlated dynamics across the sovereign bond market and stocks, exchange rates, oil, and commodities. The stock market, as Sun et al. (2020) and Ibhagui (2021) concluded, occupies the dominant position in the interactive spillovers among either the sovereign CDS, stock and commodity markets in G7 and BRICS countries or the sovereign CDS, stock indexes, and currency-basis swaps in the Euro-zone, the U.K., Australia, and Japan. In the study of Bouri et al. (2019), the commodity and energy markets are further investigated and concluded that either commodity or energy market shapes the sovereign risks in middle and upper volatility quantile for exporters (e.g. Brazil and Russia), while for importers (e.g. China), the predictability is only significant in upper volatility quantile. According to Feng et al. (2021)'s exploration on constructing the forecast error variance (FEV) decomposition based spillover index to explore the "sovereign CDS-exchange rate" system, the exchange rate is proven to have a higher spillover effect on the sovereign CDS. The oil price/volatility shocks are proven positively correlated to sovereign risks. Such correlation evidently amplifies the sensitivity to structural breaks (Tule et al., 2017), but has an asymmetric presence in the shock transmission. For instance, Bouri et al. (2018) reveal that oil exporters of BRIC[9] countries are more sensitive to positive oil shocks, whereas oil importers perform in a contrary way.

For these two types of factors, the first one (macroeconomic factors) can be directly input in the EWS predictive models[10], while **the latter one (contagious factors) cannot swagger to the predicting module since the contagion information usually conveys both risks transmitting strength and spillover directions, which is difficult to be quantized as time series containing structural patterns**.

[9]BRIC = Brazil, Russia, India and China

[10]Most accessed macroeconomic factors are time series.

2.4.3 Contagion quantification for the EWS

To solve the obstacle of quantifying contagious intensity, empirical studies of detecting risk spillover effects will be referred to. We have mentioned in the last section of 2.2 that the contagion detection based on latent factor models and covariance-based models Dungey et al. (2005) are prone to infer the variables correlated structures, not appropriate to depict the time-varying dynamics for contagious intensity[11]. The Dynamic Conditional Correlation multivariate GARCH (DCC-GARCH) (Engle, 2003) provides a feasible way to detect possible changes in conditional correlations over time. As Celik (2012) concluded, **the DCC-GARCH model not merely continuously adjusts the correlation for the time-varying volatility but accounts for the standardized residuals' heteroskedasticity in a direct way as well**. Such attributes have been verified by many studies to demonstrate the co-movements across for both macroeconomic fundamentals (Jones and Olson, 2013) and financial markets of stocks (Syllignakis and Kouretas, 2011; Hou and Li, 2016), foreign exchanges (Gomez-Gonzalez and Rojas-Espinosa, 2019) and energy products (Hou et al., 2019; Chen et al., 2020). Besides, recent studies combine the DCC-GARCH with structural break models, such as regional factor model (Bonga-Bonga and Mabe, 2020) and the Diebold–Yilmaz model (Akhtaruzzaman et al., 2021) to assess the geographic shock transmission.

The DCC model quantized time-varying correlated intensity cannot be directly input as the risk transmission factors to the EWS unless being imported the crisis transmitting direction, in other words, **the DCC can infer the dynamic changes of correlated linkages across markets but hardly affords to point out the crisis originator**. Being sparkled by the Hamilton (1989)'s work, which model infers the probability of market being in the crisis, the DCC produced correlation coefficients can be **pairwise combined with the SWARCH inferred probability of contagious source markets being in crisis periods to form the final contagious intensity dynamics**, which not only reflects the complete contagion information but accesses the EWS in-

[11] Almeida and Czado (2012) though allows the time-varying inference on the dependent structure across variables, the full transmitting pattern is so complex as to bring computing intractability in practice.

put layer directly as well. The specific exploration process for constructing such contagious intensity dynamics will be described in Chapter 6, Section 6.2.

2.5 Gaps in the literature

The research gaps for CM and EWS development that have been found in the review process will be summarized as follows.

1. **For both CM and EWS, the crisis identification is problematic.** The crisis has been numerously defined or identified in the CM and EWS literature, but the studies are still unsatisfactory in terms of defining for specific markets on high-frequency data (usually daily basis) and discussing the challenge of threshold selection and volatile levels determination, especially relating to the EWS crisis classifier's effectiveness in practice.

2. **For CM, the correlated structure across domestic asset markets should be hierarchically decomposed but not be singularly explained by a simple model.** The complexity risk transmitting channels are more appropriately modeled by hierarchical strata frameworks to implement dual aims of visualizing the contagion episodes (pairwise) and figuring out the contagious paths. Rare literature, to the best of our knowledge, neither distinguishes the difference between such two types of hybrid models nor accomplishes comparative analysis for domestic assets between China and the developed country.

3. **For EWS, each component module in the frameworks, especially the predictive model, is yet to be fully explored.** A great number of studies have developed the EWS model, few discuss each of the functional parts in an integrated framework. Furthermore, the competitive horse race among empirical and stylized predictive model generations continues, especially after the deep learning techniques being flourished in financial studies. Meanwhile, the exploration on developing the EWS that fuses the contagion information and draws the leading indicators with statistical significance testifying is still in the bud.

Chapter 3

Methodological preliminaries

Before formally introducing the CM and EWS development process, the knowledge basis for methodological preliminaries in the project will be first presented here. In this chapter, we will first formulate the crisis detection based on volatility regime switching principles in Section 3.1, and then detail the bi-variate SWARCH model, heavy tail modeling technique of EVT, and multivariate correlated structure model of vine copulas in Section 3.2 to make preparations for constructing two hybrid CMs. Last, the stylized predictive model of deep neural networks of LSTM, which are used in further EWS development, will be displayed after sketching the basic neural networks in Section 3.3.

3.1 Volatility based crisis identification

3.1.1 SWARCH model

Different from the widely-used autoregressive integrated moving average (ARIMA) models that impose stationarity to time series, the autoregressive conditionally heteroscedasticity (ARCH) class of models, including the generalized ARCH (GARCH) model, the exponential GARCH (EGARCH) model, and the quadratic GARCH (QGARCH) model, etc., deal with non-stationary data by exhibiting time-varying volatility. The univariate SWARCH model is stylized to fit innovations by clustering states of different volatility regimes (Hamilton, 1989) and is further applied to identify possible crisis episodes (A and B, 1994).

In general, an AR(p)-SWARCH(K,q) model, where the AR(p) component

captures the long-term stationary part of the process, can be expressed as follows.

$$y_t = u + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \cdots + \theta_p y_{t-p} + \epsilon_t, \epsilon_t | \mathcal{I}_{t-1} \sim N(0, h_t), \tag{3.1}$$

$$\frac{h_t^2}{\gamma_{s_t}} = \alpha_0 + \alpha_1 \frac{\epsilon_{t-1}^2}{\gamma_{s_{t-1}}} + \cdots + \alpha_q \frac{\epsilon_{t-q}^2}{\gamma_{s_{t-q}}}, s_t = \{1, \ldots, K\}, \tag{3.2}$$

where u is the constant, $\{\theta_1, ..., \theta_p\}$ and $\{\alpha_0, \alpha_1, \ldots, \alpha_q\}$ denote the parameters to be estimated, s_t is the unobserved state variable with K different regimes and γ_{s_t} is the scaling parameter which controls the variance process. The residual term ϵ_t conditional on one-step back information is assumed to follow a normal distribution with mean 0 and variance h_t. The latent state variable s_t follows a Markov chain with the transition probabilities of

$$p_{ij} = Prob(s_t = j | s_{t-1} = i), i, j = \{1, 2, ..., K\}. \tag{3.3}$$

Hence the full structure of transition matrix is

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} & \cdots & p_{K1} \\ p_{12} & p_{22} & \cdots & p_{K2} \\ \vdots & \vdots & \ldots & \vdots \\ p_{1K} & p_{2K} & \cdots & p_{KK} \end{bmatrix} \tag{3.4}$$

with the constraint $\sum_{j=1}^{K} p_{ij} = 1$. In this way, s_t governs the conditional density of the observation vector $\mathbf{y}_t$ as

$$f(\mathbf{y}_t | \mathscr{Y}_{t-1}; \boldsymbol{\theta}) := f(\mathbf{y}_t | s_t, s_{t-1}, \ldots, s_{t-q}, \mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_0; \boldsymbol{\theta}), \tag{3.5}$$

where $\boldsymbol{\theta}$ is the vector of all model parameters.

To implement the *Maximum Likelihood Estimation (MLE)* to estimate the parameters, the log-likelihood function $\mathscr{L}(\boldsymbol{\theta})$ is given by

$$\mathscr{L}(\boldsymbol{\theta}) = \sum_{t=1}^{T} log f(\mathbf{y}_t | \mathscr{Y}_{t-1}; \boldsymbol{\theta}) \tag{3.6}$$

$$= \sum_{t=1}^{T} log\big(\sum_{i=1}^{K} P(s_t = i | \mathscr{Y}_{t-1}; \boldsymbol{\theta}) f(\mathbf{y}_t | s_t = i, \mathscr{Y}_{t-1}; \boldsymbol{\theta}) \big), \tag{3.7}$$

where $P(s_t = i|\mathscr{Y}_{t-1}; \boldsymbol{\theta})$ is the prediction probability at time t based on observations until $t-1$, and can be computed based on the filtering probability.

In previous studies, the regime cases are customarily divided into two volatility states - high and low, to make the most common sense of capturing the dynamic of financial return oscillations. The model of AR(1)-SWARCH(2,1) is thus specified as follows.

$$y_t = u + \theta_1 y_{t-1} + \epsilon_t, \epsilon_t|\mathcal{I}_{t-1} \sim N(0, h_t), \tag{3.8}$$

$$\frac{h_t^2}{\gamma_{s_t}} = \alpha_0 + \alpha_1 \frac{\epsilon_{t-1}^2}{\gamma_{s_{t-1}}}, s_t = \{1, 2\}, \tag{3.9}$$

where $s_t = 1$ indicates the low volatility state, and $s_t = 2$ indicates the high volatility state.

In order to visualize the regime switching dynamics of states and to date the crisis episodes, the filtering and the smooth probabilities, written as $P(s_t = i|\mathscr{Y}_t; \boldsymbol{\theta})$ and $P(s_t = i|\mathscr{Y}_T; \boldsymbol{\theta})$, need to be computed. By definition, the filtering probability is the conditional probability based on the current information, whilst the smooth probability is the inferred probability according to the full length of observations (Kim, 1994; Kuan, 2002). The observation at time t will be judged locating into high or low volatility states as its corresponding either filtering or smooth probability is greater than the empirical value of 0.5[1].

3.1.2 Further variations on SWARCH classification

After solving the crisis identification problem in clustering financial series volatility, two questions intuitively crop up.

1. Is the fixed cutoff value of 0.5 being 'omnipotent' or the 'Jack-of-all-trades' for all SWARCH model based crisis identification cases?

2. Whether two regimes are enough to distinguish the volatility levels for all return dynamics from different markets?

To answer the questions, we explore a further step for each EWS development of Chinese stock and bond markets according to their price featured patterns,

[1]R package MSGARCH (Ardia et al., 2017) is applied to model the univariate SWARCH.

respectively. Specifically, the cutoff value will be sprinkled by introducing a variation scheme to gain the SWARCH model based crisis classifier's adaptability to different windowed samples in stocks; and then a trial of varying the regime count will be experimented to find the optimal division way to distinguish the dynamics fluctuating levels especially for China's bonds, which market volatility is not significant enough for a long period because of the monetary policy forced intervention[2]. The methodologies of two-peak dynamically thresholding and RCM determination for attaining the question points of 1 and 2 will be described in Chapter 5 for stock and bond markets' application backgrounds, respectively.

3.2 Joint tail behaviour decomposition

3.2.1 Bivariate SWARCH model

The SWARCH model is further extended to the bi-variate case given a potential volatility originator[3] in each couple of markets. It is necessary to specify the crisis originator as investigating their volatility linkage to further explore the existence of a reversed risk transmission per se. Given two-state specification of the the univariate SWARCH settings for the high- and low- volatility cases (i.e. $s = 1$ for low volatility, and $s = 2$ for high volatility.), the Bi-variate SWARCH (Bi-SWARCH) model will include four primitive states that can be written as below,

$$
\begin{aligned}
s_t = 1 &: s_{o,t} = 1, s_{r,t} = 1; \\
s_t = 2 &: s_{o,t} = 1, s_{r,t} = 2; \\
s_t = 3 &: s_{o,t} = 2, s_{r,t} = 1; \\
s_t = 4 &: s_{o,t} = 2, s_{r,t} = 2,
\end{aligned}
$$

where s_t denotes the primitive states at time t for the Bi-SWARCH model, $s_{o,t}$ and $s_{r,t}$ are the univariate states for the crisis originator and the recipient, re-

[2]China's bond market prices, before the interest rate liberalization being proposed, are strongly regulated by the national department that sets the financial policy.

[3]The volatility originator refers to the market where the crisis is originated and spills over to other markets (Ramchand and Susmel, 1998; Edwards and Susmel, 2001).

spectively. The transition probabilities of p_{ij} are rather flexible given different imposed restrictions on assumptions. For example, $p_{ij} = p_{ii}^{o} p_{ij}^{r}$ means there are two independent volatility states for the paired markets (i.e. the state independence assumption), while $p_{ij} = 0, for\ all\ i, j \in [2, 3]$ assumes that two volatility states are shared with each other in the coupled markets (i.e. the volatility synchronization assumption).

The Bi-SWARCH(2,1,2) model with 2 regime states for each variable (the first 2), 1 time lag in the variance (the middle 1) and 2 covariances (the last 2) can be written as

$$\begin{bmatrix} y_{o,t} \\ y_{r,t} \end{bmatrix} = \begin{bmatrix} u_o \\ u_r \end{bmatrix} + \begin{bmatrix} \theta_o \\ \theta_r \end{bmatrix} \begin{bmatrix} y_{o,t-1} \\ y_{r,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{o,t} \\ \epsilon_{r,t} \end{bmatrix}, [\epsilon_{o,t}, \epsilon_{r,t}]' | \mathcal{I}_{t-1} \sim BN(0, \boldsymbol{H}_{t,s_t}), \quad (3.10)$$

where BN is the bivariate normal distribution. And the time-varying state-dependent conditional variance-covariance matrix $\boldsymbol{H}_{t,s_t}$ is structured as follows,

$$\boldsymbol{H}_{t,s_t} = \begin{bmatrix} h_{o,t}^2 & \rho_{s_t} h_{o,t} h_{r,t} \\ \rho_{s_t} h_{o,t} h_{r,t} & h_{r,t}^2 \end{bmatrix}, \quad (3.11)$$

with the conditional variance terms

$$\frac{h_{o,t}^2}{\gamma_{o,s_t}} = \alpha_{o,0} + \alpha_{o,1} \frac{\epsilon_{o,t-1}^2}{\gamma_{o,s_{t-1}}}, \quad (3.12)$$

$$\frac{h_{r,t}^2}{\gamma_{r,s_t}} = \alpha_{r,0} + \alpha_{r,1} \frac{\epsilon_{r,t-1}^2}{\gamma_{r,s_{t-1}}}, \quad (3.13)$$

and the state-dependent correlation coefficients ρ_{s_t}. The value of ρ_{s_t} is assumed to be dependent on the state of crisis originator, thus can be more precisely denoted as $\rho_{o_{s_t}}$. As argued in Bollerslev (1990), this simplification not only increases the efficiency of converging the maximum likelihood, but facilitates the hypothesis testing for contagion as well. That is, if $\rho_{o_{s_t}}$ in the high volatility state is proved to be significantly higher than that in the low volatility state, the contagion effect is indicated.

3.2.2 EVT model for heavy tails

The abnormality of financial series is performed as heavy tails. In the decomposition process, to estimate the heavy tails of univariate asset, the extreme value theory (EVT) method will be adopted. According to the Peak-Over-Threshold (POT) method, exceedances over the thresholds can be captured by a generalized Pareto distribution (GPD), and the margin for standardized residuals z_t[4] can be expressed as follows,

$$F(z_t) = \begin{cases} \frac{N_L}{N}\left(1+\xi_L\frac{u_L-z_t}{\beta_L}\right)^{-\frac{1}{\xi_L}}, & z_t < u_L \\ F_{em}(z_t), & u_L < z_t < u_R \\ 1-\frac{N_R}{N}\left(1+\xi_R\frac{z_t-u_R}{\beta_R}\right)^{-\frac{1}{\xi_R}}, & z_t > u_R, \end{cases} \tag{3.14}$$

where u_L, u_R are the lower and upper thresholds assigned to be the 10% and the 90% percentiles respectively. ξ_L, ξ_R and β_L, β_R are coefficients of the GPD. N is the sample size. N_L and N_R are the counts of exceedances beyond the lower and upper thresholds. F_{em} is the empirical distribution of z_t.

Coalescing with the Bi-SWARCH model and the parametric EVT model, one of our hybrid contagion models, Bi-SAWRCH-EVT model is thus designed to measure the joint tail behavior of paired asset returns in the occurrence of extreme events. The corresponding implementation diagram will be drawn as a flowchart in Chapter 4, Section 4.1.1. The other hybrid contagion models, which contribute more to understand the risk transmitting paths across multiple markets simultaneously, rely on the copulas.

3.2.3 Vine copulas

Copulas are believed to excel in capturing the non-linear correlation with its inhabited asymptotic tail-dependence property. Sklar's Theorem (Sklar, 1959) provides the basis of copulas and argues that the n-dimensional distribution $F(x_1, x_2, \ldots, x_n)$ with margins of $F_1(x_1), \ldots, F_n(x_n)$ can be decomposed as fol-

[4] z_t is the innovated term of the $SWARCH$ model, and satisfies $\epsilon_t = z_t h_t$.

lows with a n-copula.

$$F(x_1, x_2, \ldots, x_n) = C_{1,\ldots,n}(F_1(x_1), \ldots, F_n(x_n)), \tag{3.15}$$

where the margins F_i and the copula $C_{1,\ldots,n}$ are differentiable and the multivariate density function$f(x_1, \ldots, x_n)$ can be written as,

$$f(x_1, \ldots, x_n) = \frac{\partial C_{1,\ldots,n}(F_1(x_1), \ldots, F_n(x_n))}{\partial x_1 \cdots \partial x_n} \cdot f_1(x_1) f_2(x_2) \cdots f_n(x_n) \tag{3.16}$$

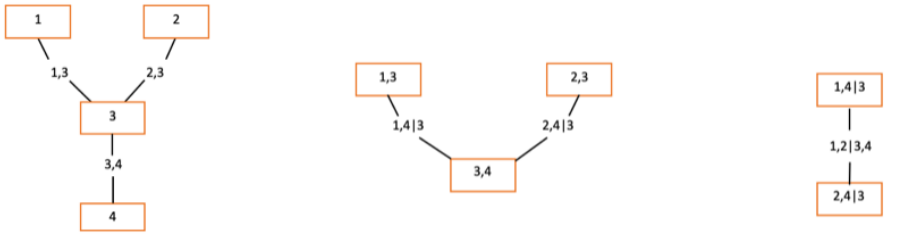
$$= c_{1,\ldots,n}(F_1(x_1), \ldots, F_n(x_n)) \cdot f_1(x_1) f_2(x_2) \cdots f_n(x_n). \tag{3.17}$$

The multivariate distributions are thus decomposed in the product of multiple copulas and individual marginal functions where the non-linear dependence across multiple financial series can be elaborately modeled.

As a widely applied family of copula functions, vine copulas provide a prominent solution to the inflexibility raised in handling multivariate data and allow for a wide variety of dependence structures in a multidimensional scale. The regular vine (R-vine) copula (Bedford and Cooke, 2002), one of the most powerful copula families that recognizes the complex patterns of dependence by applying R-vine decomposition, breaks down the multivariate probability density into bi-variate conditional copulas. Specifically, the full dependence structure across n variables is represented as a set of trees $\mathcal{V} = \{T_1, ..., T_{n-1}\}$, by regarding each variable as the node and the pairwise bi-variate copula dependency as the edge. To construct an R-vine copula, the first tree T_1 includes all variable nodes of $\mathcal{N}_1$ and all bi-variate edges of $\mathcal{E}_1$, respectively. Then, in the second tree, T_2, the number of nodes is counted as the number of edges in the first tree, i.e. $\mathcal{N}_2 = \mathcal{E}_1$ and the set of edges $\mathcal{E}_2$ contains all bi-variate dependencies conditioning on the common node that two nodes share in tree T_1. In a recursive way, the R-vine tree $T_{j,(j>2)}$ with nodes $\mathcal{N}_j = \mathcal{E}_{j-1}$ and edges $\mathcal{E}_j$ can be finally formed.

An example of the R-vine tree based on four variables of $\{1, 2, 3, 4\}$ is presented as Figure 3.1, where each copula edge is expressed in the comma-separated form of its connective nodes. The conditional structures are indicated by the conditional sign '|'.[5] There are three trees in the R-vine decomposition for four variables.

[5]Without loss of generality, the dependence structure formed by R-vine copulas is not pre-



(a) Tree 1 (b) Tree 2 (c) Tree 3

Figure 3.1: The example R-vine decomposition trees.

The nodes in the tree T_{j+1} are the edges of the tree T_j, and will be bridged by a copula edge if a common node is shared in the previous tree T_j. In our study, each node represents one market asset of interest, and the edges are estimated in bi-variate copula functions between two of them. The general formulae for the R-vine conditional marginal density can be written as follows,

$$f(x|\boldsymbol{v}) = c_{x,v_l|\boldsymbol{v}_{-l}}(F(x|\boldsymbol{v}_{-l}), F(v_l|\boldsymbol{v}_{-l})) \cdot f(x|\boldsymbol{v}_{-l}), \tag{3.18}$$

where $\boldsymbol{v}$ is the n-dimensional variable vector and $\boldsymbol{v}_{-l}$ denotes the variable vector excluding l^{th} element. Aas et al. (2009) and Almeida and Czado (2012) specialize the multivariate density formula for C-vine and D-vine copulas, as Eq. (3.19) - (3.20) show, to perform the different types of decomposed patterns for modeling the correlated structure. C-vine gives a star or radiation while D-vine shows a path, which means, in the sequential method for vine tree selection, we identify the root variable in C-vine trees by searching for strongest dependencies to all other variables (i.e. maximum column sum for a matrix of empirical Kendall's tau.) but should determine the order of variables by solving a Hamiltonian path.

$$\underset{C-vine}{f(\boldsymbol{x})} = \prod_{k=1}^{n} f(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|1,\ldots,j-1}(F(x_j|x_1,\ldots,x_{j-1}), F(x_{j+i}|x_{1,\ldots,x_{j-1}})) \tag{3.19}$$

$$\underset{D-vine}{f(\boldsymbol{x})} = \prod_{k=1}^{n} f(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i,i+j|i+1,\ldots,i+j-1}(F(x_i|x_{i+1},\ldots,x_{i+j-1}), F(x_{i+j}|x_{i+1,\ldots,x_{i+j-1}})), \tag{3.20}$$

determined. Figure 3.1 only provides one possible form as an example of the R-vine.

where j and i denote the j^{th} tree and the i^{th} edge respectively, c is the bi-variate copula density function. For n variables, the log-likelihood for joint distribution thus can be fully written as

$$l(\boldsymbol{x}) = \sum_{i=1}^{n-1} lnf_i(x_i) + lnc_{1,...,n}(F_1(x_1), F_2(x_2), ..., F_n(x_n)). \tag{3.21}$$

In the process of inferring vine copula structures for multiple variables, two central issues should be solved: i) select appropriate vine copula trees to structure dependence patterns. In other words, choose from R-vine, C-vine and D-vine to decompose the joint distribution; ii) select suitable bi-variate copulas to fit the conditional dependence for each subtrees. Dißmann et al. (2013) propose feasible selection criteria in practice, that is, the sequential method based on maximizing the sum of absolute values for empirical Kendall's taus by using spanning tree (MST) algorithm[6]. More specifically, we assume that the value for empirical Kendall's tau for the variable pair (also edge) $e_{i,j}$ in n variables is $\tau_{i,j}$, then maximize the sum of absolute values for all Kendall taus (see Eq. 3.22) to select the first spanning tree for all variables, then in the selected spanning tree, copulas and corresponding parameters will be selected and estimated by following Akaike and Bayesian Information Criteria (AIC and BIC) in maximum likelihood estimation. Note that, in this step, all opted copula families will be computed and made comparisons.[7] Then, this selection procedure for spanning tree and copulas will be iterated on all rest of conditional variable pairs for sub-trees till we reach the final sub-tree (where only one edge is left).

$$max \sum_{e_{i,j} \in \text{spanning tree}} |\tau_{i,j}|, \text{for} 1 \leq i, j \leq n. \tag{3.22}$$

The full options for copula families in our study, though have been detailed in the manual for R package 'VineCopula', such as elliptical (Gaussian and Student-t) and Archimedean (Clayton, Gumbel, Frank, Joe, BB1, BB6, BB7 and BB8) copulas that cover most dependence patterns. We summarize them in Table 3.1, where the number in the second column labels the corresponding copula family's tag, as shown in R package 'VineCopula'.

[6]MST is an important mathematical algorithm in Graph Theory

[7]A test for independence will be performed beforehand.

Copula	family number	par1	par2
Gaussian	1	(-1,1)	-
Student-t	2	(-1,1)	(2,+∞)
(Survival) Clayton	3,13	(0, +∞)	-
Rotated Clayton (90 and 270 degrees)	23,33	(-∞,0)	-
(Survival) Gumbel	4,14	[1, +∞)	-
Rotated Gumbel (90 and 270 degrees)	24,34	(-∞,-1]	-
Frank	5	$\mathcal{R}\backslash\{0\}$	-
(Survival) Joe	6,16	(1, +∞)	-
Rotated Joe (90 and 270 degrees)	26,36	(-∞,-1)	-
(Survival) Clayton-Gumbel (BB1)	7,17	(0, +∞)	[1, +∞)
Rotated Clayton-Gumbel (90 and 270 degrees)	27,37	(-∞,0)	(-∞, -1]
(Survival) Joe-Gumbel (BB6)	8,18	[1, +∞)	[1, +∞)
Rotated Joe-Gumbel (90 and 270 degrees)	28,38	(-∞,-1]	(-∞, -1]
(Survival) Joe-Clayton (BB7)	9,19	[1, +∞)	(0, +∞)
Rotated Joe-Clayton (90 and 270 degrees)	29,39	(-∞,-1]	(-∞, 0)
(Survival) Joe-Frank (BB8)	10,20	[1, +∞)	(0,1]
Rotated Joe-Frank (90 and 270 degrees)	30,40	(-∞,-1]	[-1,0)
(Survival) Tawn type 1	104,114	[1, +∞)	[0,1]
Rotated Tawn type 1 (90 and 270 degrees)	124,134	(-∞,-1]	[0,1]
(Survival) Tawn type 2	104,114	[1, +∞)	[0,1]
Rotated Tawn type 2 (90 and 270 degrees)	124,134	(-∞,-1]	[0,1]

Table 3.1: The copulas and their family number in R package of 'VineCopula'.

Thus, being combined with the fore-works of univariate SWARCH to detect the single crisis time horizon for one asset and EVT to model the heavy tails behavior, the other hybrid CM, namely paired SWARCH-EVT-Copula is constructed. Different from the Bi-SWARCH-EVT being competent to trace the contagious episodes between markets, the Copula based one underlines the structural analysis on inferring the transmission path.

3.3 Predictive models

As reviewed in the literature survey, the predictive models adopted for constructing the EWS are clustered into two generations - the empirical parametric (Eichengreen et al., 1995; Frankel and Rose, 1996; Bussiere and Fratzscher, 2006; Dawood et al., 2017) and non-parametric methods (Kaminsky and Reinhart, 1999; Berg and Pattillo, 1999; Lestano et al., 2004; Davis and Karim, 2008; Peng and Bajona, 2008), and the stylized machine learning techniques (Ahn et al., 2011; Sevim et al., 2014; Chatzis et al., 2018; Samitas et al., 2020). The machine

learning models are state-of-art techniques being more flexible than traditional econometric models to predict complex data with non-linearity. Being taken such advantages, a wide range of EWS models have constructed on machine learning techniques, such as neural networks (Nag and Mitra, 1999; Oh et al., 2006; Celik and Karatepe, 2007; Yoon and Park, 2014), decision trees (Tanaka et al., 2016; Holopainen and Sarlin, 2017), support vector machine (Ahn et al., 2011), and deep neural networks (Wang et al., 2020a,b; Ouyang et al., 2021), have been studied. This section will put more emphasis on neuron models to introduce the time-dependent deep neural networks of LSTM, the predictive model that is hired in our EWS development. For other machine learning models that are adopted as the baseline for comparison, their adept aspects will be detailed and analyzed in Appendix B.

3.3.1 Neural Networks

Big data science has unsealed a nova technology era. More powerful models are required to solve the non-linear problems with greater precision and less cost. Artificial neural networks (ANN) are born in this background. By far, despite brimming with disputes on transparency and interpretability, the neuron models are generally acknowledged as the most robust and flexible model for financial predicting work. Nag and Mitra (1999), Oh et al. (2006) and Fioramanti (2008) have successfully predicted the currency, stock and debt crises by hiring the feed-forward multi-layer perceptrons.

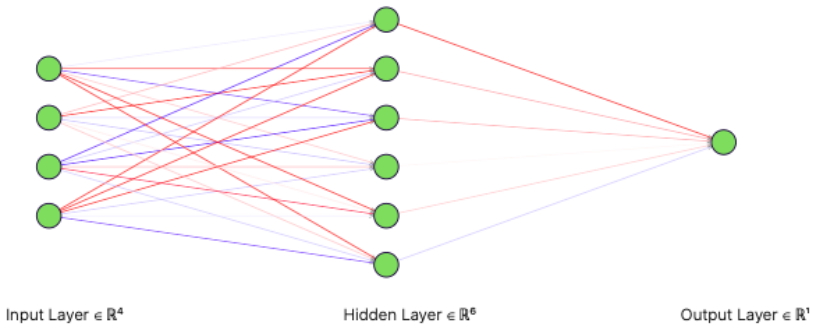


Figure 3.2: Three layered neural networks.

Figure 3.2 shows an example architecture of a three-layer feed-forward neural network embedded on a 4-cell input later, a 6-cell hidden layer, and a 1-cell output

layer. Green circles are cells of activation functions to process information before passing through the corresponding layer. The arrows represent the information flow direction from input to output, where red and blue label the positive and negative edge proportional to assigned weights. In practice, the count of neurons in the input layer is required to be the same as the dimension of input predictors, and the cell number for hidden layers will be determined in trials that start from 2 and increase at a rate of 2 power. Such expanding neurons will bring structural complexity for predicting effectiveness, especially long-winded networks that will make the model over-parameterized (i.e. less generalized beyond the trained samples).

The neurons in each layer provide a driving force to aggregate information by hiring activation functions. Plenty of activation functions, such as *sigmoid*, *ReLU* and *tanh*, are available to process various non-linear relationships according to the property of the learning target. The *sigmoid* activation function is mostly used to output the probability by confining the predicting value between 0 and 1.

$$\text{sigmoid: } f^2(x) = \frac{1}{1+e^{-x}}. \tag{3.23}$$

Before the output being processed, the weight parameters vector will be applied for each layer neurons, thus the aggregated information can be normally connected. Denote $\boldsymbol{w}^1$ and $\boldsymbol{w}^2$ to be the weight parameters for bridging between (1) input layer and hidden layer and (2) hidden layer and output layer. Thus, the predicted result for a three-layer ANN with n input variables, m-cell hidden layer and single cell output layer, can be written as follows,

$$\hat{y}_{t+1} = f^2(\sum_{j=0}^{m} w_j^2 \cdot f^1(\sum_{i=0}^{n} w_{j,i}^1 \cdot x_{i,t})), \tag{3.24}$$

where $x_{i,t}$ is the value of variable i at time t, $w_{j,i}$ is the applied weight to the i^{th} input neuron for producing the input for j^{th} hidden neuron and w_j is the applied weight to j^{th} hidden neuron output for the final singular prediction. For more than three layers model, the process can be recursively implemented by assigning various dimensional weight parameters. The parameters will then be optimized by minimizing the L_2 penalized objective function in a number of epoch iterations.

3.3.2 Deep neural networks of LSTM

The deep neural networks (DNNs) are the extended frameworks for the conventional neural networks (NNs). Subasi (2020) summarizes that the main difference between the NN and DNN is embodied in the hidden layer complexity, where the NNs have more shallow depth (of one or two layers) than DNNs. The convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long-short term memory neural networks (LSTMs) are all members of DNNs. The long-short term memory (LSTM) network (Jordan, 1997) is the extended form of recurrent neural networks (RNNs) (Hochreiter and Schmidhuber, 1997) which learns from both long- and short-term dependencies for sequential forecasting. Its advantage is not merely embodied by carrying the time-dependence between observations but being embarked on collaborating with more advanced neuron layers(such as bidirectional LSTM and attention mechanism equipped), making it more adaptive to various application scenarios.

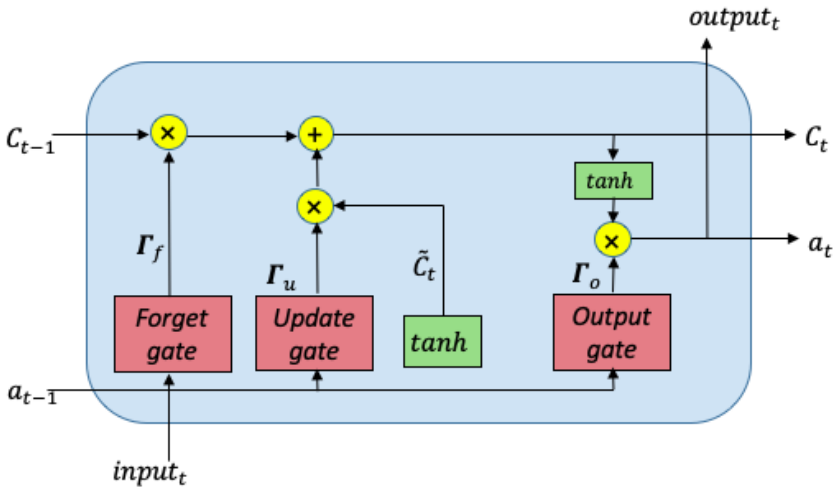


Figure 3.3: The LSTM cell inner structure at time t.

As an extension of classic RNN, LSTM keeps its merit which allows the processing of sequential data with arbitrary lengths via the hidden state vector, at the same time enhances the learning power of long-distance dependency by introducing the 'memory' cell. As Figure 3.3 displayed, the inputs of a_{t-1} (activation function) and C_{t-1} (peehole function) that carry historical information from the former cell to pass through the LSTM cell to generate the output and next input information for time t. The a_t and C_t will be recurrently employed for the sub-

sequent memory block. The initial values of C_0 and a_0 are both zero. $\mathbf{\Gamma}_f$, $\mathbf{\Gamma}_u$, $\mathbf{\Gamma}_o$ are sigmoid function of *the forget gate, the update gate* and *the output gate* that determine the information to be discarded, added and reproduced, respectively. $\mathbf{\Gamma}_{f,u,o}$ will process the information to balance between the previous activation and the current input. $\tilde{C}_t$ is the new candidate output created by the *tanh* layer on the basis of absorbing the current input and the previous hidden state. It serves for the next state C_t which is a combination of the previous memory C_{t-1} and the updated hidden state $\tilde{C}_t$. The formulation for $\mathbf{\Gamma}_f, \mathbf{\Gamma}_u, \mathbf{\Gamma}_o$ and the new candidate state $\tilde{C}_t$ can be written as follows,

$$
\begin{aligned}
\mathbf{\Gamma}_f &= \sigma(x_t U^f + a_{t-1} W^f), \\
\mathbf{\Gamma}_u &= \sigma(x_t U^u + a_{t-1} W^u), \\
\mathbf{\Gamma}_o &= \sigma(x_t U^o + a_{t-1} W^o), \\
\tilde{C}_t &= tanh(x_t U^g + a_{t-1} W^g),
\end{aligned}
$$

where σ is the constant parameter, x_t is the input observation vector, U is the weighted matrix connecting inputs to the current layer, W is the recurrent connection between the previous and current layers.

Based on the LSTM cell, the integrated EWS for stocks and more advanced EWS variants, including bidirectional information processing layer and attention mechanism stacked layer for bonds, will be further explored in Chapter 5 to satisfy the aforementioned objectives.

By far, the elementary methodology bases for the CM and EWS development have been described in terms of formulations and diagrams. The following three chapters of 4, 5 and 6 will respectively introduce the proceedings of 1) the hybrid contagion models construction and empirical analysis on the database for the U.S. and China, 2) the integrated early warning system development for China's stocks and bonds with advanced crisis identification techniques and deep learning based predictive module and 3) the contagion fused early warning system that quantized the contagious effects by proposing a new contagion intensity index and validates the machine learning drawn leading factors by designing appropriate statistical hypothesis tests.

Chapter 4

Hybrid contagion model construction

4.1 Hierarchical structure for modeling contagion - Two hybrid CMs

With the purpose of detecting augmented linkage within the financial sectors in the domestic scale of China, two hybrid contagion models, namely the Bi-SWARCH-EVT and the Paired SWARCH-EVT-Copula, are developed and implemented to handle the joint tail behaviors and dependence structure of paired assets in the occurrence of crises. The implementing process will be factorized by hierarchically arranging elementary bases of SWARCH, EVT and R-vine copulas, that have been formulated in sections of 3.1−3.2. Different from previous studies on contagion effects, our hybrid models can solve and clarify contagions among the internal markets for one country in a more comprehensive way: 1) the Bi-SWARCH model marks the co-movement period for each paired two markets, 2) the EVT allows to separately discuss the heavy tail effect of financial variables after stripping out the contagious effects, and 3) the copula melts the bergs of imposing a contagious originator and disability of inferring risk transmitting path in the Bi-SWARCH based contagion model frame.

4.1.1 Bi-SWARCH-EVT Model

The Bi-SWARCH model assumes residuals to follow the normal distribution with zero means, which is deemed to be an ideal case, especially for the chaotic financial

markets. In contrast, characterizing residuals of return series as non-symmetrical distribution with fat-tails is proven to be a reliable approach, especially in terms of handling extreme returns under financial turmoils. One of the handy ways to model heavy-tailed residuals is to utilize the so-called Peak-Over-Threshold (POT) method, together with the Block Maxima method, is one of the two major techniques covered in Extreme Value Theory (EVT). Instead of directly assuming the whole residuals follow complex distributions of normal-inverse Gaussian and hyperbolic functions, the EVT allows us to investigate the behavioral characteristics of heavy tails separately. By integrating EVT with the Bi-SWARCH model, the hybrid contagion model enables the parameterization of the tail behavior of residuals with applicable distributions.

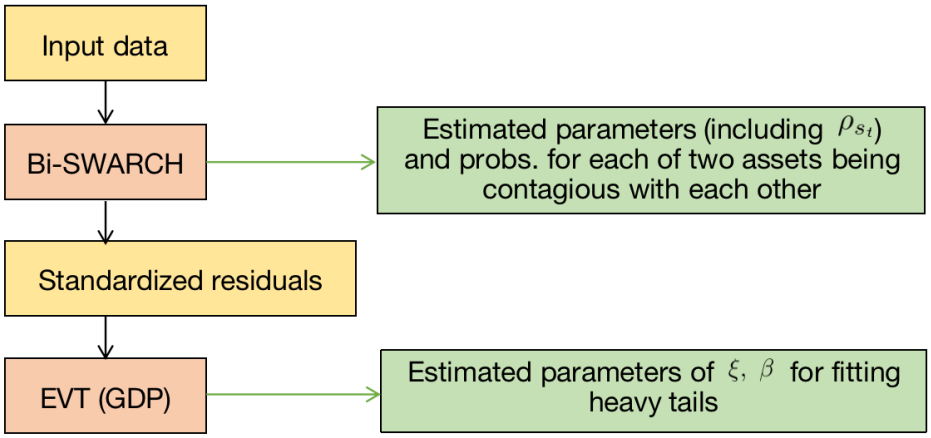


Figure 4.1: Bi-SWARCH-EVT model frame.

In detail, the Bi-SWARCH-EVT model is implemented in three steps. As Figure 4.1 shows, the log returns are first empirically fitted in the Bi-SWARCH model given the crisis originator. Then, the model processed outputs will be collected as two parts of i) bi-variate SWARCH model estimated parameters, including $\boldsymbol{\theta}$, the state-varying correlation coefficient ρ_{s_t} and the filtering/smooth probabilities, and ii) standardized residuals. In i), potential contagion episodes between the paired markets will be clarified by thresholding the filtering/smooth probabilities under 0.5 cutoff level. Last, the extracted residuals will be further fitted by the GPD. The model estimated parameters of ξ's and β's will gain an understanding of the tail behavior of asset returns under the Bi-SWARCH specifications.

4.1.2 Paired SWARCH-EVT-Copula Model

The assumption that Bi-SWARCH-EVT model imposes a crisis originator brings two problems: 1) the bi-variate model limits the study scope to the risk transmission between each of two markets - such limitation though can be resolved by reversing the crisis originator to the role of the recipient for each couple, bring the computational cost in practice; 2) it is, both theoretically and technically, challenging to extend the SWARCH model to a higher dimension than two to simultaneously study the crisis transmission across multiple internal markets for a country. The alternative hybrid CM of Paired SWARCH-EVT-Copula can fill the dual gaps by saving the cost of replicating experiments for reversing the crisis originator role and studying the risk transmission path for including multiple financial markets.

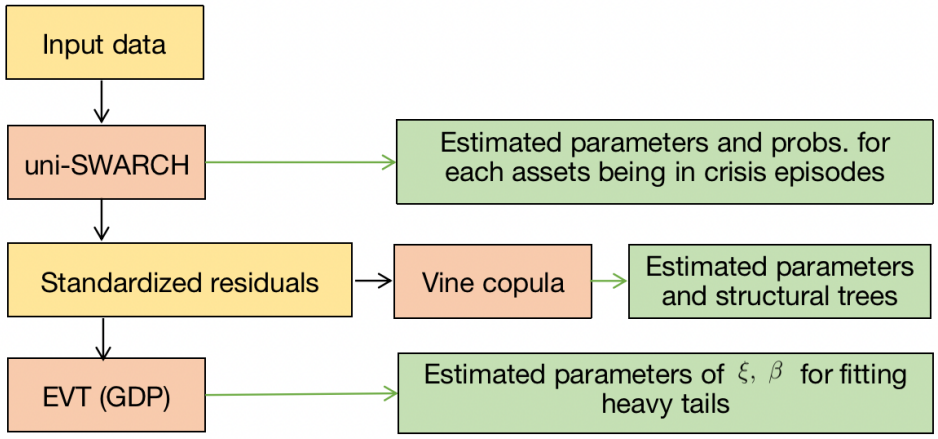


Figure 4.2: Paired SWARCH-EVT-Copula model frame.

By integrating R-vine copulas and EVT with the univariate SWARCH, the implementation process for Paired SWARCH-EVT-Copula model, as Figure 4.2 displays, will be unpacked in three steps. First, the log returns data of each asset is fitted by the univariate SWARCH model, and the model estimates, especially of filtering/smooth probabilities for both high- and low- volatility states, are extracted to date the crisis episodes for each asset, which is also regarded as the supportive evidence to discover potential crises spread as the complementary results to the Bi-SWARCH model. Then, standardized residuals from the univariate SWARCH model will be fitted in the GPD to measure the asymptotic

tail behavior of each asset return. R-vines are meanwhile applied to all univariate SWARCH extracted innovations to measure the connected structure and tail dependence channels across multiple markets.

Table 4.1: Two proposed hybrid Contagion Models and their corresponding highlights.

model	CM1: Bi-SWARCH-EVT	CM2: SWARCH-EVT-Copula
SWARCH	Bi-variate: Date the contagion episodes for paired markets based on the smooth probability of the high-high volatility state; Infer the contagion intensity via the dependent correlation coefficient.	Univariate: Date the crisis episodes for a given asset based on the smooth probability of the high volatility state.
EVT	The heavy tail of asset returns are fitted by the semi-parametric GPD.	
Copula		C-/D-Vines: The dependence structure across financial markets are investigated by pair based on the conditional bivariate dependency; Risk transmission paths are identified by estimating the vine trees.
Highlights	• Visualize contagion episodes pairwise • Infer correlation-based intensities	• Visualize crisis episodes individually • Infer the risk transmission channels

By far, two hybrid CMs constructing and implementing procedures have been fully depicted. Table 4.1 summarizes the proposed models' highlights in the comparison to show what they are competent for in the contagious effects detection, respectively. Each of model components being hired in the hybrid CM construction is displayed and described their functioning. As the bottom row concludes, the Bi-SWARCH-EVT model visualizes the contagion periods between each pair of asset markets and meanwhile estimates the contagious intensity during the crisis originator's high volatile period, whilst the latter model of SWARCH-EVT-Copula is capable of specifying the crisis period for each asset and clarifying the

risk transmitting pattern across multiple assets.

In the following section, the proposed models will be retrieved in practice by accessing data from principle markets of real estate, stocks, bonds, and forex for China and the U.S. to make comparative analysis for both countries, especially to distinguish the difference of risk transmission between developed and developing countries. To display the empirical results in a tidy way, the modeling process will be introduced by A) first implementing the SWARCH-EVT-Copula model and B) then carrying out the Bi-SWARCH-EVT model in the following application sequence:

A-1) the univariate SWARCH inferred crisis episodes for each of market assets;
A-2) the SWARCH-EVT estimated marginal effects on heavy tails;
A-3) the R-vine copulas measured the possible risk transmission channels;

B-1) the bi-variate SWARCH inferred contagious episodes pairwise;
B-2) the Bi-SWARCH-EVT measured spillover effects.

4.2 Empirical analysis for China and the U.S. markets

4.2.1 Data

There are seven assets' daily log returns being sourced to investigate the contagious effects and risk transmissions across the real estate, the stock, the bond, and the forex markets in the U.S. and China based[1]. Specifically, we include the iShare U.S. Real Estate ETF (IYR, U.S. real-estate index), Invesco China Real Estate ETF (TAO, China's real estate market index), S&P500 (U.S. stock index), Shanghai Securities Composite Index (SSEC, Chinese stock index), 3-Year Treasury Bond Yield to Maturity (T-Bond in the U.S.), 3-Year China Government Bond Yield to Maturity (G-Bond in China), and the Off-shore Foreign Exchange between the U.S. Dollar and Chinese Yuan (USD/CNH)[2]. The time span cov-

[1]The data set for the bond market is taken from the yield rate for one of the most popular government bonds. In general, the short-term bonds show greater liquidity in the secondary market tradings.

[2]USD/CNH is obtained from the Federal Reserve Economic Data. ITR, TAO, S&P500, and SSEC are from Yahoo Finance. G-bond and T-bond yields are from Wind.

ers from June 6, 2008 to April 4, 2019 with 2521 daily observations for each asset series.

Table 4.2: Data descriptive statistics.

freq: daily	U.S.			China			USD/CNH
	IYR	**S&P500**	**T-Bond**	**TAO**	**SSEC**	**G-Bond**	
min.	-0.23	-0.14	0.28	-0.21	-1.61	0.80	-0.01
max.	0.15	0.10	3.38	0.18	2.22	5.11	0.02
median	0.00*	0.00*	1.04	0.00*	0.00**	2.63	0.00*
mean	0.00*	0.00*	1.22	0.00*	0.00**	2.55	0.00*
se.mean	0.00*	0.00**	0.01	0.00**	0.00**	0.02	0.00**
std.dev.	0.02	0.01	0.71	0.02	0.17	0.80	0.00**
skew.	-0.77	-0.72	0.92	-0.02	0.99	-0.22	0.36
kurtosis	19.1	14.66	0.12	11.78	43.98	-0.21	11.59
Jarque-Bera	38645†	22836†	359.92†	14594†	203860†	24.14	14186†

* and ** indicate the values are statistically zero at the 10% and 5% confidence levels.
† indicates the Jarque-Bera test is failed at the 5% confidence level.

According to Table 4.2, the Jarque-Bera test significantly rejects the normality hypothesis of all series except the G-Bond yield, which suggests the essentials of fitting margin models to capture the innate irregularity of financial returns.

4.2.2 SWARCH-EVT-Copula estimation

Turmoil episodes for principal markets

Figure 4.3 - 4.6 show the price index and model inference plots for each asset, specifically, the 2-by-2 subplots are respectively the price (top left), ARIMA fitted residuals (top right), conditional volatility (bottom left) and smooth probabilities (bottom right) of each asset. Red regions shade the observations of $p_{s_t=2} > 0.5$ in the last panel, which have been listed in Table 4.3 as specific dates of crisis episodes in each sample. The following results are observed according to the plots and specified crisis episodes from the figures and the table.

First, the smooth probabilities give consistent indications, as the residuals and conditional volatility plots suggest. With the 0.5 threshold of the smooth probability, the turmoil and tranquil periods could be classified in a reasonable manner. Second, the off-shore USD/CNH exchange rate dynamics tend to remain at a highly fluctuating level as the Renminbi policy changed to a "crawl-like arrangement" since 2010, whilst the "crawling peg" policy sustained the stability

of Renminbi during 2008 to 2010. Third, the interest rate market exhibits a higher level of sensitivity towards risks as the bond yields appear to have more frequent and lasting turmoil periods in comparison to the real-estate and stock log returns. Last but not least, significant overlaps between the high volatility regimes of all the American and Chinese markets are observed. In particular, two major turmoils are the Subprime Mortgage Crisis and European Sovereign Crisis that took place at around 2008, and the worldwide economic turbulence during 2015-2016 as a response to various economic and political shocks, including the interest rate rise by the Federal Reserve, the Brexit voting, the collapse of the crude oil market, and the weak global economy.

Table 4.3: The univariate SWARCH inferred crisis dates for each asset.

