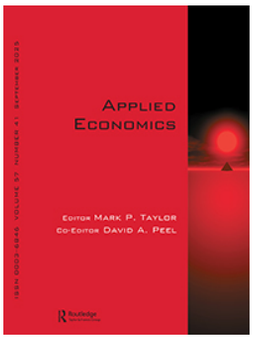

Routledge
Taylor & Francis Group

Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: www.tandfonline.com/journals/raec20

A non-linear Lasso and explainable LSTM approach for estimating tail risk interconnectedness

Tuhin Subhra De, Madeti Karthikeya & Sujoy Bhattacharya

To cite this article: Tuhin Subhra De, Madeti Karthikeya & Sujoy Bhattacharya (2025) A non-linear Lasso and explainable LSTM approach for estimating tail risk interconnectedness, Applied Economics, 57:41, 6433-6447, DOI: 10.1080/00036846.2024.2385747

To link to this article: https://doi.org/10.1080/00036846.2024.2385747

Published online: 13 Aug 2024.

APPLIED ECONOMICS
2025, VOL. 57, NO. 41, 6433–6447
https://doi.org/10.1080/00036846.2024.2385747

Routledge
Taylor & Francis Group

A non-linear Lasso and explainable LSTM approach for estimating tail risk interconnectedness

Tuhin Subhra De[a], Madeti Karthikeya[a,b] and Sujoy Bhattacharya[c]

[a]Department of Civil Engineering, Indian Institute of Technology Kharagpur, West Bengal, India; [b]Department of Mechanical Engineering, Indian Institute of Technology Kharagpur, West Bengal, India; [c]Business School, Edinburgh Napier University, Edinburgh, UK

ABSTRACT

Tail risk inter-connectivity is a significant aspect and a risk indicator that should be focused on. Many of the previous works have shown potential non-linearity in tail risk contagion. With the recent advancements in deep learning, Long-Short Term Memory (LSTM) networks have played an important role in sequential data prediction. We experiment with LASSO-based neural networks and interpretative LSTM model along with other machine learning approaches for investigating tail risk interconnectedness among the public banks of Japan. We also investigate the risk reception from large overseas banks in United States finding that medium-sized banks are more likely to receive international risks. Our studies show that LSTM-based model is an excellent fit for the scenario and total connectedness goes up during an economic crisis. The banks having larger market capitalization are more prone to emission and reception of tail risks. This is accompanied by exhibiting the impact of some major economic distresses on Japanese banking system. These results provide important information to regulators and policy makers.

KEYWORDS

Tail risk; interconnectedness; machine learning; neural networks; LSTM

JEL CLASSIFICATION

G10; G15; G17

I. Introduction

During crises, bad luck (tail risks) can spread between financial institutions, threatening the whole system. This 'systemic risk' stems from two factors: how risky individual companies are and how connected they are. A single failure can cause a domino effect, crippling the entire system. Liquidity issues, credit risks of failing partners, or even panic-driven price swings could all cause this 'spillover effect'. Regulators can't just look at individual companies anymore. They, like investors who focus on potential losses more than gains, need to consider the interconnectedness (Basel 2013) of the financial system as a key risk measure.

There are several issues with tail risk interconnectedness. First, although tail risk spillover effects are somewhat nonlinear in nature while most of the suggested solutions are based on the linear assumption. Second, risk depends on lot of factors hence a multi-dimensional setup should be used. Thirdly, if a flexible and nonlinear model is used the challenge of interpretation comes into picture. It would be challenging to assess the contribution of particular risk variables. Having an indicator that distils nonlinear and interacting risk components is useful for regulators. As a result, it demands a data-driven model that is not only adaptive but also capable of providing straightforward interpret-ability.

Nguyen, Chevapatrakul, and Yao (2020) used LASSO (Least Absolute Shrinkage and Selection Operator) to capture the important features for assessment of interconnectedness. They have directly incorporated the penalized linear coefficients of the regressor as feature importance. The existing approaches to estimating conditional tail risk have certain potential drawbacks, as pointed out by Härdle, Wang, and Yu (2016). Torri, Giacometti, and Tichý (2021) studied a network model to analyse how interconnected banks spread risk. It identifies vulnerable banks and calculates systemic risk for each one. The model considers traditional risk factors as well. Results show interconnectedness in the European banking system, with higher concentration of risk in southern Europe during economic downturns. With the

CONTACT Tuhin Subhra De tuhinsubhrade069@gmail.com Indian Institute of Technology Kharagpur, Lane no. 11, Sreepally 1, Kenduadihi, Bankura, West Bengal 722012, India

ongoing fascination of cryptocurrencies, tail risk measurement was done with several cryptocurrencies (Ahelegbey, Giudici, and Mojtahedi 2021) using Extreme Downside Hedge (EDH) and Extreme Downside Correlation (EDC) which led the authors to bifurcate cryptocurrencies into speculative and technical assets.

Hautsch, Schaumburg, and Schienle (2014) researched a new way to assess how financial firms contribute to overall risk in the financial system. They consider the interconnectedness of financial firms and how the risk of one firm can impact others. Their method, called 'realized systemic risk beta', uses statistical analysis to measure how a firm's risk (Value-at-Risk) affects the risk of the entire system. Giudici, Leach, and Pagnottoni (2022) compare two types of stablecoins: one backed by a basket of currencies and another pegged to a single major currency (like the US dollar). The argument is about basket-based stablecoin being better. They achieve this by finding the optimal mix of currencies for the basket, then showing how this basket is less volatile than individual currencies.

Since, any financial crisis shows its impact on the systemic banking network of the economy, within recent years, Dew-Becker (2022) and Nguyen and Lambe (2021) have explored the determinants of total tail risk. Capital flow, degree of trade and business linkage between countries, macroeconomic factors such as inflation, stock market volatility, exchange rates, interest rates play a crucial role in determining tail risk interconnectedness. Giudici and Parisi (2018) suggested a novel credit risk measurement model for Corporate Default Swap (CDS) spreads that combines vector autoregressive regression with correlation networks. This was applied to European countries creating a clustering effect between central and border countries. Maghyereh and Yamani (2022) found that banks with more income sources (diversified) have lower risk of causing financial system problems. This effect is stronger for Islamic banks. Even during the COVID-19 pandemic, both Islamic and conventional banks faced similar risk.

Long et al. (2022) proposed a Gradient Boosting Machine (GBM) regressor approach to measure tail risk interconnectedness. They used relative variable importance integrating with explanatory power of inputs as a measure to quantify the metric. Inspired by recent growth in deep learning, we imitate a robust and accurate fit model for estimating the tail risk within the financial institutes. A recent work developed by Lemhadri et al. (2021) which created LassoNet: A neural network with feature sparsity. LassoNet demonstrates interpretable characteristics for our estimation by applying L1 regularization (LASSO) to appropriate weights of the neural network. Another form of Neural Networks are Sequential or Recurrent Neural Networks (RNN Rumelhart, Hinton, and Williams (1986), Jordan (1986)) introduced in 1980s tend to capture the patterns and trends in sequential data form. Specially designed RNN known as Long-Short Term Memory (LSTM) Hochreiter and Schmidhuber (1997) were introduced to retain the flow of long term important dependencies within the models. Extracting tail risk interconnectedness is a fine task of predicting returns and hence we chose to go with the LSTM model added with interpretable nature termed as Interpretable-Multivariate LSTM (Guo, Lin, and Antulov-Fantulin 2019). We experiment with the above described approaches and find out that IMV-LSTM-based approach outperforms all the other methodologies.

We focus on Japan's banking system primarily on how it is affected by its own national banking system and other major international banks from United States. Japan consists of several large banks like Mitsubishi UFJ Financial Group ($2.54 billion USD), Mizuho Financial Group ($1.67 billion USD), Bank of Japan ($4.85 billion USD) in terms of total assets. Our major concern lies with studying the influence of 2011 mega tsunami that hit Japan and the drastic effect of COVID-19 pandemic. Using the LSTM approach we extract out the interconnectedness among the Japanese banks along with the risk contagion with large banks in United States. Various analyses on dynamic risk transmission and risk reception from domestic and international domain show that banks with high market capitalization are prone to get affected by tail risks and mediocre banks are highly likely to get affected by international risks. It is also found that the risk goes high up during the time of economic distress. Liu (2014) showcased that in case of high volatile situations of Japanese markets and the

necessity to control volatility in Japanese markets. We extend our risk interconnectedness to major overseas banks in United States of America. With the Japanese banks as our epicentre of analysis we only observe the risk reception from these banks into the Japanese system.

