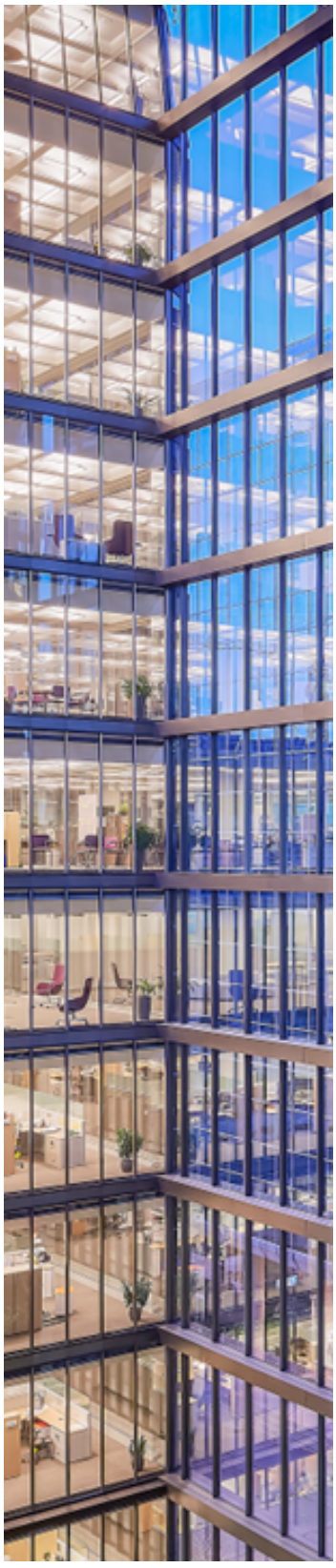
BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel – 2024-XX

Last Updated: June 25, 2024



# GLOBAL FINANCIAL TAIL SPILLOVERS

Author(s): Javier Ojea Ferreiro

Financial Stability Department
Bank of Canada
jojeaferreiro@bank-banque-canada.ca



DOI: https://doi.org/10.34989/san-2024-XX — ISSN 1914-0568

**Acknowledgements**

We would like to express our gratitude to Thibaut Duprey, Ruben Hipp and Gabriel Bruneau for their valuable comments and suggestions. Finally, I am indebted to Sascha Clazie-Thomson for his research assistance. The mistakes are, of course, my own. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Bank of Canada.

**Abstract**

This article proposes a methodology to compute international spillovers between a large set of financial institutions. We rely on a latent factor model to gather the most relevant dependence features with a score-driven dynamic to capture time evolution. The latent factor model decreases the prediction error and improves the goodness-of-fit over the standard factor model, i.e., a financial index factor. We propose a set of measures that allow us to track systemic risk and identify emerging stress scenarios.

***Topics***— systemic risk, financial stability, multivariate density forecasting, joint tail risk

***JEL Codes***— C02, C32, C58, G21

**Résumé**

Cet article propose une méthodologie pour calculer les retombées internationales entre un grand nombre d'institutions financières. Nous nous appuyons sur un modèle à facteurs latents pour rassembler les caractéristiques de dépendance les plus pertinentes et sur une dynamique axée sur les scores pour saisir l'évolution temporelle. Le modèle à facteurs latents réduit l'erreur de prédiction et améliore la qualité de l'analyse. l'erreur de prédiction et améliore la qualité de l'ajustement par rapport au modèle factoriel standard. modèle factoriel standard, c'est-à-dire le facteur de l'indice financier. Nous proposons un ensemble de mesures qui nous permettent de suivre le risque systémique et d'identifier les scénarios de stress émergents.

***Les sujets***— risque systémique, stabilité financière, prédiction de densité multivariée, risque de queue commun

***JEL Codes***— C02, C32, C58, G21

# 1 Introduction

Systemic risk can be defined as the possibility that occurrence of negative shocks, associated to single industries or firms, causes a financial market turmoil. The financial sector is potentially the greatest source of systemic risk. Since banks perform the function of inter-temporal and maturity transformation (i.e., financing long-term assets with short-term liabilities), the banking sector is vulnerable to contagion effect due to high leverage, growth of shadow banking, risk of confidence loss and the use of aggressive liquidity management strategies, i.e., high reliance on funding from the interbank market (ECB 2009). Other financial institutions can contribute to the financial instability by increasing the effect of fire sales as a consequence of following a procyclical behaviour, like mutual funds (Sydow et al. 2024), or might alleviate the impact from the financial system acting as a shock absorber, as have been the case of pension funds for some countries (Bédard-Pagé et al. 2016, Bruneau et al. 2023, ) Building a resilient financial sector able to prevent spread of a breakdown in the economic system is an important target for supervisory authorities, as irregular performance of the financial system could reduce the effectiveness of monetary policy, hampering the economic and financial well-being of the citizens ( Gravelle and Li 2013). Analytical tools enabling the timely identification of shocks and sources of risk that could lead to systemic events can help policymakers to assess the likely spread of individual problems in the financial system during crisis periods and to calibrate prudential instruments in tranquil times.

In this paper we develop a high-dimensional non-Gaussian modeling framework for financial stability assessment. We use a latent bifactorial model to capture national and international dependencies in financial markets and we model dynamics using a "generalized autoregressive score" (GAS) model. This approach allows us to overcome substantial difficulties in the econometric modeling of high-dimensional panels. We use the model to generate a set of systemic risk measures, with a special interest on the international spillovers and the sectoral decomposition of the systemic risk. Our framework captures the marginal, joint and conditional distribution of the variable of interest, which provide enough information to define which are the scenarios prone to generate the highest impact in the financial system with a certain probability of occurrence. It also allows us to decompose the tail returns of each institution, i.e. Expected Shortfall, into scenarios of joint distress with other financial firms.

We apply our general framework to weekly equity data from US ($N_{US} = 158$), Canada ($N_{CA} = 24$), Europe ($N_{EU} = 53$), Switzerland ($N_{CH} = 18$) and United Kingdom ($N_{UK} = 21$) from January 2006 to March 2024 ($T = 956$). Consequently, our database considers several crisis and periods of distress in the financial market. The 2008 global financial crisis, the sovereign European debt crisis, the COVID crisis, the UK gilt market crisis or the collapse of SVB (US) and Credit Suisse (CH) are examples of distress events that were materialized during the sample period.

Our study relates to the literature on high-dimensional tail dependence modeling as well as the literature on systemic risk measurement.
Regarding the first branch, we follow the factor copula introduced by Krupskii and Joe (2013) in a bifactor framework to take into account two sources of co-movement: domestic and international. In particular, we implement a latent nested bifactor model (Krupskii and Joe 2015). The factor model reduces the number of parameters to be estimated and the latent variable allow to capture the co-movement in a more flexible way. We also propose a test to check if the latent factor provides a better fitting of the data than the one-factor explicit model following a test similar to the likelihood ratio test. Pakel et al. (2021) follows the composite likelihood approach to use bivariate likelihoods to infer the full-dimensional structure of the data, overlooking the role of each variable in the full sample, which can be capture via the latent factor. Additionally, the latent factor model ensures the inversion of the large dimensional matrices. Engle and Kelly (2012) proposes the estimation of the correlation matrix assuming the same parameter for all the variables, i.e. equicorrelation structure. Lucas et al. (2017) employs the block equicorrelation structure to capture the dependency between Euro Area financial firms. The block equicorrelation structure could be seen as a restricted version of the latent one-factor model. Lucas et al. (2014) uses the hypersphere parameterization of the correlation matrix, where the number of parameters equals the number of correlation pairs. They employ this parameterization to model the sovereign CDS spread for 10 European countries. Although highly flexible, this approach might face the curse of dimensionality when the number of institutions becomes larger.Oh and Patton (2017) proposes a flexible linear factor model for modeling multivariate financial data and Oh and Patton (2018) employs this model to estimate the contagion risk from a set of 100 US firms using CDS log-change as a proxy of credit risk. Oh and Patton (2023) uses a cluster approach to associate firms into similar groups in bifactor model, where one factor is common and a second factor is group-specific. This is the paper closer to us in modeling the dependency in the high dimension framework. They estimate the group-specific factor via a cluster procedure, while we assume that the country defines the group-specific dependence. A second difference is the fact that they assume that both factors are independent, while we assume that the country-specific factor is nested within the global factor, which allow us to analyse how global connections spillovers into the domestic dependency structure.
Regarding the systemic risk literature, there are three main broad categories of classification depending on the data and tools employed. The first strand of research focuses on how shocks could propagate via balance sheet linkages, generating externalities between economic agents and amplifying losses (Bruneau et al. 2023, Sydow et al. 2024, Hałaj 2020, Hałaj and Hipp 2024). These studies usually rely on private or confidential data, which make them difficult to replicate. A second branch uses network theory to understand the structure of connections in the financial system and to get a better idea about the shocks that could trigger the collapse of the whole network (Chan-Lau et al. 2009, Cai et al. 2018, Blei and Ergashev 2014, Hałaj and Kok 2013). These exercises usually rely on simulations to unveil systemic connections not easy to identify due to the highly complex structure in network structure. The third strand uses information from market data to measure systemic risk. Our work belongs to this branch of the literature, where most of the measures in a one-factor model structure between an index of financial institutions and each institution (Adrian and Brunnermeier 2016, Acharya et al. 2012, Acharya et al. 2017), with multivariate models going beyong the bivariate case present scalability issues (Lehar 2005, Brownlees and Engle 2017, Segoviano and Goodhart 2010, Gravelle and Li 2013). In particular, Segoviano and Goodhart 2010 and Gravelle and Li 2013 are the studies dealing with systemic risk more similar to this piece of research. Segoviano and Goodhart (2010) uses a non-parametric copula to estimate a set of systemic risk indicators where the conditioning scenario could entail several financial firms, while Gravelle and Li (2013) relies on a semiparametric extreme value theory approach to estimate the contagion spillovers between Canadian, US, European and Asian financial firms. We also estimate the spillovers between global financial institutions following a parametric approach, which would allow us to consider changes in the estimates in a timely way by modeling the dynamics using a Generalized Autoregressive Score model (GAS). Also, we take into account the probability of occurrence when building stress scenarios involving several financial institutions. To our knowledge, most of the measures for systemic risk assumes the scenario of stress without taking into account the potential set of stress scenarios when building the risk indicator. One of the few exception is González-Rivera et al. (2021), where they estimate a quantile of the GDP-growth conditional to a scenarios (i.e. Growth-at-Risk measure), where the probability of the multivariate scenario conditions the risk measure. In other words, they are not only looking at the tail of the GDP-growth conditional distribution, but they are also selecting the scenario with a higher impact in terms of GDP-growth given a certain probability of materialization of the scenario. SImialrly, we proposed some modifications of systemic risk measures proposed by Segoviano and Goodhart (2010) to take into account the probability of occurrence of the different scenario, allowing us to track and identify emerging risks using this scenario-based analysis.

We found that the role of US is predominant in terms of effect on the systemic risk measures for the rest of jurisdictions, but the composition of the stress scenarios is changing over time. While US financial and commodity marker operator companies are present in the stress scenarios for other regions in the period 2008-2010, the banking sector lead the scenario composition in the period 2010-2018, while the US insurance sector start having more relevance in the design of stress scenario from 2018 onwards. Also we find different patterns, both

for conditional probabilities and returns, for the different jurisdictions, with a higher volatility of European systemic indicators compared to American systemic risk measures and some idiosyncratic patterns captured by some regions, like the Brexit event in the UK time series. Also, we find that the between 20 to 60% of the Expected Shortfall from the US financial system is happening at the same time that the stress scenario of Canada, which reflect the great symbiosis between these two jurisdictions. Canada also presents ties with Switzerland, being more than 40% of the Swiss Expected Shortfall is ocurring at the same time than tail losses in Canadian financial firms. This link could be explained by the role of insurance companies in two countries.

The findings of this research have implications for several economic agents. Firstly, for the supervisory authorities, who seek to monitor and track systemic risk and to identify scenarios that could generate large losses in the financial market. Secondly, policy makers, who wish to understand the interactions between financial system from different regions and how it could impact domestic economy. Thirdly, for risk management, investors and traders, who might find an interesting methodology to reduce portfolio's tail loss.

The article is laid out as follows: Section 2 present the methodology and the tail measures. Section 3 introduces the data and presents a descriptive analysis of our sample. Section 4 presents the results of the estimation and the spillovers across jurisdictions and sectors. Finally, section 5 provides the conclusion.

# 2 Methodology

## 2.1 Tail indicators

The definition of a crisis or stress scenario is not uniformly defined in the literature. While Acharya et al. (2012) consider a drop higher than 40% in the overall market capitalization over six months,Brownlees and Engle (2017) consider a decrease of 10-20% over one month period. At daily frequency, Acharya et al. (2012) establish an absolute threshold of daily losses bigger than 2%. Hautsch et al. (2015) indicates the threshold in relative terms as being below the 5% worst case scenarios while Adrian and Brunnermeier (2016) consider also the 1% worst case scenario. The methodology is flexible enough to allow the definition of the stress scenario in relative or absolute terms. We will present the threshold scenario as $\kappa$, which could be an absolute value, e.g. -10%, or a relative value, e.g. 5% worst case scenario for the returns of a specific firm. For the empirical application we assume that a firm is in a distress scenario if it is below its unconditional 5-th quantile, which would allow us to consider both marginal and dependence features when defining the distress region.

### 2.1.1 Index of joint distress

Segoviano and Goodhart (2010) proposes a systemic risk measure that looks into the probability of joint distress for the financial system ($JPoD$), i.e.

$$P(r_i < \kappa_i, \ldots, r_M < \kappa_M) = JPoD, \tag{1}$$

where $\kappa_j$ is the distress threshold for institution $j$ in a financial system of $M$ entities and $r_i$ is the variable of interest for entity $i$, which we will refer as returns hereinbelow. This measure, although informative of the general resilience of the financial system, might imply difficulties in the computation of the probability of joint distress in financial systems with a large number of institutions, as clusters of distress in a subset of institutions are hard to identify as a consequence of being faded out in the full financial network. Moreover, the threshold $\kappa$ cannot be understood aside of the joint probability, as the set of $\kappa$ values will condition the joint probability of distress. This makes difficult to understand what is driving the change in the $JPoD$ as, for instance, the decrease in the returns' volatility of one institution, might imply a lower probability of crossing the distress probability, affecting the overall probability of distress. Consequently, the increase or decrease in the $JPoD$ is not informative about the institutions driving that change. We should mention that not all the institutions or sectors have the same effects in the financial system, like systemic important institutions and small consumer lending entities, and we would like to adjust our measure to take into account the relative importance or the expert

advice on the institutions which could have a stronger effect in the market due to their size or funding connections.

We could consider which would be the maximum loss in an index made of financial institutions with a certain probability. We would like to point out which are the firms leading those losses, where we could combine the information obtained from the joint distribution with ad-hoc or prior information obtained from other sources (e.g. balance sheet information about leverage or liquidity, market capitalization, ...). We define the financial index of joint distress ($FIJD$) as,

$$\begin{aligned} FIJD(\alpha) &= \min_A E\left(\sum_{i=1}^{M} \omega_i r_i | A\right), \\ s.t. &\quad P(A) \geq \alpha \end{aligned} \tag{2}$$

where the scenario $A$ is defined by a distress scenario for a subset of firms and $\omega_i$ is the weight assigned to institution $i$.

Note that the solution to Eq. (2) is similar to the Expected Shortfall of the financial system using weights $\omega$, however, $IJD$ uses individual thresholds in the conditioning even while in the Expected Shortfall the conditioning event is defined on the convolution of firms' returns. Aside of this small difference, we could define the $FIJD$ as $FIJD = ES(A)$ where the probability of the conditioning scenario A $P(A) = \alpha$. We apply the decomposition of the Expected Shortfall in contributions from each institutions (Acharya et al. 2012, Banulescu and Dumitrescu 2015), i.e. $ES(A) = \sum_{i=1}^{M} \omega_i E(r_i|A)$.
This decomposition is extremely interesting when the set of institutions considered in the scenario $A$ and the entities with a weight $\omega$ higher than zero are different, e.g. if we consider the $JPoD$ for US institutions and we assign a weight $\omega$ of zero for the non-Canadian financial entities, we would get

(i) an estimation of the losses for the Canadian financial system given a distress scenario for the US financial system;

(ii) the US distress scenario with highest spillover into the Canadian financial system with a certain probability of occurrence;

(iii) the exposure of Canadian firms and their individual contribution to the overall losses.

### 2.1.2 Financial Spillover Index

Segoviano and Goodhart (2010) proposes a second indicator to assess the connectivity of the financial system. They define the Banking Stability Index ($BSI$) as the expected number of banks becoming distressed given that at least one bank has become distressed, i.e.

$$BSI = \frac{P(r_i \leq \kappa_i, \ldots, r_M \leq \kappa_M}{1 - P(r_i \geq \kappa_i, \ldots, r_M \geq \kappa_M)}, \tag{3}$$

where the distress scenario is given by the probability of being below the threshold $\kappa$. The numerator of Eq. (3) indicates the probability of each institution of being in distress and the denominator the probability of at least one institution becoming distressed. This measure could be seen as an indicator of banking linkage (Huang 1992). When the $BSI = 1$ in the limit, banking linkage is weak (asymptotic independence). The higher the $BSI$ is, the greater the banking linkage is (asymptotic dependence) implying a higher instability for the financial system.

Note that $P(r_i \leq \kappa_i)$ is within $P(A) = 1 - P(r_i \geq \kappa_i, \ldots, r_M \geq \kappa_M)$, i.e. firm i would be distressed if and only if at least one institution (firm i) is distressed. Consequently $P\left((r_i \leq \kappa_i) \cap A\right) = P(r_i \leq \kappa_i)$. We could see the $BSI$ as a weighted sum of the marginal probabilities, providing a decomposition of the contribution of each institution into the index. The $BSI$ could be easily generalized to a distress scenario where $N$ financial institutions are under stress. However, we are not controlling for the probability of occurrence of the conditioning event, so when dealing with heterogeneous financial systems with entities with diverse business model, a firm being in distress might not be enough to define a stressful scenario for the full network. For instance, let us suppose we have $M$ financial firms, of which $M - 1$ have perfect dependence and one institution has a negative perfect dependence with the remaining institution. In that case the $BSI$ would be the sum of the marginal probabilities of distress, as P(A)=1, without providing a good view of the dependence between institutions.

We proposed the following definition for a Financial Stability Index ($FSI$), i.e.

$$\begin{aligned} FSI &= \max_A \frac{\sum_{i=1}^{M} \mathbf{1}_{(r_i \leq \kappa_i) \cap A}}{P(A)} = \max_A \sum_{i=1}^{M} P(r_i \leq \kappa_i | A) \\ s.t. &\quad P(A), \geq \alpha \end{aligned} \tag{4}$$

where $\mathbf{1}_x$ is an indicator function that values 1 when the condition $x$ holds or zero otherwise. To distinguish the effect of the scenario from the evolution of the marginal probability of distress, we present in section 4 the results as a difference between the sum of the probabilities with and without scenario, i.e. $\sum_{i=1}^{M} P(r_i \leq \kappa_i|A) - P(r_i)$. This measure provides a couple of advantages with respect to BSI. First, the scenario conditioning our outcome is not set, like in the BSI, as "at least one firm being distressed", but the model would identify which is the worst scenario in terms of contagion of firms with a certain probability. Second, the measure should be risk averse, so the expected number of firms distressed would be less informative if we do not know the specific scenario or the probability of that scenario that we are using as conditioning event. Increases in the FSI could inform us, not only about the rise of distress in the financial system but it would give us a better view about the triggering institutions which could condition that escalation. Similarly to the $FIJD$, we could decompose the $FSI$ into contribution of each firm or subset of firms.