Country	Univariate	Obs.($p_{s_t=2} > 0.5$)	Duration	
U.S.	IYR	926	2008/6/6 - 2009/6/2	2009/7/29 - 2009/10/28
			2010/4/7 - 2010/7/26	2011/7/25 - 2011/12/12
			2013/5/20 - 2013/7/1	2014/12/16 - 2014/12/31
			2015/1/5 - 2015/5/13	2015/8/18 - 2016/3/18
			2016/4/27 - 2016/5/18	2016/9/2 - 2016/12/20
			2018/1/2 - 2018/4/20	2018/9/17 - 2019/2/12
	S&P500	797	2008/6/25 - 2009/9/5	2010/4/26 - 2010/6/17
			2011/7/27 - 2011/12/9	2014/9/23 - 2014/10/12
			2014/12/8 - 2015/3/25	2015/8/14 - 2016/3/1
			2016/6/21 - 2016/7/7	2016/9/8 - 2016/9/28
			2018/1/10 - 2018/7/9	2018/9/27 - 2019/4/4
	T-bond	1523	2008/6/6 - 2011/8/10	2013/6/6 - 2013/9/30
			2014/1/2 - 2014/3/21	2014/9/19 - 2016/9/14
			2016/11/14 - 2017/4/26	2018/2/1 - 2018/3/1
			2018/12/18 -2019/4/4	
China	TAO	860	2008/6/6 - 2009/8/7	2011/8/2 - 2011/11/14
			2015/3/26 - 2015/10/16	2017/8/7 - 2019/4/4
	SSEC	726	2008/6/6 - 2008/10/9	2008/12/5 - 2008/12/11
			2009/2/2 - 2009/4/7	2009/6/26 - 2010/5/25
			2010/9/29 - 2011/1/20	2014/11/25 - 2015/3/9
			2015/8/20 - 2015/10/22	2015/12/31 - 2016/4/19
			2018/2/5 - 2018/8/2	2018/10/10 - 2018/10/11
			2019/2/14 - 2019/4/4	
	G-bond	1830	2008/6/6 - 2008/12/18	2009/3/30 - 2009/7/23
			2010/5/14 - 2010/7/20	2010/8/25 - 2011/11/30
			2012/1/12 - 2012/2/27	2012/5/7 - 2012/5/31
			2012/7/18 - 2012/9/19	2013/5/31 - 2016/4/29
			2016/11/21 - 2018/8/27	2018/9/12 - 2019/4/4
	USD/CNH	1878	2008/6/6 - 2008/12/31	2010/6/11 - 2012/12/14
			2013/1/25 - 2013/7/1	2014/2/13 - 2015/5/8
			2015/8/10 - 2019/4/4	

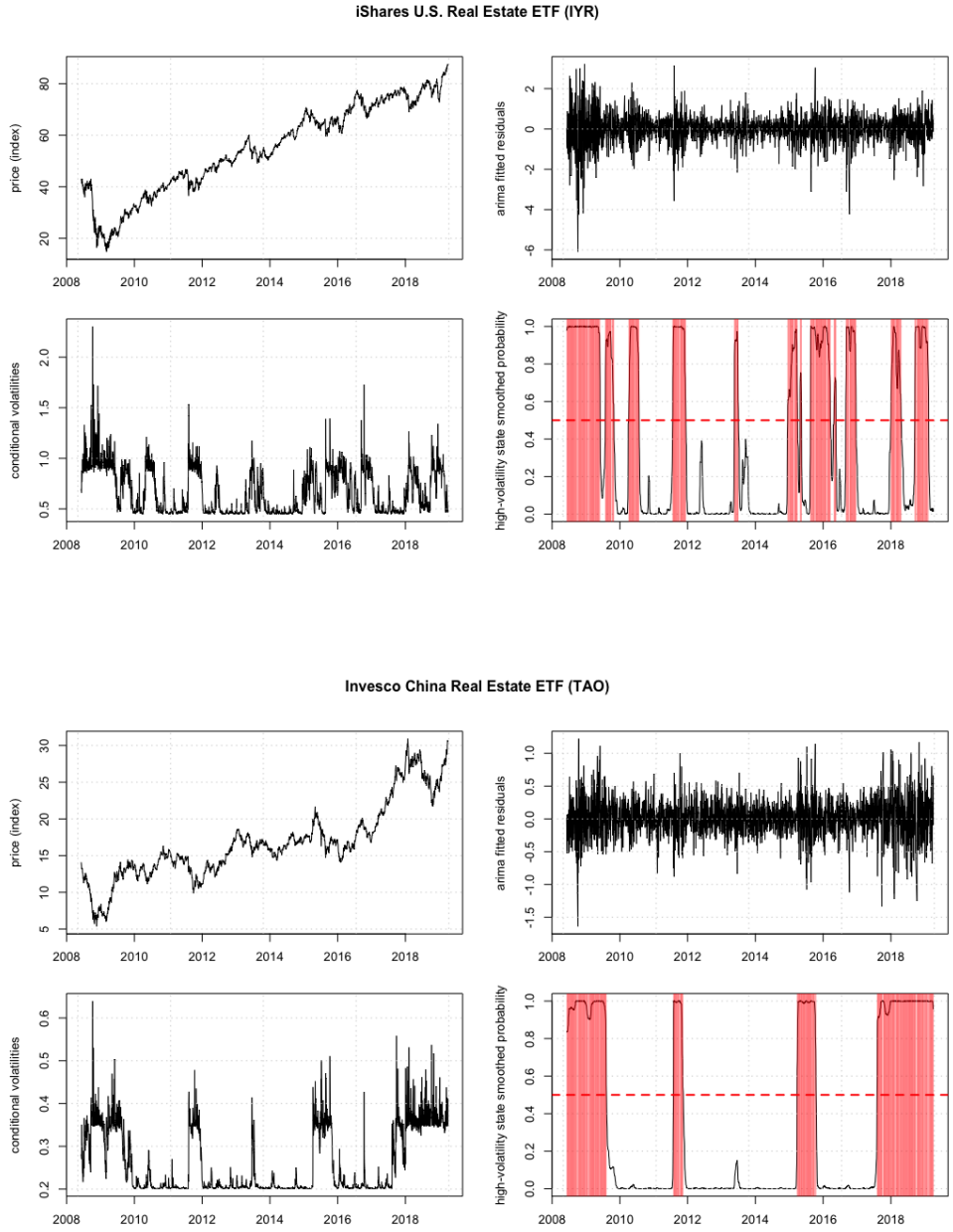


Figure 4.3: Plots of the real-estate in the U.S. (upper panel) and China (lower panel).

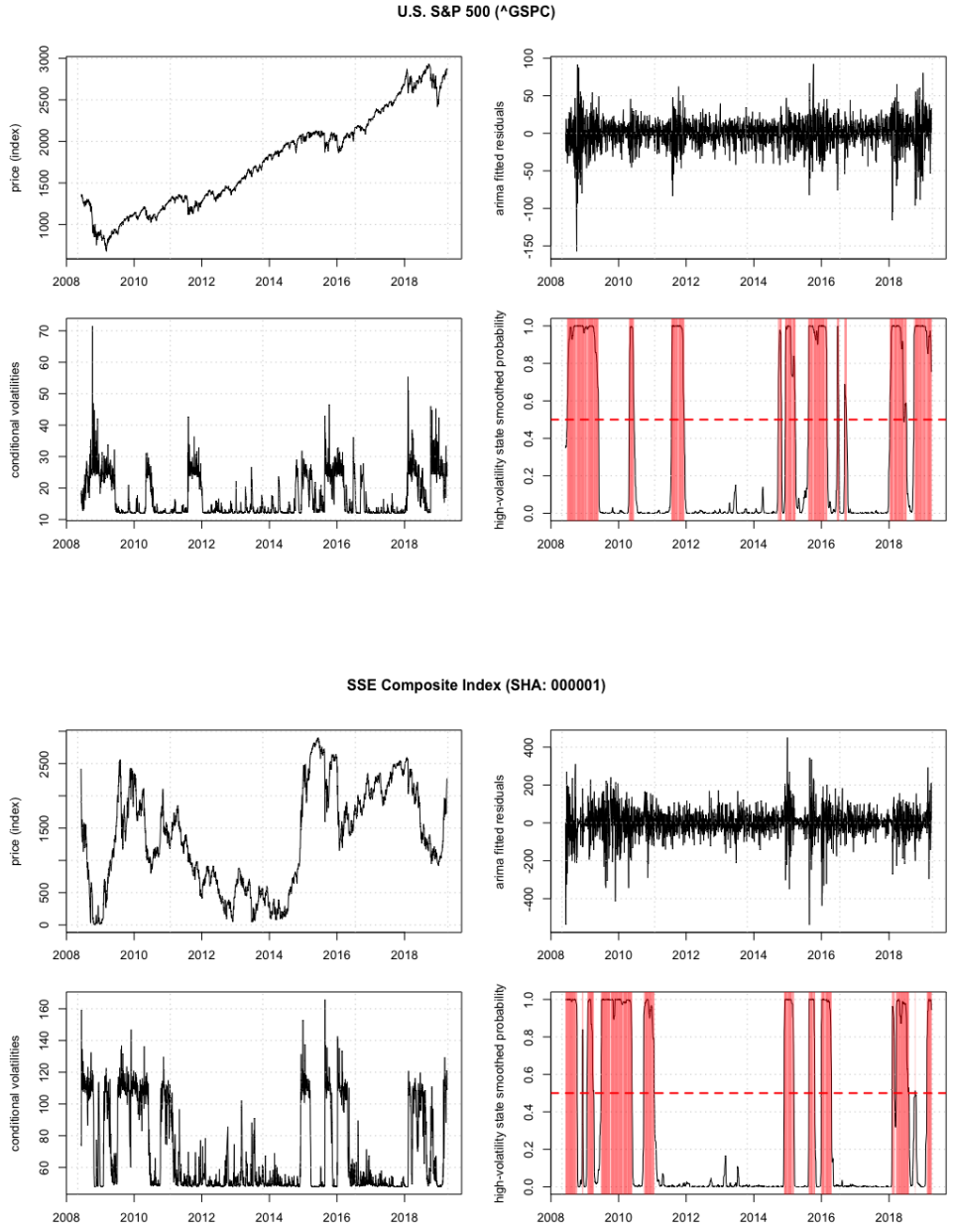


Figure 4.4: Plots of the stocks in the U.S. (upper panel) and China (lower panel).

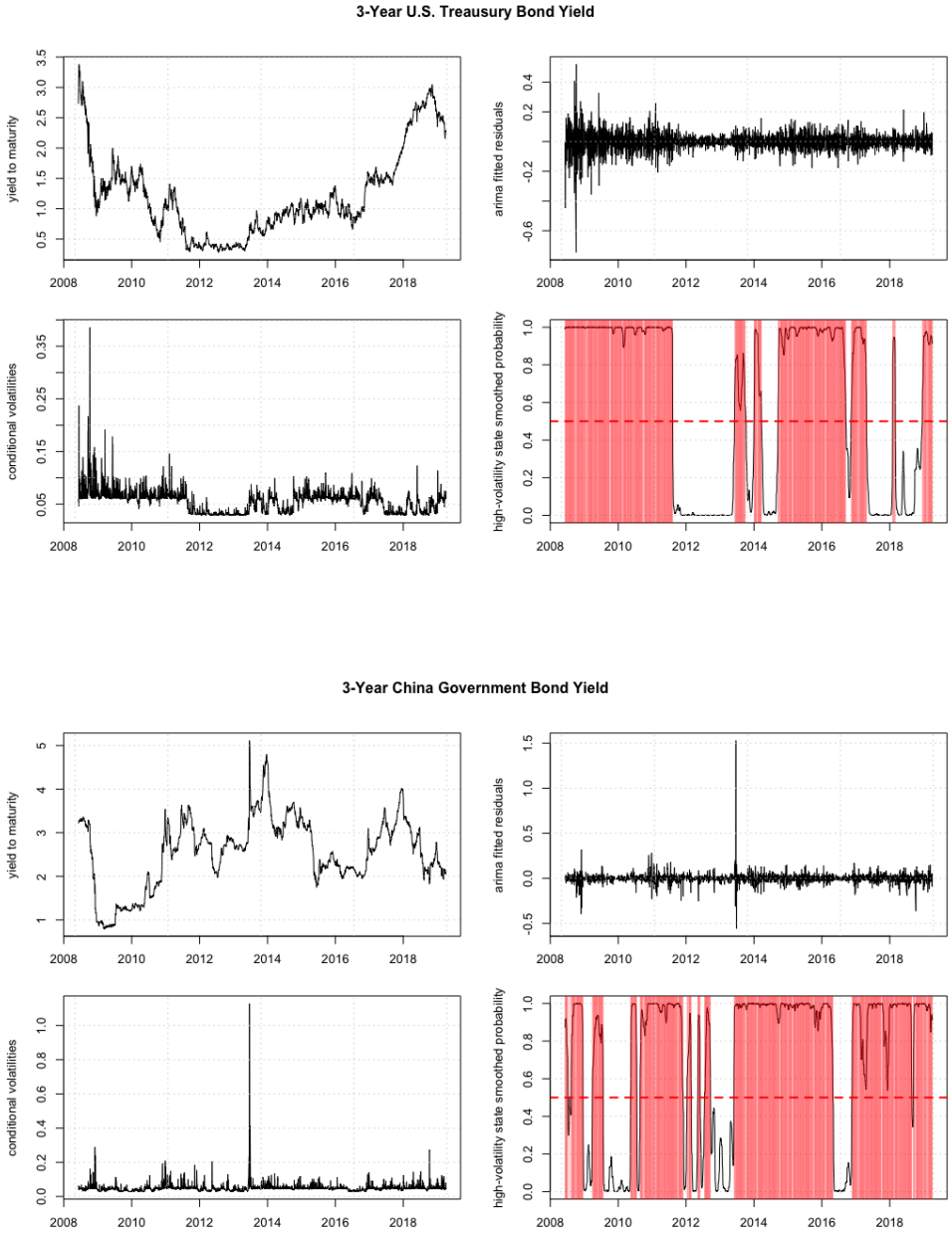


Figure 4.5: Plots of the 3-Year bond yield to maturity in the U.S. (upper panel) and China (lower panel).

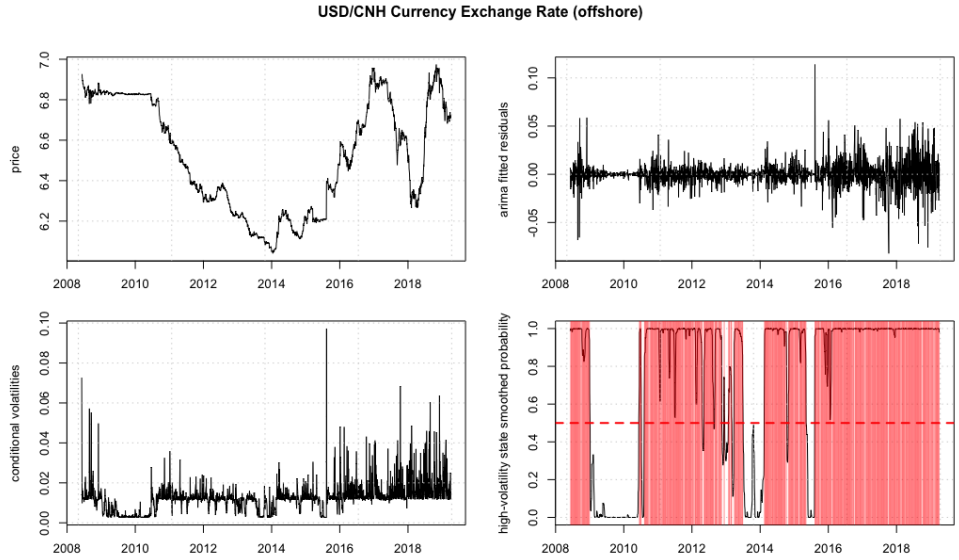


Figure 4.6: Plots of the off-shore forex rate between the U.S. dollar and the Chinese yuan.

Table 4.4: SWARCH-EVT estimated parameters.

Univariate	**IYR**	**S&P500**	**T-Bond**	**TAO**	**SSEC**	**G-Bond**	**USD/CNH**
u	0.018**	0.000***	1.037***	0.000***	0.000***	2.648***	0.000***
	(1.385)	–	(8.526)	–	–	(11.287)	–
θ_1	-1.107*	-0.609*	-0.478*	0.033***	-0.201*	-0.196*	0.000***
	(-0.073)	(-0.063)	(-0.062)	(1.638)	(-0.222)	(-0.556)	–
α_{01}	0.188***	0.415***	0.001***	0.040***	0.002***	0.001***	0.000 ***
	(13.164)	(15.743)	(9.596)	(22.907)	(3.903)	(1.181)	–
α_{02}	0.840***	3.109***	0.004***	0.123***	0.036***	0.002***	0.000***
	(9.416)	(10.277)	(12.163)	(12.576)	(4.218)	(8.095)	–
α_{11}	0.0532***	0.064**	0.210***	0.000**	0.165***	0.797***	1.000***
	(1.461)	(0.084)	(3.213)	(0.021)	(1.913)	(0.965)	–
α_{12}	0.122***	0.186***	0.262***	0.107***	0.999***	0.523***	1.000***
	(2.545)	(3.303)	(5.021)	(2.231)	(312.461)	(5.481)	–
p_1	0.622	0.678	0.381	0.597	0.665	0.283	0.174
p_2	0.378	0.322	0.619	0.403	0.335	0.717	0.826
ξ_L	0.102**	0.048***	0.202***	0.085***	0.198***	0.305***	0.229***
	(1.598)	(0.838)	(3.012)	(1.288)	(2.517)	(3.433)	(2.854)
ξ_R	-0.045**	0.078***	0.137***	-0.021**	-0.032**	0.244***	0.204**
	(-0.679)	(1.151)	(1.972)	(-0.309)	(-0.568)	(3.179)	(2.965)
β_L	0.537***	3.157***	0.136***	0.277***	5.245***	0.125***	0.059***
	(11.223)	(11.408)	(10.915)	(10.739)	(10.008)	(9.607)	(9.779)
β_R	0.490***	1.714***	0.120**	0.301***	5.379**	0.111***	0.069***
	(11.001)	(10.853)	(10.701)	(10.792)	(10.802)	(10.207)	(10.820)

The maximum likelihood estimations of the SWARCH-EVT model. The t-values are provided in the parenthesis. *,** and *** denote significant levels at 0.1, 0.05 and 0.01, respectively. − means the value is less than 10^{-5}

Marginal model estimations

The estimated results of the univariate SWARCH-EVT marginal model are displayed in Table 4.4. Basically, it supports the validity of SWARCH-EVT assumptions on the margins as all model parameters are shown to be significant. As the state probabilities in the high (p_2) and low (p_1) volatility states suggest, the bond and currency markets tend to be more volatile during the sample period with their p_2 values greater than 0.5, whilst the stock and real-estate markets tend to be more stable in both China and the U.S. .

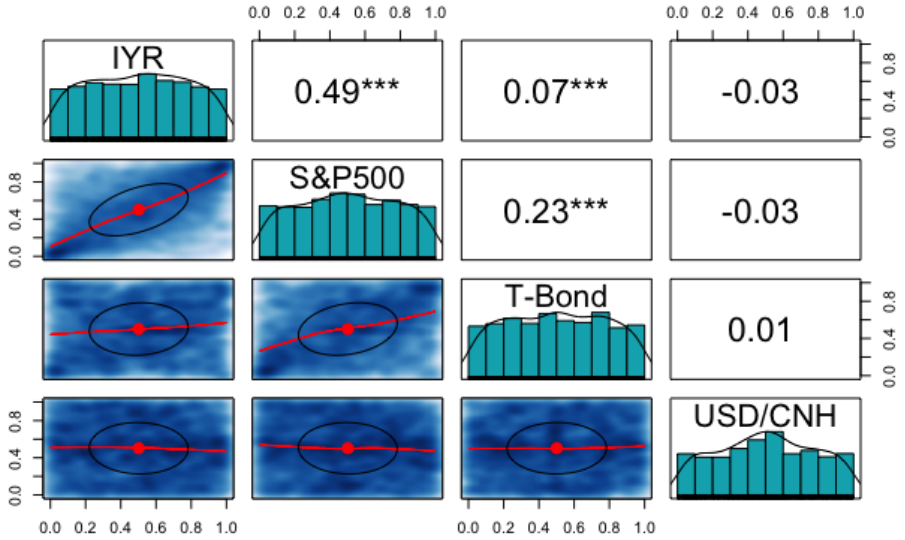
Empirical correlated patterns

On the basis of the univariate SWARCH-EVT estimation, marginal residuals are obtained and transformed to uniformly distributed series to fit the pair-wise dependence structure with bi-variate copulas. Based on the entire samples of residuals, Figure 4.7 shows the residual histograms (the diagonal subplots), as well as scatter plots (sub-plots in the lower triangle) and the Kendall's $\tau's$ (subplots in the upper triangle) between each pair of assets. The implications can be obtained regarding the general tail dependence structures.

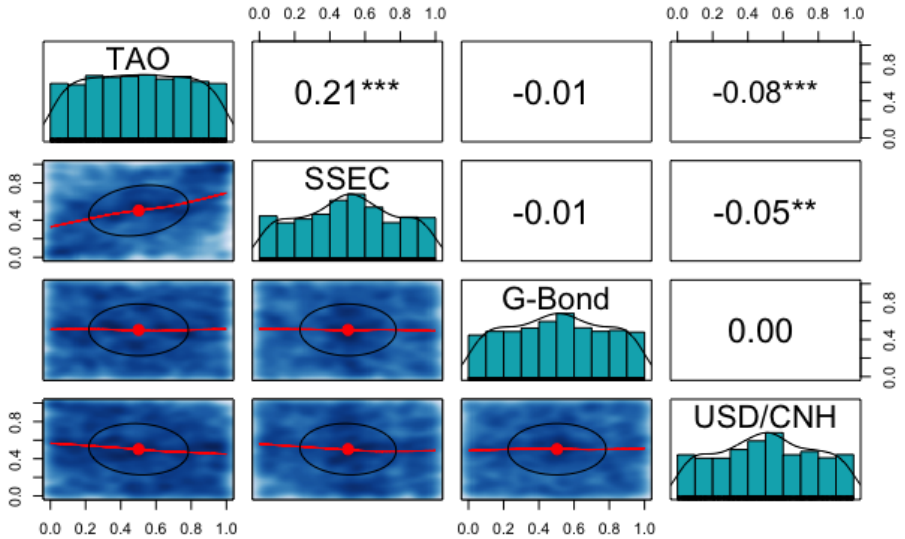
1. The stock and real-estate markets in both the U.S. and China exhibit strong tail dependencies with high Kendall's $\tau's$ (0.49 between the IYR and S&P500 in the U.S., and 0.21 between the TAO and SSEC in China), as well as clear linear relationships in the scatter plots.

2. In addition, the S&P500 and T-bond exhibit strong positive linkages in the tails with the Kendall's $\tau = 0.23$.

3. Both the U.S. dollar and Renminbi show negative tail correlations with their local real estate and stock markets suggested by negative Kendall's $\tau's$, and less significant correlations with local bond markets.



(a) The Chinese real-estate, stock, bond and forex markets.



(b) The U.S. real-estate, stock, bond and forex markets.

Figure 4.7: The scatter plots of empirical correlation between asset residuals.

The residual histograms, the pair-wise scatter plots with the linear regression results (red line) and the elliptical contours (black), and the Kendall's $\tau's$. (*,** and *** denote significant levels at 0.1, 0.05 and 0.01, respectively.)

Copulas estimated risk transmission paths

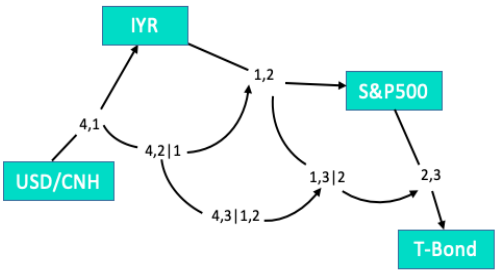
To gain a more comprehensive picture of the assets' tail behaviors under market turmoils, Figure 4.8 and 4.11 explicitly show the R-vine copula structure changes in the entire sample and the crisis episodes. In particular, subplot (a) shows the R-vine path whilst (b)(c)(d) are the decomposed trees with the edges labeled by the fitted copulas and estimated Kendall's $\tau's$ in parentheses. The major findings are threefold.

1. The R-vine copula dependence structure during the crisis periods (Figure 4.9 - 4.11) varies from that in the full sample (Figure 4.8 - 4.10). In the crisis period, the American stock market and the Chinese real-estate market tend to play central roles in spreading risks to their local financial sectors.

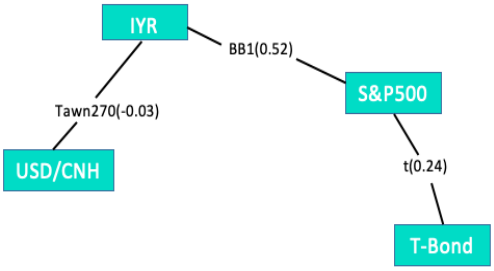
2. Moreover, strong linkages between the real-estate and stock markets are detected in both the U.S. and China during the crisis with opposite transmission directions i.e. U.S.: stock $\rightarrow$ real-estate. China: real-estate $\rightarrow$ stock.

3. In the full sample, the exchange rate is shown to be the starting node of tail dependence in both the U.S. and China, with the transmission path of the forex $\rightarrow$ real-estate $\rightarrow$ stock $\rightarrow$ bond market. However, the relatively low and negative values of the Kendall's τ suggest an overall weak spreading effect from the forex to domestic markets.

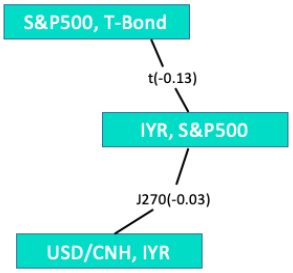
Table 4.5 lists the parameter estimations of the R-vine copulas in Figure 4.8 - 4.11. Basically, the tail dependence parameters tell the same story as the R-vine copula trees in terms of the risk transmission channels in the U.S. and China. In general, the American markets demonstrate higher levels of tail-dependent contagion effects than the Chinese markets with greater values of $\tau's$ and $\lambda's$. From the full sample estimation in panel A to the crisis-period estimation in panel B, the values of the Kendall's $\tau's$ between the real estate and stock markets increase significantly in both countries, which further supports the R-vine copula tree structures obtained in Figure 4.9 and 4.11.



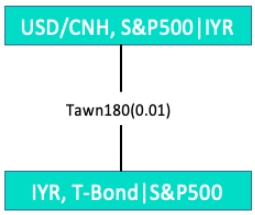
(a) R-vine path



(b) Tree 1

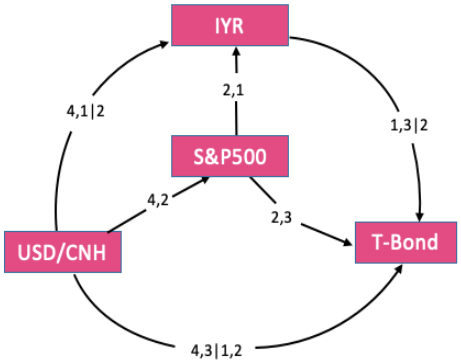


(c) Tree 2

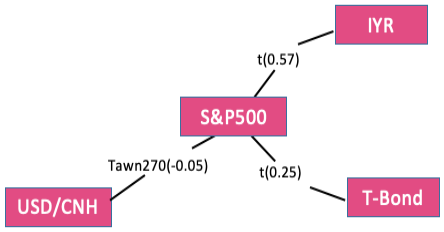


(d) Tree 3

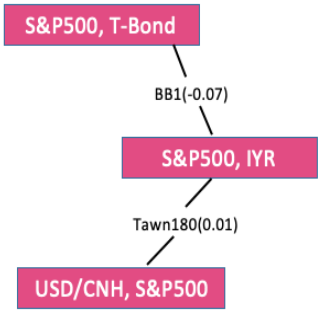
Figure 4.8: The R-vine path and decomposed trees for all residuals in the U.S..



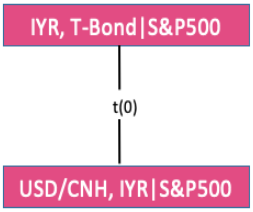
(a) R-vine path



(b) Tree 1

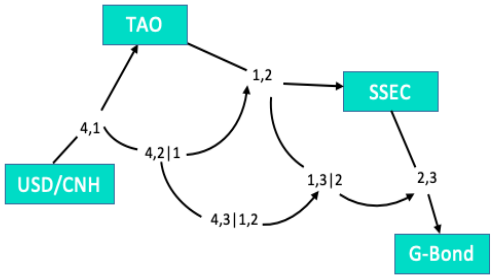


(c) Tree 2

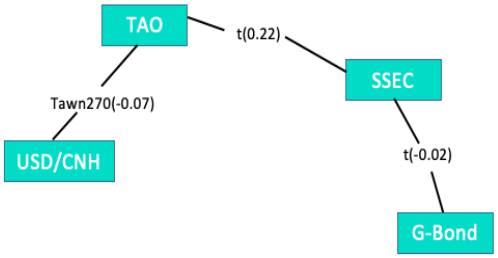


(d) Tree 3

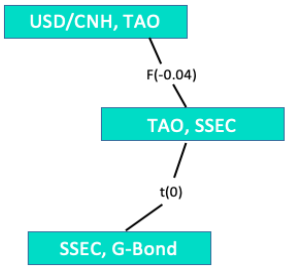
Figure 4.9: The R-vine path and decomposed trees for crisis-period residuals in the U.S..



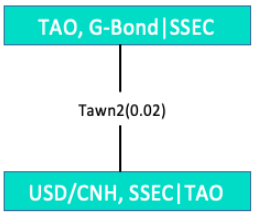
(a) R-vine path



(b) Tree 1

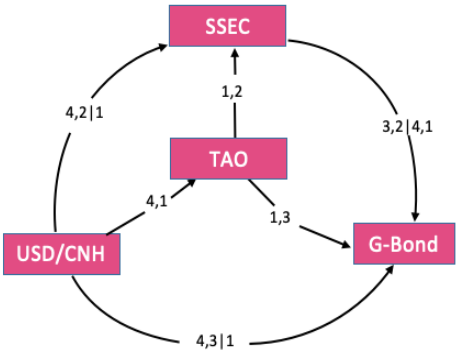


(c) Tree 2

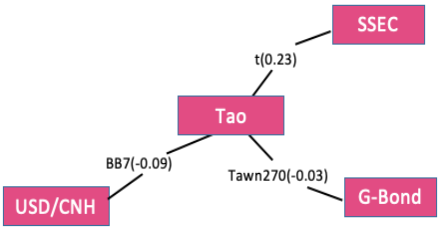


(d) Tree 3

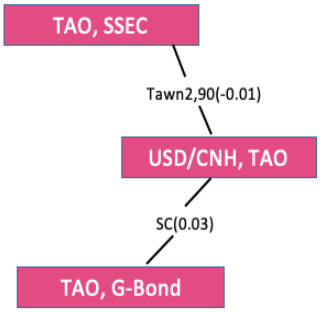
Figure 4.10: The R-vine path and decomposed trees for all residuals in China.



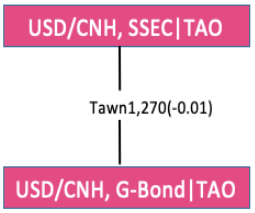
(a) R-vine path



(b) Tree 1



(c) Tree 2



(d) Tree 3

Figure 4.11: The R-vine path and decomposed trees for crisis-period residuals in China.

Table 4.5: R-Vine copula estimation for the U.S. (upper panel) and China (lower panel).

Country	Tree	Edge	Family	Parameter$_1$	Parameter$_2$	Kendall's $\tau_{est.}$	λ_U	λ_L
U.S.	panel A: full samples							
	Tree 1	1,2	BB1	0.57(1.89)	1.61(2.55)	0.52	0.461	0.469
		2,3	t	0.37(1.67)	5.08(3.89)	0.24	0.145	0.145
		4,1	Tawn 1_{270}	-1.21(-0.04)	0.06(0.02)	-0.03	0	0
	Tree 2	1,3|2	t	-0.21(-0.03)	11.33(6.56)	-0.13	0.001	0.001
		4,2|1	Joe$_{270}$	-1.06(-0.09)	–	-0.03	0	0
	Tree 3	4,3|1,2	Tawn 1_{180}	1.4(2.89)	0.02(0.65)	0.01	0	0.046
	panel B: crisis samples							
	Tree 1	2,1	t	0.67(1.43)	4.79(2.67)	0.57	0.327	0.327
		2,3	t	0.27(1.27)	13.7(7.29)	0.25	0.011	0.011
		4,2	Tawn 1_{270}	-1.31(-0.06)	0.25(0.93)	-0.05	0	0
	Tree 2	3,1|2	BB1$_{90}$	-0.07(-1.03)	-1.04(-0.02)	-0.05	0	0
		4,3|2	Tawn 1_{180}	1.9(3.19)	0.01(0.39)	0.01	0	0.010
	Tree 3	4,1|3,2	t	0(2.68)	9.31(6.95)	0	0.009	0.009
China	panel A: full samples							
	Tree 1	1,2	t	0.34(0.61)	6.80(2.71)	0.22	0.086	0.086
		2,3	t	-0.02(-0.16)	14.82(5.34)	-0.02	0.001	0.001
		4,1	Tawn 1_{270}	-1.30(-0.15)	0.15(0.06)	-0.07	0	0
	Tree 2	1,3|2	t	-0.01(-1.12)	18.16(4.72)	0	0.000	0.000
		4,2|1	Frank	-0.32(-1.57)	–	-0.04	0	0
	Tree 3	4,3|1,2	Tawn 2	1.14(0.12)	0.06(0.07)	0.02	0.026	0
	panel B: crisis samples							
	Tree 1	1,2	t	0.35(2.26)	20.24(18.67)	0.23	0.004	0.004
		1,3	Tawn 1_{270}	-1.27(-0.11)	0.05(1.43)	-0.03	0	0
		4,1	BB7$_{270}$	-1.09(-0.07)	-0.1(-0.04)	-0.09	0	0
	Tree 2	4,2|1	Tawn 2_{90}	-1.28(-1.16)	0.01(4.29)	-0.01	0	0
		4,3|1	Survival Clayton	0.06(1.73)	–	0.03	0.001	0
	Tree 3	3,2|4,1	Tawn 1_{270}	-1.8(-0.02)	0.01(3.45)	-0.01	0	0

Edges are: for the U.S. 1 $\leftrightarrow$ IYR, 2 $\leftrightarrow$ S&P500, and 3 $\leftrightarrow$ T-Bond; for China 1 $\leftrightarrow$ TAO, 2 $\leftrightarrow$ SSEC and 3 $\leftrightarrow$ G-Bond. 4 $\leftrightarrow$ USD/CNH for both countries. t-statistics are in the parenthesis for Parameter$_1$ and $_2$. λ_U and λ_L reflect the tail dependence for upper and lower sides.

4.2.3 Bi-SWARCH-EVT estimation

Contagious episodes across markets

We further investigate the Bi-SWARCH-EVT model estimated contagious episodes. Figure 4.12 - 4.13 show the bi-variate SWARCH smooth probabilities for the high-high volatility state $s = 4$. The contagious episodes are thus inferred as the estimated smooth probability $p_{s_t=4}$ exceeds 0.5 and summarized in Table 4.6. The Bi-SWARCH-EVT model estimated parameter results are provided in Table 4.7.

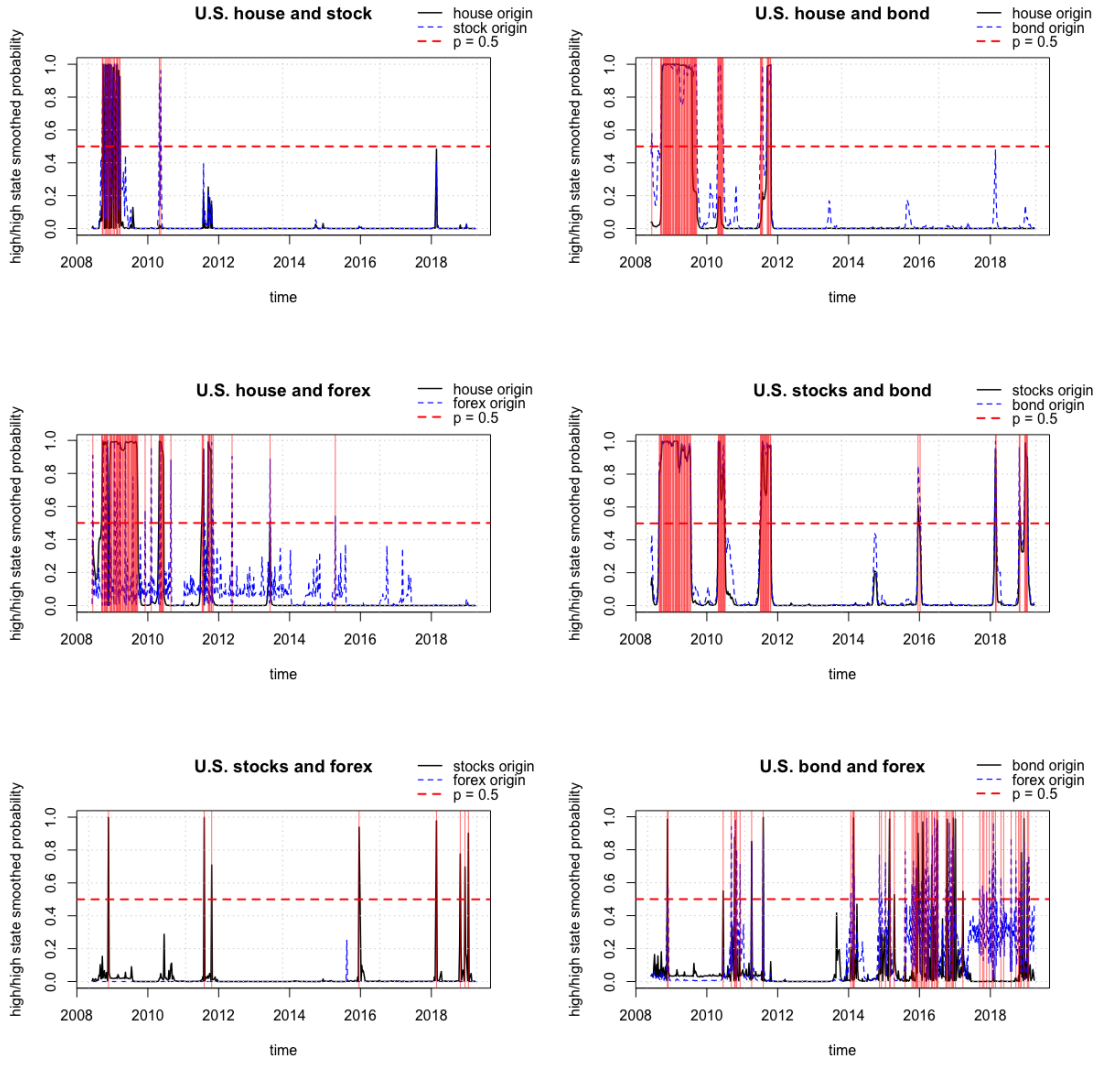


Figure 4.12: U.S.: Bi-variate SWARCH inference based on the smooth probability $p_{s_t=4} > 0.5$.

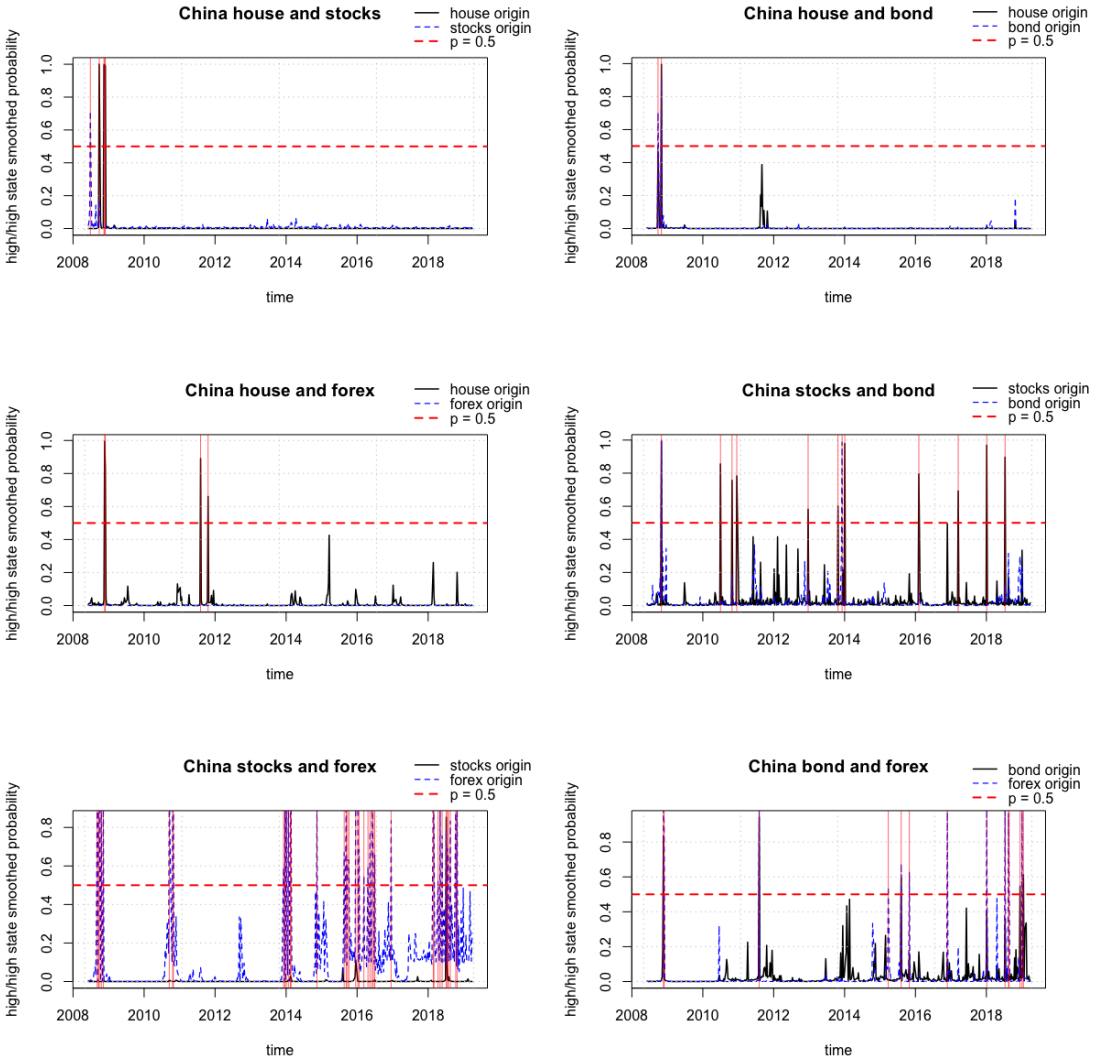


Figure 4.13: China: Bi-variate SWARCH inference based on the smooth probability $p_{s_t=4} > 0.5$.

Table 4.6: Count of contagious periods in the high-high volatility state ($s_t = 4$) for each pair of assets.

Country	Bi-variate	Obs.($p_{s_t=4} > 0.5$)	Duration	
U.S.	real-estate and stocks	112	2008/9/19 - 2009/3/20	2010/4/30 - 2010/5/14
	real-estate and bond	462	2008/6/13 - 2008/6/20	2008/9/12 - 2009/9/18
			2010/4/23 - 2010/6/18	2011/7/8 - 2011/10/21
	real-estate and forex	476	2008/6/13 - 2008/6/20	2008/9/12 - 2009/12/4
			2010/2/5 - 2010/2/12	2010/4/30 - 2010/5/7
			2010/5/7 - 2010/6/11	2010/8/27 - 2010/9/3
			2011/7/15 - 2011/7/29	2011/9/16 - 2011/11/4
			2012/5/18 - 2012/5/25	2013/6/14 - 2013/6/21
			2015/4/17 - 2015/4/24	
	stocks and bond	539	2008/8/29 - 2009/7/17	2010/4/30 - 2011/10/21
			2015/12/18 - 2015/12/25	2016/1/8 - 2016/1/15
			2018/2/23 - 2018/3/2	2018/10/26 - 2018/11/2
			2018/12/21 - 2019/1/18	
	stocks and forex	56	2008/11/21 - 2008/11/28	2011/8/5 - 2011/8/12
			2011/10/21 - 2011/10/28	2015/12/18 - 2015/12/25
			2018/2/23 - 2018/3/2	2018/10/26 - 2018/11/2
			2018/12/14 - 2018/12/21	2019/1/18 - 2019/1/25
	bond and forex	469	2008/11/21 - 2008/11/28	2010/6/18 - 2010/12/10
			2011/4/8 - 2011/4/	2011/8/5 - 2011/8/12
			2014/1/24 - 2014/2/28	2014/11/14 - 2015/4/17
			2015/8/7 - 2015/8/14	2015/10/16 - 2016/4/8
			2016/5/6 - 2016/7/8	2016/9/30 - 2017/1/6
			2017/3/24 - 2017/3/31	2017/9/15 - 2017/11/24
			2017/12/22 - 2017/12/29	2018/1/19 - 2018/2/23
			2018/4/20 - 2018/4/27	2018/5/18 - 2018/5/25
			2018/8/3 - 2018/8/10	2018/9/21 - 2019/2/1
China	real-estate and stocks	35	2008/6/27 - 2008/7/4	2008/9/26 - 2008/10/3
			2008/11/14 - 2008/11/28	
	real-estate and bond	14	2008/9/26 - 2008/10/3	2008/10/31 - 2008/11/7
	real-estate and forex	28	2008/12/21 - 2008/11/28	2011/8/5 -2011/8/12
			2011/10/21 -2011/10/28	
	stocks and bond	84	2008/10/31 - 2008/11/7	2010/7/2 - 2010/7/9
			2010/10/29 - 2010/11/5	2010/12/17 - 2010/12/24
			2012/12/21 -2012/12/28	2013/10/25 - 2013/11/1
			2013/12/6 - 2013/12/13	2014/1/3 - 2014/1/10
			2016/2/5 - 2016/2/12	2017/3/17- 2017/3/24
			2018/1/5 - 2018/1/12	2018/7/13 - 2018/7/20
	stocks and forex	280	2008/9/5 - 2008/11/7	2010/9/17 - 2010/1//29
			2013/12/6 - 2014/2/21	2014/11/14 - 2015/11/21
			2015/8/21 - 2015/10/9	2015/12/18 - 2016/1/22
			2016/3/11 - 2016/3/18	2016/4/22 - 2016/7/1
			2016/12/16 -2016/12/23	2018/2/23 - 2018/3/2
			2018/4/13 - 2018/5/25	2018/7/6 - 2018/8/17
			2018/10/12 - 2018/10/26	
	bond and forex	98	2008/11/21 - 2008/11/28	2011/8/5 - 2011/8/12
			2015/3/27 - 2015/4/3	2015/8/7 - 2015/8/14
			2015/10/31 - 2015/11/7	2016/11/25 - 2016/12/2
			2018/1/5 - 2018/1/12	2018/7/13 - 2018/7/20
			2018/8/17 - 2018/8/24	2018/12/14 - 2019/1/18

Figure 4.12 and 4.13 provides overall supportive evidences on the implications drew from the R-vine copulas with additional information in terms of dating the contagious periods.

1. There is a compelling indication of existing contagion channels in the U.S. markets, especially between the real-estate and stock markets, as well as the forex and interest rate markets. Moreover, risk contagious episodes related to the real-estate and stock markets are mostly detected during the aftermath of the 2008 Subprime Debt Crisis, whilst the bond and forex markets are more heavily influenced by the risk spreading around mid-2016 when Brexit took place. As Figure 4.12 suggests, finance-related turbulence tends to generate joint effects on the U.S. real-estate and stock markets whilst the forex and interest rate are more likely to show binding reactions towards international political shocks.

2. Although risk spreading could be observed in the China's financial markets, especially after the 2008 global crisis, contagion effects are more temporary in China comparing with the U.S.. Nonetheless, a uniform spillover effect was detected in the real estate, stock, and bond markets against the forex at around 2011, soon after the change of the Renminbi policy to the "crawling-like agreement". Furthermore, the Chinese stock market and the USD/CNH exchange rate demonstrate an increasing intensity in terms of risk transmission after 2014 mainly due to the combined effect of the speeding-up of Renminbi internationalization and the uncertainties raised from the global economic and political changes.

Comparative analysis on spillover effects

In Table 4.6, detected contagion episodes are summarized mainly regarding two cases, i.e. either the real-estate or stocks to be the crisis originator. Meanwhile, total numbers of contagious observations (in the unit of the week) between each pair of assets are provided. Table 4.6 provides supportive evidence on the arguments made according to the smooth probability plots with additional findings summarized as follows.

For the U.S. markets,

1. the risk spillover effects are shown to be more significant than those in the Chinese markets, with overall longer contagion episodes especially between the four pairs of assets, namely real-estate and bond, real-estate and forex, stock and bond, and bond and forex, of which the contagion episodes all exceed 450 weeks.

2. The stock and currency are shown to be the least connected pairs, with the total number of contagion observations being 58 weeks.

The linkages between the Chinese financial sectors in terms of contagion periods are rather different from those in the U.S.. Specifically,

1. the Chinese real-estate market appears to be more stable towards transmitted risks from other markets with overall less than 40 weeks of contagion during the 21 years of observation, whilst the rest three markets are more connected with longer contagion periods.

2. Further, it is suggested that the major risk transmission channel in the Chinese financial sectors lies between the stock market and the USD/CNH exchange rate with 280 weeks of contagion since 2008.

Table 4.7 shows the parameter estimations for Bi-SWARCH-EVT model given the crisis originator of real estate. It is observed that 1) the state-dependent correlation coefficients ρ_2 (for high-vol. state) are greater than ρ_1 (for low-vol. state) for all paired American markets, but the situation conversely performs in China; and 2) for the GPD estimations on $\boldsymbol{\xi}$ (bottom panel), all pairs of markets in both countries perform thicker tails in downside regime (since most values of $\boldsymbol{\xi_L}$ are less than $\boldsymbol{\xi_R}$, except the negative value of $\boldsymbol{\xi_R}$ already implying the light tail on the upper side. Thus, the table not only validates the necessity of hiring the EVT to estimate the heavy-tail effect leading contagion but supplements the understanding of market dependence by comparing the state-varying correlation coefficients between high- and low- volatility states.

Table 4.7: Bi-SWARCH-EVT model estimated parameters.

originator:real-estate

Country	**U.S.**			**China**		
Recipient	**stocks**	**bond**	**forex**	**stocks**	**bond**	**forex**
u_o	0.241	0.301	0.292	0.279	0.022	0.183
	(1.172)	(-0.109)	(2.113)	(1.114)	(2.101)	(-0.215)
u_r	0.217	-0.095	-0.027	0.239	0.105	-0.029
	(1.183)	(-0.101)	(-0.515)	(1.752)	(1.892)	(-0.190)
θ_o	-0.009	-0.002	-0.021	-0.004	0.010	0.002
	(-0.281)	(-0.160)	(-0.241)	(-0.178)	(2.031)	(1.512)
θ_r	-0.064	-0.027	0.060	0.121	0.241	0.056
	(-0.142)	(-0.024)	(3.140)	(0.041)	(2.122)	(0.998)
$\alpha_{o,0}$	5.80	3.725	4.245	14.15	10.701	10.941
	(2.378)	(0.426)	(1.896)	(5.118)	(2.460)	(3.662)
$\alpha_{r,0}$	4.210	6.928	0.282	44.706	33.480	0.018
	(1.974)	(-0.237)	(1.710)	(5.219)	(2.128)	(3.261)
$\alpha_{o,1}$	0.153	0.281	0.000**	0.000**	0.000**	0.000***
	(1.136)	(0.020)	(0.000)	(0.000)	(0.000)	(0.000)
$\alpha_{r,1}$	0.000**	0.132	0.297	0.645	0.638	0.216
	(0.000)	(0.005)	(-0.061)	(1.072)	(0.161)	(-0.083)
$\gamma_{o,2}$	20.469	29.913	19.118	87.421	12.342	7.630
	(1.140)	(2.963)	(1.672)	(6.718)	(5.250)	(2.314)
$\gamma_{r,2}$	14.257	15.631	0.055	95.727	17.796	16.564
	(2.343)	(1.587)	(4.727)	(4.637)	(6.730)	(3.721)
ρ_1	0.180	0.022	0.006	0.495	0.025	0.007
	(3.127)	(1.296)	(2.153)	(1.096)	(-0.128)	(0.032)
ρ_2	0.447	0.189	0.024	0.003	0.007	0.000*
	(2.141)	(1.064)	(2.011)	(1.122)	(0.150)	(0.000)
$\xi_{L,o}$	0.625	0.229	0.503	-0.073	-0.107	0.308
	(2.612)	(1.406)	(2.636)	(-0.472)	(-0.896)	(2.017)
$\xi_{L,r}$	0.522	-0.137	0.336	0.879	0.008	0.359
	(2.124)	(-1.538)	(1.749)	(3.007)	(0.169)	(1.774)
$\xi_{R,o}$	0.332	0.387	0.249	0.029	0.064	-0.073
	(1.749)	(2.045)	(1.539)	(0.179)	(0.454)	(-0.536)
$\xi_{R,r}$	0.159	0.217	0.145	0.586	0.466	0.156
	(0.954)	(1.341)	(1.008)	(2.583)	(2.257)	(1.046)
$\beta_{L,o}$	0.549	0.576	4.285	0.665	0.707	5.349
	(3.661)	(4.292)	(4.311)	(5.007)	(5.105)	(4.781)
$\beta_{L,r}$	0.434	1.547	0.466	2.212	2.797	0.451
	(4.195)	(5.592)	(3.964)	(3.375)	(2.899)	(3.953)
$\beta_{R,o}$	0.587	0.326	4.077	0.662	0.616	7.110
	(4.124)	(4.250)	(4.502)	(5.178)	(5.032)	(5.572)
$\beta_{R,r}$	0.408	1.463	0.656	1.575	1.334	0.644
	(4.308)	(3.996)	(4.636)	(3.976)	(4.155)	(4.587)

The t-value is in the parenthesis below each estimated coefficient. ** and *** denote the significance at level 0.05 and 0.01, respectively.

4.3 Implications

In this chapter, two parallel hybrid approaches are introduced by combining the SWARCH models, EVT and R-vine copulas, with the target of studying the contagion effects of the real estate, stock, bond, and forex markets in China in comparison with the U.S.. In particular, the SWARCH-EVT-Copula model is implemented to obtain information in terms of the crisis periods, tail dependence correlations and risk transmission channels, whilst the Bi-SWARCH-EVT model is conducted mainly to explore the possible dating of the contagion episodes. Regarding the model results, market behaviors in response to major events, including the 2008 global financial crisis, the modification of the Renminbi policy, the Brexit referendum, etc., are explicitly discussed. Key implications of this study are summarized as follows.

