Extremal connectedness of hedge funds*

L. Mhalla[†], J. Hambuckers[‡] and M. Lambert[‡]

Abstract

We propose a dynamic measure of extremal connectedness tailored to the short reporting period and unbalanced nature of hedge funds data. Using multivariate extreme value regression techniques, we estimate this measure conditional on factors reflecting the economic uncertainty and the state of the financial markets, and derive risk indicators reflecting the likelihood of extreme spillovers. Empirically, we study the dynamics of tail dependencies between hedge funds grouped per investment strategies, as well as with the banking sector. We show that during crisis periods, some pairs of strategies display an increase in their extremal connectedness, revealing a higher likelihood of simultaneous extreme losses. We also find a sizable tail dependence between hedge funds and banks, indicating that banks are more likely to suffer extreme losses when the hedge fund sector does. Our results highlight that a proactive regulatory framework should account for the dynamic nature of the tail dependence and its link with financial stress.

Keywords: extreme value theory, systemic measure, tail dependence measure.

1 Introduction

In the aftermath of the Global Financial Crisis, European and U.S. market authorities have become interested in the contribution of unregulated hedge funds to global financial instability (Ang et al. 2011, Billio et al. 2012). Their preoccupation stems from the high probability that these funds suffer extreme losses, which in turn would cause funds' closures, fire-sales of assets, and eventually, financial distress of systemically important institutions (King & Maier 2009).

*Additional Supporting Information may be found online at `https://github.com/lindamhalla/Extremal-connectedness-of-hedge-funds`

[†]HEC Lausanne, University of Lausanne, Switzerland
[‡]University of Liège – HEC Liège, Belgium

The high probability of extreme losses among hedge funds is mainly attributable to exotic investment strategies that make funds highly susceptible to liquidity restrictions (Agarwal et al. 2017). Typically, funds buy illiquid securities while shorting liquid ones, making them exposed to huge margin calls likely to generate large losses through fire-sales (Mitchell & Pulvino 2012). Besides, hedge funds returns are also characterized by a strong *connectedness* across investment styles, in particular in crisis periods. Boyson et al. (2010) trace back this dependence to liquidity spirals affecting all assets owned by speculators and triggered by capital providers withdrawing their funds. Thus, in case of extreme losses suffered by funds with a given investment style, funds following other styles are more likely to endure a similar fate. This connectedness amplifies the shocks transmitted to the banking sector via their brokerage activities; see for instance the recent demise of the Archegos fund (WSJ 2021).

Due to the existence of significant extreme losses, connectedness of hedge funds is undoubtedly linked to the notion of *concurrent* extreme events. However, this feature has been remarkably overlooked in previous studies on the interdependence between funds, focusing mostly on commonalities in the expected returns. Therefore, a main contribution of this paper is to develop an econometric set-up to measure the *time-varying extremal connectedness* between investment styles of hedge funds. To do so, we combine regression modeling with multivariate extreme value (MEV) theory. With our approach, we are able to compute probabilities of concurrent extremely negative returns between two investment styles, given particular market conditions and uncertainty levels. These quantities are then used to analyse the extremal network between hedge funds and construct risk indicators capturing the vulnerability of the financial system coming from funds interconnectedness. In addition, our modeling framework can be exploited to measure connectedness between hedge funds and banks, with the aim of gauging potential spillovers from hedge funds to the banking sector.

Thanks to our MEV-based approach, we overcome the limitations of recently proposed measures of connectedness that do not explicitly focus on extreme events, but rather focus on time-varying means (Billio et al. 2012), variance decompositions (Diebold & Yilmaz 2009), or conditional quantiles (Boyson et al. 2010) of the returns. Moreover, we provide correct inference on the joint tail distribution of hedge funds returns *beyond historical data*, a feature that is critical for systemic risk analysis and that only extreme value theory can handle. This is a desirable property for regulators when implementing cash and capital requirements with limited historical data. In addition, a clear mapping of the intensity of the tail dependence and its dynamic during market stress would be useful to improve risk management.

The use of MEV has been advocated as being particularly well-suited to study connectedness and financial crises, given the importance of extreme events (Hartmann et al. 2004). Conceptually speaking, classical MEV pursues the same objective as copula modeling, i.e., modeling the dependence structure between random variables, but with a focus on the tail of their joint distribution. In our proposed setting, we rely on MEV to construct risk indicators derived from probabilities of occurrence of joint extreme returns. Aggregated over various financial institutions, these indicators form risk measures. This technique was successfully applied, e.g., to stock indices (Poon et al. 2004), U.S. depository institutions (van Oordt & Zhou 2019) and crash risk of stocks (van Oordt & Zhou 2016). However, in the case of hedge funds, we face two major issues that prevent us from using standard MEV. On the one hand, the highest reporting frequency of the returns is monthly, yielding few observations. On the other hand, a large fraction of funds are only reported for a short period of time (typically a few years). Consequently, applying MEV to pairs of funds would certainly leave us with very few observations to infer the marginal and joint distributions. In addition, previous studies implicitly acknowledge the time-varying nature of tail risk and tail dependence, e.g., via the use of rolling-window estimations (Agarwal et al. 2017).

To overcome these issues, we adapt the MEV in several ways. First, we group funds across investment styles to tackle the low reporting frequency. Thus, instead of working with individual time series of hedge funds returns, we build our analysis on panels of funds yielding vectors of extreme returns from a much larger pool compared to an index-based approach. Second, we rely on univariate extreme value regression techniques (Chavez-Demoulin et al. 2016, Hambuckers et al. 2018) to filter out fund-specific, macroeconomic and financial factors affecting the marginal tails. Finally, we account for the time-varying tail dependence through an MEV regression model (Mhalla et al. 2019) where the tail dependence between losses of investment styles depends on risk factors in a smooth and flexible manner. Since our risk factors of interest are collected over time, our tail dependence is *de facto* time-varying. This is an important improvement compared to previous studies that acknowledge the time-varying nature of the tail dependence but rather assume a fixed tail dependence over rolling windows.

We apply the proposed approach to the study of hedge funds registered in the Hedge Funds Research (HFR) database, over the period 1994–2017 and grouped in 12 investment styles. Our main findings are summarized in the following points:

- We find a strong average tail dependence between *Relative value*, *Event driven*, and *Equity market neutral* investment styles. On the contrary, *Short bias* investment style appears to be weakly connected at extreme levels with the other styles.

- Several investment styles show important changes in extremal connectedness over time and market conditions, especially during the Global Financial Crisis (GFC).
- Banks and hedge funds exhibit a sizeable tail dependence, indicating a non-negligible likelihood of joint bankruptcies, in particular at the onset of the GFC.

These findings have several implications. First, they suggest that tail risk can be diversified to a certain extend by combining adequately the various investment styles. Second, our analysis reveals discrepancies in risk profiles between investment styles that can be exploited to give a focus to regulators. Although an exhaustive screening of individual funds is hardly possible, our approach provides a compass indicating where the supervisory efforts must concentrate. Last, the substantial tail connectedness between banks and funds should be taken into account by supervision authorities wishing to decrease the consequence of spillovers in the banking sector.

The rest of the paper is organised as follows: in Section 2, we describe the theoretical foundations of our approach and detail the inference techniques. In Section 3, we describe our data, conduct both univariate and multivariate tail analyses, and discuss risk management implications. We conclude in Section 4.

2 Methodology

To infer about extreme funds returns that are beyond the range of observed values, we rely on multivariate extreme value theory (EVT). As highlighted in Poon et al. (2004), multivariate EVT requires first to estimate the marginal distributions of the different components. Then, the marginal components are accordingly scaled and the dependence structure is estimated using likelihood-based procedures. However, we are likely subject to marginal non-stationarity and heterogeneity, due to economic and fund-specific variables influencing the tail distributions of hedge funds (Agarwal et al. 2017). To account for these issues, we use a dynamic extreme value regression model in the initial filtering step, where the tail distribution depends on these variables. This preliminary step is followed by the modeling of the dynamic tail dependence structure with multivariate extreme value regression models, supposing that market risk factors smoothly affect the strength of the extremal dependence. The resulting measures of extremal connectedness are thus conditioned on these risk factors.

2.1 Marginal estimation of extreme value distributions

Let Y represent a random variable with cumulative distribution function F. Here, Y refers to losses of hedge funds with the same investment style, and the spotlight is put on the large values of these losses. The key assumption underlying extreme value modeling is that the distribution F belongs to the maximum domain of attraction (MDA) of an extreme value distribution G (Fisher & Tippett 1928). Under this fundamental assumption, the tail behavior of Y can be characterized using the POT approach (Davison & Smith 1990). Namely, for a fixed finite high threshold u, the distribution of the exceedances $Y - u \mid Y > u$ satisfies

$$\Pr(Y - u > y \mid Y > u) \underset{u \to y_F}{\longrightarrow} \begin{cases} (1 + \xi y/\sigma)_+^{-1/\xi}, & \xi \neq 0, \\ \exp(-y/\sigma), & \xi = 0, \end{cases}$$

where y_F is the upper endpoint of F. The limiting distribution of the threshold exceedances is a Generalized Pareto (GP) distribution, denoted $GPD(\cdot; \sigma, \xi)$, and is defined on $\{y : y > 0 \text{ and } (1 + \xi y/\sigma) > 0\}$. The shape parameter ξ governs the behavior of the tail of F, with positive values reflecting a heavy-tailed distribution. Thus, if the MDA assumption holds, one can infer the behavior of the losses far in the tails by estimating ξ and σ, and computing risk measures such as extreme quantiles. Inference is typically performed using likelihood-based methods. Starting with a set of observations $\{y_i\}_{i=1}^n$, we fix a high threshold u, e.g., the empirical quantile at a high probability, and retain solely the subset $\{y_i : y_i > u\}$. Then, inference for $\boldsymbol{\theta}_0 = (\sigma, \xi) \in \mathbb{R}_+ \times \mathbb{R}$ consists in maximising the log-likelihood

$$l(\boldsymbol{\theta}_0) = \sum_{i=1}^{n_u} \log\{GPD(y_i - u; \sigma, \xi)\}, \tag{1}$$

where n_u is the number of exceedances above u and $\{y_i\}_{i=1}^{n_u}$ is the set of reindexed exceedances.

The validity of (1) relies on the underlying assumption that the data generating process is stationary. This assumption does not hold in our application since we collect extreme returns over time and across various hedge funds with the same investment strategy. For instance, hedge funds' extreme losses might depend on a set of covariates that reflect relevant features such as the state of the economy or characteristics of the fund, see e.g., Agarwal & Naik (2004) and Boyson et al. (2010). To remedy this issue, dynamic approaches to extreme value modeling have been proposed by integrating a regression structure to EVT models (Chavez-Demoulin et al. 2016). This technique allows pooling data with different tail characteristics, circumventing issues related to missing data and small samples. In addition, it enables a preliminary filtering of the data that removes the heterogeneity in the upper tail specific to a given investment style, similarly to what is done in Boyson et al. (2010) for the mean.

Thus, we assume $\boldsymbol{\theta}_0$ to be function of a set of predictors and denote its components by $\sigma(\mathbf{x}_0^\sigma)$ and $\xi(\mathbf{x}_0^\xi)$. For simplicity of exposition, we consider additive linear functions where the parameters of the dynamic GP distribution are described by the following equations:

$$\log\{\sigma(\mathbf{x}_0^\sigma)\} = \beta_0^\sigma + \sum_{j=1}^{k_0} \beta_j^\sigma x_{0j}^\sigma \quad \text{and} \quad \xi(\mathbf{x}_0^\xi) = \beta_0^\xi + \sum_{j=1}^{k_0} \beta_j^\xi x_{0j}^\xi, \tag{2}$$

where $\mathbf{x}_0^\sigma = \left(x_{01}^\sigma, \dots, x_{0k_0}^\sigma\right)$ and $\mathbf{x}_0^\xi = \left(x_{01}^\xi, \dots, x_{0k_0}^\xi\right)$ are the vectors of covariates for the scale and shape parameters, respectively, and $\boldsymbol{\beta}_0^\sigma = \left(\beta_0^\sigma, \beta_1^\sigma, \dots, \beta_{k_0}^\sigma\right)$ and $\boldsymbol{\beta}_0^\xi = \left(\beta_0^\xi, \beta_1^\xi, \dots, \beta_{k_0}^\xi\right)$ denote the vectors of their corresponding regression coefficients.