The rest of the paper is structured in the following format. Section II contains the details of methodology implemented. Section I shows the empirical results followed by Section IV which concludes our study.

II. Methodology

Concept of CoVaR

Value at Risk (VaR) is a measurement of the likelihood of a loss over a certain time period. Regulators and risk managers in the financial sector frequently use it to evaluate the magnitude and probability of possible loss in their assets. VaR_q^i can be defined as:

$$P\left(X^i \leq VaR_q^i\right) = q \tag{1}$$

where X^i stands for the log return of a financial institute i and q is the quantile level. Adrian and Brunnermeier (2016) gave a concept of CoVaR, which is basically the conditional format of V aR taking spillover effects and the macro state of the economy into account:

$$P\left(X^j \leq CoVaR_q^{j|V^{-j}} | V^{-j}\right) = q \tag{2}$$

where $V^{-j} \equiv (X^i, M)$ is a vector representing the influences from financial institutes other than j along with the lagged macro-state variables. More precisely to determine V aR at time t, Equation 3 can be solved using qauntile regression:

$$X_t^i = \alpha^i + \gamma^i M_{t-1} + \varepsilon_t^i \tag{3}$$

$$VaR_t^i = \hat{\alpha}^i + \hat{\gamma}^i M_{t-1} \tag{4}$$

Here γ is the sensitivity term for dependence of log return on macro-state variables and α is the offset. CoVaR can be determined by solving Equation 5 by quantile regression and substituting VaR obtained from Equation 4 into Equation 6:

$$X_t^{sys|i} = \alpha^{sys|i} + \gamma^{sys|i} M_{t-1} + \beta^{sys|i} X_t^{sys|-i} + \varepsilon_t^{sys|i} \tag{5}$$

$$CoVaR_t^i = \alpha^{\hat{sys}|i} + \gamma^{\hat{sys}|i} M_{t-1} + \beta^{\hat{sys}|i} VaR_t^i \tag{6}$$

The β in Equation 6 can interpreted as the degree of dependence of log return of institute i to the influences from institutes other than i or which we refer as a system of institutes sys. The amount of tail risk interconnectedness within a financial system is determined by using this equation iteratively for all institutes.

Adrian and Brunnermeier (2016) described a bivariate estimation of β in their work not considering the interactions within the institutes. Moreover Chao, HÃ¤rdle, and Wang (2012) showed that complex nonlinear dependency exists between financial assets, especially when the financial system is under distress. Keeping things under consideration we extend our search into a nonlinear space. We experiment with several machine learning models extending from tree based to constrained by regularization and Recurrent Neural Networks. Among all of the approaches the performance of RNN-based model was the best assisting us to derive an accurate metric for the tail risk interconnectedness.

Objective

To present our tail risk estimation problem in a mathematical formulation let us consider a set of financial institutes S, their daily log returns to be X and macro state variables which are country specific to be M. Then for some institute $j \in S$ we need to estimate an empirical value let us denote it by D which will represent the tail risk interconnectedness of j with $i \in S | i \neq j$. We start with a flexible parametric model:

$$X^j = F(V^{-j}) \tag{7}$$

Where F is a function. D is obtained from the internal parameters of the function F post optimization subjected to minimizing the residual quantile sum of loss given by:

$$RL_q^j = \sum_{t=1}^{T} L_q\left(X_t^j - F(V_t^{-j})\right) \tag{8}$$

where $L_q(u) = u(q - I(u < 0))$ is the quantile loss function for quantile q, T is total number of data samples and $I()$ is an indicator function which returns 1 if its argument condition is satisfied, else 0. Since we are focused on left tail, hence we choose to go with $q = 0.05$.

LassoNet

LASSO or least absolute shrinkage and selection operator is a commonly used regularization method used for zeroing out weights corresponding to redundant features. Generally lasso is used with linear models with applying the L1 norm on the weights of regression model.

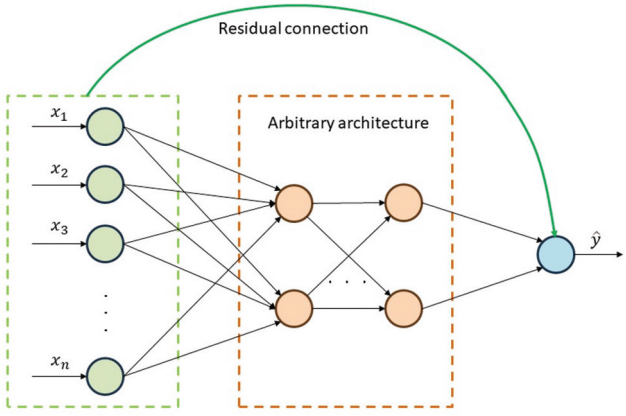

Figure 1. Architecture of LassoNet. Design comprises a single residual link (shown in green) and an arbitrary feed-forward neural network (represented in **black**).

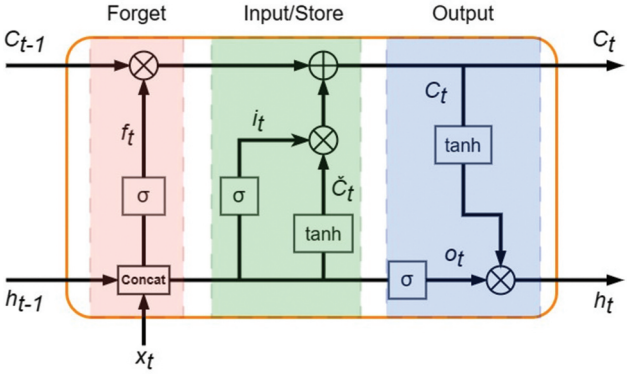

Figure 2. Architecture of vanilla LSTM model with highlighted regions of functions. ⊗ and ⊕ denote element-wise multiplication and addition respectively. *σ* and *tanh* represent sigmoid and hyperbolic-tangent activation function respectively.

Previously discussed in Equation 7 that F is a generalized parametric model hence we look at Lemhadri et al. (2021) which describes a residual neural network model framework with global feature selection (Figure 1). It is capable of selecting relative importance of input variables by applying L1-regularization in a non-linear space.

Representation

To estimate F, we chose this to be class of residual feed-forward networks:

$$F = \{f \equiv f_{\theta,W} : x \to \theta^T \mathbf{x} + h_W(\mathbf{x})\}, \tag{9}$$

where $\mathbf{x}$ is the input vector with dimension d, h_W denotes the feed-forward neural network with weights W and $\theta \in \mathbb{R}^d$ are the weights of the residual (skip) layer. Assuming K to be the number of units in the first hidden layer, then its weights; $W^{(1)} \in \mathbb{R}^{d \times K}$.

For applying regularization, the original loss function minimization is added with a penalty that promotes feature sparsity. There is involvement of a proximal algorithm during training which assists in the trade-off between linear and non-linear component of the model by updating skip and first hidden layer's weights following a customized constraint defined in Equation 11. We then choose an optimal network F^* from Equation 9 which minimizes the quantile loss function in Equation 8:

$$F^*(x) = argmin_{F(x)} RL_q^j \tag{10}$$

Formulation

The constrained objective of LassoNet is defined by:

$$\begin{aligned} \underset{\theta,W}{\text{minimize}} \quad & L(\theta, W) + \lambda ||\theta||_1 \\ \text{subject to} \quad & ||W_j^{(1)}||_\infty \leq M|\theta_j|, j = 1, \ldots, d. \end{aligned} \tag{11}$$

where $L(\theta, W)$ [1] is the loss function defined in Equation 8 and M is a hierarchical coefficient.

This type of model architecture is hybrid in nature as it jointly consists of linear and non-linear operators. Parameters from both kind of operators can be directly used for feature

[1]The default loss function for LassoNet regressor is set to mean-squared error. We modify the source code to change the loss function to quantile loss.

selection. Considering the scalibilty factor of data, sometimes deep neural networks can be difficult to train given the computation time constraints. However, residual networks like this are easy to train shown by He et al. (2016). Further Raghu et al. (2017) showed that residual neural networks act as universal approximators to many classes of functions which give us the freedom to choose LassoNet as an ideal approach for this scenario.