### 2.1.3 Vulnerability Index

Cortes et al. (2018) defines the Vulnerability Index ($VI$) as the sum of bivariate probability of stress between an institution $i$ and the rest of the sample, i.e.

$$\begin{aligned} VI_i &= P(r_i \leq \kappa_i, r_j \leq \kappa_j) + \cdots + P(r_i \leq \kappa_i, r_M \leq \kappa_M) \\ &= P(r_i \leq \kappa_i | r_j \leq \kappa_j) P(r_j \leq \kappa_j) + \cdots + P(r_i \leq \kappa_i | r_M \leq \kappa_M) P(r_M \leq \kappa_M), \end{aligned}$$

The contribution of firm $j$ to the vulnerability of firm $i$ ($CVI_{j\to i}$)is defined as

$$CVI_{j\to i} = \frac{P(r_i \leq \kappa_i, r_j \leq \kappa_j)}{VI_i}.$$

The main drawback of this measure if that if firm $i$ is independent from the remainder institutions the $CVI$ would give you a positive contribution, while it should be zero. Note that $P(r_i \leq \kappa_i, r_j \leq \kappa_j) = P(r_i \leq \kappa_i)P(r_j \leq \kappa_j)$ if firm $i$ and firm $j$ are independent $P(r_i \leq \kappa_i, r_j \leq \kappa_j) = P(r_i \leq \kappa_i)$ if the dependence is perfect and $P(r_i \leq \kappa_i, r_j \leq \kappa_j) = 0$ if the dependence is perfect and negative.

Our contribution measure should be centered around zero, with negative contribution when the dependence is negative and positive contribution when the dependence is positive. Hence, we redefine the vulnerability index as the deviation of the joint probability with respect to the independence case, i.e.

$$VI_i = \Delta_{j\to i} P\left(r_j \leq \kappa_j\right) + \cdots + \Delta_{M\to i} P\left(r_M \leq \kappa_M\right), \tag{5}$$

where $\Delta_{j\to i}$ is the change in the probability of distress of firm i if institution j is distressed, i.e. $P\left(r_i \leq \kappa_i | r_j \leq \kappa_j\right) - P\left(r_i \leq \kappa_i\right)$. The contribution to the vulnerability index would be redefined as

$$CVI_{j\to i} = \frac{\Delta_{j\to i} P(r_j \leq \kappa_j)}{VI_i}. \tag{6}$$

Under the new definition of the vulnerability index, the ratio between the $VI_i$ and the marginal probability of distress for institution $i$ provides the increase in probabilities of distress in the rest of institutions when firm $i$ is under stress, i.e.,

$$\frac{VI_i}{P(r_i \leq \kappa_i)} = \sum_{j\in L} \Delta_{i\to j} = \Delta_{*\to j} \quad \text{with } i \notin L.$$

### 2.1.4 Network of Expected Shortfall scenarios

The Expected Shortfall of a institution $i$ is defined as the mean return when the return is below a threshold $\kappa$, i.e.

$$ES_i = E(r_i | r_i \leq \kappa_i).$$

The Expected Shortfall could be divided into two component depending on a second entity $j$: 1) a section in which both firms are in distress and 2) another fraction where only institution $i$ is in distress, i.e.

$$ES_i = E(r_i | r_i \leq \kappa_i, r_j \leq \kappa_j) P(r_j \leq \kappa_j) + E(r_i | r_i \leq \kappa_i, r_j > \kappa_j)(1 - P(r_j \leq \kappa_j)).$$

The share $\gamma_{j\to i} = \frac{E(r_i | r_i \leq \kappa_i, r_j \leq \kappa_j) P(r_j \leq \kappa_j)}{ES_i}$ of the expected shortfall for entity $i$ is materialized at the same time than the stress scenario for institution $j$.
If we had a third institution we could continue the decomposition of the expected shortfall of entity $i$ in three portions.

$$\begin{aligned} ES_i &= E(r_i | r_i \leq \kappa_i, r_j \leq \kappa_j) P(r_j \leq \kappa_j) \\ &\quad + E(r_i | r_i \leq \kappa_i, r_j > \kappa_j, r_k \leq \kappa_k)(1 - P(r_j \leq \kappa_j)) P(r_k \leq \kappa_k | r_j > \kappa_j) \\ &\quad + E(r_i | r_i \leq \kappa_i, r_j > \kappa_j, r_k > \kappa_k)(1 - P(r_j \leq \kappa_j))(1 - P(r_k \leq \kappa_k | r_j > \kappa_j)). \end{aligned} \tag{7}$$

We can decompose the equation until reaching a certain number of institutions or a share of the expected shortfall, for instance more than 90%, i.e. $\sum_j \gamma_{j\to i} \geq 0.9$. The selection of the order is important and we start with that institution with higher probability of stress, so in Eq. (7) $P(r_j \leq \kappa_j) > P(r_k \leq \kappa_k)$. By doing this decomposition exercise for all the tail returns of the financial system, we can create a directed weighted network, where the arrow $j \to i$ indicates that the scenario for the expected shortfall of institution $i$ is the same than the stress scenario for entity $j$. The weight $\gamma_{j\to i}$ indicates how important is that scenario to explain the expected shortfall for firm $i$. From the network we can get a set of measure, like the weighted in-degree, which would indicate how important is each institution to explain the tail returns for the rest of the financial system.

## 2.2 Modelling distribution

The goal of building a set of systemic risk indicators implies the need of the modelling the joint distribution of some variables of interest (e.g. equity returns, log-changes in CDS spreads change,... ), which would act as a proxy of co-movements in fundamentals of entities in the financial system.

Several systemic measures like CoVaR (Adrian and Brunnermeier 2016), SRISK (Brownlees and Engle 2017), Co-Risk (Chan-Lau et al. 2009), Distressed Insurance Premium (DIP) (Huang et al. 2009), Banking Stability Indicator (Segoviano and Goodhart 2010) can be written in terms of conditional distribution or some transformation of the joint distribution. The computation of the systemic risk measures is straightforward provided that the joint distribution of the considered variables is well captured. The modelling of the distribution should consider several stylised features of the financial data. First, higher moments (skewness and kurtosis) can have a large impact on some tail measures like CoVaR or MES. Secondly, heterocedasticity and autocorrelation of the data would be captured via a flexible parametrization of the conditional mean and variance. Regarding the dependency structure, it should be flexible enough to capture tail joint dependence and asymmetries (e.g. higher correlation in downturn periods), and it should evolve over time to replicate changes in the dependence structure.

These characteristics should match with a set of restrictions when selecting the right model. The modelling should be simple enough for simulation and estimation, but realistic to capture the actual dependence pattern. We estimate the marginal distribution via a ARMA-GJR-GARCH with innovations distributed following the Hansen (1994)'s Skewed t distribution, which captures higher moments. The interdependencies between financial institutions are modelled via a Skewed-t copula structure, which reflect tail dependence and asymmetric relationships. To deal with the dimensional problem in large cross-section dataset, we assume a bifactoral latent factor structure with a country factor and a global factor, which would drive the dependence between financial entities. The bifactoral model is nested, in the sense than each set of firms in the same country relates with the domestic latent factor, while the country latent factors for different regions comove via the global factor. Additional, we add more flexibility to the model by allowing a time evolution of the correlation parameters via a Generalized Autoregressive Score approach. More details about the marginal and dependence modelling is provided in the next two subsections.

## 2.3 Marginal behaviour

We characterise the marginal densities of equity returns by an ARMA(p,q)-GARCH-GJR(h,k) model, i.e.

$$r_{i,t} = \mu_{i,t} + \epsilon_{i,t}, \tag{8}$$

$$\epsilon_{i,t} = \sigma_{i,t}\varepsilon_i, \tag{9}$$

$$\mu_{i,t} = \phi_0 + \sum_{j=1}^{p} \phi_j r_{i,t-j} + \sum_{k=1}^{q} \psi_k \epsilon_{i,t-k}, \tag{10}$$

$$\sigma^2_{i,t} = \omega + \sum_{l=1}^{k} \alpha_l \epsilon^2_{i,t-k} + \sum_{q=1}^{h} \beta_q \sigma_{i,t-q} + \delta \mathbf{1}_{\epsilon_{i,t-1}<0} \epsilon^2_{i,t-1} \tag{11}$$

where $\phi_j$ and $\psi_k$ are the parameters of the AR and MA components of the marginal model, $\omega$, $\alpha_l$, $\beta_q$ and $\delta$ are the components of the GJR-GARCH, which with the parameter $\delta$ allows for an leverage effect in the dynamics of the variance, implying a higher increase when there are negative shocks. The standardized innovation $\varepsilon_i$ follow a Hansen (1994)'s skewed t distribution, which captures the skewness and excess of kurtosis that we might find in financial returns. The density function is

$$f(\varepsilon_i; \lambda_i, \nu_i) = \begin{cases} bc\left(1 + \frac{1}{\nu_i - 2}\left(\frac{b\varepsilon_i + a}{1-\lambda_i}\right)^2\right)^{-\frac{\nu_i+1}{2}} & \text{for } \varepsilon_i < -\frac{a}{b} \\ bc\left(1 + \frac{1}{\nu_i - 2}\left(\frac{b\varepsilon_i + a}{1+\lambda_i}\right)^2\right)^{-\frac{\nu_i+1}{2}} & \text{for } \varepsilon_i \geq -\frac{a}{b} \end{cases}, \tag{12}$$

where $a = 4\lambda_i c(\frac{\nu_i - 2}{\nu_i - 1})$, $b = \sqrt{1 - 3\lambda_i^2 - a^2}$ and $c = \frac{\Gamma(\frac{\nu_i+1}{2})}{\sqrt{\pi(\nu_i-2)}\Gamma(\frac{\nu_i}{2})}$, the number of degrees of freedom $\nu_i$ must be higher than 2 and the parameter of asymmetry $\lambda_i$ could take a value between -1 and 1. This distribution converge to the Gaussian distribution when $\lambda_i = 0$ and $\nu_i \to \infty$ and the symmetric Student-t when the number of degrees of freedom are finite and $\lambda_i = 0$.

## 2.4 Dependence structure

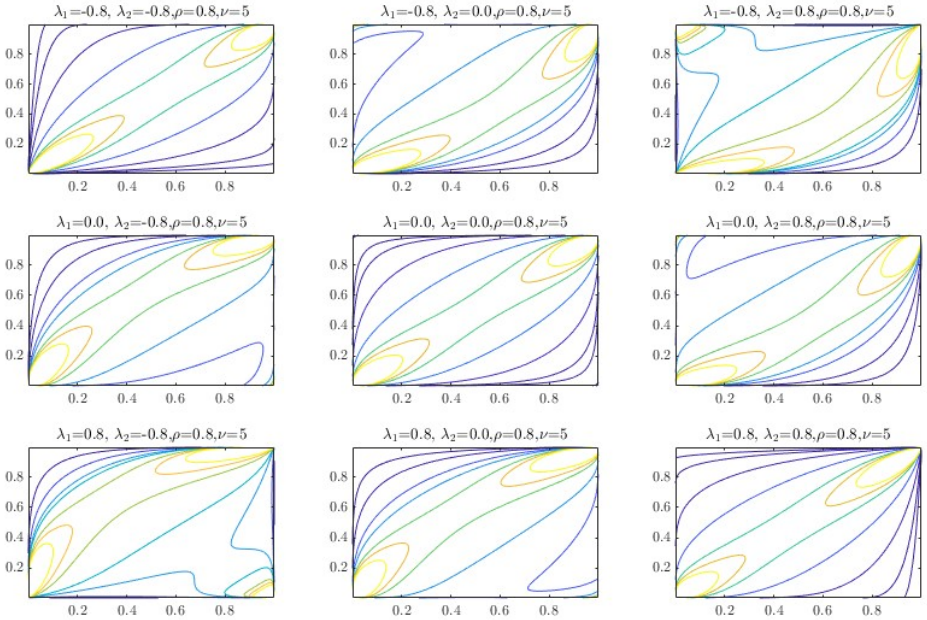
We rely on a copula methodology (Sklar 1959) to capture the joint distribution of the returns. The joint cumulative distribution function could be seem as a function build on the marginal cumulative distribution functions, i.e. $F(x, y) = C(F_x(x), F_y(y))$ Where the function that links the marginal cumulative distribution function is called copula $C(\dots)$. If we derive the joint cumulative distribution function with respect to its inputs, we get the density function, which could be express also in term of a copula, i.e. $f(x, y) = c(F_x(x), F_y(y))f_x(x)f_y(y)$. This formula becomes very useful to define conditional distributions. For instance, the conditional density of y given x would be: $f(y|x) = c(F_x(x), F_y(y))f_y(y)$ As a result of applying the Bayes' theorem on the right side of the equation. This highly simplifies the analysis of risk measures, as most of them are built in a conditional manner.

The copula approach also allows us to simplify the estimation in a two-step procedure (Joe and Xu 1996), where we first estimate the marginal features and them, the dependence structure is built based on the the pseudo-integral probability transformations of the standardized. More details about the estimation process are provided in the corresponding subsection.

### 2.4.1 The Skewed Student-t copula

We estimate the dependence between financial entities using a Skewed Student-t copula. The Skewed Student-t distribution has been widely employed to model financial and economic data (Lucas et al. 2014,Lucas et al. 2017, Oh and Patton 2023,Oh and Patton 2018) due to the flexibility to capture different tail behaviours as shown by Figures 1 and 2.

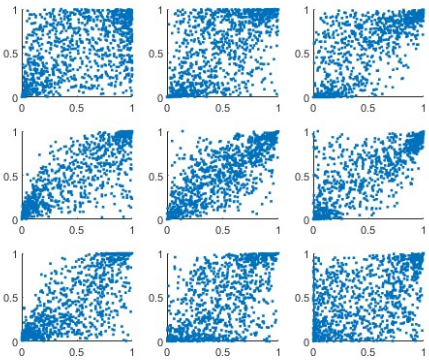
Figure 1: Copula density for different values of $\lambda_1$ and $\lambda_2$



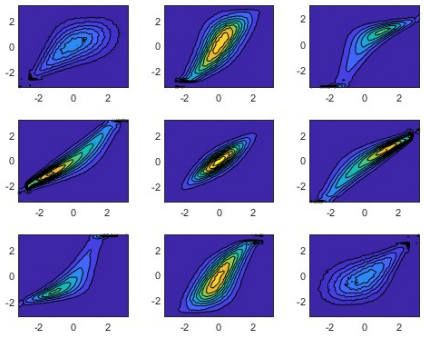

These figures show the density copula of a Skewed Student-t copula with $\rho = 0.8$, $\nu = 5$ and different values for the asymmetric parameters as shown above each picture

Figure 2: Simulation of a Skewed t copula

(a) Simulation of uniform variables with a Skewed-t dependence structure



(b) Isoquants of normal distributed variables with a Skewed-t dependence structure



These figures show the simulation from a Skewed Student-t copula with uniform marginals (Figure 2a) and Gaussian marginals (Figure 2b). The parameters selected for the Skewed-t copula are the same than in Figure 1.

The N-variate Skewed Student-t distribution discussed in Demarta and McNeil (2005) $ST(\mu, P, \lambda, \nu)$ has the following density distribution:

$$f_X(x) = c\frac{K_{\frac{\nu+N}{2}}(\sqrt{(\nu+d(x))\lambda' P^{-1}\lambda})\exp([x-\mu]' P^{-1}\lambda)}{(\sqrt{(\nu+d(x))\lambda' P^{-1}\lambda})^{-\frac{\nu+N}{2}}(1+\frac{d(x)}{\nu})^{\frac{\nu+N}{2}}}, \tag{13}$$

with $c = \frac{2^{-\nu/2}}{\Gamma(\frac{\nu}{2})\pi\nu|P|^{1/2}}$, $d(x) = [x-\mu]' P^{-1}[x-\mu]$ and $K_a(b)$ being the modified Bessel function of the second kind. The Skewed-t Student distribution becomes the Student-t distribution when $\lambda = 0$ and converges to the Gaussian distribution when $\lambda = 0$ and $\nu \to \infty$. Note that for the Skewed-t distribution $\nu > 4$ to have a defined variance, while the restriction for the Student-t distribution is $\nu > 2$.

The Skewed Student-t copula is a implicit copula (Smith 2023), in the sense that there is not an explicit formula for this dependence but it is defined as the ratio between the joint distribution and the marginal distributions. In other words, given the definition of the joint distribution as the product of density copula and marginal distributions, we defined the Skewed Student-t copula density as

$$c_{ST}(u,v) = \frac{f_{ST}(F_X^{-1}(u), F_Y^{-1}(v))}{f_X(F_X^{-1}(u)) f_Y(F_Y^{-1}(v))},$$

where $f_{ST}(x,y)$ is the Skewed-t bivariate distribution with parameters $\mu = [0,0]'$,$P = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, $\lambda = [\lambda_1, \lambda_2]'$ and $\nu$, $f_X(x)$ is an univariate Skewed-t distribution with parameters $\mu_1 = 0$, $P = 1$, $\lambda = \lambda_1$ and $\nu$ and $f_Y(y)$ is an univariate Skewed-t distribution with parameters $\mu_2 = 0$, $P = 1$, $\lambda = \lambda_2$ and $\nu$. Appendix presents more details about the Skewed-t copula, the assessment of conditional copulas and cumulative copulas, and some details about the estimation of the Skewed Student-t copula within our framework.

The estimation of a n-dimensional Skewed Student-t copula could become complicated for large n, as the number of parameters for correlation matrix are $\frac{N(N-1)}{2}$. In other to solve this problem, we take the simplifying assumption that the relationship between financial firms could be explained by a latent one-factor model, which reduces the number of parameters in the correlation matrix to n. Also, the joint copula can be built as a combination of bivariate copulas. Next subsection develops this modeling approach.

### 2.4.2 Nested factor copula

**Factor copula.** Let us assume that the joint behaviour of a matrix of $N$ variables is explained by distribution function with a correlation matrix $P$,i.e. $f(X;P)$. This model implies estimating $\frac{N(N-1)}{2}$ parameters for the

correlation matrix, i.e.

$$P = \begin{bmatrix} 1 & \rho_{1,2} & \dots & \rho_{1,N} \\ \rho_{1,2} & 1 & \dots & \rho_{2,N} \\ \dots & \dots\dots & \dots & \\ \rho_{1,N} & \rho_{2,N} & \dots & 1 \end{bmatrix}.$$

Let us assume that a one-factor model is able to explain the dynamics of the dataset, dividing the stochastic behaviour into a systematic part, driven by the factor, and a idiosyncratic part, e.g.

$$X_i = \rho_i Z + \epsilon_i.$$

Consequently, the correlation between variable $X_i$ and $X_j$ becomes $\rho_i \rho_j$, approximating the correlation matrix by:

$$\tilde{P}_Z = \begin{bmatrix} 1 & \rho_1\rho_2 & \dots & \rho_1\rho_N \\ \rho_1\rho_2 & 1 & \dots & \rho_2\rho_N \\ \dots & \dots\dots & \dots & \\ \rho_1\rho_N & \rho_2\rho_N & \dots & 1 \end{bmatrix}.$$

Note that to estimate this model, we would maximize the joint distribution conditioned on the factor, i.e. $f(X|Z;P) = \prod_{i=1}^{N} f(X_i|Z;\rho_i)$.

