**IMT Institute for Advanced Studies, Lucca**

Lucca, Italy

# Essays on Extreme Value Theory in Economics and Finance

PhD Program in Economics

XXVIII Cycle

**By**

**Luca Trapin**

**2016**

**The dissertation of Luca Trapin is approved.**

Program Coordinator: Prof. Andrea Vindigni, IMT Institute for Advanced Studies Lucca

Supervisor: Prof. Massimo Riccaboni, IMT Institute for Advanced Studies Lucca

Supervisor: Prof. Marco Bee, University of Trento

Tutor: Prof. Irene Crimaldi, IMT Institute for Advanced Studies Lucca

The dissertation of Luca Trapin has been reviewed by:

Prof. Valerie Chavez-Demoulin, HEC Lausanne

Prof. Fulvio Corsi, University of Venice

**IMT Institute for Advanced Studies, Lucca**

**2016**

*L'épaule à la roue.*

# Contents

# List of Figures

# List of Tables

# Acknowledgements

**Chapter 2** is the reproduction of the paper "U.S. stock returns: Are there seasons of excesses?", co-authored with Marco Bee and Debbie Dupuis, currently under revision for *Quantitative Finance*;

**Chapter 3** is the reproduction of the paper "Realizing the extremes: Estimation of tail-risk measures from a high-frequency perspective", co-authored with Marco Bee and Debbie Dupuis, to appear in *Journal of Empirical Finance*;

**Chapter 5** is the reproduction of the paper "The Realized Peaks Over Threshold: A high-frequency extreme value approach for financial time series", co-authored with Marco Bee and Debbie Dupuis, currently submitted to *Journal of Econometrics*.

I wish to thank my supervisor, Prof. Massimo Riccaboni, for his constant support and useful advices. He encouraged me a lot and contributed significantly to my own growth as a researcher. I am especially grateful to Prof. Marco Bee. In the last four years, he has been an endless and valuable source of inspiration and stimuli. I am deeply in debt to Prof. Debbie Dupuis who guided me into the world of extreme values and provided me much inputs on how to write a publishable paper. I also want to acknowledge the two referees, Prof. Valerie Chavez-Demoulin and Prof. Fulvio Corsi, for their helpful and insightful comments. Finally, I want to thank Prof. Roberto Renó for his suggestions on my Job Market Paper.

# Vita

| | |
|---|---|
| **December 16, 1988** | Born, Trento (TN), Italy |
| **2010** | Bachelor of Science (B.Sc.) in Economics<br>Final mark: 100/110<br>University of Trento, Trento, Italy |
| **2011** | Exchange Program<br>Maastricht University (SBE)<br>Maastricht, Netherlands |
| **2012** | Master of Science (M.Sc.) in Finance<br>Final mark: 110/110 cum laude<br>University of Trento, Trento, Italy |
| **2015** | Visiting Scholar<br>HEC Montréal<br>Montréal (QC), Canada |

# Publications

1. Alessandro Chessa, Irene Crimaldi, Massimo Riccaboni and Luca Trapin (2014). Cluster analysis of weighted bipartite networks: A new copula-based approach. *PlosOne*, 9(10): e109507.
2. Marco Bee, Debbie J. Dupuis and Luca Trapin (2016). Realizing the extremes: Estimation of tail-risk measures from a high-frequency perspective. *Journal of Empirical Finance*, Forthcoming.

# Presentations

1. Luca Trapin, "The realized peaks over threshold: A high-frequency extreme value approach for financial time series", at $9^{th}$ International Conference on Computational and Financial Econometrics, London, 2015.

# Abstract

Extreme value methods have been successfully applied in various disciplines with the purpose of estimating tail quantiles. The probabilistic results underlying the inference procedures for the extreme values rely on the assumption of independent and identically distributed (*iid*) random variables. However, empirical observations often present time variation and violate the *iid* assumption, thus the development of methods for modelling the extremes of dependent data is currently the subject of ongoing research. This thesis provides original contributions in this direction. Exploiting data on financial asset returns, we address questions regarding the tails of the conditional return distribution and propose models for them. We begin questioning whether extreme returns exhibit seasonal behaviour and develop an approach to uncover this fact. Next, we propose to employ a method based on high-frequency data to pre-whiten the returns, and then apply an extreme value model to the tails of the estimated residuals. We then study the extremal dependence inherent in financial returns, and evaluate the ability of various high-frequency based volatility processes in generating such dependence. Finally, we propose a new class of dynamic extreme value models that exploit high-frequency data to model the tails of the conditional return distribution.

# Chapter 1

# Introduction

Empirical and theoretical contributions to the analysis of heavy-tailed distributions in economics and finance date back at least to Mandelbrot (1963). He was the first to note that rare events in asset returns tend to occur more frequently than under the Gaussian paradigm. Since then, numerous studies have uncovered the presence of heavy-tailed distributions in several economically relevant variables, such as financial returns and foreign exchange rates (Fama, 1965; McFarland et al., 1982); the income and the wealth distributions (Gabaix, 2009); the city and the firm size distributions (Axtell, 2001; Gabaix, 1999); insurance claims from operational losses and natural disasters (Embrechts et al., 1997).

Since the probability of rare disasters sensibly affects the social welfare (Barro, 2006), identifying heavy-tailed phenomena and understanding their implications becomes relevant for policy and regulatory purposes as well as for the decision processes of firms and individuals. But, what does *heavy-tail* mean? As discussed in Embrechts et al. (1997), there is no a unique definition of heavy-tailed distribution, and throughout we will identify it with the notion of *regular variation*. A random variable $X$ is said to be regularly varying with index $\alpha \geq 0$ if its distribution $F(x)$ obeys to the following law,

$$\overline{F}(x) = x^{-\alpha} L(x),$$

where $\overline{F}(x) = 1 - F(x)$ and $L(x)$ is a slowly-varying function, meaning that $\lim_{x\to\infty} L(tx)/L(x) = 1$ with $t$ constant. This law states that as we move further into the tails of the distribution, the probability mass scales as a function of $\alpha$. In particular, the smaller the tail index $\alpha$, the more the probability mass in the tail, hence the probability of observing an extreme event. For example, take a worldwide investor that wants to make an investment in a certain country, and suppose that the success of such an investment is linked to the expected GDP growth of that country. Now, suppose that the distribution of the negative shocks to the GDP is regularly varying with $L(x)$ constant equal to one. It is clear that the lower is $\alpha$, the higher is the probability of experiencing a negative shocks. If, for instance, $\alpha = 3$ then the probability of observing negative shocks larger than $2\%$ is $1/2^3 = 0.125$. However, if $\alpha = 2$ then such probability becomes $1/2^2 = 0.25$. It doubles! Given this scenario, it is clear that the propensity of investing in that country strongly depends on the value of $\alpha$.

It should be now clear that the knowledge of the value of the tail index $\alpha$ is crucial to infer the behaviour of a heavy-tailed phenomenon, and the probabilistic results of the Extreme Value Theory offer strong foundations to develop statistical methods to address this issue. Fisher and Tippett (1928) show that, for a sequence of independent and identically distributed (*iid*) random variables, there exist only three possible limiting distributions to which normalized sample maxima can converge. Furthermore, these limiting distributions preserve the same tail behaviour of the distributions they attract. This result represents the backbone for the statistical approaches of the *Block Maxima* and *Peaks over Threshold*, that allow to estimate the tail index $\alpha$ (Embrechts et al., 1997). These approaches have been applied successfully in fields where extreme risk is of major concern: from environmental studies (Smith, 1989) to climatology (Dupuis, 2012), from civil engineering (Smith, 1986) to finance (McNeil and Frey, 2000).

The beauty of EVT comes forward when dealing with *iid* observations. However, variables of interest in economics and finance typically come in the form of time series presenting dependencies and structural

changes. To overcome these issues, a lot of efforts have been done to draw probabilistic results concerning the sample maxima of both stationary and non-stationary sequences (Hüsler, 1986; Leadbetter et al., 1983). Similarly, several statistical methods to deal with heteroskedasticity and trends in the extremes have been developed following the work of Davison and Smith (1990). This thesis contributes to this literature and addresses novel questions regarding the time variation in the extremes of asset returns and provide novel approaches to estimate the tails of the return distribution when the *iid* assumption does not hold.

Besides the introduction, this dissertation consists of four chapters. Each of them represents an individual article, consequently references and appendices are given at the end of each chapter, rather than at the end of the thesis. Furthermore, the necessary notation is defined in the introduction of each chapter, hence it is not necessarily consistent between them.

Chapter 2 investigates whether the tails of the conditional return distribution present seasonalities. The interest of financial economics in asset return seasonalities has long history, however this literature typically focuses on mean returns and volatility. We try to shift the attention to extreme returns and present a new algorithm to detect and estimate seasonal change-points in the tails of financial time series. Exploiting a dataset recording the daily returns of twelve U.S. sectors for almost ninety years, our empirical investigation reveal the emergence of seasonal behaviour in the upper tail.

Chapter 3 is inspired by the widely cited work of McNeil and Frey (2000), which proposes a two-step procedure to model the tails of asset returns with EVT while accounting for the intrinsic dependence. In particular, they first pre-whiten the returns with an autoregressive volatility model and then use an extreme value model for the tails of the estimated residuals. Recent advancements in the financial econometrics literature have lead to the development of models for the volatility exploiting high-frequency data. These models prove to outperform the traditional parametric models in terms of forecasts, and we propose to use them to pre-whiten the returns before applying EVT to the estimated

residuals.

Chapter 4 is partly a follow up of Chapter 3 and addresses different questions regarding the dependence between extreme observations. While the dependence in the second moment of the conditional return distribution has been widely studied, empirical investigations concerning the extremal dependence of financial returns did not receive the deserved attention. The first contribution of this chapter is to fill this gap. We then question whether processes for the daily returns exploiting high-frequency data to model the dynamic of the volatility are able to generate a degree of extremal dependence consistent with that observed empirically.

Chapter 5 is inspired by a recent literature that tries to apply EVT directly to the conditional return distribution instead of proceeding as in McNeil and Frey (2000). A shortcoming of the two-step procedure is that if the time series model used to pre-whiten the returns do not fully capture the dependence, then application of EVT to the estimated residuals could be undermined. Therefore, we develop a dynamic Peaks-over-Threshold method where measures built from high-frequency data are used to model the time-varying behaviour of daily extreme returns.

# Chapter 2

# U.S. stock returns: Are there seasons of excesses?

## 2.1 Introduction

The search for seasonal patterns in asset returns has a long history in the financial economics literature. During the seventies, economists got interested in the possible existence of seasonality in stock returns and its implications for capital market theory, capital market efficiency and the nature of the distribution of stock returns. The earliest evidence of seasonal behaviour is provided by Officer (1975), Rozeff and Kinney (1976), Roll (1983) and Keim (1983) which report evidence of a *January effect*, observing that average stock returns in the U.S. are higher in January than during the remaining eleven months of the year. Similar behaviour is also observed for the industrialized countries in Gultekin and Gultekin (1983) and the emerging markets in Claessens et al. (1995).

Although the January effect has been the most intriguing financial markets anomaly, other seasonal patterns have been reported. For instance, Cross (1973), French (1980), Jaffe and Westerfield (1985) and Lakonishok and Maberly (1990) provide evidence for the *weekend effect*, observing that Monday returns are on average lower than returns on other days. Ariel (1990) and Kim and Park (1994) give support to the *holiday*

*effect*, observing that returns are on average higher the day before a holiday. Another notable example is the *Halloween indicator*, according to which average stock returns should be lower in the May through October period than during the rest of the year (Bouman and Jacobsen, 2002). An extensive investigation of many of these seasonal effects in international markets is found in Fama (1991), Agrawal and Tandon (1994) and Hawawini and Keim (1995).

Although investigation of seasonality in *mean* returns still continues on both the empirical and theoretical front, see e.g. Heston and Sadka (2008), Heston and Sadka (2010), Kamstra et al. (2012), analysis of seasonality in stock markets has also moved forward to the second moment of the return distribution. Evidence of seasonal variation in the *variance* of stock returns at the intra-daily, daily and monthly level is reported by Wood et al. (1985), French and Roll (1986) and Gallant et al. (1992), respectively. Differences in the mean and the variance are investigated by including a dummy variable in the regression for the unconditional moment[1]. In Glosten et al. (1993) and Baillie and Bollerslev (2002), two different GARCH specifications with external regressors to capture seasonal dynamics in the conditional variance are inspected, while Bollerslev and Ghysels (1996) introduce a periodic autoregressive conditional model where parameters vary with time.

The literature sheds some light on the seasonality in equity returns at different frequency levels, but the analysis is restricted to the first two moments of the return distribution. In this paper, we investigate seasonality over the year in *extreme* returns. In the last fifteen years, there has been growing interest in the tail behaviour of returns to improve risk management practices and deepen the knowledge of financial markets (Embrechts et al., 1997, 1999; Jalal and Rockinger, 2008; Jondeau and Rockinger, 2003; McNeil, 1999; McNeil and Frey, 2000). Nonetheless, to the best of our knowledge, no attention has been paid to the possible existence of seasonality in extreme stock returns. The relevance of such an

[1] Alternatively to the dummy variable approach, differences in the mean and the variance are also tested non parametrically with the Kruskal-Wallis test. An example of a parametric test for both the unconditional mean and variance is reported in Section 2.2.

investigation is pointed out in Rozeff and Kinney (1976). These authors assume the unconditional return distribution to be fully represented by a symmetric-stable distribution and study whether the characteristic exponent of the latter varies across the months of the year.

We examine tail estimation for return series of financial data. We hypothesize that the return series follow a stationary process with time-varying conditional mean and volatility. The presence of time-varying conditional moments means that two return distributions are of interest: the conditional return distribution where conditioning is on the available information set, and the marginal distribution of returns, see McNeil et al. (2005) for a discussion. In this paper, we focus on the former for two specific reasons: financial interest is mainly directed to moments conditional on available information; probabilistic theory underlying the statistical model we outline in the next section to identify the seasons is valid for observations that satisfy the *iid* assumption or a particular mixing condition.

We take a two-stage approach. We pre-whiten the data by means of appropriate dynamics (estimated using quasi maximum likelihood methods) and model the innovation distribution using the residuals from the dynamic model as data, after controlling for seasonalities in the first two moments of the return distribution. The model for the tail distribution of the innovations is based on EVT.

We develop a new algorithm based on Kim and Lee (2009) to detect and estimate the *seasonal* change-points in the tail of financial time series. These seasonal change-points define *tail seasons*. A comprehensive simulation study establishes the good performance of our procedure and shows how *tail seasons* can be distinguished from *mean* and *volatility seasons*. An extensive empirical study examining the daily returns of 12 equally-weighted industry portfolios of stocks from the NYSE, AMEX and NASDAQ over the last 90 years reveals the emergence of coherent *tail seasons* over time for many industries.

Seasonality in the tails presents a potentially serious challenge for both financial economics and risk management practice. As to the former, tail events are the major determinant of asset risk premia (Boller-

slev and Todorov, 2011; Kelly and Jiang, 2014), thus a proper framework describing the dynamics of the tails is required. As to risk management, cycles in the tail risk have strong implications relative to capital requirements and strategic allocation, thus careful modelling of the tails is needed to avoid under- or overestimation of risks.

The remainder of the chapter is organized as follows. In Section 2.2, we present our approach to identify *tail seasons* in financial return series. In Section 2.3, we validate the approach through a comprehensive simulation experiment. In Section 2.4, we present the results of the large empirical study on returns of U.S. stocks. In Section 2.5, we try to link tail seasonalities to conditional skewness and kurtosis. In Section 2.6, we sketch some implications for financial practitioners. Finally, some discussion and concluding remarks appear in Section 2.7.

## 2.2 Methodology

We consider daily returns $r_t$ that follow the general discrete-time model

$$\begin{aligned} r_t &= f(r_{t-1}, \ldots, r_{t-u}, \epsilon_t, \ldots, \epsilon_{t-v}) \\ \epsilon_t &= \sigma_t \eta_t \\ \sigma_t^2 &= h(\sigma_{t-1}^2, \ldots, \sigma_{t-q}^2, \epsilon_{t-1}, \ldots, \epsilon_{t-p}) \end{aligned}$$

where $f(r_{t-1}, \ldots, r_{t-u}, \epsilon_t, \ldots, \epsilon_{t-v})$ and $h(\sigma_{t-1}^2, \ldots, \sigma_{t-q}^2, \epsilon_{t-1}, \ldots, \epsilon_{t-p})$ are functions describing respectively the dynamics of the conditional mean and the conditional variance, and $\eta_t$ is an *iid* random variable with zero mean and unit variance.

For the conditional variance, we use GARCH models (Bollerslev, 1986). As economic theory suggests that market information should have an asymmetric effect on volatility and empirical analyses show that positive and negative innovations to returns have a different impact on conditional volatility (Glosten et al., 1993), we use a GARCH model with leverage to model the time-varying volatility. We also include autoregressive terms in the mean dynamics. More precisely, we consider the

ARMA$(u, v)$-GJR-GARCH$(p, q)$ class of models given by

$$r_t = \mu + \sum_{k=1}^{u} \phi_k r_{t-k} + \sum_{z=1}^{v} \psi_z \epsilon_{t-z} + \epsilon_t \tag{2.1}$$

$$\epsilon_t = \sigma_t \eta_t \tag{2.2}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \left(\alpha_i + \gamma_i I(\epsilon_{t-i} < 0)\right) \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2.3}$$

where $I$ is the indicator function, $\alpha_i$, $\gamma_i$ and $\beta_i$ are constants satisfying the constraints $\alpha_i > 0$, $\beta_i > 0$ and $\sum_{i=1}^{p}(\alpha_i + \gamma_i/2) + \sum_{j=1}^{q} \beta_j < 1$, see Ling and McAleer (2002).

### 2.2.1 *Seasons* in the returns

A standard approach used in financial economics to investigate possible seasonality in the mean returns of the stocks is to test whether the returns distribution differs across the months of the year (Agrawal and Tandon, 1994; Claessens et al., 1995; Corhay et al., 1987; Gultekin and Gultekin, 1983; Rozeff and Kinney, 1976). For instance, defining a dummy variable $D_m$ which takes on value 1 only when the return $r_t$ is in month $m$, $m = 1, \ldots, 11$, we can extend (2.1) to

$$r_t = \mu + \sum_{m=1}^{11} \delta_m D_m + \sum_{k=1}^{u} \phi_k r_{t-k} + \sum_{z=1}^{v} \psi_z \epsilon_{t-z} + \epsilon_t \ .$$

One can then proceed as in Keim (1983) and Rozeff and Kinney (1976), testing for differences in the unconditional mean via the hypothesis test $\mathcal{H}_0 : \delta_1 = \cdots = \delta_{11} = 0$ vs $\mathcal{H}_1 :$ Not $\mathcal{H}_0$. Similarly, by generalizing (2.3) to

$$\sigma_t^2 = \omega + \sum_{m=1}^{11} \psi_m D_m + \sum_{i=1}^{p} \left(\alpha_i + \gamma_i I(\epsilon_{t-i} < 0)\right) \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

one can proceed as in Rozeff and Kinney (1976) and Gallant et al. (1992), testing for monthly differences in the variance via the hypothesis test $\mathcal{H}_0 : \psi_1 = \cdots = \psi_{11} = 0$ vs $\mathcal{H}_1 :$ Not $\mathcal{H}_0$.

In this paper, our interest lies in the tails of the conditional return distribution, and thus in the tails of $\eta_t \sim F$. Suppose that $F$ has upper end point $\upsilon_F := \sup\{\eta_t : F(\eta_t) < 1\}$. Given a high threshold $u$, $u < \upsilon_F$, Pickands (1975) shows that when $u \to \upsilon_F$, the distribution of the excesses $(\eta_t - u)_+$ converges to a *Generalized Pareto* (GP) distribution $G$ with shape parameter $\xi$ and scale parameter $\nu > 0$. That is, $\Pr(\eta_t - u \leq x | \eta_t > u)$ goes to

$$G(x;\xi,\nu) = \begin{cases} 1 - \{1 + \xi x/\nu\}^{-\frac{1}{\xi}} & \text{for} \quad \xi \neq 0 \\ 1 - \exp\{-x/\nu\} & \text{for} \quad \xi = 0 \end{cases} \tag{2.4}$$

as $u \to \upsilon_F$. When $\xi > 0$, $F$ has Pareto-type upper tail with tail index $1/\xi$. In this case, to test for tail index changes in the time series $\{\eta_t\}_{t=1}^T$, Kim and Lee (2009) consider the hypothesis

$$\mathcal{H}_0 : \xi_1 = \cdots = \xi_T = \xi \qquad \mathcal{H}_1 : \text{Not } \mathcal{H}_0 \ . \tag{2.5}$$

One can investigate both tails of $\eta_t$ over time by looking for changes in the tail index of the sequence $\eta_t$ and the negated sequence $-\eta_t$, respectively. However, this is insufficient for our purpose as we try to find seasonal change-points over the annual cycle by considering annual data over several years. In Section 2.2.2, we discuss how the Kim and Lee approach can be adapted to derive a test in this setup. Once the *seasons* are identified, we can study the tail behaviour using the peaks-over-thresholds method (Embrechts et al., 1997). Specifically, for each season, the excesses over a high threshold $u$ are identified, and the limiting result in Equation (2.4) is used to model the respective tails. We then obtain a tail estimator

$$\widehat{\bar{F}}(u+x) = \frac{N_u}{T}\left(1 + \widehat{\xi}\frac{x}{\widehat{\nu}}\right)^{-\frac{1}{\widehat{\xi}}}$$

where $\bar{F} = 1 - F$, $N_u$ is the number of observations exceeding the threshold $u$, and $\widehat{\xi}$ and $\widehat{\nu}$ are Maximum Likelihood (ML) estimates.

### 2.2.2 Finding *tail seasons* in $\eta_t$

Our goal is not to establish that the tail behavior of the $\eta_t$ is not the same in different months, this could easily be determined via a likelihood ratio

test. We rather wish to identify *seasons* for the tail behavior. More specifically, we adopt a change-point model and aim to identify change-points from multiple cross sectional time series of financial returns.

As the $\eta_t$ are unobservable, we have to find a suitable proxy. It is a standard approach to use the scaled residuals from the estimated conditional mean and conditional variance models as the proxy, but one must proceed cautiously: if the estimated parameters of the conditional models are not at least consistent, the resulting scaled residuals may be a very poor proxy for the $\eta_t$. Meitz and Saikkonen (2011) show that the global Gaussian quasi ML estimators (QMLE) for all the classes of models in Equations (2.1)-(2.3) is strongly consistent, even for heavy-tailed innovation distributions as long as they have finite second moment[2]. The stronger condition $\mathrm{E}\left[\eta_t^4\right] < \infty$ is required in many proofs of the consistency of QMLE in the GARCH extended family of models. If our model is correctly specified and QMLE is used to estimate the model parameters, residuals of the model fit, scaled by the level of conditional volatility, should represent approximately an *iid* sequence from the distribution function $F$. Scaled residuals $\widehat{\eta}_t$ are used both to check for the adequacy of the dynamic model and as input for the second stage of our method. To derive the scaled residuals $\widehat{\eta}_t$, we consider different ARMA($u$,$v$)-GJR-GARCH(1,1) specifications[3] and select the best filter according to the BIC criterion (Schwarz, 1978).

We relabel the $\widehat{\eta}_t$ to a corresponding $Z_{id}$ notation to facilitate the remaining discussion. Our data thus consist of observations $\{Z_{id}, i = 1, \ldots, n, d = 1, \ldots, D\}$, where $i$ denotes the year of the observation and $d$ denotes the day of the observation. The value of $n$ depends on the length of the sample (we use $n = 50$ both in the simulations and the empirical study), whereas $D$ is equal to 250 in the simulation study and $D = 365$ in the real-data analysis[4]. Let $\iota_0, \ldots, \iota_m$, $m \geq 2$ and $\iota_0 = \iota_m$, be the un-

[2]Other models can also capture the leverage effect (Nelson, 1991; Zakoian, 1994), however as the consistency of the Gaussian QMLE has only been established for the class in (3.2)-(3.3), it is our best option.

[3]We also add external regressors for the mean to explain observed behavior.

[4]Since the markets are not open on weekends and public holidays, there are not always $n$ observations for day $d$. Therefore, when it is convenient, we use the notation $n_d \in$

known change-points resulting in $m$ segments defining the *seasons*[5]. For every $i$, assume that $Z_{id} \sim f_j$ for $d \in [\iota_{j-1}, \iota_j)$, $j = 1, \ldots, m$. The $f_j$ are distributions such that $f_j$ differs from both $f_{j-1}$ and $f_{j+1}$ in the right tail.

We now test the hypothesis (2.5), but using the sequence $Z_{id}$ to draw insights regarding the *seasons*. Let $\tau$ be a high quantile level close to one. Define the test statistic

$$H(\phi) = \frac{1}{\sqrt{T\tau}} \max_{\iota^- \leq \iota \leq \iota^+} \left| M(\iota, \widehat{q}_\tau) - \frac{\iota}{T} M(D, \widehat{q}_\tau) \right|, \tag{2.6}$$

where $M(\iota, \widehat{q}_\tau) = \sum_{d=1}^{\iota} \sum_{i=1}^{n_d} \phi \{\log (Z_{id}/\widehat{q}_\tau)\}$, $T = \sum_{d=1}^{D} n_d$, $\widehat{q}_\tau$ is the $\tau$th sample quantile of $\{Z_{id}\}$, $\iota \in \{1, \ldots, 365\}$ are the candidate change-points, and $\iota^-$ and $\iota^+$ are the first candidate change-points such that $\iota^- \sum_{d=1}^{\iota^-} n_d \geq \lambda T$ and $(T - \iota^+ \sum_{d=1}^{\iota^+} n_d) \geq (1 - \lambda T)$ where $\lambda$ is set equal to a small positive value[6] Defining $\phi_1(x) = I(x > 0)$ and $\phi_2(x) = x_+$ we can exploit respectively the frequency of the excesses and their magnitude to find the change-points. Kim and Lee (2009) show that under $\mathcal{H}_0$, $H(\phi_1) \xrightarrow{d} \mathcal{B}$ and $H(\phi_2)/\sqrt{2}\xi \xrightarrow{d} \mathcal{B}$ where $\mathcal{B}$ is the distribution function of $\sup_{\lambda \leq s \leq 1-\lambda} |B(s)|$, with $B$ a Brownian bridge. We obtain an estimate of $1/\xi$ with the Hill (1975) estimator, $1/\widehat{\xi} = \left( K^{-1} \sum_{d=1}^{D} \sum_{i=1}^{n_d} \log (Z_{id}/\widehat{q}_\tau) \right)^{-1}$ where $K$ is the number of observations $Z_{id}$ exceeding the threshold $\widehat{q}_\tau$. The above adaptation of Kim and Lee (2009) was shown in Dupuis et al. (2015) to offer the best performances under different assumptions on the distribution $F$ of the $\{Z_{id}\}$ and to be computationally very efficient. Furthermore, it is completely non-parametric and offers the flexibility of using the frequency of the exceedances or their magnitude.

Dupuis et al. (2015) assume $\iota_0 = 1$ and this implies a necessary change-point on January 1. As our goal is to identify the possible seasons in the year for the tail (and one may run through the Fall and Winter period for

$\{n_1, \ldots, n_D\}$ to denote the number of observations corresponding to each day $d$.

[5]In Section 2.3 we discuss the possibility of finding only one change-point and provide an explanation for this occurrence.

[6]This constraint is needed to avoid intervals that are too small as a minimum number of observations are required to estimate the tail index. In a given iteration, some of the 365 days of the year are thus excluded from being candidate change-points. However, as our algorithm involves multiple iterations, these days become candidate change-points in other iterations.

example), this artificial constraint is problematic. We implement a new sequential strategy that returns both the optimal number $m$ of change-points and their locations $\widehat{\iota}_0, \ldots, \widehat{\iota}_m$. The algorithm has two phases: the **Initialization** phase where one seeks to establish the existence of a first change-point, and the **Multiple change-points** phase where one looks for more change-points if an initial change-point is found in the first phase. In the first step, we carry out the test in equation (2.6) over the sequence $Z_{id}$ from the beginning to the end of the year. The idea is to find a first change-point to replace January 1 as the starting point in Dupuis et al. (2015). If $H(\phi_1) < \mathcal{B}_\alpha$ or $H(\phi_2)/\sqrt{2}\xi < \mathcal{B}_\alpha$, where $\mathcal{B}_\alpha$ is the (1-$\alpha$)-quantile of $\mathcal{B}$, then there is no change-point and the procedure terminates. Otherwise, $\mathcal{H}_0$ is rejected and the first change-point is set to be $\widehat{\iota}_1 = \arg\max_\iota \left|M(\iota, \widehat{q}_\tau) - \frac{\iota}{T} M(D, \widehat{q}_\tau)\right|$. Then, the $\{Z_{id}\}$ are relabeled so that they begin and end at the identified $\widehat{\iota}_0$ and $\widehat{\iota}_0 - 1$, respectively. In the second phase, the test in equation (2.6) is performed over each interval, and the $m$th change-point is identified and set equal to $\widehat{\iota}_m = \arg\max_{\iota^- \leq \iota \leq \iota^+} \left|M(\iota, \widehat{q}_\tau) - \frac{\iota}{T} M(D, \widehat{q}_\tau)\right|$, if $\mathcal{H}_0$ is rejected at the level $\alpha/(m+1)$. This iterative procedure continues until the test fails to reject the null in each segment, or when the shortest segment defined by adding another change-point is shorter than $\lambda T$.