Although estimates of $\boldsymbol{\beta}_0^\sigma$ and $\boldsymbol{\beta}_0^\xi$ are easily obtained using maximum likelihood procedures in the parametric case, Hambuckers et al. (2018) and Groll et al. (2019) highlight the need for an efficient model selection procedure to obtain interpretable sparse models. To do so, they advocate the use of the LASSO-type penalized likelihood given by

$$l(\boldsymbol{\theta}_0; \mathbf{x}_0^\sigma, \mathbf{x}_0^\xi, \lambda_0^\sigma, \lambda_0^\xi) = \sum_{i=1}^{n_u} \log\{GPD(y_i; \sigma(\mathbf{x}_0^\sigma), \xi(\mathbf{x}_0^\xi))\} + \lambda_0^\sigma \sum_{j=1}^{k_0} |\beta_j^\sigma| + \lambda_0^\xi \sum_{j=1}^{k_0} |\beta_j^\xi|, \tag{3}$$

where $\boldsymbol{\theta}_0$ is the stacked vector of regression coefficients $\boldsymbol{\beta}_0^\sigma$ and $\boldsymbol{\beta}_0^\xi$ and $(\lambda_0^\sigma, \lambda_0^\xi)$ the regularization parameters. The resulting estimator is denoted $\hat{\boldsymbol{\theta}}_0$. One might question the validity of the oracle properties of such LASSO-type approach, since the considered data generating process exhibits both tail heaviness and temporal dependence in the covariates, which violate the usual assumptions of sub-Gaussianity and independence (see, e.g., van de Geer & Bühlmann 2009). Here, we consider the LASSO penalty in a penalized likelihood framework where, as discussed in Fan & Li (2001), Negahban et al. (2012) and Lee et al. (2015), the resulting LASSO-type regularized estimator enjoys the oracle properties under regularity conditions on the likelihood function. The regularization parameters are usually chosen from a grid via information scores, e.g., the Bayesian information criterion (BIC). In Section 3.2, we allow for a large number of risk factors affecting both parameters, and thus use the LASSO-type estimator to infer the dynamic marginal distributions of threshold exceedances. The resulting estimates of $\sigma(\mathbf{x}_0^\sigma)$ and $\xi(\mathbf{x}_0^\xi)$ are used to infer the marginal tail risk of a given investment strategy. Note that, as shown in Fan & Li (2001), this approach results in biased estimates, yielding potentially biased $\hat{\sigma}(\mathbf{x}_0^\sigma)$ and $\hat{\xi}(\mathbf{x}_0^\xi)$. To correct for this bias, we use an additional estimation step where we retain only the active set obtained from the LASSO, i.e., the subset of covariates with non-zero estimates, and fit this model with classical maximum likelihood techniques. This two-step procedure is a common bias-correction technique (Belloni & Chernozhukov 2013) and gives the so-called post-LASSO estimator, which we denote by $\hat{\boldsymbol{\theta}}_0^+$. Thus, unless specified otherwise, estimates $\hat{\sigma}(\mathbf{x}_0^\sigma)$ and $\hat{\xi}(\mathbf{x}_0^\xi)$ are to be understood as being computed with $\hat{\boldsymbol{\theta}}_0^+$.

2.2 Multivariate dependence

From an extremal connectedness perspective, we are interested in the joint tail behavior of several random variables. In this multivariate setting, quantities related to both the structure and the degree of extremal dependence are of utmost importance.

We now consider $\mathbf{Y} = (Y_1, \ldots, Y_d)^\top$ to be a d-dimensional random vector with joint distribution F. Here, we interpret Y_j as a loss related to an investment style j, and $\mathbf{Y}$ as the vector of such losses in the d investment styles under consideration[1]. Similarly to the univariate EVT, inquiries about the joint upper tail behavior of $\mathbf{Y}$ rely on F being in the MDA of a non-degenerate distribution G. The limiting distribution G is then a multivariate extreme value distribution (Resnick 1987, Chapter 5.3) with associated extreme value copula C^{EV} and non-degenerate Generalized Extreme Value margins. Inference about the dependence structure in G, and hence the tail dependence structure in $\mathbf{Y}$, relies on the following representation of G under unit-Fréchet margins

$$G(\mathbf{z}) = \exp\left\{-\int_{S_d} \max\left(\frac{\omega_1}{z_1}, \ldots, \frac{\omega_d}{z_d}\right) dH(\boldsymbol{\omega})\right\}, \quad \mathbf{z} \in (0, \infty)^d, \tag{4}$$

where H is a positive finite measure defined on the unit simplex $S_d = \{(\omega_1, \ldots, \omega_d) \in [0,1]^d : \omega_1 + \cdots + w_d = 1\}$ and verifying $\int_{S_d} \omega_j dH(\boldsymbol{\omega}) = 1$, for $j = 1, \ldots, d$. The measure H is termed the spectral measure and contains the information regarding the structure of the dependence between the components of the limiting random vector, or equivalently the joint upper tail dependence of $\mathbf{Y}$. The basic modeling principle is to transform in a first step, the marginal variables Y_j to the unit-Fréchet scale, i.e., $Y_j^F = -1/\log\{F_j(Y_j)\}$. The focus of the second step is then on the dependence structure, where we rely on the limiting independence of the radial and pseudo-angular parts of the transformed variables. That is,

$$\Pr\left(\mathbf{W} \in B, R > tr \mid R > t\right) \underset{t \to \infty}{\longrightarrow} r^{-1} H(B), \quad B \in S_d, \; r \geqslant 1,$$

[1]For simplicity, we do not index Y_j with respect to a specific fund.

where $R = \|\mathbf{Y}^F\|_1$ and $\mathbf{W} = \mathbf{Y}^F/\|\mathbf{Y}^F\|_1$ are the radial and pseudo-angular components of the transformed vector $\mathbf{Y}^F = (Y_1^F, \ldots, Y_d^F)^\top$, respectively. The strength of the tail dependence in $\mathbf{Y}$ can be summarized through the pairwise coefficients of tail dependence (Heffernan 2000)

$$\chi^{kl} = \lim_{u \to 1} \Pr\left\{F_k(Y_k) > u \mid F_l(Y_l) > u\right\}. \tag{5}$$

Under the assumption of asymptotic dependence, the coefficient of tail dependence χ^{kl} is strictly positive for all pairs of components and the existence of a spectral density $h(\boldsymbol{\omega}) = dH(\boldsymbol{\omega})/d\boldsymbol{\omega}$ is assumed whenever H is absolutely continuous. Inference for the spectral density h is typically performed using likelihood-based methods where flexible classes of parametric models are assumed (Beirlant et al. 2004, Section 9.2.2). The choice of the parametric model for the tail dependence between investment strategies is discussed in Section 3.3.

Let $\boldsymbol{\theta}$ be the p-variate vector of parameters describing the spectral density $h(\cdot; \boldsymbol{\theta})$, and assume that we have a sample of n independent copies of $\mathbf{Y}^F$. Then, likelihood-based inference for $\boldsymbol{\theta}$ is performed using the observations with a radial component exceeding a high radial threshold u_R. Denoting the radial threshold exceedances by $\left\{\tilde{\mathbf{Y}}_i^F = (\tilde{Y}_{i,1}^F, \ldots, \tilde{Y}_{i,d}^F)^\top\right\}_{i=1}^{n_R}$, inference for the parameter of interest $\boldsymbol{\theta}$ is performed by maximising the following log-likelihood

$$\ell(\boldsymbol{\theta}) \equiv -(d+1)\sum_{i=1}^{n_R} \log \|\tilde{\mathbf{Y}}_i^F\|_1 + \sum_{i=1}^{n_R} \log\left\{h\left(\frac{\tilde{Y}_{i,1}^F}{\|\tilde{\mathbf{Y}}_i^F\|_1}, \ldots, \frac{\tilde{Y}_{i,d}^F}{\|\tilde{\mathbf{Y}}_i^F\|_1}; \boldsymbol{\theta}\right)\right\}. \tag{6}$$

When non-stationarity arises in a multivariate setting, both the marginal distributions and the dependence structure are believed to vary with covariates and flexible tail models are needed. For instance, tail dependence between stock losses might increase under a stress scenario (e.g., during financial crises); see Castro-Camilo et al. (2018). To deal with the effects of multiple covariates in a flexible manner, Mhalla et al. (2019) propose to model $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ in a spline-based fashion. Specifically, for a set of covariates $\mathbf{x} = (x_1, \ldots, x_K)$, the j-th component of $\boldsymbol{\theta}(\mathbf{x})$ is given by

$$\theta_j(\mathbf{x}; \boldsymbol{\beta}) = g_j^{-1}\left\{\beta_j + \sum_{i=1}^{K} f_{j,i}(x_i)\right\}, \quad j = 1, \ldots, p, \tag{7}$$

where g_j denotes a link function restricting θ_j to its parameter space, β_j is an intercept, and $f_{j,i}$ describes the effect of the i-th covariate x_i on θ_j. The choice of the function $f_{j,i}$ is flexible and can be either linear or smooth; see Wood (2017, Chapter 5) for a review on smoothing splines. This flexibility comes however at a price as it increases considerably the number of parameters to estimate. Inference for $\boldsymbol{\theta}(\mathbf{x};\boldsymbol{\beta})$ and hence for the effects of $\mathbf{x}$ on the upper tail dependence in $\mathbf{Y}^F$ should be performed by penalizing the log-likelihood (6), i.e.,

$$\max_{\boldsymbol{\beta},\boldsymbol{\gamma}} \ell\left\{\boldsymbol{\theta}(\mathbf{x};\boldsymbol{\beta})\right\} - \boldsymbol{\beta}^\top \mathbf{P}(\boldsymbol{\gamma})\boldsymbol{\beta}, \tag{8}$$

where the penalty matrix $\mathbf{P}(\boldsymbol{\gamma})$ controls, through a vector of smoothing parameters $\boldsymbol{\gamma}$, the curvature of the smooth functions to avoid over-fitting issues. Under certain regularity conditions outlined in Mhalla et al. (2019), the resulting penalized maximum likelihood estimator (PMLE) $\hat{\boldsymbol{\beta}}$ is consistent and asymptotically normal. Therefore, confidence intervals for $h(\cdot;\hat{\boldsymbol{\theta}}) \equiv h\{\cdot;\boldsymbol{\theta}(\mathbf{x},\hat{\boldsymbol{\beta}})\}$ can be built using the asymptotic variance of $\hat{\boldsymbol{\beta}}$.

Under this dynamic setting, the flexible smooth model (7) induces a non-stationary coefficient of tail dependence

$$\begin{aligned}\chi^{kl}(\mathbf{x}) &= \lim_{u\to 1} \Pr\left\{F_k(Y_k) > u \mid F_l(Y_l) > u; \mathbf{x}\right\} \\ &= 2 - \int_0^1 \max(\omega, 1-\omega) h\{\omega;\boldsymbol{\theta}(\mathbf{x};\boldsymbol{\beta})\} d\omega. \end{aligned} \tag{9}$$

Thus, an estimate $\hat{\chi}^{kl}(\mathbf{x})$ of the *conditional tail coefficient* $\chi^{kl}(\mathbf{x})$ obtained by plugging the PMLE $\hat{\boldsymbol{\beta}}$ in (9), enjoys the same nice asymptotic properties through an application of the Delta method. When the set of considered covariates describes entities evolving through time, we use the quantity (9) as a measure of the **time-varying extremal connectedness** between investment styles' universes. This measure can be related to the unconditional tail dependence across financial institutions estimated in Poon et al. (2004) and Balla et al. (2014). A similar tail dependence measure is advocated in Agarwal et al. (2017) and van Oordt & Zhou (2019), where emphasis is put on the dependence between an institution and the market, as opposed to the dependence between institutions themselves. Here, we use $\hat{\chi}^{kl}(\mathbf{x})$ to build the extremal network across investment styles as well as several systemic risk indicators; see Sections 3.3 and 3.4.