Training

LassoNet mainly has two hyper-parameters associated with the L1 penalty coefficient λ, and hierarchical parameter M. It controls the contribution of linear and non-linear component.

The whole training process is formulated in the two parts. The first part starts with fitting a general un-regularized artificial neural network on the training set of data. The network uses early-stopping method to stop the training. The best fitted weights (the ones which are having the least value of loss function on test set) are then used to compute the regularized version of the network. The model then starts to sail from dense-to-sparse state as the penalizing coefficient λ is gradually increased. It starts filtering out less significant features and keeps on continuing with the most significant ones. Once all the active features become zero the training eventually stops.

Interpretable LSTM

IMV-LSTM

The emergence of Long Short-Term Memory (LSTM) (Hochreiter and Schmidhuber 1997) networks has revolutionized the landscape of time series forecasting. LSTMs are widely employed for forecasting stock prices, exchange rates, and other financial time series, offering superior accuracy compared to traditional methods (Bhandari et al. 2022).

However in the case of tail risk interconnectedness estimation we primarily deal with the extent of impact of features on the predicted variables. Extracting out feature importance in LSTMs is a challenging task. To handle this problem, Guo, Lin, and Antulov-Fantulin (2019) reinvented the LSTM with interpretable nature known as Interpretable Multi-Variable LSTM, or IMV-LSTM: (Figure 2)

Representation

Assuming we have $N-1$ time series and a target series variable $\mathbf{y}_T = [y_1, y_2, \ldots y_T]$ of length T. The multi variable inputs can now be stacked together to form an input tensor $\mathbf{X}_T = [\mathbf{x}_1, \ldots, \mathbf{x}_T]$ and $\mathbf{x}_t = [x_t^1, \ldots, x_t^{N-1}, y_t]$ so $\mathbf{x}_t \in \mathbb{R}^N$. Each of x_t^n is the input at time step t and can be multidimensional. Given $\mathbf{X}_T$, the objective is to learn a non-linear mapping which gives us the target value for the next time step:

$$\tilde{y}_{T+1} = \mathcal{F}(\mathbf{X}_T) \tag{12}$$

While estimating the function $\mathcal{F}$, the variable importance are learned. The importance vector $\mathbf{I} \in \mathbb{R}^N$ is non-negative and is normalized such that $\sum_{n=1}^{N} I_n = 1$.

The purpose of IMV-LSTM is to use hidden state matrices and create corresponding update schemes so that each element (such as a row) of the hidden matrix only contains data from a particular input variable. The hidden states and gates in IMV-LSTM are replaced by tilde to avoid confusion with the vanilla LSTM. The hidden state and time step t is given as $\tilde{\mathbf{h}}_t$ where $\mathbf{h}_t^n \in \mathbb{R}^d$. The input to hidden transition matrix $\mathcal{U}_j$ where $\mathbf{U}_j^n \in \mathbb{R}^{d \times d_0}$ and the hidden to hidden transition is given by $\mathcal{W}_j$, where $\mathbf{W}_j^n \in \mathbb{R}^{d \times d}$. The updated hidden state is given by:

$$\tilde{\mathbf{j}}_t = \tanh\left(\mathcal{W}_j * \tilde{\mathbf{h}}_{t-1} + \mathcal{U}_j * \mathbf{x}_t + b_j\right) \tag{13}$$

Here $\tilde{\mathbf{j}}_t \in \mathbb{R}^{N \times D}$ and $*$ is the tensor dot product between two tensors. The process of vectorization denoted by *vec* in Equation 15 involves flattening a two-dimensional matrix into one dimension by concatenating the columns. And the operation matricization$(\cdot)$ is the inverse of vectorization in Equation 16. It reshapes a vector into two-dimensional matrix, in this case it goes from $\mathbb{R}^D \rightarrow \mathbb{R}^{N \times d}$. Element wise multiplication is defined by $\odot$.

$$\begin{bmatrix} \mathbf{i}_t \\ \mathbf{f}_t \\ \mathbf{o}_t \end{bmatrix} = \sigma\Big(\mathbf{W}[\mathbf{x}_t, \tilde{\mathbf{h}}_{t-1}] + b\Big) \tag{14}$$

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot vec(\tilde{\mathbf{j}}_t) \tag{15}$$

$$\tilde{\mathbf{h}}_t = \text{matricization}(\mathbf{o}_t \odot \tanh(\mathbf{c}_t)) \tag{16}$$

In general the Equation 12 can be taken up by Bayesian method in probabilistic domain:

$$\begin{aligned} p(y_{T+1}|\mathbf{X}_T) &= \sum_{n=1}^{N} p(y_{T+1}|z_{T+1} = n, \mathbf{X}_T) \\ &\cdot p(z_{T+1} = n, \mathbf{X}_T) \cdot p(z_{T+1} = n, \tilde{\mathbf{h}}_1, \ldots, \tilde{\mathbf{h}}_T) \\ &= \sum_{n=1}^{N} p(y_{T+1}|z_{T+1} = n, [\mathbf{h}_T^n, \mathbf{g}^n]) \\ &\cdot p(z_{T+1} = n|[\mathbf{h}_t^1, \mathbf{g}^1], \ldots, [\mathbf{h}_t^N, \mathbf{g}^N]) \end{aligned} \tag{17}$$

In Equation 17 there is an introduction of latent variable z_{T+1} for estimating y_{T+1}. It is discrete variable over the set of values {1, 2, N}. The context vector $\mathbf{g}^n$ is computed as the temporal attention weighted sum of hidden states of variable n. It is given by $\mathbf{g}^n = \sum_t \alpha_t^n \mathbf{h}_t^n$ where $\alpha_t^n = \frac{\exp f_n(\mathbf{h}_t^n)}{\sum_k \exp f_n(\mathbf{h}_k^n)}$. Here $f_n(\cdot)$ can a feed forward neural network.

The first part of probability in Equation 17 can be estimated by Gaussian distribution parameterized on $[\mu_n, \sigma_n] = \varphi_n([\mathbf{h}_T^n, \mathbf{g}^n])$ and second part by taking softmax function over $\{f([\mathbf{h}_T^n, \mathbf{g}^n])\}_N$.

Estimation of tail risk interconnectedness

The importance vector $\mathbf{I}$ is learned during the training process of the IMV-LSTM. For a training data instance m, we define the variable importance as:

$$\begin{aligned} \mathbf{I} &= \frac{1}{M}\sum_m \mathbf{q}_m; \text{where} \\ q &:= p(y_{T+1,m}|z_{T+1,m} = n, [\mathbf{h}_T^n, \mathbf{g}^n]) \\ &\cdot p(z_{T+1,m} = n|[\mathbf{h}_t^1, \mathbf{g}^1], \ldots, [\mathbf{h}_t^N, \mathbf{g}^N]) \end{aligned} \tag{18}$$

However only I values do not provide any information on variation of response caused by a feature. Therefore we incorporate the quantile format of R-square defined in Davino, Furno, and Vistocco (2014) by multiplying it with Equation 18. It also measures the quality of fit of a regression model on input data:

$$R_q^j = 1 - \frac{RL_q^j}{TL_q^j} \tag{19}$$

where RL_q^j is defined in Equation 8 and $TL_q^j = \sum_{t=1}^{T} L_q\Big(X_t^j - VaR_q^j\Big)$. These metrics are derived from the linear part of LassoNet.

We also extend our estimation by incorporating the sensitivity of the outputs according to small perturbations in inputs. For a neural network with one hidden layer Yannis, Paul, and Sovan (1995) define the sensitivity as the jacobian matrix of the mapping function. Absolute value is used since we are concerned with the magnitude, or how much it changes rather than the direction of change. Härdle, Wang, and Yu (2016) also proposed the concept variable selection through a single-index model by leveraging the partial derivative of the estimator function of CoVaR in their work. Therefore as stated about the empirical value D in Section 2.2, we can combine both the metrics and the final tail risk interconnectedness metric can be stated as:

$$D^{i|j} = 0.5\left(R_q^j \cdot \mathbf{I}_i^j + \left|\frac{\partial \mathcal{F}(V^{-j})}{\partial V^{-j}}\right|_i\right) \tag{20}$$

Where $\mathcal{F}$ is defined in Equation 12, the partial derivatives for a feature i are calculated by averaging out the individual values through all the data samples.