Country-wise, the American financial markets are shown to share a higher level of integration and exhibit a clearer linkage than the Chinese markets with the stronger tail dependence and longer contagion periods in crises. The major risk transmission channels in the U.S. are within the three markets of real estate, stock, and bond, whilst the forex is less connected to its domestic markets with a minor negative correlation. In China, the real-estate market demonstrates overall high stability against major economic shocks after 2008, which could be explained by the combined effect of the urbanization process, the money supply on property purchase, and the high expectation on housing prices. Furthermore, the risk spillover effects between the foreign exchange rate and other Chinese financial sectors, especially the stock market, are shown to be temporarily aggravated as reactions to the global political and economic reshapes since 2015.

Market-wise, the real-estate and stock markets not only show the strongest linkage but play the central role in the crisis transmission among all market assets. With empirical evidence, the real-estate and stock markets tend to evolve into the driving force of financial turbulence. Moreover, the bond market receives risk spillovers from both the real-estate and stock markets and takes a longer time to recover from the interest rate triggered crises in comparison with the real-estate and stock markets. In addition, the USD/CHN exchange rate demonstrates a high degree of sensitivity towards risks with the most extended high-volatility

period in the sample and has more reactive behaviors in the Chinese markets rather than in the U.S..

On the basis of increasing linkage across markets affecting the traders' decisions during the financial crisis (Zhang and Liu, 2018; BenSaïda, 2018; Cubillos-Rocha et al., 2019), our study provides key conclusions for the market participants as well as policymakers that A) the proposed hybrid models are the reliable and thorough mechanism of understanding the contagion effects and the risk transmission channels in crises, B) the American financial markets appear to be under greater exposures of systemic risks as a result of its contagious system, and C) despite that the investors seem to be secured from the flourishing real-estate market in China during the last decade, the exchange rate risk emerges to play a more prominent role in the evolution of Chinese markets with greater scales of cross-sector co-movements and acts one of the major linkages to connect the Chinese financial system with global markets.

Chapter 5

Integrated early warning system development

5.1 Conceptual model of integrated EWS

In this chapter, the conceptual model of the integrated early warning system for specific financial market turbulence prediction will be proposed. To the best of our knowledge, this is the first study to sculpt the EWS in a compact way to distinguish different functional zones for the crisis classification and the crisis prediction based on market asset price volatility dynamics.

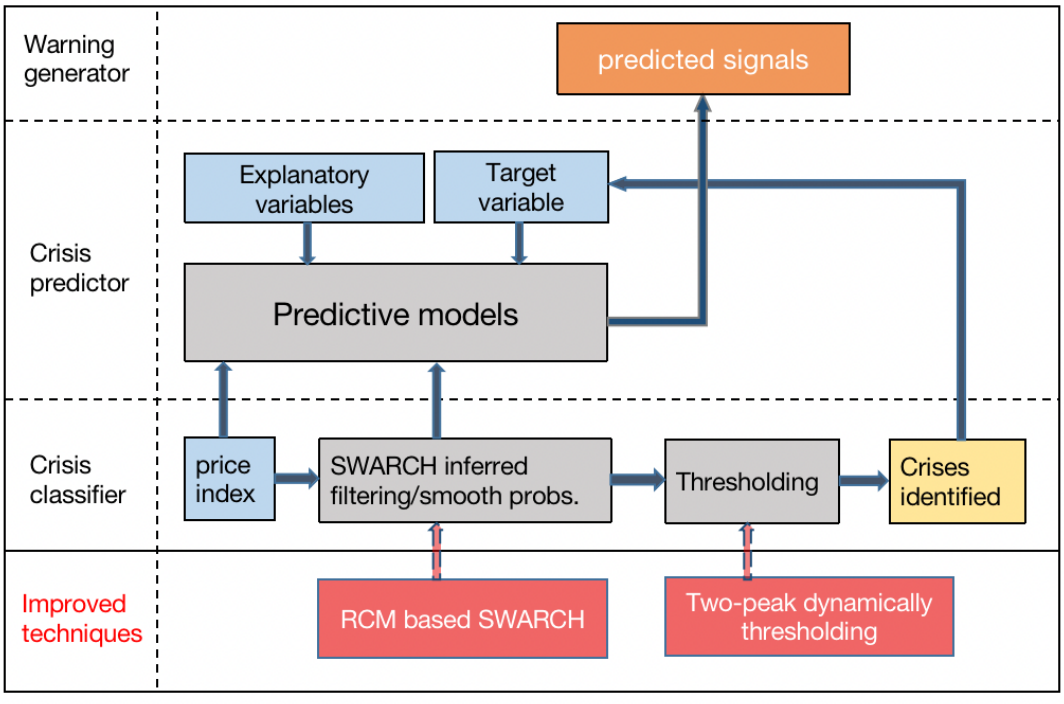


Figure 5.1: The general structure of EWS with three essential components.

Figure 5.1 shows the general frame for an integrated EWS, which is diagrammed with three essential components: *crisis classifier, crisis predictor, and warning generator*.

Starting from the bottom base of the crisis classifier, the price index dynamics will be input into the Markovian switching model to infer the probability of the observation being in either turmoils or tranquility. After being thresholded, the probabilities will be transformed into the identified crisis samples in a binary function, and these identified crisis samples will access to the predictive models as the target variable for further prediction.

In the next functional zone of crisis predictor, predictive models will take both endogenous (i.e. price index including log returns and time-lagged price, filtering or smooth probabilities inferred from SWARCH model) and exogenous factors (i.e. explanatory variables that reflect the economic situations and target market stability), and generate the final predicted warning signals as being input the target variable.

At last, the stratum of the warning generator will output the predictions for further usage in politics and investing activities. That is the normal flow to accomplish the full EWS functioning diagram yet being considered extra techniques to improve the EWS performance on identifying crisis samples and predicting crisis signals.

5.1.1 Crisis classifier: conventional versus improved

As Figure 5.1 highlights at the bottom, two red blocks can be recruited by the SWARCH based crisis classifier as improved techniques, namely regime classification measure (RCM) and two-peak methods. Comparing to the conventional way that takes the fixed number of regimes (usually for high and low two volatility states) and the arbitrary cutoff value (generally 0.5), the improved techniques provide further explorations on determining the number of volatility levels and optimizing the value of cutoffs to boost the appropriateness and robustness in the SWARCH classifying process.

To make a clarified display for improved techniques, we will exploit one for each market of China's bonds and stocks, respectively, i.e. the RCM based

SWARCH model will be applied to the bond market, and the two-peak dynamically thresholding technique to determine the cutoff values will be examined for the stocks.

Arbitrary versus dynamic cutoffs

The thresholding issue for crisis identification has been a noticeable problem for financial studies especially relating to date crisis periods. Lestano (2007) discusses the threshold variation in dating currency crises and proves that the ad hoc thresholds are less sensitive than extreme-value determining thresholds in a recursive selection scheme. However, in the SWARCH frame, rare studies query whether the ad-hoc cutoff of 0.5 is absolutely universal for all asset indexes. That is exactly the initial impetus for us to propose the question (displayed in Chapter 3, Section 3.1) on SWARCH classification ability in terms of cutoff selection and to further explore the dynamical thresholding methodology to investigate whether it will improve the volatility classification based crisis classifier. In our study, the two-peak method is adopted (in the following section of 5.2.1) to accomplish the dynamical thresholding scheme on the SWARCH inferred probabilities. The specific reason and application process of adopting the two-peak method will be described in Section 5.2.1

Intuitive versus metric-determined volatility levels

Corresponding to the other aforementioned question on SWARCH model defined crisis classifier, two regimes of high and low volatility levels seem hardly competent to adapt all price index dynamics with varied fluctuation dispersion features. In fact, the original form of SWARCH allows K (more than two) regimes (Hamilton, 1989, 1994) but most studies opt for 2 to simplify the probability inference process and avoid tangled calculation for high dimensional transition matrix[1]. However, in practice, this simplified operation will not improve the accuracy of SWARCH model's identification of crisis samples. Even worse will bring a chaotic crisis probability inference pattern for the assets that have not fluctuated enough

[1]Appendix A shows the full structure of regime-transition probability matrix for $K = 2$ with two variables, which has been expanded to 16×16 dimension. The greater regime K, the more complex the transition matrix.

in the observed period. In our study, the appropriateness of using two volatility regimes will be examined, and then the optimization on the count of regimes will be made by RCM, the measurement which is proposed to determine the count of volatile states in the SWARCH frameworks (Ang and Bekaert, 2002), based on bond asset data. The application process will be specified in Section 5.3.4.

5.1.2 Crisis predictor: classic versus stylized

Besides owning a robust crisis classifier, the other key to generate effective warning signals in the EWS frame is to acquire a powerful crisis predictor. A number of econometric models based on either parametric or non-parametric methods (Eichengreen et al., 1995; Frankel and Rose, 1996; Kaminsky and Reinhart, 1999; Abiad, 2003; Bussiere and Fratzscher, 2006; Dawood et al., 2017) have been proposed to fulfill such predicting role. These classical models, though they have strong interpretability on economic significance, the performed forecasting precision is far from the asset market investment's requirement for timeliness, especially in the contemporary data-booming era. The stylized machine learning models seem to remedy the classic models' deficiency in forecasting capability with mighty computing power and operating flexibility, even though they bear the blame of over-parameterization and in-interpretability. In following sections, a comparative analysis will be made in the horse race 1) between machine learning models with different structural complexity (such as simplistic neural networks and complex deep neural networks) and 2) between classic (such as logit regression) and machine learning models (such as attention based neural networks) for specific markets. This horse race also inspires the further exploration on improving the machine learning models' performance on detecting leading indicators and validating the detected factors' reliability, which will be discussed in Chapter 6 for the contagion fused EWS development.

The rest of this chapter's content will be arranged as follows. The EWS for China's stock market will be first introduced by experimenting with the two-peak dynamically thresholding instructed crisis classifier and deep neurons of LSTM networks embedded predictive model in Section 5.2. Then, Section 5.3 will construct the alternative EWS, which aims to solve the problems of regime number

selection and extracting leading indicator factors extraction by i) adopting the RCM based crisis classification and ii) stacking the attention into the deep neural structure. Section 5.4 will summarize the implications from experiment results and discuss unsettled questions linked to further works for constructing the contagion fused EWS.

5.2 EWS for China's stock market

In this section, we propose the stock EWS for China with threefold aims.

First, we attempt to develop a precise and robust crisis classifier to identify daily stock market turbulence on the basis of switching ARCH (SWARCH) model (Hamilton and Susmel, 1994) and two-peak (or valley-of-two-peaks) thresholding (Rosenfeld and De La Torre, 1983) technique. In our case, the classification way in two regimes of high- and low- volatility states that respectively imply the market asset turbulent and tranquil episodes (Hamilton and Susmel, 1994; Hamilton and Gang, 1996; Ramchand and Susmel, 1998; Edwards and Susmel, 2001) will be adopted. Furthermore, the two-peak method will be applied in a forward sliding approach (Jain et al., 1995) to automatically determine the threshold for more robust segmentation based on time-varying observations.

Second, to generate effective warning signals for stock turbulence on a daily basis, the dynamic scheme for EWS is implemented by integrating the improved SWARCH based crisis classifier, and LSTM neural networks (Jordan, 1997) based crisis predictor. The LSTM is proven to be a state-of-art mechanism in the general field of financial forecasting (Chen et al., 2015; Fischer and Krauss, 2018; Wu and Gao, 2018; Cao et al., 2019), including volatility forecasting (Yu and Li, 2018; Kim and Won, 2018; Liu, 2019).

Last, a comprehensive evaluation of the integrated EWS should be conducted by examining the crisis classifier and predicting robustness separately. Specifically, the integrated EWS proposed crisis classifier will be compared with the CMAX based crisis indicator function, which has been widely used to define stock crises. And then, the LSTM crisis predictor will be evaluated upon two baseline machine learning models of the back-propagation neural network (BPNN) and

support vector machine (SVM), regarding the statistical metrics including the rand accuracy (Rand, 1971), binary cross-entropy loss (Shannon, 1948), receiver operating curve (ROC), area under curve (AUC) (Metz, 1978) and the SAR score (Caruana and Niculescu-Mizil, 2004). To evaluate the stability for proposed EWS as a whole, the early warning algorithm is not only performed on the test set but testified in k-fold cross-validations as well.

5.2.1 Crisis identification: two-peak determined dynamic threshold

Stock crashes are inevitable results of volatility jumps. Thus high/low volatility regimes of the stock price oscillation are investigated based on the SWARCH model (Hamilton and Susmel, 1994) to reflect the crisis/non-crisis periods. The classification of high/low volatility regimes can be implemented on the basis of SWARCH model inferred filtering probabilities, which is a byproduct of the maximum likelihood estimation (which derivation refers to Chapter 3 and Appendix A). Here, the model of AR(1)-SWARCH(2,1) is adopted, which has been specifically formulated in Chapter 3, Section 3.1 to define the crisis classifier.

The filtering probability based on historical observations till time t can be written as

$$P(s_t = i|\mathscr{Y}_t; \boldsymbol{\theta}), \tag{5.1}$$

where $\boldsymbol{\theta}_t$ is the vector of model parameters to be estimated, can be interpreted as the conditional probability based on the current information $\mathscr{Y}_t$.

Then, the crisis classifier can be defined as the following binary function.

$$\text{Crisis}_t = \begin{cases} 1, & P(s_t = 2|\mathscr{Y}_t; \hat{\boldsymbol{\theta}}_t) \geq c, \\ 0, & \text{otherwise}, \end{cases} \tag{5.2}$$

where $\hat{\boldsymbol{\theta}}_t$ is the estimated parameter vector and c is the cutoff (that is assigned value between 0 and 1) for identifying crises.

In this way, the stock crisis prediction, as well as the early warning signaling,

are structured as a binary classification problem through the mechanism that filtering probabilities of the system being in the high volatility regimes tend to increase as the stock price becomes more volatile. There exists a threshold c that identifies the occurrence of a potential crisis once it is exceeded. The level of c indicates the lowest-level severity of the stock market turmoil that is classified as "crisis", of which the determination can be a tricky problem.

The two-peak method is thus adopted to automatically select the crisis threshold that balances the trade-off between sensitivity and false alarms (Babecký et al., 2014). It is first developed with the general purpose of finding the optimal threshold in the context of binary classification and is proven experimentally credible as solving image processing-related classification problems [2]. According to the two-peak method, the optimal threshold of a binary classification system is the minimum value between two peaks of the observation's frequency density histogram (Weszka, 1978). There are several other automatic thresholding mechanisms that are built on the histogram, such as the Otsu's method (Ohtsu, 2007) that solves the multi-threshold problem by considering the pixel variance. Here we use two-peak as it is the most straightforward and the foundation of other approaches thereafter.

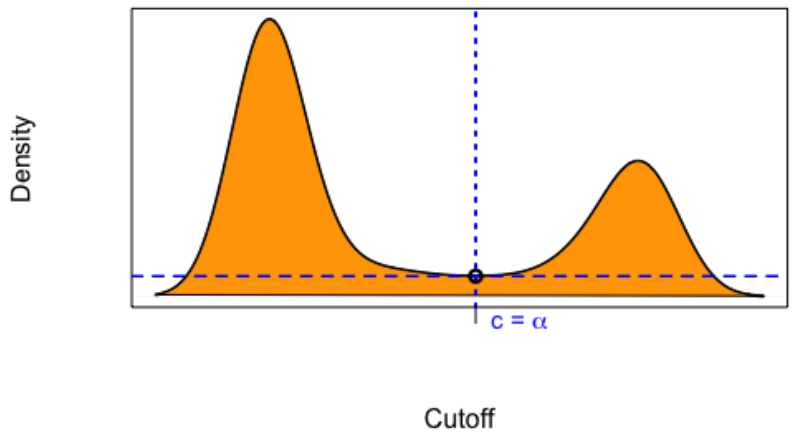


Figure 5.2: The two peak thresholded valley value for the cutoff $c = \alpha$.

[2]Prewitt and Mendelsohn (1966) first introduce the two-peak method in the cell image analysis of distinguishing the gray-level difference between the background and the object. The performance of the method is further verified in Rosenfeld and De La Torre (1983) by analyzing the histogram's concavity structure.

Given that our crisis classifier has two state classes of crisis (1) and non-crisis (0), the two-peak method is applied to determine the cutoff of our crisis classifier by the following steps:

1. Till time t, the SWARCH produces filtering probabilities for the high-volatility state $P(s_t = 2|\mathscr{Y}_t; \hat{\boldsymbol{\theta}}_t)$;

2. Sketch the histogram of all high-volatility filtering probabilities from time 0 to t, in which the valley bottom of two peaks is selected as the optimal cutoff point. (Figure 5.2 examples the histogram with a valley point of α, which will be taken as the optimal cutoff for thresholding the filtering probabilities);

3. Perform the two-peak method on a recursive basis to obtain dynamic thresholds as the prediction moves forward along the time horizon. (See Algorithm 1 in the following section).

5.2.2 Forecasting system: LSTM based crisis predictor

The LSTM is hired as the predictive model to estimate the occurrence of stock market turmoils on daily basis based on the set of historical information with a fixed window size l. Figure 5.4 shows the predicting procedure that is made from a network of l LSTM memory blocks sequentially processing the input of both the explanatory variables $\{\boldsymbol{x}_{t-l+1}, ..., \boldsymbol{x}_t\}$[3] and the vector of filtering probabilities for the high-volatility state $\{P[s_{t-l+1} = 2|\mathscr{Y}_{t-l+1}; \hat{\boldsymbol{\theta}}_{t-l+1}]$,...,$P[s_t = 2|\mathscr{Y}_t; \hat{\boldsymbol{\theta}}_t]\}$ from time $t - l + 1$ to t, for $t \geq l$.

The final output $\hat{y}_{t+1}$ is produced by a sigmoid function to squash the value in the range of [0,1], indicating the probability of the observed sample being in high-volatility at $t + 1$. Early warning signals are thus released for time $t + 1$ once the value of $\hat{y}_{t+1}$ exceeds the optimal crisis threshold selected by the two-peaks method (refer to Section 5.2.1). Given the full sample size of T days, $T - l$ predictions will be made from $t = l+1$ onward. The LSTM network consists of 21[4] input layers, 32 LSTM layers, and the output layer, which brings 5921 parameters to be trained. The batch size and epoch numbers are 20 and 100, respectively.

[3]where $\boldsymbol{x}_{t-l+1}$ is the vector $[x_{1,t-l+1}, x_{2,t-l+1}, ..., x_{N,t-l+1}]$ for N variables

[4]For the stock EWS study, the count of input features is 21 in total, thus the input layer should keep the same with it.

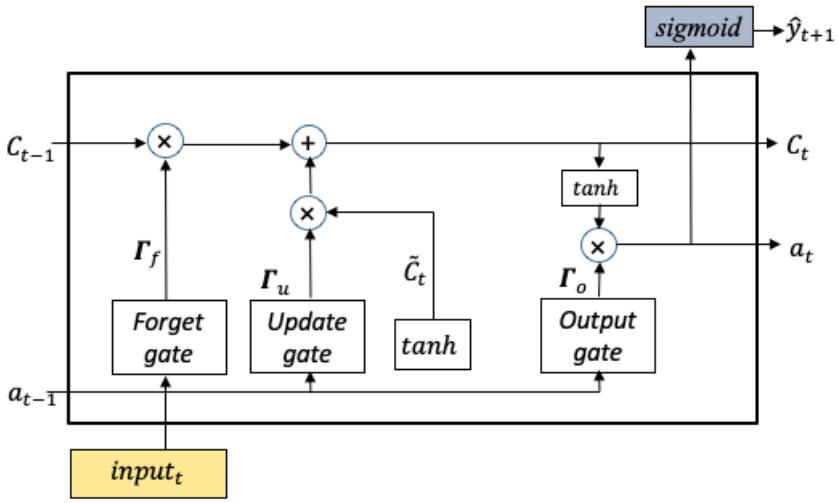


Figure 5.3: LSTM cell for stock EWS.

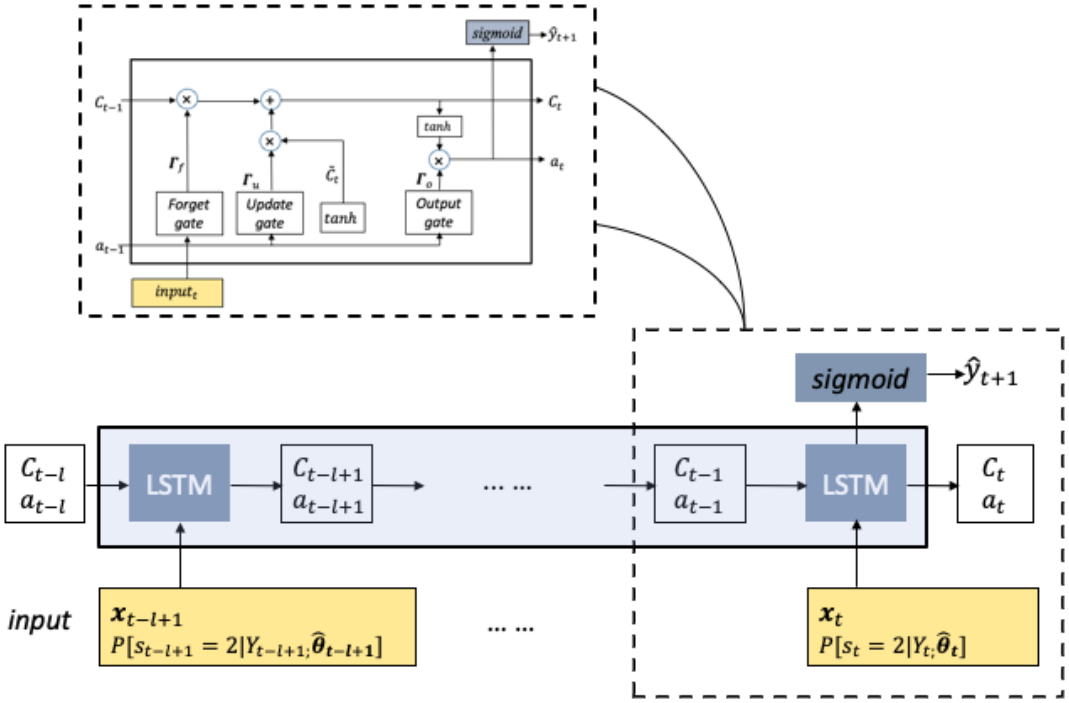


Figure 5.4: LSTM with window size l.

The LSTM cell structure (see Figure 5.3) mainly duplicates that is depicted in Chapter 3, Section 3.3 (see Figure 3.3)). Note that, the difference is the output at time t being substituted by the sigmoid function at the upper right corner, which is only included in the last cell of the full LSTM networks to produce the final prediction of $\hat{y}_{t+1}$ in $[0, 1]$.

Algorithm 1: Draw daily warning for China's stock turbulence.

Initial inputs:
The SSEC index price P_t;
The explanatory variables $\boldsymbol{x}_t$ excluding P_t;

Final output :
The predicted signals $\hat{y}_{t+1}$;

calculate log returns of SSEC index price, $\{logR_t, t = 1, ..., T\}$;
set up the window size l;
for t from l to T **do**
  repeat
    for i from 1 to $t+1$ **do**
      input $logR_i$ into SWARCH;
      output filtering probability $P[s_i = 2|\mathscr{Y}_i; \hat{\boldsymbol{\theta}}_i]$;
    end
    two-peak method selects the optimal cutoff c_t;
    for i from 1 to $t+1$ **do**
      if $P[s_i = 2|\mathscr{Y}_i; \hat{\boldsymbol{\theta}}_i] \geq c_t$ **then**
        $\text{Crisis}_i = 1$;
      end
    end
    for j from 1 to l **do**
      input explanatory variable vector $\boldsymbol{x}_{t-l+j}$, filtering probability $P[s_{t-l+j} = 2|\mathscr{Y}_{t-l+j}; \hat{\boldsymbol{\theta}}_{t-l+j}]$ and identified crisis signals Crisis_{t-l+j} into LSTM neural networks;
    end
    output the prediction $\hat{y}_{t+1}$;
  until t=T;
end

Thus, the full picture of integrated EWS for the stock can be drawn by implementing the procedure of that 1) the crisis classifier first identifies stock market turmoils according to Eq. 5.2 based on the SWARCH filtering probability of the high-volatility state and the improved technique for automatically determining the crisis cutoff by the two-peak method, 2) the identified crisis samples from the classifier enter into the target variable and are meanwhile fed into the predictive model of LSTM networks together with other explanatory variables listed in Table 5.1, and 3) early warning signals are generated as the LSTM predicted results as exceeding the determined cutoffs on a dynamically-recursive basis. The procedure is programmed by Algorithm 1 on the full sample of size T.

5.2.3 Data

In this study, the Shanghai Stock Exchange Composite (SSEC) index is hired to reflect the Chinese stock market oscillation. Explanatory variables that are incorporated to predict stock crises are described in Table 5.1 in terms of frequency, reflection and source. Specifically, endogenous factors include the close price, log return and daily realized volatility[5] of the SSEC index. The rest of the variables are exogenous factors of four genres reflecting the U.S. stock market, currency level, global and domestic economies, respectively. The samples span from Oct. 7, 2008 to Sept. 18, 2018 and are split into 70% training and 30% test sets.

Table 5.2 shows the full sample statistics of data included in this study. The correlation heat-map of explanatory variables is displayed in Figure 5.5. Most input variable factors are weakly correlated with each other, which gives the maximal orthogonality of the input information, except some empirical correlations A) between the US stock index and Chinese domestic macroeconomic factors (of CPI, M1, M2, and Fixed asset investment gain), B) between the US stock index and the Chinese real-estate price, C) between the Chinese real-estate price and the domestic macroeconomic factors (of CPI, M1, M2 and Fixed asset investment gain) and D) among the domestic macroeconomic factors (of CPI, M1, M2 and Fixed asset investment gain).

As the fixed investment gain and real-estate price index are strongly correlated with the previously entered domestic factors of CPI, M1 and M2, they may not sufficiently contribute extra information to model predictions. Furthermore, the domestic economic variables seem to be hooked with U.S. stock market index, which reflects globalization on the one hand but lessens the dependency of simultaneously referring to both domestic and overseas economic indicators to warn the Chinese stock market turmoils on the other hand.

[5]The daily realized volatility at time t is defined as $\sigma_{rv} = \sqrt{\frac{1}{N_t}\sum_{t=1}^{N_t}(p_t - \bar{p_t})^2}$, where the N_t is the count of days after time t, p_t is the log return at t and $\bar{p_t}$ is the average of log return til t.

Table 5.1: Description of explanatory variables.

Data	Freq.	Reflection	Source
Close price, Log return, Realized volatility	Daily	Local stock market	WIND database
10-year government bond index, Pledge repo rate, Difference between a) 1-week and 1-day, b) 1-month and 1-week SHIBOR rates	Daily	Local fixed-income market	WIND database
Real-estate price index	Daily	Local house market	Yahoo Finance
USD/CNY exchange rate,	Daily		
Foreign exchange reserve, Foreign exchange liability	Monthly	Local currency market	US Federal Reserve Board
Interest rate for China (published by IMF), CPI, M1, M2, Fixed-asset investment gain rate (cumulative year-on-year basis)	Monthly	Domestic macroeconomic factor	WIND database
Investors' sentiment index (published by National institute of development, Peking University)	Monthly	Domestic investors' sentiment	bbs.pinggu.org
S&P500 price index, VIX	Daily	US stock market	Fred St. Louis
Gold Price	Daily	Commodity price	World Gold Council
Oil Price			International Monetary Fund

Table 5.2: Statistics of explanatory variables.

	Size	Mean	St.Dev.	Skewness	Kurtosis	Jarque-Bera
SSEC close price	2427	2766.65	560.77	0.68	1.01	291.46***
SSEC log return	2426	0.02	1.49	-0.78	4.86	2643.2***
SSEC realized volatility	2426	1.7	0.31	1.86	4.05	3069.1***
10-year Gbond index	2433	117.46	3.08	-0.04	0.01	0.5650***
SHIBOR7d_1d	2438	0.53	0.45	1.9	9.82	10825***
SHIBOR1m_7d	2438	0.72	0.55	0.43	0.28	79.137***
Repo	2439	0.67	0.44	1.70	7.23	6228.7***
Real-estate price index	2505	16.46	4.8	0.7	0.74	242.75***
USD/CNY exchange rate	2499	76.49	0.27	0.06	-1.46	217.38***
Forex_reserve	121	31229.5	5488.29	-0.61	-0.22	113.55***
Forex_liability	121	13.94	18.15	1.08	1.74	750.47***
Interest rate for China	121	3.06	0.22	0.73	2.22	717.48**
CPI	121	95.83	6.78	-0.4	-1.04	174.59***
M1	121	3.38†	1.08†	0.47	-0.79	151.87***
M2	121	1.10†	3.91†	0.12	-1.21	154.91***
Fixed_investment	121	18.29	7.99	0.05	-1.17	133.87***
ISI	120	53.95	20.33	2.02	5.83	4906.2***
S&P500 Index	2483	1682.81	529.84	0.19	-1.04	124.03***
VIX	2489	19.43	10.08	2.48	7.48	7822.5***
Gold Price	2475	1296.08	231.33	0.24	-0.14	26.129***
Oil Price	2493	73.25	22.88	-0.12	-1.41	208.00***

St.Dev. is the standard deviation. $***$ and $**$ denote the (null normal) hypothesis test at the 1% and 5% significance level. † denotes the unit of M1 and M2 is 10^{13} Chinese yuan.

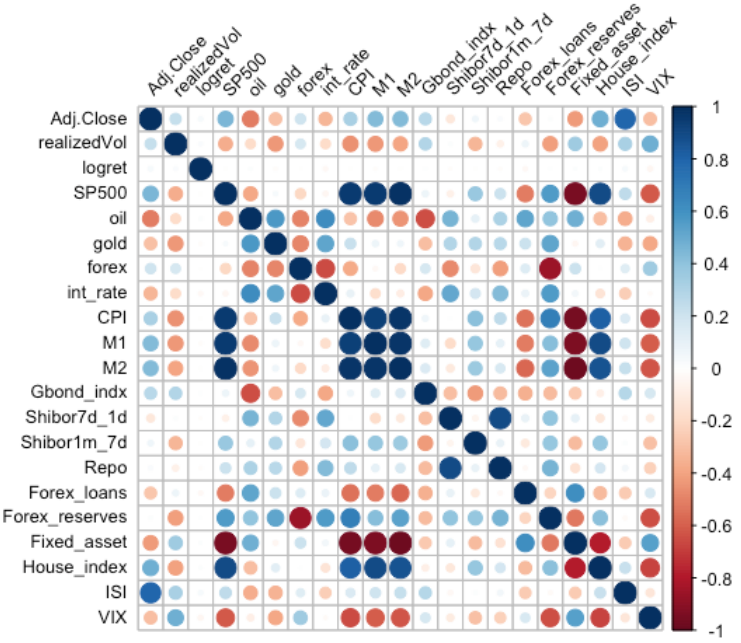


Figure 5.5: Correlogram for input variables.

5.2.4 Crisis classifier robustness

The credibility of an EWS has rooted in a precise and robust crisis classifier. According to Figure 5.4 and Algorithm 1, stock crisis cutoffs are dynamically computed for each prediction taking into account the current market condition as well as past information. To validate the reliability of the proposed classification mechanism, we analyze the crisis identification results in terms of their precision and robustness.

Chronological evidence

As crisis classification is a subjective topic heavily depending on the individual understanding of the crisis, the limited analysis could be done on quantitatively evaluating the accuracy due to the lack of true crisis labels. Similar to empirical studies for crisis prediction, we investigate the precision of the crisis classifier with emphasis on the empirical evidence related to volatility regimes.

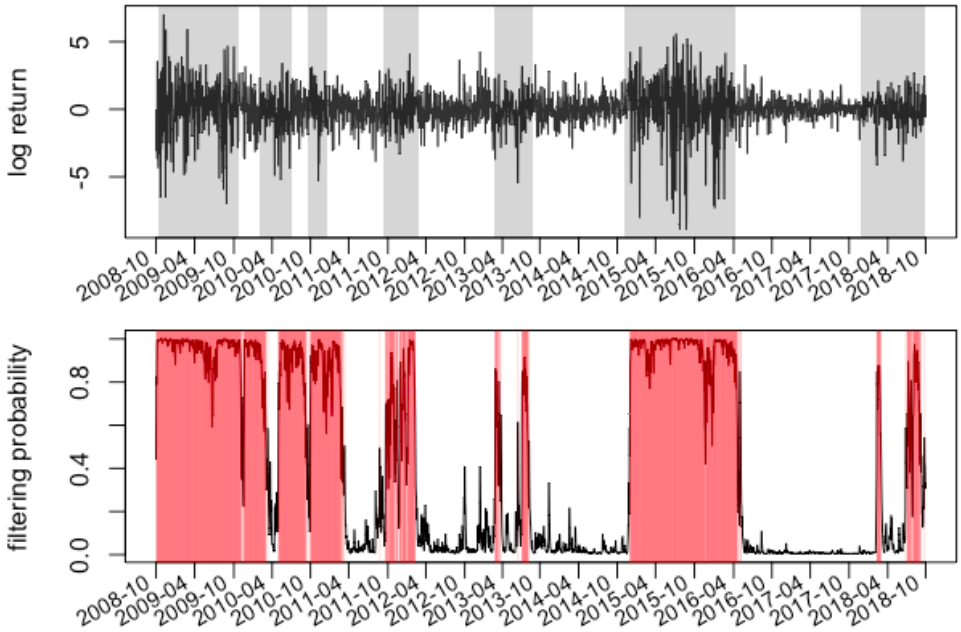


Figure 5.6: Log return of the SSEC stock index (upper panel) and the corresponding high-volatility filtering probability (lower panel).

Figure 5.6 highlights the identified crisis periods in both the log returns (grey in the upper panel) and filtering probability plots (red in the lower panel). The figure suggests that the proposed hybrid algorithm successfully capture all recorded stock crises being reflected by volatile log returns and filtering probability jumps.

Table 5.3: Stock market critical events and the EWS identified crisis episodes.

Critical Event	Sign	Identified turmoil episodes
2008 Global financial crisis	SSEC ended with 65% down in the last quarter of 2008.	2008/10/07 - 2009/11/06; 2009/11/16 - 2010/03/28.
2010 European debt crisis	SSEC dropped by 27% in the first half of 2010 and 15% in the whole year.	2010/05/06 - 2010/09/16; 2010/10/08 - 2011/03/17.
–	–	2011/10/08 - 2012/02/17.
2013 Industrial reformation	China's stock market is remarked as the weakest in the Asia with the 6.75% drop over the year of 2013.	2013/03/04 - 2013/03/28; 2013/07/11 - 2013/08/12.
2015 Chinese stock crash	A third of A-shares (on SSE) value was evaporated within one month from June 12, 2015.	2014/12/02 - 2016/04/27; 2016/05/09 - 2016/05/11.
2018 Sino-US trading war	China's stock market lost 2.3 trillion dollars (about 25% market value) in 2018.	2018/02/09 - 2018/03/06; 2018/07/02 - 2018/08/03; 2018/08/06 - 2018/08/31; 2018/09/04 - 2018/09/18.

Table 5.3 extracted the start, and end dates for the EWS classified turmoil periods by performing Algorithm 1 on the full sample in accordance with chronological shreds of evidence of critical events on China's stock market. During the recent decade from 2008 to 2018, five critical events truly emerged: 1) the 2008 global financial crisis triggered 65% decline of SSEC index, 2) the 2010 European debt crisis caused an annual drop of 15%, 3) the 2013 industrial reform led a greatest daily drop on 5.3% and the annual decline of 6.75%, 4) the 2015 China's stock market turbulence accumulated more than 30% devaluation of A-shares SSE within one month, and 5) the 2018 Sino-US trading war brought about 25% loss for the China's stock index till the end of 2018. Thus comparing to these recorded events, the crisis classifier identified turmoil periods (the last column in Table 5.3) place a promising match except the episode of 2011/10/08 - 2012/02/07, which cannot be explained by the chronology.

Comparison with CMAX based classifier in statistics

The robustness of a model broadly refers to its error-resisting strength and resilience in producing results as data change. Therefore, robust crisis classifications are subject to a dynamical thresholding mechanism to handle turbulence with limited influence from sample variations.

Table 5.4: Statistics of crisis cutoffs in the full sample and test set.

	Count	Mean	St.Dev	Median	Mode	Range
$\text{Cutoff}_{\text{full-sample}}$	2408	0.513	0.119	0.478	0.481	1.00
$\text{Cutoff}_{\text{test-set}}$	720	0.430	0.128	0.396	0.354	0.997

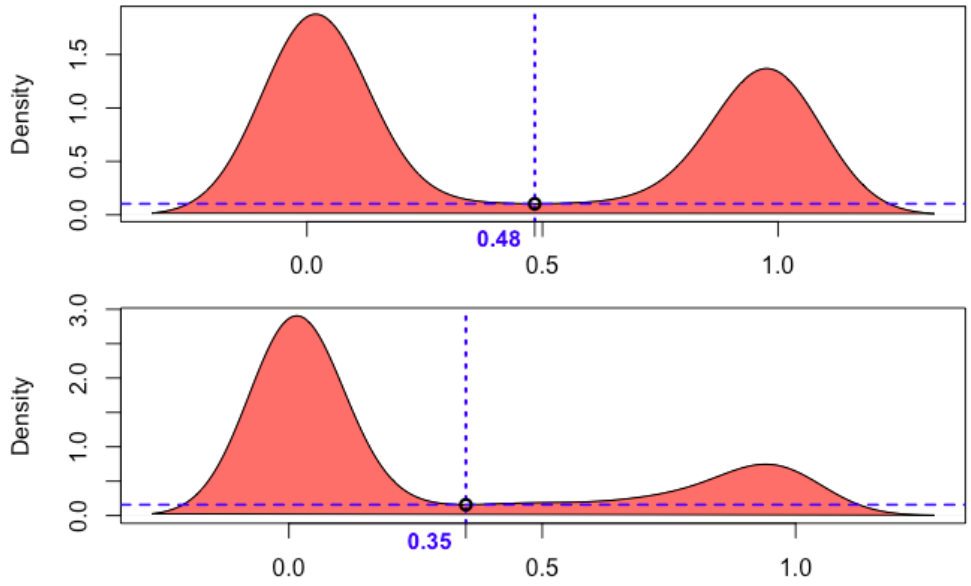


Figure 5.7: Cutoffs selected by the two-peak method in the full sample (upper panel) and test set (lower panel).

Table 5.4 summarizes the statistics of crisis cutoffs that are determined in the full sample and test set by Algorithm 1. The number of cutoffs in a sample is given by the difference between the number of observations T and the window size l. With windows of size 5 (days), this study computes 2408 and 720 cutoffs in the full sample and test set of lengths 2412 and 724 (days), respectively. As Table 5.4 displays, the cutoff distributions of the full sample and test set are both right-skewed given the greater means (0.513, 0.430) than the medians (0.478, 0.396) and modes (0.481, 0.354). In other words, the positive skewness indicates that cutoffs are more likely to take values below the mean and around the median/mode.

Moreover, test-set cutoffs exhibit lower values with mean, median, and mode approximating to 0.4, whereas those in the full sample are closer to 0.5.

To explain this difference in the crisis cutoff distributions, Figure 5.7 shows the histograms of SWARCH filtering probabilities in the full (upper panel) and test (lower panel) sets. The optimal cutoffs determined at the end of Algorithm 1 for the last day observation are circled in blue. Although the test set exhibits a greater proportion of tranquil days with a significantly higher right peak, the two-peak method detects the true valley at 0.35 to threshold the crisis.

Table 5.5: Difference between crises identified on the full sample and test set.

	Integrated EWS	$\text{CMAX}_{\lambda=1}$	$\text{CMAX}_{\lambda=1.5}$	$\text{CMAX}_{\lambda=2}$	$\text{CMAX}_{\lambda=2.5}$
On full samples	195	210	109	0	0
On test set	201	155	161	128	55
non-match days	**6**	55	52	128	55
No. of days	724	724	724	724	724
% of non-match	**2.07**	7.86	7.43	18.3	7.86

With the argument that a robust classification model ought to produce stable classification results regardless of the sampled information, Table 5.5 compares stock crises identified by Algorithm 1 with those defined on the CMAX indicator[6]. Daily classifications are computed in both the full-sample and test set for each model. To examine the level of consistency between crises identified on different samples, Table 5.5 lists the number (Row 3) and percentage (Row 5) of days that the full-sample crises differ from the test-set crises during the period from 2015/10/08 to 2018/09/18 (724 days in total)[7]. With six days of deviation in a period of almost three years and a percentage of 2.07%[8], the integrated EWS produces the most robust crisis classification result in comparison to the CMAX indicator on a range of parameters $\lambda = 1, 1.5, 2, 2.5$.

[6]The CMAX index is the most widely used crisis indicator in the literature concerning stock market early warning (Coudert and Gex, 2008; Li et al., 2015; Fu et al., 2019). It defines stock crashes with an indicator function $1_{CMAX_t<\mu_t-\lambda\sigma_t}$, where μ_t and σ_t are the mean and standard deviation of $CMAX_t$, and λ is a market-dependent constant (Kaminsky and Reinhart, 1999).

[7]This is the period when full sample and test set intersect.

[8]We believe that the percentage deviation of 2.07% could be further reduced with a larger sample of test set and cross validation. Relevant analyses on this aspect will be conducted in future study.

5.2.5 Evaluation in statistical metrics

We now evaluate the crisis predictor based on LSTM in comparison to two other predictive baselines of BPNN and SVM. The evaluation metrics of the predictive model include three classes of performance measures that are designed for classification models, (I) the rand accuracy (Rand, 1971) and binary cross-entropy loss (Shannon, 1948), (II) the receiver operating curve (ROC) and area under curve (AUC) (Metz, 1978), and (III) the SAR score (Caruana and Niculescu-Mizil, 2004). Prior to the performance evaluation, Table 5.6 lists the true or false positive and negative statistics (also known as the confusion matrix), which will be used by the ROC, SAR score and accuracy measures.

Table 5.6: Confusion matrix for daily stock early warning.

Actual \ Predicted	1: Crisis	0: Non-crisis
1: Crisis	True positive (TP)	False negative (FN)
0: Non-crisis	False positive (FP)	True negative (TN)

In general, true positive/negative corresponds to true predictions of turmoil/tranquility, whereas false positive/negative corresponds to false predictions. Moreover, the true positive rate (TPR) and false positive rate (FPR) are defined as the percentage of truly predicted crisis signals over the total number of actual crises, and the percentage of falsely predicted crisis signals over the total number of actual tranquility, respectively.

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}. \tag{5.3}$$

Evaluation Metric I: The rand accuracy of a binary classification problem is defined as the proportion of true results over the total number of cases examined as Eq. 5.4 defined. The binary cross-entropy loss measures the performance of classification models in terms of the level that the predicted probability of getting 1 deviates from the true label 0 or 1, and it is expressed as Eq. 5.5,

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \tag{5.4}$$

$$\text{Loss} = -\frac{\sum_{i=1}^{n-l+1}(y_i log(\hat{y}_i) + (1 - y_i) log(1 - \hat{y}_i))}{n - l + 1}, \tag{5.5}$$

where y_i and $\hat{y}_i$ denote the true and predicted values of either 0 or 1, n is the sample size. As we set the label for the crisis True (= 1), an EWS model that warns all the True crisis despite of the number of False alarms it creates, has zero loss indicating none of the crisis is lost. A greater level of predictive power is believed to come along with higher rand accuracy and lower binary cross-entropy loss.

Evaluation Metric II: As one of the most classic performance measures, ROC plots the FPR (x-axis) against the TPR (y-axis) for each classifier. As a higher true positive rate is always more preferable given the level of the false positive rate, models with more dominating performances are those with ROC bending closer towards the left-hand side. To offer a quantitative representation of the graphic information carried by ROC, AUC computes the total area under the ROC curve and suggests a better model with the greater AUC value.

Evaluation Metric III: Different from the widely-used F1-score, the SAR score (Caruana and Niculescu-Mizil, 2004) is developed as a more holistic performance measure due to the uncertainty of the correct evaluation metric. By taking into account, three distinctive measures, including the accuracy, AUC, and root mean-squared error (RMSE), models with higher SARs are regarded as better-performing as they produce overall high accuracy/AUC and low RMSE.

$$\text{SAR} = \frac{1}{3}(\text{Accuracy} + \text{AUC} + (1 - \text{RMSE})). \tag{5.6}$$

To compare the predictive power of LSTM with baseline models of BPNN and SVM, Table 5.7 preliminary lists the test-set rand accuracy and binary cross-entropy loss of the three models by following Algorithm 1[9]. Three window sizes $l = 22, 10, 5$ are considered to cover both long- and short- term prediction.

As Table 5.7 suggests, LSTM with window size $l = 5$ produces the optimal crisis prediction that yields the highest accuracy 0.964 and lowest loss 0.165 among all cases examined. LSTM consistently demonstrates the greatest forecasting power of stock crises among the three predictive models given different window sizes. Moreover, it is observed that with the last five days of information, all the three models achieve the best result (except the accuracy of SVM) in compar-

[9]To obtain the baseline results, Algorithm 1 is implemented by replacing the LSTM in line 16 by BPNN and SVM.

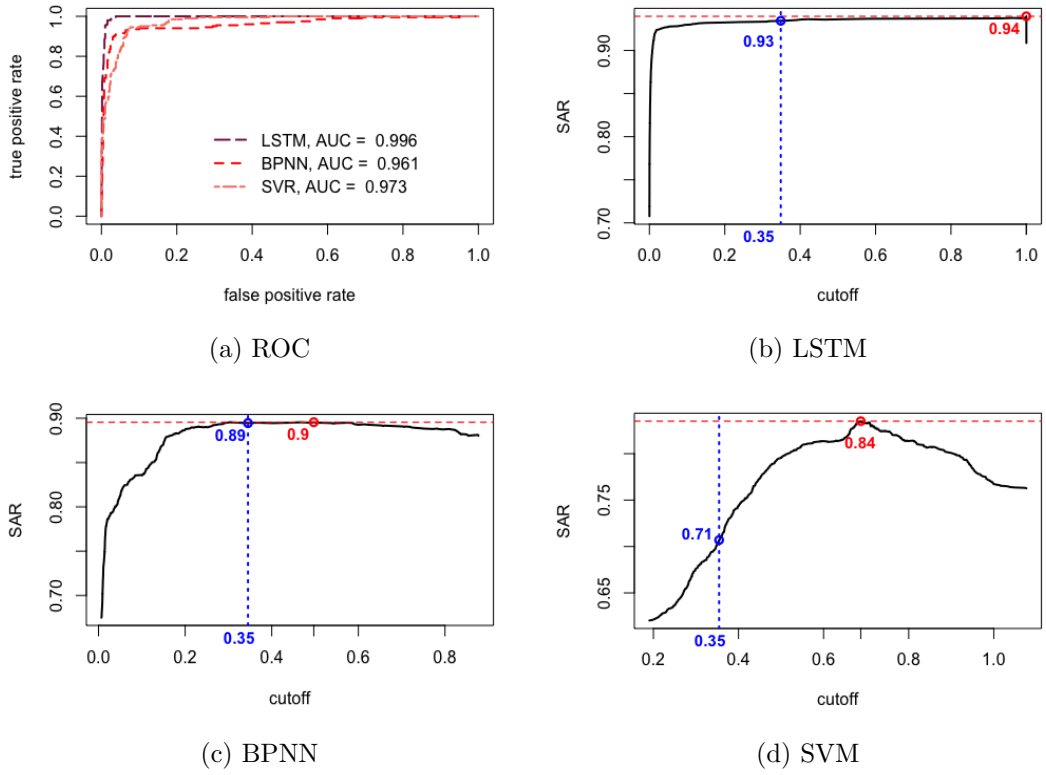
Table 5.7: Rand accuracy and binary cross-entropy loss for each model with varied window sizes on test set.

Model	EWS-LSTM	EWS-BPNN	EWS-SVM
Window size $l = 22$			
Accuracy	0.915	0.863	0.919
Binary cross-entropy loss	0.288	0.465	0.521
Window size $l = 10$			
Accuracy	0.928	0.921	0.931
Binary cross-entropy loss	0.274	0.486	0.442
Window size $l = 5$			
Accuracy	**0.964**	0.947	0.928
Binary cross-entropy loss	**0.165**	0.239	0.399

ison to the predictions made with 22 and 10 days information. Therefore, the remaining evaluation is conducted with a window size 5.

Figure 5.8 further shows ROC and SAR curves for the test set. Particularly, Panel (a) shows the ROC curves and AUC values generated from the test-set predictions. As the ROC-oriented metric tells the model's ability to classify the binary states, LSTM enhances BPNN and SVM with its outstanding capacity to distinguish turbulence/tranquility with the optimal ROC curve and AUC value of 0.996. Panels (b)-(d) display the SAR score against the crisis cutoff for the three predictive models. The score value at the test-set cutoff line is highlighted as the blue point in each panel corresponding to the last day cutoff obtained from the dynamic crisis classifier, whereas the red point is the largest score obtained by the predictive model regardless of the optimal cutoff.

From the perspective of model scores, LSTM remains its dominating state with the highest test-set score (blue) of 0.93, whereas BPNN and SVM score 0.89 and 0.71, respectively. Moreover, LSTM appears to be the most insensitive model to cutoff variations as the scores remain relatively high in a prolonged range shaped like a flat peak in Panel (b). With a similar shape in Panel (c), BPNN produces a SAR curve with a reduced test-set score 0.89 and a smaller peak 0.9. SVM produces the sharpest SAR curve among three models with the smallest peak of 0.84, which performs a vulnerable instability in predicting as the cutoff varies.



(a) ROC

(b) LSTM

(c) BPNN

(d) SVM

Figure 5.8: Test-set ROC (Panel a) and SAR (Panel b-d) curves of LSTM, BPNN and SVM.

5.2.6 Forewarned effectiveness on test set

This section examines the integrated EWS in terms of its true "early warning" power to the actual duration ahead of the correctly alarmed crises. By keeping BPNN and SVM as baselines, three experiments are implemented, including the test-set forecasting and cross-validation, hoping to gain a comprehensive understanding of the system's capacity in terms of crisis forecasting power and stability.

Figure 5.9 shows the model predicted signals against their true crisis labels (1 for crisis and 0 for tranquility). According to the figure, crisis onsets on the test set mainly occurred in 2016 and 2018 due to the lasting effect of the 2015 stock market crash and the financial instability in China. Overall, the proposed EWS-LSTM depicts the test-set crises in a relatively precise manner, with the first alarms (red line) released adequately early before the actual onsets (blue dashed line). As BPNN replaces the predictive model, the EWS tends to delay producing the first crisis signal despite its ability to capture ongoing crises. In

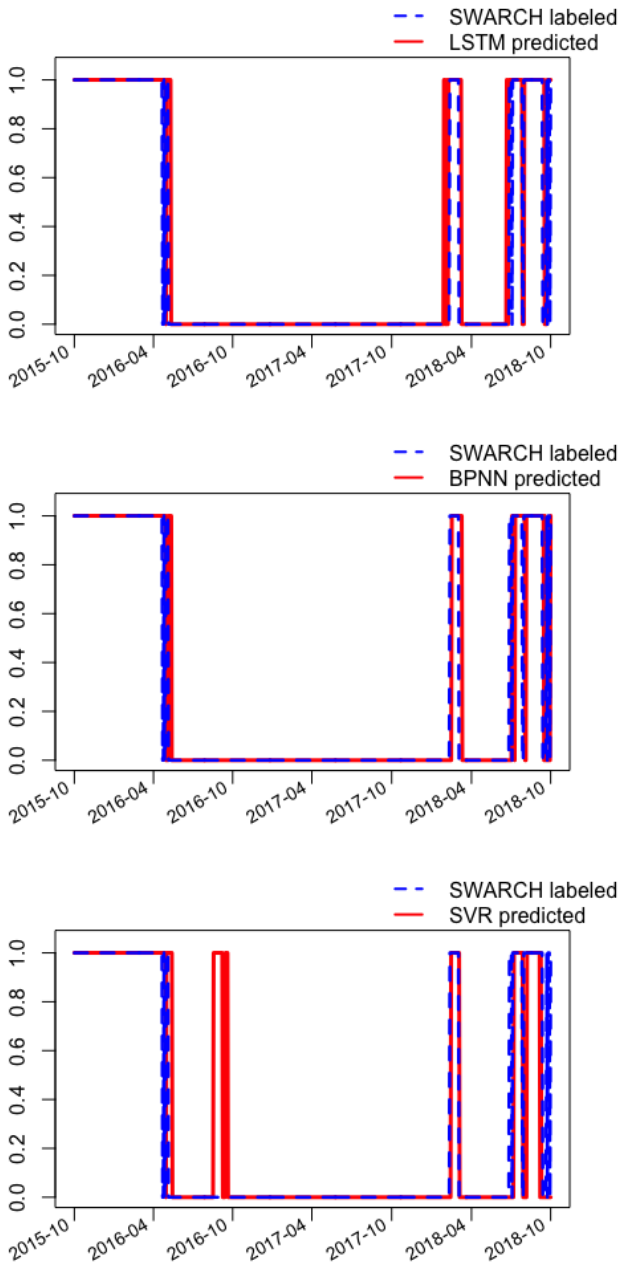


Figure 5.9: Test-set early warning signals

contrast to LSTM and BPNN, SVM appears to suffer from both delayed warnings and false alarms.