Hence, the estimation of $\tilde{P}_Z$ would depend on the factor $Z$ that is chosen to explain the dynamics of the matrix $X$. The best factor $Z^*$ to explain the joint behaviour of matrix $X$, would be that one that generates a correlation matrix $\tilde{P}^*$ for which $f(X;\tilde{P}^*) > f(X;\tilde{P}_Z)$ for any factor $Z \neq Z^*$. The key element explaining the joint behaviour is $\tilde{P}_{Z^*}$, so we could try to estimate it directly, without explicitly selection any particular factor if we make some assumptions about the distribution of the factor and the relationship between the factor and the matrix $X$, i.e.

$$f(X;\tilde{P}^*) = \int_{-\infty}^{\infty} \prod_{i=1}^{N} f(X_i|Z;\rho_i) f(Z) dZ.$$

Krupskii and Joe (2013) proposes this approach to define the factor copula, where under the assumption of a certain copula structure, the dependence structure could be estimated as

$$c(U;\tilde{P}^*) = \int_0^1 \prod_{i=1}^{N} c(u_i, v;\rho_i) dv, \tag{14}$$

where $c(\dots)$ is the density copula, $U$ is the matrix of integral distribution functions for matrix $X$, i.e. $U = F_X^{-1}(X)$ and $v$ is the latent factor driving the dependency between the variables in the dataset.
Similarly, we could obtain the copula as

$$C(U;\tilde{P}^*) = \int_0^1 \prod_{i=1}^{N} C(u_i|v;\rho_i) dv, \tag{15}$$

where $C(\dots|\dots)$ is the conditional copula

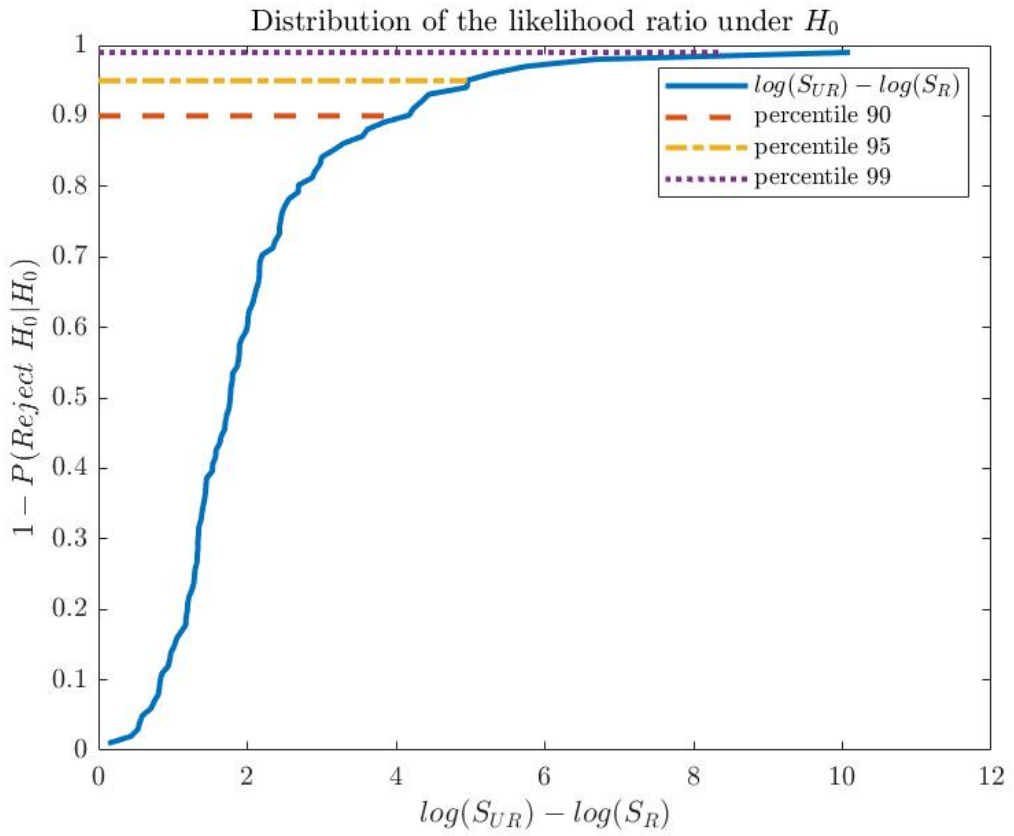
**Likelihood ratio test on the factor structure.** The presentation of the correlation matrix as a result of a factor model makes possible to distinguish a restricted model, where the factor is known, e.g. a financial index, and an unrestricted model, where the factor is unknown or latent. This distinction between restricted and unrestricted models reminds us about a likelihood ratio test. Actually, we could see the restricted model as a realization of the factor. The issue with this test is similar to what Cai (1994) faced to test a Switching Markov model with a model without structural change, i.e. the distribution function of the factor is not identified under the null hypothesis, which.makes the distribution of the likelihood ratio unknown under the null hypothesis, making the use of Monte Carlo techniques necessary to get this distribution. We apply the same idea to get the distribution under the null hypothesis, i.e. there is no significant increase in the goodness of fit of the latent factor, or in other words, the explicit factor is a realization of a stochastic variable with distribution $f(Z)$.

The MonteCarlo exercise relies in the following steps:

1. Using the explicit factor and the estimated correlation matrix simulate $w$ realizations of the matrix $X$, i.e. $\tilde{X}$.
2. Estimate the correlation matrix for each of the $w$ simulations of $X$ by using the latent factor as shown in Eq. (15). The likelihood of this model would be $S_{UR}^w$ for each simulation $w$.
3. Estimate the correlation matrix for each of the $w$ simulations of $X$ by using the explicit factor. The likelihood of this model would be $S_R^w$ for each simulation $w$.
4. Compute the log difference $log(S_{UR}) - log(S_R)$ for all the simulations, which would give us a distribution of the likelihood ratio test under the null hypothesis (see Chart 1).

5. Look at the corresponding value for a certain percentile of the distribution and compare it with the likelihood-ratio statistic from the real data. If the likelihood statistic is higher than the corresponding value for the selected percentile, we would reject the null hypothesis with a probability given by that percentile.

Chart 1: Likelihood ratio statistics distribution under the null hypothesis

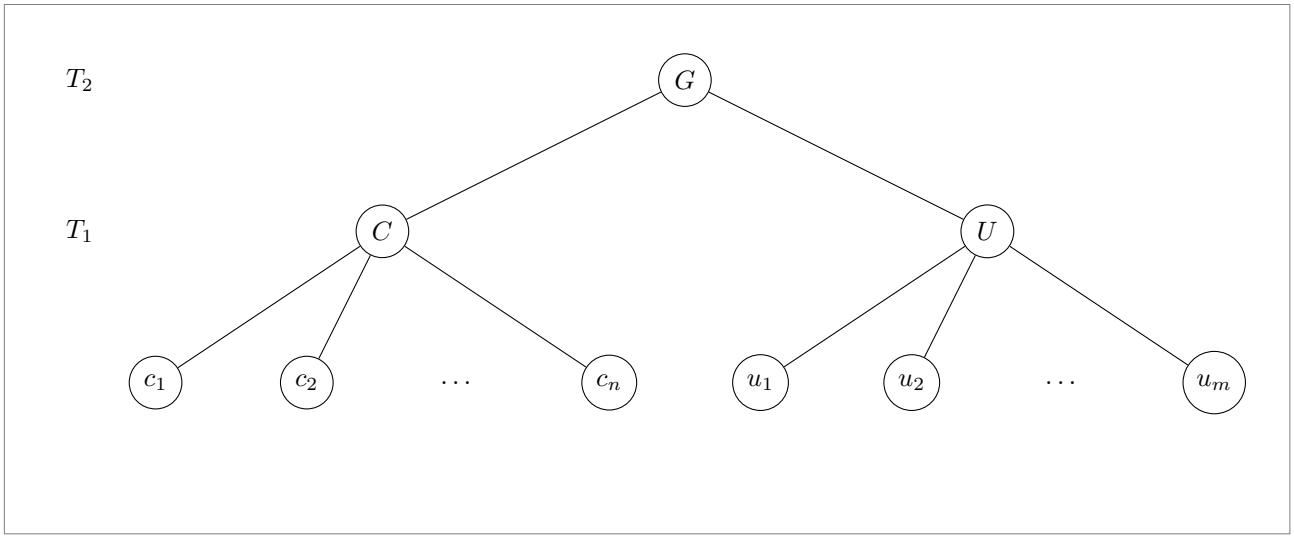



This chart shows the distribution of the likelihood ratio under the null hypothesis that the dependence structure estimated from an explicit one-factor model is a realization of the dependence structure from the latent factor. The chart is based on a simulation of 100 realizations for a dataset of length T=1000 and N=24 institution

**Nested bivariate factor structure** The structured factor copula is introduced by Krupskii and Joe (2015), where the dependence between variables in the bivariate case is explained by a common or global factor and a group-specific factor. Krupskii and Joe (2015) suggests two ways to generate dependence in non-homogeneous groups. The first approach would be assuming that the common factor and the group-specific factor are independents, which in the case of modeling dependence between financial firms might be difficult to hold. The second approach assume that the group-specific factors are correlated via the global factor. Probably the latter approach would be more realistic to replicate the financial network structure. Financial firms within the same country or region would be linked to a latent factor specific for that region, then the latent factors are linked between then via a global factor. The link with the global factor makes that stress periods in some regions would happen at the same time than stress scenario in other regions. In other words, the global latent factor proxies the current state of the international financial system, which drives other factors, reflecting the situation of domestic financial systems. These factors defines a dependence structure for the corresponding domestic financial entities.

Figure 3 shows an example for $n$ Canadian financial firms and $m$ US financial firms. Canadian firms are related to a Canadian latent factor $C$, while US firms are connected to a US latent factor $U$. Both latent factors are not independent as they are connected via a global latent factor $G$.

Figure 3: Hierarchical dependence structure of a nested bivariate copula model.




This figure shows the structure of hierarchical dependence for a nested bivariate copula. One that for variables in the bottom the dependence is always explained by a one-factor model (either factor $C$ or $U$) and the dependence between the variables with a different factor comes from the dependence between those factor via a second factor, i.e. global factor $G$, that drives the comovement between factor $C$ and factor $U$. The estimation process starts with the estimation of the dependence of the bottom layer $T_1$, followed by the dependence in the top layer $T_2$.

Eq. 15 could be rewritten to consider a set of $G$ independent group-specific factor, i.e.

$$C(U) = \prod_{g=1}^{G} \int_0^1 \prod_{i=1}^{N_g} C(u_i|v_g) dv_g,$$

where $N_g$ is the number of institutions in group $g$ and the factor copula becomes a product of factor copulas. The parameters $\rho$ and matrix $P$ of the copula from Eq. (15) are deleted here to simplify the notation.

If the group-specific factors are not independent we should take into account this relationship in a dependence structure. Hence, our previous equation becomes

$$C(U) = \int_{[0,1]^G} \prod_{g=1}^{G} \prod_{i=1}^{N_g} C(u_i|v_g) c_V(v_1 \ldots, v_G) dv_1 \ldots dv_g,$$

where $c_V(\ldots)$ is the density copula between the group-specific factors. As a simplifying assumption, the group-specific latent factors are conditional independent given a global latent factor. This is a key assumption to remove the G-dimensional integral from the previous equation, expressing the dependence structure into a double integral formula. Hence, our previous equation becomes

$$C(U) = \int_0^1 \left( \prod_{g=1}^{G} \int_0^1 \left( c_V(v_g, v_0) \prod_{i=1}^{N_g} C(u_i|v_g) \right) dv_g \right) dv_0, \tag{16}$$

where the inner parenthesis indicates the conditional probability given the specific-group factor and the global factor, while the outer parenthesis indicates the conditional probability given the global factor. The density copula would be

$$c(U) = \int_0^1 \left( \prod_{g=1}^{G} \int_0^1 \left( c_V(v_g, v_0) \prod_{i=1}^{N_g} c(u_i, v_g) \right) dv_g \right) dv_0. \tag{17}$$

Note that the structure shown in Figure 3 allows us to estimate first the group-specific copula in $T_1$, and then estimate the global copula $c_V(\ldots)$ in $T_2$, as the dependence between variables within the same group doesn't change because of the global copula, as the global copula just sets the dependence across different groups. The estimation approach subsection presents more details about the optimization process.

### 2.4.3 GAS dynamics

We consider that the parameters of correlation matrix in the Skewed Student-t copula are time-varying, following a dynamic described by a general autoregressive score (GAS) model (Creal et al. 2013). The score driven model reduces to many well-established models in financial econometrics, although capturing additional information in the data. Koopman et al. (2016) finds that GAS specification leads to similar predictive performance than state-space models, reducing the computational cost. GAS model outperforms autoregressive conditional models (e.g. GARCH).

The updating equation in the GAS model is

$$f_{i,t+1} = \omega_i + \alpha_i s_{i,t} + \beta_i f_{i,t}, \tag{18}$$

where $\omega_i$, $\alpha_i$ and $\beta_i$ are the parameters of the GAS model, $s_{i,t} = S_{i,t}\nabla_{i,t}$, $S_{i,t}$ is a scaling factor and $\nabla_{i,t}$ is the derivative of the loglikelihood function at time t with respect to the parameter $f_{i,t}$, i.e. $\nabla_{i,t} = \frac{\partial \log(c(u_{i,t},v_t;f_{i,t}))}{\partial f_{i,t}}$. Thus, $f_{i,t+1}$, is determined by an autoregressive updating function that has an innovation equal to the score of the log-likelihood with respect to $f_{i,t}$. We choose the scaling factor $S_{i,t} = \mathcal{J}_{t|t-1}$ where $\mathcal{J}_t\mathcal{J}'_{t|t-1} = \mathcal{I}_{t|t-1}$ and $\mathcal{I}_{t|t-1} = E_{t-1}(\nabla'_t\nabla_t)$. This scaling factor allows to standardize the value of $\nabla_t$ with its standard deviation. Another way to see this scaling factor is that we are using a steepest ascent method to update the value of $f_{i,t}$, where the direction is given by the gradient $\nabla_{i,t}$ and size of the step in that direction is given by the Hessian, which is approximated by outer product of the gradient. In other words, we use the local curvature of the log-density to improve the step.
The parameter $f_{i,t+1}$ is a transformation from the original parameter of correlation from the Skewed Student-t copula $\rho_{i,t}$. The transformation should increase the feasible values that the parameter could take, so we do not need to set some restrictions in the values of $\omega_i$, $\alpha_i$ and $\beta_i$. The transformation function $f_{i,t} = h(\rho_{i,t}) = -\log(\frac{1-\rho_{i,t}}{1+\rho_{i,t}})$ makes that the original feasible values in the range $(-1, 1)$ for $\rho_{i,t}$ increases to the real line under $h(\rho_{i,t})$. The score of the transformed parameters becomes

$$\frac{\partial \log(c(u_{i,t}, v_t; f_{i,t}))}{\partial f_{i,t}} = \frac{\partial \log(c(u_{i,t}, v_t; f_{i,t}))}{\partial \rho_{i,t}} \frac{\partial \rho_{i,t}}{\partial f_{i,t}},$$

where $\frac{\partial \rho_{i,t}}{\partial f_{i,t}} = \left(\frac{\partial h(\rho_{i,t})}{\partial \rho_{i,t}}\right)^{-1} = \frac{1-\rho_{i,t}^2}{2}$.

## 2.5 Estimation approach and computational details

We estimate the parameters of the model in two steps (Joe and Xu 1996): in a first step, parameters of the marginal distribution are estimated by maximum likelihood, and in a second step, copula parameters are estimated by maximum likelihood using pseudo-sample observations from the marginals as given by the integral probability transformations of standardized returns. In other words, the parameters of uniform margins are estimated at the first step and dependence parameters at the second step with parameters of the univariate margins fixed at the estimates obtained from the first step. The two-step estimation approach, also known as Inference for Margins (IMF) significantly reduces the computation time, simplifying the estimation process.

The integral from Eq. (14) and the double integral from Eq. (17) could be easily approximated using a Gauss-Legendre quadrature (Stroud et al. 1966) with a good precision using between 25 to 30 quadrature points. The Gauss–Legendre quadrature approximates the integral as a weighted combination of integrands evaluated at quadrature points, e.g., we could write Eq. (14) as

$$c(U; \tilde{P}^*) \approx \sum_{k=1}^{n_q} w_k \prod_{i=1}^{N} c(u_i, x_k; \rho_i),$$

where $x_k$ are the nodes, $w_k$ are the quadrature weights and $n_q$ is the number of quadrature points. Eq. (17) could be approximate as

$$c(U) = \sum_{k_1=1}^{n_q} w_{k_1} \prod_{g=1}^{G} \sum_{k_2=1}^{n_q} w_{k_2} c_V(x_{k_2}, x_{k_1}) \prod_{i=1}^{N_g} c(u_i, x_{k_2}).$$

An attractive property of Gauss–Legendre quadrature is that the same nodes and weights are used for different functions to compute the integral quickly and with a high precision. The same nodes also help in smooth numerical derivatives for numerical optimization.

We also follow an step-optimization approach where the dependence parameters of the copula model are estimated in steps. For our nested copula model, shown in Figure 3, parameters for the group-specific copulas

are estimated using data from the corresponding sector, as shown in the row $T_1$ from Figure 3. Within each group, the data is modeled via a latent one-factor model so the estimation is fast are stable, as shown by Krupskii and Joe (2015). Once, we have got the estimates for the group-specific copulas, we estimate the parameters of the global copula, stage $T_2$ in Figure 3, with the other parameters set equal to their estimates. This step-wise approach allow to get the estimates in a quick way, however we would need to rely on resampling methods to obtain the standard errors of the model, as this method would give us the standard errors conditional on the previous step. The resampling method would also allow us to get the multivariate distribution of the parameters, which would give us the chance to generate in-sample and out-of-sample forecast bands for the time-varying parameters using the methods shown by Blasques et al. (2016).

Some additional restrictions are made for the estimation of the Skewed Student-t copula at each stage of the process. First, we assume the same parameter $\lambda_1$, $\lambda_2$ and $\nu$ for all the variables modeled with the same factor. This simplification would speed up the estimation because we would need to simulate realization of the Skewed-t distribution, as there is not a closed formula for the inverse cumulative function of the univariate Skewed Student-t distribution. Also for each copula the number of parameters to estimate would be $N+3$, where $N$ are the number of variables modeled within the same factor. A second restriction is the use of a "variance targeting" approach to alleviate the complexity of the estimation process (Oh and Patton 2023). In particular, this approach imply using a two-step procedure where the estimates of the constant copula are estimated and, stationarity assumption, the GAS dynamics are estimated. The constant copula gives us the estimation of the asymmetry $\lambda$, the number of degrees of freedom $\nu$ and a vector of estimates $\rho$ that defines the long-term relationship between the variables and the latent factor. Similarly to the variance targeting in the GARCH literature, where the constant term $\omega$ is estimated based on the long-run variance, in the GAS approach, under the stationary assumption, the unconditional expectation of the loadings in Eq. (18) is $\bar{f}_i = \frac{\omega_i}{1-\beta_i}$ as the expectation of the score is zero (Blasques et al. 2022). The paraemters $\alpha_i$ and $\beta_i$ are assumed to be equal to all the variables linked with the same factor, to further simplify the estimation process.