If there are N financial institutes associated, for the national case we can construct a $N \times N$ directed adjacency matrix in which each element at i^{th} row and j^{th} column represents the tail risk from institute j to i.

$$Z = \begin{bmatrix} 0 & D^{2|1} & \ldots & D^{N|1} \\ D^{1|2} & 0 & D^{3|2} & D^{N|2} \\ \vdots & D^{2|3} & \ddots & \vdots \\ D^{1|N} & D^{2|N} & \ldots & 0 \end{bmatrix} \tag{21}$$

It may to be noted that Z may not be symmetric in nature or $D^{i|j}$ might not be same as $D^{j|i}$. This accounts for the spillover effect between the institutes. Further it can be concluded that each row of the matrix indicates the outgoing influence from one respective institute and each column denotes how one respective institute is affected by other institutes. Since one institute can't affect itself hence the diagonal elements are 0.

In the case of risk received from international banks, we can construct a $N \times M$ matrix where M is the number of international banks. Since we are concerned with how much risk is received by the Japanese banks from US banks hence, the matrix so formed in this case will have non-zero elements and $D^{i|j}$ will represent the risk received by i^{th} Japanese bank from j^{th} US Bank.

III. Experimental results

Investigating the tail risk spread within the public banks in Japan, we apply the described methodology to perform an country specific analysis of ten large public banks in terms of assets situated in Japan. Along with that, the connectedness is extended to large public banks of United States of America. Most of the world's major banks in terms of assets are located in US, China and Japan. Due to restrictions on data availability, the link of tail risk between Chinese and Japanese banks can not be taken up. For the case of international tail risk interconnectedness, we use the national as well as international macro state variables for analysis.

Data

All the returns data used in our work are available at Yahoo Finance. The daily log returns are defined by $X_t^j = \log(P_t^j/P_{t-1}^j)$ where P_t^j is the dividend-adjusted closing price of each bank. Due to constraints on availability of data for certain parameters from early 2000s, we stick to data obtained for over a period of almost 14 years starting from 1 April 2010 to 31 January 2024 (both dates inclusive) spanning for 5054 days. Table 1 provides a detailed description of the listed banks for Japan and United States. Several economic events affected Japan during this period, the major ones being the great tsunami of 2011 and COVID-19 pandemic staring from 2019 extending to 2021 termed as post-covid period. Other event include consumption tax increase in 2014, negative interest rates implemented by Bank of Japan in 2016 and Germany overtaking Japan in economy in 2023.

From Table 1, we see few of the Japanese banks have high daily fluctuations in returns giving rise to maximum and minimum both greater than 2 signifying unpredictable patterns in Japanese stock markets (Andersen, Todorov, and Ubukata 2021). However the patterns are total opposite in the case of US banks. The maximum values are less than 2 in

Table 1. Data description of all the public banks of Japan and US with ticker code and abbreviation. The table provides analysis of the log returns' mean, standard deviation, minimum, maximum, skewness and the kurtosis of the values.

	ticker	abbreviation	mean	std	min	max	skewness	kurtosis
Japanese Banks								
Bank of Japan	8301.T	BOJ	−0.00015	0.01765	−0.20479	0.19237	0.62832	23.72007
Mitsubishi UFJ Financial Group	8306.T	MUFG	0.00030	0.01483	−0.11973	0.10606	0.15425	6.51407
Sumitomo Mitsui Financial Group	8316.T	SMFG	0.00029	0.01398	−0.09396	0.09866	0.12703	6.34153
Mizuho Financial Group	8411.T	MZF	0.00018	0.04775	−2.33791	2.27309	−1.89512	2151.64664
Sumitomo Mitsui Trust Holdings	8309.T	SMTH	0.00020	0.04962	−2.26592	2.25132	−0.36218	1711.84205
Resona Holdings Inc	8308.T	RHC	0.00003	0.01555	−0.17841	0.12909	−0.29967	10.92338
Aozora Bank	8304.T	AZB	0.00031	0.04793	−2.30715	2.28942	−0.52247	2090.88071
The Chiba Bank	8331.T	CHB	0.00019	0.01511	−0.10612	0.10253	−0.16081	5.69270
Hokuhoku Financial Group	8377.T	HFG	0.00002	0.04809	−2.28539	2.27691	−0.24535	2001.72894
Seven Bank	8410.T	SVB	0.00018	0.13808	−6.90574	6.91444	0.08988	2479.67897
US Banks								
JPMorgan Chase and Co.	JPM	JPMC	0.00034	0.01468	−0.16211	0.16562	−0.05031	15.10439
Bank of America	BAC	BOA	0.00017	0.01793	−0.22713	0.16379	−0.31166	15.77881
Wells Fargo	WFC	WFGO	0.00017	0.01577	−0.17278	0.13571	−0.32127	13.17231
Citigroup. Inc	C	CITI	0.00010	0.04899	−2.30168	2.27931	−0.62631	1888.84682
PNC Financial Services Group	PNC	PNC	0.00026	0.01500	−0.17317	0.12172	−0.33693	13.40489
The Goldman Sachs Group	GS	GSG	0.00021	0.01510	−0.13686	0.16195	−0.36520	13.27923
Truist Financial Corporation	TFC	TFC	0.00012	0.01649	−0.21190	0.15909	−0.83341	20.05204
Capital One Financial Corporation	COF	COFG	0.00028	0.01792	−0.27278	0.17189	−0.75840	22.81533
Toronto-Dominion Bank	TD	TDB	0.00020	0.01788	−0.69720	0.69129	−0.38066	900.59292
US Bancorp	USB	USB	0.00017	0.01473	−0.15596	0.16014	−0.29582	16.25534

most of the cases and the absolute minimum is also less than maximum suggesting longer positive tails. Models especially neural network based do not expect any probability distribution for the input data. Hence, no statistical test is performed on data.

Following Adrian and Brunnermeier (2016) we obtain the macro state variables for Japan. These variables are highly impacted by the country's economic state:

(1) The change in three-month treasury bond yield
(2) The difference between the yields on 3-month and 10-year Treasury bonds
(3) The difference between the 10-year Treasury bond yield and AAA corporate bond yield
(4) The difference between the 3-month Treasury bond yield and the 3-month interbank offered rate (TIBOR-Tokyo interbank offered rate)
(5) The daily return of the Nikkei 225 stock index
(6) The Nikkei 500 real estate daily return
(7) The conditional variance of a Nikkei 225 stock index returns determined using the GARCH (1,1) model represents market volatility.

For the macro state variables of United States, we follow them as stated in Fan et al. (2018).

(1) VIX, which measures the implied volatility in the market.
(2) The daily change in the 3-month Treasury constant maturities, which can be defined as the difference between the current day and the previous day of 3-month Treasury constant maturities.
(3) The change in the slope of the yield curve, which is defined by the difference between the 10-year Treasury constant maturities and the 3-month Treasury constant maturities.
(4) The change in the credit spread between 10-year BAA corporate bonds and the 10-year Treasury constant maturities.
(5) The daily S&P500 index returns.
(6) The daily Dow Jones U.S. Real Estate index returns.

All the bond-related data are extracted for the same time span as described previously and sourced from https://www.investing.com/Investing.com and Bloomberg terminal.

Setup

We apply the proposed methodology to the the ten Japanese banks where each of them receive tail risk from nine other banks. We build ten independent models. The network architecture of LassoNet is determined by three-fold cross-validation on g_W which includes the number of neurons in the first hidden layer and the hierarchical parameter M. The procedure is also followed for IMV-LSTM which consists of number of LSTM units. We also experiment with Gradient Boosting Machine (GBM) as processed by Long et al. (2022) and TENET (Tail-Event driven NETwork risk) as proposed by Härdle, Wang, and Yu (2016). GBM is optimized by five-fold grid search cross-validation for fine tuning the hyper-parameters as mentioned in the source work. The design of TENET is presented in details at the source paper which includes Minimum Average Contrast Estimation approach (MACE) with Smooth Clipped Absolute Deviation (SCAD) as regularization.

For connecting the Japanese banks with US banks, we plug the US macro variables in the Equation 5 and Equation 6 along the existing Japanese macro variables. Interestingly, the returns $X_t^{sys|-i}$ in Equation 5 become the returns of the US banks as the independent variable. Hence the new versions of equation can be written as:

$$CoVaR_t^i = \alpha^{\hat{sys}|i} + \gamma^{\hat{sys}|i} M_{t-1}^{(J)} + \delta^{\hat{sys}|i} M_{t-1}^{(US)} + \beta^{\hat{sys}|i} VaR_t^i \tag{22}$$

Where $M_{t-1}^{(US)}$ and $M_{t-1}^{(J)}$ the lagged macro variables of US and Japan respectively.