Following common practice in the copula modeling literature, the marginal modeling is performed in a separate step prior to the dependence modeling (Genest et al. 1995). In the above definition of $\chi^{kl}(\mathbf{x})$, we assume fixed marginal distributions and omit consequently any non-stationarity that might arise in the marginal investment styles. Therefore, we assume that a set of covariates $\mathbf{x}_0$ influences the individual tail behavior of the losses Y_j of the j-th investment style, and account for their influence using equation (2). Inference on $\chi^{kl}(\mathbf{x})$ is then based on the resulting pseudo marginal observations $\hat{F}_k(Y_k \mid \mathbf{x}_0)$ and $\hat{F}_l(Y_l \mid \mathbf{x}_0)$.

3 Empirical study

Here, we first detail the hedge funds data and present our fund-specific extremal connectedness analysis. Then, we discuss risk measures obtained from our approach. Finally, we analyse how the banking sector, represented by a sample of large banks, is connected with hedge funds.

3.1 Description of the hedge funds data

Our sample of hedge funds monthly returns comes from the Hedge Fund Research (HFR) database. We cover a period between January 1994 and May 2017. Comparative analyses across databases performed by Aiken et al. (2013) and Joenväärä et al. (2021) emphasize the quality of HFR with regard to the usual survivorship and backfill biases. To obtain our final sample, we apply a series of filters to the original dataset of historical returns, which we detail in Section 1.1 of the Supplementary material. As of May 2017, our final sample consists of 7,924 funds, among which 2,544 are active and 5,380 are dead, i.e., either liquidated or not reporting. Overall, we have 673,793 monthly observations.

In addition, HFR gives us the (self-reported) investment styles of the funds. The four broad classes (equity hedge, event driven, macro, and relative value) are subdivided into 12 investment styles as depicted in Table 1. We use this classification to decompose the dataset into 12 sub-universes across which the dynamic extremal connectedness is investigated. Descriptive statistics of the returns with respect to the investment styles are reported in Section 1.2 of the Supplementary material, where we observe important differences in all aspects of their respective distributions, especially in the skewness and kurtosis dimensions.

Equity Hedge	Event Driven	Macro	Relative Value
Long/Short equity Equity market neutral Short bias	Distressed/Restructuring Merger arbitrage Other	Global trading Managed futures	Fixed income arbitrage Convertible arbitrage Yield alternatives Multi-strategy

Table 1: HFR investment styles

3.2 Marginal tail risk of investment styles

First, we model the marginal distributions F_j of the observed losses $\{y_{j,i}\}_{i=1}^{n_j}$ of the j-th investment style, for $j = 1, \ldots, 12$ and $n_j = \sum_{t=1}^{T} n_j(t)$. Here, T is the number of available distinct time periods and $n_j(t)$ the number of funds reporting at time t and following investment style j. As the focus of the analysis is on the tail behavior within investment styles, we consider, for each strategy j, a semiparametric marginal approach where

$$\hat{F}_j(y \mid \mathbf{x}_0) = \begin{cases} F_j^{\text{emp}}(y), & y \leqslant u(\mathbf{t}_j), \\ \Pr\{Y_j \leqslant u(\mathbf{t}_j)\} + \Pr\{Y_j > u(\mathbf{t}_j)\} GPD\{y - u(\mathbf{t}_j); \hat{\sigma}_j(\mathbf{x}_0^{\sigma}), \hat{\xi}_j(\mathbf{x}_0^{\xi})\}, & y > u(\mathbf{t}_j), \end{cases} \tag{10}$$

and F_j^{emp} is the empirical distribution of $\{y_{j,i}\}_{i=1}^{n_j}$ and $u(\mathbf{t}_j)$ its 95% time-varying sample quantile. That is, denoting the time vector

$$\mathbf{t}_j = (\underbrace{1, \ldots, 1}_{\times n_j(1)}, \ldots, \underbrace{T, \ldots, T}_{\times n_j(T)}),$$

we fit the following parametric quantile regression

$$u(\mathbf{t}_j) = \underset{\beta_0, \beta_t \in \mathbb{R}}{\arg\min} \sum_{i=1}^{n_j} \rho_{0.95}\{y_{j,i} - (\beta_0 + \beta_t \mathbf{t}_{j,i})\},$$

where $\rho_\tau(u) = (\tau - \mathbf{1}_{u<0})u$ is the check function. Thus, $\Pr\{Y_j > u(\mathbf{t}_j)\} = 0.05$ for all strategies j and at all time periods t. The semiparametric transformation (10) prevents making assumptions on the distribution of the bulk of the observations, while improving the marginal empirical transformation in the region of interest, i.e., the tails, through the asymptotically motivated GP distribution. Such an approach is commonly used in multivariate extreme value modeling, see, e.g., Huser & Wadsworth (2019).

Relying on this semiparametric approach, we appropriately capture any heterogeneity in the large losses that might be due to macroeconomic, fund specific, and financial factors through $\mathbf{x}_0^{\sigma}$ and $\mathbf{x}_0^{\xi}$. We fix $\mathbf{x}_0^{\sigma} = \mathbf{x}_0^{\xi} = \mathbf{x}_0$, where $\mathbf{x}_0$ describes the set of selected risk factors driving the tail distribution of a fund with investment style j and at time period t[2]. Thus, the dynamic EVT part of the marginal model (10) yields time-dependent tail risk estimates within each investment style. For instance, for a given fund, we consider the set

$$\Big\{|R_{t-1}|, \texttt{IncFee}, \texttt{ManFee}, \texttt{TotAss}_{t-1}, \texttt{IntRateST}_{t-1}, \texttt{IntRateLT}_{t-1},$$
$$\texttt{Leverage}, \texttt{IndProd}_{t-1}, \texttt{UnempRate}_{t-1}, \texttt{StockVol}_{t-1},$$
$$\texttt{PTFSBD}_t, \texttt{PTFSFX}_t, \texttt{PTFSCOM}_t, \texttt{EMF}_t, \texttt{ESF}_t, \texttt{BD10RET}_t, \texttt{BAAMTSY}_t, \texttt{SMB}_t, \texttt{HML}_t, \texttt{MOM}_t\Big\},$$

where $|R_{t-1}|$ and $\texttt{TotAss}_{t-1}$ denote the absolute loss and total asset (in mio. USD) over the

[2]For notational simplicity, mention to the investment style j and time period t is omitted.

previous time period, respectively. $\texttt{Leverage}$ is a binary variable indicating whether the fund is leveraged or not, and $\texttt{IncFee}$ and $\texttt{ManFee}$ denote its incentive and management fees (in %), respectively. These variables were provided by HFR. The explanatory variables $\texttt{IntRateST}_{t-1}$, $\texttt{IntRateLT}_{t-1}$, $\texttt{IndProd}_{t-1}$, $\texttt{StockVol}_{t-1}$, and $\texttt{UnempRate}_{t-1}$ denote the level of short term and long term interest rates, the changes in industrial production, the stock market volatility, and the unemployment rate of the registration country provided by HFR, respectively. These macro data are gathered from the OECD database. Following Fung & Hsieh (2004), $\texttt{PTFSBD}_t$, $\texttt{PTFSFX}_t$, and $\texttt{PTFSCOM}_t$ denote the bond, currency, and commodity trend-following factors. $\texttt{EMF}_t$, $\texttt{ESF}_t$, $\texttt{BD10RET}_t$ and $\texttt{BAAMTSY}_t$ denote the equity, size spread, bond market, and credit spread factors, respectively. $\texttt{SMB}_t$, $\texttt{HML}_t$ and $\texttt{MOM}_t$ are the classical size, value, and momentum factors of Fama & French (1993) and Carhart (1997).

The leverage and incentive fees structure of hedge funds influence the asymmetry of their return payoff, impacting the upper and/or lower tail behavior (Titman & Tiu 2011). Besides, high management fees and the absence of a financial involvement of the managers also contribute to an increase in tail risk exposure (Karagiannis & Tolikas 2019). Unemployment rate, industrial production, interest rate levels, and stock volatility control for the general macroeconomic context, affecting risk exposure (Bali et al. 2014). Regarding $|R_{t-1}|$, we use it as a proxy for the standard deviation of returns at time $t-1$, a factor found to impact significantly the tail risk (Agarwal et al. 2017). In Section 5.1 of the Supplementary material, we provide a simulation study confirming the suitability of this approach to handle the dynamic tail of various heteroscedastic processes (see, e.g., Bee et al. 2019). Finally, we use the traditional risk factors to control for systematic risk exposure in the tail. The aim here is to filter out their effects from the marginal distribution, in the idea of Hale & Lopez (2019).

Estimates of $\boldsymbol{\beta}_0^\sigma$ and $\boldsymbol{\beta}_0^\xi$ rely on the post-LASSO estimator where the explanatory variables in $\mathbf{x}_0$ are standardized prior to the optimisation step. The regularization parameters are selected over a 30×30 grid of values, using the BIC criterion with degrees of freedom equal to the number of active covariates. Results are displayed and discussed more extensively in Section 2.1 of the Supplementary material, along with graphical goodness-of-fit tests.

How to measure marginal tail risk in our universes of hedge funds?

Using the semiparametric marginal model given by (10) and estimated with our panel approach, we compute the value-at-risk (VaR) at level α of the loss distribution for each fund observed at any time in our sample. Defining $\mathbf{x}_0^{i,k,t}$ and $\texttt{TotAss}^{i,k,t}$ as the risk factors and the total assets of the i-th fund with investment style k, alive at time t, the corresponding value-at-risk $VaR_\alpha^k(\mathbf{x}_0^{i,k,t})$ is defined as the α-th quantile of the marginal distribution $F_k(\cdot \mid \mathbf{x}_0^{i,k,t})$, i.e.,

$$VaR_\alpha^k(\mathbf{x}_0^{i,k,t}) = F_k^{-1}(\alpha \mid \mathbf{x}_0^{i,k,t}). \tag{11}$$

Thus, we obtain a tail risk measure for each existing fund, even when few or no extreme observations are available for this particular fund at a given date. We discuss the resulting tail risk profiles for the various investment styles in Section 2.1 of the Supplementary material.

3.3 Tail dependence between investment styles

Relying on the fitted marginal models of the losses for the 12 investment strategies, we transform the observations to the unit-Fréchet scale through

$$y_{j,i}^F = -\frac{1}{\log\{\hat{F}_j(y_{j,i} \mid \mathbf{x}_0)\}}, \quad j = 1, \ldots, 12, \text{ and } i = 1, \ldots, n_j,$$

and proceed with modeling their pairwise extremal connectedness. We refer to Section 5.2 of the Supplementary material for a sensitivity analysis of this step to marginal misspecification. We now quantify the extremal connectedness by the conditional tail coefficient $\chi^{kl}(\mathbf{x})$. Here, the set of covariates $\mathbf{x}$ consists of four factors reflecting the stress level on the financial markets over time: the CBOE volatilty index `VIX` reflecting financial instability, risk aversion, and the stock market volatility; the stock market performance as measured by the loss of the `MSCI` world index; the U.S. Economic Policy Uncertainty `EPU` index of Baker et al. (2016) reflecting general macroeconomic uncertainty; and the Financial Stress Index (`FSI`) of the Federal Reserve Bank of St. Louis (2014) associated with large shifts in asset prices, abrupt increase in risk, and illiquidity of the financial system. The frequency of the selected covariates is either daily (`VIX`, `MSCI`, and `EPU`) or weekly (`FSI`). Thus, in order to match the reporting frequency of the HFR database, we use monthly averages. Following Patton & Ramadorai (2013), we select these factors as they indicate conditions likely to drive managerial decisions of portfolio rebalancing. In particular, the MSCI accounts for "benchmarking pressures" that push fund managers to increase their exposure to the market. The VIX captures the fact that high market volatility can lead fund managers to reduce their risk exposures to keep the volatility of their own returns constant (Ferson & Schadt 1996). Finally, the EPU and the FSI control for changes in general uncertainty, cost of borrowing, and liquidity, which are factors shown to affect leverage and diversification decisions by fund managers (Cao et al. 2013).