For the numerical optimization of the the constant copula, we used a modified Newton-Raphson algorithm, for which the first and second partial derivatives of the copula density with respect of the parameter vector are needed. This method allows us to get good estimates for a large number of variables and parameters if we are starting from a sensible starting point. The speed in convergence could increase if we obtain analytical expressions for the gradient and the Hessian. Using the differentiation under the integral sign, we can build the first and second derivatives with respect of the parameter vector. The appendix provides the analytical expressions of the gradient and Hessian for the Skewed Student-t copula and some details on how this derivatives would match in a factor copula framework.

# 3 Data

We download from Refinitiv EIKON equity stock prices with dividends being reinvested in the same assets, i.e. total return price, for financial firms in Canada, US, Europe (EA), Switzerland and United Kingdom.

The data has been filtered to keep only those assets that are traded enough during the period 2006-2024, as it would allow to analyse the response to several periods of stress like the global financial crisis, the European sovereign debt crisis and the COVID crisis. Also, subsectors like mutual funds, hedge funds, trust and venture capital are deleted from the sample as they have a focused investment strategy (e.g. health care sector, high-tech,...). Finally, only firms for which we can get balance-sheet information is considered.

The dataset is composed by 274 financial firms, which could be decomposed in 158 US firms, 53 EA firms, 24 Canadian firms, 21 UK firms and 18 Swiss firms. We present below some summary statistics, where the market capitalization and number of institutions is disclosed by country and subsector. The subsector classification follows the Refinitiv Business Classification (TRBC).

Table 1: Distribution of the number of financial firms by region and subsector

| # | Canada | US | UK | Europe | Switzerland |
|---|---|---|---|---|---|
| Banks | 9 | 118 | 6 | 32 | 8 |
| Insurance | 9 | 28 | 4 | 15 | 6 |
| Financial & Commodity Market Operators & Service Providers | 1 | 4 | 1 | – | 1 |
| Consumer Lending | 3 | 3 | – | – | – |
| Investment Banking & Investment Services | 2 | 5 | 10 | 5 | 3 |
| Corporate Financial Services | – | 1 | – | 1 | – |

Table 2: Distribution of the size of financial firms by region and subsector

| Million CAD | Canada | US | UK | Europe | Switzerland |
|---|---|---|---|---|---|
| Banks | 620273 | 2249974 | 417722 | 1045503 | 38102 |
| Insurance | 258321 | 907748 | 43949 | 559402 | 201209 |
| Financial & Commodity Market Operators & Service Providers | 10046 | 266189 | 84473 | – | 6234 |
| Consumer Lending | 3870 | 313874 | – | – | – |
| Investment Banking & Investment Services | 941 | 677175 | 117564 | 58457 | 145444 |
| Corporate Financial Services | – | 1526 | – | 1524 | – |
| Exchange Rate | | USDCAD | GBPCAD | EURCAD | CHFCAD |
| | | 1.3672 | 1.717 | 1.476 | 1.5074 |

Note: The table presents the market capitalization in millions of Canadian dollars at the end of sample, using the corresponding exchange rate.

# 4 Results

We present in this section some results of the four systemic risk measures aggregated at the region level. We provided in the appendix a pseudo-code to simulate the systemic risk measures of $FSI$ and $FIJD$, but also we show the copula formulas to obtain the four types of systemic risk measures proposed in Section 2. We have also performed the Likelihood Ratio type test proposed in section 2 for each region, using the static copula, where the null hypothesis has been always rejected at 99% confidence level.

## 4.1 Financial Spillover Index

The Financial Spillover indicate the difference between the sum of the conditional probabilities and the sum of the marginal probabilities of distress. The conditioning scenario is a joint stress scenario for a set of firms in another region with a probability equal or higher than 4%.

We observe different patterns in the five regions. Canada, shown in the top left subplot in Chart 2, has a peak in the period 2016-2018 although the series are quite stable between a $FSI$ 15-20. The maximum possible FSI would be 24 (the number of institutions in the sample), so these are high numbers which are decreasing after 2018, being at the end of the sample close to 10. The US financial sector is the region with the highest impact on Canadian $FSI$ with the EU being the one with a lowest impact. With respect to the role of Canadian institutions as a conditioning scenario for the FSI of other regions, Chart 3 shows that up to six institutions under stress during the global financial crisis (GFC) and 4-5 institutions during the COVID crisis. Banking Services and Investment Banking and Investment Services are the two subsectors that defines the stress scenario for most of the period, with the insurance sector being the less representative for the stress scenario in the FSI.

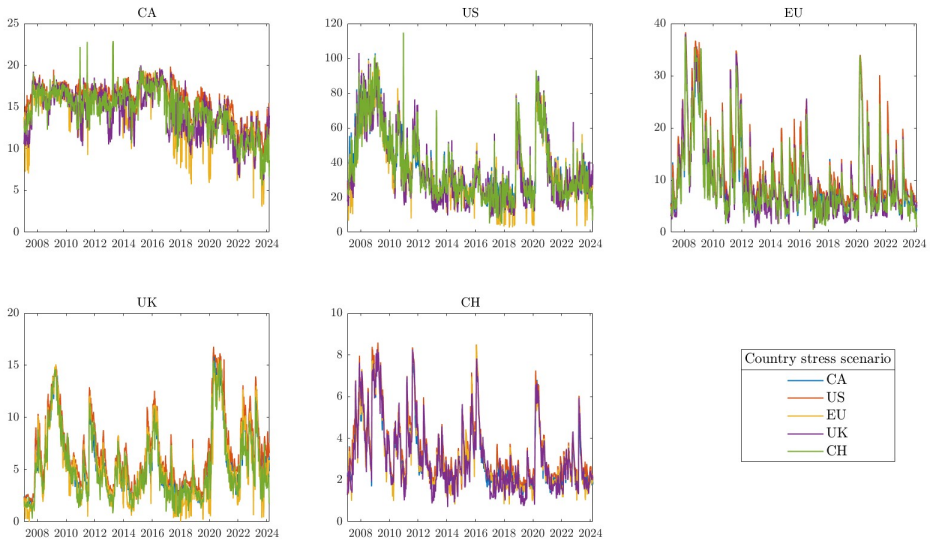
US FSI, shown the top center subplot in Chart 2, has two peaks corresponding to the GFC and the COVID crisis with maximums around 80-100 and minimum around 20. The number of institutions are 158, which indicate a very low FSI compared to the size of the financial system. The lowest exposure of US is to European financial system and the highest exposure is to UK and Switzerland. Regarding the role of US financial firms in the stress scenario, shown in Chart 4, the relevance of the different subsectors change over time. While in the period of 2008-2010 the Commodity Market Operator (in purple) has the highest relevance in the design of the stress scenario, while for the period 2010-2018 the banking sector is leading the stress scenarios for the FSI of other regions. Finally, starting from 2018, the insurance (in green) has a higher relevance when designing the stress scenario.

European FSI presents a much more volatile pattern than Canada and US with the highest exposure to US. The time series shows a maximum of 40 a minimum around 5. The number of European institutions in our sample are 53, so the average degree of distress shown by the FSI is lower than US and Canada. Interesting, we have some peaks that are not observable in other series. For instance there is a peak in 2012 coinciding with the European sovereign debt crisis. The role of European financial firms as stress scenario for the FSI in other regions is focused on banking and insurance companies, as shown by Chart 5.

UK FSI shows the highest exposure to a stress scenario in the US financial sector but the lowest exposure to a distress event for EU financial firms. The path of the indicator presents several peaks, coinciding with crisis and political events. For instance the peak in 2016 coincides with the Brexit pool. The range of values in the FSI for UK goes from 2 to 15, which represents a large gap for a set of 21 institutions. Banks, Commodity Market Operators and, in a third position, the insurance sector are the UK financial sectors that provides the highest impact on the FSI of other countries according the Chart 6.

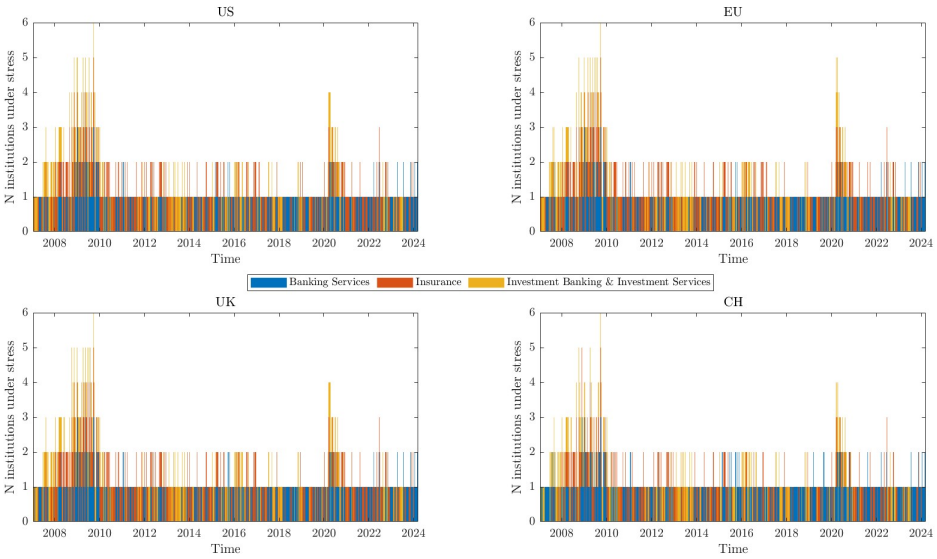
Finally, Swiss FSI shows the lowest value in comparison to its size. It shows a maximum at 8, which with a set of 18 institutions shows a stress below 50%. The European stress scenario seems to generate highest impact, together with the US, to the Swiss FSI. Chart 7 indicate that the banking sector, followed by the insurance sector, are the Swiss financial sectors with highest impact on the remainder financial regions.

Chart 2: Financial Spillover Index



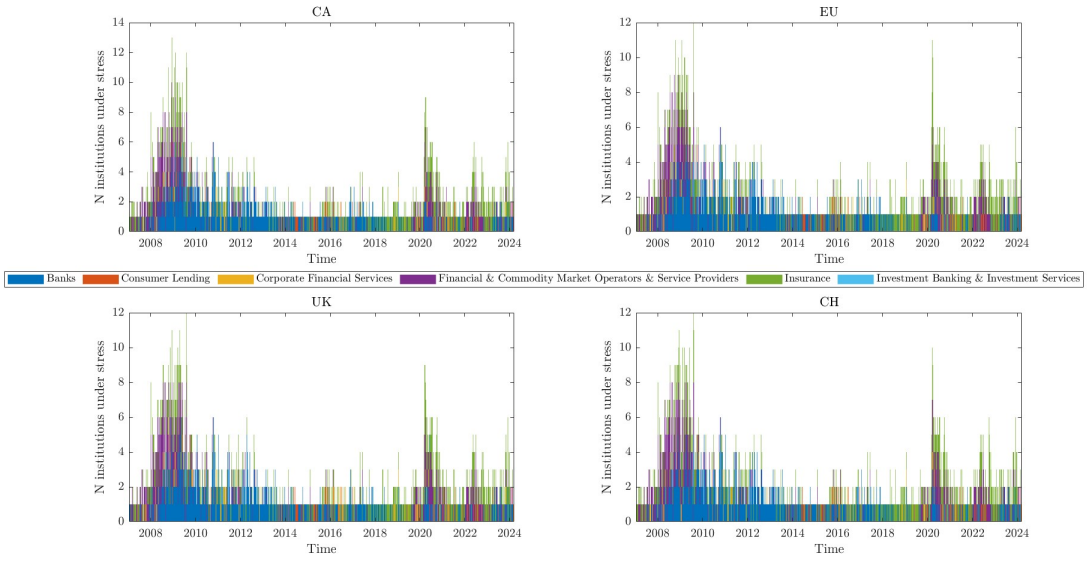

This chart shows the Financial Spillover Index as the sum of the difference between the conditional and unconditional probability of distress. We have set a $\alpha$=0.04, while the threshold $\kappa$ that defines the distress scenario is set at the empirical 5-th percentile of the unconditional marginal distribution. Each subplot indicates the FSI for the country specify in the title, while the countries from which the scenario $A$ is generated appear in the legend. Charts 3, 4, 5, 6 and 7 presents the scenarios disclosed by the subsector. Measures have been simulated using 100,000 realizations.

Chart 3: Distress scenario for Canadian firms in the $FSI$ measure



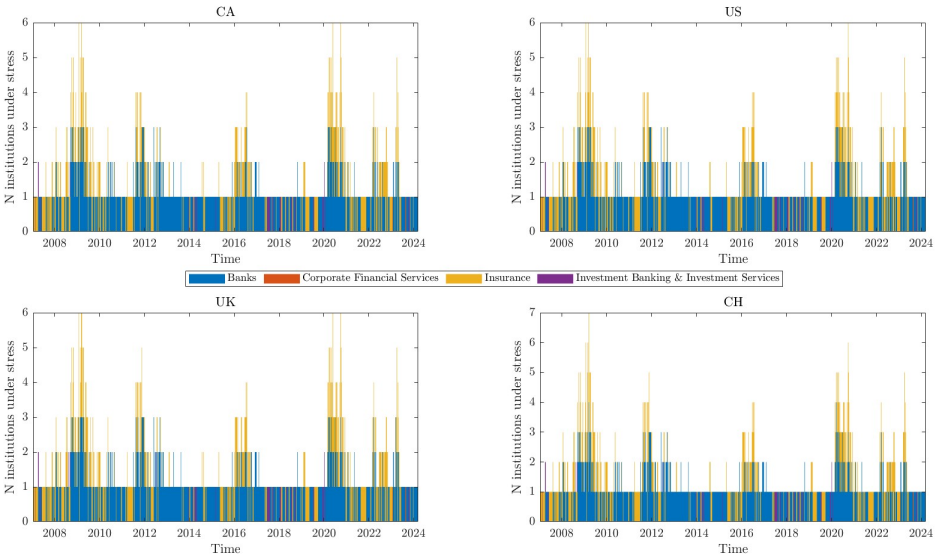

This chart shows the distress scenario for Canadian firms that generates the highest $FSI$ for each of the countries indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

Chart 4: Distress scenario for US firms in the $FSI$ measure



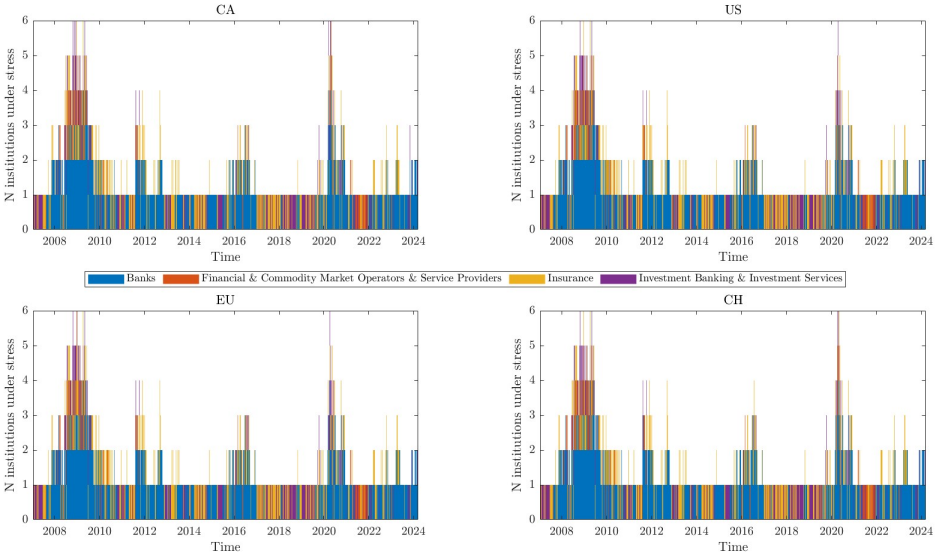

This chart shows the distress scenario for US firms that generates the highest $FSI$ for each of the countries indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

Chart 5: Distress scenario for European firms in the $FSI$ measure



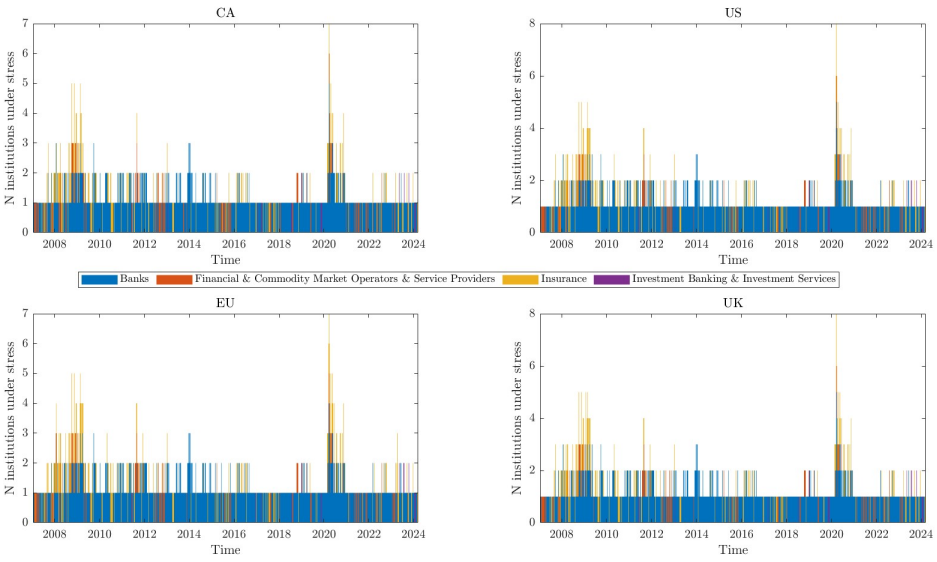

This chart shows the distress scenario for EU firms that generates the highest $FSI$ for each of the countries indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

Chart 6: Distress scenario for British firms in the $FSI$ measure




This chart shows the distress scenario for UK firms that generates the highest $FSI$ for each of the countries indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

Chart 7: Distress scenario for Swiss firms in the $FSI$ measure



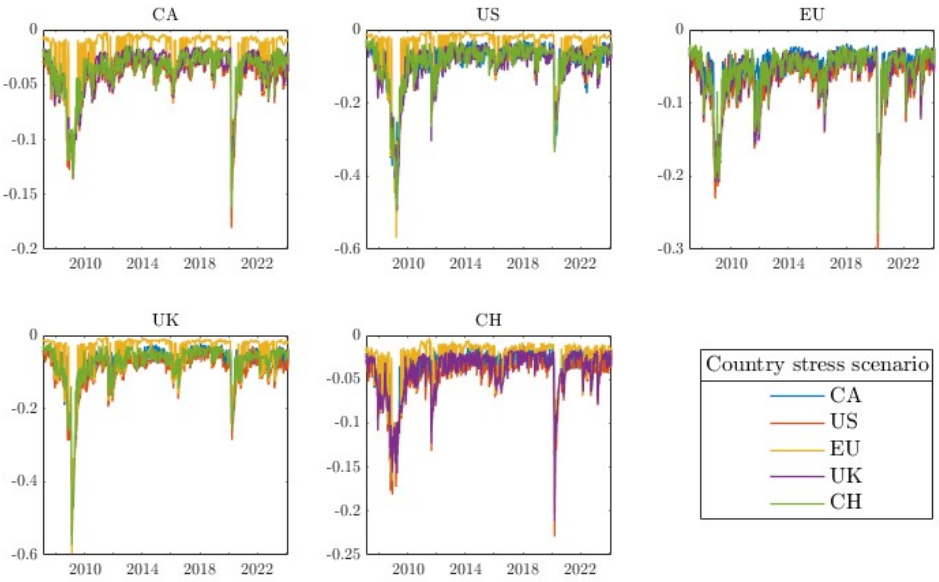

This chart shows the distress scenario for Swiss firms that generates the highest $FSI$ for each of the countries indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

## 4.2 Financial Index Of Joint Distress

The financial index of joint distress (FIJD), shown in Chart 8, is built using the market capitalization of each company as a weighting factor of the index, with the returns shown in the domestic currency, so there is no FX noisy the the pattern. Apart from the drops in 2008 and 2020, coinciding with the financial and health crisis, we observe a drop in 2012-2013 for US, Europe and Switzerland, and a drop for Europe in 2016-2017. Note that the pattern, and the stress scenarios, might be different than the ones shown by the FSI, as here we are using the market capitalization to weight the relevance of each company within the financial system. The region with lowest influence in the global financial system seems to be Europe, followed by Canada, whereas US present the highest effect in the performance of the index for the rest of regions. US financial index seems to perform worse when the stress scenario is built on the UK financial sector.