In the algorithm, the function $\phi(\cdot)$, the threshold $\tau$ and the value of the parameter $\lambda$ must be chosen. We conduct a simulation experiment to assess the sensitivity of the performance to the choice of $\phi(\cdot)$ and $\tau$. The value of $\lambda$ is set pragmatically as we have to prevent intervals shorter than say one month, otherwise a given interval may not have a sufficient number of observations to carry out the subsequent extreme value analysis. Details of the implemented algorithm are given in Appendix A.1.

Finally, once the *tail seasons* are identified, the peaks-over-threshold method is used to estimate the tail model for each. The extent to which large volatility adjusted innovations are likely throughout the year, and seasonal differences occur, can then be assessed.

### 2.2.3 Extremes of GARCH processes

Conditional and unconditional distributions are both important concepts in finance, and their differences are related to the information sets with respect to which they are measured. For a fixed point in time $t$, the conditional distribution returns information regarding the probability of an event given the time series path up to that time. On the other hand, the unconditional distribution gives the probability of an event regardless of the accumulated information. The general class of GARCH processes are known to be unconditionally heavy-tailed (Basrak et al., 2002; Davis and Mikosch, 2009a; Mikosch and Starica, 2000). From Corollary 1 in Davis and Mikosch (2009a), under appropriate conditions on $\eta_t$, the tails of the marginal distribution of a stationary GARCH process exhibit power law behaviour. In particular, for $\epsilon_t$ in equation (2.2) there exist $\kappa > 0$ and positive constants $c_{|\epsilon|}$ and $c_\sigma$ such that

$$P(|\epsilon| > x) \sim c_{|\epsilon|} x^{-2\kappa} \qquad \text{and} \qquad P(\sigma > x) \sim c_\sigma x^{-2\kappa} \,. \tag{2.7}$$

The value of $\kappa$ can be determined following Kesten (1973). In the case of a GARCH(1,1) process, direct calculation shows that $\kappa$ is the unique solution of the equation

$$E[(\alpha_1 \eta^2 + \beta_1)^\kappa] = 1, \tag{2.8}$$

and can be found numerically for known values of the GARCH parameters $\alpha_1$ and $\beta_1$, and assuming a density for $\eta$. Thus, a close relationship can be established between the parameters of the GARCH process and the tail index of the marginal distribution of $|\epsilon|$. A change in the GARCH parameters induces a change in both tails of the marginal distribution of $r_t$, whereas a change may exist in only one tail of the distribution of the innovations $\eta_t$. We can thus detect *tail seasons* in only one tail (positive or negative) of innovations and not necessarily in both. Our simulation study seeks to confirm these facts and the good performance of our approach.

Financial extremes tend to occur in clusters. The degree of dependence in the extremes of a stationary series is determined by the extremal

index $\theta$ (Leadbetter, 1983). In particular, when $\theta < 1$ the extremes are dependent, conversely for $\theta = 1$ they are independent. More precisely, following Davis and Mikosch (2009a), for $\epsilon_t$ in equations (2.2)-(2.3), it holds that $(c_{|\epsilon|}T)^{-\frac{1}{2\kappa}} M_T(|\epsilon|) \xrightarrow{d} \theta_{|\epsilon|}^{\frac{1}{2\kappa}} \Phi_{2\kappa}$ where $M_T(|\epsilon|) = \max\{|\epsilon_1|, \ldots, |\epsilon_T|\}$, $\Phi_{2\kappa}$ is a Fréchet distribution with index $2\kappa$, and $\theta_{|\epsilon|} < 1$ (Embrechts et al., 1997). This property of the GARCH process is extremely relevant and justifies the use of a GARCH model to pre-whiten the data. In fact, a correct pre-processing model is essential to obtain estimates of $\eta_t$ that are iid as required by the peaks-over-threshold approach. When it is not possible to get rid of the extremal dependence, the probability of exceedances are distorted. The application of a de-clustering procedure returns the probability of cluster exceedances and this is unintuitive in a financial context. To verify the degree of dependency left in the estimated residuals $\widehat{\eta}_t$ we use the interval estimator of Ferro and Segers (2003). This estimator does not need to be tuned and requires only a choice of threshold.

## 2.3 Simulation experiments

In this section, we set up a large Monte Carlo experiment to assess both the precision of our algorithm in identifying the change-points, and our ability to recover the innovation tail parameter of each season given the returns. Similar analysis for independent observations has been carried out in Dupuis et al. (2015), but the use of the estimated residuals $\widehat{\eta}_t$ of the pre-processing model as a proxy for the innovations $\eta_t$ adds a layer of complexity. It is worth investigating the performance for GARCH-like returns as the test statistic in Equation (2.6) may lose power as a consequence of the filtering procedure. We establish different scenarios to control for two aspects that are relevant in the analysis. Differences in the extremes of $\widehat{\eta}_t$ may arise as a consequence of changes in the mean and the variance process, when these are not appropriately accounted for in the pre-processing model. We must be able to disentangle the changes in the innovation tails from those in the first two moments of the return distribution, and we assess the behaviour of our algorithm under different assumptions regarding the nature of the change-points. Contrarily to

changes in the parameters of the mean and the variance processes, innovations may present variation in only one tail that causes changes in only one tail of the conditional distribution of returns. It is important to verify that both one-tailed variation and two-tailed variation can be recovered from the estimated residuals $\widehat{\eta}_t$.

A second purpose of these experiments is to study the performance of the algorithm for different choices of the threshold $\tau$ and the function $\phi(\cdot)$. These two parameters must be set in advance and may affect the quality of the information provided by the algorithm. In what follows, we give a detailed description of the procedures and discuss the results.

### 2.3.1 Double change-points

The most intuitive situation is the double change-points case. We consider $50$ years of $250$ days and two seasons per year defined by two change-points at $\iota_0 = 125$ and $\iota_1 = 250$. Each season presents a specific tail behaviour determined by an ARMA(1,1)-GJR-GARCH(1,1) process. Parameters in the baseline specification are varied depending on the nature of the change that we want to induce. We list the various scenarios, explaining how they relate to the different features of the algorithm that we want to test. Parameter settings are given in Table 1.

*Case A.1* Two-tailed change-points arising as a consequence of changes in both tails of the innovation process $\eta_t$.

*Case A.2* Change-points only in the upper tail of $\eta_t$ resulting in one-tailed seasons in $r_t$. To generate this pattern we rely on the Asymmetric Exponential Power (AEP) distribution introduced by Bottazzi and Secchi (2011), which allows one to generate observations following a standard Normal distribution on one side and a heavy-tailed distribution on the other side. Along with *A.1*, this scenario allows us to understand whether the algorithm recovers both one-tailed and two-tailed changes.

*Case B.* In Equation (2.7) we have shown that GARCH processes induce regular variation, thus generating heavy-tailed observations even

when the innovation process is light-tailed. A change in the GARCH parameter $\beta$ causes a variation in the parameter of both tails, hence a different extremal behaviour.

*Case C.* Similar to *Case B*, an increase in the unconditional variance $\omega$ generates more extreme observations in both tails.

*Case D.* A change in the unconditional mean $\mu$ of the process causes respectively an increase of the number of extreme observations in one tail and a reduction in the other for the marginal process.

For each scenario, we carry out 100 simulations using the two functions $\phi_1$ and $\phi_2$, and three thresholds $\widehat{q}_\tau$ at $\tau = 0.95, 0.975, 0.99$. To assess the performance of the algorithm, we estimate the scaled residuals $\widehat{\eta}_t$, apply the change-point algorithm to $Z_{id}$, and provide summary statistics measuring our ability to retrieve the true change-points $\iota_0$ and $\iota_1$. In Table 2, we report the following statistics:

**TP1:** Number of cases where one true change-point is detected. We record a true positive only if the detected change-point is within five days of a true value;

**TP2:** Number of cases where two true change-points are detected;

**FP1:** Number of cases where at least one false positive is detected. We record a false positive if any of the identified change-points lie outside the five-day radius around the true change-points;

**FP2:** Number of times that more than two change-points are detected.

Although the choice of the function $\phi(\cdot)$ does not affect the ability to recover inadequate pre-processing of mean and variance changes, the first two columns of Table 2 show that $\phi_2$ has a higher precision in finding the change-points generated by the innovation process. Analogously, the threshold choice is relevant to detect this type of change, and it seems that going further in the tails allows us to more easily pinpoint the exact position of the change-points. Sometimes the 95th quantile outperforms the 97.5th because of the proximity of the latter to the crossing point of

the densities of the two generating processes. In fact, a heavy-tailed distribution has a thinner body than a light-tailed distribution. This implies that when the 95th quantiles of the two distributions are not far enough in the tail, the difference in the central part is caught. This pattern is confirmed by the fact that the 99th quantile outperforms the others in all scenarios.

As to the different phenomena that cause differences in the tails along the year, we see that the algorithm is able to capture all of them, thus we must be able to find a way to disentangle the true generating mechanism of the change-points. In this regard, the fact that it perfectly discerns one-tailed changes from two-tailed changes is critical in the empirical analysis, as changes in the tails of the innovations process are the only ones that can lead to one-tailed change-points.

**Table 1: Double change-points**. ARMA(1,1)-GJR-GARCH(1,1) parameters used to generate the different scenarios. Unless otherwise stated, parameters values are $(\mu, \phi, \psi, \omega, \alpha, \gamma, \beta) = (10^{-4}, 0.1, 0.2, 10^{-6}, 0.1, 0.05, 0.8)$ and $\eta \sim N(0,1)$.

| Scenario | First Season | Second Season |
|---|---|---|
| Case A.1 | $\eta \sim N(0,1)$ | $\eta \sim t(4)$ |
| Case A.2 | $\eta \sim N(0,1)$ | $\eta \sim AEP(1,2,0.8,1)$ |
| Case B | $\beta = 0.8$ | $\beta = 0.85$ |
| Case C | $\omega = 10^{-6}$ | $\omega = 10^{-4}$ |
| Case D | $\mu = 10^{-4}$ | $\mu = 10^{-2}$ |

**Table 2: Double change-points.** Performance measurements for 100 replications of the scenarios described in Section 2.3, with thresholds corresponding to the quantiles $\widehat{q}_\tau$ at $\tau = 0.95, 0.975, 0.99$, and functions $\phi_2$ and $\phi_1$. The values in the table are the counts of each statistic out of 100 simulations.

| | | *Function* $\phi_2$ | | | | | | | | | | *Function* $\phi_1$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Upper tail | | | | | Lower Tail | | | | | Upper tail | | | | | Lower Tail | | | | |
| | | *A.1* | *A.2* | *B* | *C* | *D* | *A.1* | *A.2* | *B* | *C* | *D* | *A.1* | *A.2* | *B* | *C* | *D* | *A.1* | *A.2* | *B* | *C* | *D* |
| 0.99 | TP1 | 88 | 89 | 90 | 100 | 100 | 82 | 6 | 89 | 100 | 87 | 67 | 85 | 92 | 100 | 98 | 66 | 3 | 87 | 100 | 85 |
| | TP2 | 45 | 48 | 43 | 90 | 27 | 36 | 1 | 39 | 86 | 20 | 20 | 46 | 45 | 82 | 34 | 17 | 0 | 49 | 81 | 20 |
| | FP1 | 57 | 52 | 53 | 45 | 26 | 65 | 8 | 61 | 42 | 42 | 59 | 55 | 57 | 32 | 23 | 70 | 7 | 54 | 32 | 42 |
| | FP2 | 3 | 1 | 3 | 38 | 1 | 1 | 0 | 4 | 31 | 0 | 1 | 1 | 5 | 17 | 1 | 3 | 0 | 5 | 15 | 5 |
| 0.975 | TP1 | 71 | 67 | 93 | 100 | 100 | 67 | 1 | 94 | 100 | 94 | 1 | 38 | 91 | 100 | 98 | 2 | 6 | 95 | 100 | 85 |
| | TP2 | 22 | 26 | 65 | 76 | 49 | 18 | 0 | 55 | 67 | 52 | 0 | 11 | 57 | 97 | 63 | 0 | 0 | 57 | 98 | 35 |
| | FP1 | 72 | 68 | 37 | 39 | 48 | 73 | 2 | 49 | 45 | 52 | 4 | 47 | 45 | 47 | 48 | 4 | 8 | 49 | 48 | 65 |
| | FP2 | 1 | 1 | 3 | 15 | 16 | 1 | 0 | 5 | 12 | 15 | 0 | 1 | 5 | 44 | 19 | 0 | 1 | 9 | 46 | 13 |
| 0.95 | TP1 | 12 | 70 | 96 | 100 | 100 | 11 | 2 | 98 | 100 | 98 | 55 | 6 | 96 | 100 | 98 | 55 | 6 | 96 | 100 | 97 |
| | TP2 | 2 | 21 | 69 | 91 | 54 | 4 | 0 | 60 | 95 | 49 | 17 | 1 | 58 | 100 | 69 | 19 | 2 | 59 | 98 | 43 |
| | FP1 | 16 | 62 | 34 | 40 | 77 | 12 | 4 | 45 | 29 | 81 | 67 | 12 | 42 | 67 | 70 | 53 | 13 | 46 | 54 | 86 |
| | FP2 | 0 | 0 | 3 | 31 | 43 | 0 | 0 | 7 | 24 | 37 | 1 | 0 | 1 | 67 | 44 | 0 | 0 | 6 | 52 | 38 |

### 2.3.2 Triple change-points

For this simulation study, we generate $50$ years $\times$ $250$ days of observations with change-points at $\iota_0 = 83$, $\iota_1 = 166$ and $\iota_2 = 250$. We simply adapt the scenarios previously described to the three change-point case, and generate the intervals according to the variations shown in Table 3. In Table 4, we report TP1, TP2, FP1 and

**TP3:** Number of cases where three true change-points are detected;

**FP2:** Number of times that more than three change-points are detected.

The results obtained from the triple change-point experiments are qualitatively the same as those from the double change-points, though the performance deteriorates as a consequence of the lower power of the test statistic that is computed using fewer observations in each interval. Also in this case, the choices of $\phi(\cdot)$ and of the threshold $\tau$ are not relevant for the identification of change-points that originated either from misspecification of the mean or the variance processes. Using $\phi_2$ with a higher threshold level considerably improves the performance of the algorithm in the case of changes in the innovation tails.

**Table 3: Triple change-points**. ARMA(1,1)-GJR-GARCH(1,1) parameter values used to generate the different scenarios. Unless otherwise stated, parameter values are $(\mu, \phi, \psi, \omega, \alpha, \gamma, \beta) = (10^{-4}, 0.1, 0.2, 10^{-6}, 0.1, 0.05, 0.8)$.

| Scenario | First Season | Second Season | Third Season |
|---|---|---|---|
| Case A.1 | $\eta \sim N(0,1)$ | $\eta \sim AEP(1, 1.7, 1, 1.7)$ | $\eta \sim t(4)$ |
| Case A.2 | $\eta \sim N(0,1)$ | $\eta \sim AEP(1, 2, 0.95, 1.5)$ | $\eta \sim AEP(1, 2, 0.8, 1)$ |
| Case B | $\beta = 0.75$ | $\beta = 0.8$ | $\beta = 0.85$ |
| Case C | $\omega = 10^{-6}$ | $\omega = 10^{-5}$ | $\omega = 10^{-4}$ |
| Case D | $\mu = 10^{-4}$ | $\mu = 10^{-3}$ | $\mu = 10^{-2}$ |

**Table 4: Triple change-points**. Performance measurements for 100 replications of the scenarios described in Section 2.3, with thresholds corresponding to the quantiles $\widehat{q}_\tau$ at $\tau = 0.95, 0.975, 0.99$, and functions $\phi_2$ and $\phi_1$. The values in the table are the counts of each statistic out of 100 simulations.

| | | *Function* $\phi_2$ | | | | | | | | | | *Function* $\phi_1$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Upper tail | | | | | Lower Tail | | | | | Upper tail | | | | | Lower Tail | | | | |
| | | *A.1* | *A.2* | *B* | *C* | *D* | *A.1* | *A.2* | *B* | *C* | *D* | *A.1* | *A.2* | *B* | *C* | *D* | *A.1* | *A.2* | *B* | *C* | *D* |
| 0.99 | TP1 | 57 | 41 | 65 | 99 | 90 | 54 | 0 | 65 | 100 | 30 | 55 | 28 | 59 | 100 | 77 | 50 | 2 | 58 | 100 | 51 |
| | TP2 | 0 | 6 | 13 | 35 | 0 | 0 | 0 | 15 | 34 | 0 | 4 | 2 | 11 | 31 | 3 | 0 | 1 | 15 | 32 | 2 |
| | TP3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | FP1 | 98 | 99 | 99 | 99 | 55 | 99 | 4 | 99 | 97 | 100 | 82 | 95 | 100 | 100 | 0 | 73 | 8 | 100 | 100 | 100 |
| | FP3 | 0 | 0 | 1 | 25 | 0 | 0 | 0 | 1 | 16 | 0 | 0 | 0 | 6 | 31 | 94 | 0 | 0 | 1 | 26 | 1 |
| 0.975 | TP1 | 51 | 56 | 67 | 100 | 96 | 48 | 0 | 72 | 100 | 54 | 10 | 12 | 60 | 100 | 82 | 3 | 3 | 66 | 100 | 71 |
| | TP2 | 1 | 2 | 26 | 59 | 1 | 3 | 0 | 24 | 70 | 0 | 2 | 0 | 20 | 16 | 4 | 1 | 0 | 16 | 12 | 8 |
| | TP3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | FP1 | 87 | 96 | 100 | 100 | 93 | 84 | 2 | 100 | 100 | 100 | 15 | 35 | 100 | 100 | 100 | 5 | 10 | 100 | 100 | 100 |
| | FP3 | 0 | 0 | 5 | 50 | 0 | 0 | 0 | 3 | 2 | 0 | 0 | 0 | 3 | 21 | 1 | 0 | 0 | 5 | 16 | 5 |
| 0.95 | TP1 | 9 | 25 | 98 | 100 | 94 | 19 | 2 | 98 | 100 | 83 | 21 | 2 | 57 | 100 | 78 | 34 | 3 | 64 | 99 | 72 |
| | TP2 | 0 | 0 | 46 | 44 | 1 | 0 | 0 | 38 | 49 | 7 | 0 | 0 | 17 | 4 | 4 | 1 | 0 | 21 | 2 | 8 |
| | TP3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | FP1 | 27 | 66 | 90 | 100 | 100 | 36 | 3 | 97 | 100 | 100 | 67 | 20 | 100 | 100 | 100 | 68 | 14 | 100 | 100 | 100 |
| | FP3 | 0 | 0 | 3 | 40 | 1 | 0 | 0 | 4 | 46 | 1 | 0 | 0 | 6 | 11 | 2 | 0 | 0 | 6 | 8 | 5 |

### 2.3.3 Single Change-point

Even though at first glance this may seem to be the most straightforward case, the presence of a single change-point is a bit counterintuitive as it entails that a change during the year does indeed not happen. However, as empirical analysis may yield a persistent single change-point pattern, we formulate some theoretical explanations that may account for this singularity, and run simulations to verify their plausibility.

Our first proposal asserts that at a certain point of the year an abrupt change in the tail shape occurs and it then fades out over time. We try to reproduce this effect by drawing innovations $\eta_t$ from a standard Normal to the left of the single change-point, then switch to a standardized Student t with degrees of freedom slowly increasing as we move away from the change-point.

A second possible explanation is that there are two real change-points, but they are so close to each other that the test statistic is not powerful enough to detect both of them. Under this scenario, what appears as a year with a single change-point is actually a year with two change-points that the algorithm does not recognize. To account for this effect we simply draw the innovations from a $N(0,1)$ in one period, and a $t(4)$ in the other, but contrarily to the double change-point case, we set $\iota_0 = 110$ and $\iota_1 = 150$.

We report in Table 5 the performance statistics of the algorithm. Our first proposal generates too much noise: the true change-point is rarely identified and there are many false positives. The second proposal produces the sought after behaviour and thus may be a plausible explanation for the one change-point case.

### 2.3.4 A full experiment

Our simulations challenged our algorithm's ability of finding the true change-points. The full two-step approach however foresees that after having identified the intervals, we study the behaviour in the tails by fitting a GP distribution. It is important to understand how difficult it is to retrieve the true tail parameters given the noise introduced by both

**Table 5: Single change-point.** Results for 100 replications of the single change-point scenarios described in Section 2.3.3, with thresholds $\widehat{q}_\tau$ at $\tau = 0.95, 0.975, 0.99$, and functions $\phi_2$ and $\phi_1$. *Smooth* refers to a single change-point that fades out over the year, while *Short* refers to two change-points occurring at a short distance. The values in the table are the counts of each statistic out of 100 simulations. The values marked with the asterisk refer to cases in which more than one change-point was found.

| | Function $\phi_2$ | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Upper tail | | | | | | Lower Tail | | | | | |
| | Smooth | | | Short | | | Smooth | | | Short | | |
| | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.99 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.99 |
| TP | 1 | 3 | 3 | 1 | 6 | 34* | 1 | 3 | 4 | 0 | 6 | 26* |
| FP1 | 9 | 22 | 55 | 5 | 15 | 43 | 6 | 26 | 49 | 3 | 13 | 40 |
| FP2 | 9 | 22 | 55 | 0 | 0 | 2 | 6 | 26 | 49 | 0 | 0 | 2 |
| | Function $\phi_1$ | | | | | | | | | | | |
| | Upper tail | | | | | | Lower Tail | | | | | |
| | Smooth | | | Short | | | Smooth | | | Short | | |
| | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.99 | 0.95 | 0.975 | 0.99 | 0.90 | 0.95 | 0.99 |
| TP | 0 | 0 | 0 | 4 | 0 | 3 | 0 | 0 | 1 | 7 | 2 | 2 |
| FP1 | 4 | 2 | 27 | 11 | 6 | 10 | 6 | 5 | 25 | 11 | 4 | 4 |
| FP2 | 4 | 2 | 27 | 2 | 0 | 0 | 6 | 5 | 25 | 1 | 0 | 1 |

the pre-processing model and the change-point algorithm.

We consider two GARCH(1,1) processes with $\alpha = 0.1$, $\beta = 0.8$, and respectively $\eta_t \sim N(0,1)$ and $\eta_t \sim t(4)$. We generate a double change-point sequence of $50$ years $\times$ $250$ days, with $\iota_0 = 125$ and $\iota_1 = 250$. Then, we filter the observations with a GARCH(1,1) model to obtain the estimated residuals $\widehat{\eta}_t$. Finally, we apply the change-point algorithm considering the function $\phi_2$ and $\tau = 0.99$, and fit a GP distribution to the excesses over the 99th quantile of each interval. Our purpose is to verify whether we are able to recover the true tail parameters of the two innovation processes, $\xi = 0$ for the $N(0,1)$ and $\xi = 0.25$ for the $t(4)$.

We discard for clarity those cases in which the algorithm does not identify exactly two change-points, as our main interest is to judge the ability of our two-step approach in recovering the values of the true tail parameters. In Figure 1, we report the $100$ estimates of the parameter $\xi$ of the GP distribution for the two intervals. We also display the $100$ estimates of $\xi$ obtained by fitting a GP distribution to data over the 99th quantile of a sample of $50 \times 125$ iid observations from respectively a $N(0,1)$ and a $t(4)$. The figure clearly shows that the properties of the estimator of the tail parameter are not jeopardized by our two-step procedure. The estimated values using the scaled residuals are similar to those obtained using the iid sequence with only a slight loss of efficiency.

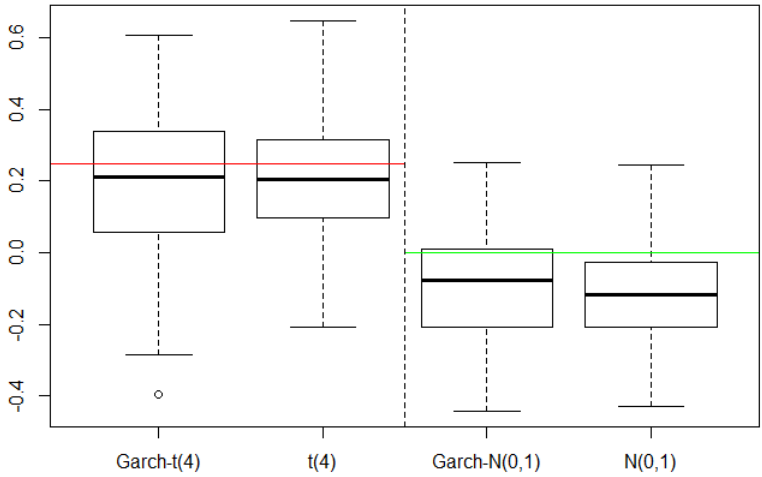
## 2.4 Tail seasonality in U.S. industries

### 2.4.1 Data description

We consider the equally-weighted returns of $12$ industry portfolios available in the Kenneth R. French data library. The dataset runs from 1926 to 2013 and was created from the CRSP database, assigning each NYSE, AMEX, and NASDAQ stock to an industry portfolio based on its four-digit SIC code.

As we benefit from long time series, we consider $39$ overlapping rolling windows of $50$ years. Precisely, we analyze data in the interval $1926 - 1975$, then consider $1927 - 1976$, and so on up to the interval $1964 -$

**Figure 1: Full experiment.** Estimated $\widehat{\xi}$ parameters for data over the 99th empirical quantile for GARCH(1,1)-$t(4)$ residuals, iid $t(4)$, GARCH(1,1)-$N(0,1)$ residuals and iid $N(0,1)$. Horizontal lines indicate the true values $\xi = 0.25$ for $t(4)$ and $\xi = 0$ for $N(0,1)$. There are always $50 \times 125$ observations in the iid case. For the filtered observations, the number varies due to estimated change-points.



2013. The use of rolling windows is common in econometrics and finance (Chan et al., 1999; Merton, 1980), but there are no clear rules regarding the length of the interval and the overlaps to include, thus making these settings case-dependent. Our choice of a 50-year rolling window is a trade-off between the need for a sample long enough to conduct an extreme value analysis and the desire to carry out many analyses to verify whether the identified seasons are coherent over time.

### 2.4.2 Pre-processing approach

Given that we consider moving windows spanning over a time interval of 50 years, it is not sensible to assume time invariance of the ARMA-GARCH parameters over that period. In particular, Mikosch and Starica (2004) discuss how time structural changes in the parameters might be harmful for the Gaussian QMLE. Consequently, we allow for a more dynamic environment by fitting an ARMA($u$,$v$)-GJR-GARCH(1,1) model to every five years of data, a period considered long enough to permit a good inference. As noted in the introduction, the financial literature documents many seasonalities in the mean and variance behavior. We found some of these anomalies in the mean but not in the variance. First, in all the moving windows of the different industry series, the average return in January tends to be much higher than during the rest of the period, particularly in the first week of the month. Second, we also noticed a daily effect where Monday typically presents lower mean returns, and Friday higher mean returns, with respect to the other days. We did not find substantial difference in the unconditional variance across the months of the year. Our investigation was performed by binding observations by month and calculating the Median Absolute Deviation. We used a robust measure of the variability to exclude extreme events, as extreme seasonalities could be interpreted as seasonalities in the volatility if we used the standard deviation.

To take into account these effects and better single out tail changes, we consider two augmented versions of the ARMA model. First, we consider a season in the mean for the first week of January and add a

dummy variable only for this period, $d_{FWJ}$, to the baseline specification, i.e.

$$r_t = \mu + \delta_1 d_{FWJ} + \sum_{k=1}^{u} \phi_k r_{t-k} + \sum_{z=1}^{v} \psi_z \epsilon_{t-z} + \epsilon_t \ . \tag{2.9}$$

Then, we further augment the model and consider all the seasonal anomalies in the mean: we add a dummy variable for Monday ($d_{Mon}$) and Friday ($d_{Fri}$), and a dummy variable for the whole month of January ($d_{Jan}$). Thus the model specification becomes,

$$r_t = \mu + \delta_1 d_{Jan} + \delta_2 d_{Mon} + \delta_3 d_{Fri} + \sum_{k=1}^{u} \phi_k r_{t-k} + \sum_{z=1}^{v} \psi_z \epsilon_{t-z} + \epsilon_t \ . \tag{2.10}$$

For both the baseline specification in (3.2) and the augmented specifications in (2.9)-(2.10), we consider various values of lags in the parameters of the conditional mean model. Specifically, while we consider only the first order specification of the conditional variance model (Hansen and Lunde, 2005), we employ ARMA($u$,$v$) models with $\{u, v \leq 2\}$ allowing for higher orders of dependence in the conditional mean. We report only the results for the model with the lowest BIC. Specifically, since in every window we fit an ARMA-GJR-GARCH model to every five years of data, we compare the BIC of the 12 models in each five-year interval, and select the specification with the highest number of lowest value of the BIC[7]. We found that in nine industries out of 12 the optimal specification is an ARMA(1,2)-GJR-GARCH(1,1) model without external regressors, while in the other three industries, the preferred model is the first order full specification that includes external regressors for the different seasonal effects in the mean. We report the results for only one specification, however the findings in Section 2.4.3 extend to the larger models.