We consider a pairwise approach to the tail connectedness where pairs of losses are constructed as follows. For a pair of strategies (k, l) and at a fixed time period $t = 1, \ldots, T$, we retain the n_t^{kl} largest rescaled losses in each strategy. Here, n_t^{kl} is the minimum between the number of funds in strategies k and l at the time period t. The reason for this pairwise approach is that it results in a more flexible tail modeling, as pairs are allowed to depend distinctly from each other on the covariates $\mathbf{x}$. The tail dependence of the resulting bivariate vector is modeled dynamically using a spectral density $h\{\cdot; \boldsymbol{\theta}^{kl}(\mathbf{x}_t)\}$, where the parameter $\boldsymbol{\theta}^{kl}(\mathbf{x}_t)$ depends smoothly on the selected global risk factors at time period t, i.e., $\mathbf{x}_t = (\mathtt{VIX}_t, \mathtt{MSCI}_t, \mathtt{EPU}_t, \mathtt{FSI}_t)$. For all the 66 pairs of investment strategies, we describe the spectral density using the Hüsler–Reiss parametric family (Hüsler & Reiss 1989) with a pair-specific dependence parameter $\theta^{kl}(\mathbf{x}_t) > 0$ describing the strength of tail dependence between the losses of the l-th and k-th investment styles at time period t. The reason for this choice of parametric family is twofold. First, the Hüsler–Reiss density is flexible as it captures different strengths of asymptotic dependence through the unique parameter $\theta^{kl}(\mathbf{x}_t)$. Second, from an inference perspective, the expected Hessian of this density can be obtained in closed form, thus avoiding numerical issues when optimizing the penalized log-likelihood; see Mhalla et al. (2019) for details. Note that, although dynamic MEV distributions are the most appropriate choice for capturing tail dependence when the behavior of the bulk is not relevant, we compared our resulting tail estimates to the ones obtained using a non-extreme but tail dependent t-copula. The entire dataset is modeled dynamically using the same covariate specification as in the Hüsler–Reiss model. We notice that this copula specification underestimates the tail dependence when compared to an empirical approach using rolling windows.

As mentioned in Section 2.2, our inference approach for the tail dependence relies on the set of bivariate pseudo-angular observations with a radial component exceeding a high radial threshold. Although the marginal data are assumed stationary after removal of the effect of marginal explanatory covariates, we model the radial threshold through a parametric quantile regression at the 95% level. The goal here is to capture any residual time non-stationarity in the data. We then retain the observations with a radial component exceeding this threshold to infer about the spectral density. Under the Hüsler–Reiss assumption, our estimate of the dynamic extremal connectedness between the pair of strategies (k, l) becomes

$$\hat{\chi}^{kl}(\mathbf{x}_t) = 2 - 2\Phi\{1/\hat{\theta}^{kl}(\mathbf{x}_t)\},$$

where $\hat{\theta}^{kl}$ is our estimate of θ^{kl} obtained by penalized log-likelihood; see Section 2.2.

We now look at our estimates of the tail dependence between the 66 pairs of investment styles. We focus first on the mean tail dependence (between each pair) over the entire period of analysis 1994–2017, represented in Figure 1 (left)[3].

[Figure 1 about here.]

Some main tendencies can be deduced. For instance, most *Relative value* investment styles seem strongly connected to each other, as well as to the *Other (Event driven)* and the *Equity market neutral* styles. An exception is the *Yield alternatives* investment style which appears to be weakly connected to the other ones. A similar observation holds for *Short bias* whose 11 associated pairwise conditional tail coefficients hint at weak asymptotic dependencies. From an economic standpoint, these results are consistent. As suggested in Boyson et al. (2010), hedge funds are strongly connected because they are all exposed to restrictions in funding liquidity which force them to reduce leverage and triggers a decrease in asset liquidity. When the initial liquidity shock is strong enough, it causes a self-reinforcing liquidity spiral. However, Sadka (2010) shows (although marginally) that exposure to liquidity shocks varies strongly across investment styles. In particular, the author finds that *Short bias* is the only investment style with a negative exposure to liquidity, thus benefiting from price decrease during liquidity shortage. This mechanism would explain why *Short bias* is found to be weakly connected with other styles. Moreover, the strong connection between *Event driven*, *Relative value*, and *Equity market neutral* can be related to the fact that all these strategies are equity-oriented (Agarwal & Naik 2004) and rely on some forms of "arbitrage positions" while not being exposed to a market of reference. For example, *Equity market neutral* encompasses statistical arbitrage techniques with no market exposure, whereas *Others (Event driven)* or *Convertible arbitrage* suppose taking long and short positions in equities and convertible debt instruments. Thus, although these strategies have been differently labelled, it is very likely that the portfolio compositions of these funds are close from one another. On the contrary, funds following a *Yield alternatives* strategy tend to enter the same bets, but targeting asset classes like real estates or infrastructure, known for their weak correlation with the equity market.

Thus, an investor aiming to diversify its tail risk might consider owning a portfolio mixing funds with, e.g., *Short bias* and *Yield alternatives* investment styles. These findings hold on average but are no longer valid under a stress scenario as discussed in what follows.

Does financial stress lead to changes in the extremal connectedness?

To investigate this question, we compute the empirical mean pairwise tail coefficients,

[3] For better visibility and unless otherwise mentioned, the subsequent network diagrams display the edges with values exceeding the median of all edges' values.

conditional on a given risk factor, e.g., $\texttt{VIX}_t$ exceeding its highest (empirical) quartile[4]. Figure 1 (right) displays these quantities for the VIX (graphs for the other factors can be found in Section 2.2 of the Supplementary material). We see that under these conditions, some pairs become more connected, e.g., *Managed futures* with *Multi-strategy* and *Global trading* with *Multi-strategy*, while others become less connected, e.g., *Equity market neutral* and *Merger arbitrage*. On the contrary, *Short bias* appears to be consistently loosely connected to the other investment styles.

However, these measures are related to a marginal stress condition and do not account for dependence across stress factors. To investigate changes in the extremal connectedness under realistic scenarios, we look at the tail dependence on 09/2008 (the beginning of the GFC) and at the mean tail coefficients over the time period 01/2013–12/2013 (a time period reflecting relatively stable market conditions, see, e.g., Lucas et al. (2017)). These quantities are displayed in Figure 2.

[Figure 2 about here.]

We clearly see differences in the intensity of the tail dependence between the various styles over the two time periods, highlighting the time-varying nature of our connectedness measure. To get an empirical measure of the variability in the tail dependence under historical stress conditions, we compute the differences between the mean tail dependence over the entire period of analysis and the tail dependence estimated during a financial distress. That is, for a given pair of investment styles[5], we compute an empirical estimate of

$$\mathrm{E}\{\chi(\mathbf{x}_t) \mid t \in \tilde{T}\} - \mathrm{E}\{\chi(\mathbf{x}_t) \mid t \in [1994-01-01, 2017-05-31]\}, \qquad (12)$$

for $\tilde{T} = 09/2008$, and $\tilde{T} = [01/2013, 12/2013]$. We display these differences in Figure 3.

[Figure 3 about here.]

Positive (resp. negative) values indicate an increase (resp. a decrease) in tail dependence with respect to the reference period. As expected, differences in the strength of the tail dependence are negligible when comparing the mean values in 2013 with the mean values over the entire period of analysis. Looking at these differences with respect to September 2008, we notice relatively large increases and decreases in tail dependence. For instance, one of

[4]This quantity is denoted $q_{0.75}(x^c)$ where x^c, for $c = 1, \ldots, 4$ describes each of our four risk factors.

[5]Mention to the pair in the superscript is omitted when no confusion can arise.

the strongest decreases in tail dependence is observed for the pair of investment styles *Fixed income arbitrage* and *Managed futures* with a difference equal to -0.42 (a relative difference of -87.8%). Among the pairs with increasing tail dependence during the GFC, the pair of investment styles *Merger arbitrage* and *Yield alternatives* displays the strongest increase with a difference equal to 0.42 (a relative difference of 121.7%).

We focus now on two pairs with a tail dependence that either decreased (*Fixed income arbitrage* and *Managed futures*) or increased (*Merger arbitrage* and *Yield alternatives*) at the beginning of the GFC. A decreasing tail dependence indicates that extreme losses are less likely to occur jointly, whereas an increasing tail dependence indicates that both styles suffer extreme losses simultaneously with a higher likelihood. Figure 4 shows the estimated tail dependence over time for these two pairs and their asymptotic (pointwise) 95% confidence intervals.

[Figure 4 about here.]

Interestingly, for the pair *Fixed income arbitrage* and *Managed futures*, a drop in tail dependence took also place in August 1998. Similarly to September 2008, August 1998 witnessed a major financial event, namely the default of Russia on its debt, and was followed by the bankruptcy of the fund LTCM. September 2008 and August 1998 are also characterised by record losses of the MSCI (-13.45% in 08/1998 and -12.08% in 09/2008). During these two time periods, the extremal connectedness reaches its smallest values (1.52×10^{-5} and 0.059 respectively), suggesting weak dependence of extremely large losses.

Examining the pair formed by *Merger arbitrage* and *Yield alternatives*, we observe a progressive build-up of the tail dependence from October 2003 onward until September 2008, where it reaches its historical maximum at 0.77. Hence, at that point in time, *Merger arbitrage* and *Yield alternatives* funds end up with the highest probability of jointly suffering extreme losses. A potential mechanism responsible for these changes is obviously the portfolio composition of the hedge funds (deriving directly from the strategy), and how they modify its content in time of market stress (Ben-David et al. 2012, Patton & Ramadorai 2013). When a large deleveraging takes place, the portfolios are concentrated on a smaller number of assets, leading either to stronger or weaker commonalities depending on how similar the portfolios are across strategies. Factors potentially responsible for magnifying this effect are dependencies and changes in leverage ratio over time. For instance, Ang et al. (2011) show that leverage decision depends strongly on anticipations of the general state of the economy, suggesting that funds share a common exposure via this mechanism. In the same idea, Patton & Ramadorai

(2013) show at the level of hedge funds indices that both the risk factors and their time-dynamics vary across investment styles. In particular, they find that a macro investment style like *Managed futures* loads mainly on the size ($\texttt{SMB}_t$) and bond market ($\texttt{BD10RET}_t$) factors, whereas relative values or event-driven strategies like *Fixed income arbitrage*, *Merger arbitrage*, and *Yield alternatives* are mostly exposed to the market and credit spread ($\texttt{BAAMTSY}_t$) factors. In addition, *Managed futures* funds have a rather constant exposure to these factors, whereas the other styles exhibit a time-varying exposure of the market (resp. credit spread) factor that is negatively (resp. positively) related to market liquidity, volatility and leverage. As a consequence, *Merger arbitrage* and *Yield alternatives* funds will deleverage their portfolio and adopt similar risk exposures in time of crisis, leading to an increase in the likelihood of joint extremely negative returns. On the contrary, funds adopting *Managed futures* or *Fixed income arbitrage* investment styles will load on different risk factors while deleveraging, leading to a decrease in the probability of joint extremes. The signs of the large variations in extremal connectedness for these two pairs (positive and negative, respectively) observed between 2013 and September 2008 would be consistent with this mechanism. Notice in addition that all our results are obtained after filtering out the dependence to risk factors in the marginal distribution of the returns.

To conclude this section, notice that our approach, although related in several ways to that of Balla et al. (2014) and van Oordt & Zhou (2019), exhibits important differences with these studies. Similarly to Balla et al. (2014), we measure fund connectedness *between financial institutions*, but not *with the market* as in van Oordt & Zhou (2019). However, similarly to van Oordt & Zhou (2019), we also measure connectedness *conditional on the market* since we use market variables in our tail dependence measures. Contrary to both approaches, our connectedness measure is explicitly time-varying. Therefore, we can use a large sample for inference, avoiding rolling-window estimations. Moreover, in Section 3.4.3, we detail how we adapt our connectedness measure to capture the extremal dependence *with the banking sector, given the state of the financial market*. The resulting measure intends to capture some of the systemic risks associated with hedge funds.

3.4 Extremal risk measures

In this section, we detail how our framework can be exploited to quantify risks stemming from interconnections between the different parts of the financial system, and connect our approach with the popular $CoVaR$ measure of Adrian & Brunnermeier (2016). The resulting tail risk measures are applied first to the study of interconnections in the funds' universe, and second to systemic interconnections, i.e., connectedness between funds and the banking sector.