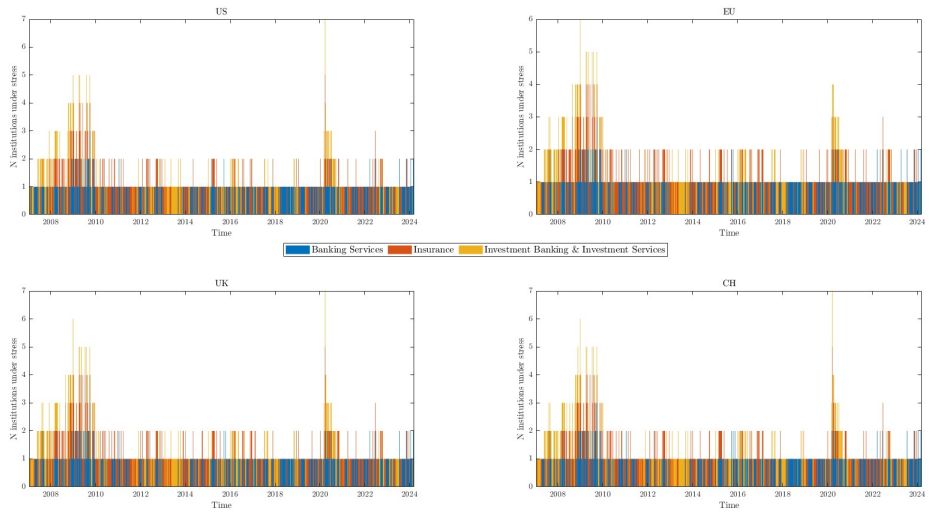
Concerning the stress scenarios, they are different from the scenarios shown for the previous measure as the large size of the banking and insurance sector would put a higher weight in generating scenarios with a worse outcome for them. For instance, in Canada, we see that the set of institution under stress for the US FIDJ shown in Chart 9 is smaller during the GFC than in Chart 3. Also, the number of institutions stressed peaks 7 for the COVID crisis for the FIDJ of all the regions. Swiss stress scenario for the Canadian FIDJ, shown in Chart 13, diminishes the relevance of Commodity Market Operators with respect to banking and insurance sectors, which have a higher relevance for Canada than in Chart 13.

Chart 8: Financial Index of Joint Distress



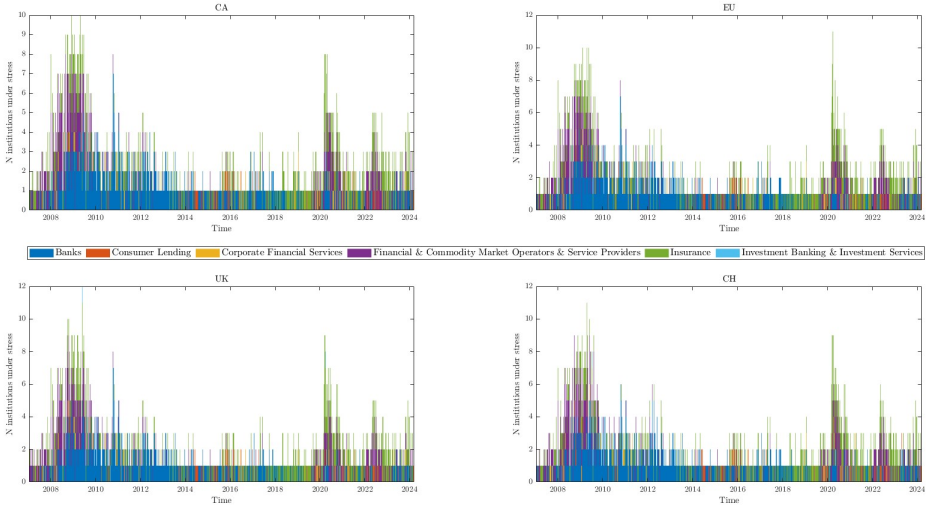

This chart shows the Financial Index of Joint Distress using the market capitalization at each time as weighting factor. We have set a $\alpha$=0.04, while the threshold $\kappa$ that defines the distress scenario is set at the empirical 5-th percentile of the unconditional marginal distribution.Each subplot indicates the FIJD for the country specify in the title, while the countries from which the scenario $A$ is generated appear in the legend. Charts 9, 10, 11, 12 and 13 presents the scenarios disclosed by the subsector. Measures have been simulated using 100,000 realizations.

Chart 9: Distress scenario for Canadian firms in the $FIJD$ measure



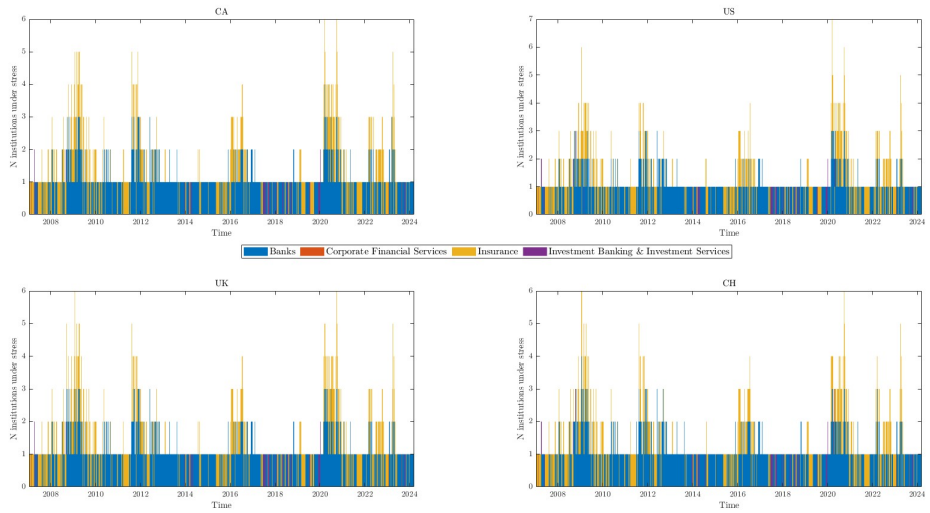

This chart shows the distress scenario for Canadian firms that generates the lowest market return for a certain set of firms, weighted by their market capitalization, aggregated by the regions indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

Chart 10: Distress scenario for US firms in the $FIJD$ measure



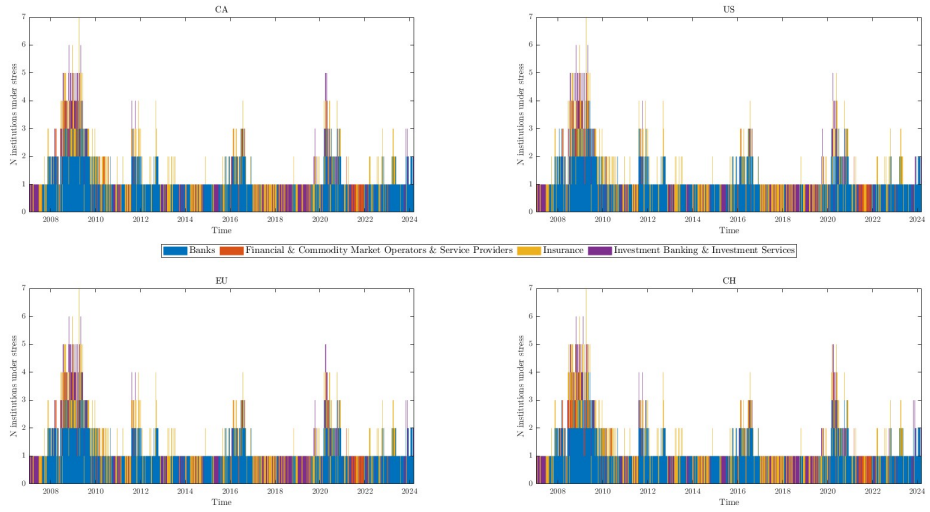

This chart shows the distress scenario for US firms that generates the lowest market return for a certain set of firms, weighted by their market capitalization, aggregated by the regions indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

Chart 11: Distress scenario for European firms in the $FIJD$ measure



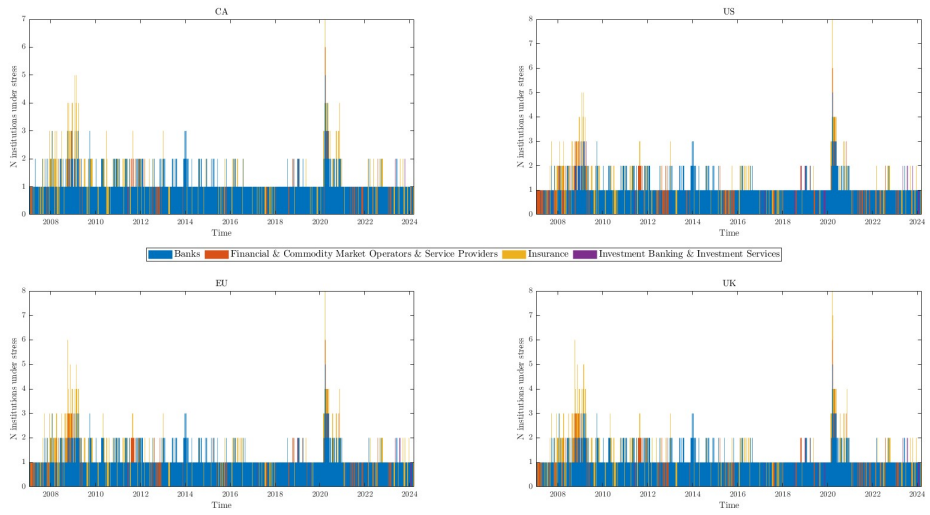

This chart shows the distress scenario for European firms that generates the lowest market return for a certain set of firms, weighted by their market capitalization, aggregated by the regions indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

Chart 12: Distress scenario for British firms in the $FIJD$ measure




This chart shows the distress scenario for British firms that generates the lowest market return for a certain set of firms, weighted by their market capitalization, aggregated by the regions indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

Chart 13: Distress scenario for Swiss firms in the $FIJD$ measure



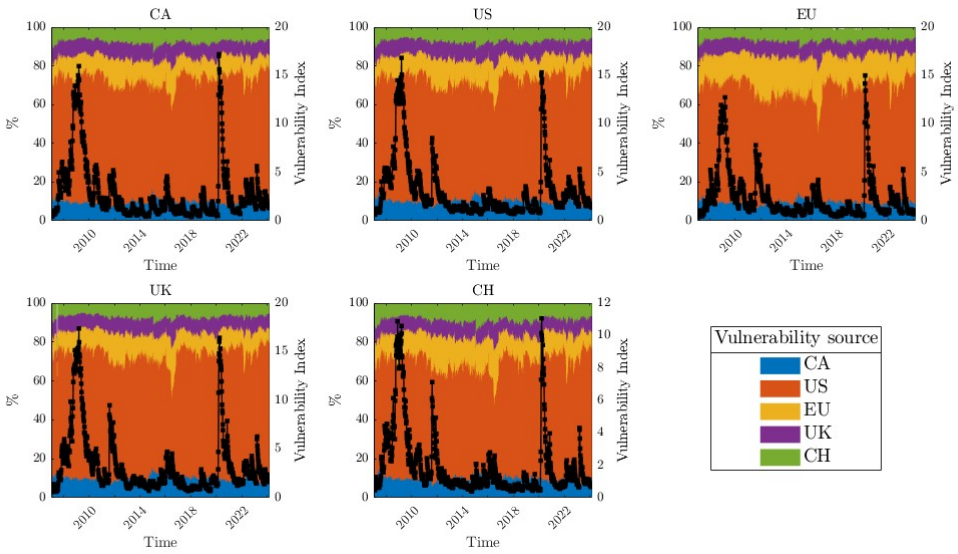

This chart shows the distress scenario for Swiss firms that generates the lowest market return for a certain set of firms, weighted by their market capitalization, aggregated by the regions indicated in the title of the subplot. Each bar shows the number of institutions under stress, with a decomposition by subsector with a joint probability higher than 4%.

## 4.3 Vulnerability Index

The Vulnerability Index is aggregated within the institutions in the same region, by using the market capitalization as a weighting factor. Chart 14 shows the Vulnerability Index (VI) in the right side of the subplots for each region, while the right side shows the contribution of the other regions to the corresponding vulnerability index of the region indicated in the title of each subplot.

The US shows a dominant role as contributor, both to the US vulnerability index and foreign vulnerability index. The contribution of Europe is higher for the European Vulnerability Index than the rest of regions.

Chart 14: Vulnerability Index and its Contribution.




This chart shows the Vulnerability Index as defined by 5, aggregated by region using the market capitalization of the financial firms as a weights. The Contribution to the Vulnerability Index, as defined by 6, indicates the role of each country shown in the legend in the regional Vulnerability Index indicated in the title of each subplot. Measures have been simulated using 100,000 realizations.

## 4.4 Network Of Expected Shortfall

We have done the decomposition of the Expected Shortfall (ES) up to 10 conditioning financial institutions (FIs) or up to gather the full ES under stress scenarios for other FIs. When selecting the firms to decompose the Expected Shortfall, we start selecting the financial firm with higher conditional probability of distress. For instance, if we have three institutions (A, B and C) and we want to decompose the Expected Shortfall of institution A, i.e. $E(r_A|r_A \leq \kappa_A)$, we would start decomposing by firm B if $P(r_B \leq \kappa_B|r_A \leq \kappa_A) > P(r_C \leq \kappa_C|r_A \leq \kappa_A)$. We could graphically represent the connection between institution A and institutions B and C by two arrows going out of A and reaching B and C. The arrow from A to B would have a width of $\frac{P(r_B \leq \kappa_B|r_A \leq \kappa_A)E(r_A|r_A \leq \kappa_A, r_B \leq \kappa_B)}{E(r_A|r_A \leq \kappa_A)}$, while the arrow from A to C would have a width of $\frac{P(r_C \leq \kappa_C|r_A \leq \kappa_A, r_B > \kappa_B)E(r_A|r_A \leq \kappa_A, r_B > \kappa_B, r_C \leq \kappa_C)}{E(r_A|r_A \leq \kappa_A)}$. Note that the sum of the two arrows would have a width of 100% if tail loss for firm A happens always when institution B or C is in their tail return, or it would be below 100% if there is a share of the Expected Shortfall for the returns of firm A that are not materialized when firms B and C are not in stress, i.e. ; $P(r_A \leq \kappa_A|r_B > \kappa_B, r_C > \kappa_C)$.

We present in Figure 4 the network of Expected Shortfall at the breakout of the COVID crisis, i.e. week of the 18 March 2020. The dots represents the financial firms in our sample (274 dots), with the color representing the region of the firm. This picture show a key role of US for the financial system, with a secondary role of European and Swiss financial firms. Then, in a second layer, outside the inner core of the network, we find Canadian and British financial institutions.

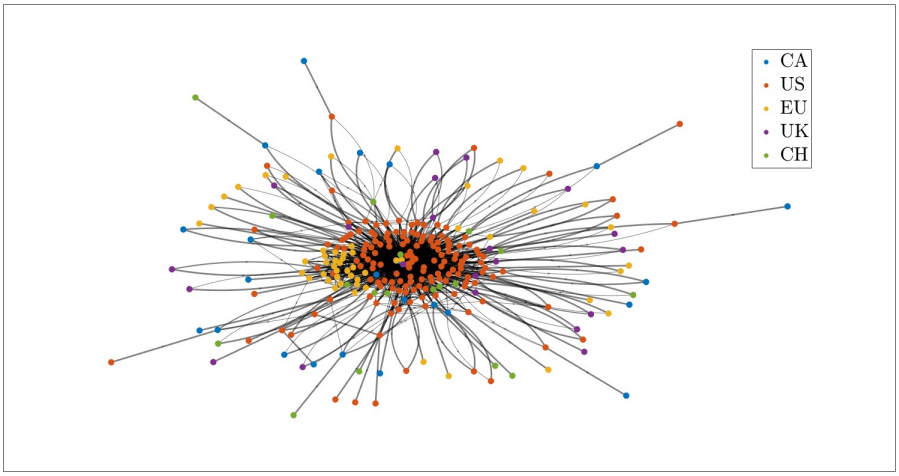
We can plot this type of graph over time, but in order to capture information in a easier way, we construct a weighted out-degree measure, e.g. we are summing the arrows we are summing the arrows going out of Canadian Institutions weighted by the width of each arrow and the market capitalization share to each region.[1]

Figure 15 show the final outcome. The last subplot in the bottom right of the figure indicate the share of the Expected Shortfall that we have been able to explain by using 10 institutions. For most of the regions we reach close to 100%, only for Switzerland we find a coverage around 95% and a small drop at the beginning of the sample for the UK financial system at 85%. Nevertheless, the coverage of the Expected Shortfall is always

[1] In the network literature the weight usually limits to the width of the arrow, however we are also applying the weight of the market capitalization to get a value that sum up to 1 for each region, telling us about which are the regions for which the stress scenarios could explain most of the ES in the financial system.