A comparative analysis between all the models is presented in Table 2. The losses presented are quantile loss (Equation 8) given by the trained models on the test set for each of the banks. It is seen that IMV-LSTM surpasses other models by a very large margin by the terms of average loss. Table 3 shows the tail risk associated among the banks using the IMV-LSTM model. It is in the form of adjacency matrix defined by Equation 21 and Table 4 shows the risk received by Japanese banks from the US banks. The top 10 highest interconnectedness are marked in bold for the both the cases.

Table 2. Comparative study of loss function by different models on test set while obtaining interconnectedness within the Japanese banks. In the cases of LassoNet and TENET, value is presented for pure quantile loss or unregularized loss.

Ticker	GBM	LassoNet	TENET	IMV-LSTM
8301.T	1.57	1.58661	0.92153	0.78995
8306.T	0.55334	0.59866	0.42556	0.55669
8316.T	0.55503	0.53268	0.57899	0.60761
8411.T	0.67964	0.66489	0.89123	0.69371
8309.T	1.23485	1.38546	0.9231	0.74236
8308.T	0.74382	0.85623	0.86441	0.70942
8304.T	0.94433	0.86548	1.02379	0.887
8331.T	0.74723	0.76312	0.75139	0.68413
8377.T	1.00089	0.95663	0.98734	0.48942
8410.T	1.03362	1.5794	0.85691	0.63823
Average	0.906275	0.978916	0.822425	**0.679782**

Further results

Tail risk analysis

We dive deep into the analysis of tail risk interconnectedness by looking at the Risk Received (RR) and Risk Transmitted (RT). They measure the amount of total risk taken in by the bank and total risk emitted to other banks in the whole financial system respectively.

$$RT_i = \sum_i D^{i|j} \tag{23}$$

$$RR_j = \sum_j D^{i|j} \tag{24}$$

The total connectedness (TC) is defined as the sum of risks along all the financial institutes within the network. It also acts as a system-level risk measure.

$$TC = \sum_i \sum_j D^{i|j} \tag{25}$$

Generally the institutes with high transmitted risk grab the attention from regulators. To visually represent the tail risk contribution among the banks we plot directed and weighted graphs. Figure 3 shows the network depicting the contribution. The width of the edges is directly proportional to the magnitude of the tail risk emitted. In Figure 4 we select a threshold of 0.018, which is greater than the 75% percentile of all the values in tail risk matrix. The risks which are less than the threshold are omitted out. We observe that MUFG, SMFG, HFG and MZF are the ones receiving and transmitting major amount of risks. All of them are top banking services in Japan and also leading global financial groups. Mitsubishi UFJ Financial Group (MUFG) is ranked as the largest bank in japan followed by Sumitomo Mitsui Financial Group (SMFG) in terms of assets and market capitalization. Hence, there exists a strong tail risk emission from SMFG to MUFG. These observations show that the results obtained by IMV-LSTM are intuitive and economically

Table 3. Tail risk contagion matrix obtained from IMV-LSTM model for interconnectedness within the Japanese banks. Each row depicts the tail risk received from corresponding bank in columns.

	BOJ	MUFG	SMFG	MZF	SMTH	RHC	AZB	CHB	HFG	SVB
BOJ	0	0.0193	0.0161	0.0205	0.0166	0.0162	0.0191	0.0161	0.0202	0.0202
MUFG	**0.0247**	0	0.0248	**0.0248**	**0.0249**	0.0222	0.0190	**0.0249**	0.0240	**0.0247**
SMFG	**0.0264**	0.0210	0	**0.0290**	0.0202	**0.0260**	0.0206	0.0188	0.0183	0.0220
MZF	0.0210	0.0231	0.0164	0	0.0226	0.0231	0.0155	0.0230	0.0214	0.0228
SMTH	0.0139	0.0134	0.0139	0.0133	0	0.0141	0.0147	0.0141	0.0135	0.0105
RHC	0.0189	0.0194	0.0183	0.0185	0.0186	0	0.0192	0.0182	0.0187	0.0193
AZB	0.0071	0.0081	0.0072	0.0085	0.0082	0.0122	0	0.0116	0.0080	0.0081
CHB	0.0218	0.0221	0.0222	0.0223	0.0220	0.0218	0.0218	0	0.0222	0.0221
HFG	0.0195	0.0154	0.0230	**0.0281**	0.0195	0.0201	**0.0277**	0.0196	0	0.0188
SVB	0.0150	0.0167	0.0145	0.0155	0.0178	0.0150	0.0171	0.0144	0.0156	0

Table 4. Tail risk contagion matrix obtained from IMV-LSTM for risk spread from US banks to Japanese banks. Each row depicts the tail risk received from corresponding US banks in columns.

	JPMC	BOA	WFGO	CITI	PNC	GSG	TFC	COFC	TDB	USB
BOJ	0.0145	0.0143	0.0157	0.0154	0.0143	0.0170	0.0171	0.0147	0.0142	0.0151
MUFG	0.0180	0.0172	0.0179	0.0171	0.0172	0.0175	0.0170	0.0166	0.0174	0.0173
SMFG	0.0167	0.0166	0.0167	0.0167	0.0166	0.0167	0.0167	0.0167	0.0167	0.0167
MZF	0.0130	0.0128	0.0127	0.0196	**0.0208**	0.0137	0.0124	0.0131	0.0131	**0.0207**
SMTH	0.0169	0.0168	0.0168	0.0169	0.0168	0.0168	0.0167	0.0168	0.0168	0.0168
RHC	**0.0216**	0.0061	**0.0181**	**0.0209**	**0.0207**	**0.0208**	**0.0207**	**0.0205**	**0.0208**	0.0155
AZB	0.0130	0.0148	0.0154	0.0161	0.0143	0.0137	0.0133	0.0145	0.0127	0.0199
CHB	0.0166	0.0164	0.0166	0.0164	0.0161	0.0162	0.0166	0.0160	0.0159	0.0163
HFG	0.0183	0.0185	0.0185	0.0182	0.0180	0.0185	0.0184	0.0186	0.0181	0.0182
SVB	0.0168	0.0169	0.0169	0.0168	0.0168	0.0172	0.0169	0.0171	0.0170	0.0171

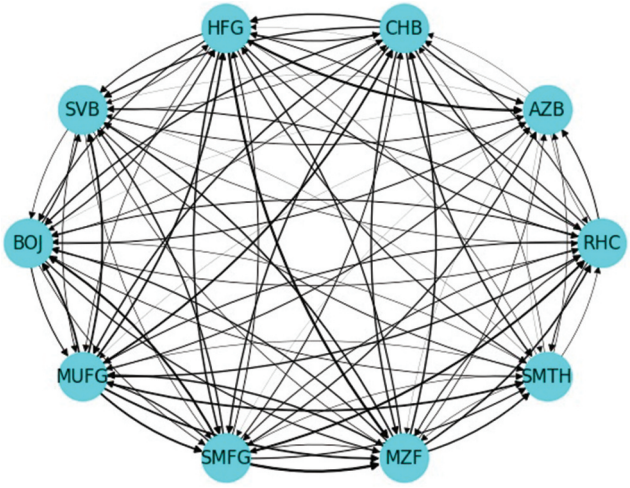

Figure 3. The directed network representation of the tail risk interconnectedness between the Japanese banks. Each node represents the listed banks and the width of the edge is directly proportional to the magnitude of risk emitted from one bank to another.