### 2.4.3 *Tail season* identification

The change-point algorithm is applied to the scaled residuals from the pre-processing model. In Section 2.3, we established the better performance of the algorithm using function $\phi_2$, the threshold $\widehat{q}_\tau$ set at the 99th

[7] $12 = 3 \times 4$ models are considered: three specifications defined by the possible inclusion of external regressors, times four possible combinations of the lags.

empirical quantile, and a significance level $\alpha = 0.05$. In particular, the test with these settings is more powerful in detecting the true change-points and this is substantiated also by the three change-points analysis reported in the supplementary appendix. We use these settings throughout, as a different configuration may increase the probability of missing a change-point or finding spurious change-points.

In Figures 2-4, we report the change-points in the upper tail of the industries that exhibit at least one change-point in any of the 39 windows. Similar results are displayed in Figure 5 for the lower tail of the only two industries that present change-points. Given that there is no evidence of change-points in the lower tail, apart for the random occurrence reported in Figure 5, we conclude that lower tail behaviour remains constant over the year. Contrarily, we report evidence of change-points in the upper tail for nine out of 12 industries, signalling a strong seasonal behaviour in the upper tail. As pointed out in Section 2.3, this evidence is due to pure tail changes given that one-tailed change-points can only occur as a consequence of variation in only one of the tails of the innovation distribution.

Apart for slight time variations, the change-points are quite stable within industries, while they tend to be different in terms of number and location across industries. We note the presence of a short **Winter** season and a long **Summer** season, with the length of these seasons depending on the industry. In particular, Winter seems to be critical in that fluctuation in the tail looks frequent during this period, as witnessed by the appearance of a third change-point that splits this season in two intervals in both the *Manufacturing* and *Shops* industries.

In a few windows of the *Other* industry, a single change-point is observed. Focusing on the whole spectrum of change-points across the 39 windows of this industry, we can notice that windows with two change-points very close to each other co-exist with windows with a single change-point. This evidence gives support to the hypothesis put forth in Section 2.3 that only one change-point may be detected if the actual season is quite small. Either the power of the test is too low, or the second change-point happens within the 30-day bound defined by the $\lambda$ parameter in

Equation (2.6).

Once we have identified the change-points, we proceed with the study of the tail behaviour within each interval. Although we filter the observations with a pre-processing model, it is important to verify the level of dependence in the extremes of the estimated residuals as it may be quite different from that which occurs at mean levels. In Figure 25 we report, for the $10$ industries presenting at least one change-point in the upper tail, the estimated values of the extremal index $\theta$ obtained with the interval estimator of Ferro and Segers (2003). Most of the estimates are in the interval $(0.8, 1)$. We argue that the pre-processing model has filtered away much of the dependence and we treat the residuals as independent.

In each identified interval, we fit the GP to the exceedances over a threshold $q$. There exists a huge literature regarding threshold selection in EVT (Scarrott and MacDonald, 2012), nonetheless a full agreement regarding the right way to proceed has not been achieved. We fit the GP distribution to data over a sequence of thresholds $q_n$, from the 90th up to the 99th empirical quantile, and qualitatively evaluate the tail behaviour.

To assess the goodness of fit, we visually inspect the QQ-plots. Given the large number of plots, we do not report them here, but note that most of the time the observations nicely adhere to the theoretical lines of the GP. We also test the null hypothesis of a GP distribution with the Anderson-Darling and the Cramer-von Mises tests of Choulakian and Stephens (2001). In Figures 6-14, we report the frequency of rejection and non rejection for both tests at the different $q_n$ thresholds.

Specifically, to clearly distinguish between Summer and Winter, we consider only windows exhibiting two seasons. The figures show that in most cases the assumption that the upper tail follows a GP distribution is not rejected. Summer seasons tend to present a higher rate of rejection as a consequence of the larger number of observations. This implies that some of the thresholds are too low for the convergence result of Pickands (1975). On a few occasions, the number of observations over the 99th quantile in Winter is insufficient to compute MLE.

In Figure 15, we show a scatter plot of the $\nu$ and $\xi$ parameter estimates, obtained over the $10$ different threshold levels in $q_n$, for the Sum-

mer and the Winter seasons. Specifically, we discard the estimated values for which the GP distribution null hypothesis was rejected at the $5\%$ level either by the Anderson-Darling or the Cramer-von Mises test. The figure shows that the two seasons are not substantially different with respect to $\xi$, but they are strongly different in terms of scale. The values of $\nu$ are much higher in Winter than in Summer. In Figures 16-24 we report the empirical quantiles of $\widehat{\eta}_t$, considering only those windows for which exactly two seasons are identified. The heat map tracks the change in the value of the quantile by changing the colour of the corresponding tile. From these plots two main messages can be extracted. First, the fact that much of the area in the higher quantiles of the Winter plots is covered by red tiles confirms that Winter is subject to more extreme observations. Second, the fact that the difference between the Summer and the Winter plots at the 90th quantile is small confirms that this behaviour is a consequence of a pure shift in the tail shape, and not a consequence of the mean or the variance process. Indeed, we compute the average difference between the quantiles at the same level in Winter and Summer, and find that it increases as we move toward the 99th quantile. We conclude that while both seasons present a similar upper tail decay, the magnitude of the tail events is higher in Winter.

**Figure 2: Upper tail change-points**. Each line corresponds to one of the 39 rolling windows. Lines 1 and 39 represent respectively the time intervals $1926 - 1975$ and $1964 - 2013$, with the other lines representing the intervals ranging in between at yearly steps. The red crosses represents the identified change-points over the year.

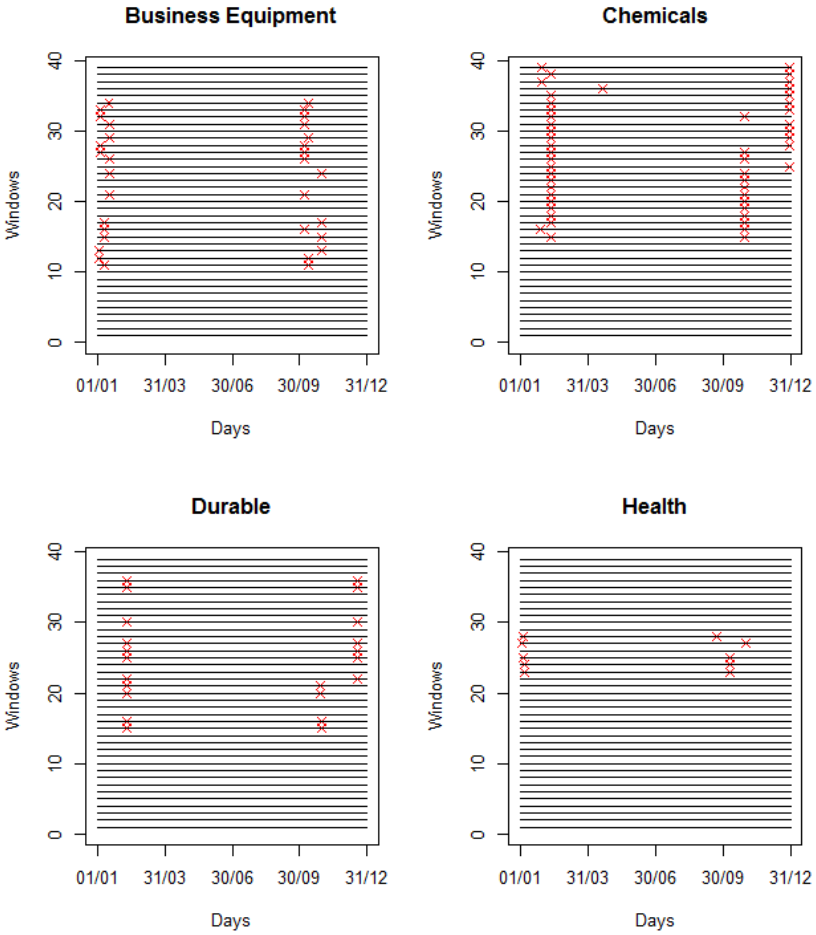


**Figure 3: Upper tail change-points**. Each line corresponds to one of the 39 rolling windows. Lines 1 and 39 represent respectively the time intervals $1926 - 1975$ and $1964 - 2013$, with the other lines representing the intervals ranging in between at yearly steps. The red crosses represents the identified change-points over the year.

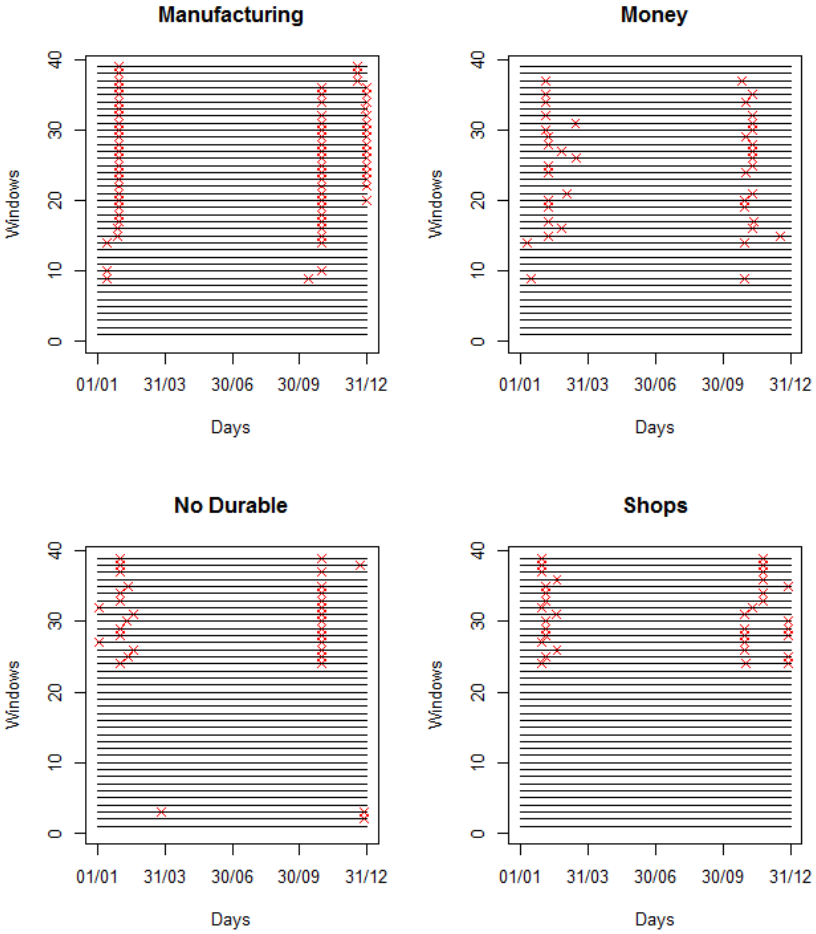


**Figure 4: Upper tail change-points**. Each line corresponds to one of the 39 rolling windows. Lines 1 and 39 represent respectively the time intervals $1926 - 1975$ and $1964 - 2013$, with the other lines representing the intervals ranging in between at yearly steps. The red crosses represents the identified change-points over the year.

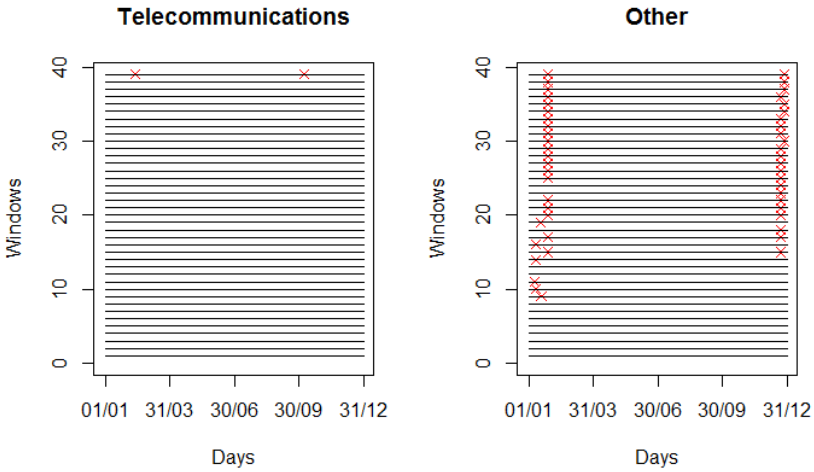


**Figure 5: Lower tail change-points**. Each line corresponds to one of the 39 rolling windows. Lines 1 and 39 represent respectively the time intervals $1926 - 1975$ and $1964 - 2013$, with the other lines representing the intervals ranging in between at yearly steps. The red crosses represents the identified change-points over the year.

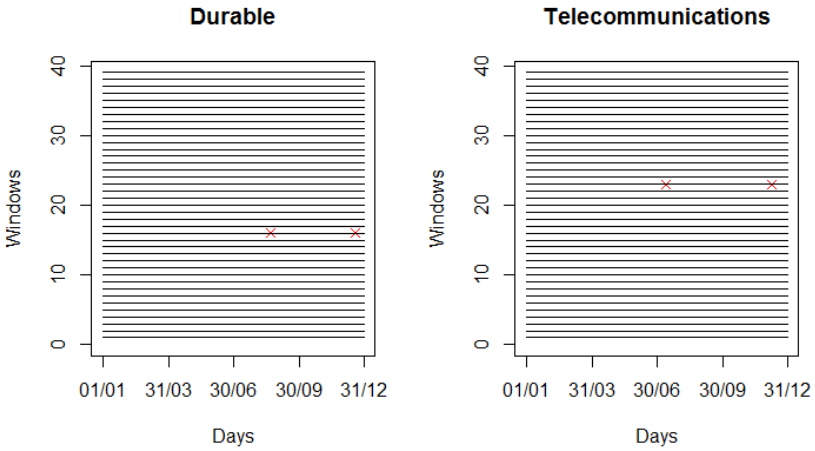


**Figure 6: Goodness of fit - Business Equipment.** Non rejection (*light*) and rejection (*dark*) frequencies for the GP null hypothesis. Only windows for which two seasons were identified are considered. The horizontal axis refers to the sequence of threshold levels $q_n$.

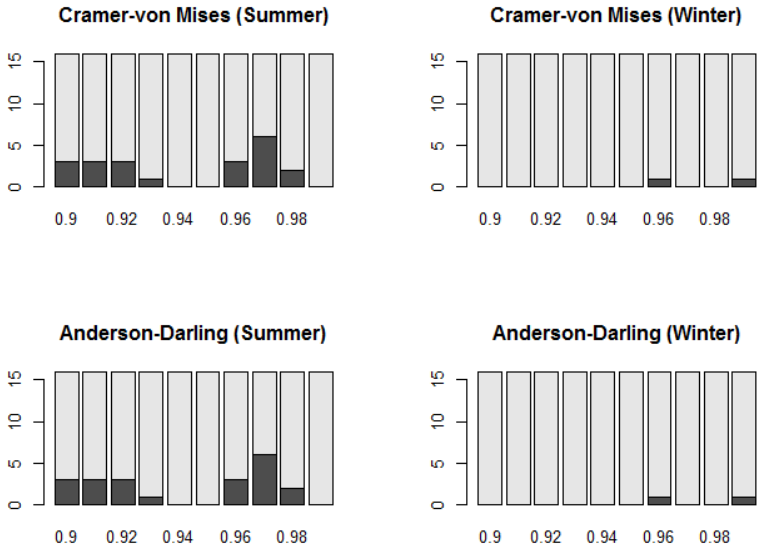


**Figure 7: Goodness of fit - Chemicals.** Non rejection (*light*) and rejection (*dark*) frequencies for the GP null hypothesis. Only windows for which two seasons were identified are considered. The horizontal axis refers to the sequence of threshold levels $q_n$. At the 99th quantile of the Winter season it was not always possible to obtain estimates of the GPD parameters because of the small number of observations.

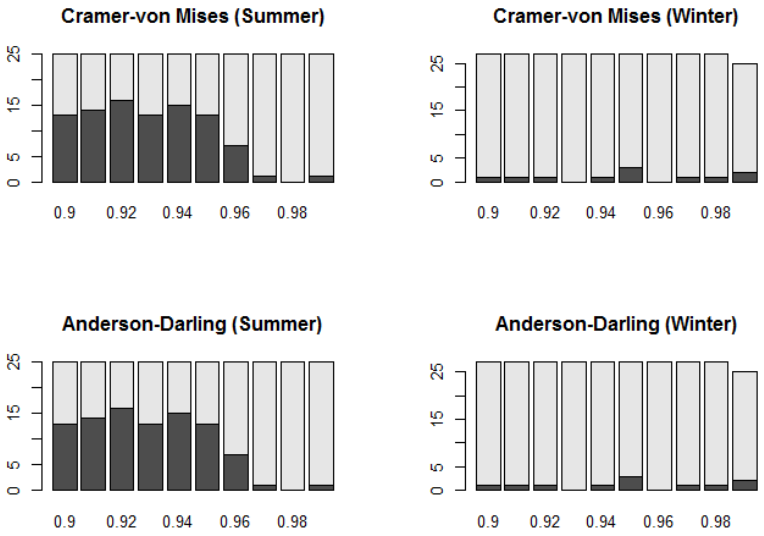


**Figure 8: Goodness of fit - Durable.** Non rejection (*light*) and rejection (*dark*) frequencies for the GP null hypothesis. Only windows for which two seasons were identified are considered. The horizontal axis refers to the sequence of threshold levels $q_n$.

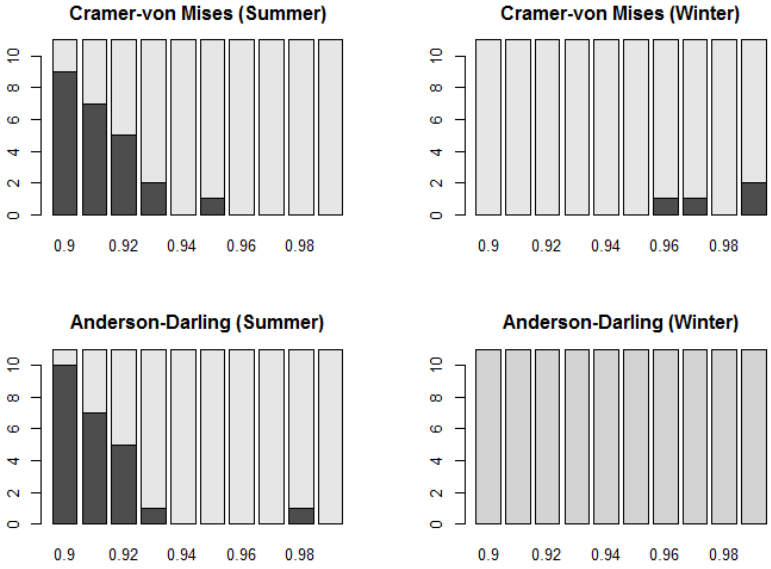


**Figure 9: Goodness of fit - Health.** Non rejection (*light*) and rejection (*dark*) frequencies for the GP null hypothesis. Only windows for which two seasons were identified are considered. The horizontal axis refers to the sequence of threshold levels $q_n$.

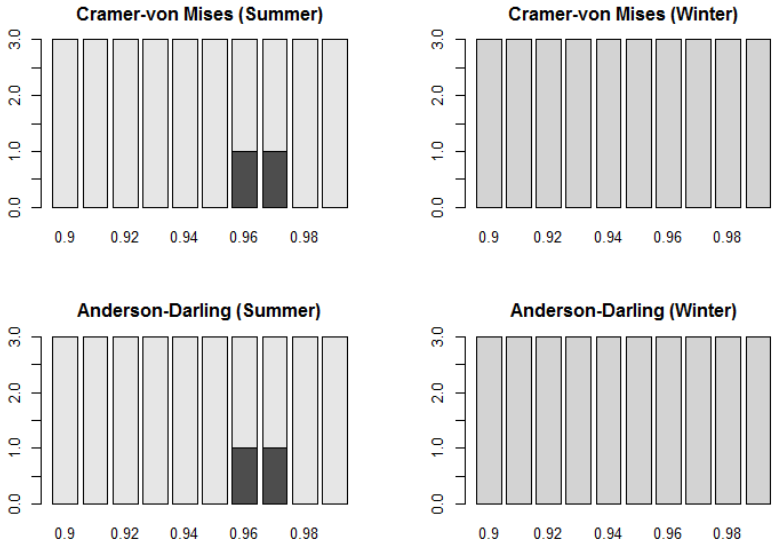


**Figure 10: Goodness of fit - Manufacturing.** Non rejection (*light*) and rejection (*dark*) frequencies for the GP null hypothesis. Only windows for which two seasons were identified are considered. The horizontal axis refers to the sequence of threshold levels $q_n$.

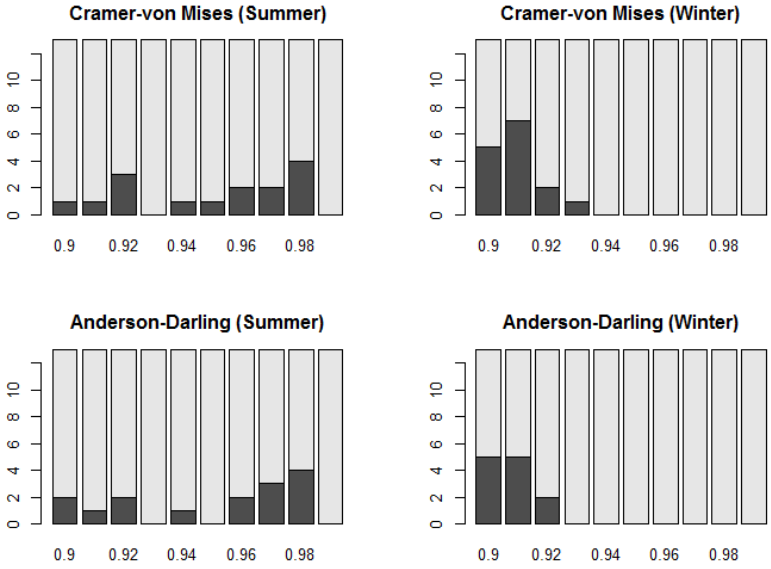


**Figure 11: Goodness of fit - Money.** Non rejection (*light*) and rejection (*dark*) frequencies for the GP null hypothesis. Only windows for which two seasons were identified are considered. The horizontal axis refers to the sequence of threshold levels $q_n$.

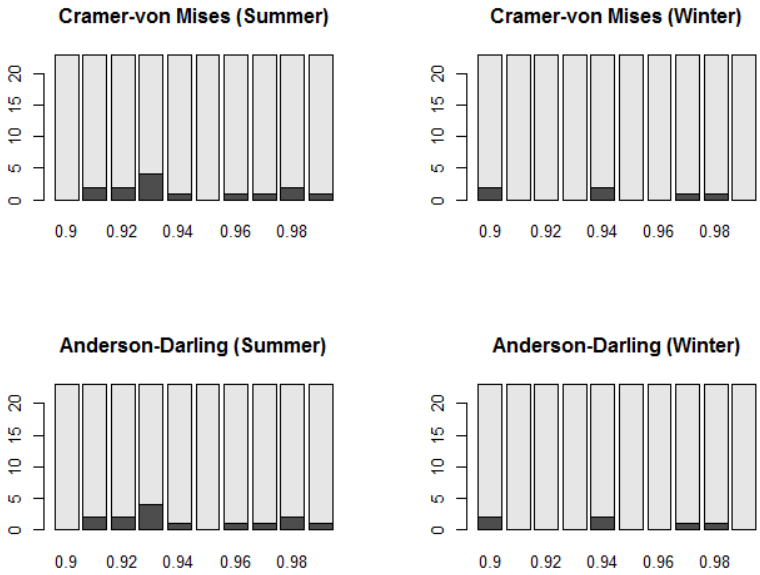


**Figure 12: Goodness of fit - No Durable.** Non rejection (*light*) and rejection (*dark*) frequencies for the GP null hypothesis. Only windows for which two seasons were identified are considered. The horizontal axis refers to the sequence of threshold levels $q_n$.

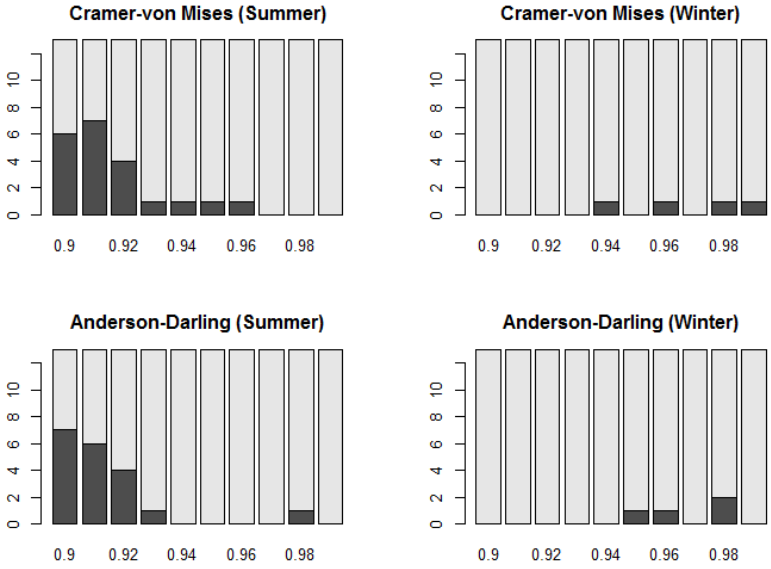


**Figure 13: Goodness of fit - Shops.** Non rejection (*light*) and rejection (*dark*) frequencies for the GP null hypothesis. Only windows for which two seasons were identified are considered. The horizontal axis refers to the sequence of threshold levels $q_n$.

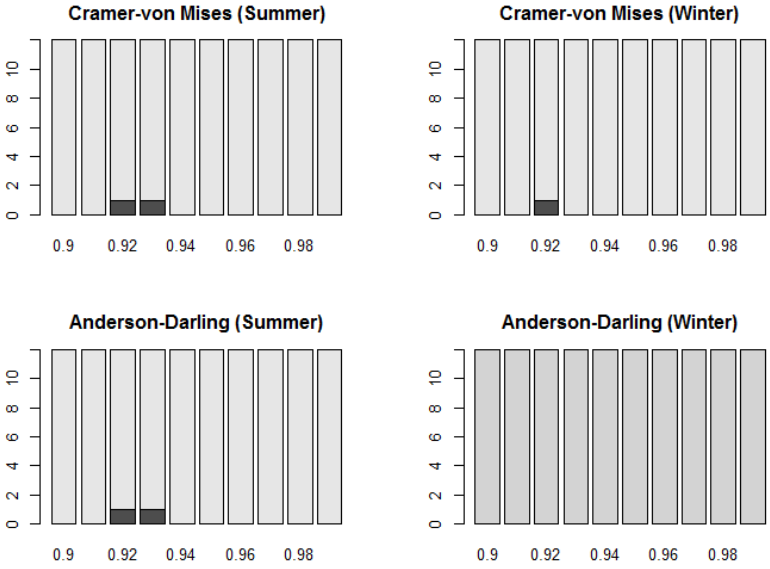


**Figure 14: Goodness of fit - Other.** Non rejection (*light*) and rejection (*dark*) frequencies for the GP null hypothesis. Only windows for which two seasons were identified are considered. The horizontal axis refers to the sequence of threshold levels $q_n$. At the 99th quantile of the Winter season it was not always possible to obtain estimates of the GPD parameters because of the small number of observations.

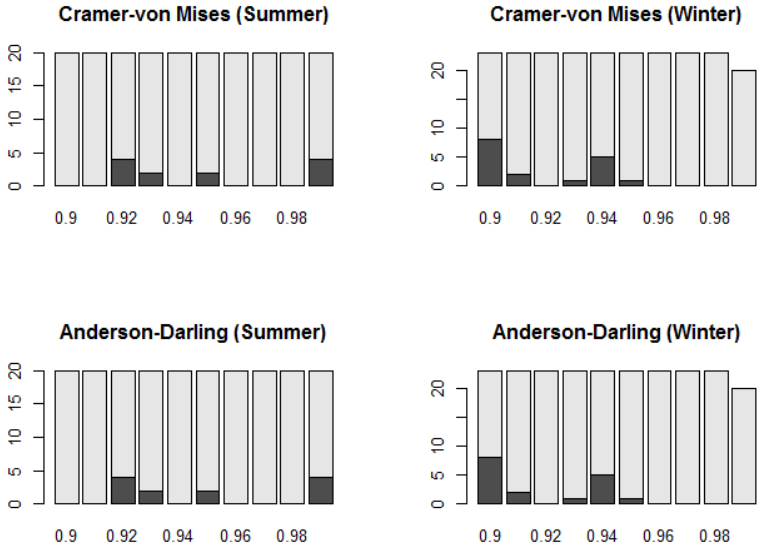


**Figure 15: GP parameters scatterplot.** Estimated parameters of the GP distribution in Equation (2.4). For each industry we consider only windows for which two change-points were identified and plot the estimates for the Winter season (blue) and the Summer season (red).