3.4.1 Definitions

Total connectedness

The extremal connectedness measure provides us with a first tool to explore tail interconnections between different entities. Here, we consider a series of transformations of $\hat{\chi}^{kl}(\mathbf{x}_t)$ for all pairs (k, l) to gauge the risk level in the considered investment universe. In the idea of Billio et al. (2012) and Balla et al. (2014), we first compute a measure of the total interconnections, referred hereafter as *total connectedness*. To do so, we compute at each point in time the sum over all pairs (equally-weighted) of tail dependence measures:

$$\tilde{\chi}_t = \sum_{k=1}^{12} \sum_{l \neq k} \hat{\chi}^{kl}(\mathbf{x}_t)/66. \tag{13}$$

We normalize the sum by the number of pairs, a quantity that indicates the maximum connectedness in the considered system if all tail dependence measures were equal to one. A natural interpretation of this quantity is that it measures the average co-crash probability across two investment styles within the hedge fund universe. Moreover, this measure accounts for the impact of *realized* market returns and uncertainty levels on extremal interconnections. To investigate the individual contribution of each investment strategy to the overall connectedness, we decompose (13) in its style-specific components given by

$$\tilde{\chi}_t^k = \sum_{l \neq k} \hat{\chi}^{kl}(\mathbf{x}_t)/11. \tag{14}$$

This measure is a time-varying version of the systemic impact index (SII) proposed by Zhou (2010). The normalization ensures that $\tilde{\chi}_t^k \in (0, 1]$, such that (14) can be interpreted as the expected proportion of investment styles registering the failure of at least one fund, given that a fund in the reference investment style goes bankrupt at the same time. For example, a value $\tilde{\chi}_t^k = .6$ implies that, conditional on a fund in investment style k going bankrupt, one would expect 60% of the other styles to observe a failure as well.

To assess the spillovers of extreme losses in the hedge fund industry to banks, we define a similar quantity between the hedge fund sector and the banking sector. More precisely, relying on a sample of banks' market data, we can estimate the dynamic extremal connectedness

$\hat{\chi}^{banks,l}(\mathbf{x}_t)$ between the banking sector and funds with an investment strategy l as a function of the set of covariates $\{\texttt{VIX}, \texttt{MSCI}, \texttt{EPU}, \texttt{FSI}\}$. Then, we define the bank-specific connectedness

$$\tilde{\chi}_t^{banks} = \sum_{l=1}^{12} \hat{\chi}^{banks,l}(\mathbf{x}_t)/12, \tag{15}$$

as the average likelihood that a bank goes bankrupt, given that a hedge fund in any investment style goes bankrupt at time t. In Section 3.4.3, we conduct an analysis of the connectedness between banks and funds using a pool of 29 banks; see Section 3.1 of the Supplementary material for details on the dataset and the marginal dynamic modeling of their large losses.

Extremal CoVaR

So far, we have defined tail risk measures that rely solely on the extremal connectedness measures. However, since we have at our disposal the joint extremal distribution either between two investment styles or between an investment style and the banks, we can compute a conditional value-at-risk far in the tail, in the idea of the popular $CoVaR$ measure of Adrian & Brunnermeier (2016).

Using the generalization in terms of copulas proposed in Girardi & Ergün (2013) and adapting their results to our setting, we define the *extremal CoVaR* (eCoVaR) of an investment style k with respect to investment style l as the quantile at a high level α of a fund's loss in style k, conditional on observing a loss in style l larger than its associated quantile at a high level $\tilde{\alpha}$. Intuitively, this quantity measures the value-at-risk of a given fund in style k when any fund is style l suffers from an extreme loss under specific market conditions.

Let (Y_k, Y_l) be the pair of negative returns in styles k and l with marginal distributions $F_k(\cdot \mid \mathbf{x}_0)$ and $F_l(\cdot \mid \mathbf{x}_0)$, as defined in (10). The quantity $eCoVar_\alpha^{k|l,\tilde{\alpha}}$ is then defined by

$$\mathbb{P}\left\{Y_k > eCoVaR_\alpha^{k|l,\tilde{\alpha}} | Y_l > VaR_{\tilde{\alpha}}^l(\mathbf{x}_0)\right\} = 1 - \alpha, \tag{16}$$

with $VaR_{\tilde{\alpha}}^l(\mathbf{x}_0)$ as defined in (11). Thus, contrary to the classical definition proposed in Adrian & Brunnermeier (2016) and Girardi & Ergün (2013) which measures the effect of a distressed institution on the VaR (at some high level) of the market, the eCoVaR considered here is the value-at-risk *of a given fund in style* k conditional on *any fund with style* l exhibiting a loss larger than its fund-specific $\tilde{\alpha}$-th quantile.

Rewriting (16) in terms of joint probabilities as

$$\frac{\mathbb{P}\left\{F_k(Y_k \mid \mathbf{x}_0) > F_k(eCoVaR_\alpha^{k|l,\tilde{\alpha}} \mid \mathbf{x}_0), F_l(Y_l \mid \mathbf{x}_0) > \tilde{\alpha}\right\}}{1-\tilde{\alpha}} = 1-\alpha,$$

our quantity of interest $eCoVaR_\alpha^{k|l,\tilde{\alpha}}$ is thus defined as the u_α-th quantile of $F_k(\cdot \mid \mathbf{x}_0)$, i.e.,

$$eCoVaR_\alpha^{k|l,\tilde{\alpha}} = F_k^{-1}(u_\alpha \mid \mathbf{x}_0),$$

where the probability u_α satisfies

$$\mathbb{P}\left\{F_k(Y_k \mid \mathbf{x}_0) > u_\alpha, F_l(Y_l \mid \mathbf{x}_0) > \tilde{\alpha}\right\} - (1-\alpha)(1-\tilde{\alpha}) = 0. \tag{17}$$

Thus, $eCoVaR_\alpha^{k|l,\tilde{\alpha}}$ depends on macroeconomic, fund-specific, and financial factors through $\mathbf{x}_0$. Additionally, as solving (17) for u_α requires the copula describing the dependence between large losses in the pair of strategies (k,l), the probability u_α and hence $eCoVaR_\alpha^{k|l,\tilde{\alpha}}$ are intrinsically varying with $\mathbf{x}_t = (\texttt{VIX}_t, \texttt{MSCI}_t, \texttt{EPU}_t, \texttt{FSI}_t)$[6]. As a matter of fact, relying on our extremal modeling assumptions, the tail dependence in the pair of losses (Y_k, Y_l) is described by the dynamic Hüsler–Reiss spectral density $h\{\cdot;\theta^{kl}(\mathbf{x}_t)\}$, or equivalently by its associated extreme value copula $C^{EV}\{\cdot;\theta^{kl}(\mathbf{x}_t)\}$. Defining the survival extreme value copula as

$$\bar{C}^{EV}\{v_1, v_2;\theta^{kl}(\mathbf{x}_t)\} = v_1 + v_2 - 1 + C^{EV}\{1-v_1, 1-v_2;\theta^{kl}(\mathbf{x}_t)\},$$

the joint survival probability $\mathbb{P}(Y_k > y_k, Y_l > y_l)$ can be approximated for large y_k and y_l by $\bar{C}^{EV}\{\mathbb{P}(Y_l > y_l), \mathbb{P}(Y_k > y_k);\theta^{kl}(\mathbf{x}_t)\}$. Therefore,

$$\begin{aligned}
\mathbb{P}\left\{F_k(Y_k \mid \mathbf{x}_0) > u_\alpha, F_l(Y_l \mid \mathbf{x}_0) > \tilde{\alpha}\right\} &= \bar{C}^{EV}\left[\mathbb{P}\{Y_k > F_k^{-1}(u_\alpha \mid \mathbf{x}_0)\}, \mathbb{P}\{Y_l > F_l^{-1}(\tilde{\alpha} \mid \mathbf{x}_0)\};\theta^{kl}(\mathbf{x}_t)\right] \\
&= \bar{C}^{EV}\left\{1-u_\alpha, 1-\tilde{\alpha};\theta^{kl}(\mathbf{x}_t)\right\} \\
&= 1-\tilde{\alpha}-u_\alpha + C^{EV}\left\{u_\alpha, \tilde{\alpha};\theta^{kl}(\mathbf{x}_t)\right\},
\end{aligned}$$

and the probability u_α solves

$$1-\tilde{\alpha}-u_\alpha + C^{EV}\left\{u_\alpha, \tilde{\alpha};\theta^{kl}(\mathbf{x}_t)\right\} = (1-\alpha)(1-\tilde{\alpha}).$$

Note that $u_\alpha \in [\alpha, 1)$ and is exactly equal to α when tail independence arises. In case of

[6]The dependence of u_α on $\mathbf{x}_t$ and that of $eCoVaR_\alpha^{k|l,\tilde{\alpha}}$ on $\mathbf{x}_0$ and $\mathbf{x}_t$, is omitted for ease of exposition.

comonotonicity, $u_\alpha = 1 - (1-\alpha)(1-\tilde{\alpha})$ and is thus very close to 1 when the quantile levels α and $\tilde{\alpha}$ are high. This highlights the need of using the appropriate tail distributions provided by EVT to ensure the reliability of very high quantile estimation.

Relying on the resulting eCoVaR quantity, one can define a dynamic measure of the *risk contribution* (RC) of investment style l at the *fund*'s level by computing

$$\Delta eCoVaR_\alpha^{k|l,\tilde{\alpha}}(\mathbf{x}_0^{i,k,t}, \mathbf{x}_t) = eCoVaR_\alpha^{k|l,\tilde{\alpha}}(\mathbf{x}_0^{i,k,t}, \mathbf{x}_t) - VaR_\alpha^k(\mathbf{x}_0^{i,k,t}), \tag{18}$$

for any fund i with style k, $k \neq l$, and associated macroeconomic and fund-specific factors $\mathbf{x}_0^{i,k,t}$. The RC summarizes the "risk premium" at any time t, that a fund with investment style k has to bear on its $VaR_\alpha^k(\mathbf{x}_0^{i,k,t})$ if any fund with investment style l exhibits extreme losses.

Finally, one could be interested in computing a similar quantity measuring the effect of a distressed hedge fund in strategy l on the VaR (at some high level) of the banking sector. We denote this quantity by $eCoVaR_\alpha^{banks|l,\tilde{\alpha}}$ and define it as verifying

$$\mathbb{P}\left\{Y_{banks} > eCoVaR_\alpha^{bank|l,\tilde{\alpha}} | Y_l > VaR_{\tilde{\alpha}}^l(\mathbf{x}_0)\right\} = 1 - \alpha, \tag{19}$$

where Y_{banks} is the losses of the banks. Then, we can show that

$$eCoVaR_\alpha^{banks|l,\tilde{\alpha}} = F_{banks}^{-1}(u_\alpha \mid \mathbf{x}_0^{banks}),$$

where the probability u_α satisfies (17) (with $k = banks$) and $F_{banks}(\cdot \mid \mathbf{x}_0^{banks})$ is the marginal distribution of the banking sector. Here, similarly to the marginal modeling of the investment styles (see Section 3.2), the marginal distribution of the banks' losses is modeled semi-parametrically with the right tail being described by a dynamic GP distribution. The selected set of covariates $\mathbf{x}_0^{banks}$ is described in details in Section 3.2 of the Supplementary material. Therefore, $eCoVaR_\alpha^{banks|l,\tilde{\alpha}}$ is the quantile at level u_α under particular financial and economic factors summarized in $\mathbf{x}_0^{banks}$ and $\mathbf{x}_t$. Similarly to what we propose for funds, we obtain the *systemic risk contribution* (SRC) of hedge funds style l at the *bank* level by computing

$$\Delta eCoVaR_\alpha^{banks|l,\tilde{\alpha}}(\mathbf{x}_0^{i,banks,t}, \mathbf{x}_t) = eCoVaR_\alpha^{banks|l,\tilde{\alpha}}(\mathbf{x}_0^{i,banks,t}, \mathbf{x}_t) - VaR_\alpha^{banks}(\mathbf{x}_0^{i,banks,t})$$

with bank- and time-specific factors $\mathbf{x}_0^{i,banks,t}$.