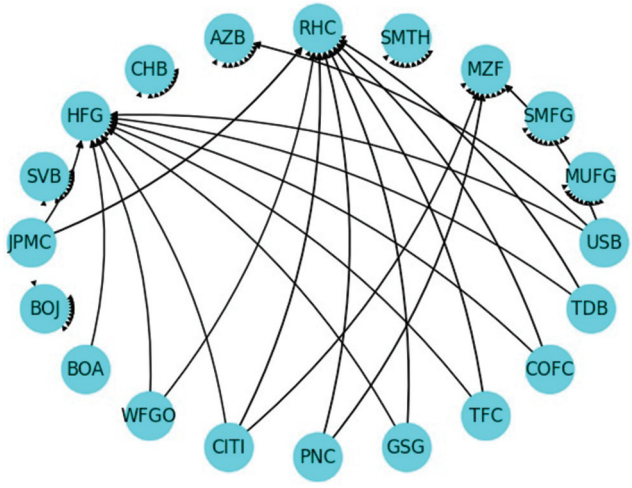

Figure 5. The directed network representation of tail risk contagion from US banks to Japanese banks corresponding to threshold = 0.018.

consistent. Figure 5 shows the risk received from US banks with the same selected threshold. Middle range banks in terms of total assets are more prone to receive tail risks from the overseas banks. Here we see that banks like Resona Holding Corporation (RHC) and HFG are the receptors of large risks from top US banks like JP Morgan and Chase, CITI Group, Wells Fargo and others.

Significance analysis

A descriptive analysis of the risk received and risk transmitted of the banks is presented in Table 5. This also contains the information of market capitalization (market cap), interbank asset ratio and interbank liability ratio for the listed banks. Basel (2013) states that market capitalization and interbank liability ratio are the important judgement parameters for interconnectedness especially when financial system is under distress. There exists a positive correlation between these indicators and the absorption and emission of risk. We perform a linear regression analysis of risk received and risk transmitted with market capitalization, asset ratio and liability ratio. Table 6 displays market information about the top US banks selected for analysis. Tables 7 and 8 show description of regression analysis.

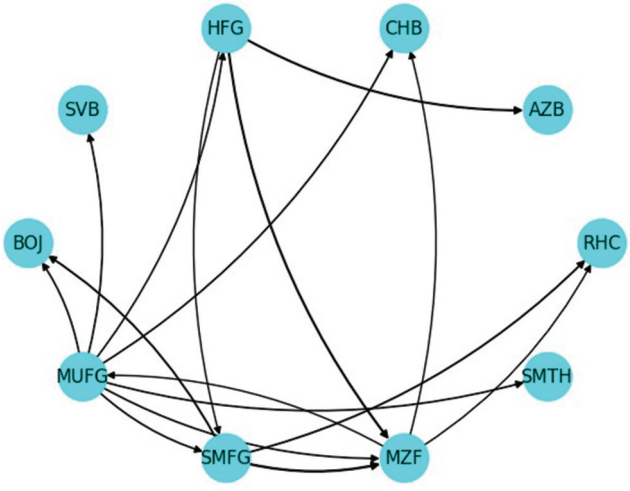

Figure 4. The directed network representation of tail risk contagion between Japanese banks corresponding to threshold = 0.018. The risks which are less than 0.018 are omitted out and the persisting strong connections are shown.

We observe that in both the cases (risk received and risk transmitted) market capitalization and interbank liability ratio have positive coefficient for regression proving they are positively correlated. This result is consistent with the propositions by Bassel committee. In case of Risk Received within the Japanese Banks we see that inter liability ratio is statistically significant from Table 7. In the case of risk received from overseas banks the coefficient of market cap is not significant however inter liability ratio is. For transmission of risk Table 8 depicts that both market capitalization and inter liability ratio are significant. The relation between the results obtained and the balance sheet variables supports our approach.

Table 5. The description of listed banks in terms of risk received, risk transmitted within the national and international domain. Also presents information on market capitalization and total assets (both in billion Japanese yen), inter asset ratio and inter liability ratio.

Bank	Market Cap	Asset Ratio	Liability Ratio	Total Assets	Risk Received (US)	Risk Transmitted	Risk Received
BOJ	30.85000	0.00284	0.16591	735.1166	0.16540	0.16430	0.16830
MUFG	18534.65740	0.00289	0.05798	386.780	0.15040	0.21400	0.15850
SMFG	11930.07504	0.00298	0.10865	270.4290	0.16530	0.20230	0.15640
MZF	7792.00628	0.00218	0.07709	254.256	0.17410	0.18890	0.18050
SMTH	2419.97840	0.00277	0.24802	690.230	0.17160	0.12140	0.17040
RHC	2295.29913	0.00214	0.07992	74.813	0.16810	0.16910	0.17070
AZB	286.09708	0.00121	0.09513	7.1840	0.16580	0.07900	0.17470
CHB	928.27950	0.00305	0.10820	19.788	0.16460	0.19830	0.16070
HFG	243.41369	0.00133	0.09397	16.173	0.16270	0.19170	0.16190
SVB	352.14291	0.01437	0.11046	1.3120	0.17360	0.14160	0.16850

Table 6. The description of listed banks of US in terms of market capitalization and total assets (both in billion US dollars), inter asset ratio and inter liability ratio.

Bank	Market Cap	Asset Ratio	Liability Ratio	Total Assets
JPMC	568.39782	0.01279	0.11264	3875.393
BOA	295.70172	0.00834	0.10512	3180.151
WFGO	201.82051	0.00991	0.11357	1932.468
CITI	119.38898	0.00383	0.13437	2411.834
PNC	62.31048	0.00993	0.12952	561.58
GSG	133.75381	0.00519	0.20300	1641.594
TFC	50.84788	0.00204	0.11150	535.349
COFC	54.17163	0.01021	0.10308	478.464
TDB	106.08992	0.00551	0.20532	1957.024
USB	67.92689	0.00818	0.09485	663.491

Table 7. Description of linear regression analysis showing how risk received is dependent on market capitalization, interbank asset (I. asset) ratio and interbank liability (I. liability) ratio.

Risk Received	Among Japanese Banks			From US banks		
Variables	coef	t value	p-value	coef	t value	p-value
Market cap	5.234e-06	1.740	0.125	5.018e-06	1.743	0.125
I. asset ratio	5.4301	1.040	0.333	5.8467	1.170	0.280
I. liability ratio	0.9384	4.253	0.004	0.9318	4.412	0.003

Table 8. Description of linear regression analysis showing how risk transmitted is dependent on market capitalization, interbank asset (I. asset) ratio and interbank liability (I. liability) ratio.

Risk Transmitted	Among Japanese Banks		
Variables	coef	t value	p-value
Market cap	9.059e-06	2.541	0.039
I. asset ratio	5.2514	0.848	0.424
I. liability ratio	0.7889	3.017	0.019

Rolling window analysis

To study the spillover effects among the institutes we execute a rolling window analysis of the tail risk interconnectedness. To effectively catch COVID-19 and the time periods of several other economic distresses we choose the rolling window size to be 250 days which is also the standard trading days in a year. Figure 6 exhibits the trend shown by total connectedness (Equation 25) within the Japanese banks over the years. A high magnitude of total connectedness is observed during the period of 2010–2011. This is the effect of great tsunami that hit Japan in 2011 due to earthquake. It deeply affected the whole financial system across the country. From 2016 onwards, the total connectedness went down, this is probably due to starting increment in taxes by government and negative interest rate started by Bank of Japan (BOJ). Starting after 2019, we again see a spike which can be simply accounted for the COVID-19 outbreak through out the world. This was the most catastrophic event badly affecting the markets and banking systems. Takahashi and Yamada (2021) showed that during COVID, Japanese firms which experienced international exposure showed less abnormal returns than which were not. This shows that the spillover effects are stronger within the national limits. From the mentioned observations, we conclude

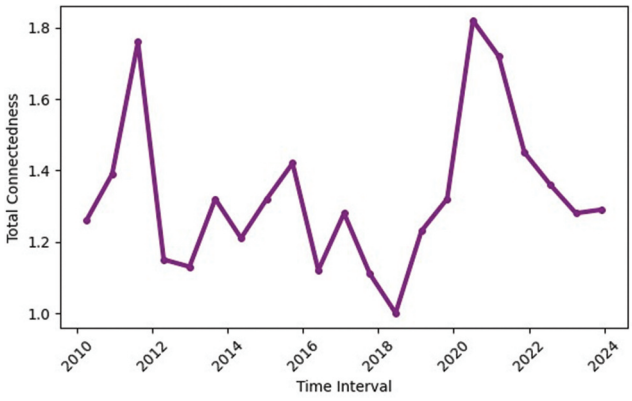

Figure 6. Total connectedness among the Japanese banks varying with 250 days time interval from the year 2010 to 2024.