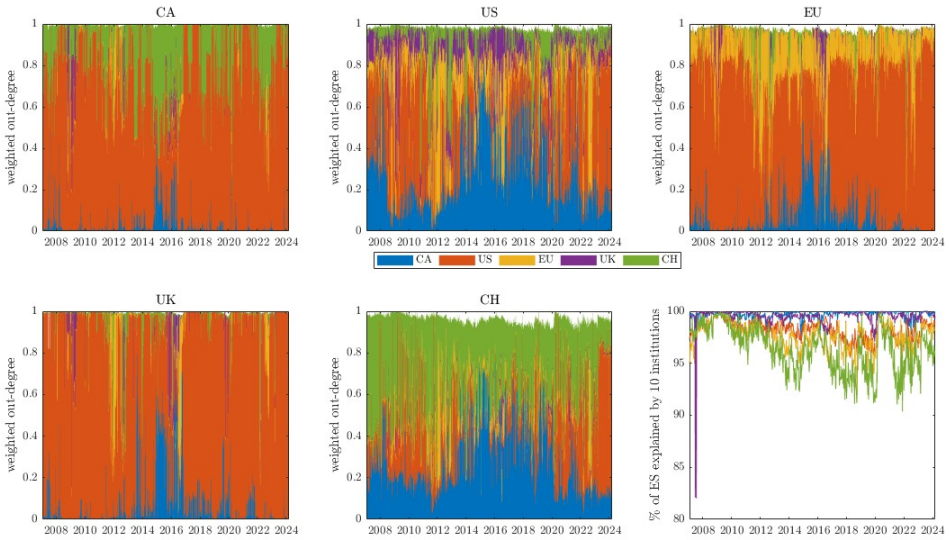
above 80%. The US dominates the explanation of the Expected Shortfall of other countries, which is a similar conclusion to the rest of the measures. European and Swiss financial systems presents a relevant decomposition of their Expected Shortfall in domestic scenarios, while for UK and Canada we find foreign regions to be more relevant in general. Interestingly, we find a couple of results that we haven't seen in the other measures. First, Canada becomes now relevant, in terms of scenarios, for the Expected Shortfall of US financial firms and Swiss financial firms. The financial links with US and the size of the insurance sectors in Canada and Switzerland might explain this results. Second, Switzerland is, after US, the region that have a higher share of the Canadian Expected Shortfall happening at the same time than a distress scenario in this country.

Figure 4: Network of Expected Shortfall at the breakout of the COVID crisis (week of 18 March 2020)




This figure shows the decomposition of Expected Shortfall as defined in Eq. (7), where the colour of the node indicate the region of the financial institution and the edge from A to B indicates that the Expected Shortfall of A happens at the same time than the Expected Shortfall of B. The width of the edge indicate which is the share of the Expected Shortfall of A happening at the same time than the Expected Shortfall of B. Measures have been simulated using 100,000 realizations.

Chart 15: Average weighted out-degree of each region




This chart shows the importance of each region to explain the Expected Shortfall of other regions (shown in the legend). The importance is measure by the average of the weighted in-degree of each region, where the in-degree indicates how many tail returns institutions from a certain region happen at the same time than the tail returns of the firms in the region indicated in the title of each subplot. The weight is given by the share of the Expected Shortfall that is explained, where we also use the market capitalization to take into account the size of each firm. Measures have been simulated using 100,000 realizations.

# 5 Conclusion

The systemic risk literature has proposed several risk measures to identify the connection between financial institutions. Most of the measures set a scenario for a single institution to generate the probability space where the measure is built. This limits the scope of the measure, as the institution and the numbers of institutions on which we build the scenario might have a different probability of occurrence over time and a different effect on the institutions under scope. In order to build a robust set of systemic risk measures, we need a methodology flexible enough to adapt to different sets of conditioning scenarios. This article uses a bi-factor latent copula approach to capture the joint dependence across institutions from different financial subsectors and regions. We also employ a Skewed Student t copula to be able to capture in a better way the tail behaviour of financial assets. This framework allows us to generate two systemic risk measures looking at conditional probabilities and conditional returns with a flexible conditioning definition, setting the probability but leaving free the number of institutions that would define the stress scenario. This modification will give us more information about the scenario from which some institutions are more vulnerable and how that scenario changes over time, complementing the information provided by the risk measure. We also proposed a decomposition of the tail loss of the financial institutions in portions where other financial institutions are under distress too, which would allow us to build network and analyse the centrality of some institutions in the network of stress scenarios between them.

We perform an empirical application using weekly equity data from 2006 up to 2023 for financial institutions from 5 different regions (Canada, United States, Euro Area, United Kingdom and Switzerland). We found that the role of US is predominant in terms of effect on the systemic risk measures for the rest of jurisdictions, but the composition of the stress scenarios is changing over time. While US financial and commodity marker operator companies are present in the stress scenarios for other regions in the period 2008-2010, the banking sector lead the scenario composition in the period 2010-2018, while the US insurance sector starts having more relevance in the design of stress scenario from 2018 onwards. Also we find different patterns, both for conditional probabilities and returns, for the different jurisdictions, with a higher volatility of European systemic indicators compared to American systemic risk measures and some idiosyncratic patterns captured by some regions, like the Brexit event in the UK time series. Also, we find that the between 20 to 60% of the Expected Shortfall from the US financial system is happening at the same time that the stress scenario of Canada, which reflect the great symbiosis between these two jurisdictions. Canada also presents ties with Switzerland, being more than 40% of the Swiss Expected Shortfall is ocurring at the same time than tail losses in Canadian financial firms. This link could be explained by the role of insurance companies in two countries.

The proposed methodology opens the floor to analyse global spillovers from and to emerging economies by

controlling the exchange rate effects, as the latent factor allow us to model the currencies through relationships between latent factor. Also, the model could be employed to generate generalizations of some well-known measures like the CoVaR in a multidimensional framework, where the conditioning event would be defined by a set of financial institutions.

The findings of this research have implications for several economic agents. Firstly, for the supervisory authorities, who seek to monitor and track systemic risk and to identify scenarios that could generate large losses in the financial market. Secondly, policy makers, who wish to understand the interactions between financial system from different regions and how it could impact domestic economy. Thirdly, for risk management, investors and traders, who might find this methodology interesting to reduce portfolio's tail loss.

## References

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

Acharya, V. V., L. H. Pedersen, T. Philippon, and M. Richardson (2017). Measuring systemic risk. *The review of financial studies 30*(1), 2–47.

Adrian, T. and M. K. Brunnermeier (2016). Covar. *The American Economic Review 106*(7), 1705.

Banulescu, G.-D. and E.-I. Dumitrescu (2015). Which are the sifis? a component expected shortfall approach to systemic risk. *Journal of Banking & Finance 50*, 575–588.

Bédard-Pagé, G., A. Demers, E. Tuer, and M. Tremblay (2016). Large canadian public pension funds: A financial system perspective. *Bank of Canada Financial System Review*, 33–38.

Blasques, F., S. J. Koopman, K. Łasak, and A. Lucas (2016). In-sample confidence bands and out-of-sample forecast bands for time-varying parameters in observation-driven models. *International Journal of Forecasting 32*(3), 875–887.

Blasques, F., J. van Brummelen, S. J. Koopman, and A. Lucas (2022). Maximum likelihood estimation for score-driven models. *Journal of Econometrics 227*(2), 325–346.

Blei, S. K. and B. Ergashev (2014). Asset commonality and systemic risk among large banks in the united states. *Available at SSRN 2503046*.

Brownlees, C. and R. F. Engle (2017). Srisk: A conditional capital shortfall measure of systemic risk. *The Review of Financial Studies 30*(1), 48–79.

Bruneau, G., J. Ojea Ferreiro, A. Plummer, M.-C. Tremblay, and A. Witts (2023). Understanding the systemic implications of climate transition risk: Applying a framework using canadian financial system data. Technical report, Bank of Canada.

Brychkov, Y. A. (2016). Higher derivatives of the bessel functions with respect to the order. *Integral Transforms and Special Functions 27*(7), 566–577.

Cai, J. (1994). A markov model of switching-regime arch. *Journal of Business & Economic Statistics 12*(3), 309–316.

Cai, J., F. Eidam, A. Saunders, and S. Steffen (2018). Syndication, interconnectedness, and systemic risk. *Journal of Financial Stability 34*, 105–120.

Chan-Lau, J., M. Espinosa, and J. Solé (2009). On the use of network analysis to assess systemic financial linkages. *IMF (International Monetary Fund) working paper*.

Chan-Lau, J. A., M. Espinosa, K. Giesecke, and J. A. Solé (2009). Assessing the systemic implications of financial linkages. *IMF global financial stability report 2*.

Cortes, F., P. Lindner, S. Malik, and M. A. S. Basurto (2018). *A comprehensive multi-sector tool for analysis of systemic risk and interconnectedness (SyRIN)*. International Monetary Fund.

Creal, D., S. J. Koopman, and A. Lucas (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics 28*(5), 777–795.

Demarta, S. and A. J. McNeil (2005). The t copula and related copulas. *International statistical review 73*(1), 111–129.

ECB (2009, December). The concept of systemic risk. Technical report, EuropeanCentral Bank.

Engle, R. and B. Kelly (2012). Dynamic equicorrelation. *Journal of Business & Economic Statistics 30*(2), 212–228.

González-Rivera, G., C. V. Rodríguez-Caballero, and E. Ruiz (2021). Expecting the unexpected: economic growth under stress. Technical report, Department of Economics and Business Economics, Aarhus University.

González-Santander, J. (2018). Closed-form expressions for derivatives of bessel functions with respect to the order. *Journal of Mathematical Analysis and Applications 466*(1), 1060–1081.

González-Santander, J. L. (2023). Reflection formulas for order derivatives of bessel functions. *Results in Mathematics 78*(1), 30.

Gravelle, T. and F. Li (2013). Measuring systemic importance of financial institutions: An extreme value theory approach. *Journal of Banking & Finance 37*(7), 2196–2209.

Hałaj, G. (2020). Resilience of canadian banks to funding liquidity shocks. *Latin American Journal of Central Banking 1*(1-4), 100002.

Hałaj, G. and R. Hipp (2024). Decomposing systemic risk: the roles of contagion and common exposures. (No 2929).

Hałaj, G. and C. Kok (2013). Assessing interbank contagion using simulated networks. *Computational Management Science 10*, 157–186.

Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review*, 705–730.

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

Huang, X. (1992). *Statistics of Bivariate Extreme Values*. Tinbergen Institute Re-search Series, Ph.D. Thesis, No 22, Erasmus University, Rotterdam, Netherlands.

Huang, X., H. Zhou, and H. Zhu (2009). A framework for assessing the systemic risk of major financial institutions. *Journal of Banking & Finance 33*(11), 2036–2049.

Joe, H. and J. J. Xu (1996). The estimation method of inference functions for margins for multivariate models. Technical report, Department of Statistics, University of British Columbia.

Koopman, S. J., A. Lucas, and M. Scharth (2016). Predicting time-varying parameters with parameter-driven and observation-driven models. *Review of Economics and Statistics 98*(1), 97–110.

Krupskii, P. and H. Joe (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis 120*, 85–101.

Krupskii, P. and H. Joe (2015). Structured factor copula models: Theory, inference and computation. *Journal of Multivariate Analysis 138*, 53–73.

Lehar, A. (2005). Measuring systemic risk: A risk management approach. *Journal of Banking & Finance 29*(10), 2577–2603.

Lucas, A., B. Schwaab, and X. Zhang (2014). Conditional euro area sovereign default risk. *Journal of Business & Economic Statistics 32*(2), 271–284.

Lucas, A., B. Schwaab, and X. Zhang (2017). Modeling financial sector joint tail risk in the euro area. *Journal of Applied Econometrics 32*(1), 171–191.

Oh, D. H. and A. J. Patton (2017). Modeling dependence in high dimensions with factor copulas. *Journal of Business & Economic Statistics 35*(1), 139–154.

Oh, D. H. and A. J. Patton (2018). Time-varying systemic risk: Evidence from a dynamic copula model of CDS spreads. *Journal of Business & Economic Statistics 36*(2), 181–195.

Oh, D. H. and A. J. Patton (2023). Dynamic factor copula models with estimated cluster assignments. *Journal of Econometrics 237*(2), 105374.

Pakel, C., N. Shephard, K. Sheppard, and R. F. Engle (2021). Fitting vast dimensional time-varying covariance models. *Journal of Business & Economic Statistics 39*(3), 652–668.

Segoviano, M. A. and C. Goodhart (2010). Banking stability measures. *IMF Working Papers 2010*(09/4).

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. In *Annales de l'ISUP*, Volume 8, pp. 229–231.

Smith, M. S. (2023). Implicit copulas: An overview. *Econometrics and Statistics 28*, 81–104.

Stroud, A. H., D. Secrest, et al. (1966). *Gaussian quadrature formulas*, Volume 374. Prentice-Hall Englewood Cliffs, NJ.

Sydow, M., A. Schilte, G. Covi, M. Deipenbrock, L. Del Vecchio, P. Fiedor, G. Fukker, M. Gehrend, R. Gourdel, A. Grassi, et al. (2024). Shock amplification in an interconnected financial system of banks and investment funds. *Journal of Financial Stability*, 101234.

Yang, Z.-H. and Y.-M. Chu (2017). On approximating the modified bessel function of the second kind. *Journal of inequalities and applications 2017*, 1–8.

# Appendix

## The Skewed t Distribution

The bivariate skewed t distribution discussed in Demarta and McNeil (2005) has the following stochastic representation:

$$X \quad = \quad \mu + \lambda W + \sqrt{W} Z \tag{19}$$

where $\mu$ and $\lambda$ are the mean and asymmetry vector parameter respectively , $W$ is an inverse gamma distributed random variable $W \sim IG(\frac{\nu}{2}, \frac{\nu}{2})$ with $\nu$ being the number of degrees of freedom, $Z$ is a multivariable normal random variable $Z \sim N(0, P)$ independent of $W$ with $P = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$, i.e. $Z = L\varepsilon$ with $\varepsilon \sim N(0, I)$ and $L = \begin{bmatrix} L_{11} & 0 \\ L_{12} & L_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix}$.
Note that $E(X) = \mu + \frac{\nu}{\nu-2}\lambda$ and $Cov(X) = \Sigma = \frac{\nu}{\nu-2}P + \frac{2\nu^2}{(\nu-2)^2(\nu-4)}\lambda\lambda'$, so $\nu > 4$ to have a defined variance. The bivariate distribution is:

$$f_X(x) \quad = \quad c\frac{K_{\frac{\nu+2}{2}}(\sqrt{(\nu + d(x))\lambda' P^{-1}\lambda}) \exp([x-\mu]' P^{-1}\lambda)}{(\sqrt{(\nu + d(x))\lambda' P^{-1}\lambda})^{-\frac{\nu+2}{2}} (1 + \frac{d(x)}{\nu})^{\frac{\nu+2}{2}}},$$

with $c = \frac{2^{-\nu/2}}{\Gamma(\frac{\nu}{2})\pi\nu|P|^{1/2}}$, $d(x) = [x-\mu]' P^{-1}[x-\mu]$ and $K_a(b)$ being the modified Bessel function of the second kind. We use the approximations from Yang and Chu (2017) when b is close to zero and b is large. In particular $\lim_{b\to 0} K_a(b) = \frac{1}{2}\Gamma(a)\left(\frac{b}{2}\right)^{-a}$ and $\lim_{b\to\infty} K_a(b) = \sqrt{\frac{\pi}{2b}}\exp(-b)\left(1 + \frac{4a^2-1}{8b}\left(1 + \frac{4a^2-9}{16b}\right)\right)$. As shown by Lucas et al. (2014), this approximation allows us to cancel the exponential term multiplying the modified Bessel function, adding numerical stability for the skewness effect in the far tails.
Note that Eq. (19) conditioned to a realization of $W$ is normally distributed, i.e.

$$X|W \sim N\left(\underbrace{\begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix} + W \otimes \begin{bmatrix}\lambda_1 \\ \lambda_2\end{bmatrix}}_{\mu_{X|W}}, \underbrace{W \otimes \begin{bmatrix}1 & \rho \\ \rho & 1\end{bmatrix}}_{\Sigma_{X|W}}\right)$$

where $\otimes$ is the Kronecker product. Eq. (13) could be rewritten as

$$f_X(x) = \int_0^\infty \phi_{X|W}(x|W) f_W(w) dw, \tag{20}$$

where $\phi_{X|W})(\dots)$ is the normal probability distribution function with mean $\mu_{X|W}$ and variance matrix $\Sigma_{X|W}$ and $f_W(\dots)$ is the probability distribution function of the inverse Gaussian with all parameters equal to $\frac{\nu}{2}$. Eq. (20) allows us to write the cumulative distribution function of the skewed t distribution as

$$F_X(x) = \int_0^\infty \Phi_{X|W}(x|w) f_W(w) dw, \tag{21}$$

where $\Phi_{X|W})(\dots)$ is the cumulative normal probability distribution function with mean $\mu_{X|W}$ and variance matrix $\Sigma_{X|W}$.