3.4.2 Risk measures in the funds' universe

Figure 5 displays the total connectedness $\tilde{\chi}_t$ over time and within the universe of funds.

[Figure 5 about here.]

The minimum total connectedness is reached in August 1998 (0.207) whereas the maximum (0.583) takes place in November 2008. We distinguish several periods: from 1994 to 1998, the total connectedness remains stable around 0.3, before dropping sharply after Russia's default in 1998 and the collapse of LTCM fund. Then, from 2003 to 2007, we observe a phase of increasing connectedness until February 2007. During the GFC, connectedness drops in April 2008 before exhibiting a sharp increase until November 2008. The post-crisis period shows a progressive increase until 2016. Interestingly, the increase in total connectedness at the end of 2008 is consistent with an increase in the dependence in assets held by funds during the GFC (see, e.g., Nickerson & Griffin (2017) regarding default risks) and with the time-varying exposures to risk factors, as highlighted in Oh & Patton (2018). The dynamics of our measure seems also to reflect well the time periods prior to the crisis identified by Agarwal et al. (2017) where hedge funds modified drastically the composition of their portfolios. In particular, from February 2007 onwards, hedge funds with long positions in stocks were found to increase their investments in long put options to hedge against a market drop. This observation is in accordance with the decrease in extremal connectedness observed over that period, indicating that joint extreme losses across styles are less likely to occur on average due to reinforced hedging. Another explanation, suggested by Baele et al. (2019), is the reduction of hedge funds' exposure to flight-to-safety phenomena. As shown by the authors, hedge funds from all investment styles are negatively exposed to this factor and tend to reduce their exposure several months before flight-to-safety events.

To investigate each investment strategy individually, we compute the style-specific measures $\tilde{\chi}_t^k$ for $k = 1, \ldots, 12$. These quantities are displayed in Figure 6.

[Figure 6 about here.]

Interestingly, we see that *Short bias* remains loosely connected with the other styles during most of the considered time period. Several styles appear strongly connected (e.g., *Fixed income arbitrage* exhibiting $\tilde{\chi}_t^k$ values around 0.45) or with large variations (e.g., *Managed futures* with $\tilde{\chi}_t^k$ values ranging from 0.15 to 0.7). Most investment styles exhibit a decreasing trend in their style-specific connectedness in summer 2008. The drop is particularly pronounced for *Long/Short equity*, in accordance with an increase in put options hold by equity funds to hedge long stock positions before the GFC. However, similarly to observations made for the total connectedness, we observe a pronounced increase in style-specific connectedness for *Distressed/Restructuring*, *Managed futures*, and *Fixed income arbitrage* between May 2008 and November 2008. Their estimated levels of connectedness during that time period imply that, if a fund with one of these strategies fails, then around 70% of the other investment styles will register bankruptcies as well. Using the style-specific measure (14), we can also study the contribution of each investment style to the total connectedness. These measures reflect the proportion of the total connectedness attributed to a particular investment style. For example, we observe that at the end of 2008, the *Distressed/Restructuring* strategy had its contribution to the total connectedness shifting from 8% to 11%, reflecting a style-specific connectedness above 70% (see Figure 5 in the Supplementary material).

Now, going beyond connectedness measures, we look at the *risk contribution* (RC) of the *Distressed/Restructuring* style to the other investment styles, as defined in (18). An overview of this risk contribution is obtained by looking at the distributional behavior of the set

$$\left\{ \Delta eCoVaR_{\alpha}^{k|l,\tilde{\alpha}}(\mathbf{x}_0^{i,k,t}, \mathbf{x}_t) \right\}_{\substack{k \neq l \\ i=1,\dots,n_k(t)}} \tag{20}$$

for the investment style *Distressed/Restructuring* and at fixed time points t. We display in Figure 7 the empirical kernel density estimate of the set (20) with *Distressed/Restructuring* as a reference investment style and at three dates (October 2008, March 2009, and January 2013), corresponding to turmoil, post-crisis, and calm periods.

[Figure 7 about here.]

Distinct characteristics of the estimated distributions can be depicted by comparing the three dates. For instance, differences in the tail heaviness of each density indicate a higher systemic contribution of *Distressed/Restructuring* during the GFC, attributable to changes in extremal connectedness and marginal tail risks. While many funds usually exhibit a $\Delta eCoVaR_{\alpha}^{k|l,\tilde{\alpha}}$ around 12%, i.e., an absolute increase over VaR_{α}^{k} by 12%, $\Delta eCoVaR_{\alpha}^{k|l,\tilde{\alpha}}$ becomes concentrated around 25% in October 2008. In addition, the support of the empirical density at that date is much wider, suggesting that several funds are susceptible to suffer from tremendous losses or go bankrupt if a fund in *Distressed/Restructuring* registers a large loss. Finally, we define a dynamic fund-aggregated measure of $\Delta eCoVaR_{\alpha}^{k|l,\tilde{\alpha}}(\mathbf{x}_0, \mathbf{x}_t)$ at the style level, given by

$$\Delta eCoVaR_{\alpha}^{k|l,\tilde{\alpha}}(t) = \frac{\sum_{i=1}^{n_k(t)} \Delta eCoVaR_{\alpha}^{k|l,\tilde{\alpha}}(\mathbf{x}_0^{i,k,t}, \mathbf{x}_t) \times \mathtt{TotAss}^{i,k,t}}{\sum_{i=1}^{n_k(t)} \mathtt{TotAss}^{i,k,t}}. \tag{21}$$

This quantity describes the weighted risk contribution (as a proportion of the total assets) of investment style l for all funds alive at time t and following an investment style $k \neq l$. As an illustration, we display this quantity in Figure 8 with the investment style *Distressed/Restructuring* as a reference and $\alpha = \tilde{\alpha} = .975$.

[Figure 8 about here.]

Most styles exhibit a sharp rise during the GFC, e.g., above 30% for *Yield alternatives*, consistent with the findings of Figure 7. For some strategies though, these variations are rather limited and the range of the observed values is relatively narrow. For example, *Equity market neutral* registered a maximum $\Delta eCoVaR_{\alpha}^{k|l,\tilde{\alpha}}$ around 15% and most values around 6%.

3.4.3 Systemic risk measures

In this section, we study the likelihood that systemically important financial institutions may be in trouble at the same time as hedge funds. To do so, we use daily stock return data from a panel of 29 international banks over the period 1994–2017, and infer the extremal connectedness between these banks and the various hedge funds styles. The full list of banks and details of the first-stage estimation can be found in Section 3 of the Supplementary material. For all styles, we find a substantial average tail dependence with the pool of banks, reflecting that hedge funds and banks tend to suffer from extreme negative returns simultaneously; see Figures 9 and 10 of the Supplementary material. The style that is the most connected with banks on average, is *Long/Short equity* ($\hat{\chi} = .47$), whereas *Short bias* is the less connected one ($\hat{\chi} = .09$). These results are consistent with the fact that *Short bias* tend to perform well in bad market conditions, contrary to banks.

We display the bank-specific connectedness in Figure 9. We observe first a period of stability from 1994 to 2003, then a progressive increase followed by a stagnation at a rather high level up to July 2007. During that month, Bear Sterns had to famously liquidate two hedge funds over-invested in CDOs, revealing the exposure of banks to subprime loans through their participation in these funds. Interestingly, after that date, we observe a sharp drop in connectedness over the next year, until a surge in September 2008 when Lehman Brothers filed for bankruptcy. These observations are consistent with the ones made for the fund-specific connectedness, suggesting a reduction in risk exposure during the period leading to Lehman Brothers' bankruptcy. However, whereas fund-specific connectedness keeps increasing *after* September 2008 and up to November 2008, we witness an earlier decrease in bank-specific connectedness. It suggests that, during this period, the likelihood of simultaneous defaults of funds across investment styles increases, but spillovers to the banking sector become less likely. These results paint a finer picture than the one portrayed by Billio et al. (2012), who discussed results aggregated over three-year periods. They observe an overall increase of their connectedness measure between funds and banks from January 2006 to December 2008. We find a contrario that extremal connectedness increases *before* the onset of the crisis (from January 2003 to March 2005), remains at a relatively high level up to July 2007, peaks in September and October 2008, then mostly decreases *during* the crisis (up to March 2009). These observations are consistent with the deleveraging phenomenon and the liquidity disappearance described in Ben-David et al. (2012) and Nagel (2012): starting from July 2007, hedge funds faced tighter liquidity constraints due to massive redemption and withdrawals of capital providers. Hedge funds had therefore to withdraw extensively from the market to meet their obligations. Consequently, banks became less susceptible to suffer simultaneously from extreme losses since they were less exposed, via their liquidity-provider activities, to the bankruptcy of a fund. In September and October 2008, with the near-collapse of the financial system, this likelihood increased again (suggesting that risk exposure was still significant) before falling back to its 2003 level. Since 2010, we observe an upward trend in connectedness that is consistent with a rise in lending activities.

[Figure 9 about here.]

Although informative regarding the existing linkage with financial institutions, our connectedness measure neglects possible interaction effects with the marginal distribution. To overcome this limitation, we turn our attention to the systemic risk contribution (SRC) of the different hedge funds' style. The objective is to measure the additional risk to be supported - at the marginal level - by banks in case of extreme losses among hedge funds. Similarly to our fund-specific analysis, we consider the distribution of the following set at several dates:

$$\left\{\Delta eCoVaR_{\alpha}^{banks|l,\tilde{\alpha}}(\mathbf{x}_0^{i,banks,t},\mathbf{x}_t)\right\}_{\substack{i=1,\ldots,n_t^{banks}\\ \alpha=\tilde{\alpha}=0.975}}, \tag{22}$$

where n_t^{banks} is the number of observations in our sample of banks at time t. For illustrative purposes, we consider *Long/Short equity* and *Short bias* as reference investment styles. Recall that in case of tail independence, the corresponding eCoVaR would be equal to its VaR and $\Delta eCoVaR$ to 0. The empirical densities of the sets (22) are displayed in Figure 10 at four different dates: July 2007, September 2008, October 2008, and March 2009.

[Figure 10 about here.]

We observe the following: in July 2007, the median SRC is around 2.5% for *Short bias*, and 5% for *Long/Short equity*, reflecting the difference in connectedness levels (0.06 for *Short bias* and 0.52 for *Long/Short equity*). This SRC value reflects also the fact that the marginal distributions of the banks are not particularly heavy-tailed at that time (see Figure 8 in the Supplementary). Thus, the probability of suffering from extreme losses among banks remains limited when taking into account the tail connectedness between banks and funds. However, in comparison to a situation where funds and banks would be tail independent, we observe a substantial risk increase, with eCoVaR being roughly from 50% to 100% larger than the VaR. This interaction between marginal distribution and connectedness becomes much more pronounced in times of crisis: in September 2008, for *Short bias*, the connectedness stays rather stable. However, banks exhibit higher marginal risks. Thus, the risk increase (in absolute term) is larger compared to July 2007, with a median value of approximately 7%. For *Long/Short equity*, the connectedness increases during that period (reaching 0.75), leading to a SRC around 15%. These effects are even more pronounced in October 2008, with median SRC values reaching around 10% for *Short bias* and 22% for *Long/Short equity*, despite connectedness levels back at 0.05 and 0.48, respectively. Additional results for the full period considered are available in Section 3.3 of the Supplementary material.

4 Conclusion

We propose an improved econometric approach to measure the extremal connectedness of hedge funds, a component playing a crucial role in the vulnerability level of the financial system. The proposed approach combines multivariate extreme value theory with advanced regression techniques. It adequately deals with extreme events and accounts for both heterogeneity and non-stationarity due to changing market conditions. To tackle low reporting frequencies and short reporting periods at the fund level, we rely on panels of hedge funds' returns grouped by investment styles. Several extremal connectedness and risk indicators are derived from these models.We estimate these various quantities for a large sample of hedge funds' returns reported monthly in the HFR database. Our results suggest that some tail risk diversification is achievable by combining funds from investment styles that are weakly connected at extreme levels. In addition, for several pairs of styles, we identify important variations of extremal connectedness during financial crises. Looking at the link between hedge funds and a panel of banks, we find a significant extremal connectedness. As a consequence, the risk that the banking sector suffers from extreme losses when hedge funds do so is much higher than supposed by a marginal analysis.