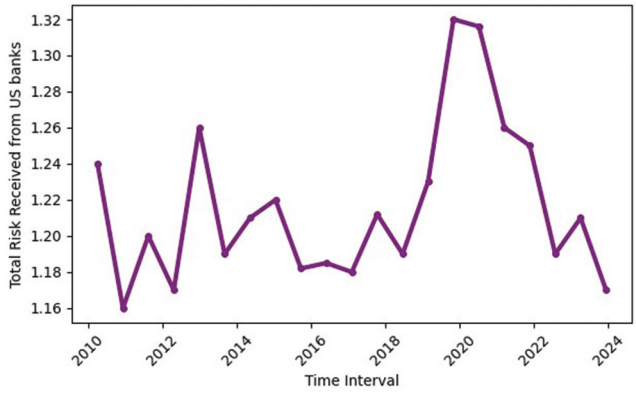

Figure 7. Total risk received from US banks to Japanese banks varying with 250 days time interval from the year 2010 to 2024.

that during the time of economic distress the total connectedness rises up and vice-versa. Skipping next to post-COVID period (after 2020) we observe sudden fall in the trend. All the economic activities resumed in normal fashion and hence system risk went down. Referring to Figure 7, we see that the total risk received by Japanese banks from the US banks also went up during the COVID-19 period. This was anticipated due to global impact of COVID to all the banks.

Banks need to be ready for a cascade effect in case of pandemics and other international calamities. As they falter, international financial institutions become more dangerous for one another. If banks have a greater understanding of how external shocks spread across the banking system, they will be more capable of managing these higher risks. This may mean preparing for an increase in loan defaults or increasing liquidity in unpredictable economic times. Banks and authorities can identify times of increased risk, such as during financial crises, when the overall risk throughout the system increases dramatically, by putting this rolling window strategy into practice.

Dynamic risk analysis

To assess the dynamic risk emitted and transmitted in the rolling windows, we look at some results in the form of heat-map to get visual estimation of the magnitude. Figures 9 and 8

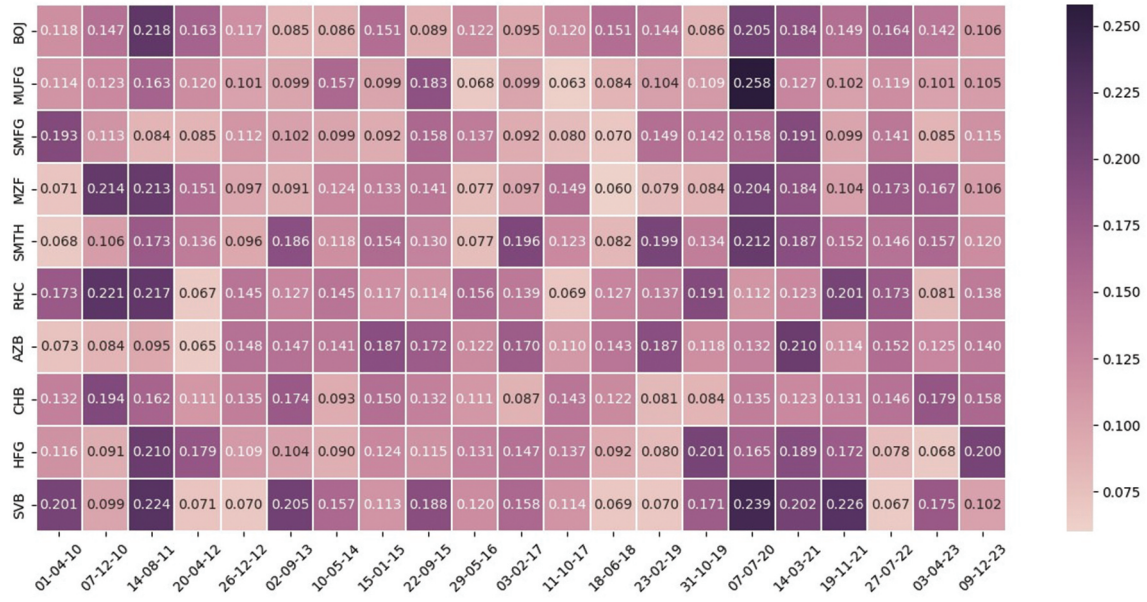

Figure 8. Risk transmitted by the Japanese banks to themselves over the course of years within 250 days interval.

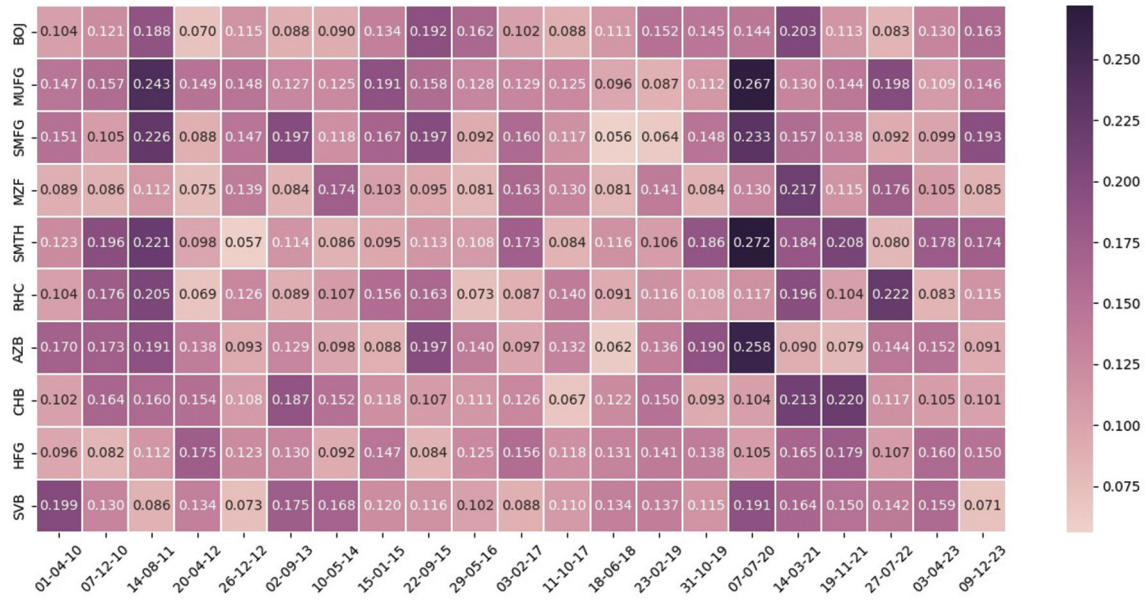

Figure 9. Risk received by the Japanese banks from themselves over the course of years within 250 days interval.

show the analysis of risk received and transmitted within the Japanese banks respectively. SMFG, MZF, BOJ and SVB are the highest risk transmitters during the 2011 period of natural calamity. We again observed high temperature values during the COVID-19 period. During that time MUFG transmitted the highest risk to other Japanese bank followed by SVB. There are specifically two high temperature regions through out the course of years as also described in the total connectedness plot. Figure 9 also depicts that banks receive high amount of risks during economic crisis. MUFG, SMTH and AZB are the banks which received the highest amount of risks during the 2019 to 2020 period. During the unfortunate event of tsunami MUFG and SMFG were the highest risk receivers.

IV. Conclusion

In this work we propose an interpretable LSTM-based approach using the model named IMV-LSTM to estimate the tail risk contagion among the Japanese banks and overseas banks at US. We primarily focus on the 5% quantile level of the left tail. IMV-LSTM works in multidimensional space and accounts for the potential linearity and non-linearity in the tail risk contagion metrics. An indicator $D^{i|j}$ is presented which is equally harnessed from the linear and non-linear settings of the model taking all the numerical dependencies like variable importance and rate of variation between the institutes and the system. Institutes with high tail risk connectedness are worth attention by the regulators. Gofman (2017) stated that such institutes can be too interconnected to fail. We find that institutes with large market capitalization are more prone to receive and transmit tail risk to other institutes. Large banks in Japan like MUFG, SMFG, SMTH, etc., are the ones which are highly susceptible to risks.

From the 2011 tsunami crisis we observe that MZF and RHC transmitted the high risk to others. Talking about the risk received from US banks, we conclude that global economic distress affects the banking systems due to good linkage due to which during the COVID period the risk reception from international banks soared high. Before this not much research was done on impact of COVID-19 on Japan's banking system. Our work depicted that MUFG received that largest risk during COVID times and results are intuitively and economically consistent.