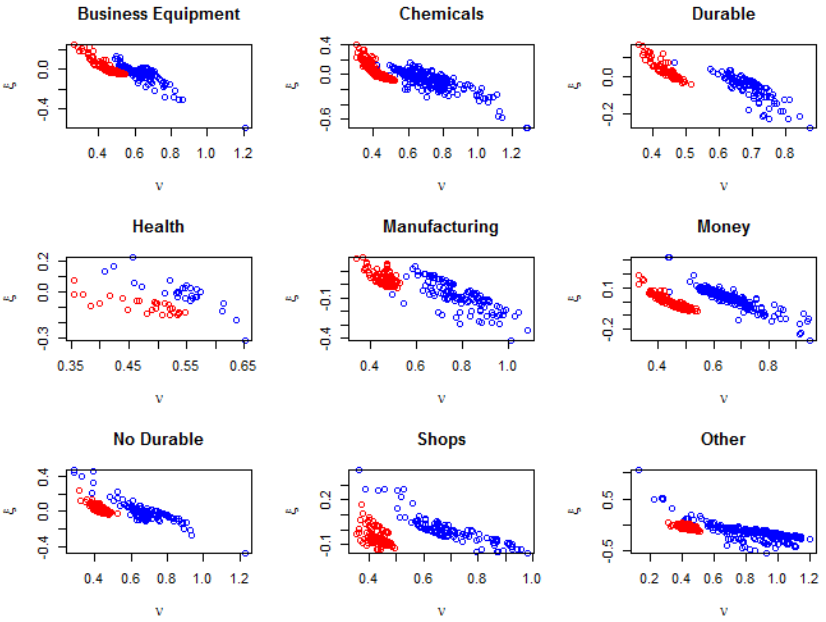


**Figure 16: Empirical tails - Business Equipment.** 90th to 99th empirical quantiles for the Summer and Winter seasons. Only the windows with exactly two change-points are considered. The tile corresponding to a fixed window and threshold shifts from green to red as the value of the empirical quantile increases.

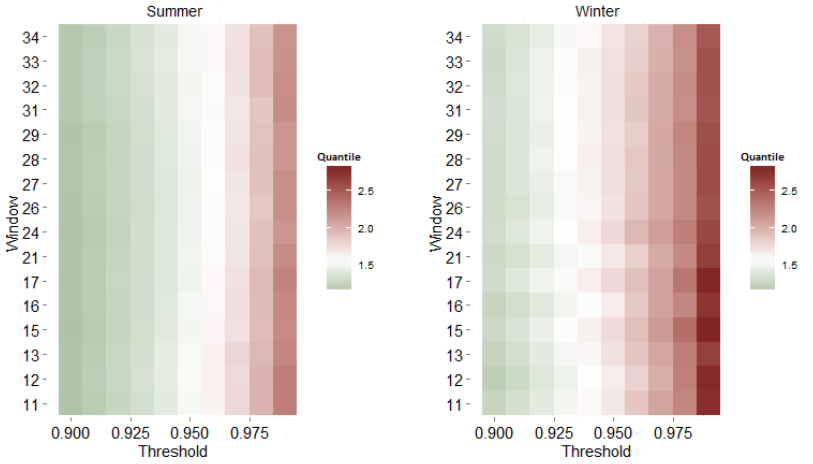


**Figure 17: Empirical tails - Chemicals.** 90th to 99th empirical quantiles for the Summer and Winter seasons. Only the windows with exactly two change-points are considered. The tile corresponding to a fixed window and threshold shifts from green to red as the value of the empirical quantile increases.

**Figure 18: Empirical tails - Durable.** 90th to 99th empirical quantiles for the Summer and Winter seasons. Only the windows with exactly two change-points are considered. The tile corresponding to a fixed window and threshold shifts from green to red as the value of the empirical quantile increases.

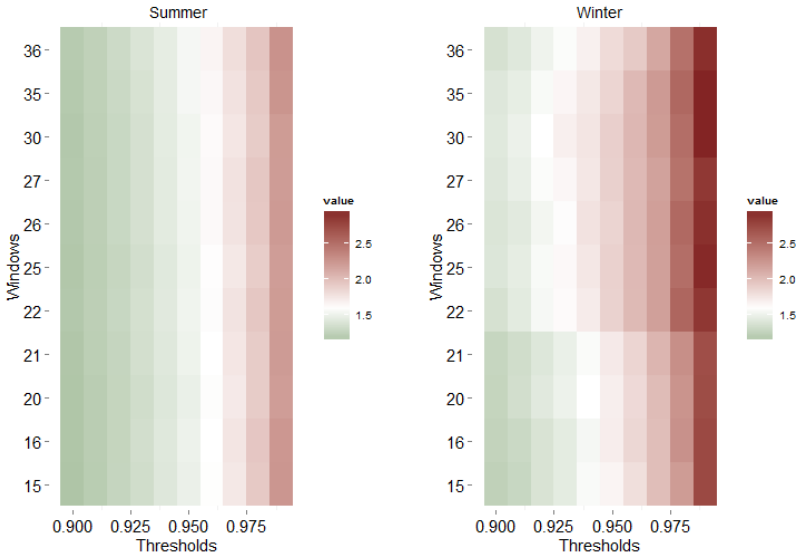


**Figure 19: Empirical tails - Health.** 90th to 99th empirical quantiles for the Summer and Winter seasons. Only the windows with exactly two change-points are considered. The tile corresponding to a fixed window and threshold shifts from green to red as the value of the empirical quantile increases.

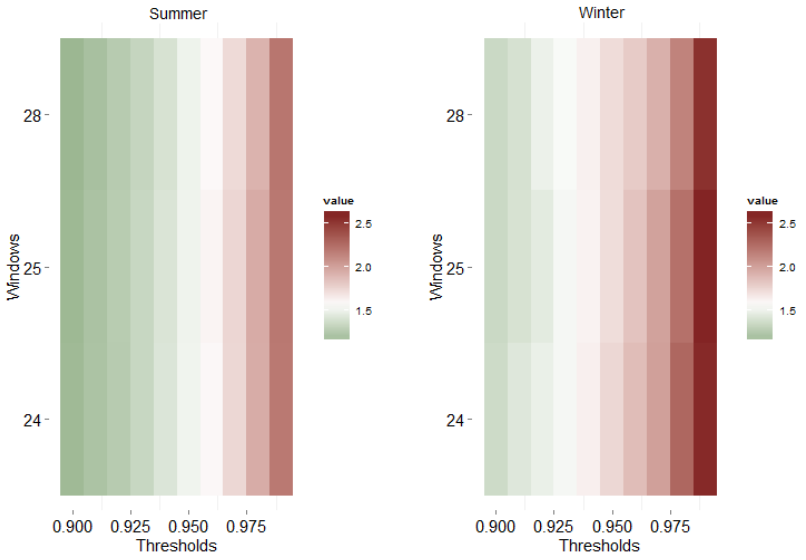


**Figure 20: Empirical tails - Manufacturing.** 90th to 99th empirical quantiles for the Summer and Winter seasons. Only the windows with exactly two change-points are considered. The tile corresponding to a fixed window and threshold shifts from green to red as the value of the empirical quantile increases.

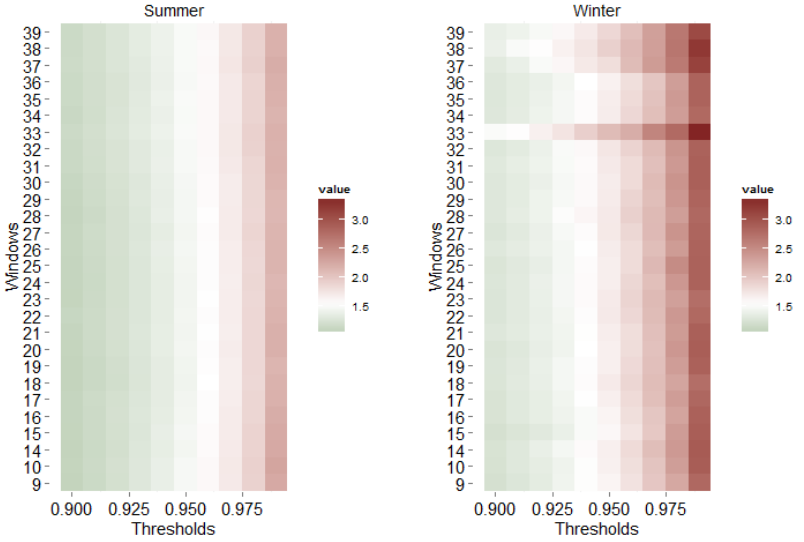


**Figure 21: Empirical tails - Money.** 90th to 99th empirical quantiles for the Summer and Winter seasons. Only the windows with exactly two change-points are considered. The tile corresponding to a fixed window and threshold shifts from green to red as the value of the empirical quantile increases.

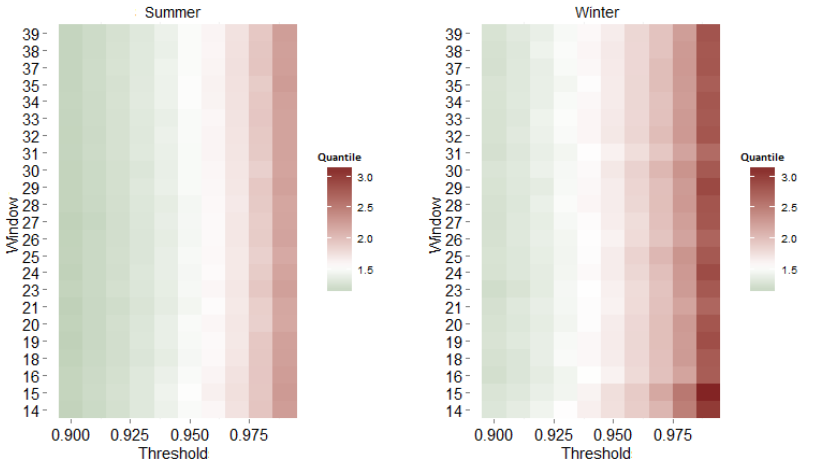


**Figure 22: Empirical tails - No Durable.** 90th to 99th empirical quantiles for the Summer and Winter seasons. Only the windows with exactly two change-points are considered. The tile corresponding to a fixed window and threshold shifts from green to red as the value of the empirical quantile increases.

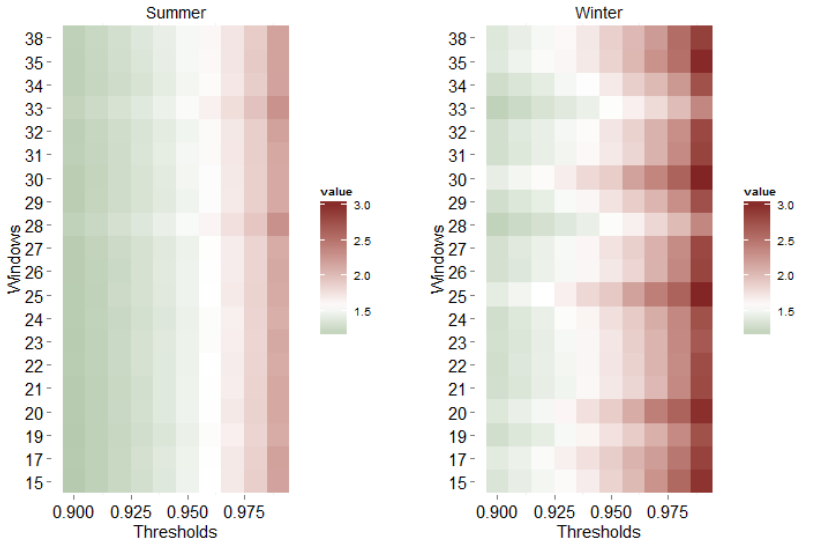


**Figure 23: Empirical tails - Shops.** 90th to 99th empirical quantiles for the Summer and Winter seasons. Only the windows with exactly two change-points are considered. The tile corresponding to a fixed window and threshold shifts from green to red as the value of the empirical quantile increases.

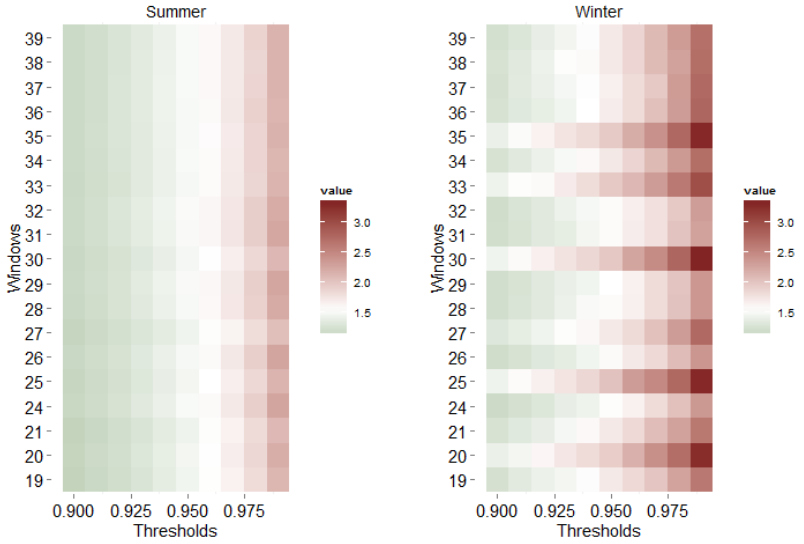


**Figure 24: Empirical tails - Other.** 90th to 99th empirical quantiles for the Summer and Winter seasons. Only the windows with exactly two change-points are considered. The tile corresponding to a fixed window and threshold shifts from green to red as the value of the empirical quantile increases.

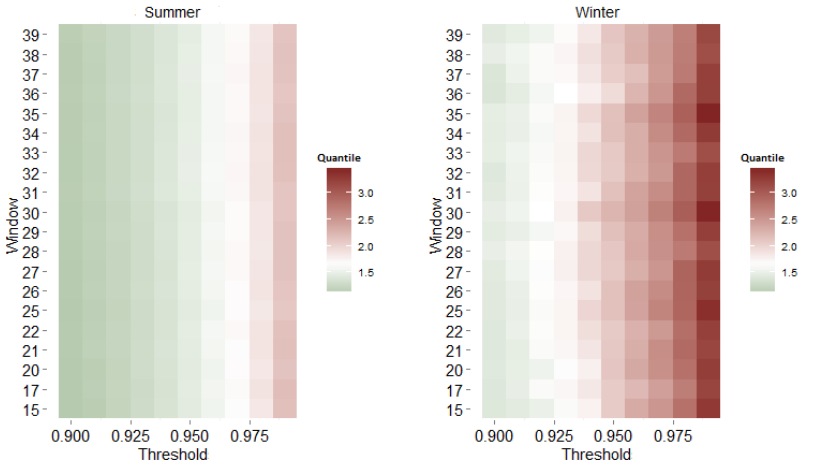
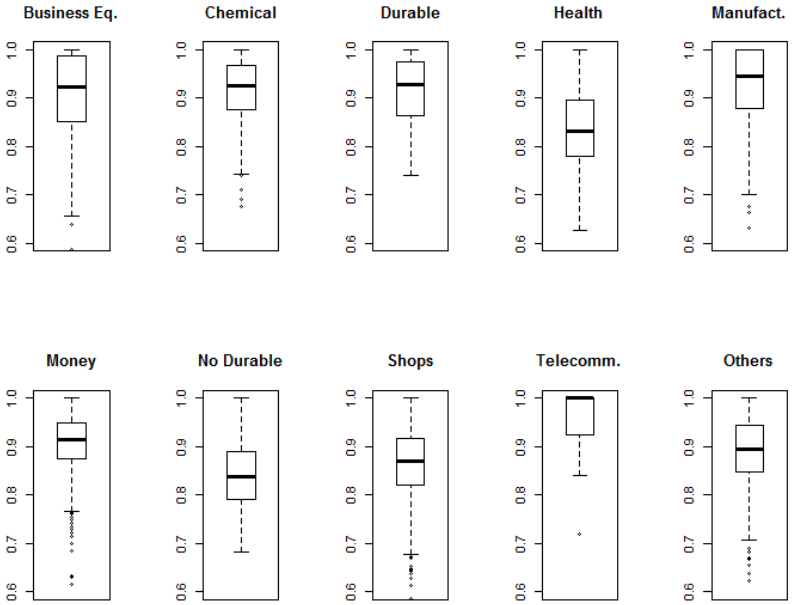


**Figure 25: Extremal index estimates**. Only sectors presenting at least one change-point in any of the 39 windows are reported. The box-plots pool together the tail index estimates over ten thresholds (from the 90th to the 99th quantiles) of both the Summer and the Winter seasons.




## 2.5 Tails vs skewness and kurtosis

The finance literature is typically concerned with the moments of the return distribution, and in particular, skewness and kurtosis. In this section, we try to establish a link between our findings on tail seasonalities and these two moments. Although it is difficult to determine an elegant mathematical relationship (if possible), the intuition is that the seasonal variation that we observed in the upper tail should somehow be reflected in the skewness and kurtosis.

The traditional approach to identify seasonalities computes sample statistics for observations binned by month. In Figure 26, we report the empirical skewness and kurtosis computed on the unconditional return for each month in every window. Each month has a natural seasonal colour, i.e. reds for Summer, browns for Autumn, blues for Winter, and greens for Spring. The different sectors present approximately the same U-shape behaviour indicating that there is probably a market-wide component generating this pattern. Though it is quite difficult to say whether a particular month presents higher skewness or kurtosis than the others, as different colours contribute to create the whole smile, two insights emerge. Brown points on the extreme right correspond to two events in the October following the Great Depression, so it's hard to argue that is a seasonality. Blue points generally lie to the right of red points, suggesting that unconditional returns might be more right-skewed in Winter. We pursue a qualitative investigation of this possibility and compute the skewness and kurtosis of the unconditional returns for the *Summer* and *Winter* seasons identified in Section 2.4. Figure 27 shows these measures computed for each 50-year window. Although there is no statistical foundation for adopting a decomposition of the year which follows the *seasons* arising from a specific model on the conditional tails, the figure suggests that there is probably also some seasonal behaviour in the skewness and kurtosis of the unconditional returns. In particular, the fact that *Winter* is right-skewed is consistent with the idea that this period is more prone to positive tail events.

Is the dynamic model for mean and variance able to capture the seasonalities in Figure 27? Figure 28 shows the monthly skewness and kurtosis computed on the estimated residuals from the ARMA-GARCH model considered in Section 2.4. We can see that the model captures the observations corresponding to the re-bound after the Great Depression, and the figure now looks more like a smirk than a smile. Also in this case, binning the observations by month does not return clear information, but one can still notice the tendency of the blue points to favor the right-side of the plot. In Figure 29 we compute the skewness and kurtosis of the residuals of the ARMA-GARCH model for the *Summer* and *Win-*

*ter* seasons. Once again, there is no justification for the use of these two seasons, but the figure still suggests that the innovations may present a seasonal component in both skewness and kurtosis. Figure 29 is similar to Figure 15 but not as clear. Our approach which focusses on the tails allows us to disentangle possible differences between the upper and the lower tails. Indeed, our empirical analysis strongly indicates that seasons in the extremes exist only in the upper tail, while the lower tail tends to be constant over time. It is difficult to appreciate this distinction while pursuing a study of the skewness and kurtosis, as the blurred clusters in Figure 29 show.

Relying on a monthly-based approach to identify seasonality does not return complete information, thus performing a sound statistical analysis is certainly worthwhile. The apparent seasonal behaviour in skewness and kurtosis is not limited to the unconditional returns, and it is not fully captured by a conditional model for the mean and the variance. This suggests that a direct analysis of the seasonal variation of these two moments could be interesting for several applications. We discuss some of them in the next section.

## 2.6 Implications for practices in finance

### 2.6.1 Impact on asset pricing

There is compelling evidence that higher moments of the returns distribution are relevant in asset pricing. Harvey and Siddique (2000) present a quadratic pricing kernel which allows the skewness to be priced, while Dittmar (2002) proposes a cubic pricing kernel which accounts also for the risk-aversion to kurtosis. Considering these frameworks, the impact of seasonalities in the skewness and kurtosis on asset pricing can be assessed through the definition of an appropriate econometric model which accounts for seasonal components in the skewness and kurtosis. An appealing econometric framework was developed in Harvey and Siddique (1999) where they use a non-central $t$ distribution to assess the impact of time variation in the conditional skewness on asset pricing.

**Figure 26: Unconditional monthly skewness and kurtosis.** Each colour represents a different month and the tones follow the natural seasons. Multiple points of the same colour refer to the same month over the 39 windows.

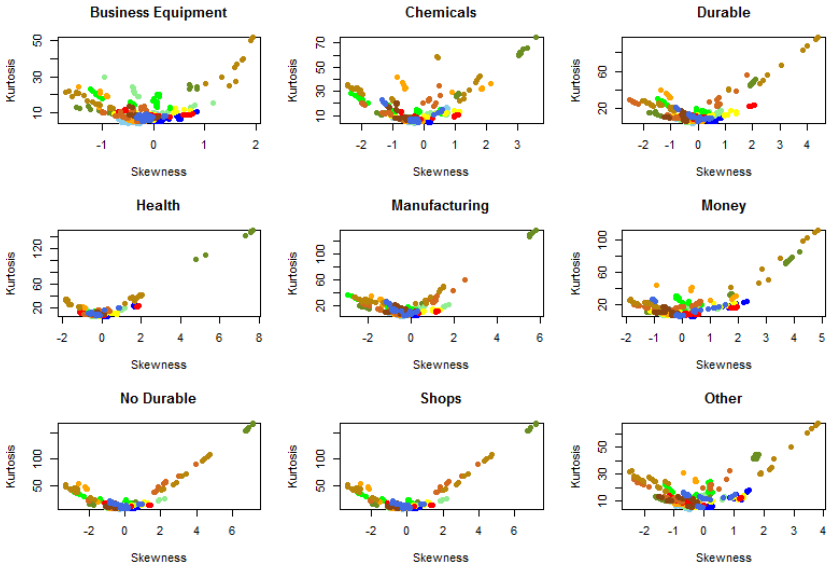


**Figure 27: Unconditional skewness and kurtosis in *Summer* and *Winter*.** For each industry we consider only windows for which two change-points were identified and plot the estimates for the Winter season (blue) and the Summer season (red).

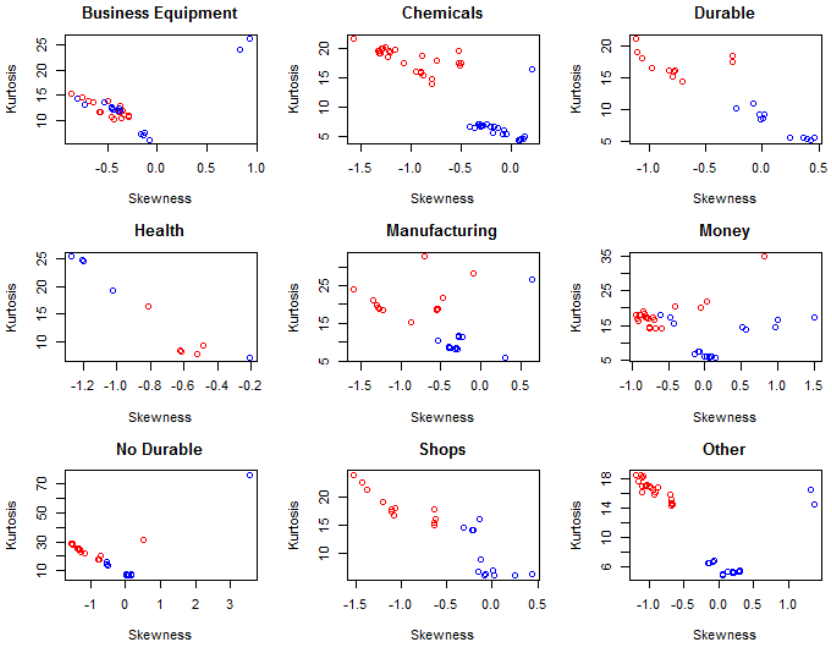


**Figure 28: Conditional monthly skewness and kurtosis as calculated on the residuals of the ARMA-GARCH model.** Each colour represents a different month and the tones follow the natural seasons. Multiple points of the same colour refer to the same month over the 39 windows.

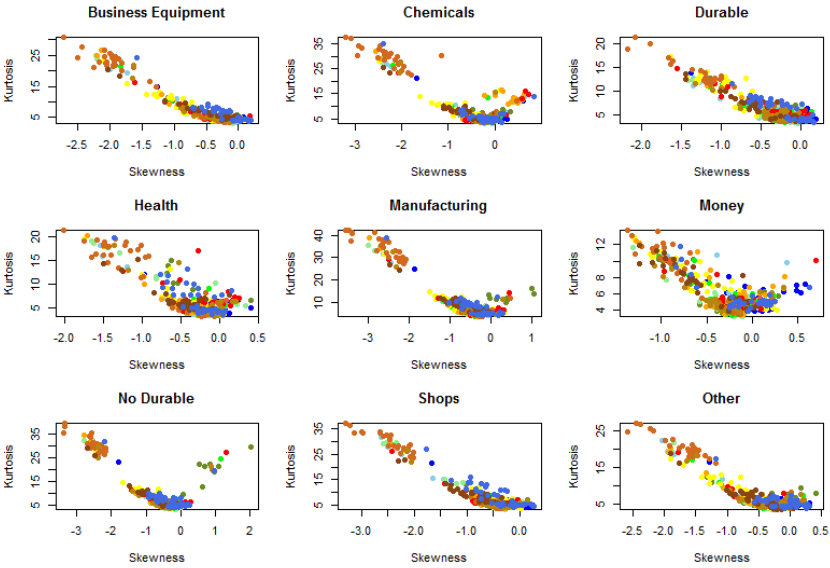
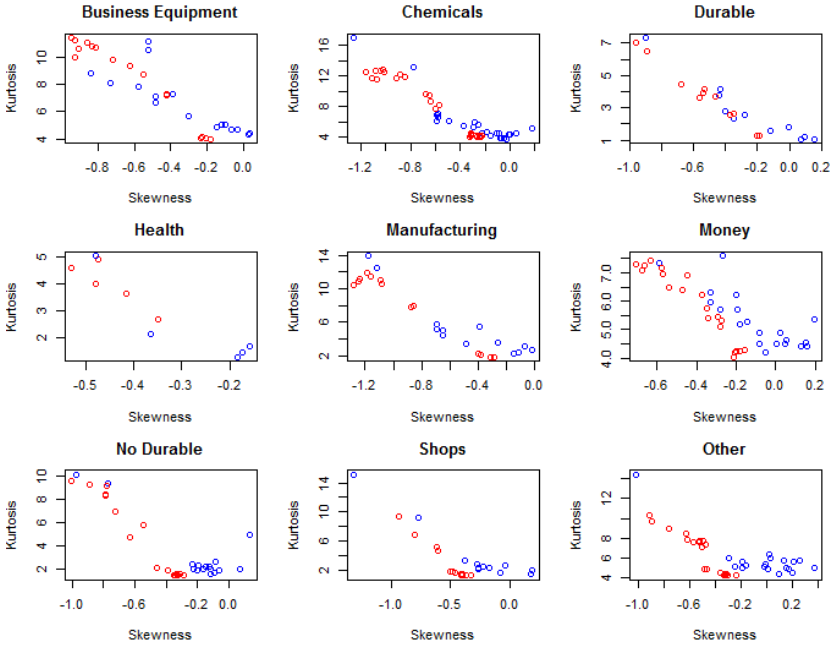


**Figure 29: Conditional skewness and kurtosis for *Summer* and *Winter* as calculated on the residuals of the ARMA-GARCH model.** For each industry we consider only windows for which two change-points were identified and plot the estimates for the Winter season (blue) and the Summer season (red).



In particular, they consider GARCH-type dynamics for both the conditional variance and skewness. One could extend this model to include a seasonal component in the skewness, using for example a periodic model as in Bollerslev and Ghysels (1996).

An alternative approach is followed by Maheu et al. (2013) who combine a cubic pricing kernel with a GARCH model with time-varying jumps. They show that the jump process affects both skewness and kurtosis, and this in turn has an impact on the pricing kernel. Following this path, one can define a jump intensity varying according to the seasons and investigate whether the jump component may be the source of seasonality in the tails. In the same spirit, another possible framework to investigate seasonalities in the jumps is the one of Bollerslev and Todorov (2011).

### 2.6.2 Impact on option pricing

Christoffersen et al. (2006) introduce conditional skewness in a discrete-time GARCH model to capture the volatility smirk. In particular, they use an inverse Gaussian distribution on top of an asymmetric GARCH model to reinforce the leverage parameter. They present a closed-form option pricing formula and show that there are some benefits from modelling the conditional skewness. Extending this framework to allow for a seasonal component in the conditional skewness could be interesting. Intuitively, given the evidence of increasing positive extreme events during *Winter*, one might find an increment in the demand for in-the-money call options, and subsequently, a seasonal change in the volatility smirk.

### 2.6.3 Impact on risk management

We found the emergence of seasonality in the upper tail of the return distribution. Although most of the analysis from a risk management perspective deals with the lower tail, the upper tail is still relevant to investors who pursue strategies involving short positions. Fully rational investors would take advantage of both bearish and bullish trends in the market. Long-traders hold a stock while it appreciates in price, and fear

the downside risk. On the other hand, short-traders react in the opposite way and dread any upside movements. When two different measures $\mathbb{P}_W$ and $\mathbb{P}_S$ determine the tail behaviour over the year, then the use of a constant measures $\mathbb{P}$ results in both an over- and underestimation of capital requirements and a mis-interpretation of risk.