Acknowledgements

Financial support of the Swiss National Science Foundation (LM) and the National Bank of Belgium (JH) is gratefully acknowledged. Parts of the research for this paper were conducted while LM was affiliated to HEC Montréal (Canada) and JH to the University of Göttingen (Germany), Chair of Statistics. The authors wish to thank the Editor (E. Ghysels) and two anonymous referees for helpful comments that improved the paper.

References

Adrian, T. & Brunnermeier, M. K. (2016), 'CoVaR', *American Economic Review* **106**(7), 1705–41.

Agarwal, V. & Naik, N. (2004), 'Risks and portfolio decisions involving hedge funds', *The Review of Financial Studies* **17**(1), 63–98.

Agarwal, V., Ruenzi, S. & Weigert, F. (2017), 'Tail risk in hedge funds: A unique view from portfolio holdings', *Journal of Financial Economics* **125**(3), 610–636.

Aiken, A. L., Clifford, C. P. & Ellis, J. (2013), 'Out of the dark: Hedge fund reporting biases and commercial databases', *The Review of Financial Studies* **26**(1), 208–243.

Ang, A., Gorovyy, S. & van Inwegen, G. (2011), 'Hedge fund leverage', *Journal of Financial Economics* **102**(1), 102–126.

Baele, L., Bekaert, G., Inghelbrecht, K. & Wei, M. (2019), 'Flights to Safety', *The Review of Financial Studies* **33**(2), 689–746.

Baker, S. R., Bloom, N. & Davis, S. J. (2016), 'Measuring economic policy uncertainty', *The Quarterly Journal of Economics* **131**(4), 1593–1636.

Bali, T. G., Brown, S. J. & Caglayan, M. O. (2014), 'Macroeconomic risk and hedge fund returns', *Journal of Financial Economics* **114**(1), 1–19.

Balla, E., Ergen, I. & Migueis, M. (2014), 'Tail dependence and indicators of systemic risk for large US depositories', *Journal of Financial Stability* **15**, 195–209.

Bee, M., Dupuis, D. J. & Trapin, L. (2019), 'Realized peaks over threshold: A time-varying extreme value approach with high-frequency-based measures', *Journal of Financial Econometrics* **17**, 254–283.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., De Waal, D. & Ferro, C. (2004), *Statistics of Extremes: Theory and Applications*, Wiley, New York.

Belloni, A. & Chernozhukov, V. (2013), 'Least squares after model selection in high-dimensional sparse models', *Bernouilli* **19**(2), 521–547.

Ben-David, I., Franzoni, F. & Moussawi, R. (2012), 'Hedge fund stock trading in the financial crisis of 2007–2009', *The Review of Financial Studies* **25**(1), 1–54.

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

Boyson, N., Stahel, C. & Stulz, R. (2010), 'Hedge fund contagion and liquidity shocks', *The Journal of Finance* **65**(5), 1789–1816.

Cao, C., Chen, Y., Liang, B. & Lo, A. W. (2013), 'Can hedge funds time market liquidity?', *Journal of Financial Economics* **109**(2), 493–516.

Carhart, M. M. (1997), 'On persistence in mutual fund performance', *The Journal of Finance* **52**(1), 57–82.

Castro-Camilo, D., de Carvalho, M. & Wadsworth, J. (2018), 'Time-varying extreme value dependence with application to leading European stock markets', *The Annals of Applied Statistics* **12**(1), 283–309.

Chavez-Demoulin, V., Embrechts, P. & Hofert, M. (2016), 'An extreme value approach for modeling operational risk losses depending on covariates', *Journal of Risk and Insurance* **83**(3), 735–776.

Davison, A. C. & Smith, R. L. (1990), 'Models for exceedances over high thresholds', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **52**(3), 393–442.

Diebold, F. X. & Yilmaz, K. (2009), 'Measuring financial asset return and volatility spillovers, with application to global equity markets', *The Economic Journal* **119**(534), 158–171.

Fama, E. & French, K. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**(1), 3–56.

Fan, J. & Li, R. (2001), 'Variable selection via nonconcave penalized likelihood and its oracle properties', *Journal of the American Statistical Association* **96**(456), 1348–1360.

Federal Reserve Bank of St. Louis (2014), 'St. Louis Fed Financial Stress Index [STLFSI]'.

Ferson, W. & Schadt, R. (1996), ‘Measuring fund strategy and performance in changing economic conditions’, *The Journal of Finance* **51**(2), 425–461.

Fisher, R. A. & Tippett, L. H. C. (1928), ‘Limiting forms of the frequency distribution of the largest or smallest member of a sample’, *Mathematical Proceedings of the Cambridge Philosophical Society* **24**(2), 180–190.

Fung, W. & Hsieh, D. (2004), ‘Hedge fund benchmarks: A risk-based approach’, *Financial Analysts Journal* **60**(5), 65–80.

Genest, C., Ghoudi, K. & Rivest, L.-P. (1995), ‘A semiparametric estimation procedure of dependence parameters in multivariate families of distributions’, *Biometrika* **82**(3), 543–552.

Girardi, G. & Ergün, A. (2013), ‘Systemic risk measurement: Multivariate GARCH estimation of CoVaR’, *Journal of Banking & Finance* **37**(8), 3169 – 3180.

Groll, A., Hambuckers, J., Kneib, T. & Umlauf, N. (2019), ‘LASSO-type penalization in the framework of generalized additive models for location, scale and shape’, *Computational Statistics & Data Analysis* **140**, 59–73.

Hale, G. & Lopez, J. A. (2019), ‘Monitoring banking system connectedness with big data’, *Journal of Econometrics* **212**(1), 203–220.

Hambuckers, J., Groll, A. & Kneib, T. (2018), ‘Understanding the economic determinants of the severity of operational losses: A regularized generalized Pareto regression approach’, *Journal of Applied Econometrics* **33**(6), 898–935.

Hartmann, P., Straetmans, S. & de Vries, C. G. (2004), ‘Asset market linkages in crisis periods’, *The Review of Economics and Statistics* **86**(1), 313–326.

Heffernan, J. E. (2000), ‘A directory of coefficients of tail dependence’, *Extremes* **3**(9), 279–290.

Huser, R. & Wadsworth, J. L. (2019), ‘Modeling spatial processes with unknown extremal dependence class’, *Journal of the American Statistical Association* **114**(525), 434–444.

Hüsler, J. & Reiss, R.-D. (1989), ‘Maxima of normal random vectors: Between independence and complete dependence’, *Statistics & Probability Letters* **7**(4), 283–286.

Joenväärä, J., Kauppila, M., Kosowski, R. & Tolonen, P. (2021), ‘Hedge fund performance: Are stylized facts sensitive to which database one uses?’, *Critical Finance Review* **10**(2).

Karagiannis, N. & Tolikas, K. (2019), ‘Tail risk and the cross-section of mutual fund expected returns’, *Journal of Financial and Quantitative Analysis* **54**(1), 425–447.

King, M. & Maier, P. (2009), ‘Hedge funds and financial stability: Regulating prime brokers will mitigate systemic risks’, *Journal of Financial Stability* **5**(3), 283–297.

Lee, J. D., Sun, Y. & Taylor, J. E. (2015), ‘On model selection consistency of regularized M-estimators’, *Electronic Journal of Statistics* **9**(1), 608 – 642.

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

Mhalla, L., de Carvalho, M. & Chavez-Demoulin, V. (2019), 'Regression-type models for extremal dependence', *Scandinavian Journal of Statistics* **46**(4), 1141–1167.

Mitchell, M. & Pulvino, T. (2012), 'Arbitrage crashes and the speed of capital', *Journal of Financial Economics* **104**(3), 469–490.

Nagel, S. (2012), 'Evaporating Liquidity', *The Review of Financial Studies* **25**(7), 2005–2039.

Negahban, S. N., Ravikumar, P., Wainwright, M. J. & Yu, B. (2012), 'A unified framework for high-dimensional analysis of M-estimators with decomposable regularizers', *Statistical Science* **27**(4), 538–557.

Nickerson, J. & Griffin, J. M. (2017), 'Debt correlations in the wake of the financial crisis: What are appropriate default correlations for structured products?', *Journal of Financial Economics* **125**(3), 454 – 474.

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

Patton, A. & Ramadorai, T. (2013), 'On the high-frequency dynamics of hedge fund risk exposures', *The Journal of Finance* **68**(2), 597–635.

Poon, S., Rockinger, M. & Tawn, J. A. (2004), 'Extreme value dependence in financial markets: Diagnostics, models, and financial implications', *The Review of Financial Studies* **17**(2), 581–610.

Resnick, S. I. (1987), *Extreme values, regular variation and point processes*, Springer, New York.

Sadka, R. (2010), 'Liquidity risk and the cross-section of hedge-fund returns', *Journal of Financial Economics* **98**(1), 54–71.

Titman, S. & Tiu, C. (2011), 'Do the best hedge funds hedge?', *The Review of Financial Studies* **24**(1), 123–168.

van de Geer, S. A. & Bühlmann, P. (2009), 'On the conditions used to prove oracle results for the Lasso', *Electronic Journal of Statistics* **3**, 1360 – 1392.

van Oordt, M. R. C. & Zhou, C. (2016), 'Systematic tail risk', *Journal of Financial and Quantitative Analysis* **51**(2), 685–705.

van Oordt, M. & Zhou, C. (2019), 'Systemic risk and bank business models', *Journal of Applied Econometrics* **34**(3), 365–384.

Wood, S. N. (2017), *Generalized Additive Models: An Introduction with R*, Chapman and Hall/CRC.

WSJ (2021), 'What is Archegos and how did it rattle the stock market?', *The Wall Street Journal* .

Zhou, C. (2010), 'Are banks too big to fail? Measuring systemic importance of financial institutions', *International Journal of Central Banking* **6**(34), 205–250.

Figures

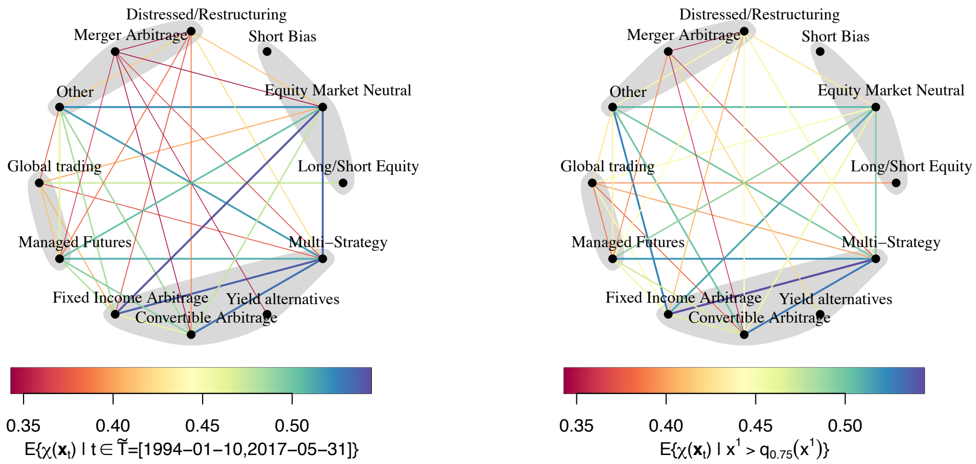

Figure 1: Left: Unconditional empirical mean of the pairwise conditional tail coefficient $\chi(\mathbf{x}_t)$. Right: Empirical mean of the pairwise conditional tail coefficient $\chi(\mathbf{x}_t)$ conditional on `VIX` exceeding its third quartile. The width of each edge is proportional to the estimated intensity of the tail dependence between two styles, used as nodes in the graph. Styles are grouped in the four broader HFR categories (see Table 1).