Rare, severe events (tail risks) threaten big banks and the financial system. These risks quickly spread between connected banks. To prevent this 'contagion', banks need better risk management. This means anticipating and assessing risks, then using tools like reserves or diversification to reduce them. But focusing just on individual risks isn't enough.

By understanding how banks are connected, they can identify threats that could crash the entire system. This knowledge allows for targeted risk management, like reducing exposure to risky banks. In short, improved, interconnected risk management is vital for banks to protect themselves and the system from financial crises.

Regulators should consider how connected banks are, not just individual risk. Banks with high 'tail risk connectedness' (heavily reliant on borrowing and lending with others) are especially worrisome. These 'too interconnected to fail' banks could cause major financial crises if they collapse. Regulators need to focus on these connections and tailor rules to address them. For highly connected banks, this might mean requiring more capital reserves to handle unexpected events. Rising of total connectedness is a sign of financial distress. Accuracy in determination of tail risk interconnectedness is a very crucial aspect in financial networks of banks and deep learning based neural networks can do so.

In future, the inclusion of other financial factors (Dew-Becker 2022) and with better and advanced machine learning models results can be further improved. More explainable models can also help in formation of robust metrics. Currently, due to constraints on data availability on the Chinese banking system from our end, the study inhibits to the domestic Japanese financial system and its linking with the US banking systems. For using IMV-LSTM, the data needs to be presented in sequential form hence, the choice of appropriate sliding window still remains a hyper-parameter to be played around with. Given this deep learning approach, cross-validation methods for optimal neural architecture determination are time consuming. The variable importance as presented in IMV-LSTM is stochastic in nature unlike of tree-based methods of LASSO methods which are deterministic. This can be improved in future with deterministic models with accompanied with explainable nature fit for sequential data.

Disclosure statement

No potential conflict of interests were reported by any of the author(s).

References

Adrian, T., and M. K. Brunnermeier. 2016. "Covar." *The American Economic Review* 106:1705–1741. https://doi.org/10.1257/aer.20120555.

Ahelegbey, D. F., P. Giudici, and F. Mojtahedi. 2021. "Tail Risk Measurement in Crypto-Asset Markets." *International Review of Financial Analysis* 73:101604. https://doi.org/10.1016/j.irfa.2020.101604.

Andersen, T. G., V. Todorov, and M. Ubukata. 2021. "Tail Risk and Return Predictability for the Japanese Equity Market." *Journal of Econometrics* 222:344–363. https://doi.org/10.1016/j.jeconom.2020.07.005.

Basel. 2013. "Global Systemically Important Banks: Updated Methodology and the Higher Loss Absorbency Requirement." *Banking for International Settlements*. https://www.bis.org/publ/bcbs255.pdf.

Bhandari, H. N., B. Rimal, N. R. Pokhrel, R. Rimal, K. R. Dahal, and R. K. Khatri. 2022. "Predicting Stock Market Index Using Lstm." *Machine Learning with Applications* 9:100320. https://doi.org/10.1016/j.mlwa.2022.100320.

Chao, S. K., W. K. HÃ¤rdle, and W. Wang. 2012. "Quantile Regression in Risk Calibration." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2894219.

Davino, C., M. Furno, and D. Vistocco. 2014. *Quantile Regression: Theory and Applications*. Vol. 988. John Wiley & Sons. https://onlinelibrary.wiley.com/doi/book/10.1002/9781118752685.

Dew-Becker, I. 2022. Tail Risk in Production Networks. *Technical Report 30479*. National Bureau of Economic Research.

Fan, Y., W. W. Wolfgang Karl Härdle, L. Zhu, and L. Zhu. 2018. "Single-Index-Based Covar with Very High-Dimensional Covariates." *Journal of Business & Economic Statistics* 36 (2): 212–226. https://doi.org/10.1080/07350015.2016.1180990.

Giudici, P., T. Leach, and P. Pagnottoni. 2022. "Libra or Librae? Basket Based Stablecoins to Mitigate Foreign Exchange Volatility Spillovers." *Finance Research Letters* 44:102054. https://doi.org/10.1016/j.frl.2021.102054.

Giudici, P., and L. Parisi. 2018. "Corisk: Credit Risk Contagion with Correlation Network Models." *Risks* 6 (3): 95. https://doi.org/10.3390/risks6030095.

Gofman, M. 2017. "Efficiency and Stability of a Financial Architecture with Too-Interconnected-To-Fail Institutions." *Journal of Financial Economics* 124:113–146. https://doi.org/10.1016/j.jfineco.2016.12.009.

Guo, T., T. Lin, and N. Antulov-Fantulin. 2019. "Exploring Interpretable Lstm Neural Networks Over Multi-Variable Data." https://doi.org/10.48550/arXiv.1905.12034.

Härdle, W. K., W. Wang, and L. Yu. 2016. "Tenet: Tail-Event Driven Network Risk." *Journal of Econometrics* 192:499–513. https://doi.org/10.1016/j.jeconom.2016.02.013.

Hautsch, N., J. Schaumburg, and M. Schienle. 2014. "Financial Network Systemic Risk Contributions*." *Review of Finance* 19 (2): 685–738. https://doi.org/10.1093/rof/rfu010.

He, K., X. Zhang, S. Ren, and J. Sun. 2016. "Deep Residual Learning for Image Recognition." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. https://doi.org/10.1109/CVPR.2016.90.

Hochreiter, S., and J. Schmidhuber. 1997. "Long Short-Term Memory." *Neural Computation* 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Jordan, M. I. 1986. Serial Order: A Parallel Distributed Processing Approach. Technical Report, June 1985-March 1986. https://www.osti.gov/biblio/6910294 .

Lemhadri, I., F. Ruan, L. Abraham, and R. Tibshirani. 2021. "Lassonet: A Neural Network with Feature Sparsity." *Journal of Machine Learning Research* 22:1–29. http://jmlr.org/papers/v22/20-848.html .

Liu, L. 2014. "Extreme Downside Risk Spillover from the United States and Japan to Asia-Pacific Stock Markets." *International Review of Financial Analysis* 33:39–48. https://doi.org/10.1016/j.irfa.2013.07.009.

Long, Y., L. Zeng, J. Wang, X. Long, and L. Wu. 2022. "A Gradient Boosting Approach to Estimating Tail Risk Interconnectedness." *Applied Economics* 54:862–879. https://doi.org/10.1080/00036846.2021.1969002.

Maghyereh, A. I., and E. Yamani. 2022. "Does Bank Income Diversification Affect Systemic Risk: New Evidence from Dual Banking Systems." *Finance Research Letters* 47:102814. https://doi.org/10.1016/j.frl.2022.102814.

Nguyen, L. H., T. Chevapatrakul, and K. Yao. 2020. "Investigating Tail-Risk Dependence in the Cryptocurrency Markets: A LASSO Quantile Regression Approach." *Journal of Empirical Finance* 58:333–355. https://doi.org/10.1016/j.jempfin.2020.06.006.

Nguyen, L. H., and B. J. Lambe. 2021. "International Tail Risk Connectedness: Network and Determinants." *Journal of International Financial Markets, Institutions and Money* 72:101332. https://doi.org/10.1016/j.intfin.2021.101.

Raghu, M., B. Poole, J. Kleinberg, S. Ganguli, and J. Sohl-Dickstein. 2017. "On the Expressive Power of Deep Neural Networks." In *Proceedings of the 34th International Conference on Machine Learning*, edited by D. Precup and Y. W. Teh, 2847–2854. PMLR. https://proceedings.mlr.press/v70/raghu17a.html.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams. 1986. "Learning Internal Representations by Error Propagation." https://api.semanticscholar.org/CorpusID:62245742.

Takahashi, H., and K. Yamada. 2021. "When the Japanese Stock Market Meets Covid-19: Impact of Ownership, China and Us Exposure, and Esg Channels." *International Review of Financial Analysis* 74:101670. https://doi.org/10.1016/j.irfa.2021.101670.

Torri, G., R. Giacometti, and T. Tichý. 2021. "Network Tail Risk Estimation in the European Banking System." *Journal of Economic Dynamics and Control* 127:104125. https://doi.org/10.1016/j.jedc.2021.104125.

Yannis, D., B. Paul, and L. Sovan. 1995. "Use of Some Sensitivity Criteria for Choosing Networks with Good Generalization Ability." *Neural Processing Letters* 2 (6): 1–4. https://doi.org/10.1007/BF02309007.