We obtain the copula from the ratio between the joint density function and the product of the marginal distributions. The univariate density of the skewed t distribution is

$$f_X(x) = c \frac{K_{\frac{\nu+1}{2}}(\sqrt{(\nu + d(x))\frac{\lambda^2}{\sigma^2}}) \exp\left(\frac{(x-\mu)}{\sigma^2}\lambda\right)}{(\sqrt{(\nu + d(x))\frac{\lambda^2}{\sigma^2}})^{-\frac{\nu+1}{2}} (1 + \frac{d(x)}{\nu})^{\frac{\nu+1}{2}}}, \tag{22}$$

with $c = \frac{2^{1-\frac{\nu+1}{2}}}{\Gamma(\frac{\nu}{2})(\pi\nu)^{\frac{1}{2}}\sigma}$ and $d(x) = \frac{(x-\mu)^2}{\sigma^2}$, and where $\sigma.^2$ is given by the diagonal elements of $P$. The copula density function is defined implicitly via

$$c_(u, v) = \frac{f_X(F_{X_1}^{-1}(u), F_{X_2}^{-1}(v))}{f_{X_1}(F_{X_1}^{-1}(u)) f_{X_2}(F_{X_2}^{-1}(v))}. \tag{23}$$

Copula function could be easily obtained from Eq. (21) as $C(u, v) = F(F_1^1(u), F_2^{-1}(v))$, i.e.,

$$C(u, v) = \int_0^\infty \Phi_{X|W}(\Phi_{X_1|W}^{-1}(u), \Phi_{X_2|W}^{-1}(v)|w) f_W(w) dw. \tag{24}$$

The conditional copula could be obtained following Eq. (21) as

$$C(u|v) = \int_0^\infty \Phi\left(\frac{F_{X_1}^{-1}(u) - \mu_{X_1|X_2,W}}{\sigma_{X_1|X_2,W}}\right) f_W(w) dw. \tag{25}$$

with $\Phi(\dots)$ being the standardized normal cumulative distribution function and

$$\mu_{X_1|X_2,W} = \mu_{X_1|W} + \frac{\sigma_{12|W}}{\sigma_{X_2|W}^2}\left(F_{X_2|W}^{-1}(v) - \mu_{X_2|W}\right)$$

with $\sigma_{12|W}$ being the element in column 1 row 2 from matrix $\Sigma_{X|W}$ and

$$\sigma_{X_1|X_2,W} = \sqrt{\sigma_{X_1|W}^2 - \frac{\sigma_{12|W}^2}{\sigma_{X_2|W}^2}}$$

## Derivatives of the factor copula model

The minus loglikelihood function is defined as

$$S = -\sum_{t=1}^{T} \log(c(u_t, v_t; \Theta)),$$

where $\Theta$ is the set of copula parameters. The gradient with respect to the parameters $\Theta$ would be

$$\frac{\partial S}{\partial \Theta} = -\sum_{t=1}^{T} \frac{\partial \log(c(u_t, v_t; \Theta))}{\partial \Theta}.$$

Note that using the chain rule we can find the following relationship

$$\frac{\partial \log(c(u_t, v_t; \Theta))}{\partial \Theta} = \frac{1}{c(u_t, v_t; \Theta)} \frac{\partial c(u_t, v_t; \Theta)}{\partial \Theta}. \tag{26}$$

The one-factor copula is defined

$$c(u,v;\Theta) = \int_0^1 c_{u,z}(u,z;\theta_{u,z})c_{v,z}(v,z;\theta_{v,z})dz$$

where $\Theta = [\theta'_{u,z}, \theta'_{v,z}]'$. In the skewed t framework we would have six parameters, i.e. $\Theta = [\rho_{u,z}, \lambda_u, \rho_{v,z}, \lambda_v, \lambda_z, \nu]$. Obviously, the factor model would provide a good alternative for the estimation when the number of variables N is large ($N >> 2$). Following Leibniz integral rule, we can apply the differentiation under the integral sign, e.g.,

$$\begin{aligned}
\frac{\partial c(u,v;\Theta)}{\partial \rho_{u,z}} &= \int_0^1 \frac{\partial c_{u,z}(u,z;\theta_{u,z})}{\partial \rho_{u,z}} c_{v,z}(v,z;\theta_{v,z})dz \\
\frac{\partial c(u,v;\Theta)}{\partial \lambda_u} &= \int_0^1 \frac{\partial c_{u,z}(u,z;\theta_{u,z})}{\partial \lambda_u} c_{v,z}(v,z;\theta_{v,z})dz \\
\frac{\partial c(u,v;\Theta)}{\partial \rho_{v,z}} &= \int_0^1 c_{u,z}(u,z;\theta_{u,z}) \frac{\partial c_{v,z}(v,z;\theta_{v,z})}{\partial \rho_{v,z}} dz \\
\frac{\partial c(u,v;\Theta)}{\partial \lambda_v} &= \int_0^1 c_{u,z}(u,z;\theta_{u,z}) \frac{\partial c_{v,z}(v,z;\theta_{v,z})}{\partial \lambda_v} dz \\
\frac{\partial c(u,v;\Theta)}{\partial \lambda_z} &= \int_0^1 \left( \frac{\partial c_{u,z}(u,z;\theta_{u,z})}{\partial \lambda_z} c_{v,z}(v,z;\theta_{v,z}) + c_{u,z}(u,z;\theta_{u,z}) \frac{\partial c_{v,z}(v,z;\nu)}{\partial \lambda_z} \right) dz \\
\frac{\partial c(u,v;\Theta)}{\partial \nu} &= \int_0^1 \left( \frac{\partial c_{u,z}(u,z;\theta_{u,z})}{\partial \nu} c_{v,z}(v,z;\theta_{v,z}) + c_{u,z}(u,z;\theta_{u,z}) \frac{\partial c_{v,z}(v,z;\nu)}{\partial \nu} \right) dz
\end{aligned}$$

For the derivatives under the integral sign we follow Eq. (26) to get the derivative of the copula as the product between the derivative of the log-copula multiplied by the copula density, e.g. $\frac{\partial c_{u,z}(u,z;\theta_{u,z})}{\partial \rho_{u,z}} = \frac{\partial \log(c_{u,z}(u,z;\theta_{u,z}))}{\partial \rho_{u,z}} c_{u,z}(u,z;\theta_{u,z})$. Once we get the derivative of the factor copula with respect to the parameter, we apply again Eq. (26) to get the derivative of the log-copula with respect to the parameter, e.g. $\frac{\partial \log(c(u_t,v_t;\Theta))}{\partial \rho_{u,z}} = \frac{1}{c(u_t,v_t;\Theta)} \frac{\partial c(u,v;\Theta)}{\partial \rho_{u,z}}$.

The computation of the derivative of the log-copula hugely simplifies the analytical assessment of the gradient vector. The next subsection presents the analytical derivatives with respect to the parameters of the static copula model, as they are needed to compute the static factor copula for a large dimensional dataset. We use a modified Newton-Rapshon method to estimate the constant factor copula, as suggested by Krupskii and Joe (2015), to be used as "variance targeting" to anchor the long-term dependence like Oh and Patton (2023) and Lucas et al. (2014).

## Derivative of the GHST log-copula

The log-likelihood copula is defined as

$$\log\left(c(u_t,v_t;\Theta)\right) = \log\left(f(F_{X_1}^{-1}(u), F_{X_2}^{-1}(v);\Theta)\right) - \log\left(f(F_{X_1}^{-1}(u);\Theta)\right) - \log\left(f_{X_2}(F_{X_2}^{-1}(v);\Theta)\right), \quad (27)$$

with its derivative being

$$\begin{aligned}
\frac{\partial \log\left(c(u_t,v_t;\Theta)\right)}{\partial \Theta} &= \frac{\partial \log\left(f(F_{X_1}^{-1}(u), F_{X_2}^{-1}(v);\Theta)\right)}{\partial \Theta} - \frac{\partial \log\left(f_{X_1}(F_{X_1}^{-1}(u);\Theta)\right)}{\partial \Theta} \\
&\quad - \frac{\partial \log\left(f_{X_2}(F_{X_2}^{-1}(v);\Theta)\right)}{\partial \Theta},
\end{aligned} \quad (28)$$

with $f(\ldots,\ldots;\Theta)$ and $f_{\ldots}$ provided by Eq. (13) with N=2 and N=1 respectively.
We provide the derivative of the log-likelihood with respect of each parameter for the marginal and the joint distribution.

### Derivative of the joint distribution with respect to the GHST parameters.

$$\begin{aligned}
\log\left(f(\tilde{x}_1,\tilde{x}_2;\Theta)\right) &= -\frac{\nu}{2}\log(2) - \log(\Gamma(\frac{\nu}{2})) - \log(\pi\nu) - \log(\sqrt{1-\rho^2}) \\
&\quad + \log\left(K_{\frac{\nu+2}{2}}\left(((\nu+A)B)^{1/2}\right)\right) + \mathcal{L} + \frac{\nu+2}{2}\log\left(((\nu+A)B)^{1/2}\right) \\
&\quad - \frac{(\nu+2)}{2}\log\left(1+\frac{A}{\nu}\right)
\end{aligned}$$

with $x_1 = F_{X_1}^{-1}(u)$, $x_2 = F_{X_2}^{-1}(v)$, $A = a_1 + a_2 + a_3$ with $a_1 = \frac{(x_1-\mu_1)^2}{1-\rho^2}$, $a_2 = \frac{(x_2-\mu_2)^2}{1-\rho^2}$ and $a_3 = 2(x_1-\mu_1)(x_2-\mu_2)\frac{-\rho}{1-\rho^2}$. $\mu_1 = 0$, $\mu_2 = 0$, $B = b_1 + b_2 + b_3$ with $b_1 = \frac{\lambda_1^2}{1-\rho^2}$, $b_2 = \frac{\lambda_2^2}{1-\rho^2}$ and $b_3 = 2\lambda_1\lambda_2\frac{-\rho}{1-\rho^2}$. $\mathcal{L} = \ell_1 + \ell_2 + \ell_3$ with $\ell_1 = \frac{(x_1-\mu_1)\lambda_1}{1-\rho^2}$, $\ell_2 = \frac{(x_2-\mu_2)\lambda_2}{1-\rho^2}$ and $\ell_3 = ((x_1-\mu_1)\lambda_2 + (x_2-\mu_2)\lambda_1)\frac{-\rho}{1-\rho^2}$.

**Derivative of the joint distribution with respect to the correlation parameter.**

$$
\begin{aligned}
\frac{\partial a_1}{\partial \rho} &= 2\rho\frac{(x_1-\mu_1)^2}{(1-\rho^2)^2} & \frac{\partial b_3}{\partial \rho} &= 2\lambda_1\lambda_2\left(\frac{-1}{1-\rho^2} + \frac{-2\rho^2}{(1-\rho^2)^2}\right)\\
\frac{\partial a_2}{\partial \rho} &= 2\rho\frac{(x_2-\mu_2)^2}{(1-\rho^2)^2} & \frac{\partial B}{\partial \rho} &= \frac{\partial b_1}{\partial \rho} + \frac{\partial b_2}{\partial \rho} + \frac{\partial b_3}{\partial \rho}\\
\frac{\partial a_3}{\partial \rho} &= 2(x_1-\mu_1)(x_2-\mu_2)\left(\frac{-1}{1-\rho^2} + \frac{-2\rho^2}{(1-\rho^2)^2}\right) & \frac{\partial \ell_1}{\partial \rho} &= 2\rho\frac{\lambda_1(x_1-\mu_1)}{(1-\rho^2)^2}\\
\frac{\partial A}{\partial \rho} &= \frac{\partial a_1}{\partial \rho} + \frac{\partial a_2}{\partial \rho} + \frac{\partial a_3}{\partial \rho} & \frac{\partial \ell_2}{\partial \rho} &= 2\rho\frac{\lambda_2(x_2-\mu_2)}{(1-\rho^2)^2}\\
\frac{\partial b_1}{\partial \rho} &= 2\rho\frac{\lambda_1^2}{(1-\rho^2)^2} & \frac{\partial \ell_3}{\partial \rho} &= (\lambda_1(x_2-\mu_2) + \lambda_2(x_1-\mu_1))\left(\frac{-1}{1-\rho^2} + \frac{-2\rho^2}{(1-\rho^2)^2}\right)\\
\frac{\partial b_2}{\partial \rho} &= 2\rho\frac{\lambda_2^2}{(1-\rho^2)^2} & \frac{\partial \mathcal{L}}{\partial \rho} &= \frac{\partial \ell_1}{\partial \rho} + \frac{\partial \ell_2}{\partial \rho} + \frac{\partial \ell_3}{\partial \rho}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \log\left(f(x_1,x_2;\Theta)\right)}{\partial \rho} &= \frac{\rho}{1-\rho^2} + \left(\frac{1}{K_{\frac{\nu+2}{2}}\left(((\nu+A)B)^{1/2}\right)}k_{\frac{\nu+2}{2}}\left(((\nu+A)B)^{1/2}\right) + \frac{(\nu+2)/2}{((\nu+A)B)^{1/2}}\right)\\
&\quad \frac{1}{2}\left((\nu+A)B\right)^{-1/2}\left((\nu+A)\frac{\partial B}{\partial \rho} + B\frac{\partial A}{\partial \rho}\right) + \frac{\partial \mathcal{L}}{\partial \rho} +\\
&\quad -\frac{(\nu+2)/2}{\left(1+\frac{A}{\nu}\right)}\frac{1}{\nu}\frac{\partial A}{\partial \rho}
\end{aligned}
$$

**Derivative of the joint distribution with respect to the asymmetric parameter of the first variable.** The derivatives $\frac{\partial x_1}{\partial \lambda_1} = \frac{\partial F_{X_1}^{-1}(u;\lambda_1,\nu)}{\partial \lambda_1}$ is computed numerically.

$$
\begin{aligned}
\frac{\partial a_1}{\partial \lambda_1} &= \frac{2(x_1-\mu_1)\frac{\partial x_1}{\partial \lambda_1}}{1-\rho^2} & \frac{\partial b_3}{\partial \lambda_1} &= = 2\lambda_2\frac{-\rho}{1-\rho^2}\\
\frac{\partial a_2}{\partial \lambda_1} &= 0 & \frac{\partial B}{\partial \lambda_1} &= \frac{\partial b_1}{\partial \lambda_1} + \frac{\partial b_2}{\partial \lambda_1} + \frac{\partial b_3}{\partial \lambda_1}\\
\frac{\partial a_3}{\partial \lambda_1} &= 2\frac{\partial x_1}{\partial \lambda_1}(x_2-\mu_2)\frac{-\rho}{1-\rho^2} & \frac{\partial \ell_1}{\partial \lambda_1} &= \frac{\left(\frac{\partial x_1}{\partial \lambda_1}\lambda_1 + (x_1-\mu_1)\right)}{1-\rho^2}\\
\frac{\partial A}{\partial \lambda_1} &= \frac{\partial a_1}{\partial \lambda_1} + \frac{\partial a_2}{\partial \lambda_1} + \frac{\partial a_3}{\partial \lambda_1} & \frac{\partial \ell_2}{\partial \lambda_1} &= 0\\
\frac{\partial b_1}{\partial \lambda_1} &= \frac{2\lambda_1}{1-\rho^2} & \frac{\partial \ell_3}{\partial \lambda_1} &= \left(\frac{\partial x_1}{\partial \lambda_1}\lambda_2 + (x_2-\mu_2)\right)\frac{-\rho}{1-\rho^2}\\
\frac{\partial b_2}{\partial \lambda_1} &= 0 & \frac{\partial \mathcal{L}}{\partial \lambda_1} &= \frac{\partial \ell_1}{\partial \lambda_1} + \frac{\partial \ell_2}{\partial \lambda_1} + \frac{\partial \ell_3}{\partial \lambda_1}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \log\left(f(x_1,x_2;\Theta)\right)}{\partial \lambda_1} &= \\
&\quad + \left(\frac{k_{\frac{\nu+2}{2}}\left(((\nu+A)B)^{1/2}\right)}{K_{\frac{\nu+2}{2}}\left(((\nu+A)B)^{1/2}\right)} + \frac{(\nu+2)/2}{((\nu+A)B)^{1/2}}\right)\frac{1}{2}\left((\nu+A)B\right)^{-1/2}\left((\nu+A)\frac{\partial B}{\partial \lambda_1} + B\frac{\partial A}{\partial \lambda_1}\right) + \frac{\partial \mathcal{L}}{\partial \lambda_1}\\
&\quad -\frac{(\nu+2)/2}{\left(1+\frac{A}{\nu}\right)}\frac{1}{\nu}\frac{\partial A}{\partial \lambda_1},
\end{aligned}
$$

**Derivative of the joint distribution with respect to the asymmetric parameter of the second variable.** The derivative $\frac{\partial x_2}{\partial \lambda_2} = \frac{\partial F_{X_2}^{-1}(u;\lambda_2,\nu)}{\partial \lambda_2}$ is computed numerically.

$$
\begin{aligned}
\frac{\partial a_1}{\partial \lambda_2} &= 0 & \frac{\partial b_3}{\partial \lambda_2} &= = 2\lambda_1 \frac{-\rho}{1-\rho^2} \\
\frac{\partial a_2}{\partial \lambda_2} &= \frac{2(x_2-\mu_2)\frac{\partial x_2}{\partial \lambda_2}}{1-\rho^2} & \frac{\partial B}{\partial \lambda_2} &= \frac{\partial b_1}{\partial \lambda_2} + \frac{\partial b_2}{\partial \lambda_2} + \frac{\partial b_3}{\partial \lambda_2} \\
\frac{\partial a_3}{\partial \lambda_2} &= 2\frac{\partial x_2}{\partial \lambda_2}(x_1-\mu_1)\frac{-\rho}{1-\rho^2} & \frac{\partial \ell_1}{\partial \lambda_2} &= 0 \\
\frac{\partial A}{\partial \lambda_2} &= \frac{\partial a_1}{\partial \lambda_2} + \frac{\partial a_2}{\partial \lambda_2} + \frac{\partial a_3}{\partial \lambda_2} & \frac{\partial \ell_2}{\partial \lambda_1} &= \frac{\left(\frac{\partial x_2}{\partial \lambda_2}\lambda_2 + (x_2-\mu_2)\right)}{1-\rho^2} \\
\frac{\partial b_1}{\partial \lambda_2} &= 0 & \frac{\partial \ell_3}{\partial \lambda_2} &= \left(\frac{\partial x_2}{\partial \lambda_2}\lambda_1 + (x_1-\mu_1)\right)\frac{-\rho}{1-\rho^2} \\
\frac{\partial b_2}{\partial \lambda_2} &= \frac{2\lambda_2}{1-\rho^2} & \frac{\partial \mathcal{L}}{\partial \lambda_2} &= \frac{\partial \ell_1}{\partial \lambda_2} + \frac{\partial \ell_2}{\partial \lambda_2} + \frac{\partial \ell_3}{\partial \lambda_2}
\end{aligned}
$$

**Derivative of the joint distribution with respect to the number of degrees of freedom**
González-Santander (2023) shows that the n-th derivative of the modified Bessel function of the second kind is

$$
\frac{\partial^n}{\partial \nu^n} K_\nu(t) = \frac{1}{2}\int_{-\infty}^{\infty} x^n \exp(\nu x - t\cosh x)dx,
$$

where cosh is the hyperbolic cosine. For n=1, we can find a closed-form solution for non-integral $\nu$ (see Brychkov 2016, González-Santander 2018) defined as

$$
\begin{aligned}
\frac{\partial K_\nu(t)}{\partial \nu} = {} & \frac{\pi}{2}\csc(\pi\nu)\left\{\pi\cot(\pi\nu)I_\nu(z) - [I_\nu(z)+I_{-\nu}(z)]\left(\frac{z^2}{4(1-\nu^2)}\,{}_3\mathbf{F}_4\left(\begin{bmatrix}1,1,\frac{3}{2}\\ 2,2,2,-\nu,2+\nu\end{bmatrix}|z^2\right) + \log(\frac{z}{2}) - \psi(\nu) - \frac{1}{2\nu}\right)\right\} \\
& + \frac{1}{4}\left\{I_{-\nu}(z)\Gamma^2(-\nu)\left(\frac{z}{2}\right)^{2\nu}{}_2\mathbf{F}_3\left(\begin{bmatrix}\nu,\frac{1}{2}+\nu\\ 1+\nu,1+\nu,1+2\nu\end{bmatrix}|z^2\right) - I_\nu(z)\Gamma^2(\nu)\left(\frac{z}{2}\right)^{-2\nu}{}_2\mathbf{F}_3\left(\begin{bmatrix}-\nu,\frac{1}{2}-\nu\\ 1-\nu,1-\nu,1-2\nu\end{bmatrix}|z^2\right)\right\},
\end{aligned}
$$

where ${}_p\mathbf{F}_q$ is the generalized hypergeometric function and $\psi(\dots)$ is the digamma function. The derivatives $\frac{\partial x_1}{\partial \lambda_1} = \frac{\partial F_{X_1}^{-1}(u;\lambda_1,\nu)}{\partial \nu}$ and $\frac{\partial x_2}{\partial \nu} = \frac{\partial F_{X_2}^{-1}(u;\lambda_2,\nu)}{\partial \nu}$ are computed numerically.