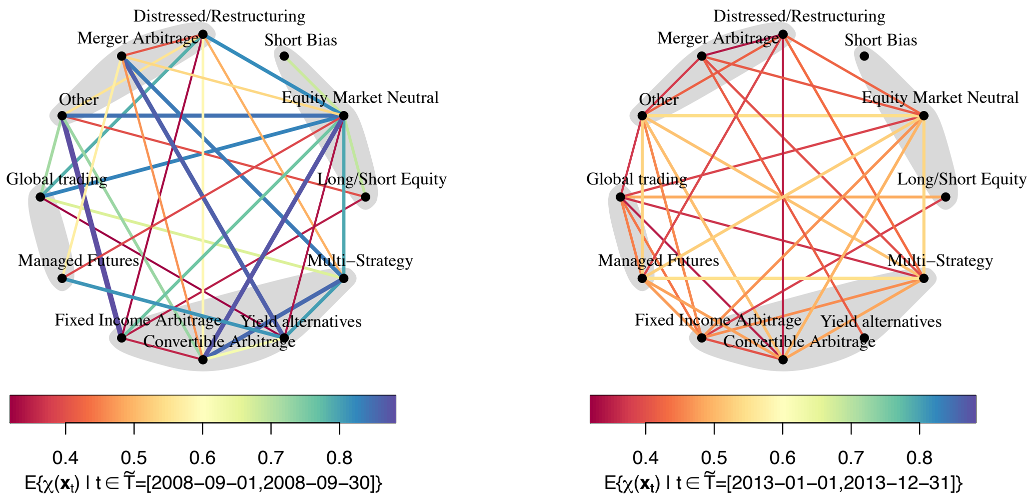

Figure 2: Estimated pairwise conditional tail coefficient $\chi(\mathbf{x}_t)$ for $t = 09/2008$ (left), and $t \in [01/2013, 12/2013]$ (right). The width of the edges is proportional to its estimated value.

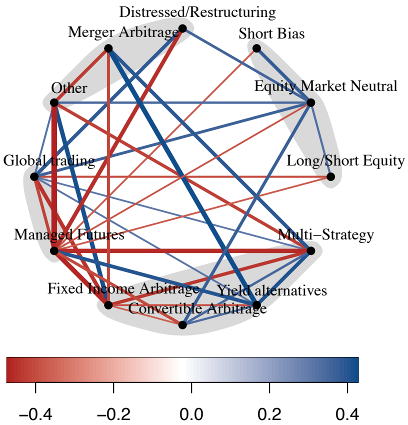

Figure 3: Differences (12) in estimated pairwise conditional tail coefficients for $\tilde{T} = 09/2008$ (left), and $\tilde{T} = [01/2013, 12/2013]$ (right). Edges with values between the first and third quartiles are removed for better visibility. The width of the remaining edges is proportional to the absolute value of the difference in tail coefficients.

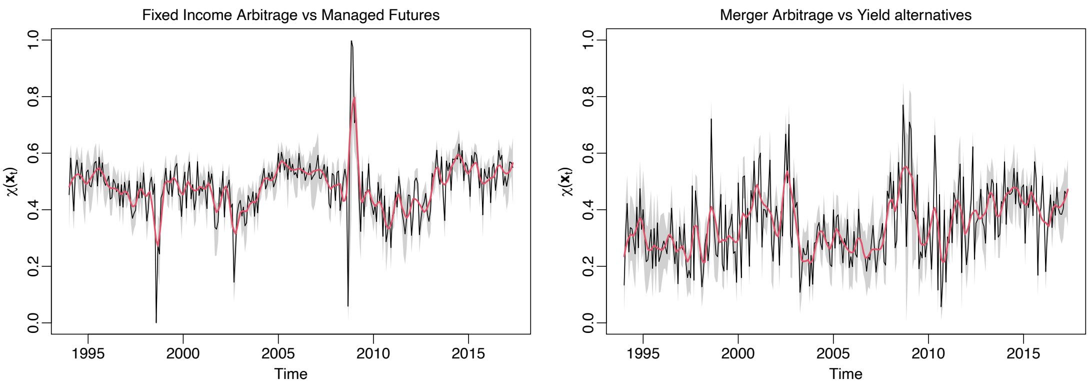

Figure 4: Estimates of the tail dependence over time, for two pairs of strategies. Gray: 95% confidence intervals. Red: smoothed lowpass moving average filter.

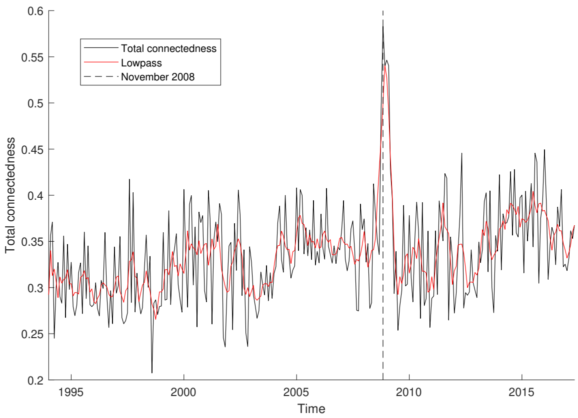

Figure 5: Total connectedness computed from (13) (black). Smoothed approximation using a lowpass moving average filter is represented in red.

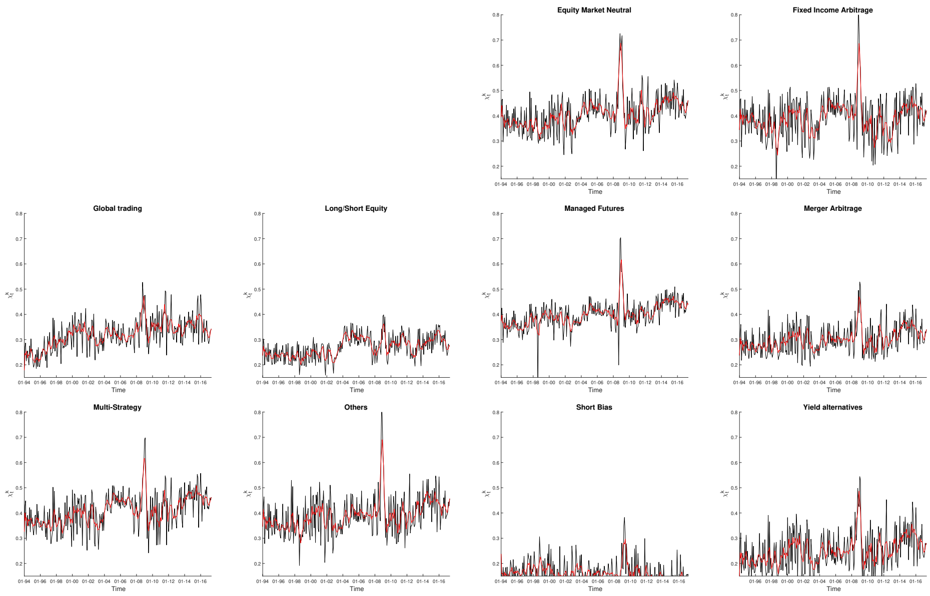

Figure 6: Measures $\tilde{\chi}_t^k$ of the style-specific connectedness for investment style $k = 1, \ldots, 12$ (black). Smoothed approximations using a lowpass moving average filter are shown in red.

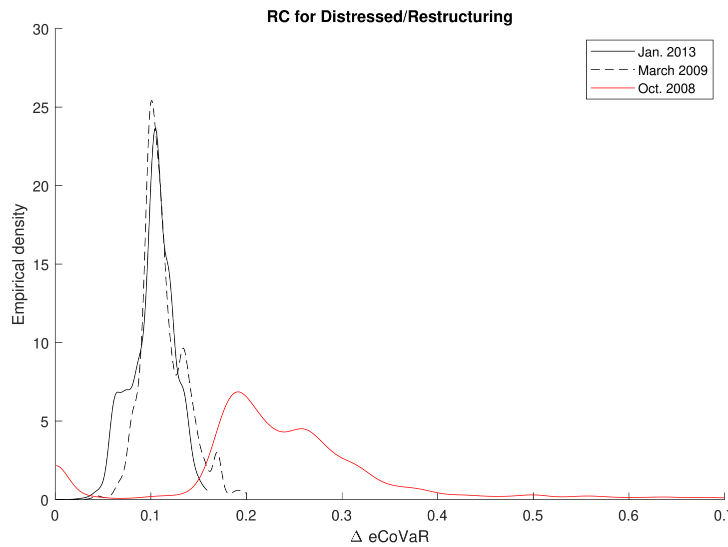

Figure 7: Empirical density of $\Delta eCoVaR_{\alpha}^{k|l,\tilde{\alpha}}$ given by (18), with *Distressed/Restructuring* as a reference investment style. We fix $\alpha = \tilde{\alpha} = .975$ and use three different dates: October 2008 (red), March 2009 (dashed black), and January 2013 (solid black).

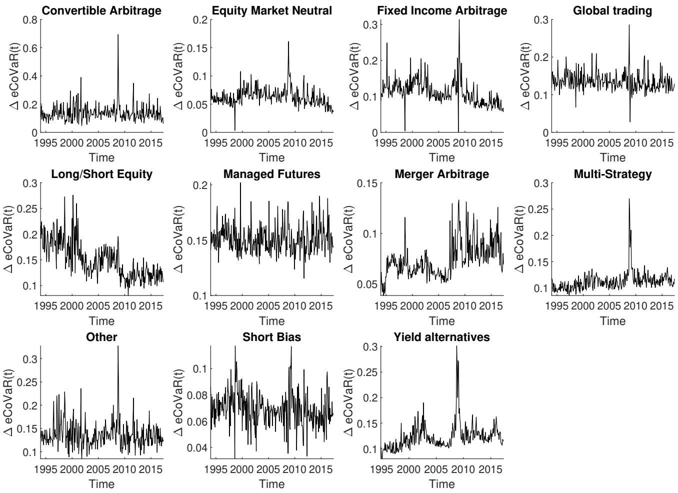

Figure 8: $\Delta eCoVaR_{\alpha}^{k|l,\tilde{\alpha}}(t)$ with *Distressed/Restructuring* as a reference and $\alpha = \tilde{\alpha} = .975$.

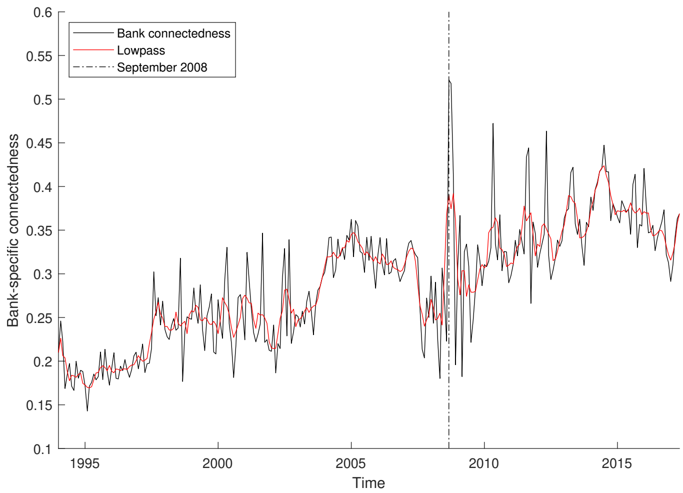

Figure 9: Black: $\tilde{\chi}_t^{banks}$ (bank-specific connectedness). Red: lowpass moving average filter.

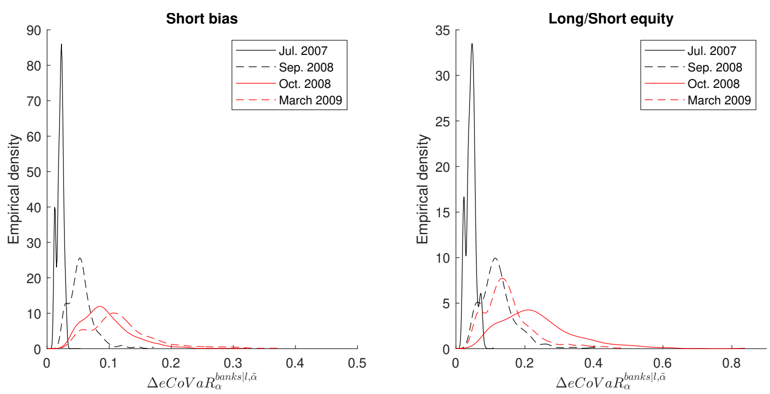

Figure 10: Empirical density of the SRC for *Short bias* and *Long/Short equity*.