To support the preceding claims with evidence, Table 5.8 summarizes the numerical results related to the test-set forecasting. The test set consists of 724 days with 201 crisis days (Row 2, Table 5.8) and 5 crisis onsets (Row 6, Table 5.8).

Concerning Table 5.8, EWS-LSTM demonstrates a promising capability of warning stock turbulence, which dominating results have been fully reflected in

Table 5.8: Summary of test-set forecasting.

Model	EWS-LSTM	EWS-BPNN	EWS-SVM
Total crises	201	201	201
Correct predictions	**200**	179	186
% of correct predictions[a]	**99.5**	89.6	92.5
Total onsets	5	5	5
Predicted onsets	**5**	3	2
% of correct predicted onsets[b]	**100**	60.0	40.0
% of false onset alarms	**0.0**	0.0	20.0
Avg. days-ahead onsets	**2.8**	2.2	1.4

[a] % of correct predictions is the percentage of total crisis signals correctly predicted,
[b] % of correct predicted onsets is the percentage of correctly forewarned onsets.

the all-around examined. In particular, LSTM-based EWS improves the baselines with 200 days of correct predictions, which yield a rate of 99.5%. On average, the model alerts stock turbulence 2.8 days ahead of the actual crises and successfully warns 100% of the onsets with 0% false alarm. In line with the observations made from Figure 5.9, the major weakness of the BPNN-based EWS is revealed due to its delay in generating crisis signals. Though the hit-ratio for correct predictions is relatively low at 89.6%, its capability to sense the crisis onsets keeps a passable level of 60%. Besides the delays, the high percentage of 20% false alarms and the fewest 1.4 forewarned days make SVM the least reliable model for the early warning task compared to LSTM and BPNN.

Table 5.9: Average rand accuracy and binary cross-entropy loss in the k-fold cross validation.

Model	EWS-LSTM	EWS-BPNN	EWS-SVM
$k = 3$			
Accuracy (avg.)	0.954	0.887	0.919
Binary cross-entropy loss (avg.)	0.251	0.432	0.639
$k = 5$			
Accuracy (avg.)	**0.961**	0.902	0.918
Binary cross-entropy loss (avg.)	**0.144**	0.310	0.566
$k = 8$			
Accuracy (avg.)	0.920	0.901	0.889
Binary cross-entropy loss (avg.)	0.167	0.319	0.508

To analyze the stability of the EWS, a k-fold cross validation is further conducted in the test set with varying values $k = 3, 5, 8$.[10] Rand accuracy and cross-entropy loss are used as the performance measures. According to Table 5.9, the LSTM-based EWS is proven to be prominently robust in the cross validation, since LSTM invariably produces the greatest accuracy and lowest loss in comparison to the baselines given different k values. In the 5-fold validation, EWS-LSTM achieves the best test-set accuracy of 96.1% and loss of 14.4%. This predominant performance keeps a stable consistency on 3-fold validation as well. Even in the 8-fold data splitting test, the LSTM performs a considerable gap with BPNN and SVM in accuracy and loss.

5.2.7 Retrieving in practice

The constructed EWS is required to be retrieved in practice to investigate its efficiency to direct the real trading world. Back-testing and reality check are two standard examination ways to achieve such an aim.

Back-testing

In the study, a simple trading strategy is adopted to the SSEC stock index to verify the proposed EWS's effectiveness, given that the information between the market and the investors (who perform a fair level of risk aversion) is efficiently symmetric. A market portfolio of SSEC index is constructed and held until the EWS alerts crises and repurchased as the EWS suggests tranquility. We set the p as the constructed portfolio. $E[R_p]$ is the expected return rate by adopting the portfolio p, and σ_p is the corresponding standard deviation. SharpeRatio is thus formulated as Sharpe Ratio=$\frac{E[R_p]-R_f}{\sigma_p}$, where R_f denotes the risk free interest rate which is set to be zero[11] in our study. Here, we use the simplistic buy-and-hold strategy, which means we buy one portion of the SSEC index and hold it till EWS warning emerges (i.e. hold zero portion during warning periods), and then repurchase it as the warnings end.

[10] Given the selection of k deals with the trade-off between bias and variance, the cross validation is conducted up to 8 folds to ensure the size of the test set is large enough to offer statistically representative of the model's forecasting power.

[11] The zero risk-free rate is the theoretical one for the perfect market, while in practice, the zero risk is not true but adopting the Treasury bond minus the inflation rate.

Table 5.10 summarizes the expectation and standard deviation of returns together with Sharp ratios in the full sample and test set. In the absence of early warning mechanisms, the market portfolio without any EWS access yields expected returns of 2.5% and -3.9% and standard deviations of 1.495 and 1.197 in the full sample and the test set, respectively. The corresponding Sharp ratios are 0.017 and-0.033. By exiting the market position for early warned turbulence, the strategy significantly reduces the systematic risk (indicated by the σ_p), which naturally results in a higher level of Sharp ratio regardless of the predictive model.

Table 5.10: Back-testing in the full sample and test set.

	$E[R_p]$	σ_p	Sharpe Ratio
panel (a): full			
market portfolio	0.025	1.495	0.017
EWS-LSTM	0.033	0.699	**0.047**
EWS-BPNN	0.025	0.761	0.033
EWS-SVM	0.026	0.734	0.035
panel (b): test			
market portfolio	-0.039	1.197	-0.033
EWS-LSTM	0.091	0.619	**0.147**
EWS-BPNN	0.014	0.744	0.019
EWS-SVM	0.045	0.617	0.073

More importantly, back-testing once more verifies that the LSTM-based EWS outperforms the baselines and holds the greatest effectiveness and stability. Specifically, the effectiveness of LSTM is proven by its dominating Sharp ratios, which improve the market portfolio by 0.03 and 0.18 in the full sample and test set, respectively. Meanwhile, its stability is suggested by the monotonous positive impact on the market portfolio regarding the three portfolio measures in the risk-return horizon. Albeit the moderate improvements achieved by BPNN (Sharp ratios of 0.033 and 0.019 in the full sample and test set) and SVM (Sharp ratios of 0.035 and 0.073), the two models exhibit limitations due to their weaker and fluctuating results.

Reality check

To testify whether the proposed EWS model forewarned signals can be comparable in practical application to mitigate the investment risk brought by the drastic stock market turbulence, the Reality Check test based on White (2000) is thus applied to release the suspicion of data-snooping[12]. The test is pervasively recognized to be statistically rigorous and practically reliable and has been generally used in the market forecasting studies for both econometric models comparison (Kim and Swanson, 2014; Yang et al., 2010) and trade strategies verification (Arévalo et al., 2017).

To more comprehensively inspect the proposed EWS superiority to either the SWARCH crisis detection model or non-model adoption in terms of diminishing the impact from stock market turbulence, the reality check in our study will generally follow the procedure proposed by Arévalo et al. (2017) but re-define the performance series f. Specifically, to make the p-value available for both positive and negative returns, the holding cash strategy will not be applied as benchmarks, but the portfolio based on SWAWRCH(2,1) with arbitrary cutoff 0.5 will be exercised without being instructed by any predictive models output.

The realized variance will be hired to calculate the performance series as to notify the EWS impact on diminishing the market turbulence. Hence we have

$$f_{t+1}^{B_1} = RV_{EWS,t+1} - RV_{SWARCH,t+1}, \tag{5.7}$$

$$f_{t+1}^{B_2} = RV_{EWS,t+1} - RV_{t+1}, \tag{5.8}$$

where RV_{t+1} is the realized variance for stock price returns. B_1 denotes the benchmark model of SWARCH(2,1) with 0.5 cutoff and B_2 denotes the market portfolio returns without any predictive model participation. The bootstrapping will repeated 10000 times for calculating the mean of f given the null hypothesis of $H_0 : E(f) \geq 0$, being explained as the proposed EWS cannot outperform the benchmark in diminishing the intensity of market turbulence. The specific implementing process can be referred to Appendix D.1.

Table 5.11 reports the p-values of rejecting the null based on reality check

[12]In other words, to inspect whether the superior model predicting performance is produced by chance not a constant result.

Table 5.11: Bootstrapped reality check p-value for model comparisons.

	EWS-LSTM	EWS-BPNN	EWS-SVM
full sample			
p-value$_{B1}$	**0.001*****	0.012**	0.004***
p-value$_{B2}$	**0.000*****	0.000***	0.001***
test set			
p-value$_{B1}$	**0.023****	0.069	0.034**
p-value$_{B2}$	**0.010****	0.011**	0.012**

***, ** denote the null hypothesis is rejected at 1% and 5% significance level, respectively.

as comparing two benchmarks of B_1 and B_2. The three hybrid EWS models based on LSTM, BPNN and SVM are separately investigated in the full and test sets. In the comparison between the SWARCH and EWS models, the null for B_1, i.e. the SWARCH outperforms EWS, is rejected with the significance level of 1% on the full sample and 5% on the test set. As for the free market portfolio without considering the alarming scheme, the reality check suggests that all three EWS models are superior to the benchmarks with remarkable significance levels. The EWS-LSTM remains most favorable for forecasting crises with the lowest p-values compared to the other two EWS models. Thus, the proposed EWS frameworks are believed robust in reducing the portfolio risks, especially during stock turmoils, to help take early actions in investment.

5.2.8 Concluding remarks

In this chapter, an effective EWS with a dynamic architecture that integrates the SWARCH model, two-peak thresholding and LSTM is developed to identify and predict China's stock market turbulence. According to the models' performance on the ten-year sample of Shanghai Stock Exchange Composite index, the following concluding remarks emerged.

1. As one of the most powerful models handling sequential data, LSTM remains in its outstanding position in the daily prediction task of stock crises. To be specific, the reliability of LSTM in this study is not only reflected by the high accuracy of 99.5% and on average 2.8 days of forewarned pe-

riod, but also its stability of outperforming the baselines throughout the evaluation process in the test-set, cross-validation as well as back-testing.

2. In addition to a high-performing predictive model, a precise and robust crisis identification mechanism also plays a central role in facilitating the effectiveness and reliability of an EWS. By adopting the two-peak method to determine crisis cutoffs, the proposed EWS suggests a constructive alternative to current existing approaches and yields promising crisis classifications in China's stock market compared to the classic indicator function based on CMAX.

3. Stock market turbulence described by the SWARCH volatility regimes is proven to be a good crisis indicator in both theory and practice, as the proposed EWS depicts all the recorded major stock crises in the sample with significantly improved back-testing results than the market portfolio.

In a nutshell, the proposed integrated EWS is recommended for two means of application. First, the EWS crisis classifier could be directly used for post-date crisis identification and characterization based on historical data without prior assumptions on crisis dates or thresholds. Second, with adequate lengths of forewarned period and high-degree accuracy, the proposed EWS allows market participants at all levels to make early decisions to react towards the potential crisis. Last, the proposed EWS can be used for the more general purpose of warning financial turbulence/turmoils for other specific markets on a daily basis.

In the following section, the integrated EWS will be applied to China's bond market by further discussing the role of improved technique − RCM for SWARCH based crisis classifier and the stacked layer − attention mechanism for LSTM based crisis predictor to extend the market-oriented EWS development in terms of gaining the crisis identification precision and the machine learning model's interpretability.

5.3 EWS for China's bond market

This section will develop the EWS for China's bond market by posting technical variations on both the crisis classifier and crisis predictor. A series of representative bond indexes published by China Central Depository & Clearing CO., LTD will be studied to accomplish such variations on the EWS development: a) other than adopting the two-peak thresholding technique, SWARCH frameworks will be introduced the RCM selection scheme to directly refine the bond index volatility levels in more than two regime states to identify the crisis samples[13]; b) the predictive model will be substituted in the attention based Bidirectional LSTM (attn-BiLSTM) to investigate the improved model performance on prediction and feature selection to make further comparative to other feature-selection-allowed predictive models.

5.3.1 Crisis identification: RCM based crisis classifier

As mentioned in Chapter 3, there are two essential inferences on observation's volatility states in SWARCH frameworks, namely filtering and smooth probabilities. The former probability has been used in the CM development to detect the turmoil episodes for univariate market assets and the EWS construction for China's stock market to identify crisis samples. In this project, the targeted crisis variable will be defined on smooth probabilities, which is believed to carry the information from the full length of observations to make the inferred probability less noisily troubled by the future uncertainty.

The expression of smooth probability for state i is written as follows,

$$P(s_t = i|\mathscr{Y}_T; \hat{\boldsymbol{\theta}}) \quad \text{for} \quad i \in \{1, ..., K\}, \tag{5.9}$$

where $\mathscr{Y}_T$ denotes the full observations and $\hat{\boldsymbol{\theta}}$ is the vector of estimated parameters. Kim (1994) and Kuan (2002) have provided the full deriving process for Eq.5.9. The R package 'MSGARCH' provides the programming code to infer the

[13]In practice, two-peak dynamical thresholding technique cannot be simultaneously used with the RCM determination technique. For example, as the number of determined regimes is greater than two, crisis samples are hard to be identified since the SWARCH estimated probabilities in the highest volatility state are too placid to form distinctive twin peaks in histogram.

smooth probability of SWARCH model.

According to previous experience, for most asset prices and indexes in financial markets, signals of crashes or crises intuitively result in volatility jumps. The binary crisis variable between 0 and 1 (1 means the crisis produces) is thus defined on the volatility level distinguishing between the 'tranquil' and 'turbulent' states in the SWARCH frameworks. The number of regimes for volatility levels in the SWARCH model is thus generally taken high and low two cases, but it is not appropriate to fit all market asset price dynamics, especially for the bond. Unlike other securities whose dynamics can be intuitively classified into two volatility regimes, the bond does not perform enough dispersion range to be distinctively clustered into significant-high for 'turmoils' and low for 'quiescence'. Thus, the appropriate selection for the number of states in the SWARCH model will be required to effectively characterize different volatility levels to avoid underestimating or overestimating the volatility intensity and further affecting the crisis definition precision. In other words, the price dynamics which have K latent volatility states will be mis-specified by SWARCH with $K+1$ (overestimated) or $K-1$ (underestimated) regimes if the regime number K is imprecisely determined.

Ang and Bekaert (2002) propose the regime classification measure (RCM) to diminish the mis-specification chance in characterizing volatility levels in the SWARCH model. Eq. 5.10 formulates these numerical metrics as follows.

$$RCM(K) = 100K^2 \frac{1}{T} \sum_{t=1}^{T} (\prod_{i=1}^{K} p_{i,t}), \tag{5.10}$$

where $p_{i,t} = P(s_t = i|\mathscr{Y}_T; \boldsymbol{\theta})$, the smooth probability. The value range of RCM is between 0 and 100, where 0 means the regimes are perfectly classified. Thus, the RCM deviates more from 0, the worse the classification produces[14].

The RCM-SWARCH determined crisis variable is defined as following crisis

[14]In practice, to avoid over-fitting, the number of states will be determined via balancing between the RCM value and the plot for smoothed probability.

binary function,

$$\text{Crisis} = \begin{cases} 1, & P(s_t = K_{rcm} | \mathscr{Y}_t; \boldsymbol{\theta}) \geq c, \\ 0, & \text{otherwise}, \end{cases} \tag{5.11}$$

where K_{rcm} is the RCM selected optimal value for regimes' number, which is usually the value for highest volatility level in SWARCH frameworks. c is the threshold for producing appropriate number of warning signals[15]

5.3.2 Forecasting system: Attention-BiLSTM based crisis predictor

The hybrid deep neural networks, attention mechanism based Bidirectional LSTM (Attention-BiLSTM) model, will be hired to predict warning signals for the bond.

First of all, comparing to the single-layer based LSTM deep neural networks, the bidirectional architecture makes a breakthrough in boosting the deep neural networks ability to fully utilizing both historical and future information via feeding dual direction of both backward and forward in the streamline of compressing data from two flow directions into the same network layer. It elegantly solves the difficulty of mixing two independent neural network layers. It processes all possible available time-dependent sequences and merges their results into one output layer. The bidirectional LSTM has been effectively applied and benchmarked with simpler structured LSTM in the study of evaluating and comparing deep learning models in the stock market prediction study (Althelaya et al., 2018).

Furthermore, to identify the factors contributing degree to the prediction and lessen the annoying noise from useless factors (Liu et al., 2017), the attention layer is essentially included to gain the comparability of our EWS model to regression and indicator-based predictive models. The attention will first assign initial weights on each factor, and then adaptive optimize the value of weights based on previous cells output in the training process to ascertain the optimal combination for the final predicted result. In other words, the attention mecha-

[15]The warning signals production is sensitive to the selection of classifier threshold. If the value is too low, warnings will cover the whole time period; and vice versa. Here, we adopt grid search to determine the threshold value.

nism allows to distinguish and leverage the convincing degree of each input factor by dynamically updating the model learned results in the training process.

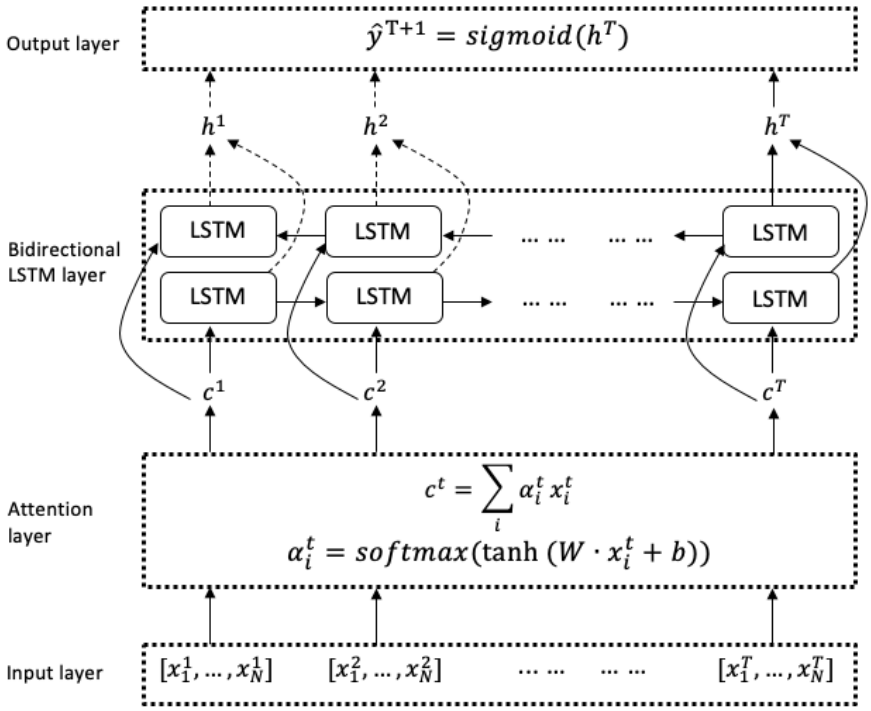


Figure 5.10: Networks frame structure of attention-BiLSTM.

The full structure for attention-BiLSTM networks includes four layers from bottom to top, and as Figure 5.10 shows, they are Input layer, Attention layer, Bidirectional LSTM layer and Output layer. The arrow directs information flow from the bottom input layer to the top layer. Each layer is framed in dotted lines. The dashed arrow means no output will be produced from cells, and N and T represent the size of features and time steps. Starting from the input layer, each piece of sample vectors $\{x_1^t, ..., x_N^t\}$ for $t = 1, ..., T$ will enter into the attention and bidirectional LSTM layers to generate the final prediction of $\hat{y}$ for time point $T+1$. Then, the next sample piece $\{x_1^t, ..., x_N^t\}$ as $t = 2..., T+1$ will slide forward for producing the prediction of $\hat{y}^{T+2}$ till the end of observations[16]. The samples will be trained for 64 batches each time, and the training session for full samples will be repeated in 200 epochs to make the modeling results stable.

The bidirectional LSTM networks include an extra layer than the unidirectional LSTM networks, which allows the connection flows backward to exploit

[16]This is the sliding window proceeding for time sequence, shuffling is not allowed.

information from the past, thus in the model training session, it will bring in higher accuracy than single-layered LSTM for predicted results[17]. As the diagram (Figure 5.10) shows, the output from the attention layer will enter into two LSTM cells simultaneously and then generate h^t that combines two outputs from each of the cells. Specifically, in the forward information flow (i.e. the lower LSTM cells layer), the input sequence will be read as the time sequence from 1 to T and thus the hidden state output $\overrightarrow{h_t}$ will be generated in this sequence. In contrast, in the backward flow (i.e. the upper LSTM cells layer), the processing order will is reversed, and the hidden output will denote as $\overleftarrow{h_t}$. Thus the output h^t for the Bidirectional LSTM layer is the concatenate result from both forward and backward outcomes, i.e. $h^t = [\overrightarrow{h_t}, \overleftarrow{h_t}]$.

In our study, the soft (global) attention mechanism (Bahdanau et al., 2014) is adopted to the Attention layer, where the weight α_i^t that defines how much of each source should be counted for each output will be aligned to the entire input space. The attention layer is arranged ahead of all other neural layers to detect which features will be distributed more attention weights in the predicting procedure. To implement the soft attention mechanism, there are three separate steps: scoring, aligning weights and summing up, and functions can be written as follows

$$s_i^t = tanh(W \cdot x_i^t + b) \quad \text{(score)}, \tag{5.12}$$

$$\alpha_i^t = \frac{exp(s_i^t)}{\sum_{i'=1}^{N} exp(s_{i'}^t)} \quad \text{(weight)}, \tag{5.13}$$

$$c^t = \sum_{i=1}^{N} \alpha_i^t x_i^t \quad \text{(summation)}, \tag{5.14}$$

where $tanh$ is the non-linear activation function to transfer value into range (-1,1). W and b are weight and bias parameters for each given attention, and they both are initialized from uniform distribution. In the process, input features will be first activated by score function and then weighted by the weight function. The final output of Attention layer is c^t, the summation of activated and weighted inputs. It will be further forwarded to the Bidirectional LSTM layer.

[17]The BiLSTM, however, consumes more time than LSTM for training model considering the number of neurons doubles.

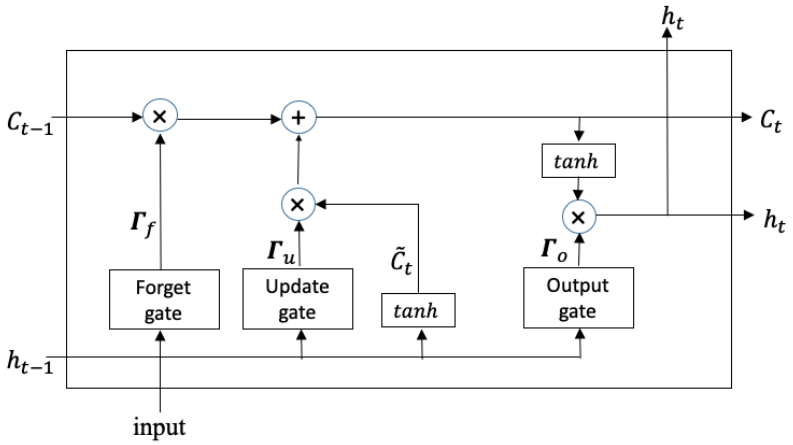


Figure 5.11: LSTM cell in bond EWS.

The LSTM cell inner calculation procedure will follow the order of equations of Eq. 5.15−5.20. The *input* here in our case is c^t, the output from Attention layer. Γ_f, Γ_u and Γ_o are sigmoid functions of the forget, update, and output gate that determine the information to be discarded added and reproduced. The updated output will be produced by the peehole function $\tilde{C}_t$. W denotes the weight parameter, and b is the bias term. All parameters are initialized at the start and will be trained in the learning process.

$$\Gamma_f = sigmoid(W_f \cdot [h_{t-1}, input] + b_f) \tag{5.15}$$

$$\Gamma_u = sigmoid(W_u \cdot [h_{t-1}, input] + b_u) \tag{5.16}$$

$$\tilde{C}_t = tanh(W_C \cdot [h_{t-1}, input] + b_C) \tag{5.17}$$

$$C_t = \Gamma_f * C_{t-1} + \Gamma_u * \tilde{C}_t \tag{5.18}$$

$$\Gamma_o = sigmoid(W_o \cdot [h_{t-1}, input] + b_o) \tag{5.19}$$

$$h_t = \Gamma_o * tanh(C_t) \tag{5.20}$$

5.3.3 Data

The data for studying China's bond market will be parted into two components: for time t, the targeted crisis variable y_t for typical bond indices and both of endogenous and exogenous feature variable vector of $\boldsymbol{x}_t$ relating to bond crisis predictions. The time span covers from 2007/2/28 to 2019/4/30, which is believed to be adequately long to include abundant turmoil information for generating

bond crisis signals, such as the 2008-2009 Global financial crisis, 2010 Sovereign Debt crisis, and the 2016 economic downturn in China. There are thus 2964 observations for each variable[18], and 2936 samples[19] by taking 28[20] time steps and 23 input factors. Full samples will be split into 65% for the train set and 35% for the test set[21], thus the count for each set are 1908 and 1028 respectively[22].

Target bond indexes

According to the type of issuers, the bond in China are mainly categorized into eight genres: government bonds, central bank bills, government-backed institutional bonds, financial bonds, corporate credit bonds, asset-backed securities, panda bonds and inter-bank certificates of deposit[23]. Our research objects to the bonds that are most traded in the secondary market, i.e. government bond and corporate bond, considering their price indexes are not costly accessed from the bond custodian agency, and they both share high liquidity and sufficient trading volume, which makes them entrusted more guiding significance in practice.

The China Central Depository & Clearing CO., LTD. (CCDC) published a series of bond indexes that can objectively project China's bond market price behavior since 2002. The indexes are the weighted average market value price for counting each bond balance in various screening conditions. The calculation is formulated as follows.

$$I_T = I_{T-1} \times \sum_i \big(\frac{P_{i,T} + Pri_{i,T}}{P_{i,T-1}} \times W_{i,T-1}\big), \tag{5.21}$$

where I_T is the price index value on T day, $P_{i,T}$ is the price bond i on T day, $Pri_{i,T}$ is the principal payment of bond i on T day, $W_{i,T-1}$ is the price market cap weight of bond i on $T-1$ day. The composed index includes the bonds

[18]The low-frequency data will be interpolated to daily frequency.

[19]Number of samples = Count of total observations (N) - Time step (T)

[20]28 days generally cover the observations in one month and will be used for predicting the crisis in the future one-day.

[21]The integer will be taken for each percentage calculation.

[22]The start date for the test set is 2015/02/09

[23]In addition to government bonds and corporate credit bonds, a limited amount of financial bonds and asset-backed securities being traded in the secondary market, the remaining bonds are only permitted to be traded in the inter-bank market, foreign exchange trading centers (such as inter-bank certificates of deposit) and overseas markets (such as panda bonds).

with one-year and above maturities. The basis point of 100 is the price data on December 31, 2001.

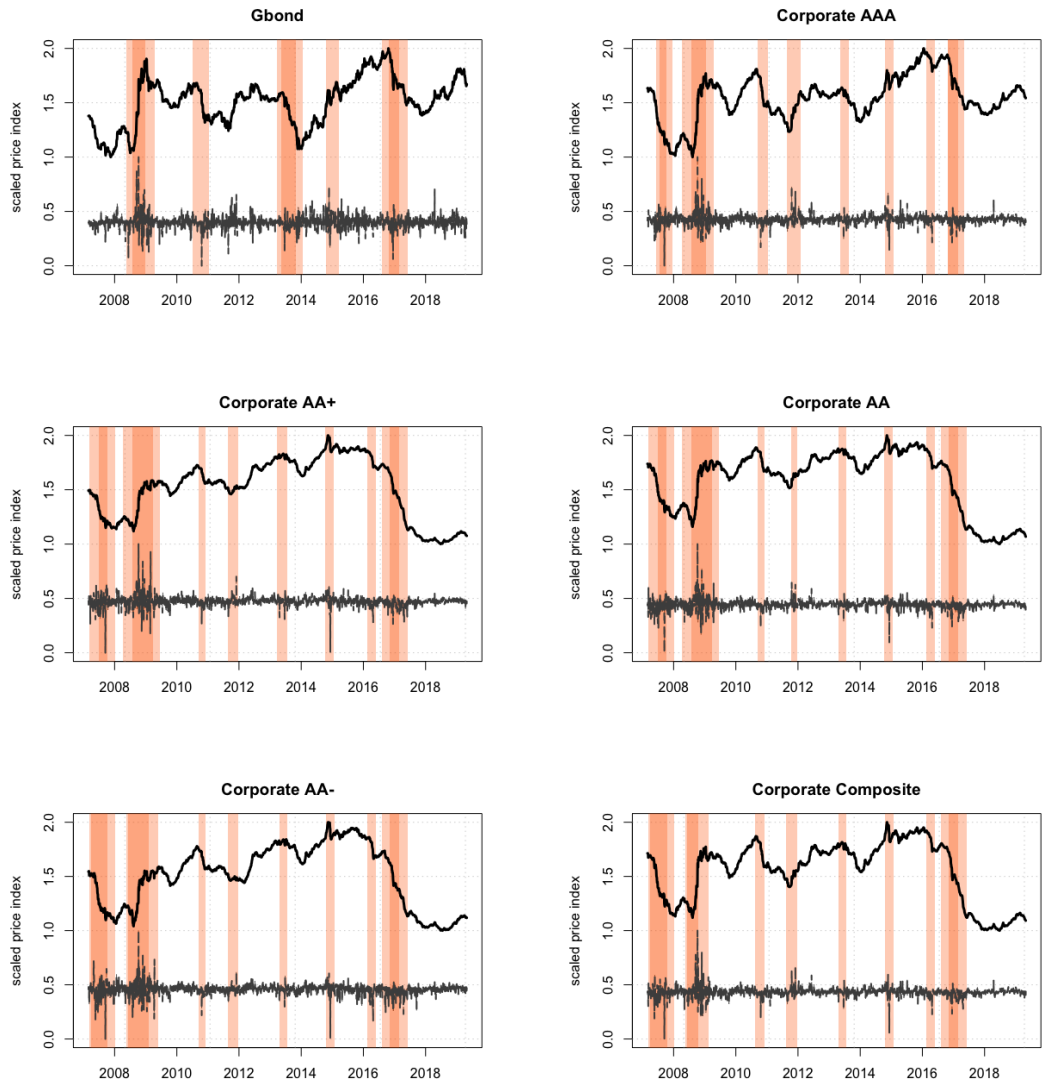


Figure 5.12: The raw plot for the bond index and the corresponding log returns.

In each subfigure, the upper black curves show the index dynamics and the lines below show the log returns for each index. Both lines are scaled in the range of $(0, 1)$.

In our study, six typical government bond indexes, the corporate composite bond, the corporate bonds with AAA rating, AA+ rating, AA rating and AA- rating levels, are sourced to cover different credit risks. They will be abbreviated as "Gbond", "Corp Comp", "Corp AAA", "Corp AA+", "Corp AA" and "Corp AA-" in following sections. Except for the substantial trading volume, we select such six bond representatives as considering 1) these index dynamics carry ma-

jor risks (from the term of structure, liquidity and credit) that affect the bond yields; and 2) the corporate bonds with different credit rating levels contribute to examine the forecasting system's functional universality as the bond underlay credit risks being diversified.

Figure 5.12 displays raw plots for scaled price indexes and corresponding log return dynamics. The orange shadow segments highlight the most fluctuated periods of the log return dynamics at the moments of price indexes abruptly dropping or rising, and the fluctuating greater, coloring darker. The raw plots empirically characterize the bonds with different level risks in turmoil patterns of that 1) the government bond covers minor vibrations in the 2008-2009 global financial crisis as compared with other corporate bonds even at the right start of 2008, 2) the corporate bonds with lower credit ratings are more intuitively fragile than Gbond and Corporate AAA bond in crises as they share longer and more frequent turmoil slots, and 3) at the end of the year 2016, China suffered from a temporary domestic economic downturn, which "dip" brought a more lasting effect for lower credit rating corporate bonds than government bond and high credit rating corporate bonds, till the sign of weak re-coverage emerged in the middle of 2018.

Exogenous features

The exogenous factors are mainly from following seven risk sources to affect the turmoil in bond market: A) the China' stock market that brims with investors sentiment and acts as the totemic investment for the financial market is accessed by introducing Shanghai Composite index and Shenzhen Component index; B) the interest rate spread of SHIBOR/Interbank lending rate/Pledge-style Repo between different short terms in the monetary market reflect the financial crisis awareness; C) macroeconomic indicators, such as CPI and PPI that reflect inflation rate in consumption and production, M2 that indicates the degree of economic prosperity, industrial value added rate and fixed-asset investment gain rate that mirror the investing levels in economic activities, will sculpt the portrait of overall national economic development; D) the stability and flexibility of Chinese yuan in the foreign exchange market will be cast in exports/imports ratios,

foreign exchange loan balance and foreign exchange reserves; E) The discount rate for China, gross currency liability and GDP index published by the international monetary fund (IMF) show the Chinese economic development level in international evaluation standards. Table 5.12 describes the statistical summary for the raw data of all exogenous features.

Table 5.12: Summary statistcs for exogenous features.

variable	Mean	S.D.	Min.	Max.	skew.	kurtosis
Sha_Comp	2926.34	754.87	1706.70	6092.06	1.35	2.38
Shen_Comp	1339.17	463.52	404.71	3140.66	0.68	-0.06
Shibor7d_1d	0.54	0.49	-2.44	4.43	2.28	10.65
Shibor1m_7d	0.67	0.57	-1.60	4.10	0.68	1.61
Tongye	0.71	0.50	-1.58	4.61	2.28	9.98
Zhiya	0.62	0.49	-1.83	4.42	2.36	10.38
CPI	102.74	2.05	98.20	108.70	0.65	0.73
PPI	101.14	4.68	91.80	110.06	-0.06	-1.18
M2	14.78	5.09	8.00	29.77	1.09	1.06
INDUSTRIAL_gain	10.16	4.87	-2.93	29.20	0.56	0.66
Fixed_asset	18.57	8.31	5.30	33.60	-0.16	-1.32
Imp_Exp	9.03	16.51	-29.08	48.37	0.03	-0.46
FOREX_Loan	15.44	19.12	-16.80	73.77	0.88	0.64
FOREX_Reserves	29131.42	7355.51	11573.72	39932.13	-0.74	-0.41
Currency_liability	285478.52	64534.67	135591.47	372492.06	-0.68	-0.73
GDP_index	100.10	0.79	97.84	101.81	-0.33	0.79
DISCOUNT_rate	4.10	1.51	1.00	7.00	0.10	-1.06

Sha_Comp and Shen_Comp are stock market indexes of Shanghai Composite and Shenzhen Component. Shibor7d_1d and Shibor1m_7d are the interest spread of SHIBOR between one week and overnight, and between one omnth and one week, Tongye and Zhiya represent the interbank lending rate and the pledge-style repo rate, respectively. INDUSTRIAL_gain denotes the industrial value added rate. Fixed_asset is the fixed asset gain rate on the year-on-year basis. Imp_Exp is the exports/imports ratio. FOREX_Loan and FOREX_Reserves denote the foreign currency loans and reserves respectively. Currency_liability is the IMF gross currency liability for China. GDP_index is provided by the IMF to reflect China's total GDP level. DISCOUNT_rate is the discount rate published by the People's Bank of China. The economic factors relating to currency, such as FOREX_Loan, FOREX_Reserves, Currency_liabilitym use millions of Chinese yuan as the unit.

Table 5.13 summarizes all input features for China's bond EWS. The conditional volatility, filtered and smoothed probabilities are the inferred results from SWARCH model for each index. All input data will be normalized in the range of $(0, 1)$ to eliminate the order of magnitude difference between data of different dimensions and to avoid the large network prediction error caused by the large

Table 5.13: Input features.

Variable	Frequency	Dimensionality	Source
Index price (including first and second difference)	Daily	Endogenous factors	CCDC
conditional volatility, filtered and smoothed probabilities			SWARCH
Shanghai Composite Index, Shenzhen Component Index	Daily	Stock market	WIND
Interest rate spread of a) SHIBOR between 1 week and overnight, b) SHIBOR between 1 month and 1 week, c) Interbank lending rate between 7 days and overnight, d) Pledge-style Repo between 7 days and overnight	Daily	Monetary market	
CPI, PPI, M2, Industrial value added rate (year-on-year basis), Fixed-asset investment gain rate (cumulative year-on-year basis)	Monthly	Domestic economy	
Exports/imports ratio (year-on-year basis), Foreign exchange loan balance, Foreign exchange reserves	Monthly	Foreign market	
IMF discount rate for China (annually), IMF gross currency liability for China, GDP index	Monthly	External evaluation	IMF

order difference between inputs and outputs. Except for the IMF gross currency liability for China and discount rate for China, data are all accessed from WIND database.

Figure 5.13 shows the empirical correlogram across exogenous factors, where

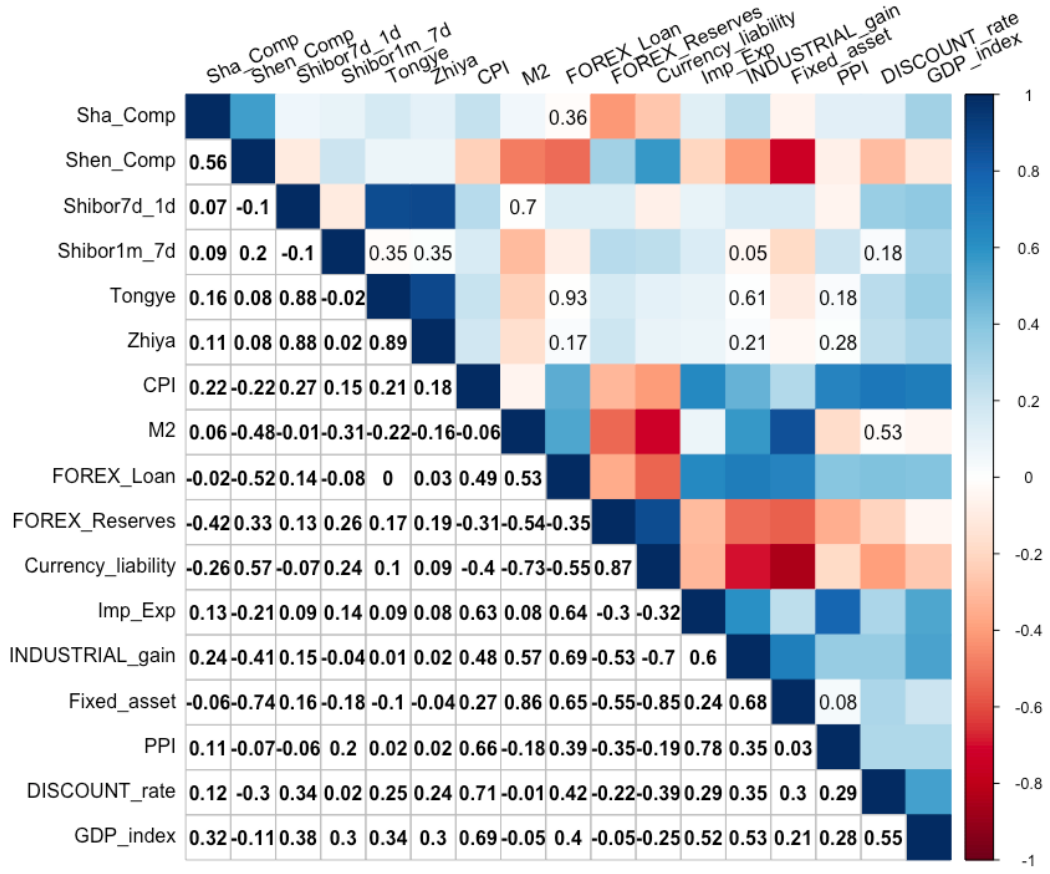


Figure 5.13: The correlogram for exogenous feature variables.

the lower triangle is the correlation coefficient matrix, and the upper mosaic-like triangle colorizes the value of the coefficient matrix with (negative correlated) red and (positive correlated) blue gradient ramp. The p-value for the independence test at 5% significance level is dropped in the (white) colored block. From the correlogram, the orthogonality across input exogenous factors can be maximized since few are tightly correlated with each other[24].

5.3.4 RCM-SWARCH identified crises

Following the construction process for the targeted variable in Section 5.3.1, we first determine the value of K in the SWARCH model to classify the volatility states. From Table 5.14, the value of RCM will not significantly decrease as K

[24]1) between the short-term Shibor spread (Shibor7d_1d) and the repo pledge rates (Tongye and Zhiya), 2) between the money supply (M2) and the fixed asset gain (Fixed_asset), 3) between the Shenzhen component stocks index (Shen_Comp) and the fixed asset gain (Fixed_asset) and 4) between the gross currency liabilities (Currency_liability) and the fixed asset gain (Fixed_asset)

increases to 4 for all bond indexes. Thus the SWARCH model will take $K = 3$ for the targeted crisis variable[25].

Table 5.14: RCM value for varied K.

Index	$K = 2$	$K = 3$	$K = 4$
Gbond	37.239	0.693	0.028
Corp AAA	26.173	2.201	0.326
Corp AA+	30.118	0.003	0.000***
Corp AA	24.709	0.437	0.354
Corp AA-	24.712	1.365	0.046
Corp Comp	17.705	0.078	0.005

*** denotes the significant level at 0.01.

Table 5.15: Descriptive statistics of bond price indexes in full, turmoil and tranquil periods.

bond index	obs.	Mean	S.D.	Min.	Max.	skew.
panel (a): full period						
Gbond	2964	116.73	3.71	108.14	125.08	-0.27
Corp AAA	2964	98.64	3.57	89.23	106.46	-0.36
Corp AA+	2964	100.20	7.51	86.08	113.43	-0.45
Corp AA	2964	96.82	6.23	83.87	106.02	-0.82
Corp AA-	2964	98.60	6.83	86.31	110.37	-0.35
Corp Comp	2964	95.76	5.27	85.40	104.33	-0.61
panel (b): turmoil period						
Gbond	210	117.41	3.77	109.41	123.45	-0.69
Corp AAA	426	97.57	4.11	89.23	104.10	-0.53
Corp AA+	523	98.49	6.44	86.08	113.43	-0.63
Corp AA	675	95.99	4.78	87.40	106.02	0.96
Corp AA-	490	97.31	6.65	87.29	110.37	0.53
Corp Comp	447	95.48	4.42	85.40	104.33	-0.65
panel (c): tranquil period						
Gbond	2754	116.73	3.70	108.14	125.08	-0.24
Corp AAA	2538	98.82	3.44	89.45	106.46	-0.24
Corp AA+	2441	100.56	7.68	89.33	111.22	-0.28
Corp AA	2289	97.06	6.57	83.87	104.60	-0.06
Corp AA-	2474	98.85	6.83	86.31	109.10	-0.31
Corp Comp	2517	95.82	5.24	87.66	103.41	-0.06

[25]The best value is 3 not 4 since when $K = 4$, the smooth probability goes to deteriorate in distinguishing the hidden turmoils from volatility fluctuations, which thus will produce fewer signals under this niggling classification.

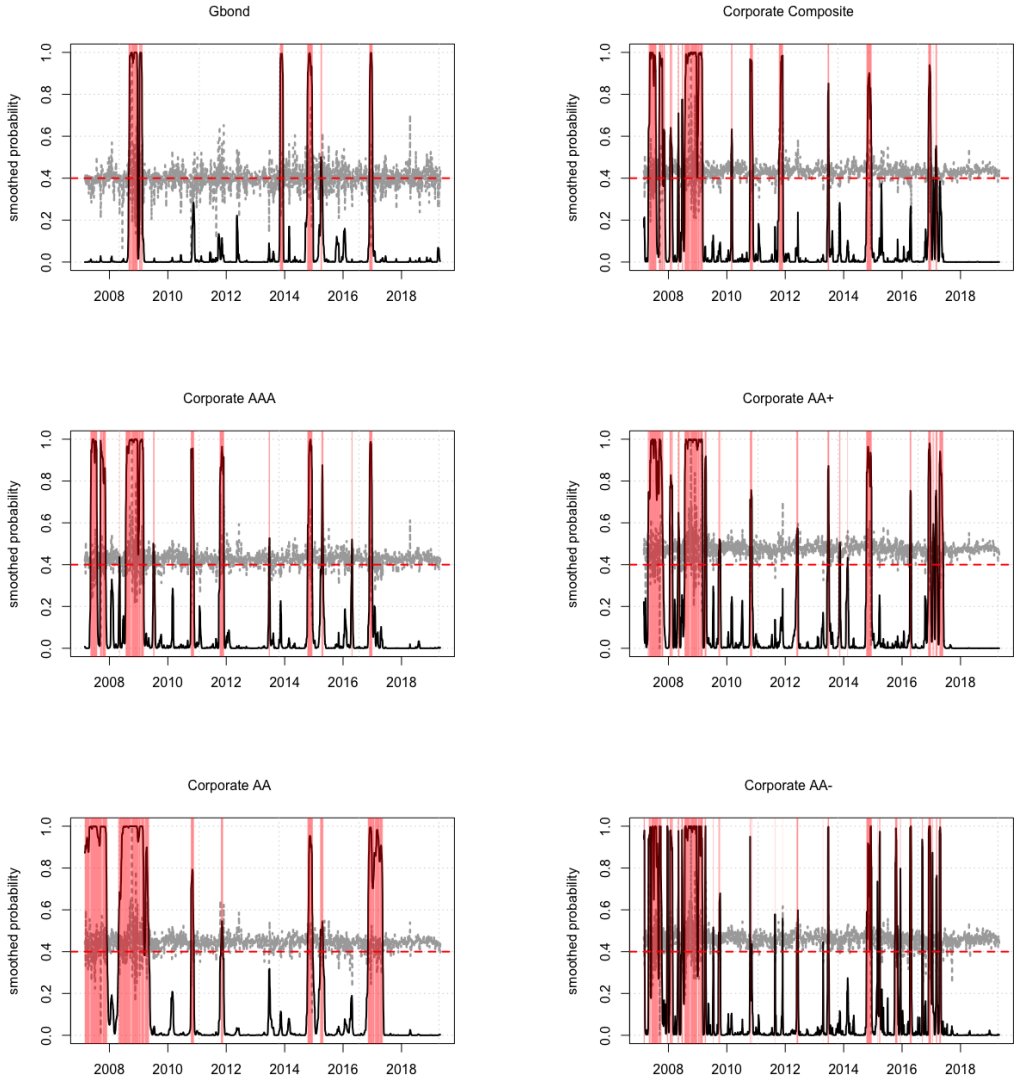


Figure 5.14: RCM-SWARCH inferred crisis periods.

The crisis samples will be produced by Eq 5.11 with grid searched threshold $c = 0.4$. Figure 5.14 highlights the RCM-SWARCH model identified turmoil periods in red shades. The gray dashed lines are the scaled log returns for each index in the range of $(0, 1)$, and the black lines are the smooth probability for $p_{s_t=3}$. The red dashed horizontal line labels the threshold 0.4. Compared with the scaled log returns, the SWARCH model with three regimes can precisely identify the turmoil periods and distinguish each bond's idiosyncratic turbulence pattern. The precision of detecting crises can be generally embodied by locating the historical financial crisis events, such as the global financial crisis of 2008 (covers from the end of 2007 to 2009), the European debt crisis (covers from

the end of 2009 to 2010), the China's stock market crash of 2015 (starts from the middle of 2015 and lasts two more months) and the China's bond liquidity crisis (at the end of 2016), that the bond index dynamics stumblingly behave. In addition, the bond turbulence degree seems to be related to the credit level. Specifically, the government bonds with higher credibility show less detected crisis turbulence, while the corporate bonds with lower credibility show more.

Table 5.15 summarizes the statistics for bond index observations in full length of observed periods (panel (a)), classified turmoil periods (panel (b)) and the rest of tranquil periods (panel (c)). As compared to the values of skewness in the last column across three panels, the classified turmoil observations for all bonds perform severer unsymmetrical characteristics than both of the full and the tranquil, which verifies the effectiveness of RCM-SWARCH in distinguishing crisis samples.

5.3.5 Forecasting performance evaluation

The model predicting performance will be testified in the following two aspects: the count of pre-called signaling days before the true crisis onsets appear and the value of rand accuracy and binary cross-entropy loss in sliding forward cross validation. To make the proposed EWS more comparative in applications, three clusters of baseline models of 1) the attention combined single-layered LSTM, 2) the deep neural networks without attention mechanism (i.e. RNN, LSTM and BiLSTM), and 3) the ordinary machine learning models and classic regression (i.e. BPNN, SVR and binary logit regression) will be implemented in terms of accuracy, loss and ROC statistical metrics.

Table 5.16 summarizes the count of predicted signals that effectively forewarn the true crisis onsets in advance for each bond index. There proves that the EWS based on both RCM-SWARCH crisis classifier and attention-BiLSTM predictor does not produce any delayed signals for both in-sample and out-of-sample sets with 100% hit ratio for the intraday prediction. For 1-day ahead prediction, the train and test sets' hit ratios decrease to 90.5% and 84.6% respectively, remain effective at a credible level above 80%. As for predicting an earlier time horizon (including more than 2 days ahead), the hit ratios gradually deteriorate from

Table 5.16: Effective count of predicted days in advance of the true crisis onsets for in-sample and out-of-sample sets.

		Days in advance								
		0	1	2	3	4	5	6	7	8
	count of onsets	count of predicted onsets								
panel (a): in-sample										
Gbond	4	4	4	4	2	1	1	1	1	1
Corp AAA	9	9	9	8	8	7	7	7	7	7
Corp AA+	12	12	10	5	3	3	2	1	0	0
Corp AA	5	5	4	4	3	2	2	2	2	2
Corp AA-	21	21	21	21	21	21	20	20	19	19
Corp Comp	12	12	9	4	2	0	1	2	1	1
total count	63	57	46	39	34	34	33	33	30	30
hit ratio(%)	–	100	90.5	73.0	61.9	54.0	52.4	52.4	47.6	47.6
panel (b): out-of-sample										
Gbond	2	2	2	1	0	0	0	0	0	0
Corp AAA	3	3	3	3	3	3	3	3	3	2
Corp AA+	6	6	5	2	3	2	2	2	1	1
Corp AA	2	2	2	2	2	2	2	2	2	1
Corp AA-	10	10	8	5	4	4	4	3	2	2
Corp Comp	3	3	2	0	0	0	0	0	0	0
total count	26	26	22	13	12	11	11	10	8	6
hit ratio(%)	–	100	84.6	50	46.2	42.3	42.3	38.5	30.8	23.1

The left column 'count of onsets' summarizes the count of SWARCH turmoil periods for each bond index. The columns of count of predicted onsets summarize the early-warning system successfully captured ones before the real turmoils begin in various 'Days in advance' situations. For example, the model succeeds to predict 2 out of 10 entering turmoils signals for Corp AA- in the situation of 9 days in advance. The bottom panel rows of total count and hit ratios refer to summing up the numbers true crisis periods for each index and calculating the percentage of predicted outputs in the total count of true crisis periods.

90.5% to 47.6% for the in-sample set and from 84.6% to 23.1% for the test set. It thus infers that the designed EWS effectively forewarn in the short term but becomes reckless to lose the credibility for long-term prediction. The table results also show the dependence of the model performance on the complexity of the bond index volatility pattern. Specifically, the government bond and corporate bond with AA credit level that indexes have aggregated turmoil observations (in other words, the RCM-SWARCH detected crisis samples are consecutively distributed) perform a slower deteriorating rate of predicting effectiveness than corporate bonds with other credit ratings that have scattered and swiftly altered

turmoil observations.