## 2.7 Conclusions

Abundant literature has uncovered seasonality in both the mean and the variance of financial returns. This article further expands our knowledge through the development of a two-step procedure to identify yearly seasonality in the tails. In particular, we focus on changes in the tails of the innovation process. We first pre-whiten the returns to obtain a proxy for the innovations, and then identify *seasons* with a change-point algorithm for the extremes. We confirm the good performance of the new estimation procedure through a series of Monte Carlo simulation experiments and illustrate its applicability using 12 industry portfolios of US stocks data.

The results of our empirical investigation show that while the lower tail is approximately constant over the year, the upper tail of innovations is larger in *Winter* than in *Summer*, in nine out of 12 industries. We find that this difference is imputable to specific changes in the tail of the innovations, and not to variation in either the mean or the variance process. This evidence is consistent with the idea of the *Santa Claus rally*, a surge in the price of stocks that often occurs in the last five trading days of the year. The extended *season* that we identified could be the result of an anticipating behaviour of investors trying to take advantage of the rise in stock during January. Investigation of this possibility remains open to future research.

Future research should address questions regarding extreme seasonality over non-annual frequencies, such as weekly and intra-daily, and investigate analogous behaviour in other markets, such as exchange rates and commodities. Similarly, from a multivariate perspective it may be interesting to assess whether different stocks present seasonal co-movements,

or, in other words, to study the seasonal dynamics of the extremal dependence.

# Appendix A

## A.1 Change-points algorithm

**Algorithm 1**

**procedure** INITIALIZE
  $\lambda \leftarrow 0$
  $T \leftarrow \mathbf{length}(Z_{id})$
  $\widehat{q}_\tau \leftarrow \mathbf{quantile}(Z_{id}, 0.99)$
  **compute** $H(\phi_1)$
  **if** $H(\phi_1) > G_\alpha$ **then**
    $\widehat{\iota}_0 \leftarrow \arg\max_\iota \left| M(\iota, \widehat{q}_\tau) - \frac{\iota}{T} M(D, \widehat{q}_\tau) \right|$
    $Z_{id} \leftarrow \left\{ \{Z_{id}\}_{d=\widehat{\iota}_0}^{D}, \{Z_{id}\}_{d=1}^{\widehat{\iota}_0 - 1} \right\}$
  **else**
    **stop**
**procedure** MULTIPLE CHANGE POINTS
  $\lambda \leftarrow \frac{T}{12}$
  $I_1 \leftarrow \widehat{\iota}_1$
  **while** $I_m > I_{m-1}$ **do**
    **for** $j$ **in** $I_m$ **do**
      $T \leftarrow \mathbf{length}(\{Z_{id}\}_{\widehat{\iota}_{j-1}}^{\widehat{\iota}_j})$
      $\widehat{q}_\tau \leftarrow \mathbf{quantile}(\{Z_{id}\}_{\widehat{\iota}_{j-1}}^{\widehat{\iota}_j}, 0.99)$
      **compute** $H(\phi_1)$
      **if** $H(\phi_1) > G_{\alpha/m}$ **then**
        $\widehat{\iota}_j \leftarrow \arg\max_{\iota^- \le \iota \le \iota^+} \left| M(\iota, \widehat{q}_\tau) - \frac{\iota}{T} M(D, \widehat{q}_\tau) \right|$
        $I_m \leftarrow \widehat{\iota}_j$
**return** $I_m$

# Chapter 3

# Realizing the extremes: Estimation of tail-risk measures from a high-frequency perspective

## 3.1 Introduction

Accurate assessment of the tail behaviour of asset returns is of the utmost importance for financial market practitioners and regulators. Extreme Value Theory (EVT) is very useful as it provides probabilistic results which characterize the tail behaviour of any distribution, without requiring knowledge of the main body of the distribution.

McNeil and Frey (2000) develop a two-step procedure to model the tails of the conditional returns distribution with EVT: first, the returns are pre-whitened with a GARCH-type model which explicitly accounts for the heteroskedasticity; then the tails of the standardized residuals from the GARCH model are fitted using the Peaks-over-Threshold (POT) method (Davison and Smith, 1990). McNeil and Frey (2000) backtest this approach on different time series and provide evidence that it outperforms both the unconditional EVT model (Danielsson and de Vries, 1997)

and the GARCH models with normal and Student's t distributions. This two-step conditional EVT (C-EVT) approach is now considered standard in the financial community.

In a large simulation experiment, Jalal and Rockinger (2008) study the performance of C-EVT under different hypotheses regarding the underlying data generating process (DGP). They conclude that C-EVT performs fairly well in terms of one-day-ahead predictions of the conditional quantiles under misspecification of the conditional mean and variance dynamics.

In this paper, we develop a *realized* EVT (RV-EVT) approach which exploits high-frequency information to pre-whiten the returns in the first step, and uses the standardized residuals of the high-frequency based model in the second step. Recent work on realized volatility has emphasized how the use of high-frequency information can enhance the forecast of the conditional variance of the returns (Shephard and Sheppard, 2010). We propose a class of high-frequency based volatility models that combines reduced form models for the realized volatility (Corsi, 2009) with a *link function* relating the conditional return volatility with the prediction of the realized volatility. We consider three different *link functions* of increasing complexity and six reduced form models with both symmetric and asymmetric structures.

It is important to filter out the dependence in the first step before applying the POT approach in the second step. We investigate whether a high-frequency based volatility model produces standardized residuals closer to the ideal iid than those obtained under a GARCH-type model. We compare the degree of extremal dependence left in the standardized residuals of our models and the GARCH model for 17 time series of international stock indices from 2000 to 2014.

We then add to the simulation experiment of Jalal and Rockinger (2008) by examining the out-of-sample C-EVT forecasts of both Value-at-Risk (VaR) and Expected Shortfall (ES) for a 10-day horizon[1]. This period is relevant from the regulatory perspective as the risk capital of a bank

[1]Note that with ten-day-ahead risk measures, we mean the risk measures estimated on the conditional distribution of the sum of the next ten-day returns.

must be sufficient to cover losses on the bank's trading portfolio over a 10-day holding period. We also consider an additional DGP where observations are generated according to the parametric model of Bandi and Renò (2015). The latter not only accommodates several stylized facts of the asset returns, but also allows us to draw intra-day observations and produce forecasts of the risk measures with the RV-EVT approach.

Finally, we compare C-EVT and RV-EVT for forecasting VaR and ES on the 17 international indices time series. The backtesting exercise is fully out-of-sample, with a training sample for the models (in-sample) of two different sizes, respectively 2000 and 500 observations. This part of the work is close to that of Giot and Laurent (2004), Clements et al. (2008), and Brownlees and Gallo (2010), in the sense that we assess the merit of using high-frequency data, but we do so within the context of EVT approaches.

The remainder of the chapter is organized as follows: in Section 3.2, we present the C-EVT of McNeil and Frey (2000); in Section 3.3 we introduce the RV-EVT; in Section 3.4 we compare the two approaches, looking separately at the filtering and forecasting components; in Section 3.5 we perform robustness checks aimed at consolidating the evidence from the main analysis; in Section 3.6 we give concluding remarks. Some technical details appear in the appendix and further outcomes can be found in the supplementary material.

## 3.2 The conditional EVT approach

Let $p_t$ be the logarithmic price at time $t$ and define the conditional log-returns $r_t$ as,

$$\begin{aligned} p_t - p_{t-1} = r_t &= \mu_t + \sigma_t \epsilon_t, \\ \mu_t &= f(\mathcal{F}_{t-1}), \\ \sigma_t^2 &= h(\mathcal{H}_{t-1}), \end{aligned} \tag{3.1}$$

where $\mu_t$ and $\sigma_t^2$, are respectively the conditional mean and variance, functions of the information sets $\mathcal{F}_{t-1}$ and $\mathcal{H}_{t-1}$, and $\epsilon_t$ an iid process with zero mean and unit variance. A notable amount of empirical research in financial markets shows time-variation and heaviness in the

tails of the conditional returns distribution. To account for this evidence, Bollerslev (1986) proposes to model $\sigma_t^2$ as a function of its past values and the past values of $\epsilon_t$, assuming $\epsilon_t$ to be normally distributed. This model is commonly referred to as GARCH.

McNeil and Frey (2000) propose to pre-whiten the returns using a standard GARCH and then model the tails of the estimated residuals by means of EVT. The theoretical justification is that Gaussian Quasi-Maximum Likelihood (QML) estimation of a GARCH model is consistent as long as $\mathbb{E}(\epsilon_t^2) < \infty$. Formally, they consider an AR(1)-GARCH(1,1),

$$r_t = \mu + \phi_1 r_{t-1} + \sigma_t \epsilon_t \tag{3.2}$$

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{3.3}$$

where $\epsilon_t \sim F$ with zero mean and unit variance, and $\alpha_1 + \beta_1 < 1$ to guarantee stationarity. Suppose[2] that $F$ has upper endpoint $v_F := \sup\{\epsilon_t : F(\epsilon_t) < 1\}$. Given a high threshold $u$, $u < v_F$, Pickands (1975) shows that when $u \to v_F$, the distribution of the excesses $(\epsilon_t - u)_+$ converges to a Generalized Pareto (GP) distribution G with shape parameter $\xi$ and scale parameter $\nu > 0$. That is, $\Pr(\epsilon_t - u \leq x | \epsilon_t > u)$ goes to

$$G(x; \xi, \nu) = \begin{cases} 1 - \{1 + \xi x/\nu\}^{-\frac{1}{\xi}} & \text{for} \quad \xi \neq 0 \\ 1 - \exp\{-x/\nu\} & \text{for} \quad \xi = 0 \end{cases} \tag{3.4}$$

as $u \to v_F$. When $\xi > 0$, $F$ has Pareto-type upper tail with tail index $1/\xi$.

A model for the tail of the residuals $\widehat{\epsilon}$ can be defined following the POT method. The tail estimator of $F$ is then

$$\widehat{\bar{F}}(\widehat{\epsilon}) = \frac{N_u}{T} \left(1 + \widehat{\xi} \frac{\widehat{\epsilon} - \widehat{u}}{\widehat{\nu}}\right)^{-\frac{1}{\widehat{\xi}}}, \tag{3.5}$$

where $\bar{F} = 1 - F$, $\widehat{u}$ is an appropriately chosen threshold, $\widehat{\xi}$ and $\widehat{\nu}$ are maximum likelihood (ML) estimates, and $N_u$ is the number of observations exceeding the threshold $\widehat{u}$.

[2]The following argument equally applies to $\epsilon_t$ and the negated series $-\epsilon_t$. Throughout we will typically refer to the latter term or its distribution function that we call *loss distribution.*

For $\alpha > 1 - N_u/T$, it is possible to use Equation (3.5) to obtain one-day-ahead prediction of the VaR and ES at level $\alpha$,

$$\widehat{VaR}^{\alpha}_{t,t+1} = \widehat{\mu}_{t,t+1} + \widehat{\sigma}_{t,t+1}\left(\widehat{u} + \frac{\widehat{\nu}}{\widehat{\xi}}\left(\left(\frac{1-\alpha}{N_u/S}\right)^{-\widehat{\xi}} - 1\right)\right), \tag{3.6}$$

$$\widehat{ES}^{\alpha}_{t,t+1} = \widehat{\mu}_{t,t+1} + \widehat{\sigma}_{t,t+1}\widehat{\epsilon}_{\alpha}\left(\frac{1}{1+\widehat{\xi}} + \frac{\widehat{\nu} - \widehat{\xi}\widehat{u}}{(1-\widehat{\xi})\widehat{\epsilon}_{\alpha}}\right), \tag{3.7}$$

where $\widehat{\mu}_{t,t+1}$ and $\widehat{\sigma}_{t,t+1}$ are the forecast of the conditional mean and the conditional variance, and $\widehat{\epsilon}_{\alpha}$ is the $(1-\alpha)$quantile of the residuals at time $t$.

Financial decisions are often predicated on accurate multi-period-ahead forecasts of the risk measures, even though the risk management literature surprisingly focuses almost exclusively on the accuracy of one-period-ahead forecasts. The dominant long-horizon forecasting approach consists of scaling the one-period-ahead forecasts by $\sqrt{k}$ where $k$ is the horizon of interest. Its popularity among practitioners stems mostly from its use in RiskMetrics. Alternatively, we follow the iterative approach of McNeil and Frey (2000) to obtain the regulatory relevant ten-day-ahead VaR and ES forecasts, as they were shown to outperform the $\sqrt{k}$ rule. We fit a GP distribution to both tails of the GARCH residuals and model the body by means of the empirical distribution function. Then, combining bootstrapping and GP simulation it is possible to obtain a semi-parametric estimate of the innovation distribution:

$$\widehat{F_{\epsilon}(\epsilon)} = \begin{cases} \frac{N_u^L}{S}\left(1 + \xi^L \frac{|\epsilon - \widehat{u}^L|}{\nu^L}\right)^{-1/\widehat{\xi}^L} & \epsilon < \widehat{u}^L \\ \frac{1}{S}\sum_{i=1}^{S} 1_{\epsilon_i \leq \epsilon} & \widehat{u}^L \leq \epsilon \leq \widehat{u}^H \\ 1 - \frac{N_u^H}{S}\left(1 + \xi^H \frac{|\epsilon - \widehat{u}^H|}{\nu^H}\right)^{-1/\widehat{\xi}^H} & \epsilon > \widehat{u}^H \end{cases} \tag{3.8}$$

where $S$ is the sample size, $N_u^i$ is the number of exceedances, $\widehat{\xi}^i$ and $\widehat{\nu}^i$ are GP parameter estimates, $\widehat{u}^i$ is a high threshold and $i = L, H$ indicate respectively the lower and upper tails. Through an iterative algorithm which consists of drawing observations from this distribution, see Appendix B.1, and updating the GARCH equation, it is possible to simulate several sample paths for the next ten days and obtain an estimate

of the multi-period distribution. Throughout the analysis we generate 1000 sample paths. The ten-day-ahead VaR is obtained from the inversion of Equation (3.5) and ES as conditional expectation of exceedances over VaR.

## 3.3 The realized EVT approach

Our proposal is to incorporate high frequency (intra-daily) information in the sets $\mathcal{F}_t$ and $\mathcal{H}_t$ of Equation (3.1), and use the POT approach to model the tails of the residuals. To the extent that high-frequency based volatility models provide better forecasts, they should produce residuals with a lower degree of extremal dependence and allow for a better estimation of the tail. As the POT approach is based on the assumption of iid observations, residuals $\widehat{\epsilon}_t$ that better proxy the innovations $\epsilon_t$ improve the inference on the GP parameters and the estimation of the probability of exceedances. At the same time, improved volatility predictions provide better forecasts of the conditional density and more accurate estimates of tail-related risk measures.

Assuming $\mu_t$ to be constant in Equation (3.1) (its effect is negligible), we specify a class of high-frequency based volatility models for the latent variable $\sigma_t^2$, using as a proxy the realized volatility (RV). The latter is a non-parametric estimator of the variation of the price path of an asset obtained as the sum of squares of intra-day returns recorded during the open hours of the stock exchange. Formally, let $p_t$ denote the log-price of an asset at time $t$ and $r_{t,\Delta} = p_t - p_{t-\Delta}$ the discretely sampled $\Delta$-period return. The RV on day $t$ is then defined as

$$RV_t = \sum_{j=1}^{N} r_{t-1+(j\cdot\Delta)}^2. \tag{3.9}$$

If market microstructure noise is absent then, as $\Delta \to 0$, $RV_t$ consistently estimates the quadratic variation of the price process on day $t$ (Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2002). In practice, market microstructure noise plays an important role, and econometricians usually resort to 1- to 5-minute return data to mitigate the effect of the noise.

Let $\mathbb{E}(RV_t|\mathcal{H}_{t-1})$ be the conditional expectation of the RV given the information set $\mathcal{H}_{t-1}$ and assume that $\sigma_t^2$ is defined as a function of $\mathbb{E}(RV_t|\mathcal{H}_{t-1})$, $\sigma_t^2 = g(\mathbb{E}(RV_t|\mathcal{H}_{t-1}))$. We need to specify a model for the conditional expectation of the RV and a *link* function $g$ that maps the conditional mean of the RV to the conditional variance of the returns. In what follows, we discuss these two issues separately and then show how to obtain forecasts of the risk measures.

### 3.3.1 The link function

Different approaches have been attempted to provide a connection between the conditional variance of the returns and RV but, to the best of our knowledge, a comparative study of these *link* functions has not been pursued so far. In this paper, we consider and discuss three different *link* functions, those of Giot and Laurent (2004), Clements et al. (2008) and Brownlees and Gallo (2010), and compare their performances.

The simplest approach is the one used in Clements et al. (2008), where the conditional variance of returns is defined by

$$\widehat{\sigma}^2_{t,t+h} = \exp\left( \widehat{\log(RV_{t,t+h})} \right). \tag{3.10}$$

We call this *link* function **type-I**. This specification bears two different sources of bias: one related to the log-transformation and the other one due to the imprecision of the volatility proxy. Specifically, given that returns are measured close-to-close while intra-daily observations neglect overnight information, this information mismatch may be a source of bias (Andersen et al., 2011). To circumvent this problem, Brownlees and Gallo (2010) assume the conditional variance $\sigma_t^2$ to be an affine function of the RV. This model is nested in the HEAVY class of Shephard and Sheppard (2010). Our logarithmic specification of the HAR function prohibits us from applying the idea directly, but we adapt it and call it **type-II**. We have

$$\widehat{\sigma}^2_{t,t+h} = c + d\exp\left( \widehat{\log(RV_{t,t+h})} \right), \tag{3.11}$$

where $c$ and $d$ are coefficients to be estimated. Finally, similarly to Giot and Laurent (2004), we incorporate a correction for the logarithmic trans-

formation and call it **type-III**. We have

$$\widehat{\sigma}^2_{t,t+h} = c + d \exp\left( \widehat{\log(RV_{t,t+h})} + 0.5\widehat{\sigma}^2_{\widehat{\eta}} \right), \tag{3.12}$$

where $\widehat{\sigma}^2_{\widehat{\eta}}$ is the estimated variance of the logarithmic HAR regression residuals.

A word of caution is in order as concerns models (3.10), (3.11) and (3.12), because they are reduced-form models that are misspecified in the presence of jumps and/or stochastic volatility, and this may have severe consequences on the estimation procedures; see, e.g., Nelson (1992).

### 3.3.2 Reduced form models

Several approaches have been proposed in the financial econometrics literature to model the dynamics of $\mathbb{E}(RV_t|\mathcal{H}_{t-1})$, see Andersen et al. (2003), Bollerslev et al. (2009), and Andersen et al. (2011). In this paper, we rely on the heterogeneous autoregressive (HAR) class of models, initially proposed by Corsi (2009), and now considered as the standard approach in the high-frequency literature. In particular, we use the logarithmic specification of these models for two specific reasons: constraints on the parameters to guarantee positive volatility are not necessary; the estimated residuals of the logarithmic specification are closer to normality, hence more amenable to standard time series procedures.

Let the multi-period normalized RV be denoted by

$$RV_{t,t+h} = h^{-1}[RV_{t+1} + \cdots + RV_{t+h}]. \tag{3.13}$$

Note that $RV_{t,t+1} \equiv RV_{t+1}$. The daily HAR model of Corsi (2009) can then be expressed as

$$\log(RV_{t,t+h}) = \beta_0 + \beta_D \log(RV_t) + \beta_W \log(RV_{t-5,t}) + \beta_M \log(RV_{t-22,t}) + \eta_{t,t+h}, \tag{3.14}$$

where $t = 1, \ldots, T$. This parsimonious specification exploits the information from the past 1-day, 5-day and 22-day average RV, reflecting the idea that heterogeneity in investor behaviour creates different volatility components having different impacts on the future volatility.

The intuition that the disentanglement of the continuous and discontinuous sample paths could be a valuable source of information leads Andersen et al. (2007) to extend the HAR model to include a jump component. This is important also in the light of the findings of Todorov and Tauchen (2011) and Bandi and Renò (2015) that volatility jumps are relevant. Adding the realized jump measure $J_t$, obtained as the difference between the realized variance and the bipower variation (Barndorff-Nielsen and Shephard, 2004), as an explanatory variable, one obtains their HAR-J model

$$\begin{aligned}\log(RV_{t,t+h}) = \quad & \beta_0 + \beta_D \log(RV_t) + \beta_W \log(RV_{t-5,t}) + \beta_M \log(RV_{t-22,t}) \\ & + \beta_J \log(J_t + 1) + \eta_{t,t+h}.\end{aligned} \tag{3.15}$$

The jump component usually adds a significant contribution in terms of $R^2$, see Andersen et al. (2007) and Corsi et al. (2010).

The models in Equations (3.14)-(3.15) are both symmetric in the sense that they do not have a parameter accounting for the so called leverage effect, i.e. the generally negative correlation between an asset's returns and its changes of volatility (Nelson, 1991). To generate the desired leverage effect, a first possibility is to use the signed jump variation obtained from the realized semi-variance (Patton and Sheppard, 2015). By isolating the sign of the jump, we can differentiate its impact on RV. Letting $RS_t^+$ and $RS_t^-$ be the positive and negative realized semi-variance respectively, the HAR-SJ model has specification

$$\begin{aligned}\log\left(RV_{t,t+h}\right) = \quad & \beta_0 + \beta_{J+} \log(\Delta J_t^{2+} + 1) + \beta_{J-} \log(\Delta J_t^{2-} + 1) \\ & + \beta_D \log(RV_t) + \beta_W \log(RV_{t-5,t}) \\ & + \beta_M \log(RV_{t-22,t}) + \eta_{t,t+h},\end{aligned} \tag{3.16}$$

with signed jump variations

$$\begin{aligned}\Delta J_t^{2+} &= (RS_t^+ - RS_t^-) I_{\{(RS_t^+ - RS_t^-) > 0\}}, \\ \Delta J_t^{2-} &= (RS_t^+ - RS_t^-) I_{\{(RS_t^+ - RS_t^-) < 0\}}.\end{aligned}$$

Assuming that jumps are rare, these two quantities broadly capture the variation of positive and negative jumps. Nonetheless, if jumps of different signs occur on the same day, then these measures account only for a

small part of the total jump variation. In this case, the HAR-SJ can miss valuable information coming from the actual size of the jumps. To overcome this issue, we propose a simple extension of this model and refer to it as the HAR-SJ*aug* model, see Appendix B.2 for a discussion.

Another asymmetric model is the LHAR of Corsi and Renò (2012) that adds to the simple HAR specification three leverage components capturing the persistence of the leverage effect. In contrast to Patton and Sheppard (2015), they use the lagged negative returns to generate the desired effect,

$$\begin{aligned}\log(RV_{t,t+h}) \quad &= \beta_0 + \beta_D \log(RV_t) + \beta_W \log(RV_{t-5,t}) + \beta_M \log(RV_{t-22,t}) \\ &+\gamma_d r_t^- + +\gamma_w r_{t-5,t}^- + \gamma_m r_{t-22,t}^- + \eta_t\end{aligned} \tag{3.17}$$

where $r_{t-h,t} = \frac{1}{h}\sum_{j=1}^{h} r_{t-j}$, and $r_{t-h,t}^- = \min(r_{t-h,t}, 0)$.

### 3.3.3 Forecasting risk measures

Once we have a high-frequency based model for the conditional variance, we can filter the returns and model the tails of the residuals with the estimator in Equation (3.5). Whereas to obtain one-day-ahead predictions for C-EVT we rely on a closed-form approach (see Equations (3.6)-(3.7)), to obtain multiple-day-ahead predictions we have to use an iterative approach based on Equation (3.8) since the multiple-day conditional distribution of a daily GARCH is not available. We can proceed differently for RV-EVT. Considering that it is standard in the financial econometrics literature to fit the HAR model directly to the multi-period RV, we can compute explicitly the forecasts at both horizons for the RV-EVT. The $h$-day-ahead predictions of the VaR and ES at level $\alpha$ are then defined as

$$\widehat{VaR}_{t,t+h}^{\alpha} = \widehat{\mu}_{t,t+h} + \widehat{\sigma}_{t,t+h}\left(\widehat{u} + \frac{\widehat{\nu}}{\widehat{\xi}}\left(\left(\frac{1-\alpha}{N_u/T}\right)^{-\widehat{\xi}} - 1\right)\right) \tag{3.18}$$

$$\widehat{ES}_{t,t+h}^{\alpha} = \widehat{\mu}_{t,t+h} + \widehat{\sigma}_{t,t+h}\widehat{\epsilon}_{\alpha}\left(\frac{1}{1+\widehat{\xi}} + \frac{\widehat{\nu} - \widehat{\xi}\widehat{u}}{(1-\widehat{\xi})\widehat{\epsilon}_{\alpha}}\right) \tag{3.19}$$

where $\widehat{\mu}_{t,t+h}$ and $\widehat{\sigma}_{t,t+h}$ are respectively the $h$-period forecasts of the conditional mean and variance, and $\widehat{\epsilon}_\alpha$ is the $(1-\alpha)$ quantile of the residuals at time $t-1$.

## 3.4 Conditional EVT vs. Realized EVT

### 3.4.1 Data

The empirical analysis is based on the Oxford-Man Institute "Realised Library" version 0.2, Heber et al. (2009). We consider 17 different stock indices from the beginning of 2000 to the end of 2014, see Table 6. For each asset, the library currently records the daily returns and several daily realized measures. Among the latter, we consider the 5-min RV and use the 5-min Bipower Variation (BV) and the 5-min RSV to extract the jump components. The 5-min sampling frequency is standard in the literature as it mitigates the microstructure noise and assures good performance of the estimators (Liu et al., 2015). To control whether the microstructure noise is still of any concern, we repeat the analysis with the *Second Best* estimator of Zhang et al. (2005) and the Realized Kernel of Barndorff-Nielsen et al. (2008) instead of the RV estimator, see Section 3.5.2.

**Table 6: Data description**. Time series of indices start January 2, 2000 and end December 31, 2014. Stock exchanges respect different holidays and the number of observations $T$ subsequently differs.

| Asset | Abbr. | T | Asset | Abbr. | T |
|---|---|---|---|---|---|
| Amsterdam Exchange Index | AEX | 3816 | IBEX35 | IBX | 3782 |
| All Ordinaries Index | AOI | 3743 | IPC Mexico | IPC | 3748 |
| Bovespa Index | BVP | 3664 | Korea Composite Index | KCI | 3690 |
| CAC40 | CAC | 3817 | Nasdaq 100 | NSQ | 3747 |
| DAX30 | DAX | 3795 | Nikkei 225 | NK | 3630 |
| Dow Jones Industrial | DJ | 3746 | Russel 2000 Index | RUS | 3745 |
| Euro Stoxx 50 | ESX | 3794 | SP500 | SPX | 3744 |
| FTSE MIB | MIB | 3778 | Swiss Market Index | SMI | 3749 |
| FTSE100 | FT | 3764 | | | |

### 3.4.2 Methodology

Given the long series in Table 6 and the out-of-sample nature of our investigation, a rolling-window approach is used to obtain a time series of performance measures to compare the C-EVT and the RV-EVT. In particular, we use a fixed rolling-window length of $S = 2000$ days to train our models (in-sample) and daily updates.

We consider the exceedances over the threshold $\widehat{u}$ corresponding to the $95$th quantile of the negated residuals $-\widehat{\epsilon}_t$. We perform Anderson-Darling and Cramér-von Mises goodness-of-fit tests at a level of significance of $0.05$ to check the validity of the GP convergence (Choulakian and Stephens, 2001). Recall from Pickands (1975) that convergence of the distribution of the excesses to the GP distribution occurs when the threshold $u$ goes to the right endpoint of the support. Selecting a high empirical quantile as a threshold, we are using the results pre-asymptotically as an approximation only. One can thus expect the quality of the approximation to differ across the stock indices. When it is poor according to the goodness-of-fit tests, we raise the threshold $u$ to the $98$th quantile. These cases are flagged with the * symbol in the figures and the tables.

We assume the conditional mean of all models to be constant and equal to zero as in Brownlees and Gallo (2010). In the C-EVT, we use a GJR-GARCH(1,1) specification instead of the GARCH(1,1) to account for the evidence of leverage effect in the financial returns.

We compare our EVT-based methods along two separate but related dimensions: the **filtering** and **forecasting** components. On the one hand, we assess how much of the extremal dependence inherent in the returns is captured by the different models. To this end, we rely on the concept of extremal index $\theta$ (Leadbetter, 1983) which is a parameter measuring this dependence. As for forecasts, we produce one-day- and ten-day-ahead predictions of the VaR and ES at level $\alpha = 0.01$ and use standard performance measures to evaluate them, see Christoffersen (1998) and Engle and Manganelli (2004) for the VaR and McNeil and Frey (2000) for the ES.