$$
\begin{aligned}
\frac{\partial a_1}{\partial \nu} &= \frac{2(x_1-\mu_1)\frac{\partial x_1}{\partial \nu}}{1-\rho^2} & \frac{\partial b_3}{\partial \nu} &= 0 \\
\frac{\partial a_2}{\partial \nu} &= \frac{2(x_2-\mu_2)\frac{\partial x_2}{\partial \nu}}{1-\rho^2} & \frac{\partial B}{\partial \nu} &= \frac{\partial b_1}{\partial \nu} + \frac{\partial b_2}{\partial \nu} + \frac{\partial b_3}{\partial \nu} \\
\frac{\partial a_3}{\partial \nu} &= 2\left((x_2-\mu_2)\frac{\partial x_1}{\partial \nu} + (x_1-\mu_1)\frac{\partial x_2}{\partial \nu}\right)\frac{-\rho}{1-\rho^2} & \frac{\partial \ell_1}{\partial \nu} &= \frac{\left(\frac{\partial x_1}{\partial \nu}\lambda_2\right)}{1-\rho^2} \\
\frac{\partial A}{\partial \nu} &= \frac{\partial a_1}{\partial \nu} + \frac{\partial a_2}{\partial \nu} + \frac{\partial a_3}{\partial \nu} & \frac{\partial \ell_2}{\partial \nu} &= \frac{\left(\frac{\partial x_2}{\partial \nu}\lambda_2\right)}{1-\rho^2} \\
\frac{\partial b_1}{\partial \nu} &= 0 & \frac{\partial \ell_3}{\partial \nu} &= \left(\frac{\partial x_2}{\partial \nu}\lambda_1 + \frac{\partial x_1}{\partial \nu}\lambda_2\right)\frac{-\rho}{1-\rho^2} \\
\frac{\partial b_2}{\partial \nu} &= 0 & \frac{\partial \mathcal{L}}{\partial \nu} &= \frac{\partial \ell_1}{\partial \nu} + \frac{\partial \ell_2}{\partial \nu} + \frac{\partial \ell_3}{\partial \nu}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \log\left(f(x_1,x_2;\Theta)\right)}{\partial \nu} = {} & \frac{1}{2}\log(2) + \psi(\frac{\nu}{2})/2 - \frac{1}{\nu} \\
& \frac{\frac{\partial K_{\frac{\nu+2}{2}}\left(((\nu+A)B)^{\frac{1}{2}}\right)}{\partial \nu}\frac{1}{2} + \frac{1}{2}k_{\frac{\nu+2}{2}}\left(((\nu+A)B)^{\frac{1}{2}}\right)((\nu+A)B)^{\frac{-1}{2}}\left[(\nu+A)\frac{\partial B}{\partial \nu} + (1+\frac{\partial A}{\partial \nu})B\right]}{K_{\frac{\nu+2}{2}}\left(((\nu+A)B)^{\frac{1}{2}}\right)} + \frac{\partial \mathcal{L}}{\partial \nu} \\
& + \frac{\log\left([(\nu+A)B]^{\frac{1}{2}}\right)}{2} + (\nu+2)\frac{[(\nu+A)B]^{\frac{-1}{2}}\left\{(\nu+A)\frac{\partial B}{\partial \nu} + (1+\frac{\partial A}{\partial \nu})B\right\}}{[(\nu+A)B]^{\frac{1}{2}}} \\
& - \frac{\log\left(1+\frac{A}{\nu}\right)}{2} - \frac{(\nu+2)/2}{\left(1+\frac{A}{\nu}\right)}\left[\frac{\partial A}{\partial \nu}\nu + A\right]
\end{aligned}
$$

## Hessian of the factor copula model

We could apply the rule of chain on Eq. (26) to get the second derivative

$$
\frac{\partial^2 \log(c(u_t,v_t;\Theta))}{\partial\theta_1\partial\theta_2} = \frac{-1}{c(u_t,v_t;\Theta)^2}\frac{\partial c(u_t,v_t;\Theta)}{\partial\theta_1}\frac{\partial c(u_t,v_t;\Theta)}{\partial\theta_2} + \frac{1}{c(u_t,v_t;\Theta)}\frac{\partial^2 c(u_t,v_t;\Theta)}{\partial\theta_1\partial\theta_2}. \tag{29}
$$

We apply again the differentiation under the integral sign as we did for the gradient to obtain the second derivative with respect to the parameters of the Skewed t distribution. For numerical minimization in a high dimensional dataset we need the gradient and hessian to use a modified Newton-Raphson algorithm, where the gradient indicates the direction to follow to minimize the minus log-likelihood and the hessian indicates the size of the step in that direction.

## Construction of systemic risk measure

### Simulation-based

Pseudo-code of the computation of the systemic indicators by simulation (1/4)

```
U=rand(W,M+6); % simulate W realizations of M firms, 5 regional latent factors
               % and one global latent factor
for t=1:T+1 % for each time period of length T and the one-week forecast of
            % the model
    Uf=Cinv(U(:,end-5:end-1),U(:,end)); %generate the dependence pattern
        % between the 5 regional latent factors and the global latent factor
    for k=1:5 % for each region we get the dependence structure of the
              % corresponding financial firms, M_g is a vector indicating the
              % number of institutions under each region so sum(M_g)=M
        Ui(:,1+sum(M_g(1:k-1)):sum(M_g(1:k)))=...
        Cinv(U(:,1+sum(M_g(1:k-1)):sum(M_g(1:k))),Uf(:,k));
    end
    for k=1:M % now for the financial firms get the returns, applying the
              % inverse marginal cumulative distribution function and the
              % mean and standard deviation from the ARMA-GARCH-GJR
        R(:,k)=Finv(Ui(:,k))sigma(t,k)+mu(t,k)
    end
    % now we start building the FSI
    for i_scen=1:5 %for each region, we create a set of scenarios
        set_shockedregions=[1:5];
        set_shockedregions(i_scen)=[]; % we consider the effect of a region
                                       % on the rest of the regions
        for i_shocked=set_shockedregions
            pts=ones(M_g(i_scen),M_g(i_scen))*100 % we set a high threshold
                        % such that the probability of being below the
                        % threshold is close to 100%
            for j=[1:M_g(i_scen)]
                pts(j,j)=kappa(1,j+sum(M_g(1:i_scen-1)));
                % kappa is a vector 1xM indicating the unconditional lowest
                % 5-th percentile of returns
            end
            % each row of pts indicates a scenario and the columns indicate
            % the institutions in distress on that scenario
            P=zeros(1,M_g(i_scen)) % vector where we will gather the
                                   % probability of that scenario
            for j=[1:M_g(i_scen)]
                P(j)=sum(all(R(:,1+sum(M_g(1:i_scen-1)):sum(M_g(1:i_scen)))<...
                pts(j,:),2))./W;
            end
            group_feasible=find(P>=alpha_threshold)'; % we would consider only
                        % those scenarios where the probability of occurance
                        % is higher than alpha threshold, which in our
                        % empirical exercise we set at 4%
            n_g=size(group_feasible,1); % here we aggregate the number of
                                        % feasible scenarios to be considered
            pts=ones(n_g,M_g(i_scen))*100;
            check_max=[]; % sum of conditional
                                              % probabilities
            check_return=[]; % index return of the
                                        % weighted-average financial firms
            filter_sim=[]; % realizations where the conditioning event is
                           % materialized
```

Pseudo-code of the computation of the systemic indicators by simulation (2/4)

```
for j=[1:n_g] % for each of the feasible scenarios
    pts(j,group_feasible(j,:))=kappa(1,group_feasible(j,:)+...
        sum(M_g(1:i_scen-1))); %generate the stress scenario
    Index=(ww(end-(TT+1)+t,1+sum(M_g(1:i_shocked-1)):...
        sum(M_g(1:i_shocked)))*(R(:,1+sum(M_g(1:i_shocked-1)):...
        sum(M_g(1:i_shocked)))'))';
    % we build the index for the firms to generate the FIJD
    % we use the market cap in ww as weighting factor
    filter_sim=find(all(R(:,1+sum(M_g(1:i_scen-1)):...
        sum(M_g(1:i_scen))<pts(j,:),2)); % we keep the realizations
        % within the W simulations where the scenario is happening
    check_max(j)=sum(sum([R(filter_sim,1+...
        sum(M_g(1:i_shocked-1)):sum(M_g(1:i_shocked)))<...
        (ones(length(filter_sim),1)* kappa.(names_fl{i1}))],1)./...
        length(filter_sim));
    % we get the sum of conditional probabilities
    check_return(j)=mean(Index(filter_sim,1));
    % we get the performance of the financial index
    % under the selected scenario
end
% now we keep the scenario that generates the maximum sum of
% conditional probabilities
% (for the FSI) and the minimum return (for the FIJD)
pos_best_fsi=min(find(check_max==max(check_max)));
set_fsi_best=max(check_max);
pos_best_fijd=min(find(check_return==min(check_return)));
set_fijd_best=min(check_return);
set_fsi_group=group_feasible(pos_best_fsi);
set_fijd_group=group_feasible(pos_best_fijd);
% we use a couple of indicators to start a loop to look for the
% optimal scenario fro the FSI and the FIJD
goset_fsi=1;
goset_fijd=1;
% we start the search of the scenario for the FSI
   while goset_fsi==1
    g=length(set_fsi_group);
    comb_set=[1:M_g(i_scen)];
    comb_set(set_fsi_group)=[];
    n_g=length(comb_set);
    pts=ones(n_g,M_g(i_scen))*100;
    P=[];
    for j=1:size(pts,1)
        pts(j,[set_fsi_group,comb_set(j)])=...
            kappa([set.fsi.group.(names_fl{i1}),comb_set(j)]+...
            sum(M_g(1:i_scen)));
        P(j)=sum(all(R(:,1+sum(M_g(1:i_scen-1)):...
            sum(M_g(1:i_scen)))<pts(j,:),2))./W;
    end
    group_feasible=find(P>=alpha_threshold)';
    % if we don't find any feasible scenario we leave the loop
    if isempty(group_feasible)
        goset_fsi=0;
    else
        n_g=size(group_feasible,1);
        comb_set=comb_set(group_feasible);
        pts=ones(n_g,M_g(i_scen))*100;
        check_max=[];
        filter_sim=[];
        % we start analysing the evolution of the conditional
        % probabilities under the feasible scenarios
```

Pseudo-code of the computation of the systemic indicators by simulation (3/4)

```
            for j=[1:n_g];
                pts(j,[set_fsi_group,comb_set(j)])=...
                    kappa([set_fsi_group,comb_set(j)]+...
                    sum(M_g(1:i_scen-1)));
                filter_sim=find(all(R(:,1+sum(M_g(1:i_scen-1)):...
                    sum(M_g(1:i_scen)))<pts(j,:),2));
                check_max(j)=sum(sum([R(filter_sim,1+...
                    sum(M_g(1:i_shocked-1)):sum(M_g(1:i_schoked)))<...
                    (ones(length(filter_sim),1)*kappa(1+...
                    sum(M_g(1:i_shocked-1)):...
                    sum(M_g(1:i_schoked))))],1)./length(filter_sim));
            end
            % if the new scenario is not better than the first
            % one we stop the loop
            if max(check_max)<set_fsi_best
                goset_fsi=0;
            else
                pos_best_fsi=min(find(check_max==max(check_max)));
                set_fsi_best=max(check_max);
                set_fsi_group=[set_fsi_group,comb_set(pos_best_fsi)];
            end
        end
        if M_g(i_scen)==length(set_fsi_group)
                goset_fsi=0;
         end
    end
    % now we follow the same process for the FIDJ
    while goset_fijd==1
        g=length(set_fijd_group);
        comb_set=[1:M_g(i_scen)];
        comb_set(set_fijd_group)=[];
        n_g=length(comb_set);
        pts=ones(n_g,M_g(i_scen))*100;
        P=[];
        for j=1:size(pts,1)
            pts(j,[set_fijd_group,comb_set(j)])=...
                kappa([set.fijd.group.(names_fl{i1}),comb_set(j)]+...
                sum(M_g(1:i_scen-1)));
            P(j)=sum(all(R(:,1+sum(M_g(1:i_scen-1)):...
                sum(M_g(1:i_scen)))<pts(j,:),2))./W;
        end
        group_feasible=find(P>=alpha_threshold)';
        % if there is no feasible scenario anymore, we stop the loop
        if isempty(group_feasible)
            goset_fijd=0;
        else
            n_g=size(group_feasible,1);
            comb_set=comb_set(group_feasible);
            pts=ones(n_g,M_g(i_scen))*100;
            check_return=[];
            filter_sim=[];
            for j=[1:n_g];
                pts(j,[set_fijd_group,comb_set(j)])=...
                    kappa([set_fijd_group,comb_set(j)]+...
                    sum(M_g(1:i_scen-1)));
                filter_sim=find(all(R(:,1+sum(M_g(1:i_scen-1)):...
                    sum(M_g(1:i_scen)))<pts(j,:),2));
                check_return(j)=mean(Index(filter_sim,1));
            end
            % if adding another institution doesn't increase the
            % losses of the index, we stop the loop
```

Pseudo-code of the computation of the systemic indicators by simulation (4/4)

```
                    if (min(check_return)>set_fijd_best)
                        goset_fijd=0;
                    else
                        pos_best_fijd=...
                            min(find(check_return==min(check_return)));
                        set_fijd_best=min(check_return);
                        set_fijd_group=[set_fijd_group ,...
                            comb_set(pos_best_fijd)];
                    end
                    if M_g(i_scen)==length(set_fijd_group)
                        goset_fijd=0;
                    end

                end
            end
            % we save the measures
            FSI(t).(sprintf("CountryUnderStress_%i",i_scen)). ...
                (sprintf("CountryAffected_%i",i_shocked)). ...
                scenario=set_fsi_group;
            FIJD(t).(sprintf("CountryUnderStress_%i",i_scen)). ...
                (sprintf("CountryAffected_%i",i_shocked)). ...
                scenario=set_fijd_group;
            FSI(t).(sprintf("CountryUnderStress_%i",i_scen)). ...
                (sprintf("CountryAffected_%i",i_shocked)). ...
                outcome=set_fsi_best;
            FIJD(t).(sprintf("CountryUnderStress_%i",i_scen)). ...
                (sprintf("CountryAffected_%i",i_shocked)). ...
                outcome=set_fijd_best;
        end % this closes the loop over the conditioned contries
    end % this closes the loop over the stress regions
end % this closes the loop over time
```

### Based on the copula structure

The probability of a subset $A$ in Eqs. (2) and (4) is defined as

$$P(A) = \int_0^1 \prod_{i \in A} C(u_i^* | v_g) dv_g, \tag{30}$$

where we assume that the institutions in the set $A$ are within the same regional latent factor. $u_i^* = F_i(\frac{\kappa_i - \mu_{i,t}}{\sigma_{i,t}})$. If the set of institution were in different $M$ regional latent factors the formula would be

$$P(A) = \int_0^1 \int_0^1 \prod_{i \in A_1} C(u_i^* | v_{a_1}) c(v_{a_1}, v_G) dv_{a_1} \cdots \int_0^1 \prod_{j \in A_M} C(u_i^* | v_{a_M}) c(v_{a_M}, v_G) dv_{a_M} dv_G.$$

The formula for the $FSI$ in terms of copulas would be

$$FSI = \frac{1}{\alpha} \sum_{j \in B} \int_0^1 \int_0^1 C(u_j^* | v_b) c(v_b, v_G) dv_b \int_0^1 \prod_{i \in A} C(u_i^* | v_a) c(v_a, v_G) dv_a dv_G. \tag{31}$$

The formula for the $FIJD$ in terms of copulas would be

$$\begin{aligned} FIJD = & \sum_{j \in B} \omega_j E(r_j | A) \\ & \frac{1}{\alpha} \sum_{j \in B} \omega_j \int_0^1 \left[ F_j^{-1}(u_j) \sigma_{j,t} + \mu_{j,t} \right] \int_0^1 \int_0^1 c(u_j, v_b) c(v_b, v_G) dv_b \\ & \int_0^1 \prod_{i \in A} C(u_i^* | v_a) c(v_a, v_G) dv_a dv_G du_j, \end{aligned} \tag{32}$$

where $\omega_j$ is the weight of institution $j$ in the set $B$.

The vulnerability index is defined in Eq. (5) as the sum of the pairwise probability of distress. If two institutions $i$ and $j$ belong to the same regional latent factor, i.e. $i, j \in A$, the formula would be

$$P(r_{i,t} \leq \kappa_i, r_{j,t} \leq \kappa_j) \quad = \quad \int_0^1 C(u_j^*|v_a)C(u_i^*|v_a)dv_a, \tag{33}$$

where $u_j^* = F_j(\frac{\kappa_j - \mu_{j,t}}{\sigma_{j,t}})$ and $u_i^* = F_i(\frac{\kappa_i - \mu_{i,t}}{\sigma_{i,t}})$. If two institutions $i$ and $j$ belong to different regional latent factors, e.g. $i \in A$ and $j \in B$, the formula would be

$$P(r_{i,t} \leq \kappa_i, r_{j,t} \leq \kappa_j) \quad = \quad \int_0^1 \int_0^1 C(u_i^*|v_a)c(v_a, v_G)dv_a \int_0^1 C(u_j^*|v_b)c(v_b, v_G)dv_b dv_G. \tag{34}$$

The Expected Shortfall decomposition in Eq. (7) relies in the conditional probability $E(r_i|r_i \leq \kappa_i, r_j \leq \kappa_j)$ and the conditional expected return $P(r_j \leq \kappa_j|r_i \leq \kappa_i)$. If two institutions $i$ and $j$ belong to the same regional latent factor, i.e. $i, j \in A$, the formula for the conditional probability would be

$$P(r_{j,t} \leq \kappa_j|r_{i,t} \leq \kappa_i) \quad = \quad \frac{1}{u_i^*} \int_0^1 C(u_j^*|v_a)C(u_i^*|v_a)dv_a, \tag{35}$$

while the formula for the conditional expected return is

$$\begin{aligned} E(r_{i,t}|r_{i,t} \leq \kappa_i, r_{j,t} \leq \kappa_j) \quad = \quad & \frac{1}{\int_0^1 C(u_j^*|v_a)C(u_i^*|v_a)dv_a} \\ & \int_0^{u_i^*} \left[F_i^{-1}(u_i)\sigma_{i,t} + \mu_{i,t}\right] \int_0^1 c(u_i, v_a)C(u_j^*|v_a)dv_a du_i. \end{aligned} \tag{36}$$

If two institutions $i$ and $j$ belong to different regional latent factors, e.g. $i \in A$ and $j \in B$, the formula for the conditional probability would be

$$P(r_{j,t} \leq \kappa_j|r_{i,t} \leq \kappa_i) \quad = \quad \frac{1}{u_i^*} \int_0^1 \int_0^1 C(u_i^*|v_a)c(v_a, v_G)dv_a \int_0^1 C(u_j^*|v_b)c(v_b, v_G)dv_b dv_G, \tag{37}$$

while the formula for the conditional expected return is

$$\begin{aligned} E(r_{i,t}|r_{i,t} \leq \kappa_i, r_{j,t} \leq \kappa_j) \quad = \quad & \frac{1}{\int_0^1 \int_0^1 C(u_i^*|v_a)c(v_a, v_G)dv_a \int_0^1 C(u_j^*|v_b)c(v_b, v_G)dv_b dv_G} \\ & \int_0^{u_i^*} \left[F_i^{-1}(u_i)\sigma_{i,t} + \mu_{i,t}\right] \int_0^1 \int_0^1 c(u_i, v_a)c(v_a, v_G)dv_a \int_0^1 C(u_j^*|v_b)c(v_b, v_G)dv_b dv_G du_i. \end{aligned} \tag{38}$$