Table 5.17: Sliding forward cross validation for the Attention-BiLSTM model

count of folds	k = 3		k = 4		k = 5		k = 10	
index	acc.	loss	acc.	loss	acc.	loss	acc.	loss
Gbond	0.935	0.030	0.989	0.076	0.984	0.236	0.958	0.277
Corp AAA	0.969	0.147	0.988	0.044	0.983	0.054	0.946	0.242
Corp AA+	0.984	0.063	0.984	0.061	0.988	0.052	0.943	0.374
Corp AA	0.965	0.095	0.972	0.130	0.946	0.198	0.934	0.506
Corp AA-	0.984	0.044	0.981	0.069	0.982	0.064	0.960	0.223
Corp Comp	0.974	0.064	0.977	0.061	0.979	0.067	0.947	0.240

To test the stability of the EWS predicting performance, the model is cross-validated by varying the value of data splitting fold k, and display the results of the average value for accuracy and loss on each test sets in Table 5.17. Either the 4-fold or the 5-fold cross validations bring the greatest accuracy for most bond indexes except the bond index of Corp AA-, which performs less accuracy and greater loss as $k = 4$ and $k = 5$. As the fold count continuously gains (till $k = 10$), the model performance will become less robust with decreasing accuracy (below 0.96) and increasing loss (beyond 0.2).

Then, compared with other stylized predictive models that are constantly hired in the previous studies of forecasting the financial crisis, two levels of examining measurements are implemented: accuracy and loss on out-of-samples and the ROC curve with AUC. To make the model comparison reliable, we hold the following three conditions that 1) the target crisis variable same defined, 2) all factor variables same included in each predictive model, and 3) the samples are covered in the same time horizon.

Three groups of models will be compared, as Table 5.18 shows, (1), (2) and (3) label the comparative models of attention mechanism combined deep neural networks (Attention-BiLSTM, Attention-LSTM), the pure deep neural networks (BiLSTM, LSTM, RNN) and the other prominent parametric models (BPNN, SVR, BL), respectively[26].

[26]The non-parametric model of KLR which approach focuses more on distinguishing the leading economic factors to forewarn the crisis will not be included here but will be hired as baseline model in contagion fused EWS project.

Table 5.18: Predictive models comparison in terms of accuracy and loss for out-of-samples.

	(1) Attention--BiLSTM	Attention--LSTM	(2) BiLSTM	LSTM	RNN	(3) BPNN	SVR	BL
panel (a): acc.								
Gbond	0.991	0.985	0.962	0.908	0.929	0.937	0.895	0.854
Corp AAA	0993	0.973	0.845	0.870	0.857	0.833	0.855	0.891
Corp AA+	0.995	0.990	0.949	0.929	0.952	0.890	0.796	0.917
Corp AA	0.988	0.992	0.988	0.972	0.981	0.878	0.878	0.863
Corp AA-	0.995	0.953	0.913	0.933	0.964	0.763	0.877	0.910
Corp Comp	0.989	0.930	0.835	0.870	0.877	0.858	0.859	0.905
panel (b): loss								
Gbond	0.030	0.045	0.038	0.079	0.042	0.285	0.675	0.801
Corp AAA	0.016	0.268	0.107	0.092	0.125	0.453	0.335	0.303
Corp AA+	0.012	0.057	0.045	0.065	0.122	0.289	0.431	0.476
Corp AA	0.045	0.038	0.011	0.025	0.066	0.256	0.441	0.771
Corp AA-	0.024	0.038	0.350	0.101	0.106	0.611	0.440	0.313
Corp Comp	0.028	0.220	0.136	0.136	0.142	0.417	0.366	0.238

From the accuracy and loss value table results, the attention combined deep neural networks group is the most effective one among all groups. In particular, the attention mechanism boosts the prediction accuracy, and the effectiveness surpasses the bi-directional layer. It is evidently proven as comparing the results of between attention-LSTM model and BiLSTM model, and the bidirectional LSTM model produces inferior performance than the attention based single layer LSTM model.

The bidirectional structure seems to bring improvements for deep neurons. Comparing models in the group of (1), the Attention-BiLSTM results in greater accuracy and lower loss than the Attention-LSTM for all bond indices except for the Corp AA. For the Corp AA bond, the attention-BiLSTM model has an interior performance than the attention-LSTM with smaller accuracy ($0.988 < 0.992$) and greater loss ($0.038 < 0.045$).

While the dual-directional information processing layer will lose its advantage without combining with the attention, in particular, comparing the results in the group of (2), the accuracy value for the Corp AAA index is 0.845 in the BiLSTM, lower than 0.870, the accuracy values brought from the LSTM, even lower than the value results from the RNN, 0.857.

In summary, deep neural networks with the attention mechanism perform best. That is, the data idiosyncrasy has the least impact on the model prediction. In the group of (1), both the accuracy and the loss hold in a stable prominent level, while neither (2) nor (3) groups can bring stable predicting performance for different credit rating level bonds.

To clarify how the bond idiosyncratic credit rating level affects the model predicting performance, the ROC curves and AUC values are obtained for all model groups predicting results on each bond. As Figure 5.15 shows, all predictive models perform steadily for the bonds issued by the government and the corporate bonds with higher credit ratings (generally above AA) with high AUC values and left upper corner bent ROC curves. The situation deteriorates for the corporate bond with AA, since the ROC curvature to the upper left sharply lessens, and the AUC value meanwhile drops. Though the model seems less effective in predicting corporate bonds with lower credit levels, it is inaccurate to conclude that the

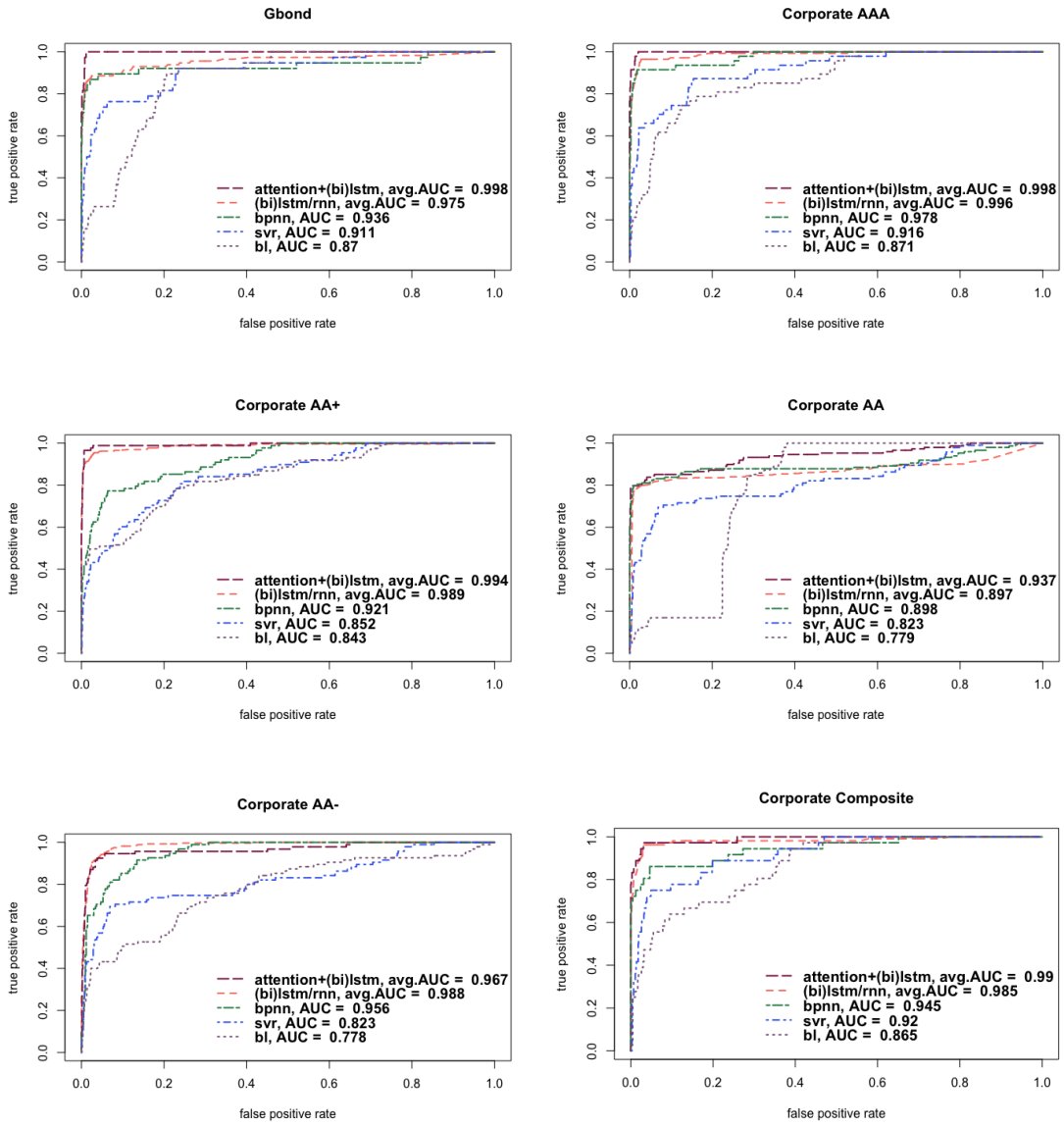


Figure 5.15: ROC curves for model comparisons.

The colored dotted lines are the curves for each model (group) and their names are texted at the right blank corner, from top to bottom are Attention-LSTM/BiLSTM, LSTM/BiLSTM, BPNN, SVR and BL, respectively. The avg.AUC refers to the mean of AUC values for models in the group of (1) attention combined deep neural networks and (2) pure deep neural networks.

bonds with higher credit risks will bother the model forecasting power because the model performance on AA- corporate bond is not worse than that on AA. It can be explained by the fact that a number of AA corporate bonds are either being or prepared to be degraded from higher credit levels to lower credit levels by rating agencies since most AA corporate bonds have been rated artificially high, which may misdirect the final predictions.

5.3.6 Attention drawn leading factors

The attention mechanism evaluates the impact of each factor. Figure 5.16 presents the attention mechanism drawn weights for each input features.

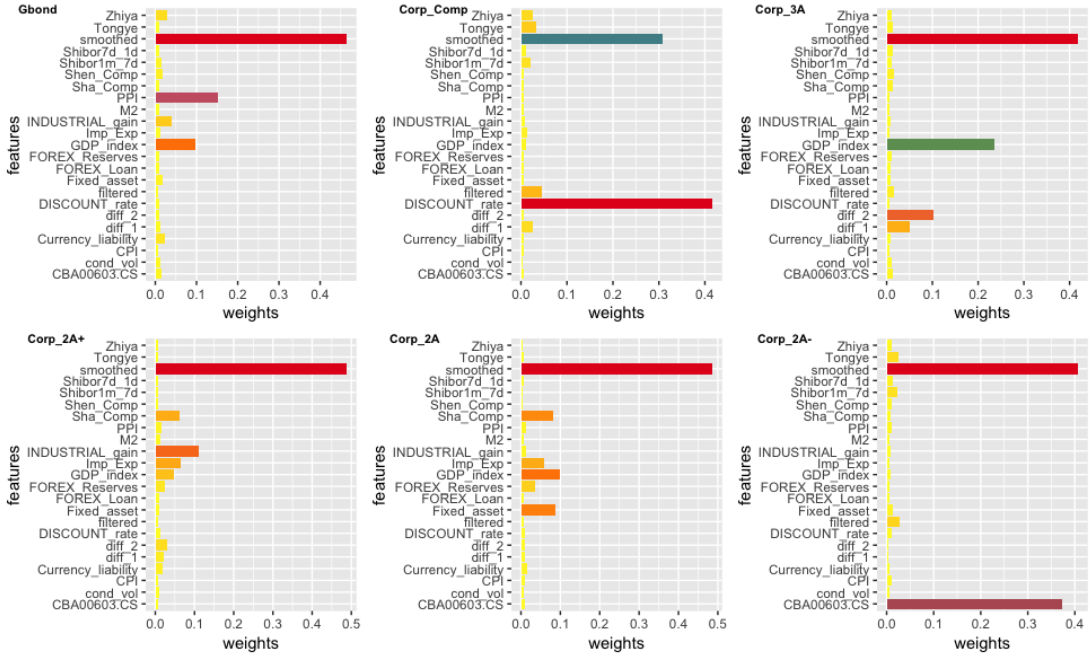


Figure 5.16: Bar chart for the attention vector on input features.

The vertical axis labels the name of factor variables and the horizontal axis scales the value of weights. A range of colors is used to color the bars with different weights, where red means the value of weight is high, green/blue implies the medium and yellow/orange indicates the light.

From the figure bars, the RCM-SWARCH estimated smooth probability for the high-volatility state is regarded as the first leading factor, except the composite corporate bond index, whose the most leading factor is estimated to be the national discount rate. In essence, the composite corporate bond index represents the diversification of credit risks of different credit rating leveled bonds. Thus, it can serve as a benchmark of de-risk investment in corporate bonds, that is, the exact role of discount rate in balancing the investment risks in the nationwide financial markets.

The price index is unlikely to be the key factor in forewarning the bond crisis, except for the AA- corporate bonds. It may imply either of that 1) the published price index for the AA- credit rating bonds is more credible to reflect the bond market turbulence as comparing with other bonds that heavily rely on the leading technical indicators (such as smooth probability), and 2) the price dynamics of

lowering credit rating bonds with higher credit risks is sensitively driven by the market turbulence, which guides the model to account for more weights on the price itself other than external economic factors.

Among the external economic factors, the GDP index has been assigned with noticeable weights on predicting credible bonds, such as government bonds and corporate bonds with credit ratings above AA-. In addition, PPI and industrial gain also play essential roles for the government bonds, whilst for the corporate bonds, the stock market factor of Sha_comp is more influential. It is in line with the conclusion of previous studies on the macroeconomic determinants for the Treasury bond and the corporate bond credit spreads in the U.S. (Liuren et al., 2008), which concludes the inflation factors, like GPD, CPI and PPI, effect on both the Treasury bond and all credit rating levels corporate bonds. In contrast, the market volatility factors, like the stock index S&P 500, have a small impact on bonds issued by the government but an increasingly strong impact on lower credit rating classes.

From the perspective of financial return decomposition (Long and Zhao, 2009), the discount rate component (DR) economic factors that reflect the time-varying risk aversion, such as GDP and PPI, will help to make forecasting the government bonds with higher credibility. The cash flows component (CF) economic factors related to firm fundamentals, such as industrial gain and fixed asset returns, will contribute more to predict corporate bonds with higher underlying credit risks.

5.3.7 Extra investigation on time steps

In our study, the attention mechanism is being extra used to examine two unresolved questions of 1) whether the change of time steps size will affect the predicting precision, and 2) whether the attention weight will not change its distribution on the time horizon as the size of time steps is either extended or shrank. In the experiment, we make the time step size vary from 5 to 10, 28, 50, the most popular horizon intervals for financial studies[27].

From the Table 5.19 shown results, extending the size of time steps generally

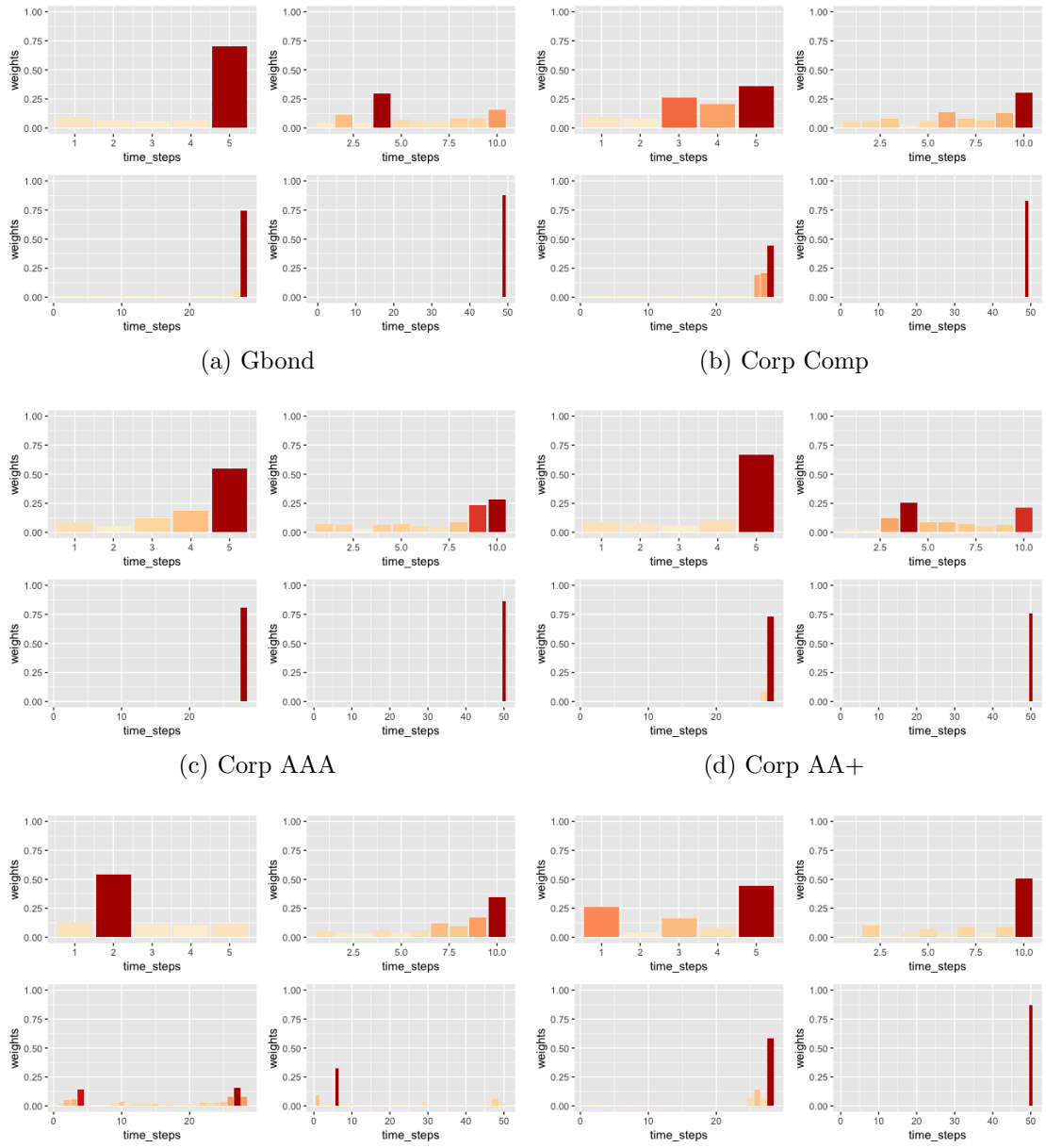
[27] 5=effective trading week, 10=effective trading twin weeks, 28=one lunar month, 50 =two effective trading months.

Table 5.19: Accuracy and loss value as time step varies for predicting on out-of-sample set.

	T=5	T=10	T=28	T=50
panel (a): acc.				
Gbond	0.9826	0.9903	0.9950	0.9949
Corp AAA	0.9654	0.9643	0.9624	0.9949
Corp AA+	0.9913	0.9825	0.9871	0.9858
Corp AA	0.9855	0.9836	0.9789	0.9858
Corp AA-	0.9659	0.9757	0.9871	0.9838
Corp Comp	0.9610	0.9659	0.9723	0.9818
panel (b): loss				
Gbond	0.0584	0.0347	0.0233	0.0208
Corp AAA	0.1344	0.1886	0.1527	0.0459
Corp AA+	0.0247	0.0418	0.0536	0.0661
Corp AA	0.0396	0.0218	0.0564	0.0379
Corp AA-	0.0554	0.0806	0.0476	0.0678
Corp Comp	0.0939	0.0913	0.0804	0.0676

improves the predicting effectiveness on Gbond, Corp Comp, Corp AAA, the bond with low credit risks, in gaining accuracy and reducing loss as more information from the past can be included. For the corporate bond of AA+, AA and AA- ratings with greater credit risks, the best performance will not uniformly appear on $T = 50$. The model with $T = 5$ and $T = 28$ perform best on Corp AA+ and Corp AA- respectively. We may illustrate this in two perspectives: on the one hand, the volatility makes time series inherently less predictable, early warning systems cannot extract more useful information from data over longer periods of time, in other words, updating more recent data is most helpful for predictions; on the other hand, the longer time may include implied default information, which has no significant influence on the prediction of bonds with high credit rating but has impacts on forecasting bonds with high credit risks. That is to say, the reason why the longest past information for the volatile bond indexes cannot boost the predicting effectiveness is that the default information contributes the unpredictability to bond indexes with higher credit risks.

Figure 5.17 shows the extracted value for model inferred attention weights on time steps for each bond index. It suggests three bond indexes of the corporate composite index, corporate bonds with AAA and AA- ratings, be allocated the



(a) Gbond (b) Corp Comp

(c) Corp AAA (d) Corp AA+

(e) Corp AA (f) Corp AA-

Figure 5.17: Attention vectors as time step varies.

The horizontal axis is the time steps and the vertical axis is the attention weights, and dark red colors the high weight bar and the dilute red marks the low ones. In each sub panel of (a)-(f), the bar plots are arranged from left to right, top to bottom as the size of time steps $T = 5, T = 10, T = 28, T = 50$. The red bar labels the greatest value of weights in attention vectors.

heaviest attention weight on the latest day regardless of the time step changes, which means the up-to-date information will do the greatest favor on predicting these three bonds, while the earlier information from the past will not. Then, the government bond and corporate bond with AA+ rating share a similar weight

pattern with the aforementioned three bond indexes that their attentions are subject to the rule of focusing on the latest information, except at $T = 10$, where the attention mechanism distributes the weights on the time horizon more evenly. The attention weights on corporate bond with AA rating are unique since changing the value of T will re-sketch the attention distribution in an irregular pattern.

In general, for the bond with high security (government bonds), the corporate bonds with diversified credit risks (corporate composite), and the corporate bonds with high credit rating level (corporate with AA+ and above ratings), the attention distributes even weights on short term information (from 5-10 days). Introducing long-past information can improve the predicting accuracy for such bonds since they are relatively stable and have a low probability of default. The more information is recruited on the timeline, the better the predicting effectiveness is, especially considering the auxiliary economic factors contained on each time step. Increased credit risk will weaken the forecasting power by including long-term past information. Specifically, the bonds with low credit ratings have greater default rates, which leads to the persistence of long past experience increase the uncertainty of our forecasts as no default information is included.

5.3.8 Concluding remarks

In this project, the EWS for China's bond market is constructed with technical improvements for the crisis classifier and the predictive model by including the RCM volatility regime determination metric and the attention mechanism. The experimented results on a series of bond indexes show that this improved EWS version produces forewarning signals with prominent versatility and stability. The regime determination provides more appropriateness to classify the variations on volatility levels, and the attention mechanism combined deep neural networks is verified to be the best performed predictive model in terms of statistical metrics, cross-validations, and model comparisons.

Furthermore, the attention mechanism intuitively pictures the contributing degree of various economic factors to predictions and laterally shows the relationship between credit risks and economic factors. For example, high credit risk

bonds are significantly more affected by cash flow factors (like industrial gain and fixed asset) and market volatility than low credit risk bonds. As putting the attention on the time horizon, we find the long-term information cannot improve the prediction accuracy for high credit risk bonds since the model does not account for indicators (for example, the default rate, which may appear in the long term prediction for bonds with low credit rating) that directly reflect the credit risk. Such inference provides a comprehensible way to facilitate the governors and market participants to prevent unexpected loss from the crisis triggering adverse impacts.

The construction of an effective early warning system for China's bond market thus has three layers of guiding significance. First, the national authority should expand the issuance of government bonds, further promote green bonds, diversify bond products, and accelerate the circulation of low-risk bonds in the secondary market. Second, the financial regulators should strengthen the monitoring of cash flows of corporations and enterprises and provide timely financial support to those in difficulties. Last but not least, banks and bond trading markets should improve the credit rating system, unify bond rating standards, update and publish the newest rating information for corporations and enterprises in time, and eliminate the asymmetric information among different circulation markets.

5.4 Implications and discussions

In this chapter, the integrated EWS models for China's stock and bond markets are designed to alarm their respective asset turbulence driven crises to 1) provide practical information for market participants to make the decision in investing activities and 2) offer indicative evidence for the policymakers to assess the market vulnerability in regulating actions.

These two proposed EWS models share the same prototype as framing the composed architecture of crisis classifier, crisis predictor, and warning generator. However, have respective variants in each of the functional zones in the EWS frame. The EWS for stocks hires the two-peak dynamically thresholding-based SWARCH with two volatility regimes and the LSTM networks to identify and

predict stock market turbulence for providing timely practical instructions for investors. While for bonds, the volatility regimes are RCM determined not being intuitively given by high- and low- two volatility states anymore, and the LSTM neurons are stacked with attention mechanism and bidirectional layer to gain the EWS's ability in terms of subdividing the volatility degrees and boosting the forecasting precision.

According to the results of applying the developed EWS on the ten-year sample of China's stock index of Shanghai Stock Exchange Composite, we have the following contributions.

- On the one hand, the crisis indicator function based on SWARCH classification method could be directly used for post-date crisis identification in a dynamic way given no assumptions on crisis dates or thresholds.

- On the other hand, such an integrated EWS module can effectively alarm stock turbulence with adequate reacting forewarned periods and sufficient high-degree accuracy. Thus regardless of the market investors' risk aversion levels, this EWS can make early decisions to help react towards the potential market crashes.

By assessing China's bond market vulnerability, the other improved form of EWS based on the RCM-SWARCH crisis classifier and the attention combined bidirectional LSTM deep neural networks is recognized to contribute the literature in the following aspects.

- For one thing, the extra measure clustered high-volatility state can depict the turmoil periods for the bond market more precisely since it not only successfully detects critical events but notifies the credit risks place idiosyncratic influences on the model forecasting power in a disadvantageous way as well.

- For the other, the policymakers and other market participants can be sensibly instructed by appropriately referring to the attention drawn leading indicator factors that are proven linked with the level of underlying bond credit risks. The government bond and high credit rating bonds are more

likely to be informed by the factor variables relating to the national economy, while those which are firm personalized or stock market related will be more persuasive to notify the corporate bonds with lower credit rating.

The current two versions of EWS are proposed as the prototype of a hybrid combination of the volatility model and the deep neural networks, which is certainly yet finalized considering that further explorations will be constituted to adapt diverse research aims and situations.

Model-wise, the leading indicators drawn by the attention is neither compared with other leading-factor-detected models' (such as logit regression, KLR indicator approach and classification trees) performance nor examined in statistically validating metrics. Target-wise, the study of applying the developed integrated EWS on predicting systematic collapse-related crises, for example, sovereign crisis, is scant. Moreover, contagious factors have not been quantized in the EWS accepted form to be investigated their impact on predicting the crisis. The following chapter will propose the contagion fused EWS to estimate the probability of sovereign crises emerging in China to fill the gaps.

Chapter 6

Contagion fused early warning system

6.1 Link the contagion effect to crisis prediction

The initial motivation to merge the contagion effect into the crisis warning system is two-fold.

From the research topic perspective, numerous studies regard the contagion as the aftermath of the financial crisis and endeavor to emphasize the transmitting pattern across different financial sectors or regions when the crisis occurs (Fernández-Rodríguez et al., 2015; Kim et al., 2015; Yu, 2017; Bostanci and Yilmaz, 2020), however, few scholars survey the reverse relationship that indicates the role of risk transmitting conduction on the crisis prediction (Samitas et al., 2020), even some recent evidence have supported the fact that specific financial turbulence is also likely to be the transmitting risks triggered (Ibhagui, 2021; Feng et al., 2021). As Dawood et al. (2017) puts forward at the end of their study, the crisis forecast should include contagious factors as well as all macroeconomic factors within the scope of leading factors.

From the technical model perspective, to infer the propagation directions, the applicable models that detect risk transmission generally output structured results (Fink et al., 2016; Tiwari et al., 2019), which cannot contribute to most EWS frameworks as being directly input time series. Therefore, converting the risk transmission effects into the acceptable time-series input through appropriate quantitative methods will greatly save the cost of reconstructing the EWS, which allows for incorporating the structured risk transmission patterns. As yet, there

are no known crisis forewarning systems quantifying the contagion information in an EWS-assimilated way.

6.1.1 Time-varying contagious information input

To make our research consistent with the previous study on contagion detection, we first examine whether the constructed two hybrid CM in Chapter 4 can be directly fused in the EWS. It seems not very possible since neither the EVT nor the Copula can produce time-varying dynamics to reflect the changes in risk transmission over time. Though the bi-variate SWARCH model can estimate the probability of two markets simultaneously being in the high volatility state in a dynamic way, it does not mean that the pairwise generated probability can infer the change of real-time correlation degree of two assets for the high-and-high volatility state. In addition, the programming for BiSWARCH currently relies on the GAUSS platform, which software requires a fee for copyright. Such a non-free platform will increase the experimenting cost and is less efficient in saving computing time in practice.

Given the deficiency of existing research methods in contagion information transformation, the DCC-GARCH model seems to be the preferred choice since the output of dynamic correlation coefficients (DCC) can reflect the time-varying change of correlated degree between markets. The model coding has been well developed by optional packages in R, the free-open programming tool to implement models, henceforth the dual-requirements on time-varying and computing efficiency appear to be fulfilled.

As the contagion factor input to the EWS, the sole DCC does not fully provide the contagion information flow between the contagious source and the target markets unless the risk transmitting direction can be specified to determine whether the contagion is taking place. In other words, the sudden increase of the correlation between two markets does not necessarily mirror the occurrence of spillover effects until the crisis origin can be certain in the turmoil state. Thus, the DCC will be associated with the univariate SWARCH, which model infers whether the asset is in the turmoil or tranquil episodes as estimate the switching between high- and low- volatility states, to ease the embarrassment of using the sole DCC

caused blur in missing out the risk spreading direction. The newly proposed dynamics is called contagious intensity index and will be specified in the following section of 6.2.1.

6.1.2 Leading indicators horse race of between contagious and macroeconomic factors

After quantifying the contagion information into the EWS allowed input form, the following question considers the competition between leading indicators, specifically, between the newly recruited contagion factors and the acquainted macroeconomic factors that judge their impacts on producing crisis warning signals. In practice, both the classic and the stylized EWS models can implement such comparing functionality. However, as compared with the classic models, the deep neuron networks are less powerful in terms of validating the convincing degree of the learned impacting weights on each factor. Such shortage will make the inferred leading factors less persuasive, even though the stylized machine learning models have been verified to bring a higher precision on crisis forecasting.

To enable our proposed EWS model to be testified whether the inferred leading factors are significantly credible, we likewise put forward the hypothesis testing on the attention mechanism learned weight results for each input factors as selecting appropriate statistics[1]. In this way, it is believed that the stylized neurons based EWS models in terms of grasping leading factors and the horse race between contagious and macroeconomic factors are more convincing to policymakers and practitioners.

6.2 EWS for China's sovereign crisis

In the background introduction and literature review (refer to Chapter 1, Section 1.1 and Chapter 2, Section 2.4), we have basically depicted the situation of China's sovereign crisis and the debate on sovereign crisis definitions of that 1) there is not an explicit chronological list for China's sovereign crises to be

[1]In addition, the selection of statistics depends on the sample size and the corresponding empirical distribution.

referred[2] and 2) due to the complexity of the political and economic system, the way to directly transplant the definition quantified tools for either developed or emerging countries to China regardless of the database deficiency is not rigorously applicable.

The other challenge for the sovereign crisis EWS study is to make a choice from plentiful leading factors. The traditional way to synthesize a composite indicator index by thresholding leading factors makes the selecting procedure overwhelmingly cumbersome. Therefore, to avoid the verbose reckoning for each factor's threshold level, constructing a composite signaling index seems not a good option for us.

Chapter 2, Section 2.4 has discussed the advantages of using the volatility to define the sovereign crisis and analyzed the methodological techniques to quantify the contagion information as well as categorized most empirically studied determinant factors for the sovereign crisis. To terminate the debate and liberate the bear of brunt for uncertainties on definitions and leading factors, the contagion fused EWS for predicting China's sovereign crisis is proposed mainly on the basis of developed EWS for China's bond market, meanwhile relies on additional improvements for each functional module of EWS as follows.

1. The crisis variable for sovereign bond market will be quantified by clarifying the representative index's different volatile status in RCM-determined SWARCH model with more varied values for the count of regimes K.

2. The contagious information will be quantified in the contagion intensity index, a time-varying dynamics, by multiplying the DCC-GARCH estimated dynamic correlation coefficients (between the contagious sources and the sovereign bond), and the SWARCH inferred filtering probabilities of high volatile state (for contagious factors).

3. The attention mechanism will be not only applied to estimate contributing degree of input factors for both fundamental macro-economic and contagious factors but first examined its estimation's credibility in statistical hypothesis tests as well.

[2]As far as we know, there is yet an officially recognized list of sovereign default events for China

Hereafter, the formulated models for the RCM-SWARCH crisis classfier[3] and the attention-BiLSTM crisis predictor will not be repeatedly depicted. However, the variant changes of 1) proposing the contagion intensity index to combine the correlation linkages with the possibility of crisis originator being in crises, and 2) designing the hypothesis tests to validate the attention estimated leading factors' significance for credibility, will be clarified.

6.2.1 Contagious intensity index

Engle (2003) first proposed the dynamic correlation coefficient GARCH model to gain the flexibility to infer the pairwise time-varying correlated relationship in a continuous way, where the volatility will be dynamically adjusted and hence no bias will be brought, being a superior to either the CCC-GARCH Bollerslev (1990) or the BEKK-GARCH Engle and Kroner (1995) as fitting the multivariate class over time. The multivariate DCC-GARCH model is formulated as follows,

$$\mathbf{Y}_t = H_t^{\frac{1}{2}} \boldsymbol{\epsilon}_t, \tag{6.1}$$

$$\begin{cases} H_t &= D_t R_t D_t \\ D_t &= diag\{h_{i,t}^{\frac{1}{2}}\} \\ R_t &= (\sqrt{diag(Q_t)})^{-1} Q_t (\sqrt{diag(Q_t)})^{-1} \end{cases}, \tag{6.2}$$

where $\mathbf{Y}_t = \{Y_{1t}, Y_{2t}, ..., Y_{nt}\}$ is the vector of n asset returns, H_t is the conditional variance matrix, $\boldsymbol{\epsilon}_t = \{\epsilon_{1t}, \epsilon_{2t}, ..., \epsilon_{nt}\}$ is the vector of standardized residuals, D_t is the diagonal matrix of conditional standardized deviations for returns with the i^{th} diagonal element of $\sqrt{h_{ii,t}}_{i=\{1,...,n\}}$ being estimated from the univariate GARCH for each asset. R_t is the $n \times n$ symmetric dynamic correlations matrix being constructed through exponential smoothing estimators as follows,

$$Q_t = S(1 - \alpha - \beta) + \alpha(\epsilon_{t-1}\epsilon_{t-1}^{'}) + \beta Q_{t-1}, \tag{6.3}$$

where $Q_t = [q_{ij,t}]$ is the $n \times n$ time-varying covariance matrix for standardized residuals and S is the unconditional correlations for standardized residuals. α

[3]The binary function for crisis variable is same with Eq. 5.11, which has been clarified in 5, Section 5.3

and β are (non-negative) scalar parameters constraint by $(\alpha + \beta) < 1$. Appendix C will display the specific derivation for the dynamic correlation.

Then the time-varying correlation for each pair of assets, i.e. the dynamic correlation coefficient of $\rho_{ij,t} = [R_t]_{i,j}$, can be estimated to display the dynamic contentedness degree. However, as aforementioned, it is not appropriate enough to straightly use such correlation as contagious intensity given non-specified contagious sources and non-determined crisis periods. In other words, the contagious factors should be the deterministic origin of transmitting turbulence to the target receivers as long as being in crisis or turmoil episodes as least. The problem is *how to ensure such 'deterministic' crisis originator role for contagious factors?*

The contagious intensity index is thus proposed to resolve the embarrassment by first hiring the SWARCH to estimate the possibility of contagious factors being in crisis periods, and then multiplying the estimated pairwise DCCs between the possible contagious source and the receiver and the SWARCH produced filtering probabilities for contagious source factors over time. The formulation is written as follows,

$$Intense_t^{x,y} = \rho_{xy,t} \cdot Pr(s_{x,t} = high | \mathscr{X}_{t-1}; \hat{\theta}_x), \tag{6.4}$$

where x is the contagious source asset and y is the receiver. $\rho_{xy,t}$ denotes the estimated dynamic correlation coefficient between x and y. $Pr(s_{x,t} = high | \mathscr{X}_{t-1}; \hat{\theta}_x)$ is the SWARCH model inferred filtering probability for x based on previous observations[4].

6.2.2 Evaluation metrics on attention learnt weights

To make the attention learned weights more comparable to the classic regression model, which can infer the impacting factor significance for each estimate by implementing normal/t-test on each parameter, we do three levels of hypothesis tests on both in-sample and out-of-sample attention vector pieces to verify,

I) "whether they both significantly deviate from zero",

II) "whether they both have little variations from the full sample set",

[4]We do not use the smooth one here to prevent the future information leakage from contagious assets perturbing the crisis prediction.

III) "whether they perform steadily over time and show little difference with each other".

The corresponding hypotheses are aligned as follows

$$\mathrm{H}_0^I : \alpha_i = 0 \text{ v. } \mathrm{H}_1^I : \alpha_i \neq 0, \tag{6.5}$$

$$\mathrm{H}_0^{II} : \alpha_i = \mu_i \text{ v. } \mathrm{H}_1^{II} : \alpha_i \neq \mu_i, \tag{6.6}$$

$$\mathrm{H}_0^{III} : \alpha_{out,i} = \alpha_{in,i} \text{ v. } \mathrm{H}_1^{III} : \alpha_{out,i} \neq \alpha_{in,i}, \tag{6.7}$$

where α_i denotes the attention weight for each factor of x_i (i.e. the estimated contributing degree for factor of x_i). μ_i denotes the overall attention weight on the full population for x_i[5].

Thus the credibility on attention mechanism inferred factors' contributing degree can be testified in three ways: 1) inspecting the significance level for factors that are counted as non-zero contributors, 2) reviewing the consistency of attention drawn weights between train/test and the full set, and 3) examining the effectiveness of attention trained results working on out-of-sample estimations. The ideal result is to expect that the null I can be significantly rejected and the rest of the two nulls II and III can be accepted with supportive evidence.

To implement the hypotheses, corresponding test statistics of Z_I, Z_{II} and Z_{III} are designed as follows conditional on imposed essential assumptions[6].

$$Z_I = \frac{\bar{\alpha}_i}{S/\sqrt{T-1}} \sim t_{(T-1)}, \tag{6.8}$$

$$Z_{II} = \frac{\bar{\alpha}_i - \mu_i}{S/\sqrt{T-1}} \sim t_{(T-1)}, \tag{6.9}$$

$$Z_{III} = \frac{\bar{\alpha}_{in,i} - \bar{\alpha}_{out,i}}{\sqrt{\frac{1}{T_1}+\frac{1}{T_2}} \cdot \sqrt{\frac{T_1 S_1^2 + T_2 S_2^2}{T_1+T_2-2}}} \sim t_{(T_1+T_2-2)}, \tag{6.10}$$

where T denotes the count of observations, and S is the sample corresponding standard deviation. $\bar{\alpha}_i = \frac{1}{T}\sum_t \hat{\alpha}_i^t$ are mean values for estimated attention weights on each factor of x_i. The p-values for each factor will be calculated and inspected

[5]The overall mean of μ_i will be calculated via bootstrapping the attention extracted weights' empirical distribution from the full sample set.

[6]The assumption has to be requested before implementing the hypothesis testing since any statistic test must be given appropriate distribution. For example, the sample size should be large enough to allow the normal distribution to be approximated. In practice, the histogram will be pictured to visualize samples' normality, and Jarque-Bera statistic will be computed to check the imposed assumption's rationality.

at both 5% and 1% significance levels.

The preliminary condition for implementing the third hypothesis is that train and test sets share the same standard deviation. In practice, we first do F-test on the two sample standard deviations, if the test can be passed, then being transferred to the next step for testing the equivalence of the mean value. Otherwise, the t-test can not be implemented, and we alter the Scheffé test, which is more specific to test two different size sample sets with unknown standard deviations. This testing process will be described in the Appendix D.2.

6.2.3 Diagram the contagion fused EWS

The full implementing process for the contagion fused EWS frameworks is displayed in a packed diagram of Figure 6.1.

Step (1): At start, data for the target observations of $\{y_t\}_{t=1,\dots,T}$ and m contagious factors of $\{x_t^j\}_{t=1,\dots,T;j=1,\dots,m}$ (among all n features) will be put into the RCM regime selected SWARCH model to get (1) smooth probabilities for y and (2) filtering probabilities for each x^j, at high volatility state. Meanwhile, they will also input into DCC-GARCH model pairwise as (y_t, x_t^j) to get the dynamic correlation coefficients series of dcc_t^{y,x^j}.

Step (2): Then, the first layer output (1) for y can be used to define the crisis variables of C_t and hence C_t^f as including the post-effect window for one trading month prediction. The output (2) filtering probabilities for each contagious factor will be multiplied to dcc_t^{y,x^j} by point-to-point to generate the contagious intensities given the risk transmitting from contagious factors to the target asset.

Step (3): Last, the crisis observations and the intensity features as well as other macro-fundamental factors will be input into the predictive model, attention-BiLSTM, which has been structured in Figure 5.10 of last Chapter 5, to draw the prediction of $\hat{y}_{t+1}$ for probability being a crisis in next twenty days and the contributing degree of $\hat{w}$ (that is the estimated value for attention parameter of α in Figure 5.10) for each input feature.

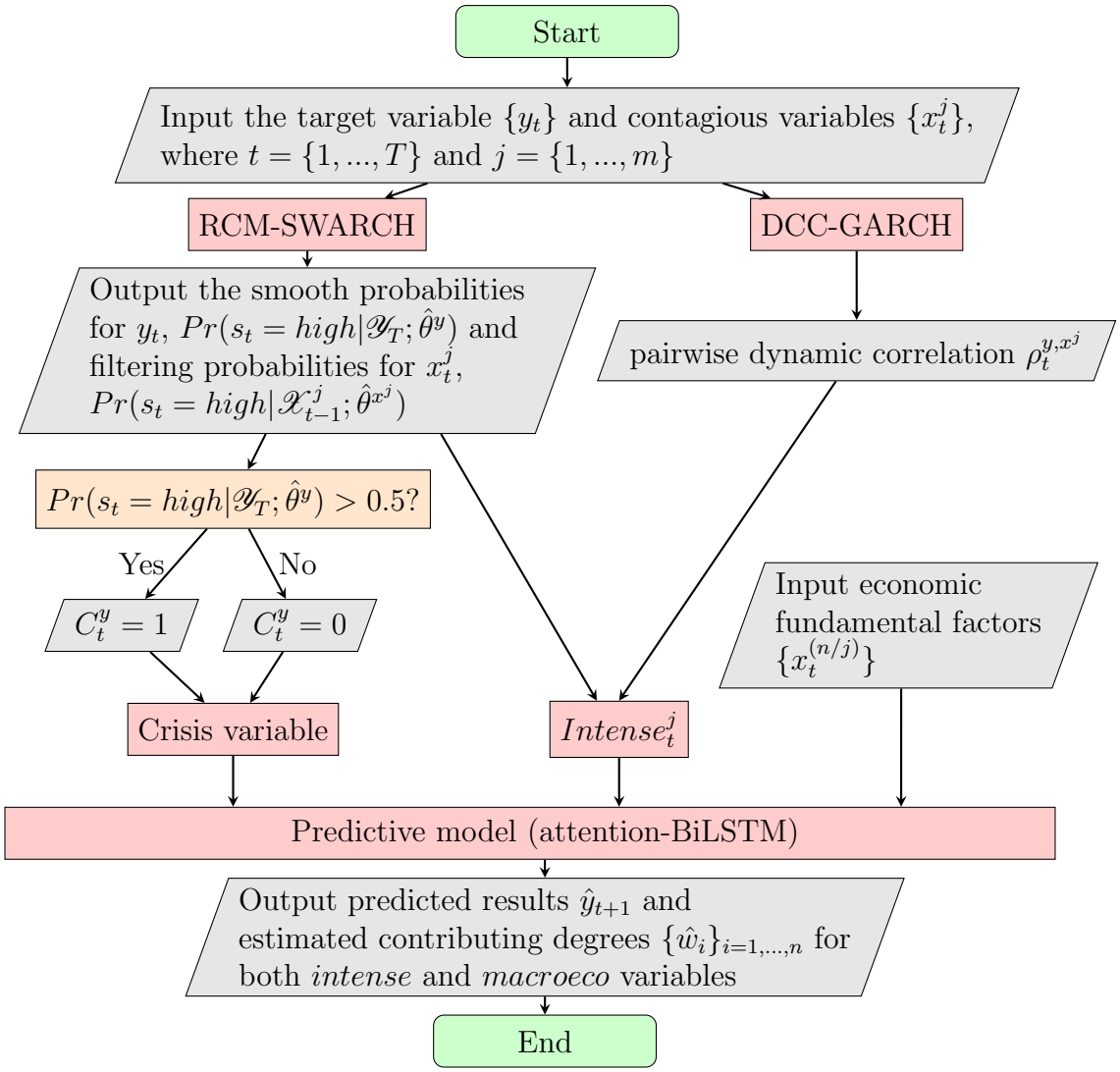


Figure 6.1: The diagram of implementing fused-EWS for sovereign crises.

Note: 1. The input factors are divided into two groups of contagious ones $\{x_t^j\}$ and economic fundamental ones $\{x_t^{(n/j)}\}$, given $j = 1, ..., m$ contagious factors.
2. $Intense_t^j = Pr(s_t = high|\mathscr{X}_{t-1}^j; \hat{\theta}^{x^j}) \cdot \rho_t^{y,x^j}$ is the quantified contagion intensity being the product of between filtering probabilities from SWARCH model and dynamic correlation coefficients from DCC-GARCH model.

6.3 Data

The daily China development bond (CDB) data, issued by the State Development Bank and actively traded in the secondary markets, will be hired to define the crisis for the sovereign bond. Unlike previous studies for the sovereign crises generally hiring either government bonds, bonds from treasury markets, or credit default swap (CDS) spreads, this quasi-sovereign bond from policy banks shows prominence in two aspects. One side, it shares the high credit level on the in-

ternational stage and has been accredited by the Standard & Poor's and other professional rating agencies to keep consistent with China's sovereign rating. On the other side, it has outstanding stock volume and high liquidity in the secondary trading market. According to China's Central Clearing House report in 2019, the CDB bonds' stock reached 3.8 trillion yuan, and the trading volume reached 60.86 trillion yuan.

As for the exogenous factors, four contagious sources are included. They are C1) seven-day weighted repos[7] that tightly correlate to the short-term interest rate level and investment sentiment in the bond trading market, C2) Shanghai and Shenzhen Composite Index (SSEC) that displays the Chinese stock market dynamics and acts as the totemic investment for other Chinese financial markets, C3) international gold price index and Brent oil price that represent the commodities in international trading activities, and C4) U.S. currency index and VIX that are recognized to lead the fluctuation for other countries' financial market turbulence.

In addition, five dimensions of macro-economic factors will instrument as the explanatory variables[8] to be aware of the sovereign crises.

E1) The overall portrait of economic development will be pictured by gross domestic production (GDP) on the year-on-year basis, macro-economic climate index, economic growth index, GDP constant prices for real estate cumulative on the year-on-year basis, capital and financial account balance, real economic leverage ratio, government sector leverage, financial sector (on debtors and asset sides) leverage, urban unemployment ratio and Engel's coefficient (EC) for urban residents;

E2) domestic production and consumption level will be depicted in industrial added value, tax revenue[9], consumer price index (CPI), consumer confidence index (CCI), and China investors' composite sentiment index (CSI)[10];

E3) domestic fiscal policy control intensity will target the money supply (M2),

[7]Repos are the repurchase rates for government bonds.

[8]We try to include all contributing factors that previous studies have mentioned. However, the availability for these macro-economic factors in the database confines our options.

[9]Tax revenue plays the crucial role for adjusting the commodities price and then further acting on the consumption level.

[10]The (CI)CSI is published by *National School of Development, to depict the sentiment changes of investors in the financial markets.*

RMB deposit-reserve ratio (RDR), and interest rate on demand deposit, as well as is synthetically counted into the monetary policy index by 9M Technologies[11];

E4) international economic communicating activities will be embodied into RMB real effective exchange rate, reserve assets and foreign exchange balance, ratio of total purchased bonds to total sold bonds by foreign investors, trade balance, and ratio of total imports and exports;

E5) ability to service debts will be marked into the ratio of short-term external debt to the foreign exchange reserves, portion of short-term debt to the balance of external debt, portion of total international balance of payments to GDP, debt service ratio, and debt-to-GDP ratio. All data are accessed from WIND database and their statistics are described in Table 6.2[12].

Table 6.2: Statistic descriptions for input factors.

Factors	Mean	St.D.	Skew.	Kurtosis	Freq.	Type
pr_idx	100.91	2.27	-0.35	-0.40	daily	
log_r	0.00	0.00	0.17	23.2	daily	endogenous
filt_k3	0.17	0.32	1.80	1.53	daily	
sev_repo	2.75	1.06	1.11	3.18	daily	
sev_repo_chg	0.00	0.04	0.35	15.57	daily	
ssec_chg	0.00	1.00	-0.34	4.34	daily	
gold	10905.4	5405.13	-0.22	-0.07	daily	
gold_chg	0.00	0.03	0.25	2.46	daily	
oil	73.97	25.81	0.44	-0.89	daily	
oil_chg	0.00	0.02	-0.23	11.21	daily	
dxy	86.86	8.01	0.14	-1.21	daily	contagious
dxy_chg	0.00	0.01	0.40	6.82	daily	
vix	18.76	9.37	2.68	9.54	daily	
vix_chg	0.00	0.08	2.12	16.69	daily	
ctg_brepo	0.00	0.03	-2.81	54.57	daily	
ctg_bstk	0.00	0.02	-8.19	92.86	daily	
ctg_bgold	0.00	0.01	-2.43	27.90	daily	
ctg_boil	0.00	0.01	-2.46	5.62	daily	
ctg_bdxy	0.00	0.01	-1.69	18.33	daily	
ctg_bvix	0.00	0.01	0.61	16.86	daily	
gdp	8.73	7.71	-0.38	-0.60	quarterly	
macro_idx	99.58	3.55	-1.34	2.66	monthly	

[11]The 9M Technologies is a limited company which provides comprehensive analytic solutions based on the world-class industry experience and advanced quantitative research to serve for the risk forecasting and portfolio management.