### 3.4.3 The filtering component

Financial extremes tend to occur in clusters, and the degree of dependence in the extremes of a stationary series is measured by the extremal index $\theta$. In particular, when $\theta < 1$ the extremes are dependent, conversely with $\theta = 1$ they are independent. The first step of C-EVT or RV-EVT should filter out this dependence so that the residuals used in the respective second steps are iid. If they are not, both the inference on the GP parameters and the probability of threshold exceedances will be affected. Given our finite sample size, we consider the standardized residuals of the models to be close to independent when $0.85 \leq \widehat{\theta} \leq 1$, where $\widehat{\theta}$ is the Ferro and Segers (2003) intervals estimate.

In what follows, we compare the performance of the GJR-GARCH(1,1) and the HAR models of Section 3.3.2 as filters. In particular, we estimate the extremal index $\widehat{\theta}$ of the model residuals obtained over each window. We repeat the analysis with the different *link* functions in Equations (3.10)-(3.12) to assess the merit of the complexity added to the high-frequency based model for the conditional volatility.

Figure 30 displays the estimates of $\theta$ for the type-I *link* function. The figure clearly shows that, in general, the residuals of each model exhibit a very low degree of extremal dependence and that the GJR-GARCH(1,1) model appears as the best performer. On a few occasions, the high-frequency based volatility models fail to capture the extremal dependence, particularly on the RUSSEL time series. Adding jumps and leverage does not seem to contribute substantially.

In Figures 31-32, we report the estimates of $\theta$ respectively for the type-II and type-III *link* functions. The figure suggests that there is no real benefit from increasing the complexity of the conditional volatility model. The degree of extremal dependence left in the residuals by the high-frequency based models is in line with that observed in Figure 30 and GJR-GARCH(1,1) is still the best choice.

Although the results are compelling, they might seem counterintuitive. Indeed, to the extent that a model provides superior forecasts of the volatility, it should also provide a better proxy of the innovations.

We argue that the observed difference in the estimated residuals might be explained by the different ability of the two approaches in generating extremal dependence.

Most of the evidence on the higher performance of the high-frequency based volatility models is based on comparisons of the $R^2$ statistic or other loss functions. But this does not say anything regarding the extremal dependence. The GARCH class of models is supported by theoretical arguments which show their ability to generate extremal dependence in the volatility process and consequently in the return process (Basrak et al., 2002). The HAR class of models is a simple equation that fits nicely the stylized facts of the financial time series and is supported by economic arguments regarding the heterogeneous behaviour of different investors. It is difficult (if possible) to derive a theoretical result on the extremal behaviour of this model as it is for GARCH processes.

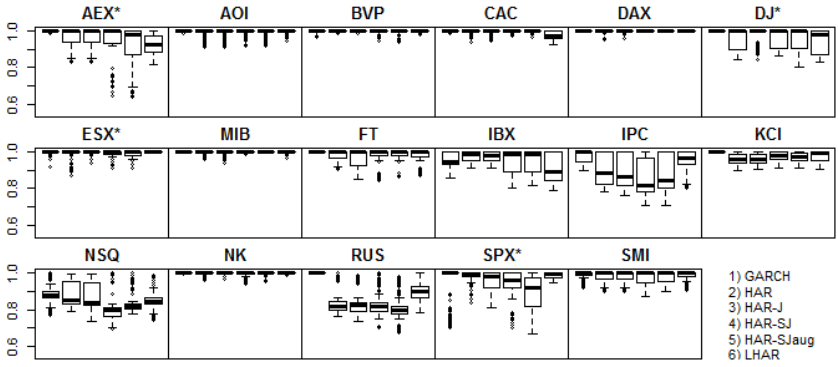
In order to get some insights on this issue, we can build upon the comparison between GARCH processes and stochastic volatility (SV) models from an extreme perspective. While GARCH processes exhibit extremal dependence, SV may or may not fail to account for this property depending on the specification (Davis and Mikosch, 2009b; Fasen et al., 2006). As the HAR class of models can be considered in the domain of SV models (Corsi, 2009), then it may either capture or not the extremal dependence in the returns.

To assess this possibility, we follow Liu and Tawn (2013) and rely on the conditional tail probability $c(x) = \mathbb{P}(X_{t+1} > q | X_t > q)$. Theoretically, $c(x) \to 0$ as $q \to \infty$ when the extremes are independent and $c(x) \to c > 0$ as $q \to \infty$ when they are dependent. As it is not possible to study the asymptotic behaviour of $c(x)$ for $q \to \infty$, we approximate it considering a sequence of high quantiles $q$. Figure 33 displays the empirical conditional tail probability at different levels $q$ for three sub-samples of the S&P500. Moreover, we report the same coefficients estimated on simulated data produced by the GJR-GARCH(1,1) and the HAR models fitted to these three samples. As expected, the tail probability implied by the GARCH model is positive at every quantile pointing toward the presence of extremal dependence. The conditional probability for the HAR

model does not reach zero at the considered levels, but it is clear that its decay is much faster than the one of the GJR-GARCH model, and it fails to capture the empirical tail dependence.

Based on this evidence, it is worth investigating whether the time series for which the HAR provides a good filter exhibit a lower degree of extremal dependence. A positive response would be consistent with the idea that GARCH and HAR models imply two different degrees of dependence in the extremes. Figure 34 displays the estimates of $\theta$ obtained on the return series of the 17 stock indexes in the different windows. The results are somewhat blurred: some cases are consistent with the findings in Figure 30 and some others are not. For instance, a high degree of dependence is shown by the AEX and the SPX and a low degree of dependence is shown for the BVP and the NK, which is in line with the dependence left in the residuals of the HAR models. But at the same time, IPC and RUS present lower dependence than other series, yet HAR models could not filter to iid.

**Figure 30: Extremal index estimates for the type-I *link* function**. Values of $\theta$ estimated on the standardized residuals obtained from the different windows of length $S = 2000$.




### 3.4.4 The forecast component

Jalal and Rockinger (2008) carry out an extensive simulation experiment

**Figure 31: Extremal index estimates for the type-II *link* function**. Values of $\theta$ estimated on the standardized residuals obtained from the different windows of length $S = 2000$.

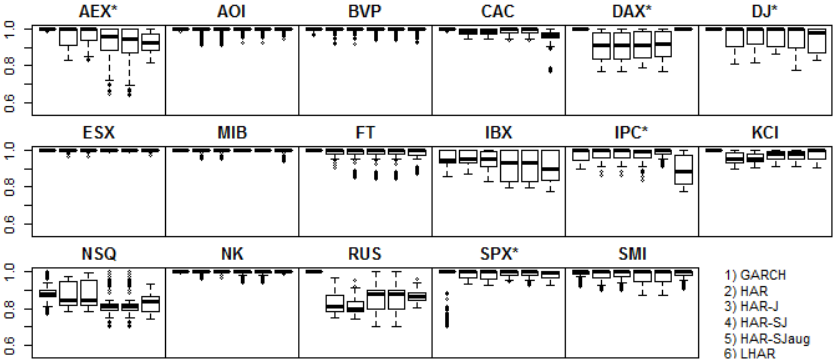



**Figure 32: Extremal index estimates for the type-III *link* function**. Values of $\theta$ estimated on the standardized residuals obtained from the different windows of length $S = 2000$.

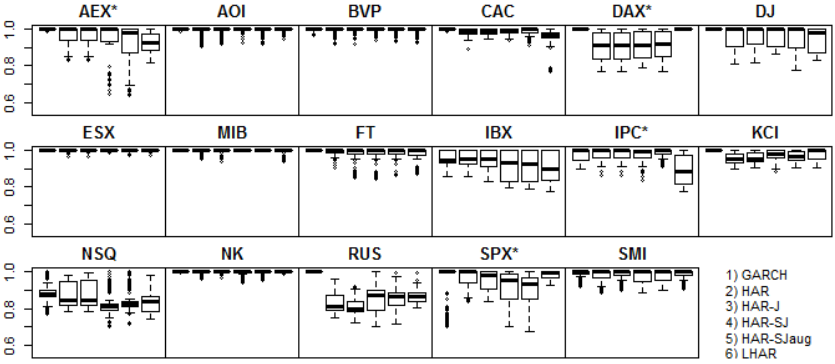


**Figure 33: Conditional tail probaility**. Estimates of $c(x)$ for the lower tail of three sub-samples of the S&P500 at 10 different quantiles: Empirical (*points*); simulation from the GJR-GARCH(1,1) model (*solid*) and tolerance levels (*dotted*); simulation from the HAR model (*dashed*) and tolerance levels (*dash-dotted*).

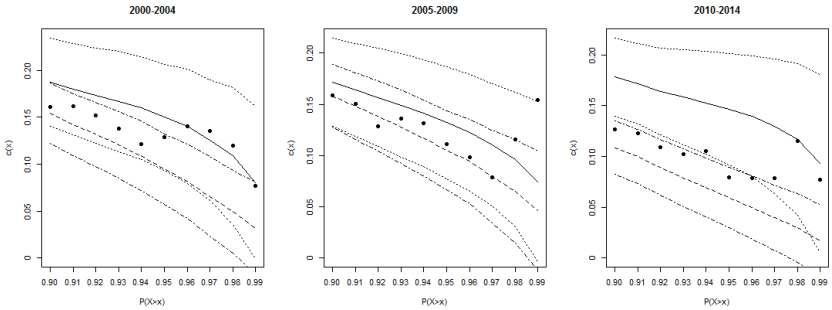
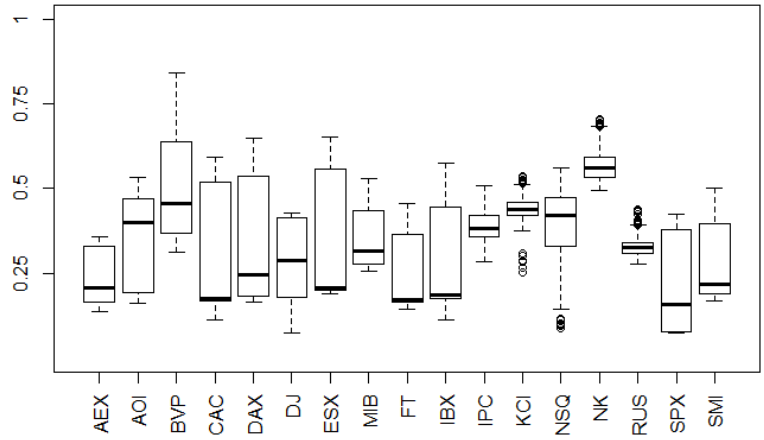



**Figure 34: Extremal index on the raw data**. Estimates of $\theta$ on the lower tail of the return series over moving windows of length $S = 2000$.



to investigate the performance of C-EVT when the DGP is not the standard GARCH(1,1) model of Bollerslev (1986). Their results emphasize how C-EVT provides accurate one-day-ahead forecasts of the VaR and ES regardless of the assumed DGP, thus strengthening the conclusions of McNeil and Frey (2000).

We suspect that the performance of the C-EVT deteriorates as estimation moves to longer time horizons. The quality of the estimation of the tail-risk measures depends on the quality of the volatility forecast. As it is shown that the volatility forecasting performance of methods based on daily returns deteriorates with increasing horizon length (Brownlees et al., 2011), similar deterioration should be seen in tail-risk forecasts.

We perform both a simulation experiment and an empirical data analysis. In both cases, we provide out-of-sample forecasts of the one-day- and ten-day-ahead VaR and ES at level $\alpha = 0.01$.

**Simulation experiment.**

First, we reproduce the analysis of Jalal and Rockinger (2008). We consider six different DGP: the GARCH(1,1) with Gaussian and $t$ innovations; the EGARCH(1,1) of Nelson (1991); the regime-switching model of Schaller and Norden (1997); the discrete-time jump diffusion model of Pan (1997) and its pure jump version. For each process, we generate $B = 100$ samples of $T = 4000$ daily returns. Figure 35 shows the results for both the one-day- and ten-day-ahead forecasts. To evaluate the performance of the VaR, we count the violations in each window. When the length of the widow is $S = 2000$, the expected number of violations is $(T - S) \times \alpha = 20$. As a performance measure for the ES, we compute its distance from the actual return when the VaR is exceeded, and expect this to be mean zero and symmetrically distributed.

Although we do not report any formal test, the results for the one-day-ahead predictions in Figure 35 are consistent with those of Jalal and Rockinger (2008). However, focussing particularly on the VaR, it is evident that the accuracy of the C-EVT is reduced at the longer time horizon. Indeed, for three out of six DGP the average number of violations moves further away from the expected number.

Can RV-EVT provide more accurate predictions? We set up another simulation experiment where the two approaches can be compared. We need to generate high-frequency observations, so we consider the stochastic volatility model of Bandi and Renò (2015). The latter involves 21 parameters and accommodates several attributes of financial time series. It accounts for time-variation in both the volatility and leverage, and allows for idiosyncratic and common jump discontinuities in the price process and the volatility, determined from a complex dependence structure.

We generate $B = 100$ samples of $T = 4000$ daily observations for both the returns and the 5-min realized measures. We use the Euler scheme to generate $N = 2520$ equally-spaced observations per day, assuming 7 hours of trading and changes in the price level occurring every ten seconds, leading to 6 ticks per minute $(6 \times 60 \times 7 = 2520)$. This discretization method is standard in the literature (Huang and Tauchen, 2005), while the choice of the frequency is mainly for computational reasons. The parameters of the model are set according to the GMM estimates obtained on the S&P500 futures from April 1982 to February 2009 in Bandi and Renò (2015).

In Figure 36, we report the one-day- and ten-day-ahead VaR and ES forecasts obtained over each window. We consider only a Type-I *link* function for RV-EVT. Consistent with the results in Figure 35, the C-EVT performs well for one-day-ahead forecasts but deteriorates as we extend the time horizon. Contrarily, the RV-EVT approach seems to be less affected by the increasing time horizon and performs well regardless of the high-frequency based volatility model for the RV. In particular, the LHAR model emerges as the best one since it presents a smaller range of variability in terms of number of violations, and it is less upward biased in the estimation of the ES.

**Empirical analysis.**

We now consider empirical data and assess whether the results suggested by the simulation experiment are confirmed.

In Table 7, we compare one-day-ahead VaR forecasts from the C-EVT and RV-EVT for the 17 indices. To evaluate the accuracy of the predic-

**Figure 35: Extending experiments in Jalal and Rockinger (2008)**. One- and ten-day-ahead predictions at level $\alpha = 1\%$. VaR violations (left panels) and actual returns minus ES when VaR is exceeded (right panels). Expected number of violations is 20. ES differences should be mean zero and symmetrically distributed. DGP are: Gaussian-GARCH(1,1) (*a*); Student's-GARCH(1,1) (*b*); EGARCH(1,1) (*c*); Switching-Markov (*d*); Jump-diffusion (*e*); Pure jumps (*f*).

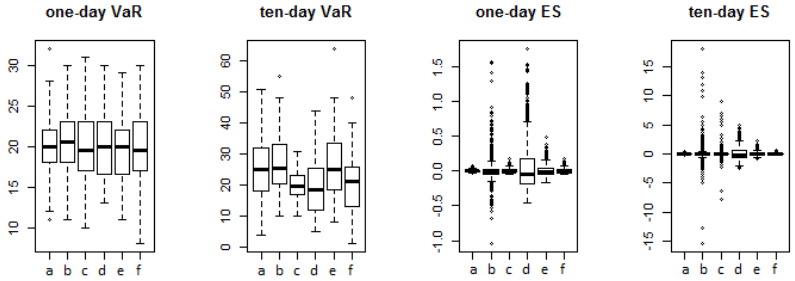
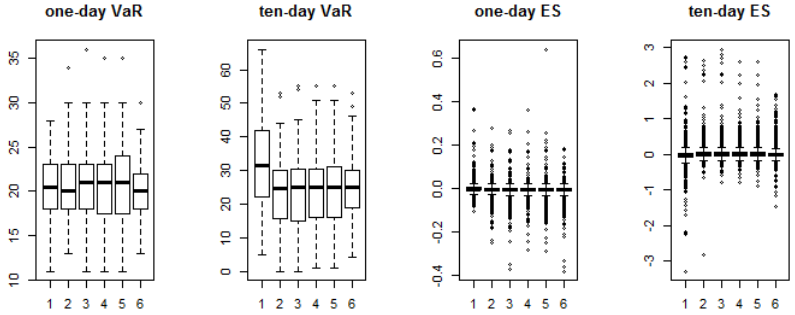



**Figure 36: Comparison of C-EVT and RV-EVT for simulated data following Bandi and Renó (2014) model**. One- and ten-day-ahead predictions at level $\alpha = 1\%$. VaR violations (left panels) and actual returns minus ES when VaR is exceeded (right panels). Expected number of violations is 20. ES differences should be mean zero and symmetrically distributed. Models used in first step: (*C-EVT*) 1 - GARCH ; (*Symmetric RV-EVT*) 2 - HAR; 3 - HAR-J; (*Asymmetric RV-EVT*) 4 - HAR-SJ; 5 - HAR-SJaug; 6 - LHAR.



tions, we use the standard tests of unconditional coverage (UC), independence assumption (IN) and conditional coverage (CC) suggested by Christoffersen (1998), and the dynamic quantile test (DQ) of Engle and Manganelli (2004). These results suggest that the C-EVT and the RV-EVT perform equally well for one-day-ahead. In most of the cases, the number of violations for both approaches is very close to the expected, and the null hypothesis of independence is typically not rejected at the standard levels. Noteworthy is the DQ statistic, reflecting the predictability of a violation given the past information. The null hypothesis of independence between consecutive violations appears to be the one most frequently rejected. In sum, these results are in line with those obtained by Giot and Laurent (2004) and Brownlees and Gallo (2010) using non EVT-based modelling assumptions.

To compare the one-day-ahead ES predictions, we test the hypothesis that conditional upon exceeding the 99th quantile of the loss distribution, the difference between the actual return and the predicted ES has mean zero. We conduct a one-sided test with the alternative that the mean is greater than zero using a bootstrap that makes no assumption about the distribution of the differences, see Section 16.4 of Efron and Tibshirani (1994). The small number of p-values below 0.05 in Table 8 indicate that the good results for the VaR extend to ES.

Turning to the ten-day-ahead predictions of these tail-risk measures, hypothesis testing is complicated by the fact that the violations are inherently autocorrelated. Nonetheless, we provide two simple statistical tests for the VaR and the ES. First, let $I^{\alpha}_{t,t+10}$ take value one when the ten-day-ahead VaR prediction at level $\alpha$ is violated, and zero otherwise. $\{I_{t,t+10}, t = 1, 2, \dots\}$ is a sequence of correlated Bernoulli random variables. Letting $\Phi(\cdot)$ be a Probit link function, we assume that $I_{t,t+10} = \Phi(\gamma)$ and test the hypothesis of correct unconditional coverage $\mathcal{H}_0 : \gamma = \Phi^{-1}(\alpha)$ against the alternative $\mathcal{H}_1 = \text{not } \mathcal{H}_0$. We perform a simple Wald test with robust standard errors computed with the Newey-West estimator to account for the autocorrelations. Table 9 shows the number of violations of each model along with the p-value of the test. The results suggest that at longer time horizons the RV-EVT approach

is more accurate than C-EVT: in three cases the null of correct unconditional coverage is rejected for C-EVT, and most of the time the RV-EVT is more accurate in terms of number of violations.

To compare the ten-day-ahead ES predictions, we employ again the bootstrap test used for the one-day-ahead ES predictions, but we base it on a block bootstrap because of the autocorrelations in the violations. In Table 10 we report the p-values of the test. Contrasts are not as clear as in the case of the VaR as the number of rejections of the null hypothesis of mean zero difference between the ten-day-ahead ES forecasts and the actual ten-day returns is very similar in all cases.

Results obtained with the Type-II and Type-III *link* functions and reported in Appendix B.3 lead to the same conclusions.

## 3.5 Further analysis

The results from the preceding sections highlight the ability of the GARCH model as a filter, and the merits of the high-frequency based volatility models for the forecast of the tail-risk measures. Now, as standard in finance, we verify that the results obtained are robust to the choices made.

### 3.5.1 An investigation during a period of turmoil

We evaluate the performance of C-EVT and RV-EVT when the training sample used to fit the models is limited to $500$ days instead of $2000$. Furthermore, we consider the time period from the beginning of 2007 to the end of 2011.

Given the reduced sample size, we consider the exceedances over the threshold $\widehat{u}$ corresponding to the $90$th quantile of the negated residuals $-\widehat{\epsilon}_t$ of the considered model. Goodness-of-fit tests are carried out as described in Section 3.4.2. When the threshold seems to be still at the sub-asymptotic level, it is raised to the $92$nd quantile. One-day- and ten-day-ahead forecasts of the VaR and the ES are produced at level $\alpha = 0.05$.

Results available in Appendix B.4 confirm the findings of Section 3.4, although the difference are less pronounced when looking only at the

**Table 7: Performance measures for the one-day-ahead VaR forecast**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at the level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behaviour.

| Index | PM | (a) | (b) | (c) | (d) | (e) | (f) | Index | PM | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX* | NV | 22 | 18 | 18 | 19 | 19 | 17 | IBX | NV | 26 | 19 | 19 | 20 | 21 | 18 |
| (18) | UC | 0.36 | 1 | 1 | 0.81 | 0.81 | 0.81 | (18) | UC | 0.07 | 0.81 | 0.81 | 0.64 | 0.49 | 1 |
| | IN | 0.46 | 0.55 | 0.55 | 0.52 | 0.52 | 0.57 | | IN | 0.38 | 0.20 | 0.20 | 0.22 | 0.25 | 0.54 |
| | CC | 0.50 | 0.83 | 0.83 | 0.79 | 0.79 | 0.82 | | CC | 0.14 | 0.42 | 0.42 | 0.42 | 0.40 | 0.83 |
| | DQ | 0.65 | 0.72 | 0.72 | 0.75 | 0.75 | 0.99 | | DQ | **0.003** | 0.19 | 0.18 | **0.001** | **0.002** | 0.54 |
| AOI | NV | 16 | 13 | 15 | 16 | 17 | 11 | IPC | NV | 13 | 7 | 7 | 7 | 9 | 7 |
| (17) | UC | 0.80 | 0.30 | 0.62 | 0.81 | 1 | 0.12 | (17) | UC | 0.31 | **0.005** | **0.005** | **0.005** | **0.03** | **0.005** |
| | IN | 0.58 | 0.66 | 0.61 | 0.59 | 0.56 | 0.71 | | IN | 0.66 | 0.81 | 0.81 | 0.81 | 0.76 | 0.81 |
| | CC | 0.82 | 0.54 | 0.77 | 0.83 | 0.84 | 0.27 | | CC | 0.54 | **0.02** | **0.02** | **0.02** | 0.10 | **0.02** |
| | DQ | 0.99 | 0.34 | 0.51 | 0.61 | 0.63 | 0.12 | | DQ | 0.27 | 0.46 | 0.49 | 0.49 | 0.75 | 0.47 |
| BVP | NV | 13 | 12 | 12 | 14 | 14 | 13 | KCI | NV | 13 | 20 | 20 | 19 | 21 | 17 |
| (17) | UC | 0.31 | 0.30 | 0.20 | 0.45 | 0.45 | 0.31 | (17) | UC | 0.31 | 0.48 | 0.48 | 0.63 | 0.35 | 1 |
| | IN | 0.65 | 0.66 | 0.68 | 0.63 | 0.63 | 0.65 | | IN | 0.65 | 0.23 | 0.23 | 0.21 | 0.26 | 0.56 |
| | CC | 0.53 | 0.54 | 0.40 | 0.66 | 0.66 | 0.53 | | CC | 0.53 | 0.38 | 0.38 | 0.40 | 0.34 | 0.83 |
| | DQ | 0.25 | 0.34 | 0.27 | 0.06 | 0.07 | 0.31 | | DQ | 0.97 | 0.36 | 0.34 | 0.16 | 0.27 | 0.42 |
| CAC | NV | 29 | 26 | 27 | 26 | 27 | 22 | NSQ | NV | 20 | 22 | 22 | 20 | 20 | 17 |
| (18) | UC | **0.01** | 0.08 | **0.05** | 0.08 | **0.05** | 0.36 | (17) | UC | 0.48 | 0.24 | 0.24 | 0.48 | 0.48 | 1 |
| | IN | 0.33 | 0.38 | 0.37 | 0.38 | 0.37 | 0.46 | | IN | 0.50 | 0.45 | 0.45 | 0.50 | 0.49 | 0.56 |
| | CC | **0.03** | 0.14 | 0.09 | 0.14 | 0.09 | 0.49 | | CC | 0.61 | 0.38 | 0.38 | 0.61 | 0.61 | 0.84 |
| | DQ | **0.02** | **0.05** | **0.04** | **0.05** | **0.04** | 0.09 | | DQ | 0.0008 | 0.13 | 0.13 | 0.18 | 0.40 | 0.16 |
| DAX | NV | 19 | 18 | 17 | 16 | 16 | 15 | NK | NV | 15 | 16 | 16 | 15 | 16 | 15 |
| (18) | UC | 0.81 | 1 | 0.81 | 0.63 | 0.63 | 0.46 | (16) | UC | 0.80 | 1 | 1 | 0.80 | 1 | 0.80 |
| | IN | 0.52 | 0.54 | 0.57 | 0.59 | 0.59 | 0.61 | | IN | 0.13 | 0.15 | 0.15 | 0.13 | 0.15 | 0.13 |
| | CC | 0.78 | 0.82 | 0.81 | 0.76 | 0.76 | 0.66 | | CC | 0.30 | 0.34 | 0.34 | 0.30 | 0.35 | 0.29 |
| | DQ | 0.16 | **0.04** | 0.14 | 0.11 | 0.42 | 0.29 | | DQ | 0.50 | **0.007** | **0.008** | **0.01** | **0.02** | **0.007** |
| DJ* | NV | 21 | 22 | 21 | 20 | 21 | 17 | RUS | NV | 23 | 25 | 25 | 23 | 24 | 20 |
| (17) | UC | 0.35 | 0.25 | 0.35 | 0.48 | 0.34 | 1 | (17) | UC | 0.16 | 0.07 | 0.07 | 0.16 | 0.11 | 0.48 |
| | IN | 0.47 | 0.28 | 0.48 | 0.50 | 0.47 | 0.56 | | IN | 0.43 | 0.39 | 0.39 | 0.43 | 0.41 | 0.49 |
| | CC | 0.49 | 0.28 | 0.49 | 0.61 | 0.49 | 0.83 | | CC | 0.28 | 0.13 | 0.13 | 0.28 | 0.19 | 0.61 |
| | DQ | 0.52 | **0.008** | **0.01** | 0.95 | 0.69 | 0.99 | | DQ | **0.005** | **0.001** | **0.001** | **0.001** | **0.02** | 0.42 |
| ESX* | NV | 31 | 18 | 18 | 19 | 19 | 15 | SPX* | NV | 27 | 25 | 28 | 25 | 25 | 23 |
| (18) | UC | **0.01** | 1 | 1 | 0.82 | 0.81 | 0.46 | (17) | UC | **0.02** | 0.07 | **0.01** | 0.07 | 0.07 | 0.16 |
| | IN | 0.29 | 0.54 | 0.55 | 0.53 | 0.53 | 0.61 | | IN | 0.36 | 0.40 | 0.34 | 0.40 | 0.40 | 0.43 |
| | CC | **0.01** | 0.82 | 0.82 | 0.79 | 0.79 | 0.66 | | CC | 0.05 | 0.13 | 0.03 | 0.13 | 0.13 | 0.27 |
| | DQ | **0.001** | 0.37 | 0.37 | 0.13 | 0.12 | 0.41 | | DQ | 0.11 | 0.40 | 0.003 | 0.20 | 0.28 | 0.73 |
| MIB | NV | 31 | 23 | 23 | 24 | 25 | 23 | SMI | NV | 17 | 19 | 19 | 18 | 18 | 15 |
| (18) | UC | **0.01** | 0.26 | 0.26 | 0.17 | 0.12 | 0.26 | (17) | UC | 1 | 0.63 | 0.63 | 0.81 | 0.81 | 0.62 |
| | IN | 0.29 | 0.44 | 0.44 | 0.42 | 0.40 | 0.43 | | IN | 0.56 | 0.20 | 0.20 | 0.54 | 0.54 | 0.62 |
| | CC | **0.01** | 0.38 | 0.38 | 0.29 | 0.20 | 0.38 | | CC | 0.84 | 0.39 | 0.39 | 0.79 | 0.79 | 0.77 |
| | DQ | **0.003** | **0.02** | **0.02** | **0.04** | 0.22 | **0.04** | | DQ | 0.68 | 0.13 | 0.13 | 0.17 | 0.17 | 0.93 |
| FT | NV | 20 | 16 | 16 | 15 | 15 | 13 | | | | | | | | |
| (18) | UC | 0.64 | 0.63 | 0.63 | 0.46 | 0.46 | 0.26 | | | | | | | | |
| | IN | 0.50 | 0.59 | 0.59 | 0.61 | 0.61 | 0.43 | | | | | | | | |
| | CC | 0.70 | 0.76 | 0.76 | 0.67 | 0.67 | 0.38 | | | | | | | | |
| | DQ | 0.71 | 0.96 | 0.97 | 0.98 | 0.98 | 0.34 | | | | | | | | |