[12]The factors in Table 6.2 have been abbreviated in Table 6.1.

eco_gr_idx	0.44	0.18	0.33	-0.87	monthly	
gdp_realest	7.98	6.28	2.03	10.66	quarterly	
cap_fin_bal	253.55	426.01	0.56	0.16	quarterly	
realeco_lev	192.85	39.79	0.28	-1.06	quarterly	
gov_lev	33.05	6.29	5.01	44.90	quarterly	
fin_lev	1.37	3.95	11.85	148.44	quarterly	
unemploy	1.58	5.14	7.39	73.74	quarterly	
ec	34.09	3.57	-0.82	-1.02	annually	
indus_add	10.60	5.95	-1.08	6.42	monthly	
tax	13.4	14.55	0.57	1.45	monthly	
cpi	2.74	1.90	0.65	0.93	monthly	macro-
cci	109.11	7.21	0.75	-0.13	monthly	-economic
csi	36.85	7.71	-0.38	-0.6	monthly	
m2	14.68	4.81	0.95	1.16	monthly	
rdr	15.38	4.35	-0.63	-0.77	monthly	
depos_r	0.47	0.17	0.86	-1.12	monthly	
mon_pol_idx	0.15	0.82	-0.18	-1.64	monthly	
rmb_forex	107.58	15.3	-0.18	-1.39	monthly	
res_forex_bal	-433.46	756.7	0.72	0.49	quarterly	
buy_sell_overseas_r	1.48	1.39	5.37	45.80	monthly	
trade_bal	46.69	164.71	3.04	13.19	monthly	
imp_exp_r	11.71	16.52	-0.12	-0.55	monthly	
exdebt_fores_r	22.40	10.86	1.86	6.97	quarterly	
short_bal_exdebt_p	63.59	9.34	0.37	2.02	quarterly	
intbal_pay_gdp_p	5.69	6.19	2.64	18.05	quarterly	
debt_ser_r	3.41	1.83	0.64	-1.15	annually	
debt_gdp_r	45.74	17.96	0.62	-1.28	annually	

Figure 6.2 visualizes the pairwise empirical correlations between input factors in the heat map. According to the depth of colored dots, we can judge their degree of orthogonality. The darker colored, the stronger correlated. Some factors will be suspiciously notified since they perform heavy correlation to more than 3 other factors. First, almost all contagious factors price indexes and their quantized intensities are rarely correlated with each other as well as to other macroeconomic scoped factors, except the U.S. currency index (dxy), which correlates more to the Engel's coefficient of urban households (ec), the debt to GDP ratio (debt_gdp_r) and the balance of reserve assets and foreign exchange (res_forex_bal). In contrast, the situation is quite different since most macroeconomic factors are correlated with each other more or less. The most noticeable correlations aggregate on following three groups of factors: i) the monetary policy (mon_pol_idex), the industrial gain (indus_add), the forex for Chinese yuan (rmb_forex), the money supply (m2), ii) Engel's coefficient (ec), the debt to GDP ratio (debt_gdp_r) and

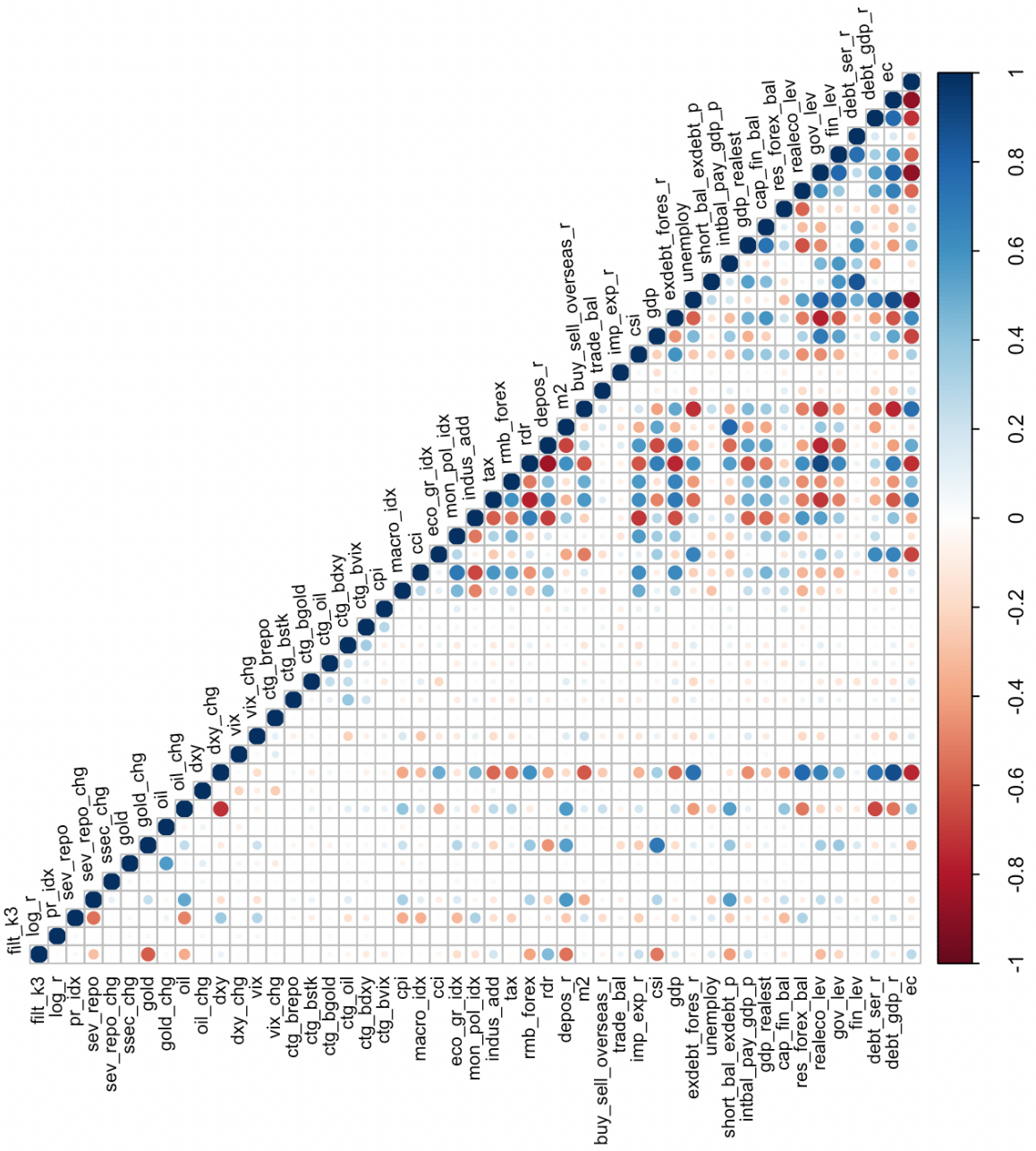


Figure 6.2: Correlation matrix of input endogenous and exogenous factors.

iii) the leverage ratio for the government sector (gov_lev), the (realeco_lev), the leverage ratio for the real economy sector (realeco_lev).

We frame sourced data from March 31, 2004 until September 1, 2020, the period that allows for covering several critical moments that once kicked in China's economic development. The target variable of C_t^f and the feature vector of $\boldsymbol{x}_t$ will be first neaten by their corresponding pre-processing methods (mentioned in section 6.2.1). By removing all not available values, we have the data frame

with 3858 rows (time horizon axis) and 50 columns (1 target and 49 features on variable axis). To make the data frame into the format that attention-BiLSTM model allows input, the time step is first set, in our case, to be 20, and hence the data will be converted into 3838[13] pieces of samples. It is essential to cover at least one crisis hump in the out-of-sample set for validations. Thus we split the data into 85% for the train and 15% for the test, specifically 3259 in-sample and 579 out-of-sample pieces, respectively. The test set spans from 2018/03/19 to 2020/09/01, which covers the most recent pandemic COVID-19 outbreak.

6.4 Preliminary analysis

In this section, two preliminary analyses of 1) the regime count determination by *RCM* in SWARCH inferred smooth and filtering probabilities for the target sovereign bond index and the contagious factors respectively, and 2) the contagious intensity index between the sovereign bond and the appointed crisis origins, are called before accessing the input for EWS.

6.4.1 RCM optimized K

Table 6.3: RCM values as K varies in SWARCH model frameworks.

$K =$	2	3	4	5	6	7	8
(a) target:							
cdb	50.0	0.000**	1.56	0.157	0.000**	0.000***	0.000***
(b) ctg sources:							
repo	12.8	0.309	0	0.000***	0.000***	0	0
ssec	10.1	0.000**	0.072	0.000***	0.000***	0.000***	0.000***
gold	5.22	0.000**	0.000***	0.000***	0	0	0
oil	10.4	0.000**	0.000**	0.000***	0.000***	0.000**	0
dxy	13.5	0.057	0.000**	0	0	0	0
vix	18.1	2.47	0.155	0.000**	0.000***	0	0

*** and ** denote the significance levels of 1% and 5%, respectively. The values below 0.01% will be noted as 0. The specific values for zeros with significance levels are shown in the Appendix.

[13]number of samples = count of observations - time step length

Table 6.3 lists the corresponding RCM value under each K[14] for the log returns of (a) target CDB index and (b) price index for each contagious source. From the table, $K = 2$ brings high RCM's for all indexes (especially 50 for the sovereign bond index), which, beyond questions, departs from previous studies regularly used 'high' and 'low' two classification cases as manifest that simply splitting two volatility states can afford to clarify volatile states no-whither. Furthermore, gaining the regime value K in a lengthy way will unnecessarily improve the model classifying performance in the volatility count testifying since the RCM does not shrink to zero continuously or steadily. For example, from panel (a), RCM equals to 0.023 under $K = 3$, it is smaller than 1.56 and 0.157 under $K = 4$ and $K = 5$. Even though the value keeps stepping down thereafter, limiting the number of regimes is essential in practice since over-refining the volatility level will blur the distinctive boundary across identified states. In other words, the greater K is given (especially after $K = 4$), the weaker susceptibility is performed in differentiating volatile levels, not to mention the expensive calculating process brought by assigning large K in SWARCH model estimation)[15]. To prevent such advantage-offset effects from overwhelming the contribution as adopt RCM in K's determination, K is chosen by obeying at least either two of the following three rules.

- The RCM value is approaching to but not necessarily be zero;
- The decrease between two adjacent values (from the front to the latter) is sharp enough to indicate the impact of gaining K;
- The plot for either smooth or filtering probabilities in high volatility state (among all distinguished volatility regimes) can intuitively differentiate the 'crisis' and 'non-crisis' samples without being aggregated around the 0.5 cutoff horizontal line or being too drastically distributed to form consecutive 'crisis' and 'non-crisis' segments.

In short, the RCM is not the dictatorial way to determine the optimal K, but incorporates other essential references in K's selection. In our case, the 'best'

[14]The regime count of K, in our case, varies from 2 to 8, which range allows us to distinctly identify different volatility cases for the index dynamics shifting among several turmoil levels.

[15]In practice, the plot for smoothed probabilities in high-volatility state performs more evenly as large K is provided, which will not benefit from drawing enough the crisis samples.

K seems uniform to be 3 for all indexes after screening conditions listed above. RCM value for VIX at $K = 3$ is 2.47, though noticeably high compared with 0.155 under $K = 4$, adorable as observing the high volatility state probability plot and the dropping steepness across each adjacent pair of increased K's. Thus the crisis variable and contagious intensity index are further brought forth after determining the K.

6.4.2 Identified crisis episodes

China has persistently held a special political control mode in financial markets for decades, making historical evidence for sovereign default or debt crisis events difficult for China to search. In this section, we tabulate the RCM-SWARCH model detected crisis episodes for sovereign bond index and list the critical events' timeline relating to the bond market as the auxiliary proof to validate the model detection's reliability.

Figure 6.3 visualizes the RCM-SWARCH model detected sovereign crisis episodes as the red shadowed region. There are eight RCM-SWARCH detected crisis segments unevenly scattering over the recent fifteen years time span. We associate the detected sovereign crisis episodes to critical events that possibly impact on the bond market in Table 6.4, where the key turning points have been clarified to unveil that the RCM-SWARCH detected march with these milestone events synchronously.

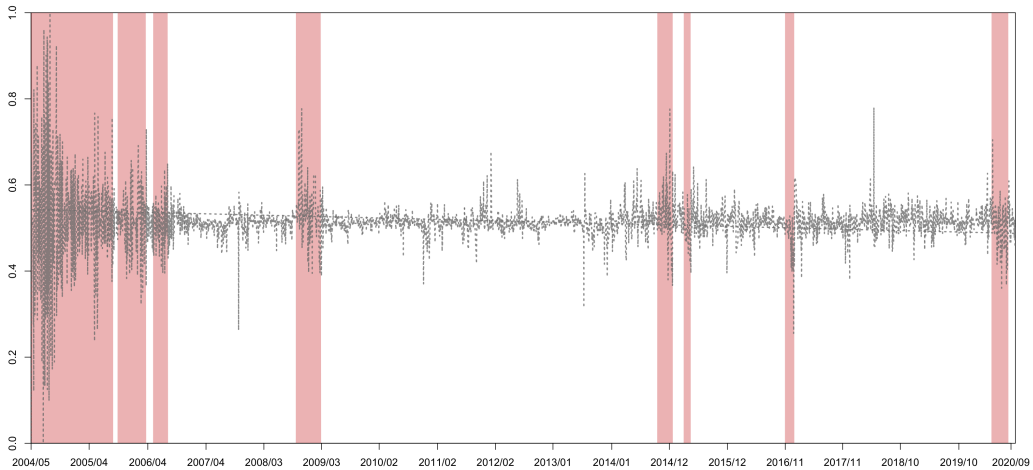


Figure 6.3: Identified sovereign crisis episodes.

Table 6.4: Chronologically contrast to the critical events.

identified crisis episodes	relevant critical events
2004/05/10−2005/09/20;	In late April of 2004, the authorities published a number of regulatory policies including increasing of the deposit reserve ratio, and meanwhile the continuous decline of the market exposed risks hidden by the previous national debt repurchase system's deficiency, which triggered the average net price of Treasury bonds fell more than 7.1%.
2005/10/18−2006/04/07;	Affected by the prudent fiscal policy, the capital in the Treasury bond market was excessive, and the interest rate fell sharply. The price of Treasury bonds fluctuated greatly as being influenced by investment sentiment since mid-October of 2005, with the overall index dropping from 112.5 to 110.9 within a month, and then rising back to 112.3 from November 22 to the end of the year.
2006/05/22−2006/08/18;	From April to July, tight monetary policies were issued intensively, the bond index momentum went down from 114.3 to 113.0 by dropping 1.3 points. From August to October, however, being supported by the huge force formed by a large number of idle funds, the bond market was continued to be pushed up with the China Bond Composite Index rising by 2.8213 points (2.5%) to 115.762 points.

2008/10/16−2009/03/16;	Since the fourth quarter of 2008, the central bank frequently cut interest rates, the 1-year fixed deposit rate was lowered by 108bp in October, which made the bond yield curve sharply and steeply declined. For instance, the yield curve for key inter-bank bonds moved down by 202 basis points (bp) on average. At the start of 2009, rapid credit growth triggered a shunt of bank liquidity and further shook the bond market financing area abundance to avert the market risk appetite to selling bonds buying stocks. Such actions made the Treasury yield curve go upward by 20 bp within one month.
2014/10/13−2015/01/15;	In the last quarter of 2014, the downward pressure on the economy was mounting, even though the central bank loosened monetary policy in December, it hardly eased the squeezed market liquidity. Meanwhile, the China Securities Depository and Clearing Co. Ltd. (CSDC) suspended the corporate debt being pledged to the Treasury. The bond market thus appeared frequent shocks till the start of 2015.
2015/03/23−2015/05/04;	China promoted the process of interest rate marketization and introduced the deposit insurance system in April 2015, but the financial market is underdeveloped, the benchmark interest rate system has yet been established, which made the transmission from the interest rate mechanism and monetary policy mechanism severely hindered, forming a "financial accelerator"[16].

2016/11/28−2017/01/23;	In the late November of 2016, the rapid devaluation of RMB increased the outflow of foreign reserves, which brought heavy capital tension. In addition, the central bank continued the prudent monetary policy by locking short and releasing long in the open market to raise the cost of funds. In December, the burst of credit default event, the Sealand Securities fraud incident also deepened the investor's distrust in the bond market. In the triple effects from policy, capital and credit crisis, bond yields accelerated upward, and credit spreads widened.
2020/05/07−2020/08/17.	After the reform of LPR in China, the interest rate transmission mechanism was basically formed. However, under the impact of COVID-19, the downward pressure on the real economy increased since the first quarter of 2020, and the non-performing loans of banks accumulated. The one-year LPR was intensively cut by 20 basis points from February to April to 3.85%, which caused the bond market turmoils.

6.4.3 Correlated pattern

In this section, the contagious effects are first analyzed by visualizing the dynamics for the DCC-GARCH inferred time-varying correlation coefficients. Then, the pairwise risk transmission intensity that synthesizes with the SWARCH estimated probabilities for contagious source being in the high volatility state are pictured by displaying how recognized contagious origins effect varies against time as the transmitting direction is pinned from the sources (contagious factors) to the receiver (sovereign bond).

[24]financial accelerator is the principle of the occurrence and transmission of the U.S. 2008 financial crisis.

Figure 6.4 pattern shows the pure correlated relationship regardless of the risk transmission direction given the contagious factors being crisis origin. Each pairwise correlation fluctuates drastically except that between stocks and the sovereign bond showing significant trend (with little deviated variations) of slowly crawling up till the year of 2018 (though still hovering around zero) and then sharply falling to negative within two years. The seven-day repo rate correlates to the sovereign bond most tightly since it performs the greatest correlation among all pairs for both positive and negative sides. International commodities of gold and oil and the U.S. currency generally play a heavier negative role than positive impact on China's sovereign bond, especially the crude oil almost holds all values below zero during the inspected episode. The VIX shares a median level correlation with two sharp bumps. One appears around 2008-2009, the time of the Global financial crash, and the other locates in 2013, the year of China sinking into severe money shortage, but no notable turbulence happened for the VIX.

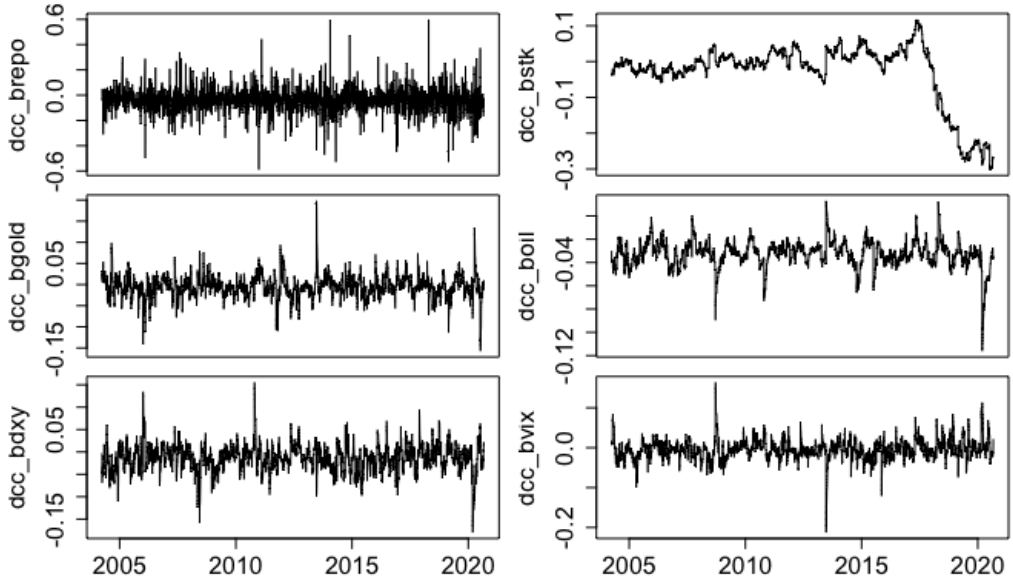


Figure 6.4: The pairwise dynamic correlation coefficients between the sovereign bond and each contagious source markets.

In such pure DCC inference, the correlated relationship that goes either positive or negative can be clarified, but the transmission intensity from specified contagious sources to the sovereign bond cannot be demonstrated. Without being

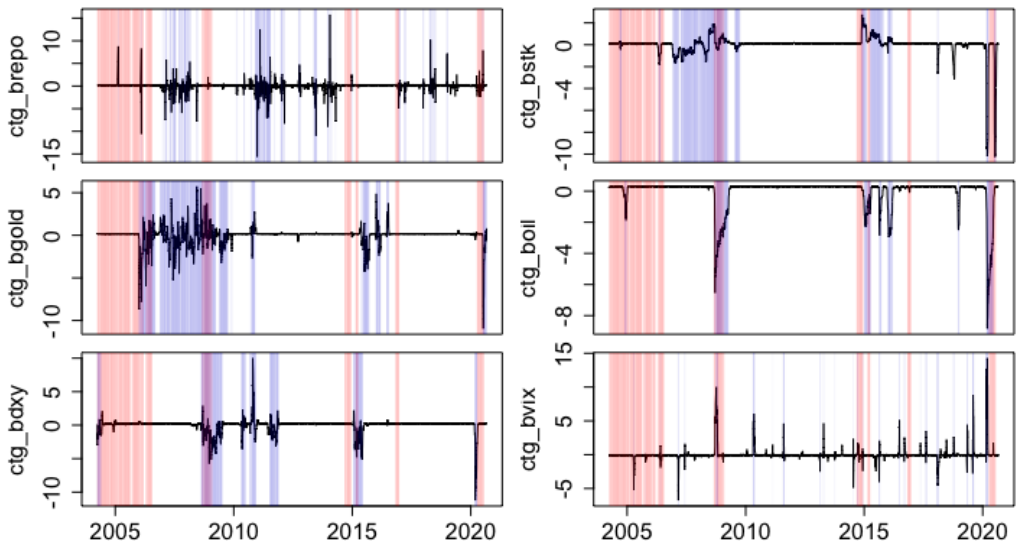


Figure 6.5: The pairwise contagion intensity indexes between the sovereign bond and each contagious source markets.

From the top left to bottom right, the subplots are aligned for repos, stocks, gold, oil, the US dollar index and VIX crisis transmitting intensity to the sovereign bond. Red region shadows the detected sovereign crisis episodes and the purple highlights the probability for each corresponding contagious factor being in high volatility state above 0.5.

provided the crisis origin or assessing the crisis origin bursting led risk transmission, it is unlikely to judge the role of contagious factors in the sovereign bond crises prediction by footing them in a fair place as macro-economic inputs[17].

Figure 6.5 shows the scaled contagion intensity index dynamics between the sovereign bond and the contagious source markets. From the figure, more distinct implications than the pure dynamic correlation coefficients can be inferred.

First, the combined index filters out some irrelevantly correlated vibrations from the co-movement dynamics. The product of correlation coefficients and the filtering probabilities for contagious origins staying in a high-volatility regime will minimize some inferred non-zero correlations (even high-level correlations) to zero when the origin is recognized as being in crises with slight possibility.

Second, not all contagious source turbulence will successfully transmit to the sovereign bond since the overlapping segments between the purple (crisis episodes for contagious sources) and the red (sovereign crisis episodes) do not perfectly

[17]The macro-economic factors including the authorities imposed fiscal policies directly intervene in market fluctuations, the impact of risk transmission, however, works on it in a more obscure and tortuous way.

equate to the turbulent intensity segments (i.e. the non-zero intensity values). For example, the sovereign bond and the gold are tightly correlated given the gold transmitting risks during 2006-2010, while the overlapping regions in that period are scattered as several slim bars at mid-2006 and mid-2008.

Third, the 2008 worldwide financial shock significantly augments the contagious effects between gold, U.S. currency, China's stocks, oil, and sovereign bond, because their abnormal vibrated intensity dynamics coincide with the overlapped crisis episodes for paired markets in the year of 2008. The COVID-19 impact, however does not share the same phenomenon even though the intensity index dynamics appear sharp valleys at the start of 2020.

Last, among six contagious sources, the VIX presents the least crisis observations and lightest transmission degree. The repo rate, which factor shows the strongest correlation in the above DCC inference, also shares the minimal overlapped crisis shadows even though neither contagious intensity nor detected crisis samples are weakest.

In a nutshell, we reasonably infer that contagious effects from the domestic repo and the VIX hardly perform as leading factors in sovereign bond crisis prediction. In contrast, the international commodities of gold and oil, their turbulence probably contribute more to the prediction during global crashes, such as 2008 financial shocks and 2020 COVID-19 impacts. The U.S. currency and national stocks' abnormal vibrations, though they become the impacting factors for sovereign bond turmoils, they severely lagged in 2015[18].

6.5 Contagion fused EWS predicting performance

In this section, to evaluate the proposed EWS model performance, it will be put in contrast with baseline models of 1) (static and dynamic) logit regression[19], 2) KLR indicator approach and 3) random forests[20], which three types of predictive

[18]Oppositely, the sovereign bond turmoils run ahead of the U.S. currency and the stock market in 2015, which implies the transmission direction perhaps revert.

[19]The dynamic EWS is first proposed by Candelon et al. (2014) to forewarn the currency crisis based on the binary logit regression. The dynamic scheme in that study is implemented either through the lagged crisis variable or through the time index to reinforce the endogenous indicators' persistence in prediction.

[20]Random forest is one of the classification trees aiming for diminishing the bias on the full set.

models i) have been accredited in previous EWS model development studies, ii) can estimate the indicators' contributing degree, and iii) cover different algorithmic routines in distinctive mathematical principles.

Two levels of measurements will be hired to evaluate the EWS forecasting power.

I. Statistical evaluation metrics of Quadratic probability score (QPS), Log probability score (LPS), Youden J and SAR, as well as the hit-ratios calculation on correct predictions, will be applied to test the predicting precision;

II. Forewarned days ahead of true crisis labels will be tabulated to inspect the forewarning effectiveness in practice.

In level I measurement, the SAR follows the formulation in last Chapter 5, Section 5.2, Eq. 5.6. The statistical metrics of QPS, LPS and Youden J are formulated as follows,

$$\text{QPS} = \frac{1}{T}\sum_{t=1}^{T} 2(\hat{y}_t - y_t)^2, \tag{6.11}$$

$$\text{LPS} = \frac{1}{T}\sum_{t=1}^{T}((1-y)log(1-\hat{y}) + ylog(\hat{y})), \tag{6.12}$$

$$\text{Youden J} = Sensitivity + Specificity - 1, \tag{6.13}$$

where QPS is the mean square error of comparing the predicted results to the true crisis labels. LPS is the logarithm of the probability estimate for the actual outcome. The Youden index measures the percentage of correctly predicted observations. The perfect model will have 0, 0, and 1 for the QPS, LPS and Youden J index, respectively.

Table 6.5 compares the four types of models in listing both of the statistical metrics and the hit ratios. From the table, the attention-BiLSTM, though it loses the horse race for the in-sample set compared to logit regressions and random forests, shares the least over-fitting effects among all predictive models by producing the greatest stable forecasting results between in-sample and out-of-sample sets.

Table 6.5: The statistical metrics results for comparative predictive models.

	static logit	dynamic logit	KLR	random forest	attn-bilstm
(a) in-sample					
QPS	0.038	0.038	0.371	0.003	0.065
LPS	-0.066	-0.065	−−	−−	-0.094
Youden J	0.918	0.919	0.237	0.996	0.907
SAR	0.944	0.945	0.678	0.987	0.929
hit ratio(calm)[a]	0.983	0.984	0.827	1.000	0.952
hit ratio(crisis)[a]	0.934	0.936	0.410	0.996	0.926
miss-out ratio[b]	0.014	0.014	0.128	0.001	0.027
false-alarm ratio[c]	0.012	0.013	0.136	0.00*	0.037
(b) out-of-sample					
QPS	1.12	1.27	0.265	0.261	0.135
LPS	-11.2	-10.4	−−	−−	-0.326
Youden J	-0.525	-0.517	-0.068	0.579	0.851
SAR	0.392	0.405	0.730	0.805	0.894
hit ratio(calm)	0.460	0.463	0.931	0.814	0.924
hit ratio(crisis)	0.015	0.014	0.00	0.765	0.875
miss-out ratio	0.116	0.116	0.118	0.028	0.010
false-alarm ratio	0.477	0.462	0.061	0.164	0.067

* denotes the significance level at 0.1%.
[a] hit ratios for calm and crisis are the percentages of corrected predictions for non-crisis and crisis samples, respectively.
[b] miss-out ratio counts the percentage of non-caught crisis signals in all predictions.
[c] false-alarm ratio is the percentage of generated warning signals in the actual tranquil periods. It should be notified that the false-alarm ratio does not distinguish the effective true early warnings (followed by actual crises) from the false alarms. Such ambiguous confusion will be clarified in Table 6.6.

Specifically, the greatest Youden J (0.996), SAR (0.987), and highest hit ratios for corrected predictions (100%) are brought by the random forest on the train data set. Such metrics values (of 0.579, 0.805 and 81.4% respectively) strikingly decrease to lower than that for attention based Bi-LSTM networks on the test set, not to mention the logit regression, which models severely lose the predicting accuracy on out-of-sample data by dropping to almost zero hit ratios (0.015 for static logit and 0.014 for the dynamic) for calling warning signals.

Table 6.6: The count of called onsets that effectively forewarned crises.

	Days in advance[a]								(called/total)[b]	forewarned[c]
	0	1	2	3	4	5	6	7	×100%	days
(a) in-sample									total = 6[d]	
static logit	2	2	2	2	2	2	2	2	33%	1.17
	(2)[e]	(4)	(5)	(5)	(6)	(5)	(3)	(3)	(68.8%)	(2.41)
dynamic logit	2	2	2	2	2	2	2	2	33%	1.17
	(2)	(4)	(5)	(5)	(6)	(6)	(4)	(3)	(72.9%)	(2.70)
KLR	3	3	3	3	3	2	0	0	35.4%	2.59
	(3)	(4)	(4)	(4)	(4)	(3)	(2)	(0)	(50%)	(2.79)
random forest	5	5	5	5	5	5	5	5	83.3%	2.92
	(5)	(6)	(6)	(6)	(6)	(6)	(6)	(6)	(97.9%)	(3.52)
attn-bilstm	6	6	6	6	6	6	6	6	100%	3.50
	(6)	(6)	(6)	(6)	(6)	(6)	(6)	(6)	(100%)	(3.50)
(b) out-of-sample									total = 1	
static logit	0	0	0	0	0	0	0	0	0%	0
	(0)	(0)	(0)	(1)	(1)	(1)	(1)	(1)	(62.5%)	(3.13)
dynamic logit	0	0	0	0	0	0	0	0	0%	0
	(0)	(0)	(0)	(1)	(1)	(1)	(1)	(1)	(62.5%)	(3.13)
KLR	0	0	0	0	0	0	0	0	0%	0
	(0)	(0)	(0)	(0)	(0)	(1)	(1)	(1)	(37.5%)	(2.25)
random forest	1	1	1	1	1	1	1	1	100%	3.50
	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(100%)	(3.50)
attn-bilstm	1	1	1	1	1	1	1	1	100%	3.50
	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(1)	(100%)	(3.50)

[a] The days in advance denotes the forewarned days before the crisis begins and the filled numbers count how many crisis onsets are correctly continuously called with specific forewarned days. For example, 6 crisis onsets are correctly called with 7 continuous days in advance for attn-bilstm model while only 2 can be called by logit regression.

[b] The average correct percentage is calculated by $\frac{\sum_{i=0}^{7}(s_i)}{\text{total}\times 8}$, where s_i denotes the count of continuously predicted onsets in advance of i days.

[c] The average forewarned days are calculated by $\frac{\sum_{i=0}^{7}(i\times s_i)}{\sum_{i=0}^{7} s_i}$.

[d] The train set actually contains 7 crisis onsets. We count as 6 since the train set starts with the 'crisis' observations, which cannot be counted as an effective crisis onset in model comparisons.

[e] The bracket here denotes the count of singular day, such as the 1^{st}, the 2^{nd} and so on, ahead of the crisis onsets, which is different from the un-bracket ones that count for the days in advance continuously.

Table 6.6 verifies the attention-BiLSTM prominence in terms of crisis onsets prediction. From the panel (b), out-of-sample statistics say both attention-BiLSTM and random forests model calling crisis onsets reaches to 100 percent correct. However, the random forest misses out one crisis onset for the in-sample

set (see panel (a)) which brings 83.3% correct crisis onset prediction ratio, less than the attention based BiLSTM model brought 100% correct prediction ratio. The classic logit regression and KLR poorly perform on the test set with zero correct prediction rates. From the last column, the average forewarned days ahead of the true crisis onsets, the attention based model outperforms other contrast groups with 3.5 forewarned days for both in-sample and out-of-sample sets, which means it generates the earliest reassuring warning signals before the crisis truly happens. All in all, the attention mechanism winged deep neuron networks are verified effectively to predict sovereign crises in terms of gaining the predicting precision, diminishing the false ratios, and drawing adequate early warnings in contrast with the classic predictive models of the static/dynamic logit regressions and the KLR indicator approach. The tree model of random forests brings comparatively prominent results in calling crisis onsets, however, produces more false alarms on out-of-sample set, making the random forests less persuasive from practical perspectives daily.

To sense the danger of sovereign crises, both investors and governors need to inspect the key contributors, which are customarily referred to as leading indicators. In practice, the operation of a complex EWS model and the access of EWS generated warning signals are usually pricey for the public. Inspecting the model suggested leading factors behaviors is the cost-effective way to be aware of the abnormal quakes for the sovereign bond in advance.

6.6 Leading factors for sovereign crisis in comparative analysis

The classic EWS models based on regression model Frankel and Rose (1996) and KLR indicator approach Kaminsky et al. (1998) have pioneered in leading factors detection and validated the estimated impacting degree for each factor. As mentioned before, though they have been popularly explored in the EWS development, the machine learning models yet achieve such comparative parametric credibility inspection on factors' marginal effects. This section will first visualize the attention mechanism drawn leading factors and then testify these contrib-

utors' significance to gain the machine learning-based EWS model's cogency in interpretability.

6.6.1 Attention drawn leading indicators

According to Figure 6.6, the weight vector distributions on the train (upper panel) and the test (lower panel) share a similar pattern. The leading factors are ranked in the following order.

1^{st} ranked: The gold price grows to the tallest (both bars above 0.35);

2^{nd} ranked: The constant price for real estate (counted in GDP) stands secondly (0.188 for the train and 0.258 for the test);

3^{rd} ranked: The oil originated crisis transmission takes the most noticeable role (0.126 for the train and 0.143 for the test) among all composing contagious index factors;

4^{th} ranked: The endogenous factor of filtering probabilities gets the equivalent importance level to the oil contagious index for the train (with 0.142) but performs weaker for the test (with 0.08);

5^{th} ranked: The CPI, the macro-economic factor that mirrors the inflation level, seems to provide some reference meaning to forewarn the sovereign abnormality with around 0.08 and 0.07 weight values for the train and test sets, respectively;

6^{th} ranked: The monetary policy index though it gets limited attention weights of 0.03 and 0.02 for in-sample and out-of-sample data sets, it is worthy of being noticed as compared with the rest of zero shared factors.

In summary, the endogenous filtering probabilities, the contagious sources of gold price and oil transmission impacts, and the macro-economic aspects of inflation level (CPI), real estate price, and monetary policy, play the leading roles to forewarn sovereign turbulence.

Then, two questions relating to the attention mechanism effectiveness on discerning the leading factors are explored: (1) can such six attentive indicators come through the proposed hypothesis tests (refer to Section 6.2.2) to testify their significance level of deviating from zero, consistency and stability? (2) whether their dynamics abnormally behave before and during the sovereign crisis episodes?

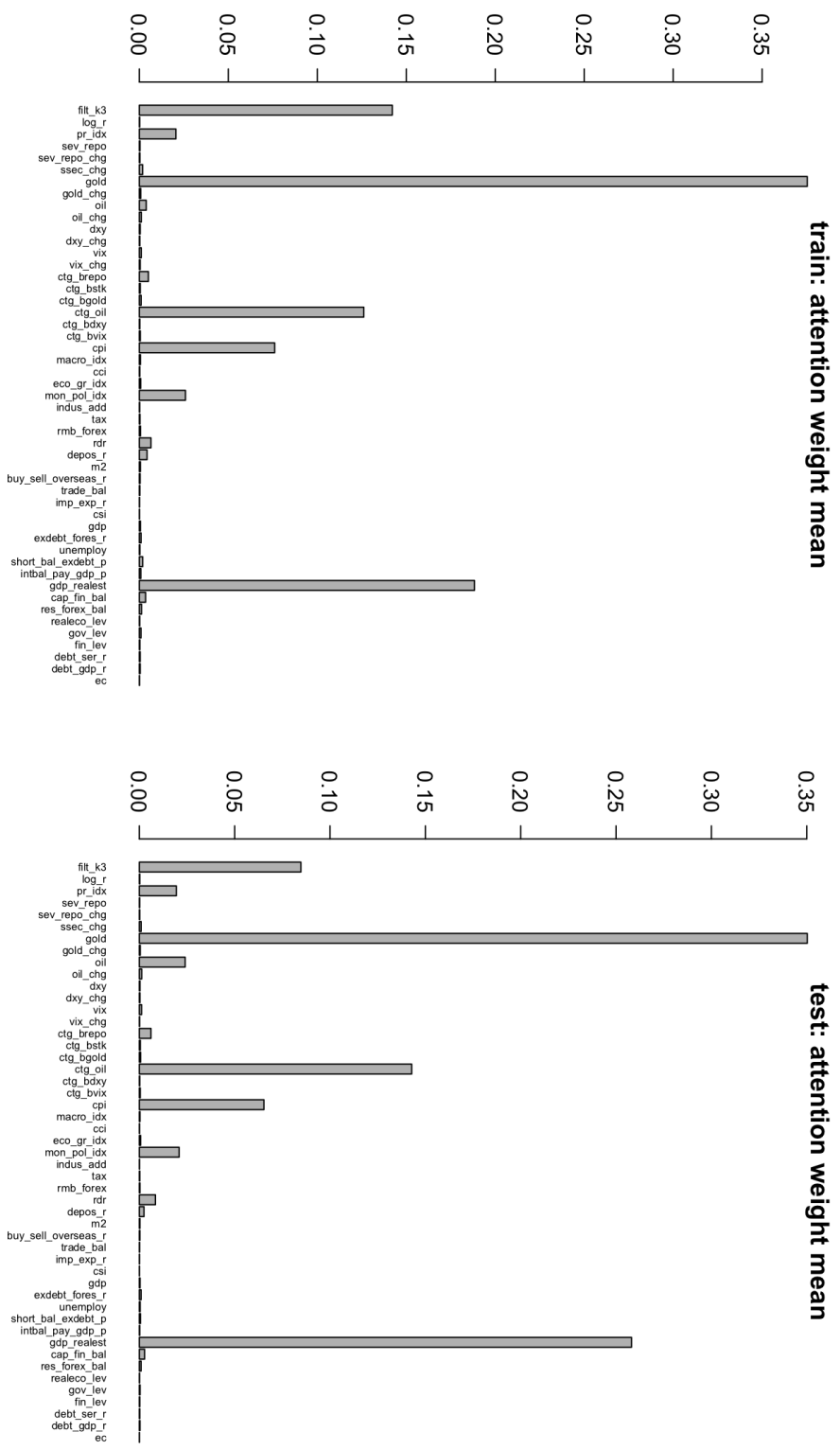


Figure 6.6: The mean of attention drawn weights for each input factors.

Table 6.7: Hypothesis tests on attention mechanism drawn weights for each input feature

			hypothesis test				
feature	est. attention weight		p_I		p_{II}		p_{III}
	in:	out:	in:	out:	in:	out:	
filt_k3	0.142	0.085	0.000	0.000	0.000	0.000	0.000
log_r	0.000*	0.000	0.115	0.479	0.004	0.000	0.137
pr_idx	0.021	0.019	0.000	0.000	0.278	0.023	0.378
sev_repo	0.000*	0.000	0.043	0.105	0.000	0.000	0.000
sev_repo_chg	0.000*	0.000	0.016	0.102	0.000	0.000	0.000
ssec_chg	0.002	0.001	0.000	0.000	0.000	0.000	0.000
gold	0.375	0.350	0.000	0.000	0.011	0.000	0.014
gold_chg	0.001	0.000	0.000	0.126	0.000	0.000	0.000
oil	0.004	0.024	0.000	0.195	0.000	0.000	0.053
oil_chg	0.001*	0.001	0.007	0.000	0.223	0.000	0.010
dxy	0.000*	0.000	0.040	0.236	0.000	0.000	0.000
dxy_chg	0.000*	0.000	0.025	0.232	0.000	0.000	0.021
vix	0.001	0.001	0.000	0.005	0.119	0.001	0.168
vix_chg	0.000*	0.000	0.072	0.117	0.000	0.000	0.000
ctg_brepo	0.005	0.006	0.000	0.000	0.000	0.000	0.000
ctg_bstk	0.001*	0.001	0.002	0.041	0.968	0.948	0.977
ctg_bgold	0.001	0.001	0.006	0.012	0.000	0.000	0.000
ctg_boil	0.126	0.143	0.000	0.000	0.011	0.001	0.177
ctg_bdxy	0.000	0.000	0.036	0.301	0.000	0.000	0.001
ctg_bvix	0.001	0.000	0.001	0.026	0.016	0.000	0.004
cpi	0.076	0.065	0.000	0.000	0.095	0.000	0.000
macro_idx	0.001	0.000	0.005	0.079	0.000	0.000	0.000
cci	0.000	0.000	0.164	0.462	0.000	0.000	0.000
eco_gr_idx	0.001	0.001	0.001	0.032	0.037	0.000	0.000
mon_pol_idx	0.026	0.021	0.000	0.000	0.000	0.000	0.000
indus_add	0.000	0.000	0.049	0.219	0.000	0.000	0.000
tax	0.000	0.000	0.004	0.118	0.000	0.000	0.000
rmb_forex	0.001	0.000	0.006	0.052	0.000	0.000	0.000
rdr	0.007	0.008	0.000	0.000	0.000	0.000	0.001
depos_r	0.004	0.002	0.000	0.000	0.000	0.000	0.000
m2	0.001	0.000	0.012	0.048	0.000	0.000	0.000
buy_sell_overseas_r	0.000	0.000	0.014	0.203	0.000	0.000	0.000
trade_bal	0.000	0.000	0.052	0.386	0.000	0.000	0.000
imp_exp_r	0.000	0.000	0.190	0.410	0.000	0.000	0.000
csi	0.000	0.000	0.186	0.351	0.000	0.000	0.000
gdp	0.001	0.000	0.019	0.162	0.000	0.000	0.000
exdebt_fores_r	0.000	0.001	0.000	0.009	0.277	0.005	0.291
unemploy	0.000	0.000	0.081	0.246	0.000	0.000	0.001
short_bal_exdebt_p	0.002	0.001	0.001	0.009	0.000	0.000	0.000
intbal_pay_gdp_p	0.001	0.000	0.006	0.111	0.000	0.000	0.000
gdp_realest	0.188	0.258	0.000	0.000	0.000	0.000	0.000
cap_fin_bal	0.004	0.003	0.000	0.000	0.004	0.000	0.000

res_forex_bal	0.001	0.001	0.000	0.027	0.000	0.000	0.000
realeco_lev	0.000	0.000	0.180	0.308	0.000	0.000	0.000
gov_lev	0.001	0.000	0.005	0.011	0.000	0.000	0.000
fin_lev	0.000	0.000	0.013	0.081	0.000	0.000	0.000
debt_ser_r	0.000	0.000	0.072	0.321	0.000	0.000	0.000
debt_gdp_r	0.000	0.000	0.015	0.125	0.000	0.000	0.000
ec	0.000	0.000	0.153	0.359	0.000	0.000	0.000

To answer the two questions, we first refer to the results of the right three columns in Table 6.7. According to the value for p_I, the null hypothesis I can be readily rejected by zero p's for the top-ranked leading indicators, which manifests the attention mechanism inferred contributing degrees are statistically significant to impact on the final prediction. However, the opt ranked indicators seem hard to pass through the type II and III hypothesis tests since none of them performs high p values to accept the nulls. In particular, the gold price is either hardly consistent with the attention estimated result on full samples (with p-value of 0.011), or unlikely performed the same pattern between the train and the test data sets (with p-value of 0.014), which both p values are below 0.05 significance level to reject the nulls of H_0^{II} for consistency and H_0^{III} for stability. Such results may expose the attention mechanism's weakness in transplanting inferred weight parameters with steadiness and compatibility, wildly as the sample varies.

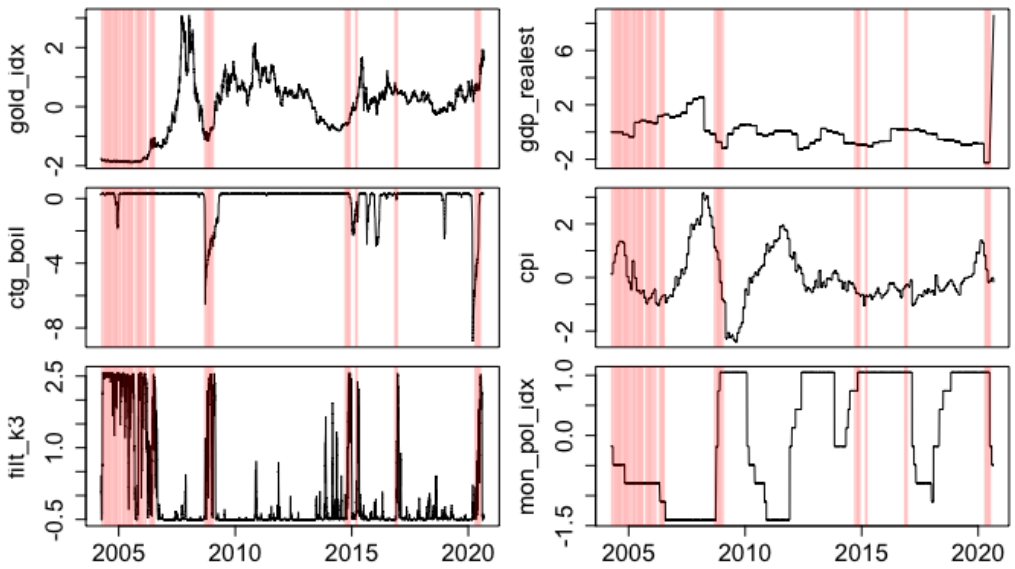


Figure 6.7: The sovereign crisis episodes imprinted leading indicators.

The second question can be answered by revisiting the dynamics of the leading factors in the observed periods to inspect whether any abnormalities appear during the crisis. Figure 6.7 shows the model distinguished the top six leading factors and imprints them with the sovereign crisis episodes (red segments).

From left panels to the right, we have observed that 1) the large humps of filtering probability for the most significant volatile state (i.e. filt_k3) almost perfectly coincides with the sovereign crisis periods, which is not surprising that the filtering probability shares a high similarity to the smooth one used to classify the crisis samples, 2) the international gold price and the quantized contagion information transmitted from the oil price seem to experience sharp falls before the sovereign bond being exposed to turmoils, which phenomenon is more significant during the 2008-09 global crash and the COVID19 outbreak, and 3) the macroeconomic factors of real estate price, CPI and monetary policy index show drastic either sharp rises or sudden sinks as to forewarn the sovereign turbulence, even though such leading effects sometimes are stamped with uncertainty which is suggested being bred by the distorted relationship between China's sovereign bond market volatility and the domestic economic development level as well as the imposed policy market-involvement degree. Regardless of poor hypothesis testing results on consistency and stability between the train and the test samples, the attention drawn leading indicators precede the sovereign turmoils as behaving exceptional fluctuations or sudden bounces.

6.6.2 Estimated leading factors from baseline models

Baseline models of random forests, KLR indicator approach, and logit regressions are implemented to be compared with the attention mechanism estimation.

Figure 6.8 bar plots show the random forest drawn feature importance on each factor. From the bottom panel for MDG, the most important is assigned to the deposit interest rate, which factor is regularly hired as an adjustment monetary policy tool to effect the sovereign risk pattern change (Afonso et al., 2018). The endogenous indicator of filtering probability is ranked second and followed by the exogenous factor of gold price.

The estimation from KLR approach and logit regression models, however,

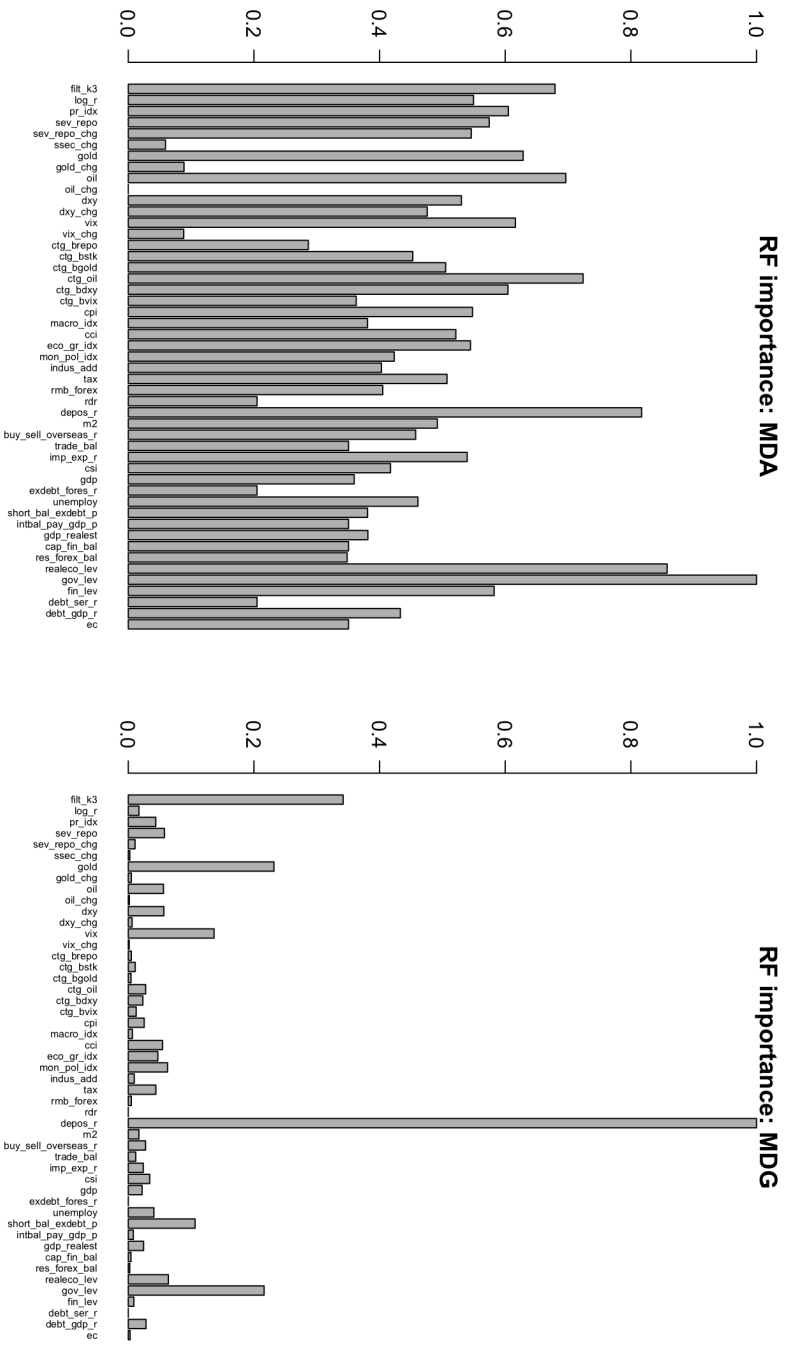


Figure 6.8: The Mean Decrease Accuracy (MDA) and Mean Decrease Gini (MDG) for Random forest inferred feature importance.

show different patterns on leading indicators. From Figure 6.9, the factors of the U.S. dollar index and the filtering probability are KLR approach recommended top two leading impacts[21] with the smallest noise-to-signal ratios.

[21]The red line and blue line are 0.5 and 0.75 cutoffs to threshold the noise-to-ratio value for

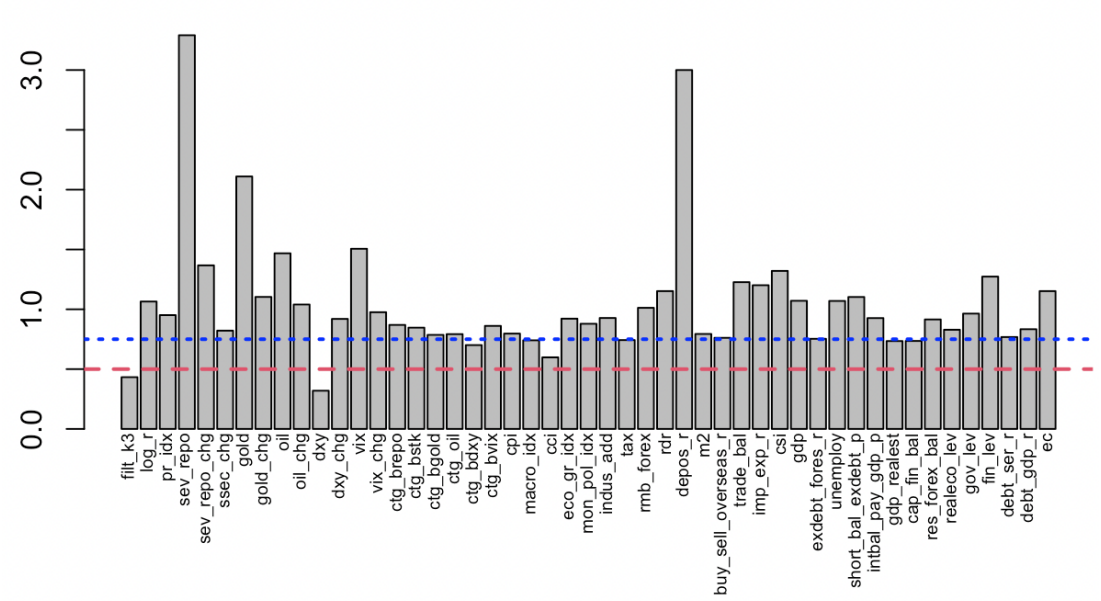


Figure 6.9: The noise-to-signal ratios (NSR) calculated by KLR approach.

According to Table 6.8, the logit regressions emphasize more on macroeconomic fundamentals relating to the fiscal stance, such as the balance of payments to GDP, real estate constant price in GDP, money supply, the balance between reserve assets and foreign exchange and financial sector leverage, which stressful state will gain the sovereign risk sensitivity to financial deficit (Beqiraj et al., 2021). However, given the ineffective predicted results, both types of contrasting models detected leading factors are suspicious of implying significant reference value in practice.

Table 6.8: The static and dynamic logit regression models estimated coefficients for input factor variables.

	static logit			dynamic logit		
term	est. coef.	t-stat.	p	est. coef.	t-stat.	p
(intercept)	-39.2	-4.84	0.00	-35.1	-4.53	0.000
y_20	−−	−−	−−	0.311	0.654	0.513
filt_k3	0.433	1.82	0.068	0.384	1.59	0.111
log_r	-0.184	-1.36	0.173	-0.191	-1.44	0.151
pr_idx	3.26	4.20	0.00	3.45	4.35	0.000
sev_repo	-1.71	-2.40	0.016	-1.56	-2.22	0.026
sev_repo_chg	-0.196	-0.846	0.398	-0.207	-0.875	0.382
ssec_chg	-0.051	-0.339	0.735	-0.057	-0.373	0.709

each factor in KLR. Moreover, the lower below the cutoff, the greater significance performs on leading crisis signals.

gold	-2.19	-2.599	0.009	-2.28	-2.77	0.006
gold_chg	0.127	1.09	0.273	0.138	1.18	0.236
oil	8.30	7.62	0.000	8.50	7.83	0.000
oil_chg	-0.172	-1.337	0.181	-0.171	-1.32	0.188
dxy	1.82	1.67	0.095	1.86	1.75	0.081
dxy_chg	-0.006	-0.043	0.966	-0.015	-0.112	0.910
vix	4.13	5.35	0.000	4.21	5.42	0.000
vix_chg	-0.159	-0.938	0.348	-0.166	-0.973	0.331
ctg_brepo	0.276	0.742	0.458	0.274	0.725	0.469
ctg_bstk	0.671	0.802	0.422	0.806	0.958	0.338
ctg_bgold	-1.18	-2.26	0.024	-1.25	-2.36	0.018
ctg_boil	0.991	1.41	0.158	0.939	1.36	0.174
ctg_bdxy	-1.67	-2.59	0.009	-1.64	-2.51	0.012
ctg_bvix	-0.064	-0.142	0.887	-0.023	-0.053	0.958
cpi	3.53	2.50	0.012	3.75	2.77	0.005
macro_idx	-3.28	-3.09	0.002	-3.60	-3.51	0.000
cci	15.6	5.88	0.000	14.5	5.67	0.000
eco_gr_idx	-8.91	-5.84	0.000	-9.17	-5.95	0.000
mon_pol_idx	4.12	2.15	0.031	3.60	1.91	0.056
indus_add	-4.67	-5.15	0.000	-4.96	-5.55	0.000
tax	0.803	0.559	1.44	0.632	1.18	0.236
rmb_forex	-10.4	-2.91	0.004	-10.4	-3.26	0.001
rdr	5.39	3.19	0.001	4.99	2.98	0.003
depos_r	2.81	0.969	0.333	2.02	0.696	0.486
m2	-23.0	-6.85	0.000	-23.2	-6.99	0.000
buy_sell_overseas_r	2.01	4.76	0.000	2.00	5.13	0.000
trade_bal	-4.37	-5.22	0.000	-4.44	-5.27	0.000
imp_exp_r	2.04	2.59	0.010	2.25	2.83	0.005
csi	-6.38	-4.56	0.000	-6.28	-4.61	0.000
gdp	24.1	6.53	0.000	21.9	6.51	0.000
exdebt_fores_r	-16.1	-3.12	0.002	-15.2	-3.08	0.002
unemploy	13.9	4.37	0.000	13.6	4.45	0.000
short_bal_exdebt_p	-21.5	-3.94	0.000	-18.9	-3.71	0.000
intbal_pay_gdp_p	27.8	5.46	0.000	26.6	5.61	0.000
gdp_realest	-21.2	-4.56	0.000	-18.6	-4.94	0.000
cap_fin_bal	1.52	1.25	0.211	1.51	1.26	0.206
res_forex_bal	23.7	6.62	0.000	22.4	7.05	0.000
realeco_lev	8.53	1.77	0.077	8.17	1.74	0.082
gov_lev	19.8	4.08	0.000	19.7	4.05	0.000
fin_lev	-24.3	-6.67	0.000	-20.8	-3.25	0.001
debt_ser_r	-10.2	-2.19	0.028	-8.44	-2.03	0.042
debt_gdp_r	-16.1	-5.34	0.000	-14.8	-5.11	0.000
ec	-17.9	-3.96	0.000	-16.8	-4.14	0.000

6.7 Concluding remarks and implications

In this chapter, we propose contagion information fused EWS for predicting China's sovereign crises on the high-frequency data basis for the first time to cook three hot potatoes in terms of both research issue and model improvement: 1) the scant reference for evidence China's sovereign crises, 2) rare risk spillover information being included as crisis determinants in the EWS frameworks and 3) bare significance tests being applied to attention based neurons model to earn the credibility on feature selections. We contribute the literature by addressing the three problems via, first of all, hiring the CDB index volatility defined sovereign crisis in RCM-SWARCH frameworks, second, quantifying the contagious intensity index to depict the risk transmission factors from contagious source, and last, implementing three hypotheses tests on attention weights vectors to verify the drawn leading indicators significance and steadiness. In the implementing procedure, the main findings can be summarized as follows.

1. RCM can effectively gain the crisis detecting precision by pre-determining the regime count value, and the custom regime number 2 is verified to be obsolete for both of the target CDB bond index and the contagion sourced market indexes;

2. The quantified contagious intensity index not only amplifies the spillover information that is proven to be an essential element in sovereign debt crisis prediction (Dawood et al., 2017), but specifies the risks transmitting impact from crisis originators to the sovereign bond market in a more detailed way as well;

3. As for the attention mechanism and the applied hypothesis tests on attention drawn weights vectors, they are considered to bring breakthroughs for time-dependent machine learning based predictive models in terms of interpretability and credibility.

The proposed EWS model's performance is investigated in contrast with both classic and stylized machine learning models. On the forecasting effectiveness side, the warning system based on complex designed deep neurons wins out in

the horse-race of competing with econometric predictive models of logit regression and classification tree of random forests on out-of-sample set by performing most excellent forecasting precision on crisis onset calls and highest hit ratios for both calm and crisis episodes. On the contributing degree side, the international gold price dynamics, the risk transmission from the oil, the domestic real estate price, and the inflation level (that CPI mirrors) are the most crucial indicator variables to forewarn China's sovereign crises according to attention mechanism estimation. It implies that the potential sovereign crises in China are not self-fulfilling but synthesized by both domestic forward-looking perspective indicators (excluding the agent rating assessment) and international risk transmissions.

Regarding the previous EWS studies of predicting the sovereign crisis, the debt exposure variables, banking sector variables, and foreign exchange markets are leading to the sovereign abnormality for most developed and emerging countries (Fuertes and Kalotychou, 2006; Savona and Vezzoli, 2015; Dawood et al., 2017; Ghulam and Derber, 2018). The signaling indicators for 'Chinese-style' sovereign crisis, however perform distinctively. First, China faces low exposure risks on external debts. According to the published China's national balance sheet in almost five years, China's net sovereign assets keep positive to cover its sovereign liabilities, indicating that the possibility of China being exposed to heavy external debts is low in the long term. In addition, the banking sector and the foreign currency market are strictly regulated in China's special centralized administrative management. Thus, the imposed financial policies have greatly released the banking pressure and stabilized exchange rate fluctuations (by pinning the U.S. dollar) in a timely and effective manner. Our study provides different leading channels to indicate sovereign turbulence.

Channel 1: The real estate market, one of the pillars in the national economy, has accumulated many bubbles in China and attracted massive social capital. Once bubbles burst, involved capitals will face severe devaluation risks and further make the local government debt condition worrisome.

Channel 2: Being the inflation level barometer, the (continuous) decrease of CPI also severely impedes production incentives, further restricts the social demand, and inhibits investment enthusiasm. Then, to release the capital liquidity,

the monetary policy of lowering the interest rate will be imposed and rapidly act on the CDB bond fluctuations as it is empowered with high liquidity in trading markets.

Channel 3: The gold price dynamics are unexpectedly revealed as the leading indicator for the Chinese sovereign bond market, which is rarely found in previous studies for both emerging and developed countries. The underlying mechanism can be explained as the gold being the safe-haven passive investment when investors question the sovereign bond stability for a country, tends to raise its price sharply.

Channel 4: The other international commodity of oil also shows the leading impact via its risk transmission intensity, as Figure 6.7 visualizes, the sudden falls of contagious intensity index of between the oil and CDB bond go ahead of detected Chinese sovereign turmoils. Like other BRICS countries (Chuffart and Hooper, 2019), China heavily depends on the crude oil resources import, which means the oil subsidies take a number of government expenditures, thus without a robust fiscal framework, mainland China will become fragile to resolve external shocks from the oil market.

Thus, some enlightened pieces of advice to politicians and investors will be suggested. On the one hand, to prevent the sovereign crisis in China, the authorities should note the excess accumulation of local debts in the real estate market, formulate reasonable debt risk control standards, and strengthen the supervision on the capital flows of local financing platforms. To enhance the sensitivity to the domestic inflation level, it is necessary to consider the impact on the liquidity of the secondary bond market when issuing fiscal policies on adjusting the interest rate level to avoid the bias of market capitals to highly liquid bonds. Meanwhile, the impact of international gold price on the domestic sovereign bond market should be paid more attention to, and the dependence on crude oil imports should be reduced by developing technologies for clean energies. On the other hand, the domestic investors should avoid being blinded by fanatical infatuation for investing the real estate and diversify risks of investing in the sovereign bond via opting industrial metals (for example, cooper) besides (precious metals of) golds to safer haven properties (Agyei-Ampomah et al., 2014).

Table 6.1: Abbreviations for input factors.

pr_idx	CDB price index
log_r	logreturns for the CDB price index
filt_k3	SWARCH inferred flitering probabilities for CDB index
sev_repo	average-weighted 7-day inter-bank repurchase rate
sev_repo_chg	return change for average-weighted 7-day inter-bank repurchase rate
ssec_chg	return change for the Shanghai and Shenzhen Composite Index
gold	worldwide gold price index
gold_chg	return change for the gold price index
oil	Brent oil price
oil_chg	return change for the Brent oil price
dxy	US dollar index
dxy_chg	return change for the US dollar index
vix	CBOE volatility index
vix_chg	return change for the CBOE volatility index
ctg_brepo	dynamic correlation coefficient between CDB and 7-day repos
ctg_bstk	dynamic correlation coefficient between CDB and ssec stock index
ctg_bgold	dynamic correlation coefficient between CDB and gold price
ctg_boil	dynamic correlation coefficient between CDB and oil price
ctg_bdxy	dynamic correlation coefficient between CDB and US dollar
ctg_bvix	dynamic correlation coefficient between CDB and volatility index
gdp	gross domestic product: constant-price on year-on-year basis
macro_idx	macroeconomic climate index: consistent index
eco_gr_idx	9M published economic growth index
gdp_realest	GDP: constant price for real estate cumulative on year-on-year basis
cap_fin_bal	balance: capital and financial items: quarterly value
realeco_lev	leverage ratio for the real economy sector
gov_lev	leverage ratio for the government sector
fin_lev	financial sector leverage: debtor side/ asset side
unemploy	Urban unemployment on year-on-year basis
ec	Engel's coefficient of urban households
indus_add	industrial added value on month-on-year basis
tax	tax revenue on month-on-year basis
cpi	consumer price index on month-on-year basis
cci	consumer confidence index
csi	consumer sentiment index
m2	amount of money in circulation: compared to the previous month
rdr	Renminbi deposit-reserve ratio: large depository institutions
depos_r	demand deposit interest rate
mon_pol_idx	9M published monetary policy index
rmb_forex	Renminbi real effective exchange rate index
res_forex_bal	balance: reserve assets and foreign exchange: quarterly value
buy_sell_overseas_r	ratio of between the total amount of purchased and sold Chinese government bonds by foreign investors
trade_bal	trade balance on month-on-year basis
imp_exp_r	import and export value on month-on-year basis
exdebt_fores_r	ratio of short-term external debt to foreign exchange reserves
shor_ex_debt_p	portion of short-term debt to foreign exchange balance
intbal_pay_gdp_p	portion of total balance of payments to GDP: quarterly value
debt_ser_r	external debt risk index: debt service ratio
debt_gdp_r	external debt risk index: debt to GDP ratio

Chapter 7

Conclusions and discussions

In this chapter, the main works of the PhD journey will be summarized in Section 7.1 and the innovated points and contributions in terms of the modeling exploration and practical significance to the literature will be concluded in Section 7.2. The study limitations and further exploration works will be discussed in Section 7.3. The enlightenment and some suggestions will be provided in Section 7.4.

7.1 Study review

This thesis mainly explores the tools to prevent the market turbulence-based financial crisis that brought damage for China by 1) figuring out the underlying transmission networks across domestic markets and 2) predicting the turmoils that potentially lead to crushing shocks for principal markets.

With such two aims, the projects of contagion models development and early warning systems construction are proposed in the context of China, in terms of solving four research questions - crisis identification, contagion channels detection, forecasting system design for high-frequency data, and contagion information fused crisis forecasting system development. In the persistent research survey, some existing methodologies are found to provide supportive bases for solvable ideas to the proposed questions. These methodologies bases are 1) SWARCH model to label the crises with observations being classified in high volatility state, 2) bi-variate SWARCH model and copulas to infer the contagious effects in terms of displaying co-movement episodes and structural transmission, respectively, 3) the time-dependent deep neural networks LSTM to learn warning signals through high-frequency factors, 4) attention mechanism stacked layer to gain the time-

dependent neuron networks inference on leading factors, 5) DCC-GARCH model to quantify the contagion into time-varying correlation dynamics, and 6) statistical hypothesis tests to validate the attention learned factors impacting degree.

On the method basis, we design four task modules (as Figure 1.1 diagrammed), practically three major projects (i.e. hybrid contagion models development, early warning system construction, and contagion fused early warning system construction), to complete the full PhD progress. For each of projects, we conduct the research sequence of *Background analysis* $\Rightarrow$ *Theoretical support* $\Rightarrow$ *Model construction* $\Rightarrow$ *Data analysis* $\Rightarrow$ *Empirical study*[1] to proceed with striving for innovations in research methods as well as endeavoring to contribute the practical significance.

7.2 Innovations and contributions

This thesis's innovative points and contributing remarks to the literature of financial risk transmission and early warning system development are embodied in aspects of methodological improvement and practical significance.

7.2.1 Methodological improvements

1. The study fills the gap of using dual-purposed CM models for developed and developing countries to compare risk transmitting mechanisms across principal domestic markets. In particular, two hybrid models of the BiSWARCH-EVT and paired SWARCH-EVT-Copula are simultaneously developed to detect spillover effects across domestic financial markets of stocks, bonds, the forex, and the real-estate in comparative analysis between China and the U.S.. On the one hand, it facilitates specifying the contagion periods given the crisis originator market and, on the other hand, explicating the transmitting channels without imposing the assumption on the crisis originator.

2. The study first proposes the prototype of an integrated early warning system

[1]For the last two projects on EWS construction, the empirical study includes the model validation.

to forewarn crises for specific markets. In the EWS frameworks, the research not only improves the crisis classifier in two aspects of dynamic threshold selection based on two-peak method and pre-classification of volatility fluctuation levels based on regime classification measurement but also includes time-dependent deep neural networks of (Bi-)LSTM as the predictive model, to finally produce the warning signals with practically guiding significance to the market through precisely learning and analyzing high-frequency data. The attention mechanism is further stacked with the time-dependent deep neurons to gain the stylized predictive models' interpretability on leading factors detection.

3. As an extended work, an early warning model fused with risk transmission information is developed to judge the exogenous factors leading to crisis prediction. The fused model is highlighted in terms of making the contagion intensity index (i.e. the product of DCC-GARCH model inferred pairwise time-varying correlation coefficients between the target market and the contagious sources, and the SWARCH model estimated filtering probabilities for contagious sources) to allow the inter-market spillover effects being quantified as an acceptable input factor in the EWS frame, and generating effective early warning signals as well as verifying the significance of attention mechanism selected crisis leading indicators by designing the hypothesis tests for inspecting the output credibility and stability.

7.2.2 Practical significance

1. By detecting the internal markets' linking dynamics during risk transmission, the different contagion paths and timing of crisis spillover effect in China (the country with a low degree of financial liberalization) and in the United States (the country with full financial liberalization) are first comparatively analyzed. In contrast with the mature markets of developed countries, the contagion periods of China's markets are relatively short, and considering imposed financial policies in China's special monetary regulating mode, the risk transmitting paths are meanwhile reshaped during the crisis. For both the United States and China, the stock market and the

real estate market are more likely to be the central risk-transmitters to spill risks over the interest rate market in turbulent episodes.

2. The proposed integrated early warning system offers advisory guidance to manage investment risks and provides assistant evidence to control market vulnerabilities. For China's stock market, the proposed early warning system output warning signals have provided a feasible, practical basis for the risk feasibility assessment by successfully reducing the risks of securities portfolio being exposed to the stock market turbulent period. Meanwhile, the early warning system based on attention mechanism plays an essentially supervisory role in preventing credit risks in China's bond market. According to the attention inspected results on leading indicators for China's bond market turbulence, the key risk prevention channels for the government bond and corporate bonds with high credit rating levels should focus on the discount rate economic factors (of GDP, CPI and PPI). In contrast, the supervision on lower credit-rated corporate bonds is prone to through the cash flow factors (of industrial gain, fixed asset investment).

3. This is the first time to reasonably quantify the financial risk transmission as a co-factor of the early warning system to investigate its preemptive effects on producing crisis signals for predicting China's sovereign bond turmoils. From the results of leading factors detection by attention mechanism, the external impact factors of the risk transmitting effects from the international crude oil market and the global gold price dynamics are pivotal regulatory objects to forewarn the nation's exposure to sovereign crises. Furthermore, the domestic economic factors of real-estate price and inflation level are the notable channels of the local government monitoring the underlying risks of sovereign bond anomalous behaviors. The government published information relating to fiscal policy adjustment should not be ignored since it always provides prompt tips to hint at the abnormal fluctuations in sovereign bonds[2].

[2]There is a significant information asymmetry between the government departments acquired and market participants held, generally speaking, the government's policy adjustment is executed in advance once the risks are sensed after collecting all accessible information from financial sectors, while the market participants' reaction always lags behind the government

In brief, rooted in China's specific markets, this study mainly explores and innovates the models for 1) the critical problem of crisis outbreaks, 2) the risk transmission periods and paths between markets during the crisis, and 3) the timely warning signals generation for market turbulence. The developed models are implemented by covering sufficient samples in long enough time intervals and ensuring the timeliness of research on a daily data basis. The results provide practical significance for risk transmission path monitoring and crisis burst blocking.

7.3 Limitations and further works

To clarify the research obstacles that we currently face in the progress, study limitations will be divided into two aspects: topic extensibility and technical challenges.

7.3.1 Possible extensions

On one side, our study is progressed and concluded for China's financial markets, which results will embody the specificity of China's economic surroundings − the limited openness of the financial markets. However, since April 1, 2020, China officially announced the full opening-up to the outside, the research results remain to be verified in generalizing the newly opened financial surroundings. As a matter of fact, both the financial crisis outbreaks and the risk transmission mechanism are various for different countries during different eras, which is related to the degree of financial liberalization and affected by the economic structure steadiness. For example, Thailand realized financial liberalization in 1989. During the subsequent development, the rapidly accumulated market bubbles and the strong attacks from international speculative capitals finally led to the Thai baht's abandonment of the fixed exchange rate in July 1997. After experiencing a sharp depreciation of over 30% on currency and a series of bubbles crushing, the fragile financial system collapsed, a total wipe-out that was unexpected. However, China has declared the opening up, yet free the exchange rate

department.

between the mainland currency (Renminbi) and other foreign currencies. Thus, it is hard to judge whether the open financial market leads to systemic risks if the Renminbi becomes an international currency in the future and whether the risk transmission will start from bubble-rich markets, such as real estate and the stock. These questions need to be testified in the further works of 1) updating the latest data for China and 2) applying the proposed models to updated data for China in comparison with other Asian emerging countries to reanalyze the commonalities and dissimilarities.

On the other side, in our study, the risk transmission has been discussed to assist the early warning system in making predictions. In practice, the effectiveness of the early warning system and the availability of its output signals will, in turn, affect the crisis transmission. We will explain the reversed influence in a catechetical way.

Question 1: How does the effectiveness of EWS affect crisis transmission?

Signal deliveries and response actions can influence the effectiveness of EWS. The warning signal delivering effectiveness has been verified in this paper by counting the confusion table and calculating the accuracy, SAR, and other statistical metrics. However, in practice, it is full of uncertainties in taking actions to prevent the crisis: on the one hand, it is uncertain that whether the EWS produced warning signals can be trusted and adopted; on the other hand, it is uncertain that whether the financial crisis will be effectively blocked as authorities taking precautionary measures. Given no information gap, after getting crisis warning signals, the governors will take preventive actions to alleviate the crisis, while the investors are more likely to have a series of panic behaviors to drive the crisis to burst reversely. In the combat between two forces, which side performs stronger, the pressure of the financial crisis will be more likely inclined to that party. Especially, as the investors' panics come from their distrust of the government power to prevent crises, the pressure on crisis transmission will sharply rise and eventually impact other markets and regions through these accumulated uncertainties.

Question 2: How does the availability of EWS output signal affect

crisis transmission?

In this paper, the availability of EWS output signals is assumed to be free-accessed for both governors and all market participants. In practice, there exists information asymmetry between the government and market participants and investors from different markets. If the government first obtains effective crisis signals and then takes risk intervention measures, the financial crisis thus may be successfully managed in such imposed prevention as long as the investors do not take any actions (such as adjusting portfolios, flight-to-quality, etc.) to offset the government's efforts. However, like what we analyzed in *Question 1*, investors usually behave in panics to make the situation counter-run to the crisis burst and spill over.

Therefore, in subsequent studies, we will further explore how to upgrade the EWS real-time forecasting capability by including the investors' behavioral uncertainty impacted financial contagion augmentation or mitigation. The conceptual model is diagrammed in Figure 7.1 given that 1) the EWS produced warnings will be freely available for all investors and 2) the government will intervene in the financial contagion as receiving the forewarned signals. However, achieving such full progress will bring new challenges (that are labeled as ①, ② and ③ in the diagram), which will be further discussed in the following section of 7.3.2.

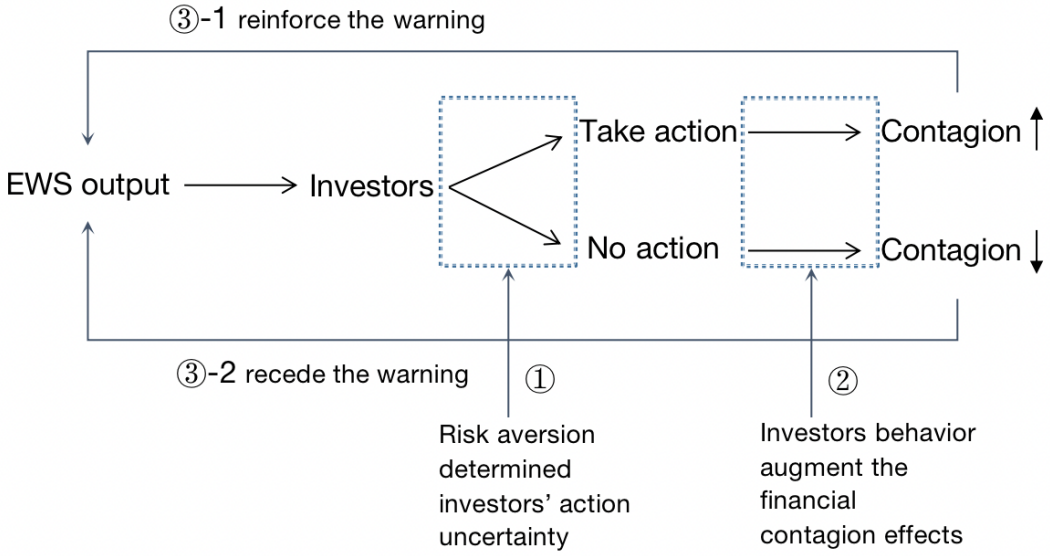


Figure 7.1: The backflow mechanism to boost the EWS real-time forecasting power.

7.3.2 Technical challenges

In the technical approach aspect, dichotomous topics will be discussed relating to 1) the higher requirement to improve the proposed contagion fused EWS model and 2) technical challenges for the extended projects mentioned above.

Dynamically structural contagion information input

In current progress for the proposed contagion fused EWS frameworks, the DCC-GARCH model and SWARCH model inferences are combined to quantize the dynamic contagion information input given the risk transmitting direction. To make the contagion quantification adaptive to the general case as either the transmitting direction changes or the potential linkage across multiple market assets varies during the observed periods, the further improvements thus will mainly focus on tracking the multivariate time-varying structural linkages without predetermining the risk transmission direction.

In the progress of CM development (See Chapter 4), the static copula models' robustness of inferring the risk transmission paths across multiple financial markets has been realized in terms of searching for the central transmitters. The dynamic copula models are thought to be used in further study to make the time-varying structural dependence simultaneously estimated in the copula frameworks. The model is proposed and mainly clustered into two types[3] of parametric specified time series model combined time-varying conditional copulas (Dias et al., 2004; Patton, 2006), and the non-parametric inferred dependence time-varying copulas (Hafner and Reznikova, 2010; Vatter and Nagler, 2018). Patton (2013) and Manner and Reznikova (2012) have comprehensively surveyed the multivariate time-varying copula-based models and made comparisons therein. In modern practice, the dynamic copulas are also successfully exploited for high-frequency data applications (De Lira Salvatierra and Patton, 2015; Acar et al., 2019). This method, therefore, provides a plausible research basis for our further explorations.

[3]There are other types of dynamic copula models but will not be included herein. For example the Markov switching model combined copulas (Stöber and Czado, 2014; Fink et al., 2016), which though implement the time-varying dependence parameter estimations, the copulas are time-invariant.

Behavioral uncertainty driven contagion backflow to the EWS

To extend the contagion backflow fed EWS driven by investor behavioral uncertainties, we mainly have three layers of technical challenges (which have been labeled in Figure 7.1): ① how to evaluate the behavioral uncertainty according to investors' varied risk preferential tastes for different markets; ② how to estimate the investors' behavior influence on contagious effects; and ③ how to adjust the warning signal strength according to the dynamically changed contagion effects.

For the first challenge, it is not easy to quantify the investment behavioral uncertainty by functionally mapping to the investors' risk aversion level[4], especially as the investing attitudes and sentiments are varied for different market participants. Based on the scanty literature, the studies relating to the risk aversion and investment uncertainty relationship may provide solving ideas heuristic. Chronopoulos et al. (2011) confirm that the risk aversion lowers the probability of investment and increases the likelihood to abandon the project. Particularly, during the worst time in crisis, as the investor risk tolerance decreases (i.e. risk aversion gains), substantial swings in trading and risk-taking behavior will be driven by such risk perception changes (Hoffmann et al., 2013). The recent study of González-Sánchez et al. (2020) validates the claim that risk aversion amplifies the effects of uncertainty on real activity exposure. Thus, intuitively, the greater the risk aversion level, the more uncertainty of investor behaviors, which implies the investor taking action can be directly proportional to the risk aversion level. Such proportional relation will be augmented once the crisis warning signals are obtained.

To solve the second technical problems, behavioral finance studies should be referred to relating to changing the market price volatile status and gaining the market correlation during the financial crisis. In the perspective of market volatility, the irrationality of investor's behavior and information asymmetry have been presented to bring significant spikes in returns and volume and gain the correlation of market returns in the highly volatile, theoretically (Gabaix et al., 2006) and practically (WATANABE, 2008). The most recent work of Bello (2019)

[4]The risk-preferential tastes though generally include three types of risk-seeking, risk-neutral and risk-averse, the risk aversion is the most discussed situation for most inefficient markets in practice.

proves that the investors' herding behavior is the main drive for contagious effects during financial crisis by examining contagion effect on 10 African financial markets. Moreover, referring to the modeling techniques, markets' contagious effect has been analyzed in terms of modeling the investor behavior in the epidemic model (Shive, 2010) and the epidemic model based information networks (Wang et al., 2019) considering the similarity between the infection of disease and the investors' interaction of spreading information. Wu (2020) designs an artificial information based model to simulate the relationship between investor behavior and risk contagion, which finds the contagion coefficient of investors affected by neighbors' sentiment plays an important role in risk contagion as new information being injected. We can attempt either the epidemic model based topological networks or the information based artificial models to quantize the contagious effect changes on the market volatility basis as investors take irrational actions to produce asymmetric information spreading.

The last technical difficulty is converting the retrieved investors' behavior-dependent contagion information into the importable input for EWS to influence the produced warning signal strength for the future time point. We have not found any available literature to provide solvable ideas to tackle such a problem, thus altering to list the challenge points here for open discussion. On the one hand, the time lag between the contagion dynamics is updating and the EWS signals generating. For example, considering the simplistic case as time lag equals to 1, the EWS produced signal for time t will affect the investor's behavior for $t+1$. The contagion pattern will be accordingly changed at time $t+1$, which means the contagion input is required to be timely updated and recursively accessed to the EWS input layer for $t+1$. However, in practice, such an ongoing information updating procedure is hardly implemented for the short time lag. On the other hand, accessing the EWS output is indeterminate between directly lowering or gaining the EWS predicted crisis probability or indirectly being contagion input to make the EWS reproduce warning signals. The former way is more intuitive to prevent the severe risks from spreading by reinforcing signals to alert for crisis danger. However, the latter one is not deterministic since updated contagion input will contribute to the reproduced warning signal in a dichotomous way.

7.4 Enlightment and prospect

We hope to bring an enlightening guide to the financial supervisory division to reinforce financial security and emphasize technical innovation through the study.

Domestically, with the rapid development of non-banking financial institutions and gradual enrichment of financial derivatives, the price volatility led risk transmission channels across markets show a trend of diversification. Thus, it puts forward higher requirements for China's financial regulatory authorities to establish a systematic monitoring scheme for implementing overall supervision on entire financial surroundings and refine the regulatory functionalities according to the specific operating modes for different financial industries. In addition, the financial supervision system should improve the information-sharing mechanism and extend the compatibility of data interaction between different platforms to boost monitoring and regulating efficiency.

In the global view, as China's financial markets have announced to fully open, the volatile risk transmission will gradually turn from the endogenous mechanism to the expected impacts of both endogenous factors and external shocks. Therefore, the markets' exposure to financial crises will currently be regarded as the joint result of unstable domestic factors to fluctuated international factors (such as foreign currencies and trading commodities). The regulatory authorities should be aware of the financial security problems brought about by the increased market openness, and then put coping strategies (such as strengthening the supervision on international short-term capital flows and controlling the high volume of overseas investment in the domestic financial security market) to handle with the external-risk-sensitive situation in an early-adaptive way.

By flexibly using policy tools to prevent financial risks and actively attracting stable and high-quality funds from abroad, we believe that the balance between financial openness and risk prevention is attainable. It will ultimately serve the bilateral promoting pattern between the domestic and international cycles.

Bibliography

A, A. C. and B, H. Z. (2014). The determinants of sovereign default: A sensitivity analysis. International Review of Economics & Finance, 33(3):300–318.

A, A. L. D., B, M. G., and c, A. L.-V. (2012). Has the cds market influenced the borrowing cost of european countries during the sovereign crisis? Journal of International Money and Finance, 31(3):481–497.

A, J. D. H. and B, R. S. (1994). Autoregressive conditional heteroskedasticity and changes in regime. Journal of Econometrics, 64(1–2):307–333.

A, R. C. and B, F. R. (2020). The relative pricing of sovereign credit risk after the eurozone crisis. Journal of International Money and Finance.

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics, 44(2):182–198.

Abiad, A. D. (2003). Early warning systems: A survey and a regime-switching approach. IMF Working Papers, 03(32):993–1052.

Acar, E. F., Czado, C., and Lysy, M. (2019). Flexible dynamic vine copula models for multivariate time series data. Econometrics and Statistics, 12:181–197.

Afonso, A., Arghyrou, M. G., Gadea, M. D., and Kontonikas, A. (2018). "whatever it takes" to resolve the european sovereign debt crisis? bond pricing regime switches and monetary policy effects. Journal of International Money and Finance, 86:1–30.

Agyei-Ampomah, S., Gounopoulos, D., and Mazouz, K. (2014). Does gold offer a better protection against losses in sovereign debt bonds than other metals? Journal of Banking & Finance, 40:507–521.

Ahmad, W., Sehgal, S., and Bhanumurthy, N. R. (2013). Eurozone crisis and bricks stock markets: Contagion or market interdependence? Economic Modelling, 33(jul.):209–225.

Ahmed, K., El-Alfy, E.-S., and Mohammed, S. (2018). Evaluation of bidirectional lstm for short-and long-term stock market prediction. pages 151–156.

Ahn, J. J., Oh, K. J., Kim, T. Y., and Kim, D. H. (2011). Usefulness of support vector machine to develop an early warning system for financial crisis. Expert Systems with Applications, 38(4):2966–2973.

Akhtaruzzaman, M., Abdel-Qader, W., Hammami, H., and Shams, S. (2021). Is china a source of financial contagion? Finance Research Letters, 38:101393.

Allen, D. E., Singh, A. K., and Powell, R. J. (2013). Evt and tail-risk modelling: Evidence from market indices and volatility series. The North American Journal of Economics and Finance, 26:355 – 369.

Almeida, C. and Czado, C. (2012). Efficient bayesian inference for stochastic time-varying copula models. Computational Statistics and Data Analysis, 56(6):1511–1527.

Aloui, R. and Ben Aïssa, M. S. (2016). Relationship between oil, stock prices and exchange rates: A vine copula based garch method. The North American Journal of Economics and Finance, 37:458–471.

Alter, A. and Beyer, A. (2014). The dynamics of spillover effects during the european sovereign debt turmoil. Journal of Banking & Finance, 42:134–153.

Althelaya, K. A., El-Alfy, E.-S. M., and Mohammed, S. (2018). Evaluation of bidirectional lstm for short-and long-term stock market prediction. In 2018 9th International Conference on Information and Communication Systems (ICICS), pages 151–156.

Ang, A. and Bekaert, G. (2002). Regime switches in interest rates. Journal of Business & Economic Statistics, 20(2):163–182.

Ardia, D., Bluteau, K., Boudt, K., Catania, L., and Trottier, D.-A. (2017). Markov-Switching GARCH Models in R: The MSGARCH package.

Arghyrou, M. G. and Kontonikas, A. (2012). The emu sovereign-debt crisis: Fundamentals, expectations and contagion. Journal of International Financial Markets, Institutions and Money, 22(4):658–677.

Arthur, W. B. (1999). Complexity and the economy. Science, 284(5411):107–109.

Arévalo, R., García, J., Guijarro, F., and Peris, A. (2017). A dynamic trading rule based on filtered flag pattern recognition for stock market price forecasting. Expert Systems with Applications, 81:177 – 192.

Babecký, J., Havránek, T., Matějů, J., Rusnák, M., Šmídková, K., and Vašiček, B. (2014). Banking, debt, and currency crises in developed countries: Stylized facts and early warning indicators. Journal of Financial Stability, 15:1–17.

Bahdanau, D., Cho, K., and Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.

Beckmann, R. (2007). Profitability of Western European banking systems: panel evidence on structural and cyclical determinants. Discussion Paper Series 2: Banking and Financial Studies 2007,17, Deutsche Bundesbank.

Bedford, T. and Cooke, R. M. (2002). Vines–a new graphical model for dependent random variables. Ann. Statist., 30(4):1031–1068.

Bekaert, G., Harvey, C., and Ng, A. (2005). Market integration and contagion. The Journal of Business, 78(1):39–70.

Bello, J. (2019). Financial contagion effect and investor behavior in african financial markets during the 2007–09 global financial crisis. Available at SSRN: https://dx.doi.org/10.2139/ssrn.3442636.

Beltratti, A. and Stulz, R. M. (2019). Why is contagion asymmetric during the european sovereign crisis? Journal of International Money and Finance, 99:102081.

BenMim, I. and BenSaïda, A. (2019). Financial contagion across major stock markets: A study during crisis episodes. The North American Journal of Economics and Finance, 48:187–201.

BenSaïda, A. (2018). The contagion effect in european sovereign debt markets: A regime-switching vine copula approach. International Review of Financial Analysis, 58:153–165.

Beqiraj, E., Patella, V., and Tancioni, M. (2021). Fiscal stance and the sovereign risk pass-through. Economic Modelling, page 105573.

Berg, A. and Pattillo, C. (1999). Predicting currency crises: The indicators approach and an alternative. Journal of international Money and Finance, 18(4):561–586.

Beutel, J., List, S., and [von Schweinitz], G. (2019). Does machine learning help us predict banking crises? Journal of Financial Stability, 45:100693.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized arch model. The Review of Economics and Statistics, 72(3):498–505.

Bonga-Bonga, L. and Mabe, Q. M. (2020). How financially integrated are trading blocs in africa? The Quarterly Review of Economics and Finance, 75:84–94.

Bostanci, G. and Yilmaz, K. (2020). How connected is the global sovereign credit risk network? Journal of Banking & Finance, 113:105761.

Bouri, E., Jalkh, N., and Roubaud, D. (2019). Commodity volatility shocks and bric sovereign risk: A garch-quantile approach. Resources Policy, 61:385–392.

Bouri, E., Shahzad, S. J. H., Raza, N., and Roubaud, D. (2018). Oil volatility and sovereign risk of brics. Energy Economics, 70:258–269.

Bussiere, M. and Fratzscher, M. (2006). Towards a new early warning system of financial crises. journal of International Money and Finance, 25(6):953–973.

Candelon, B., Dumitrescu, E.-I., and Hurlin, C. (2014). Currency crisis early warning systems: Why they should be dynamic. International Journal of Forecasting, 30(4):1016–1029.

Cao, J., Li, Z., and Li, J. (2019). Financial time series forecasting model based on ceemdan and lstm. Physica A: Statistical Mechanics and its Applications, 519:127–139.

Caporin, M., Gupta, R., Ravazzolo, F., et al. (2019). Contagion between real estate and financial markets: A bayesian quantile-on-quantile approach. Technical report.

Caprio, G. and Klingebiel, D. (2002). Episodes of systemic and borderline banking crises. Managing the real and fiscal effects of banking crises, World Bank Discussion Paper, 428:31–49.

Caruana, R. and Niculescu-Mizil, A. (2004). Data mining in metric space: an empirical analysis of supervised learning performance criteria. In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pages 69–78. ACM.

Celik, A. E. and Karatepe, Y. (2007). Evaluating and forecasting banking crises through neural network models: An application for turkish banking sector. Expert Systems with Applications, 33(4):809 – 815.

Celik, S. (2012). The more contagion effect on emerging markets: The evidence of dcc-garch model. Economic Modelling, 29(5):1946–1959.

Chamon, M., Manasse, P., and Prati, A. (2007). Can We Predict the Next Capital Account Crisis? IMF Staff Papers, 54(2):270–305.

Chan, K. F., Treepongkaruna, S., Brooks, R., and Gray, S. (2011). Asset market linkages: Evidence from financial, commodity and real estate assets. Journal of Banking & Finance, 35(6):1415–1426.

Chatterjee, S. and Eyigungor, B. (2019). Endogenous political turnover and fluctuations in sovereign default risk. Journal of International Economics, 117:37–50.

Chatzis, S. P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., and Vlachogiannakis, N. (2018). Forecasting stock market crisis events using deep and statistical machine learning techniques. Expert Systems with Applications, 112:353–371.

Chen, D. R. (2017). Sovereign debt default and bank risk in China. PhD thesis, Lincoln University.

Chen, K., Zhou, Y., and Dai, F. (2015). A lstm-based method for stock returns prediction: A case study of china stock market. In 2015 IEEE International Conference on Big Data (Big Data), pages 2823–2824. IEEE.

Chen, Y., Zheng, B., and Qu, F. (2020). Modeling the nexus of crude oil, new energy and rare earth in china: An asymmetric var-bekk (dcc)-garch approach. Resources Policy, 65:101545.

Chong, E., Han, C., and Park, F. C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. Expert Systems with Applications, 83:187–205.

Christiansen, C., Schmeling, M., and Schrimpf, A. (2012). A comprehensive look at financial volatility prediction by economic variables. Journal of Applied Econometrics, 27(6):956–977.

Chronopoulos, M., De Reyck, B., and Siddiqui, A. (2011). Optimal investment under operational flexibility, risk aversion, and uncertainty. European Journal of Operational Research, 213(1):221–237.

Chuffart, T. and Hooper, E. (2019). An investigation of oil prices impact on sovereign credit default swaps in russia and venezuela. Energy Economics, 80:904–916.

Ciarlone, A. and Trebeschi, G. (2005). Designing an early warning system for debt crises. Emerging Markets Review, 6(4):376–395. Assessing Risk in Emerging Markets.

Correa, Ricardo, Sapriza, and Horacio (2014). Sovereign debt crises. International Finance Discussion Papers, SSRN, 78(2):192–205.

Coudert, V. and Gex, M. (2008). Does risk aversion drive financial crises? testing the predictive power of empirical indicators. Journal of Empirical Finance, 15(2):167–184.

Cronin, D., Flavin, T. J., and Sheenan, L. (2016). Contagion in eurozone sovereign bond markets? the good, the bad and the ugly. Economics Letters, 143:5–8.

Cubillos-Rocha, J. S., Gomez-Gonzalez, J. E., and Melo-Velandia, L. F. (2019). Detecting exchange rate contagion using copula functions. The North American Journal of Economics and Finance, 47:13 – 22.

Daude, C. (2012). Sovereign default risk and volatility. Economics Letters, 114(1):47–50.

Davis, E. P. and Karim, D. (2008). Comparing early warning systems for banking crises. Journal of Financial Stability, 4(2):89 – 120.

Dawood, M., Horsewood, N., and Strobel, F. (2017). Predicting sovereign debt crises: An early warning system approach. Journal of Financial Stability, 28:16–28.

De Lira Salvatierra, I. and Patton, A. J. (2015). Dynamic copula models and high frequency data. Journal of Empirical Finance, 30:120–135.

Deev, O. and Hodula, M. (2016). Sovereign default risk and state-owned bank fragility in emerging markets: evidence from china and russia. Post-Communist Economies, 28(2):232–248.

Demirgüç-Kunt, A. and Detragiache, E. (1998). The determinants of banking crises in developing and developed countries. Staff Papers (International Monetary Fund), 45(1):81–109.

Detragiache, M. E. and Spilimbergo, M. A. (2001). Crises and liquidity: evidence and interpretation. Number 1-2. International Monetary Fund.

Dias, A., Embrechts, P., et al. (2004). Dynamic copula models for multivariate high-frequency data in finance.

Dißmann, J., Brechmann, E., Czado, C., and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics & Data Analysis, 59:52 – 69.

Dongho, S. (2017). Bond market exposures to macroeconomic and monetary policy risks. Review of Financial Studies.

Duan, P. and Bajona, C. (2008). China's vulnerability to currency crisis: A klr signals approach. China Economic Review, 19(2):138–151.

Dungey, M., Fry, R., González-Hermosillo, B., and Martin, V. L. (2005). Empirical Modelling of Contagion: A Review of Methodologies. Quantitative Finance, 5(1):9–24.

Dungey, M. and Martin, V. (2001). Contagion across financial markets: An empirical assessment. pages 16–17.

Dungey, M. and Martin, V. L. (2004). A Multifactor Model of Exchange Rates with Unanticipated Shocks: Measuring Contagion in the East Asian Currency Crisis. Journal of Emerging Market Finance, 3(3):305–330.

Edison, H. J. (2003). Do indicators of financial crises work? an evaluation of an early warning system. International Journal of Finance & Economics, 8(1):11–53.

Edwards, S. and Susmel, R. (2001). Volatility Dependence and Contagion in Emerging Equity Markets. Journal of Development Economics, 66(2):505–532.

Eichengreen, B. and Rose, A. K. (1998). Staying afloat when the wind shifts: External factors and emerging-market banking crises. Working Paper 6370, National Bureau of Economic Research.

Eichengreen, B., Rose, A. K., and Wyplosz, C. (1995). Exchange market mayhem: the antecedents and aftermath of speculative attacks. Economic policy, 10(21):249–312.

Engle, R. F. (2003). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. Journal of Business & Economic Statistics, 20(3):339–350.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. Econometric Theory, 11(1):122–150.

Feng, Q., Sun, X., Liu, C., and Li, J. (2021). Spillovers between sovereign cds and exchange rate markets: The role of market fear. The North American Journal of Economics and Finance, 55:101308.

Fernández-Rodríguez, F., Gómez-Puig, M., and Sosvilla-Rivero, S. (2015). Volatility spillovers in emu sovereign bond markets. International Review of Economics & Finance, 39:337–352.

Filippopoulou, C., Galariotis, E., and Spyrou, S. (2020). An early warning system for predicting systemic banking crises in the eurozone: A logit regression approach. Journal of Economic Behavior & Organization, 172:344–363.

Fink, H., Klimova, Y., Czado, C., and Stober, J. (2016). Regime switching vine copula models for global equity and volatility indices. Econometrics, 5.

Fioramanti, M. (2008). Predicting sovereign debt crises using artificial neural networks: A comparative approach. Journal of Financial Stability, 4(2):149–164.

Fischer, T. and Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2):654–669.

Fleming, J. (1998). The quality of market volatility forecasts implied by s&p 100 index option prices. Journal of Empirical Finance, 5(4):317–345.

Forbes, K. J. and Rigobon, R. (2002). No Contagion, Only Interdependence: Measuring Stock Market Comovements. The Journal of Finance, 57(5):2223–2261.

Frankel, J. and Rose, A. (1996). Currency crashes in emerging markets: An empirical treatment. Journal of International Economics, 41(3-4):351–366.

Fu, J., Zhou, Q., Liu, Y., and Wu, X. (2019). Predicting stock market crises using daily stock market valuation and investor sentiment indicators. The North American Journal of Economics and Finance.

Fuertes, A.-M. and Kalotychou, E. (2006). Early warning systems for sovereign debt crises: The role of heterogeneity. Computational Statistics & Data Analysis, 51(2):1420–1441.

Fuertes, A.-M. and Kalotychou, E. (2007). Optimal design of early warning systems for sovereign debt crises. International Journal of Forecasting, 23(1):85–100.

Gabaix, X., Gopikrishnan, P., Plerou, V., and Stanley, H. E. (2006). Institutional investors and stock market volatility*. The Quarterly Journal of Economics, 121(2):461–504.

Genberg, H. and Sulstarova, A. (2008). Macroeconomic volatility, debt dynamics, and sovereign interest rate spreads. Journal of International Money and Finance, 27(1):26–39.

Gennaioli, N., Martin, A., and Rossi, S. (2014). Banks, government bonds, and default; what do the data say? IMF Working Papers.

Georgoutsos, D. and Moratis, G. (2017). Bank-sovereign contagion in the eurozone: A panel var approach. Journal of International Financial Markets, Institutions and Money, 48:146–159.

Gerali, A., Locarno, A., Notarpietro, A., and Pisani, M. (2017). The sovereign crisis and italy's potential output. Journal of Policy Modeling, 40(2).

Ghosh, A. R. and Basurto, G. (2006). The interest rate-exchange rate nexus in the asian crisis countries. Imf Working Papers, 00(19).

Ghulam, Y. and Derber, J. (2018). Determinants of sovereign defaults. The Quarterly Review of Economics and Finance, 69:43–55.

Giovanis, E. (2012). Study of discrete choice models and adaptive neuro-fuzzy inference system in the prediction of economic crisis periods in usa. Economic Analysis and Policy, 42(1):79–96.

Gomez-Gonzalez, J. E. and Rojas-Espinosa, W. (2019). Detecting contagion in asian exchange rate markets using asymmetric dcc-garch and r-vine copulas. Economic Systems, 43(3):100717.

González-Sánchez, M., Nave, J., and Rubio, G. (2020). Effects of uncertainty and risk aversion on the exposure of investment-style factor returns to real activity. Research in International Business and Finance, 53:101236.

Granger, C. W. J. and Poon, S. H. (2003). Forecasting volatility in financial markets: A review (revised edition). Journal of Economic Literature, 41(2):478–539.

Hafner, C. M. and Reznikova, O. (2010). Efficient estimation of a semiparametric dynamic copula model. Computational Statistics & Data Analysis, 54(11):2609–2627. The Fifth Special Issue on Computational Econometrics.

Hamill, P. A., Li, Y., Pantelous, A. A., Vigne, S. A., and Waterworth, J. (2021). Was a deterioration in 'connectedness' a leading indicator of the european sovereign debt crisis? Journal of International Financial Markets, Institutions and Money, page 101300.

Hamilton, J. and Gang, L. (1996). Stock Market Volatility and the Business Cycle. Journal of Applied Econometrics, 11(5):573–93.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica, 57(2):357–384.

Hamilton, J. D. (1994). Time Series Analysis. Priceton University Press.

Hamilton, J. D. and Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. Journal of Econometrics, 64(1):307 – 333.

Harvey, C. R. and Whaley, R. E. (1992). Market volatility prediction and the efficiency of the s&p 100 index option market. Journal of Financial Economics, 31(1):43–73.

Hergott, M. J. (2018). How an lstm attention model views the 2013 bond market 'taper tantrum'. https://hergott.github.io/lstm-attention-bond-market-taper-tantrum/. Published on Jun. 23, 2018.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. Neural Comput., 9(8):1735–1780.

Hoesli, M. and Reka, K. (2015). Contagion channels between real estate and financial markets. Real Estate Economics, 43(1):101–138.

Hoffmann, A. O., Post, T., and Pennings, J. M. (2013). Individual investor perceptions and behavior during the financial crisis. Journal of Banking & Finance, 37(1):60–74.

Holopainen, M. and Sarlin, P. (2017). Toward robust early-warning models: a horse race, ensembles and model uncertainty. Quantitative Finance, 17(12):1933–1963.

Hou, Y. and Li, S. (2016). Information transmission between u.s. and china index futures markets: An asymmetric dcc garch approach. Economic Modelling, 52:884–897.

Hou, Y., Li, S., and Wen, F. (2019). Time-varying volatility spillover between chinese fuel oil and stock index futures markets based on a dcc-garch model with a semi-nonparametric approach. Energy Economics, 83:119–143.

Hu, Y., Ni, J., and Wen, L. (2020). A hybrid deep learning approach by integrating lstm-ann networks with garch model for copper price volatility prediction. Physica A: Statistical Mechanics and its Applications, 557:124907.

Huang, R. and Ali, S. (2012). Governing financial disputes in china: What have we learned from the global financial crisis of 2008? Social ence Electronic Publishing, 7(1):195.

Hui, E. C.-m. and Chan, K. K. K. (2014). The global financial crisis: Is there any contagion between real estate and equity markets? Physica A: Statistical Mechanics and its Applications, 405:216–225.

Ibhagui, O. (2021). How do sovereign risk, equity and foreign exchange derivatives markets interact? Economic Modelling, 97:58–78.

Jain, R., Kasturi, R., and Schunck, B. G. (1995). Machine vision, volume 5. McGraw-Hill New York.

Jiang, W. (2021). Applications of deep learning in stock market prediction: Recent progress. Expert Systems with Applications, 184:115537.

Jondeau, E. and Rockinger, M. (2006). The copula-garch model of conditional dependencies: An international stock market application. Journal of International Money and Finance, 25(5):827–853.

Jones, P. M. and Olson, E. (2013). The time-varying correlation between uncertainty, output, and inflation: Evidence from a dcc-garch model. Economics Letters, 118(1):33–37.

Jordan, M. (1997). Serial order: A parallel distributed processing approach. Advances in Psychology, 121:471–495.

Julianne Ams, Reza Baqir, A. G. and Trebesch, C. (2018). Chapter7 - sovereign debt: A guide for economists and practitioners. In IMF Conference, Washington. DC, U.S.

Kalbaska, A. and Gatkowski, M. (2012). Eurozone sovereign contagion: Evidence from the cds market (2005–2010). Journal of Economic Behavior & Organization, 83(3):657–673. The Great Recession: motivation for re-thinking paradigms in macroeconomic modeling.

Kaminsky, G., Lizondo, S., and Reinhart, C. (1998). Leading indicators of currency crises. IMF Staff Papers, 45(1):1–48.

Kaminsky, G. L. (1998). Currency and banking crises: the early warnings of distress. International Finance Discussion Papers 629, Board of Governors of the Federal Reserve System (U.S.).

Kaminsky, G. L. (2006). Currency crises: Are they all the same? Journal of International Money and Finance, 25(3):503–527.

Kaminsky, G. L. and Reinhart, C. M. (1999). The twin crises: The causes of banking and balance-of-payments problems. The American Economic Review, 89(3):473–500.