**Table 8: Performance of one-day-ahead ES forecast**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|
| AEX* | **0.020** | 0.090 | 0.113 | 0.097 | **0.048** | 0.135 |
| AOI | 0.920 | 0.577 | 0.480 | 0.787 | 0.630 | 0.96 |
| BVP | 0.414 | 0.988 | 0.982 | 1.000 | 1.000 | 1.000 |
| CAC | 0.103 | 0.599 | 0.457 | 0.425 | 0.148 | 0.655 |
| DAX | 0.262 | 0.315 | 0.307 | 0.303 | 0.376 | 0.235 |
| DJ* | 0.998 | 1.000 | 1.000 | 0.993 | 1.000 | 1.000 |
| ESX* | 0.737 | 0.406 | 0.431 | 0.555 | 0.567 | 0.491 |
| MIB | 0.821 | 0.807 | 0.840 | 0.857 | 0.844 | 0.906 |
| FT | 0.410 | 0.647 | 0.719 | 0.791 | 0.693 | 0.719 |
| IBX | 0.328 | 0.435 | 0.317 | 0.257 | 0.213 | 0682 |
| IPC | 0.827 | 0.991 | 0.989 | 0.999 | 0.956 | 0.953 |
| KCI | 0.063 | 0.254 | 0.135 | 0.177 | 0.076 | 0.289 |
| NSQ | 0.490 | 0.900 | 0.910 | 0.797 | 0.755 | 0.678 |
| NK | **0.023** | **0.034** | **0.023** | **0.017** | **0.008** | **0.023** |
| RUS | 0.960 | 0.781 | 0.671 | 0.523 | 0.500 | 0.539 |
| SPX* | 0.999 | 0.995 | 0.993 | 0.999 | 1.000 | 1.000 |
| SMI | 0.198 | 0.860 | 0.875 | 0.776 | 0.799 | 0.638 |

**Table 9: Performance measures for the ten-day-ahead VaR forecast**. Actual and expected number of violations of the VaR.Cases in which the null hypothesis of correct unconditional coverage is rejected at level $\alpha = 10\%$ are in bold. The best performer is in italic. The * indicates series for which a higher threshold was required for good GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 18 | **34 (0.08)** | 18 (1.00) | *16 (0.81)* | 19 (0.89) | 20 (0.79) | 18 (0.99) |
| AOI* | 17 | 25 (0.43) | 25 (0.43) | 25 (0.43) | 23 (0.60) | 22 (0.68) | *21 (0.63)* |
| BVP | 17 | 12 (0.49) | 22 (0.54) | 22 (0.54) | 21 (0.60) | 22 (0.54) | *19 (0.77)* |
| CAC | 18 | 29 (0.23) | *27 (0.36)* | *27 (0.36)* | *27 (0.34)* | *27 (0.34)* | *27 (0.30)* |
| DAX | 18 | 27 (0.33) | 22 (0.66) | *20 (0.82)* | 21 (0.73) | 21 (0.73) | 22 (0.66) |
| DJ | 17 | 12 (0.51) | 20 (0.71) | 20 (0.71) | *18 (0.92)* | *18 (0.92)* | 19 (0.83) |
| ESX | 18 | 32 (0.17) | 21 (0.72) | 21 (0.72) | 21 (0.72) | 22 (0.64) | *20 (0.81)* |
| MIB | 18 | **50 (0.02)** | *25 (0.36)* | *25 (0.36)* | *25 (0.36)* | *25 (0.36)* | *25 (0.34)* |
| FT | 18 | 33 (0.14) | 19 (0.95) | *18 (0.95)* | *18 (0.95)* | 19 (0.95) | 24 (0.41) |
| IBX | 18 | **39 (0.07)** | *24 (0.87)* | 25 (0.78) | 26 (0.69) | 26 (0.69) | 25 (0.35) |
| IPC | 17 | *14 (0.68)* | 7 (0.16) | 7 (0.16) | 7 (0.16) | 7 (0.16) | 9 (0.30) |
| KCI* | 17 | *11 (0.45)* | 26 (0.21) | 27 (0.25) | 27 (0.25) | 28 (0.23) | 26 (0.21) |
| NSQ | 17 | 20 (0.73) | *18 (0.92)* | 18 (0.92) | 15 (0.75) | 15 (0.75) | 14 (0.62) |
| NK | 16 | *24 (0.57)* | 25 (0.47) | 25 (0.47) | 26 (0.39) | 26 (0.39) | 25 (0.48) |
| RUS | 17 | 18 (0.94) | 18 (0.93) | *17 (0.98)* | *17 (0.98)* | 16 (0.88) | 18 (0.94) |
| SPX | 17 | 10 (0.38) | *17 (0.98)* | *17 (0.97)* | 16 (0.88) | 16 (0.88) | 19 (0.83) |
| SMI | 17 | 24 (0.41) | *11 (0.35)* | *11 (0.35)* | *11 (0.35)* | *11 (0.35)* | *11 (0.35)* |

**Table 10: Performance of ten-day-ahead ES forecast**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.114 | 0.124 | 0.071 | 0.123 | 0.333 | 0.294 |
| AOI* | 0.946 | 0.244 | 0.249 | 0.243 | 0.345 | 0.122 |
| BVP | 0.673 | 0.128 | 0.138 | 0.118 | **0.035** | 0.073 |
| CAC | 0.283 | 0.534 | 0.562 | 0.414 | 0.391 | 0.601 |
| DAX | **0.045** | 0.101 | **0.035** | 0.060 | 0.059 | 0.056 |
| DJ | 0.345 | 0.288 | 0.246 | 0.232 | 0.217 | 0.211 |
| ESX | 0.153 | 0.170 | 0.203 | 0.153 | 0.057 | **0.025** |
| MIB | 0.288 | 0.743 | 0.763 | 0.776 | 0.787 | 0.816 |
| FT | 0.119 | 0.451 | 0.396 | 0.312 | 0.379 | 0.622 |
| IBX | 0.738 | 0.157 | 0.196 | 0.291 | 0.238 | 0.183 |
| IPC | 0.220 | 0.767 | 0.741 | 0.908 | 0.868 | 1.000 |
| KCI* | **0.027** | 0.123 | 0.140 | 0.251 | 0.281 | 0.319 |
| NSQ | 0.062 | **0.028** | **0.022** | **0.000** | **0.000** | **0.003** |
| NK | **0.007** | **0.021** | **0.030** | **0.019** | **0.024** | **0.035** |
| RUS | 0.089 | 0.156 | 0.085 | 0.134 | 0.052 | 0.178 |
| SPX | 0.282 | **0.047** | **0.043** | **0.038** | **0.038** | 0.092 |
| SMI | 0.479 | 0.511 | 0.483 | 0.493 | 0.484 | 0.341 |

period of turmoil and the poorer performance of the GJR-GARCH(1,1) is clearer in the estimates of ES.

### 3.5.2 Microstructure noise

To investigate whether the microstructure noise has an impact on the analysis of Section 3.4, we repeat it on both the full and restricted samples using two estimators of the quadratic variation that mitigate the effect of the microstructure noise: the *Second-Best Approach: Sub-sampling and Averaging* estimator of Zhang et al. (2005), which is biased in the presence of microstructure noise but presents a lower variance than the RV estimator; and the Realized Kernel (RK) of Barndorff-Nielsen et al. (2008).

We assess both the filtering and forecasting component of the RV-EVT using first the sub-sampled measures and then substituting the sub-sampled RV with the RK. Results reported in Appendix B.5 lead to the same conclusions of Section 3.4.

## 3.6 Concluding remarks

This article questions whether combining the recent advances in the high-frequency financial econometrics with results from the Extreme Value Theory can improve the fit of the tails of the conditional returns distribution.

We propose an RV-EVT approach where returns are pre-whitened with a high-frequency based volatility model and the POT approach is applied to the tails of the standardized residuals. We use three different functions to link the predictions of the realized volatility to the conditional variance of returns, and employ six different econometric specifications to model the realized volatility.

This approach is compared to the standard C-EVT technique both through simulation and with an extensive empirical analysis on 17 international indexes. We assess both approaches' ability of filtering the dependence in the extremes and of producing stable out-of-sample VaR and ES predictions for one- and ten-day time horizons.

Summarizing the results, it seems that a GARCH-type filter performs slightly better than a high-frequency based filter, even though they both tend to produce estimated residuals which are close to independent. From a risk management perspective however, the RV-EVT approach seems preferable, especially at the longer time horizons which are of interest for regulatory purposes.

# Appendix B

## B.1 Sampling from the semi-parametric distribution of the residuals

1. Randomly select with replacement a residual from the sample of $S$ residuals;

2. If the residual exceeds an upper threshold $\widehat{u}^H$ sample a GP($\widehat{\xi}^H, \widehat{\eta}^H$) distributed observation $y^H$ from the right tail and return $\widehat{u}^H + y^H$;

3. If the residual is less than a lower threshold $\widehat{u}^L$ sample a GP($\widehat{\xi}^L, \widehat{\eta}^L$) distributed observation $y^L$ from the left tail and return $\widehat{u}^L + y^L$;

4. Otherwise return the residual itself;

5. Repeat.

## B.2 Augmented HAR-RV model with signed jumps

Patton and Sheppard (2015) propose to extend the HAR-J model of Andersen et al. (2007), isolating the information coming from the sign of the jumps. They use a measure called signed jump variation $\Delta J_t^2 \equiv RS_t^+ - RS_t^-$ which is positive when a day is dominated by an upward jump and negative when a day is dominated by a downward jump. They

explicitly decompose this measure into two components

$$\begin{aligned} \Delta J_t^{2+} &= (RS_t^+ - RS_t^-) I_{\{(RS_t^+ - RS_t^-)>0\}}, \\ \Delta J_t^{2-} &= (RS_t^+ - RS_t^-) I_{\{(RS_t^+ - RS_t^-)<0\}}. \end{aligned} \tag{B.1}$$

If jumps are rare then these measures should respectively correspond to the jump variation when either a positive or a negative jump occurs. However, if jumps tend to cluster, there might be more than one jump in a day and the measures in Equation (B.1) will fail to capture the whole jump contribution. To verify whether the actual size of the jumps could still generate valuable information, we extend the model of Patton and Sheppard (2015) to obtain the following HAR-SJ*aug* model,

$$\begin{aligned} log(RV_{t,t+h}) &= \beta_0 + \beta_J \log(1+J_t) + \beta_{J+} \log(\Delta J_t^{2+} + 1) \\ &+ \beta_{J-} \log(\Delta J_t^{2-} + 1) + \beta_{JJ+} \log(1+J_t) \log(\Delta J_t^{2+} + 1) \\ &+ \beta_{JJ-} \log(1+J_t) \log(\Delta J_t^{2-} + 1) + \beta_C \log(RV_t) \\ &+ \beta_W \log(RV_{t-5,t}) + \beta_M \log(RV_{t-22,t}) + \eta_{t,t+h}. \end{aligned}$$

## B.3 Forecasting with other link functions

In this section, we report the performance evaluation for the risk-measure forecasts obtained with the type-II and type-III link functions.

The results for both the one-day (Tables 11 to 14) and ten-day-ahead (Tables 15 to 18) predictions are very similar to those obtained with the type-I link function. At the one-day level, the performance of the C-EVT and RV-EVT are very similar, but as the time horizon increases the higher accuracy of the high-frequency based volatility estimates results in improved predictions of the risk-measures.

Comparing the outcomes across the three different link functions, it seems that the higher complexity of the conditional volatility model does not bear any significant impact on forecasting. Given that the type-II and type-III link functions require more effort to get an estimate of the volatility, only type-I link function results are included in the main paper. Practitioners can decide whether a more complex link function is necessary.

**Table 11: Performance measures for the one-day-ahead VaR forecast with type-II link function**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | PM | (a) | (b) | (c) | (d) | (e) | (f) | Index | PM | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX* | NV | 22 | 18 | 18 | 19 | 19 | 18 | IBX* | NV | 26 | 20 | 21 | 21 | 21 | 22 |
| (18) | UC | 0.36 | 1 | 1 | 0.81 | 0.81 | 1 | (18) | UC | 0.08 | 0.64 | 0.49 | 0.49 | 0.49 | 0.36 |
| | IN | 0.46 | 0.55 | 0.55 | 0.53 | 0.53 | 0.54 | | IN | 0.38 | 0.22 | 0.25 | 0.25 | 0.25 | 0.27 |
| | CC | 0.50 | 0.83 | 0.83 | 0.79 | 0.79 | 0.83 | | CC | 0.14 | 0.42 | 0.40 | 0.40 | 0.40 | 0.36 |
| | DQ | 0.65 | 0.72 | 0.72 | 0.75 | 0.75 | 0.72 | | DQ | **0.003** | 0.20 | 0.24 | **0.002** | **0.002** | 0.12 |
| AOI | NV | 16 | 13 | 16 | 14 | 15 | 17 | IPC* | NV | 13 | 9 | 9 | 9 | 9 | 9 |
| (17) | UC | 0.81 | 0.31 | 0.81 | 0.45 | 0.62 | 1 | (17) | UC | 0.31 | **0.03** | **0.03** | **0.03** | **0.03** | **0.03** |
| | IN | 0.59 | 0.66 | 0.59 | 0.63 | 0.61 | 0.56 | | IN | 0.66 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 |
| | CC | 0.83 | 0.54 | 0.83 | 0.67 | 0.77 | 0.84 | | CC | 0.54 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |
| | DQ | 1 | 0.34 | 0.00 | 0.44 | 0.48 | 0.64 | | DQ | 0.27 | **0.04** | 0.74 | 0.74 | 0.74 | **0.04** |
| BVP | NV | 13 | 14 | 14 | 13 | 14 | 14 | KCI | NV | 13 | 20 | 20 | 19 | 20 | 19 |
| (17) | UC | 0.31 | 0.45 | 0.45 | 0.31 | 0.45 | 0.45 | (17) | UC | 0.31 | 0.48 | 0.48 | 0.63 | 0.48 | 0.63 |
| | IN | 0.65 | 0.63 | 0.63 | 0.65 | 0.63 | 0.62 | | IN | 0.65 | 0.23 | 0.23 | 0.21 | 0.23 | 0.21 |
| | CC | 0.53 | 0.66 | 0.66 | 0.53 | 0.66 | 0.66 | | CC | 0.53 | 0.38 | 0.38 | 0.40 | 0.38 | 0.40 |
| | DQ | 0.25 | 0.26 | 0.26 | 0.17 | 0.27 | 0.21 | | DQ | 0.98 | 0.36 | 0.34 | 0.16 | 0.27 | 0.32 |
| CAC | NV | 29 | 27 | 28 | 26 | 26 | 27 | NSQ | NV | 20 | 21 | 21 | 21 | 21 | 23 |
| (18) | UC | **0.02** | **0.05** | **0.03** | 0.08 | 0.08 | **0.05** | (17) | UC | 0.48 | 0.35 | 0.35 | 0.35 | 0.35 | 0.16 |
| | IN | 0.33 | 0.37 | 0.35 | 0.38 | 0.38 | 0.36 | | IN | 0.50 | 0.47 | 0.47 | 0.47 | 0.47 | 0.43 |
| | CC | **0.04** | 0.09 | 0.06 | 0.14 | 0.14 | 0.09 | | CC | 0.61 | 0.49 | 0.49 | 0.49 | 0.49 | 0.28 |
| | DQ | **0.02** | 0.11 | 0.09 | 0.14 | 0.07 | 0.11 | | DQ | **0.0009** | 0.16 | 0.16 | 0.26 | 0.41 | 0.24 |
| DAX* | NV | 19 | 21 | 21 | 20 | 21 | 26 | NK | NV | 15 | 16 | 16 | 15 | 16 | 16 |
| (18) | UC | 0.81 | 0.49 | 0.49 | 0.64 | 0.49 | 0.08 | (16) | UC | 0.80 | 1 | 1 | 0.80 | 1.00 | 1 |
| | IN | 0.52 | 0.48 | 0.48 | 0.50 | 0.48 | 0.38 | | IN | 0.13 | 0.15 | 0.15 | 0.13 | 0.15 | 0.15 |
| | CC | 0.79 | 0.61 | 0.61 | 0.71 | 0.61 | 0.14 | | CC | 0.30 | 0.35 | 0.35 | 0.30 | 0.35 | 0.35 |
| | DQ | 0.16 | 0.10 | 0.11 | 0.30 | 0.31 | **0.01** | | DQ | 0.50 | **0.01** | **0.01** | **0.01** | **0.02** | **0.01** |
| DJ* | NV | 21 | 21 | 21 | 20 | 21 | 21 | RUS | NV | 23 | 22 | 24 | 22 | 22 | 26 |
| (17) | UC | 0.35 | 0.35 | 0.35 | 0.48 | 0.35 | 0.35 | (17) | UC | 0.17 | 0.24 | 0.11 | 0.24 | 0.24 | **0.04** |
| | IN | 0.47 | 0.47 | 0.47 | 0.50 | 0.47 | 0.25 | | IN | 0.43 | 0.45 | 0.41 | 0.45 | 0.45 | 0.37 |
| | CC | 0.49 | 0.49 | 0.49 | 0.61 | 0.49 | 0.33 | | CC | 0.28 | 0.38 | 0.19 | 0.38 | 0.38 | 0.08 |
| | DQ | 0.52 | **0.02** | **0.02** | 0.95 | 0.69 | 0.43 | | DQ | **0.01** | 0.66 | **0.01** | 0.66 | 0.66 | 0.31 |
| ESX* | NV | 31 | 16 | 16 | 18 | 16 | 24 | SPX* | NV | 27 | 24 | 25 | 21 | 21 | 26 |
| (18) | UC | **0.01** | 0.63 | 0.63 | 1 | 0.63 | 0.17 | (17) | UC | **0.02** | 0.11 | 0.07 | 0.35 | 0.35 | **0.05** |
| | IN | 0.30 | 0.59 | 0.59 | 0.55 | 0.59 | 0.42 | | IN | 0.36 | 0.41 | 0.39 | 0.47 | 0.47 | 0.37 |
| | CC | **0.01** | 0.76 | 0.76 | 0.82 | 0.76 | 0.28 | | CC | 0.05 | 0.19 | 0.13 | 0.49 | 0.49 | 0.09 |
| | DQ | **0.001** | 0.17 | 0.17 | 0.21 | 0.12 | **0.01** | | DQ | 0.11 | 0.51 | 0.21 | 0.45 | 0.69 | 0.31 |
| MIB | NV | 31 | 23 | 24 | 26 | 25 | 29 | SMI | NV | 17 | 18 | 18 | 18 | 18 | 22 |
| (18) | UC | **0.01** | 0.26 | 0.18 | 0.08 | 0.12 | **0.02** | (17) | UC | 1 | 0.81 | 0.81 | 0.81 | 0.81 | 0.24 |
| | IN | 0.29 | 0.44 | 0.42 | 0.38 | 0.40 | 0.32 | | IN | 0.56 | 0.18 | 0.18 | 0.54 | 0.54 | 0.28 |
| | CC | **0.01** | 0.38 | 0.28 | 0.14 | 0.20 | **0.04** | | CC | 0.84 | 0.39 | 0.39 | 0.80 | 0.80 | 0.28 |
| | DQ | **0.003** | **0.02** | **0.01** | 0.04 | 0.22 | **0.01** | | DQ | 0.68 | 0.14 | 0.14 | 0.18 | 0.18 | 0.12 |
| FT | NV | 20 | 15 | 16 | 15 | 15 | 15 | | | | | | | | |
| (18) | UC | 0.64 | 0.46 | 0.63 | 0.46 | 0.46 | 0.46 | | | | | | | | |
| | IN | 0.50 | 0.61 | 0.59 | 0.61 | 0.61 | 0.61 | | | | | | | | |
| | CC | 0.71 | 0.67 | 0.76 | 0.67 | 0.67 | 0.67 | | | | | | | | |
| | DQ | 0.71 | 0.98 | 0.97 | 0.98 | 0.98 | 0.97 | | | | | | | | |

**Table 12: Performance measures for the one-day-ahead VaR forecast with type-III link function**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | PM | (a) | (b) | (c) | (d) | (e) | (f) | Index | PM | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX* | NV | 22 | 18 | 18 | 19 | 19 | 18 | IBX | NV | 26 | 20 | 20 | 21 | 21 | 22 |
| (18) | UC | 0.36 | 1 | 1 | 0.81 | 0.81 | 1 | (18) | UC | 0.08 | 0.64 | 0.64 | 0.49 | 0.49 | 0.36 |
| | IN | 0.46 | 0.55 | 0.55 | 0.53 | 0.53 | 0.55 | | IN | 0.38 | 0.22 | 0.22 | 0.25 | 0.25 | 0.27 |
| | CC | 0.50 | 0.83 | 0.83 | 0.79 | 0.79 | 0.83 | | CC | 0.14 | 0.42 | 0.42 | 0.40 | 0.40 | 0.36 |
| | DQ | 0.65 | 0.72 | 0.72 | 0.75 | 0.75 | 0.72 | | DQ | **0.003** | 0.20 | 0.25 | **0.002** | **0.002** | 0.12 |
| AOI | NV | 16 | 13 | 16 | 14 | 15 | 17 | IPC* | NV | 13 | 9 | 9 | 9 | 9 | 9 |
| (17) | UC | 0.81 | 0.31 | 0.81 | 0.45 | 0.62 | 1 | (17) | UC | 0.31 | **0.03** | **0.03** | **0.03** | **0.03** | **0.03** |
| | IN | 0.59 | 0.66 | 0.59 | 0.63 | 0.61 | 0.56 | | IN | 0.66 | 0.76 | 0.76 | 0.76 | 0.76 | 0.76 |
| | CC | 0.83 | 0.54 | 0.83 | 0.67 | 0.77 | 0.83 | | CC | 0.54 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |
| | DQ | 1 | 0.34 | 0.00 | 0.44 | 0.48 | 0.67 | | DQ | 0.27 | **0.04** | 0.74 | 0.74 | 0.74 | **0.04** |
| BVP | NV | 13 | 14 | 14 | 13 | 14 | 14 | KCI | NV | 13 | 20 | 20 | 19 | 20 | 19 |
| (17) | UC | 0.31 | 0.45 | 0.45 | 0.31 | 0.45 | 0.45 | (17) | UC | 0.31 | 0.48 | 0.48 | 0.63 | 0.48 | 0.63 |
| | IN | 0.65 | 0.63 | 0.63 | 0.65 | 0.63 | 0.63 | | IN | 0.65 | 0.23 | 0.23 | 0.21 | 0.23 | 0.21 |
| | CC | 0.53 | 0.66 | 0.66 | 0.53 | 0.66 | 0.66 | | CC | 0.53 | 0.38 | 0.38 | 0.40 | 0.38 | 0.40 |
| | DQ | 0.25 | 0.26 | 0.26 | 0.17 | 0.27 | 0.27 | | DQ | 0.98 | 0.36 | 0.34 | 0.16 | 0.27 | 0.32 |
| CAC | NV | 29 | 27 | 28 | 26 | 26 | 27 | NSQ | NV | 20 | 21 | 21 | 21 | 21 | 23 |
| (18) | UC | **0.02** | **0.05** | **0.03** | 0.08 | 0.08 | **0.05** | (17) | UC | 0.48 | 0.35 | 0.35 | 0.35 | 0.35 | 0.17 |
| | IN | 0.33 | 0.37 | 0.35 | 0.38 | 0.38 | 0.37 | | IN | 0.50 | 0.47 | 0.47 | 0.47 | 0.47 | 0.43 |
| | CC | **0.04** | 0.09 | 0.06 | 0.14 | 0.14 | 0.09 | | CC | 0.61 | 0.49 | 0.49 | 0.49 | 0.49 | 0.27 |
| | DQ | **0.02** | 0.11 | 0.09 | 0.14 | 0.07 | 0.11 | | DQ | **0.0009** | 0.16 | 0.16 | 0.26 | 0.41 | 0.23 |
| DAX* | NV | 19 | 21 | 21 | 20 | 21 | 26 | NK | NV | 15 | 16 | 16 | 15 | 16 | 16 |
| (18) | UC | 0.81 | 0.49 | 0.49 | 0.64 | 0.49 | 0.08 | (16) | UC | 0.80 | 1 | 1 | 0.80 | 1 | 1 |
| | IN | 0.52 | 0.48 | 0.48 | 0.50 | 0.48 | 0.38 | | IN | 0.13 | 0.15 | 0.15 | 0.13 | 0.15 | 0.15 |
| | CC | 0.79 | 0.61 | 0.61 | 0.71 | 0.61 | 0.14 | | CC | 0.30 | 0.35 | 0.35 | 0.30 | 0.35 | 0.35 |
| | DQ | 0.16 | 0.10 | 0.11 | 0.30 | 0.31 | **0.01** | | DQ | 0.50 | **0.01** | **0.01** | **0.01** | **0.02** | **0.01** |
| DJ* | NV | 21 | 21 | 21 | 20 | 21 | 21 | RUS | NV | 23 | 22 | 25 | 21 | 22 | 26 |
| (17) | UC | 0.35 | 0.35 | 0.35 | 0.48 | 0.35 | 0.35 | (17) | UC | 0.17 | 0.24 | 0.07 | 0.35 | 0.24 | **0.05** |
| | IN | 0.47 | 0.47 | 0.47 | 0.50 | 0.47 | 0.25 | | IN | 0.43 | 0.45 | 0.39 | 0.47 | 0.45 | 0.37 |
| | CC | 0.49 | 0.49 | 0.49 | 0.61 | 0.49 | 0.33 | | CC | 0.28 | 0.38 | 0.13 | 0.49 | 0.38 | 0.08 |
| | DQ | 0.52 | **0.02** | **0.02** | 0.95 | 0.69 | 0.43 | | DQ | **0.01** | 0.66 | **0.001** | 0.57 | 0.66 | **0.01** |
| ESX | NV | 31 | 16 | 16 | 18 | 16 | 24 | SPX* | NV | 27 | 25 | 27 | 24 | 23 | 26 |
| (18) | UC | **0.01** | 0.63 | 0.63 | 1.00 | 0.63 | 0.18 | (17) | UC | **0.02** | 0.07 | **0.02** | 0.11 | 0.17 | **0.04** |
| | IN | 0.30 | 0.59 | 0.59 | 0.55 | 0.59 | 0.42 | | IN | 0.36 | 0.40 | 0.37 | 0.41 | 0.43 | 0.37 |
| | CC | **0.01** | 0.76 | 0.76 | 0.82 | 0.76 | 0.29 | | CC | 0.05 | 0.13 | 0.05 | 0.19 | 0.28 | 0.08 |
| | DQ | **0.001** | 0.17 | 0.17 | 0.21 | 0.12 | **0.01** | | DQ | 0.11 | 0.40 | 0.00 | 0.22 | 0.36 | 0.31 |
| MIB | NV | 31 | 23 | 24 | 26 | 25 | 29 | SMI | NV | 17 | 18 | 18 | 18 | 18 | 22 |
| (18) | UC | **0.01** | 0.26 | 0.18 | 0.08 | 0.12 | **0.02** | (17) | UC | 1 | 0.81 | 0.81 | 0.81 | 0.81 | 0.24 |
| | IN | 0.29 | 0.44 | 0.42 | 0.38 | 0.40 | 0.33 | | IN | 0.56 | 0.18 | 0.18 | 0.54 | 0.54 | 0.28 |
| | CC | **0.01** | 0.38 | 0.28 | 0.14 | 0.20 | **0.04** | | CC | 0.84 | 0.39 | 0.39 | 0.80 | 0.80 | 0.28 |
| | DQ | **0.003** | **0.02** | **0.01** | **0.04** | 0.22 | **0.01** | | DQ | 0.68 | 0.14 | 0.14 | 0.18 | 0.18 | 0.12 |
| FT | NV | 20 | 15 | 16 | 15 | 15 | 15 | | | | | | | | |
| (18) | UC | 0.64 | 0.46 | 0.63 | 0.46 | 0.46 | 0.46 | | | | | | | | |
| | IN | 0.50 | 0.61 | 0.59 | 0.61 | 0.61 | 0.61 | | | | | | | | |
| | CC | 0.71 | 0.67 | 0.76 | 0.67 | 0.67 | 0.67 | | | | | | | | |
| | DQ | 0.71 | 0.98 | 0.97 | 0.98 | 0.98 | 0.97 | | | | | | | | |

**Table 13: Performance of one-day-ahead ES forecast with type-II link function**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-Sjaug | LHAR |
|---|---|---|---|---|---|---|
| AEX* | **0.024** | 0.104 | 0.114 | 0.098 | 0.065 | 0.051 |
| AOI | 0.919 | 0.517 | 0.603 | 0.642 | 0.407 | 0.458 |
| BVP | 0.387 | 0.880 | 0.839 | 0.963 | 0.997 | 0.576 |
| CAC | 0.102 | 0.572 | 0.485 | 0.395 | 0.059 | 0.193 |
| DAX* | 0.261 | 0.543 | 0.558 | 0.636 | 0.644 | 0.717 |
| DJ* | 0.996 | 1.000 | 1.000 | 0.969 | 0.988 | 0.988 |
| ESX | 0.720 | 0.464 | 0.484 | 0.521 | 0.233 | 0.503 |
| MIB | 0.810 | 0.649 | 0.738 | 0.850 | 0.791 | 0.894 |
| FT | 0.435 | 0.560 | 0.673 | 0.720 | 0.633 | 0.365 |
| IBX | 0.354 | 0.427 | 0.419 | 0.252 | 0.151 | 0.318 |
| IPC* | 0.822 | 0.556 | 0.550 | 0.464 | 0.460 | 0.181 |
| KCI | 0.071 | 0.324 | 0.183 | 0.252 | 0.084 | 0.132 |
| NSQ | 0.481 | 0.790 | 0.829 | 0.814 | 0.826 | 0.718 |
| NK | **0.020** | **0.027** | **0.027** | **0.009** | **0.006** | **0.007** |
| RUS | 0.959 | 0.165 | 0.125 | 0.514 | 0.216 | 0.180 |
| SPX* | 0.999 | 0.988 | 0.974 | 0.983 | 0.945 | 0.887 |
| SMI | 0.199 | 0.781 | 0.754 | 0.719 | 0.784 | 0.823 |