Keddad, B. and Schalck, C. (2020). Evaluating sovereign risk spillovers on domestic banks during the european debt crisis. Economic Modelling, 88:356–375.

Kibritcioglu, B., Kose, B., and Ugur, G. (1999). A leading indicators approach to the predictability of currency crises: the case of turkey. Hazine Dergisi, Sayi Working Paper, (1998/12).

Kim, C.-J. (1994). Dynamic linear models with markov-switching. Journal of Econometrics, 60(1-2):1–22.

Kim, H. H. and Swanson, N. R. (2014). Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence. Journal of Econometrics, 178:352 – 367. Recent Advances in Time Series Econometrics.

Kim, H. Y. and Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating lstm with multiple garch-type models. Expert Systems with Applications, 103:25–37.

Kim, K. (2013). Modeling financial crisis period: A volatility perspective of credit default swap market. Physica A: Statistical Mechanics and its Applications, 392(20):4977–4988.

Kim, S.-J., Salem, L., and Wu, E. (2015). The role of macroeconomic news in sovereign cds markets: Domestic and spillover news effects from the u.s., the eurozone and china. Journal of Financial Stability, 18:208–224.

Kim, T. Y., Hwang, C., and Lee, J. (2004a). Korean economic condition indicator using a neural network trained on the 1997 crisis. Journal of Data Science, 2(4):371–381.

Kim, T. Y., Oh, K. J., Sohn, I., and Hwang, C. (2004b). Usefulness of artificial neural networks for early warning system of economic crisis. Expert Systems with Applications, 26(4):583–590.

Koyuncugil, A. S. and Ozgulbas, N. (2012). Financial early warning system model and data mining application for risk detection. Expert Systems with Applications, 39(6):6238–6253.

Kraussl, R., Lehnert, T., and Stefanova, D. (2016). The european sovereign debt crisis: What have we learned? Journal of Empirical Finance, 38:363–373.

Kuan, C.-M. (2002). Lecture on the markov switching model.

Laeven, L. and Valencia, F. (2012). Systemic banking crises database: An update.

Laeven, M. L. and Valencia, F. (2010). Resolution of banking crises: The good, the bad, and the ugly. Number 10-146. International Monetary Fund.

Lestano, J. P. A. M. (2007). Dating currency crises with ad hoc and extreme value-based thresholds: East asia 1970-2002 [dating currency crises];. International Journal of Finance & Economics, 12(4):371–388.

Lestano, L., Jacobs, J., and Kuper, G. (2004). Indicators of financial crises do work! an early-warning system for six asian countries. International finance, University Library of Munich, Germany.

Li, W.-X., Chen, C. C.-S., and French, J. J. (2015). Toward an early warning system of financial crises: What can index futures and options tell us? The Quarterly Review of Economics and Finance, 55:87–99.

Lina, El-Jahel, Lara, Cathcart, Saad, and Badaoui (2013). Do sovereign credit default swaps represent a clean measure of sovereign default risk? a factor model approach. Journal of banking & finance, 37(7):2392–2407.

Liu, J., Chen, Y., Liu, K., and Zhao, J. (2017). Attention-based event relevance model for stock price movement prediction. In Knowledge Graph and Semantic Computing. Language, Knowledge, and Intelligence, pages 37–49, Singapore. Springer Singapore.

Liu, Q., Cheng, X., Su, S., and Zhu, S. (2018). Hierarchical complementary attention network for predicting stock price movements with news. pages 1603–1606.

Liu, Y. (2019). Novel volatility forecasting using deep learning–long short term memory recurrent neural networks. Expert Systems with Applications, 132:99–109.

Liuren, W. U., ZHANG, and Xiaoling, F. (2008). A no-arbitrage analysis of macroeconomic determinants of the credit spread term structure. Management Science, 54(6):1160–1175.

Long, C. and Zhao, X. (2009). Return decomposition. Review of Financial Studies, 22(12):5213–5249.

Long, W., Lu, Z., and Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. Knowledge-Based Systems, 164:163–173.

Malkiel, B. G. and Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work*. The Journal of Finance, 25(2):383–417.

Manner, H. and Reznikova, O. (2012). A survey on time-varying copulas: Specification, simulations, and application. Econometric Reviews, 31(6):654–687.

McNeil, A. J. and Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance, 7(3-4):271–300.

Meier, S., Rodriguez Gonzalez, M., and Kunze, F. (2021). The global financial crisis, the emu sovereign debt crisis and international financial regulation: lessons from a systematic literature review. International Review of Law and Economics, 65:105945.

Metz, C. E. (1978). Basic principles of roc analysis. Seminars in Nuclear Medicine, 8(4):283 – 298.

Mody, A. and Sandri, D. (2012). The eurozone crisis: how banks and sovereigns came to be joined at the hip. Economic Policy.

Nag, A. K. and Mitra, A. (1999). Neural networks and early warning indicators of currency crisis. In: Reserve Bank of India occasional papers.

Nagayasu, J. (2001). Currency crisis and contagion: evidence from exchange rates and sectoral stock indices of the philippines and thailand. Journal of Asian Economics.

Nitschka and Thomas (2018). Bond market evidence of time variation in exposures to global risk factors and the role of us monetary policy. Journal of International Money & Finance.

Nowak, S., Andritzky, J., Jobst, A., and Tamirisa, N. (2011). Macroeconomic fundamentals, price discovery and volatility dynamics in emerging markets. Journal of Banking & Finance, 35(10):2584–2597.

Oh, K. J., Kim, T. Y., and Kim, C. (2006). An early warning system for detection of financial crisis using financial market volatility. Expert Systems, 23(2):83–98.

Ohtsu, N. (2007). A threshold selection method from gray-level histograms. IEEE Transactions on Systems Man & Cybernetics, 9(1):62–66.

Ouyang, Z., Yang, X., and Lai, Y. (2021). Systemic financial risk early warning of financial market in china using attention-lstm model. The North American Journal of Economics and Finance, 56:101383.

Patton, A. (2013). Chapter 16 - copula methods for forecasting multivariate time series. In Elliott, G. and Timmermann, A., editors, Handbook of Economic Forecasting, volume 2 of Handbook of Economic Forecasting, pages 899–960. Elsevier.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence*. International Economic Review, 47(2):527–556.

Peng, D. and Bajona, C. (2008). China's vulnerability to currency crisis: A klr signals approach. China Economic Review, 19(2):138 – 151.

Pericoli, M. and Sbracia, M. (2003). A primer on financial contagion. Journal of Economic Surveys, 17(4):571–608.

Pescatori, A. and Sy, A. N. R. (2007). Are debt crises adequately defined? IMF Staff Papers, 54(2):306–337.

Prewitt, J. and Mendelsohn, M. (1966). The analysis of cell images. Annals of the New York Academy of Sciences, 128:1035–53.

Raimbourg, P. and Salvadè, F. (2020). Rating announcements, cds spread and volatility during the european sovereign crisis. Finance Research Letters, page 101663.

Ramchand, L. and Susmel, R. (1998). Volatility and Cross Correlation across Major Stock Markets. Journal of Empirical Finance, 5(4):397–416.

Rand, W. M. (1971). Objective criteria for the evaluation of clustering methods. Journal of the American Statistical Association, 66(336):846–850.

Reinhart, Carmen, M., Rogoff, and Kenneth, S. (2011). From financial crash to debt crisis. American Economic Review.

Reinhart, C. M. and Rogoff, K. S. (2013). Banking crises: an equal opportunity menace. Journal of Banking & Finance, 37(11):4557–4573.

Ren, Y., Liao, F., and Gong, Y. (2020). Impact of news on the trend of stock price change: an analysis based on the deep bidirectiona lstm model. Procedia Computer Science, 174:128–140. 2019 International Conference on Identification, Information and Knowledge in the Internet of Things.

Reusens, P. and Croux, C. (2017). Sovereign credit rating determinants: a comparison before and after the european debt crisis. Journal of Banking & Finance, 77:108–121.

Rho, C. and Saenz, M. (2021). Financial stress and the probability of sovereign default. Journal of International Money and Finance, 110:102305.

Ribeiro, P. P., Cermeño, R., and Curto, J. D. (2017). Sovereign bond markets and financial volatility dynamics: Panel-garch evidence for six euro area countries. Finance Research Letters, 21:107–114.

Rodriguez, J. C. (2007). Measuring financial contagion: A copula approach. Journal of Empirical Finance, 4(3):401 – 423.

Roman, Krussl, Thorsten, Lehnert, Denitsa, and Stefanova (2016). The european sovereign debt crisis: What have we learned? Journal of Empirical Finance, 38(Part A):363–373.

Rosenfeld, A. and De La Torre, P. (1983). Histogram concavity analysis as an aid in threshold selection. IEEE Transactions on Systems, Man, and Cybernetics, (2):231–235.

Samitas, A., Kampouris, E., and Kenourgios, D. (2020). Machine learning as an early warning system to predict financial crisis. International Review of Financial Analysis, 71:101507.

Samuel, Y. M. Z.-t. (2008). Value at risk and conditional extreme value theory via markov regime switching models. Journal of Futures Markets, 28(2):155–181.

Savona, R. and Vezzoli, M. (2015). Fitting and forecasting sovereign defaults using multiple risk signals. Oxford Bulletin of Economics and Statistics, 77.

Schiller, R. (2000). Irrational Exuberance. Wiley Online Library.

Schimmelpfennig, A., Roubini, N., and Manasse, P. (2003). Predicting sovereign debt crises. IMF Working Papers, 03(221):192–205.

Schuster, M. and Paliwal, K. K. (1997). Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing, 45(11):2673–2681.

Sevim, C., Oztekin, A., Bali, O., Gumus, S., and Guresen, E. (2014). Developing an early warning system to predict currency crises. European Journal of Operational Research, 237(3):1095–1104.

Sezer, O. B., Gudelek, M. U., and Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning : A systematic literature review: 2005–2019. Applied Soft Computing, 90:106181.

Shalendra, D. and Sharma (2010). Dealing with the contagion: China and india in the aftermath of the subprime meltdown. China & World Economy, (2):1–14.

Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27(4):623–656.

Shive, S. (2010). An epidemic model of investor behavior. Journal of Financial and Quantitative Analysis, 45(1):169–198.

Singh, M. K., Gómez-Puig, M., and Sosvilla-Rivero, S. (2021). Quantifying sovereign risk in the euro area. Economic Modelling, 95(C):76–96.

Sklar, A. (1959). Fonctions de Répartition à n Dimensions et Leurs Marges. Institut Statistique de I'Université de Paris, 8.

Sornette, D. (2009). Dragon-kings, black swans and the prediction of crises. Swiss Finance Institute, Swiss Finance Institute Research Paper Series, 2.

Souto, M., Keller, C., and Kunzel, P. (2007). Measuring sovereign risk in turkey: An application of the contingent claims approach. IMF Working Papers, 07(7/233):1–27.

Sriboonchitta, S., Liu, J., Vladik, K., and Nguyen, H. T. (2014). A vine copula approahc for analyzing financial risk and co-movement of the indonesian, philippine and thailand stock markets. In Modeling Dependence in Econometrics, pages 245–257, Cham. Springer International Publishing.

Stöber, J. and Czado, C. (2014). Regime switches in the dependence structure of multidimensional financial data. Computational Statistics & Data Analysis, 76:672–686. CFEnetwork: The Annals of Computational and Financial Econometrics.

Subasi, A. (2020). Chapter 3 - machine learning techniques. In Subasi, A., editor, Practical Machine Learning for Data Analysis Using Python, pages 91–202. Academic Press.

Sun, X., Wang, J., Yao, Y., Li, J., and Li, J. (2020). Spillovers among sovereign cds, stock and commodity markets: A correlation network perspective. International Review of Financial Analysis, 68:101271.

Sy, A. N. R. (2004). Rating the rating agencies: Anticipating currency crises or debt crises? Journal of Banking & Finance, 28(11):2845–2867.

Syllignakis, M. N. and Kouretas, G. P. (2011). Dynamic correlation analysis of financial contagion: Evidence from the central and eastern european markets. International Review of Economics & Finance, 20(4):717–732.

Tanaka, K., Kinkyo, T., and Hamori, S. (2016). Random forests-based early warning system for bank failures. Economics Letters, 148:118 – 121.

Tiwari, A., Andre, C., Gupta, R., et al. (2019). Spillovers between us real estate and financial assets in time and frequency domains. Technical report.

Tule, M. K., Ndako, U. B., and Onipede, S. F. (2017). Oil price shocks and volatility spillovers in the nigerian sovereign bond market. Review of Financial Economics, 35:57–65.

Valencia, F. and Laeven, M. L. (2008). Systemic banking crises: A new database. Number 8-224. International Monetary Fund.

Vapnik, V. (1999). An overview of statistical learning theory. IEEE Transactions on Neural Networks, 10(5):988–999.

Vatter, T. and Nagler, T. (2018). Generalized additive models for pair-copula constructions. Journal of Computational and Graphical Statistics, 27(4):715–727.

Wang, H., Yuan, Y., Li, Y., and Wang, X. (2021). Financial contagion and contagion channels in the forex market: A new approach via the dynamic mixture copula-extreme value theory. Economic Modelling, 94:401–414.

Wang, L., Li, S., and Chen, T. (2019). Investor behavior, information disclosure strategy and counterparty credit risk contagion. Chaos, Solitons & Fractals, 119:37–49.

Wang, P., Zong, L., and Ma, Y. (2020a). An integrated early warning system for stock market turbulence. Expert Systems with Applications, 153:113463.

Wang, P., Zong, L., and Yang, Y. (2020b). Predicting chinese bond market turbulences: Attention-bilstm based early warning system. In Proceedings of the 2020 2nd International Conference on Big Data Engineering, BDE 2020, page 91–104, New York, NY, USA. Association for Computing Machinery.

WATANABE, M. (2008). Price volatility and investor behavior in an overlapping generations model with information asymmetry. The Journal of Finance, 63(1):229–272.

Weszka, J. S. (1978). A survey of threshold selection techniques. Computer Graphics & Image Processing, 7(2):259–265.

White, H. (2000). A Reality Check for Data Snooping. Econometrica, 68(5):1097–1126.

Wu, B. (2020). Investor behavior and risk contagion in an information-based artificial stock market. IEEE Access, 8:126725–126732.

Wu, Y. and Gao, J. (2018). Adaboost-based long short-term memory ensemble learning approach for financial time series forecasting. Current Science (00113891), 115(1).

Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A. C., Salakhutdinov, R., Zemel, R. S., and Bengio, Y. (2015). Show, attend and tell: Neural image caption generation with visual attention. CoRR, abs/1502.03044.

Yan, H. and Ouyang, H. (2017). Financial time series prediction based on deep learning. Wireless Personal Communications, (4):1–18.

Yang, J., Cabrera, J., and Wang, T. (2010). Nonlinearity, data-snooping, and stock index etf return predictability. European Journal of Operational Research, 200(2):498 – 507.

Yeyati, E. L. and Panizza, U. (2011). The elusive costs of sovereign defaults. Journal of Development Economics, 94(1):95–105.

Yoon, W. and Park, K. (2014). A study on the market instability index and risk warning levels in early warning system for economic crisis. Digital Signal Processing, 29:35 – 44.

Yu, L., Wang, S., Lai, K. K., and Wen, F. (2010). A multiscale neural network learning paradigm for financial crisis forecasting. Neurocomputing, 73(4-6):716–725.

Yu, S. (2017). Sovereign and bank interdependencies—evidence from the cds market. Research in International Business and Finance, 39:68–84.

Yu, S. and Li, Z. (2018). Forecasting stock price index volatility with lstm deep neural network. In Recent Developments in Data Science and Business Analytics, pages 265–272. Springer.

Yuan, C. and Qiang, F. U. (2010). Time-varying characteristics of flights and market contagion in chinese financial markets——an empirical analysis of the correlations between stocks, bonds and gold. Systems Engineering, 28(5):1–7.

Zhang, G. and Liu, W. (2018). Analysis of the international propagation of contagion between oil and stock markets. Energy, 165:469 – 486.

Zhang, W., Zhang, G., and Helwege, J. (2020). Cross country linkages and transmission of sovereign risk: Evidence from china's credit default swaps. Journal of Financial Stability, page 100838.

Zheng, Z. (2018). China doesn't face a debt-default risk. `http://www.chinadaily.com.cn/a/201804/17/WS5ad52f46a3105cdcf65189bb.html` Accessed March 3, 2021.

Zolfaghari, M. and Gholami, S. (2021). A hybrid approach of adaptive wavelet transform, long short-term memory and arima-garch family models for the stock index prediction. Expert Systems with Applications, 182:115149.

Appendix A

Filtering and smooth probability derivation in SWARCH model

A.1 Calculation for Smooth and Filtering probabilities

The formula of computing filtering probabilities

$$P(s_t = i|\mathscr{Y}_t;\boldsymbol{\theta}) = \frac{P(s_t = i|\mathscr{Y}_{t-1};\boldsymbol{\theta})f(\mathbf{y}_t|s_t = i,\mathscr{Y}_{t-1};\boldsymbol{\theta})}{f(\mathbf{y}_t|\mathscr{Y}_{t-1};\boldsymbol{\theta})}$$

And the prediction probability for the next time point $t+1$ can be calculated with the following relationship with the filtering probability

$$P(s_{t+1} = i|\mathscr{Y}_t;\boldsymbol{\theta}) = \sum_{j=1}^{K} p_{ki}P(s_t = j|\mathscr{Y}_t;\boldsymbol{\theta}) \tag{A.1}$$

Kim (1994) derived the computing procedure[1]for the *smooth probability* that inferred from full-sample observations as

$$P(s_t = i|\mathscr{Y}_T;\boldsymbol{\theta}) = \sum_{j=1}^{K} P(s_{t+1} = j|\mathscr{Y}_T;\boldsymbol{\theta})P(s_t = i|s_{t+1} = j,\mathscr{Y}_T;\boldsymbol{\theta}) \tag{A.2}$$

$$= \sum_{j=1}^{K} P(s_{t+1} = j|\mathscr{Y}_T;\boldsymbol{\theta})P(s_t = i|s_{t+1} = j,\mathscr{Y}_t;\boldsymbol{\theta}) \tag{A.3}$$

$$= \sum_{j=1}^{K} P(s_{t+1} = j|\mathscr{Y}_T;\boldsymbol{\theta}) \times \frac{p_{ij}P(s_t = i|\mathscr{Y}_t;\boldsymbol{\theta})}{P(s_{t+1} = j|\mathscr{Y}_t;\boldsymbol{\theta})} \tag{A.4}$$

$$= P(s_t = i|\mathscr{Y}_t;\boldsymbol{\theta}) \times \Big(\sum_{j=1}^{K} \frac{p_{ij}P(s_{t+1} = j|\mathscr{Y}_T;\boldsymbol{\theta})}{P(s_{t+1} = j|\mathscr{Y}_t;\boldsymbol{\theta})}\Big) \tag{A.5}$$

A.2 Transition Matrix for Intermediate Variable s_t^*

There are four primitive states for $VAR(1) - SWARCH(2,1,2)$ model

$$s_t = 1 : s_{r,t} = 1, s_{o,t} = 1$$
$$s_t = 2 : s_{r,t} = 1, s_{o,t} = 2$$
$$s_t = 3 : s_{r,t} = 2, s_{o,t} = 1$$
$$s_t = 4 : s_{r,t} = 2, s_{o,t} = 2$$

where s_t denotes the primitive state for bivariate case, $s_{r,t}$ and $s_{o,t}$ represent the state at time t for the real-estate and the other asset respectively. [2]

The primitive transition probability matrix for s_t is

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} & p_{31} & p_{41} \\ p_{12} & p_{22} & p_{32} & p_{42} \\ p_{13} & p_{23} & p_{33} & p_{43} \\ p_{14} & p_{24} & p_{34} & p_{44} \end{bmatrix} \tag{A.6}$$

We introduce the intermediate state varible s_t^* to include the state at time $t-1$

$$s_t^* = \begin{cases} 1, & s_t = 1, s_{t-1} = 1 \\ 2, & s_t = 2, s_{t-1} = 1 \\ 3, & s_t = 3, s_{t-1} = 1 \\ \ldots, & \ldots\ldots \\ 15, & s_t = 3, s_{t-1} = 4 \\ 16, & s_t = 4, s_{t-1} = 4 \end{cases} \tag{A.7}$$

[1]The recursive calculation is backward with initial value of filtered probability at time T, $P(s_T = i|\mathscr{Y}_T; \boldsymbol{\theta})$.

[2]state 1 means low and 2 means high.

Hence the transition matrix for s_t^* can be written as

$$\mathbf{P}^* = \begin{bmatrix} p_{11} & 0 & 0 & 0 & \dots & p_{11} & 0 & 0 & 0 \\ p_{12} & 0 & 0 & 0 & \dots & p_{12} & 0 & 0 & 0 \\ p_{13} & 0 & 0 & 0 & \dots & p_{13} & 0 & 0 & 0 \\ p_{14} & 0 & 0 & 0 & \dots & p_{14} & 0 & 0 & 0 \\ 0 & p_{21} & 0 & 0 & \dots & 0 & p_{21} & 0 & 0 \\ 0 & p_{22} & 0 & 0 & \dots & 0 & p_{22} & 0 & 0 \\ 0 & p_{23} & 0 & 0 & \dots & 0 & p_{23} & 0 & 0 \\ 0 & p_{24} & 0 & 0 & \dots & 0 & p_{24} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & p_{43} & \dots & 0 & 0 & 0 & p_{43} \\ 0 & 0 & 0 & p_{44} & \dots & 0 & 0 & 0 & p_{44} \end{bmatrix}_{16\times 16} \tag{A.8}$$

In the calculations of filtered and smoothed probabilities, we need to sum up corresponding cases, for instance, the filtering probability for $s_t = 1$ is

$$Pr[s_t = 1|\mathscr{Y}_t; \hat{\boldsymbol{\theta}}] = Pr[s_t^* = 1|\mathscr{Y}_t; \hat{\boldsymbol{\theta}}] + Pr[s_t^* = 5|\mathscr{Y}_t; \hat{\boldsymbol{\theta}}] + Pr[s_t^* = 9|\mathscr{Y}_t; \hat{\boldsymbol{\theta}}] + Pr[s_t^* = 13|\mathscr{Y}_t; \hat{\boldsymbol{\theta}}]. \tag{A.9}$$

Appendix B

Comparative predictive models adopted in EWS

B.1 Logit regression

The logistic regression is one of most toiling parametric models that is empirically used to construct EWS for predicting curency crisis (Eichengreen et al., 1995; Frankel and Rose, 1996; Bussiere and Fratzscher, 2006), debt crisis (Dawood et al., 2017) and financial crisis based on market index and option (Li et al., 2015). The advantage of logit regression model sticks two benefits: the latent assupmtion that the dependent variable is linearly linked to other explanatory variables by adding a logistically distributed error, can distinctly convey the relationship among variables and explain the model uncertainty; on the other side, coefficients (with p-value of t-test) magnify the model interpretability and reliability in discovering influential factors.

The logit regression for modeling the probability of binary crisis variable y_t at time $t \in \{1, ..., T\}$ can be formulated as follows,

$$Pr(y_t = 1) = \frac{e^{\boldsymbol{x}_t \boldsymbol{\beta}}}{1 + e^{\boldsymbol{x}_t \boldsymbol{\beta}}},$$

where the $\boldsymbol{x}_t$ is the vector of explanatory variables at time t, $\boldsymbol{\beta}$ is the vector of coefficients. Coefficients will be obtained by maximum likelihood estimation and the joint log likelihood function is written as

$$logL = \sum_{t=1}^{T} (y_t log(Pr(y_t = 1)) + (1 - y_t) log(1 - Pr(y_t = 1))).$$

To mitigate the curse of dimensionality in regressing large explanatory variables on one dependent variable, we adopt the stepwise regression to extract and retain the effective combination of variables that maximally explain the dependent variable variation. As mentioned in Beutel et al. (2019), the fixed effect will be removed from the regression since it should be more comparable to other predictive models without extra terms.

In the stepwise backward algorithm, assumed m is the dimension of parameter vector, following steps will be attempted to search for the optimal model.

1. Establish the regression model between y and all explanatory variables of $\boldsymbol{x} = \{x_1, x_2, ..., x_m\}$, and do $F-$test for each x, take the minimum as $F_{l_1} = min\{F_1, F_2, ..., F_m\}$.

2. If $F_{l_1} > F_\alpha(1, T - m + 1)$, no variable will be eliminated, the regression model is the optimal one. Otherwise, we elminate x_{l_1} and denote the rest of variables as $\boldsymbol{x}_{-l_1} = \{x_1^1, x_2^1, ..., x_{m-1}^1\}$.

3. Establish the regression model between y and $\boldsymbol{x}_{-l_1}$, again do the $F-$test for each x and take the minimum as $F_{l_2} = min\{F_1^1, F_2^1, ..., F_{m-1}^1\}$.

4. If $F_{l_2} > F_\alpha(1, (T - m + 1) - 1)$, no variable will be eliminated. Otherwise, we elminate x_{l_2} and repeat the steps of $F-$test, comparing minimum with the margin and elimination, till not further variable is eliminated from the regression.

In general, the backward stepwise first put all variables into the model, and then attempt to remove one variable to examine whether significant change appears after the elimination. If there is no significant change, this elimination will be retained until all factors that lead significant change to the model are left. Thus, explanatory variables will be eliminated in turn and finally reordered according to their contribution degree to the model from small to large.

B.2 KLR signal extraction

KLR indicator approach is introduced in Kaminsky and Reinhart (1999) on the basis of nonparametric methodology, which also stands out in the EWS developing

realm as this signal extraction approach not only directly assesses the abnormality of single variable behaviour before or during the crisis period without the linearity assumption constraint, but provides a more comprehensive way to policy makers without training background of econometric and statistical modelling as well (Kaminsky, 1998; Lestano et al., 2004; Davis and Karim, 2008; Peng and Bajona, 2008), even though in the EWS model comparison study of Berg and Pattillo (1999) and Davis and Karim (2008), the improved logit regression[1] is proven to perform better than the signal extraction approach for predicting currency and banking crises. The approach monitors economic variables in a specified period and detects the ones deviates from the noise-to-signal ratio (NSR) minimized threshold as leading factors. The factors that are detected to anticipate the crisis will be counted into constructing the composite indicator by weighing each variable by their respective inverse of NSR (Kaminsky and Reinhart, 1999; Davis and Karim, 2008).

Table B.1: Confusing table for calculating the noise-to-signal ratio for each cutoff.

	Crisis	No crisis
Signal was issued	A	B
No signal was issued	C	D

The implementing process of KLR methodology is presented as a flowchart diagram, i.e. Figure B.1, to simplify words described steps in a more concise way.

In the diagram, we first take 80% to 90% percentile of observations for each variable and gradient search the optimal cutoff by producing the confusing matrix, as Table B.1 shows, calculating (adjusted) noise-to-signal ratio (NSR) of $\frac{B/(B+D)}{A/(A+C)}$, and searching for the minimal NSR corresponding cutoff value. Then, factor variables will be sifted by the extracted minimal NSR of NSR^j_{min} and the optimized cutoff of cutoff^j_{opt} for variable X^j. As green blocks label in the diagram, two decision conditions are (1) whether the variable value is greater than the optimized cutoff and (2) whether the noise-to-signal ratio is smaller than 0.75[2]. Then condition filtered m out of n factor variables will be synthesized by assigning the

[1]Both study adopt the multivariate logit regression.

[2]The significant level could be varied for specific markets according to the range of NSR values. Some studies (Davis and Karim, 2008) use 0.5 but find the strict value lead none of factors can be drawn as leading factors.

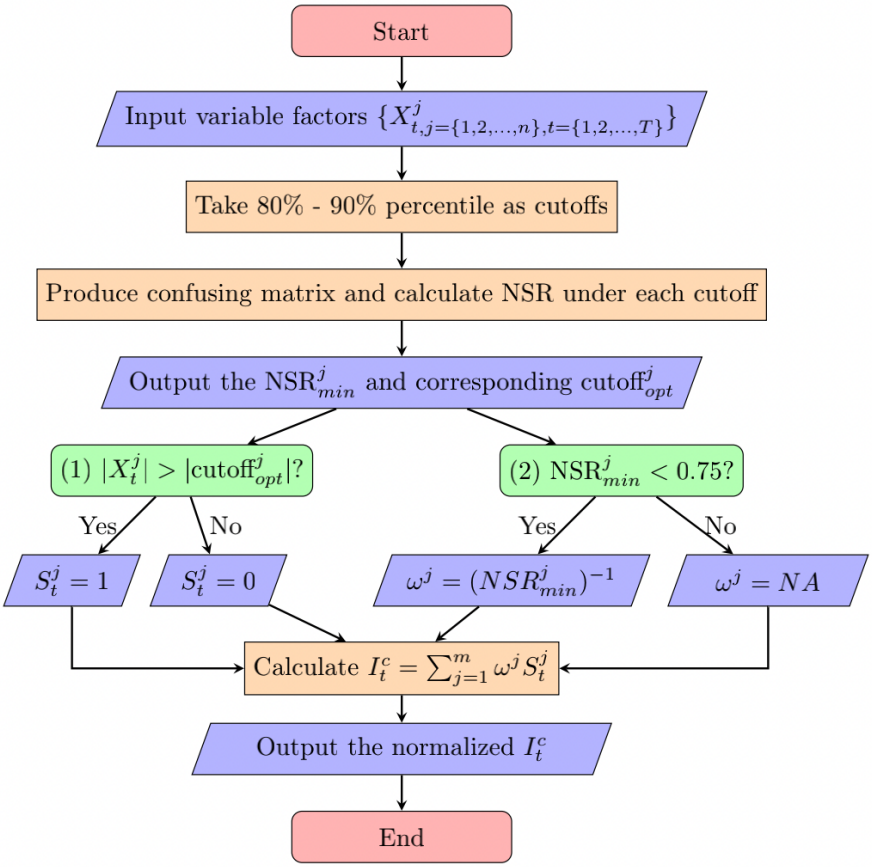


Figure B.1: KLR implementing diagram.

corresponding inverse of NSR_{min} as their weights to compose the final output, normalized crisis indicator of I^c_t.

B.3 Support Vector Machine

The support vector machine (SVM) is a statistical learning model to mainly solve the binary classification problem based on Vapnik (1999) proposed theory. The basic model is a linear classifier that seeks for the maximal separation of hyperplanes in the feature space. Its superioty of sloving nonlinear classification problems has made it speedily promoted in financial crisis predictions. Ahn et al. (2011) extend Oh et al. (2006) study on exploring the EWS for financial crisis and proved the SVM's effeciency in predicting long-term deterioration of economic fundamentals. Samitas et al. (2020) combine the traditional econometric model of GJR-GARCH and SVM to investigate the possible contagion risks across stocks, bond and CDS as the early indicator for financial crises, and show that the SVM

model makes extremely accurate predictions with 98.8% precision.

To precisely separate the samples of one class from those of the other class, the best hyperplane should be searched to make the largest margin between two classes. In our case, two classes are given as the "crisis" and "non-crisis", which is assigned to the our target variable of $y_t = \{-1, 1\}$. Along with the target, feature input vector[3] of $\mathbf{X}_t \in R^N$ will be paired to form the sample set of $\{(\mathbf{X}_t, y_t), t = 1, 2, ..., T\}$. The hyperplane to seperate the sample set will be expressed in the function of $\mathbf{w}'\mathbf{X} + b = 0$, where b is a constant number and $\mathbf{w}$ is the parameter vector sharing the same dimension with the feature vector of $\mathbf{X}$. Thus, the separating hyperplane is constraint by following boundary condition,

$$y_t(\mathbf{w}'\mathbf{X}_t + b) \geq 1. \tag{B.1}$$

The maximal distance between two planes of $\mathbf{w}'\mathbf{X} + b = 1$ and $\mathbf{w}'\mathbf{X} + b = -1$ is $\frac{2}{||\mathbf{w}||}$. To find the best hyperplane, the parameter optimization is thus converted to minimize the norm of $||\mathbf{w}||$,

$$\min_{b,\mathbf{w}} \frac{1}{2}||\mathbf{w}||^2 \tag{B.2}$$

with the constraint condition of Eq. B.1. To solve such quadratic programming problem, Lagrangian function is constructed including both objective function and corresponding constraints in assistance with Lagrange multipliers of α's.

$$\min_{b,\mathbf{w}} L = \min_{b,\mathbf{w}} \{\frac{1}{2}||\mathbf{w}||^2 - \sum_t \alpha_t[y_t(\mathbf{w}'\mathbf{X}_t + b) - 1]\} \tag{B.3}$$

where $\alpha_t \geq 0$ are Lagrange multipliers. To reach the optimal point, the saddle point equations will be differentiated by $\mathbf{w}$ and b respectively.

$$\frac{\partial L}{\partial \mathbf{w}} = \mathbf{w} - \sum_{t=1}^{T}(\alpha_t y_t \mathbf{X}_t) = 0 \tag{B.4}$$

$$\frac{\partial L}{\partial b} = \sum_{t=1}^{T}(\alpha_t y_t) = 0 \tag{B.5}$$

[3] N denotes the count of input features, also the dimension of the input feature vector at time t.

Put Eq. B.5 into B.3, we will have the dual quadratic optimization as follows,

$$\max L_D = \max(-\frac{1}{2}\sum_{t=1}^{T}\sum_{s=1}^{T}\alpha_t\alpha_s y_t y_s \mathbf{X}_t'\mathbf{X}_s + \sum_{t=1}^{T}\alpha_t) \tag{B.6}$$

with constraints of $\sum_{t=1}^{T}(\alpha_t y_t) = 0$ and $\alpha_t \geq 0$ for all t.

Such functional optimization, however, is not feasible to the non-separable cases with errors. Thus, the concept of 'soft margin' loss funciton is proposed by adding the slack variables of ξ_t and penalty term of C. The Eq. B.2 will be changed to

$$\min_{b,\mathbf{w},\xi} \frac{1}{2}||\mathbf{w}||^2 + C\sum_{t}\xi_t \tag{B.7}$$

with the constraint condition of $y_t(\mathbf{w}'\mathbf{X}_t + b) \geq 1 - \xi_t$ and $\xi_t \geq 0$ for all t. Furthremore, considering the strong non-linearity embedded in most real-life cases, the kernel function is further introduced to make the separating plane more flexible. Denoting the function that allows for mapping the samples of $\mathbf{X}_t$ from current space to the higher dimensional space as Φ: $R^d \rightarrow R^{d+1}$, then applying the kernel function $k(\cdot)$ to the mapped samples, i.e. $k(\mathbf{X}_t, \mathbf{X}_s) := (\Phi(\mathbf{X}_t'), \Phi(\mathbf{X}_s))$, the Lagrange transformation for Eq. B.7 will be restated as the following quadratic optimization problem,

$$\max L_D = \max(-\frac{1}{2}\sum_{t=1}^{T}\sum_{s=1}^{T}\alpha_t\alpha_s y_t y_s k(\mathbf{X}_t, \mathbf{X}_s) + \sum_{t=1}^{T}\alpha_t) \tag{B.8}$$

with constraints of $\sum_{t=1}^{T}(\alpha_t y_t) = 0$ and $\alpha_t \in [0, C]$ for all t.

The decision function for the separating hyperplane will be written as follows,

$$f(\mathbf{X}) = sign(\sum_{t=1}^{T}\sum_{s=1}^{T}\alpha_t y_t k(\mathbf{X}_t, \mathbf{X}_s) + b) \tag{B.9}$$

We summarize the full progress of using SVM to make classification as follows.

1. Determine the input samples and corresponding kernels.

2. Transform the objective optimization function into Lagrange function with

constraint conditions and calculate the Lagrange multipliers of α's to maximize the dual quadratic target function.

3. Put all optimized α's into expression for parameter vector of $\hat{\mathbf{w}}' = \sum_{t=1}^{T} \hat{\alpha}_t y_t \Phi(\mathbf{X}_t)$.

4. Put the value of $\hat{\mathbf{w}}'$ into $y_t = \mathbf{w}'\Phi(\mathbf{X}_t) + b$ to get $\hat{b}$.

5. Get the final expressions for the separating hyperplane of $\hat{\mathbf{w}}'\Phi(\mathbf{X}_t) + \hat{b} = 0$ and decision function for classification of $\hat{f}(\mathbf{X}) = sign(\hat{\mathbf{w}}'\Phi(\mathbf{X}_t) + \hat{b})$.

B.4 Random Forest and Gradient Boosting Tree

Both random forest and gradient boosting tree are tree-based model on the ensemble learning base. The core idea for decision tree is to continuously partition data into homogeneous clusters by refining the selection rules as either building or pruning tree branches to get the optimal tree structure. The tree-based model can naturally visualize the catergorizing rules and extract the variable importance, the model interpretability is thus more remarkable than other machine learning techniques. Koyuncugil and Ozgulbas (2012) and Tanaka et al. (2016) use the tree-based model construct EWS for predicting risk pressure for small enterprises and nationwide bank failures, which alters the practitioners' perspective in nonparametric models' predicting power. We thus put two advanced tree-based models in the stylish ensemble learning technique, random forest and gradient boosting tree, into model contrasts.

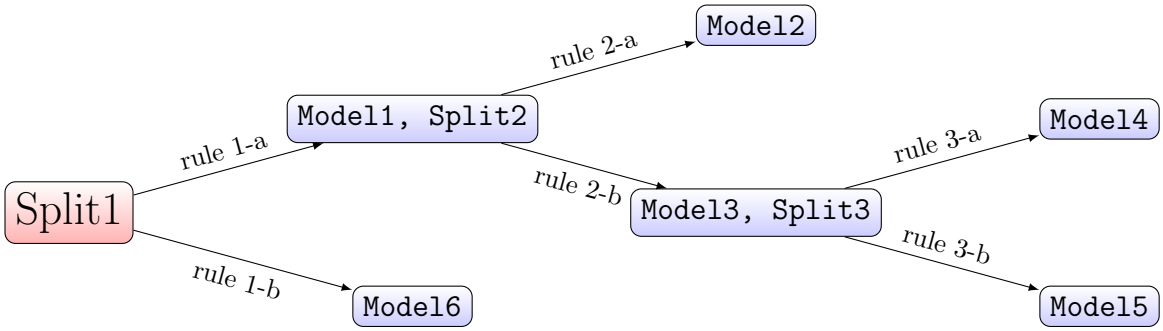


Figure B.2: An example of tree model structure with three clustering rules, three splits and six regression models.

Gradient boosting machines make the tree-based model algorithm more adaptive. It shares the similarity of random forest that the final prediction is produced

on an ensemble of tree models, but its constructing way is substantially different. Trees in random forests are built independently and each one will reach the maximum depth, while in gradient boosting, the trees are dependent on previous fitted trees by allowing the minimum depth. The computation steps are listed as follows.

1. Initialize D and M to be the tree depth and number of iterations. Compute the average of response $\bar{y}$ as the initial predicted value.

2. Start from the first iteration 1, calculate the residual, the difference between predicted value and observed value, and fit a D depth tree by setting the residuals as response.

3. Produce new predictions by using the fitted tree.

4. The predicted value will thus be updated by recursively implementing the step 2 and 3 and adding up the previous predicted value from past iterations.

Similarly, D and M are the tuning parameters for the gradient boosting machine. In the study, we across validate $D = \{1, 2, ..., 5\}$ and $M = \{50, 100, 150, ..., 500\}$, find the combination of $D = 3$ and $M = 100$ performs best by maximizing the AUC value for binary classification[4]. The time cost of implementing the gradient boosting machine is more pricey than random forest since the random forest constructs independent trees in parallel, the gradient boosting, though restricts the tree grown depth, aggregates previous results in an adaptive recurring process.

[4]R package of 'xgboost' helps to fit the gradient boosting model.

Appendix C

Dynamic correlation derivation in DCC-GARCH model

C.1 Derivation for dynamic correlation coefficient

For the multivariate DCC-GARCH model, the variance-covariance matrix H_t is

$$H_t = D_t R_t D_t \tag{C.1}$$

where $D_t = diag\{\sqrt{h_{i,t}}\}$ and R_t is the correlation matrix which can be expressed in the exponential smoother as a geometrically weighted average of standardized residuals.

$$[R_t]_{i,j} = \rho_{ij,t} = \frac{\sum_{s=1}^{t-1} \lambda^s \epsilon_{i,t-s}\epsilon_{j,t-s}}{\sqrt{(\sum_{s=1}^{t-1} \lambda^s \epsilon_{i,t-s}^2)(\sum_{s=1}^{t-1} \lambda^s \epsilon_{j,t-s}^2)}} \tag{C.2}$$

where λ's are scalar parameter and ϵ's [1] are standardized residuals.

Then, in a more simplistic way, the correlation can be written through the exponential smoothing as

$$\rho_{i,j,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}} \tag{C.3}$$

$$q_{ij,t} = (1-\lambda)(\epsilon_{i,t-l}\epsilon_{j,t-l}) + \lambda(q_{ij,t-l}), 1 \leq l \leq t-1 \tag{C.4}$$

[1] ϵ's have been denoted in Chapter 4, Eq. 6.2.

GARCH(1,1) model suggests a natural alternative to express $q_{ij,t}$ as following recursive equation,

$$\begin{aligned}
q_{ij,t} &= \bar{\rho}_{ij} + \alpha(\epsilon_{i,t-1}\epsilon_{j,t-1} - \bar{\rho}_{ij}) + \beta(q_{ij,t-1} - \bar{\rho}_{ij}) \\
&= \bar{\rho}_{ij} + \alpha\epsilon_{i,t-1}\epsilon_{j,t-1} - \alpha\bar{\rho}_{ij} + \beta q_{ij,t-1} - \beta\bar{\rho}_{ij} \\
&= (1-\alpha-\beta)\bar{\rho}_{ij} + \alpha\epsilon_{i,t-1}\epsilon_{j,t-1} + \beta q_{ij,t-1} \qquad \text{(C.5)} \\
&= (1-\alpha-\beta)\bar{\rho}_{ij} + \alpha\epsilon_{i,t-1}\epsilon_{j,t-1} + \beta((1-\alpha-\beta)\bar{\rho}_{ij} + \alpha\epsilon_{i,t-2}\epsilon_{j,t-2} + \beta q_{ij,t-2}) \\
&= (1-\alpha-\beta)(1+\beta)\bar{\rho}_{ij} + \alpha(\epsilon_{i,t-1}\epsilon_{j,t-1} + \beta\epsilon_{i,t-2}\epsilon_{j,t-2}) + \beta^2 q_{ij,t-2} \\
&= [\text{recursively expansion on } q_{ij,t-l}] \\
&= \frac{1-\alpha-\beta}{1-\beta}\bar{\rho}_{ij} + \alpha\sum_{s=1}^{\infty}\beta^s\epsilon_{i,t-s}\epsilon_{j,t-s}
\end{aligned}$$

where $\bar{\rho}_{ij}$ is the unconditional expectation of the cross product, thus for the variance $\bar{\rho}_{ii} = 1$.

The matrix forms of estimators C.4 and C.5 will be written as follows,

$$\lfloor q_{ij,t} \rfloor = Q_t = (1-\lambda)(\epsilon_{t-l}\epsilon_{t-l}^{'}) + \lambda Q_{t-l} \qquad \text{(C.6)}$$

$$Q_t = S(1-\alpha-\beta) + \alpha(\epsilon_{t-l}\epsilon_{t-l}^{'}) + \beta Q_{t-l} \qquad \text{(C.7)}$$

where S is the unconditional correlation matrix of ϵ's.

C.2 Estimation

The multivariate DCC-GARCH model can be statistically specified as

$$r_t|\mathscr{I}_{t-1} \sim N(0, H_t) \qquad \text{(C.8)}$$

$$\begin{cases}
H_t = D_t R_t D_t \\
\epsilon_t = D_t^{-1} r_t \\
D_t = diag(h_{11,t}^{1/2}, h_{22,t}^{1/2}, ..., h_{nn,t}^{1/2}) \\
Q_t = S(1-\alpha-\beta) + \alpha(\epsilon_{t-l}\epsilon_{t-l}^{'}) + \beta Q_{t-l} \\
R_t = (diag(Q_t))^{-1/2} Q_t (diag(Q_t))^{-1/2}
\end{cases} \qquad \text{(C.9)}$$

Given the assumption of normality, the parameters will be estimated by max-

imum likelihood[2]. The log likelihood can be written as

$$
\begin{aligned}
L &= -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + log|H_t| + r_t' H_t^{-1} r_t) \\
&= -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + log|D_t R_t D_t| + r_t' D_t^{-1} R_t^{-1} D_t^{-1} r_t) \\
&= -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2log|D_t| + log|R_t| + \epsilon_t' R_t^{-1} \epsilon_t) \\
&= -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2log|D_t| + log|R_t| + \epsilon_t' R_t^{-1} \epsilon_t + \epsilon_t'\epsilon_t - \epsilon_t'\epsilon_t) \\
&= -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2log|D_t| + log|R_t| + \epsilon_t' R_t^{-1} \epsilon_t + r_t' D_t^{-1} D_t^{-1} r_t - \epsilon_t'\epsilon_t) \\
&= -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2log|D_t| + r_t' D_t^{-2} r_t + log|R_t| + \epsilon_t' R_t^{-1} \epsilon_t - \epsilon_t'\epsilon_t) \\
&= -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2log|D_t| + r_t' D_t^{-2} r_t) - \frac{1}{2}\sum_{t=1}^{T}(log|R_t| + \epsilon_t' R_t^{-1} \epsilon_t - \epsilon_t'\epsilon_t) \\
&= L_V(\theta) + L_C(\theta, \phi)
\end{aligned}
$$

where $L_V(\theta)$ and $L_C(\theta, \phi)$ denote the likehood for volatility and correlation parts[3] respectively. θ is the parameters in D and ϕ is the parameters for R. The parameters will be estimated in two separate steps of

$$
\begin{aligned}
&(1):\ \hat{\theta} = \arg\max\{L_V(\theta)\}, \\
&(2):\ \hat{\phi} = \arg\max\{L_C(\hat{\theta}, \phi)\}.
\end{aligned}
$$

[2]If the Gaussian assumption is not satisfied, Quasi-Maximum Likelihood (QML) still works.

[3]$L_V(\theta) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2log|D_t| + r_t' D_t^{-2} r_t)$ denotes the volatility term, and $L_C(\theta, \phi) = -\frac{1}{2}\sum_{t=1}^{T}(log|R_t| + \epsilon_t' R_t^{-1}\epsilon_t - \epsilon_t'\epsilon_t)$ denotes the correlation term.

Appendix D

Hypothesis test

D.1 Reality check with revised benchmarks

The null for reality check can be written as

$$H_0 : E(f) \geq 0 \tag{D.1}$$

where f is the performance series of the difference between two portfolios' realized variances based on the EWS and the benchmark model.

On day of $t+1$, the performance series is produced by following

$$f_{t+1} = RV_{EWS,t+1} - RV_{benchmark,t+1}, \tag{D.2}$$

where $RV_{EWS,t+1}$ and $RV_{benchmark,t+1}$ are the realized variance of the price returns produced for the EWS model and the benchmarks respectively. Their specific calculating formulae are

$$RV_{EWS,t+1} = r_{t+1}^2 I_{EWS,t+1}, \tag{D.3}$$

$$RV_{benchmark,t+1} = r_{t+1}^2 I_{benchmark,t+1}. \tag{D.4}$$

where $r_{t+1} = ln(\frac{P_{t+1}}{P_t})$ denotes the log return of price index, and I_{t+1} is the indicator function for non-crisis observations that are complementary to forewarned signals for the EWS model, detected crises for SWARCH benchmark and always equal to one for market portfolio benchmark. As the performance statistic is the sample mean $\bar{f}$, boostrap samples $\bar{f}^*_{k,\{k=1,\ldots,10000\}}$ will be generated by applying

stationary bootstrap technique to calculate the p-values. The statistics of $\bar{V}$ will be set to compare with the quantile of simulated $\bar{V}_k^*$ for rejecting the null as follows,

$$\bar{V} = \sqrt{n}(\bar{f}), \tag{D.5}$$

$$\bar{V}_k^* = \sqrt{n}(\bar{f}_k^* - \bar{f}), \tag{D.6}$$

where n is the sample size for performance series f.

D.2 Scheffé test on attention mechanism

The third hypothesis test in Chapter 6, Section 6.2.2 on equivalence between train and test sets inferred attention weights will be implemented as follows.

Step 1: Construct F-test on the standard deviations of train and test attention weight vectors. The null and the F-test statistic are

$$\mathrm{H}_0 : \sigma_{in,i} = \sigma_{out,i} \text{ v. } \mathrm{H}_1 : \sigma_{in,i} \neq \sigma_{out,i}$$
$$\text{F-test: } F = \frac{max\{(S_{in,i}^*)^2, (S_{out,i}^*)^2\}}{min\{(S_{in,i}^*)^2, (S_{out,i}^*)^2\}} \sim F_{(\text{num, denom})}$$

where σ's denote the standard deviation for train and test two sets and S^*'s for in and out are sample unbiased standard deviations being defined as $(S_{in,i}^*)^2 = T_1 S_{in,i}^2/(T_1-1)$ and $(S_{out,i}^*)^2 = T_2 S_{out,i}^2/(T_2-1)$, respectively. The freedom degrees for F distribution are respective sizes of the numerator and denominator sets in F statistic, which depends on the maximum and minimum value for sample standard deviations.

Step 2: If the null in F-test can not be rejected with significant evidence, we construct t-test as H_0^{III} and Z_{III} (in Section 6.2.2) described. While, if the null is rejected, we alter to the Scheffé statistic as follows,

$$Z_i^t = \alpha_{out,i}^t - \sqrt{\frac{T_2}{T_1}}\alpha_{in,i}^t + \frac{1}{\sqrt{T_1 T_2}}\sum_{l=1}^{T_2}\alpha_{in,i}^l - \frac{1}{T_1}\sum_{l=1}^{T_1}\alpha_{in,i}^l, \text{ for } t=\{1,2,...,T_2\}, i=\{1,2,...,n\},$$

where all notations follow as defined in Section 6.2.2, T_1 and T_2 are the sample size for train and test sets respectively, n is the number of features. After pairing in

this way, a series of new samples of Z's will be generated with following adorable statistic summaries,

$$\begin{aligned} E(Z_i^t) &= \alpha_{out,i} - \sqrt{\frac{T_2}{T_1}}\alpha_{in,i} + \frac{1}{\sqrt{T_1 T_2}} T_2 \alpha_{in,i} - \frac{1}{T_1} \cdot T_1 \alpha_{in,i} \\ &= \alpha_{out,i} - \alpha_{in,i}, \text{ and} \\ Var(Z_i^t) &= \sigma_{out}^2 + \frac{T_2}{T_1}\sigma_{in}^2, \end{aligned}$$

then, given large size, Z's are assumed to be normally distributed as

$$Z_i \sim N(\alpha_{out,i} - \alpha_{in,i},\ \sqrt{\sigma_{out,i}^2 + \frac{T_2}{T_1}\sigma_{in,i}^2}\ )$$

Step 3: The original hypothesis thus can be transformed as the equivalent one to test the normality of Z's. Thus, the converted hypothesis and t-test can be constructed as follows,

$$\begin{aligned} &\mathrm{H}_0 : \alpha_z = \alpha_{out,i} - \alpha_{in,i} = 0 \text{ v. } \mathrm{H}_1 : \alpha_z = \alpha_{out,i} - \alpha_{in,i} \neq 0 \\ &\text{T-test} : Z = \frac{\bar{\alpha}_{out,i} - \bar{\alpha}_{in,i}}{S_Z/\sqrt{T_2 - 1}} \sim t_{(T_2-1)} \end{aligned}$$

where S_Z^2 is the sample variance being formulated as $S_Z = S_2^2 + \frac{T_2}{T_1}S_1^2 - 2\sqrt{\frac{T_2}{T_1}}S_{12}$, with the sample variance for test set $S_2^2 = \frac{1}{T_2}\sum_{t=1}^{T_2}(\alpha_{out}^t - \bar{\alpha}_{out})^2$, the truncated (till T_2) sample variance for train set $S_1^2 = \frac{1}{T_2}\sum_{t=1}^{T_2}(\alpha_{in}^t - \bar{\alpha}_{in^{T_2}})^2$ and the co-variance term $S_{12} = \frac{1}{T_2}\sum_{t=1}^{T_2}(\alpha_{out}^t - \bar{\alpha}_{out})(\alpha_{in}^t - \bar{\alpha}_{in^{T_2}})$.

Step 4: Calculated the p-value and the null will be rejected at 5%[1] level if the p-value is significantly smaller than it.

[1]Double-side hypothesis will take half of significance level value in calculation.