**Table 14: Performance of one-day-ahead ES forecast with type-III link function**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-Sjaug | LHAR |
|---|---|---|---|---|---|---|
| AEX* | **0.013** | 0.090 | 0.105 | 0.090 | 0.062 | 0.051 |
| AOI | 0.948 | 0.504 | 0.611 | 0.654 | 0.403 | 0.478 |
| BVP | 0.408 | 0.872 | 0.812 | 0.965 | 0.991 | 0.534 |
| CAC | 0.130 | 0.589 | 0.467 | 0.422 | 0.055 | 0.161 |
| DAX* | 0.241 | 0.565 | 0.537 | 0.595 | 0.658 | 0.731 |
| DJ* | 0.997 | 0.999 | 1.000 | 0.973 | 0.995 | 0.987 |
| ESX | 0.703 | 0.460 | 0.453 | 0.512 | 0.256 | 0.498 |
| MIB | 0.810 | 0.680 | 0.762 | 0.875 | 0.748 | 0.886 |
| FT | 0.409 | 0.512 | 0.690 | 0.676 | 0.634 | 0.358 |
| IBX | 0.326 | 0.431 | 0.368 | 0.246 | 0.161 | 0.347 |
| IPC* | 0.828 | 0.552 | 0.522 | 0.468 | 0.487 | 0.183 |
| KCI | 0.065 | 0.293 | 0.156 | 0.246 | 0.080 | 0.130 |
| NSQ | 0.458 | 0.804 | 0.813 | 0.794 | 0.839 | 0.730 |
| NK | **0.025** | **0.020** | **0.021** | **0.010** | **0.006** | **0.009** |
| RUS | 0.964 | 0.175 | 0.252 | 0.192 | 0.208 | 0.167 |
| SPX* | 0.999 | 0.982 | 0.979 | 0.992 | 0.959 | 0.887 |
| SMI | 0.211 | 0.754 | 0.741 | 0.759 | 0.778 | 0.810 |

**Table 15: Performance measures for the ten-day-ahead VaR forecast with type-II link function**. Actual and expected number of violations of the VaR. In bold, the cases in which the null hypothesis of correct unconditional coverage is rejected at 10%. In *italics*, the best performer. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 18 | **34 (0.08)** | 17 (0.89) | 16 (0.79) | 19 (0.89) | 21 (0.69) | *18 (0.99)* |
| AOI* | 17 | 25 (0.43) | 24 (0.45) | 24 (0.45) | 24 (0.45) | 27 (0.28) | *23 (0.43)* |
| BVP | 17 | 12 (0.49) | 24 (0.45) | 24 (0.45) | *22 (0.57)* | *22 (0.57)* | *22 (0.56)* |
| CAC* | 18 | 29 (0.23) | *25 (0.45)* | *25 (0.45)* | 26 (0.41) | 26 (0.41) | 27 (0.36) |
| DAX | 18 | 27 (0.33) | 22 (0.66) | *21 (0.74)* | *21 (0.74)* | 22 (0.66) | *21 (0.73)* |
| DJ | 17 | 12 (0.51) | *17 (0.96)* | 18 (0.93) | *17 (0.96)* | *17 (0.96)* | 16 (0.86) |
| ESX | 18 | 32 (0.17) | *20 (0.82)* | *20 (0.82)* | *20 (0.82)* | 21 (0.73) | *20 (0.81)* |
| MIB | 18 | **50 (0.02)** | *25 (0.36)* | *25 (0.36)* | *25 (0.36)* | *25 (0.36)* | *25 (0.33)* |
| FT | 18 | 33 (0.14) | *18 (0.95)* | *18 (0.95)* | 16 (0.82) | 17 (0.93) | 23 (0.48) |
| IBX | 18 | **39 (0.07)** | *29 (0.23)* | *29 (0.23)* | 31 (0.15) | 32 (0.12) | 32 (0.11) |
| IPC | 17 | 14 (0.68) | 15 (0.77) | *17 (0.96)* | 16 (0.86) | 14 (0.66) | 15 (0.75) |
| KCI | 17 | *11 (0.45)* | 25 (0.45) | 25 (0.45) | 25 (0.45) | 24 (0.53) | 27 (0.34) |
| NSQ | 17 | 20 (0.73) | 22 (0.59) | 22 (0.59) | 20 (0.76) | *19 (0.84)* | 20 (0.75) |
| NK | 16 | 24 (0.57) | 27 (0.42) | 27 (0.42) | 27 (0.39) | 28 (0.30) | *23 (0.45)* |
| RUS | 17 | *18 (0.94)* | 22 (0.68) | 22 (0.68) | 21 (0.72) | 21 (0.72) | 19 (0.87) |
| SPX | 17 | 10 (0.38) | 18 (0.94) | 18 (0.94) | *17 (0.97)* | *17 (0.97)* | 20 (0.73) |
| SMI* | 17 | 24 (0.41) | *12 (0.42)* | 11 (0.34) | 11 (0.29) | 11 (0.34) | 10 (0.28) |

**Table 16: Performance measures for the ten-day-ahead VaR forecast with type-III link function**. Actual and expected number of violations of the VaR. In bold, the cases in which the null hypothesis of correct unconditional coverage is rejected at 10%. In *italics*, the best performer. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 18 | **34 (0.08)** | 17 (0.89) | 16 (0.79) | 19 (0.89) | 21 (0.69) | *18 (0.98)* |
| AOI* | 17 | 25 (0.43) | 24 (0.46) | 24 (0.46) | 24 (0.46) | 26 (0.32) | *23 (0.43)* |
| BVP | 17 | 12 (0.49) | 24 (0.45) | 24 (0.45) | 22 (0.57) | *21 (0.65)* | 22 (0.57) |
| CAC* | 18 | 29 (0.23) | *25 (0.45)* | *25 (0.45)* | 26 (0.41) | 26 (0.41) | 27 (0.36) |
| DAX | 18 | 27 (0.33) | 22 (0.66) | *21 (0.74)* | *21 (0.74)* | 22 (0.66) | *21 (0.74)* |
| DJ | 17 | 12 (0.51) | *17 (0.96)* | *17 (0.96)* | *17 (0.96)* | *17 (0.96)* | 16 (0.86) |
| ESX | 18 | 32 (0.12) | *20 (0.82)* | *20 (0.82)* | *20 (0.82)* | 21 (0.73) | *20 (0.81)* |
| MIB | 18 | **50 (0.02)** | *25 (0.36)* | *25 (0.36)* | *25 (0.36)* | *25 (0.36)* | *25 (0.33)* |
| FT | 18 | 33 (0.14) | *18 (0.95)* | 17 (0.93) | 16 (0.82) | 17 (0.93) | 23 (0.47) |
| IBX | 18 | **39 (0.07)** | *29 (0.23)* | *29 (0.23)* | 30 (0.18) | 32 (0.13) | 32 (0.11) |
| IPC | 17 | 14 (0.68) | 15 (0.77) | *17 (0.96)* | 16 (0.86) | 14 (0.66) | 15 (0.75) |
| KCI | 17 | 11 (0.45) | *25 (0.45)* | *25 (0.45)* | 26 (0.37) | 24 (0.53) | 27 (0.34) |
| NSQ | 17 | 20 (0.73) | 22 (0.59) | 22 (0.59) | 20 (0.76) | *19 (0.84)* | 20 (0.75) |
| NK | 16 | *24 (0.57)* | 27 (0.42) | 27 (0.42) | 27 (0.36) | 28 (0.30) | 26 (0.45) |
| RUS | 17 | *18 (0.94)* | 22 (0.68) | 22 (0.68) | 21 (0.73) | 21 (0.73) | 19 (0.88) |
| SPX | 17 | 10 (0.38) | 18 (0.94) | 18 (0.94) | *17 (0.97)* | *17 (0.97)* | 20 (0.74) |
| SMI* | 17 | 24 (0.41) | *12 (0.42)* | 11 (0.33) | 11 (0.28) | 11 (0.33) | 10 (0.29) |

**Table 17: Performance of ten-day-ahead ES forecast with type-II link function**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.114 | 0.099 | 0.051 | 0.121 | 0.397 | 0.268 |
| AOI* | 0.946 | 0.176 | 0.165 | 0.286 | 0.597 | 0.222 |
| BVP | 0.673 | 0.053 | **0.045** | **0.021** | **0.022** | **0.007** |
| CAC | 0.283 | 0.344 | 0.308 | 0.362 | 0.283 | 0.585 |
| DAX | **0.045** | 0.110 | 0.074 | 0.070 | 0.103 | 0.124 |
| DJ | 0.345 | 0.213 | 0.215 | 0.215 | 0.137 | 0.092 |
| ESX | 0.153 | 0.069 | 0.072 | 0.069 | 0.059 | 0.052 |
| MIB | 0.288 | 0.694 | 0.708 | 0.742 | 0.743 | 0.868 |
| FT | 0.119 | 0.439 | 0.495 | 0.290 | 0.360 | 0.644 |
| IBX | 0.738 | **0.003** | **0.006** | **0.046** | 0.051 | **0.037** |
| IPC | 0.220 | 0.882 | 0.882 | 0.816 | 0.713 | 0.828 |
| KCI* | **0.027** | 0.213 | 0.207 | 0.293 | 0.170 | 0.347 |
| NSQ | 0.062 | **0.043** | 0.055 | **0.044** | **0.032** | **0.018** |
| NK | **0.007** | **0.029** | **0.020** | **0.024** | **0.030** | **0.015** |
| RUS | 0.089 | **0.014** | **0.031** | **0.035** | **0.044** | **0.009** |
| SPX | 0.282 | 0.057 | 0.071 | **0.031** | **0.042** | 0.085 |
| SMI | 0.479 | 0.389 | 0.359 | 0.255 | 0.339 | 0.325 |

**Table 18: Performance of ten-day-ahead ES forecast with type-III link function**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.114 | 0.095 | 0.057 | 0.120 | 0.357 | 0.283 |
| AOI* | 0.946 | 0.190 | 0.171 | 0.284 | 0.521 | 0.199 |
| BVP | 0.673 | 0.050 | 0.065 | **0.030** | **0.011** | **0.011** |
| CAC | 0.283 | 0.381 | 0.307 | 0.341 | 0.289 | 0.598 |
| DAX | **0.045** | 0.109 | 0.065 | 0.071 | 0.116 | 0.118 |
| DJ | 0.345 | 0.246 | 0.189 | 0.216 | 0.145 | 0.080 |
| ESX | 0.153 | 0.073 | 0.072 | 0.061 | **0.038** | **0.048** |
| MIB | 0.288 | 0.740 | 0.694 | 0.761 | 0.735 | 0.863 |
| FT | 0.119 | 0.423 | 0.375 | 0.271 | 0.344 | 0.639 |
| IBX | 0.738 | **0.005** | **0.005** | **0.027** | **0.047** | **0.042** |
| IPC | 0.220 | 0.893 | 0.847 | 0.798 | 0.736 | 0.840 |
| KCI* | **0.027** | 0.245 | 0.234 | 0.324 | 0.184 | 0.311 |
| NSQ | 0.062 | **0.032** | **0.045** | **0.034** | **0.032** | **0.010** |
| NK | **0.007** | **0.027** | **0.033** | **0.031** | **0.032** | **0.016** |
| RUS | 0.089 | **0.023** | **0.017** | **0.030** | **0.034** | **0.013** |
| SPX | 0.282 | 0.065 | 0.052 | **0.037** | **0.027** | 0.116 |
| SMI | 0.479 | 0.419 | 0.372 | 0.243 | 0.356 | 0.315 |

## B.4 An investigation during a period of turmoil

As in the full sample analysis, we focus on the filtering and forecasting components separately to understand the merits of the C-EVT and RV-EVT with regard to both aspects.

### B.4.1 Filtering component

Figures 37-39 report the estimates of the extremal index $\widehat{\theta}$ obtained on the restricted sample (beginning of 2007 to end of 2011) with the three different link functions. The degree of extremal dependence left in the residuals of the different models is generally quite low. Results seem unaffected by the choice of the link function and the GARCH model appears to be a better filter.

**Figure 37: Extremal index estimates for the type-I link function**. Values of $\widehat{\theta}$ estimated on the standardized residuals obtained from the different windows of length $S = 500$. Data from beginning of 2007 to end of 2011.

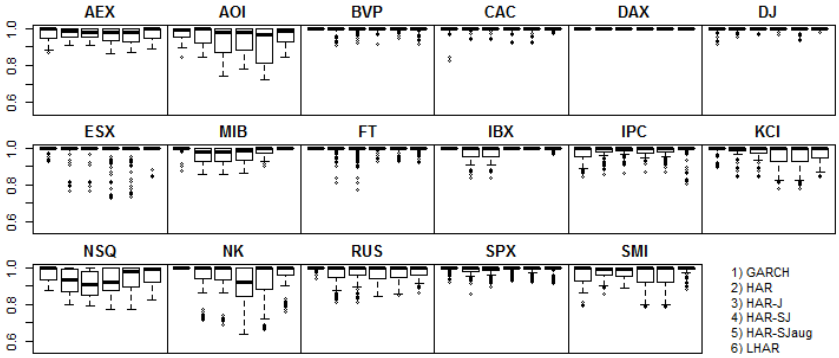


**Figure 38: Extremal index estimates for the type-II link function**. Values of $\widehat{\theta}$ estimated on the standardized residuals obtained from the different windows of length $S = 500$. Data from beginning of 2007 to end of 2011.

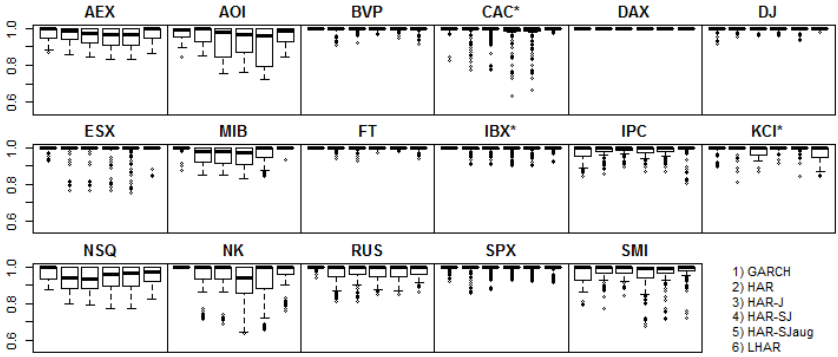
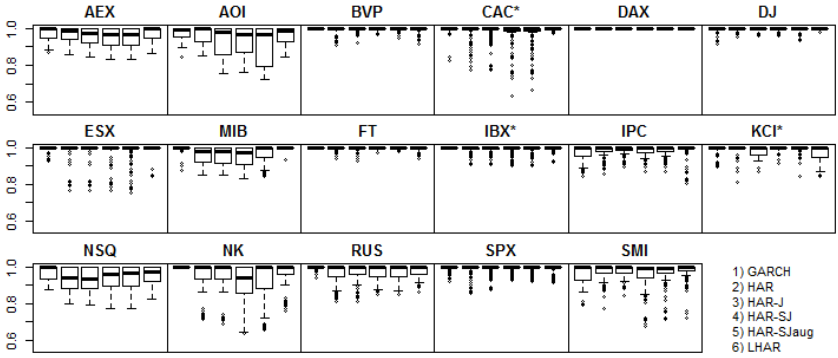



**Figure 39: Extremal index estimates for the type-III link function**. Values of $\widehat{\theta}$ estimated on the standardized residuals obtained from the different windows of length $S = 500$. Data from beginning of 2007 to end of 2011.



### B.4.2 Forecasting component

Conclusions based on the one-day-ahead VaR (Tables 19-21) and ES forecasts (Tables 22-24) obtained during the period of turmoil are the same as those from the full sample analysis: the C-EVT and RV-EVT present similar performances. Focussing on the ten-day-ahead prediction, the evidence from the restricted sample is less clear but still seems to suggest a benefit from the use of the RV-EVT. In terms of VaR violations (Tables 25-27), the GJR-GARCH(1,1) is the worst performer on half of the cases, and its poorer accuracy is emphasized in the estimates of the ES (Tables 28-30).

**Table 19: Performance measures for the one-day-ahead VaR forecast during a period of turmoil with type-I link function**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | PM | (a) | (b) | (c) | (d) | (e) | (f) | Index | PM | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX | NV | 28 | 22 | 22 | 22 | 20 | 19 | IBX | NV | 29 | 28 | 28 | 29 | 30 | 25 |
| (26) | UC | 0.69 | 0.41 | 0.41 | 0.41 | 0.21 | 0.14 | (26) | UC | 0.55 | 0.69 | 0.69 | 0.55 | 0.43 | 0.84 |
| | IN | 0.07 | 0.93 | 0.93 | 0.93 | 0.79 | 0.23 | | IN | 0.76 | 0.69 | 0.69 | 0.76 | 0.84 | 0.49 |
| | CC | 0.18 | 0.68 | 0.68 | 0.68 | 0.42 | 0.16 | | CC | 0.76 | 0.81 | 0.81 | 0.77 | 0.98 | 0.73 |
| | DQ | 0.64 | 0.51 | 0.52 | 0.52 | 0.37 | 0.74 | | DQ | 0.39 | 0.60 | 0.60 | 0.70 | 0.32 | 0.92 |
| AOI | NV | 17 | 21 | 20 | 21 | 21 | 19 | IPC | NV | 22 | 18 | 18 | 17 | 17 | 17 |
| (25) | UC | 0.08 | 0.40 | 0.29 | 0.40 | 0.40 | 0.20 | (25) | UC | 0.53 | 0.13 | 0.13 | 0.08 | 0.08 | 0.09 |
| | IN | 0.27 | 0.19 | 0.21 | 0.19 | 0.19 | 0.23 | | IN | 0.97 | 0.25 | 0.25 | 0.27 | 0.27 | 0.27 |
| | CC | 0.12 | 0.28 | 0.25 | 0.28 | 0.28 | 0.21 | | CC | 0.78 | 0.16 | 0.16 | 0.12 | 0.12 | 0.12 |
| | DQ | 0.42 | 0.57 | 0.61 | 0.57 | 0.56 | 0.40 | | DQ | 0.68 | 0.46 | 0.45 | 0.59 | 0.59 | 0.61 |
| BVP | NV | 19 | 18 | 20 | 19 | 19 | 18 | KCI | NV | 18 | 22 | 22 | 23 | 22 | 18 |
| (24) | UC | 0.28 | 0.19 | 0.39 | 0.28 | 0.28 | 0.19 | (25) | UC | 0.13 | 0.53 | 0.53 | 0.68 | 0.53 | 0.13 |
| | IN | 0.77 | 0.70 | 0.85 | 0.77 | 0.77 | 0.70 | | IN | 0.67 | 0.33 | 0.33 | 0.10 | 0.35 | 0.68 |
| | CC | 0.51 | 0.38 | 0.65 | 0.51 | 0.51 | 0.38 | | CC | 0.28 | 0.49 | 0.49 | 0.23 | 0.49 | 0.28 |
| | DQ | 0.12 | 0.45 | 0.76 | 0.85 | 0.85 | 0.87 | | DQ | 0.47 | 0.22 | 0.22 | 0.03 | 0.30 | 0.31 |
| CAC | NV | 24 | 24 | 24 | 24 | 24 | 23 | NSQ | NV | 20 | 22 | 23 | 22 | 22 | 20 |
| (26) | UC | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.54 | (25) | UC | 0.28 | 0.53 | 0.68 | 0.53 | 0.53 | 0.29 |
| | IN | 0.13 | 0.92 | 0.92 | 0.92 | 0.92 | 0.15 | | IN | 0.20 | 0.97 | 0.96 | 0.15 | 0.16 | 0.20 |
| | CC | 0.28 | 0.87 | 0.87 | 0.87 | 0.87 | 0.27 | | CC | 0.24 | 0.78 | 0.87 | 0.29 | 0.29 | 0.24 |
| | DQ | 0.28 | 0.28 | 0.29 | 0.31 | 0.30 | 0.22 | | DQ | 0.67 | 0.29 | 0.44 | 0.31 | 0.32 | 0.55 |
| DAX | NV | 26 | 27 | 28 | 27 | 25 | 25 | NK | NV | 22 | 14 | 15 | 16 | 16 | 10 |
| (26) | UC | 1 | 0.84 | 0.69 | 0.84 | 0.84 | 0.84 | (23) | UC | 0.83 | **0.04** | 0.07 | 0.11 | 0.11 | **0.01** |
| | IN | 0.09 | 0.69 | 0.66 | 0.72 | 0.12 | 0.83 | | IN | 0.14 | 0.35 | 0.33 | 0.30 | 0.30 | 0.50 |
| | CC | 0.23 | 0.86 | 0.79 | 0.87 | 0.27 | 0.91 | | CC | 0.30 | 0.07 | 0.12 | 0.16 | 0.16 | **0.01** |
| | DQ | 0.79 | 0.70 | 0.36 | 0.42 | 0.50 | 0.85 | | DQ | 0.35 | 0.55 | 0.53 | 0.64 | 0.63 | 0.21 |
| DJ | NV | 24 | 25 | 25 | 24 | 26 | 20 | RUS | NV | 21 | 22 | 21 | 23 | 21 | 18 |
| (25) | UC | 0.84 | 1 | 1 | 0.84 | 0.83 | 0.29 | (25) | UC | 0.39 | 0.53 | 0.39 | 0.68 | 0.40 | 0.13 |
| | IN | 0.13 | 0.51 | 0.51 | 0.45 | 0.58 | 0.82 | | IN | 0.17 | 0.15 | 0.17 | 0.14 | 0.17 | 0.24 |
| | CC | **0.01** | 0.77 | 0.77 | 0.70 | 0.79 | 0.53 | | CC | 0.27 | 0.28 | 0.27 | 0.29 | 0.27 | 0.16 |
| | DQ | 0.61 | 0.40 | 0.41 | 0.71 | 0.90 | 0.81 | | DQ | 0.72 | 0.78 | 0.86 | 0.85 | 0.67 | 0.54 |
| ESX | NV | 29 | 23 | 22 | 21 | 22 | 23 | SPX | NV | 21 | 20 | 21 | 22 | 22 | 20 |
| (25) | UC | 0.42 | 0.68 | 0.53 | 0.40 | 0.53 | 0.68 | (25) | UC | 0.40 | 0.28 | 0.40 | 0.53 | 0.53 | 0.29 |
| | IN | 0.54 | 0.14 | 0.16 | 0.18 | 0.16 | 0.14 | | IN | 0.17 | 0.82 | 0.90 | 0.97 | 0.97 | 0.82 |
| | CC | 0.57 | 0.30 | 0.30 | 0.28 | 0.30 | 0.29 | | CC | 0.27 | 0.53 | 0.67 | 0.78 | 0.78 | 0.53 |
| | DQ | 0.07 | 0.32 | 0.26 | 0.52 | 0.59 | 0.36 | | DQ | 0.73 | 0.89 | 0.89 | 0.93 | 0.92 | 0.87 |
| MIB | NV | 30 | 30 | 30 | 29 | 29 | 25 | SMI | NV | 26 | 29 | 29 | 30 | 29 | 27 |
| (25) | UC | 0.32 | 0.32 | 0.32 | 0.42 | 0.12 | 1.00 | (25) | UC | 0.83 | 0.42 | 0.42 | 0.32 | 0.42 | 0.69 |
| | IN | 0.05 | 0.05 | 0.05 | 0.06 | 0.06 | 0.11 | | IN | 0.10 | 0.76 | 0.58 | 0.52 | 0.57 | 0.71 |
| | CC | 0.09 | 0.09 | 0.09 | 0.12 | 0.12 | 0.26 | | CC | 0.24 | 0.65 | 0.59 | 0.46 | 0.59 | 0.81 |
| | DQ | **0.01** | **0.01** | **0.01** | **0.11** | **0.01** | 0.05 | | DQ | 0.10 | **0.01** | **0.01** | **0.01** | **0.01** | **0.01** |
| FT | NV | 25 | 27 | 26 | 25 | 26 | 23 | | | | | | | | |
| (25) | UC | 1 | 0.68 | 0.84 | 1 | 0.84 | 0.68 | | | | | | | | |
| | IN | 0.85 | 0.71 | 0.78 | 0.85 | 0.79 | 0.99 | | | | | | | | |
| | CC | 0.93 | 0.81 | 0.89 | 0.93 | 0.89 | 0.87 | | | | | | | | |
| | DQ | 0.38 | 0.11 | 0.58 | 0.67 | 0.13 | 0.34 | | | | | | | | |

**Table 20: Performance measures for the one-day-ahead VaR forecast during a period of turmoil with type-II link function**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | PM | (a) | (b) | ( c) | (d) | (e) | (f) | Index | PM | (a) | (b) | ( c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX | NV | 28 | 20 | 20 | 18 | 18 | 21 | IBX* | NV | 29 | 29 | 29 | 29 | 30 | 31 |
| (26) | UC | 0.69 | 0.21 | 0.21 | 0.09 | 0.09 | 0.30 | (26) | UC | 0.55 | 0.55 | 0.55 | 0.55 | 0.43 | 0.33 |
| | IN | 0.08 | 0.21 | 0.21 | 0.26 | 0.26 | 0.19 | | IN | 0.77 | 0.77 | 0.77 | 0.77 | 0.84 | 0.92 |
| | CC | 0.18 | 0.20 | 0.20 | 0.12 | 0.12 | 0.23 | | CC | 0.76 | 0.76 | 0.76 | 0.76 | 0.68 | 0.58 |
| | DQ | 0.64 | 0.57 | 0.57 | 0.68 | 0.67 | 0.72 | | DQ | 0.39 | 0.15 | 0.15 | 0.72 | 0.33 | 0.29 |
| AOI | NV | 17 | 21 | 21 | 21 | 21 | 20 | IPC | NV | 22 | 18 | 18 | 17 | 17 | 19 |
| (25) | UC | 0.08 | 0.40 | 0.40 | 0.40 | 0.40 | 0.29 | (25) | UC | 0.53 | 0.13 | 0.13 | 0.08 | 0.08 | 0.20 |
| | IN | 0.27 | 0.19 | 0.19 | 0.19 | 0.19 | 0.21 | | IN | 0.97 | 0.25 | 0.25 | 0.27 | 0.27 | 0.22 |
| | CC | 0.12 | 0.28 | 0.28 | 0.28 | 0.28 | 0.25 | | CC | 0.78 | 0.16 | 0.16 | 0.12 | 0.12 | 0.20 |
| | DQ | 0.42 | 0.57 | 0.57 | 0.57 | 0.56 | 0.49 | | DQ | 0.68 | 0.49 | 0.48 | 0.59 | 0.59 | 0.37 |
| BVP | NV | 19 | 19 | 19 | 19 | 19 | 19 | KCI* | NV | 18 | 21 | 21 | 23 | 22 | 21 |
| (24) | UC | 0.28 | 0.28 | 0.28 | 0.28 | 0.28 | 0.28 | (25) | UC | 0.13 | 0.40 | 0.40 | 0.68 | 0.53 | 0.40 |
| | IN | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | | IN | 0.68 | 0.28 | 0.28 | 0.10 | 0.33 | 0.28 |
| | CC | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | | CC | 0.28 | 0.38 | 0.38 | 0.23 | 0.49 | 0.38 |
| | DQ | 0.12 | 0.68 | 0.68 | 0.85 | 0.86 | 0.68 | | DQ | 0.47 | 0.18 | 0.17 | 0.03 | 0.30 | 0.22 |
| CAC* | NV | 24 | 24 | 23 | 22 | 23 | 22 | NSQ | NV | 20 | 21 | 22 | 22 | 22 | 23 |
| (26) | UC | 0.68 | 0.68 | 0.54 | 0.41 | 0.54 | 0.41 | (25) | UC | 0.29 | 0.40 | 0.53 | 0.53 | 0.53 | 0.68 |
| | IN | 0.13 | 0.92 | 0.99 | 0.94 | 0.99 | 0.94 | | IN | 0.20 | 0.89 | 0.97 | 0.16 | 0.16 | 0.95 |
| | CC | 0.28 | 0.87 | 0.79 | 0.68 | 0.79 | 0.68 | | CC | 0.24 | 0.67 | 0.78 | 0.29 | 0.29 | 0.87 |
| | DQ | 0.28 | 0.27 | 0.25 | 0.28 | 0.35 | 0.26 | | DQ | 0.67 | 0.20 | 0.33 | 0.27 | 0.27 | 0.37 |
| DAX | NV | 26 | 27 | 27 | 26 | 25 | 29 | NK | NV | 22 | 14 | 14 | 15 | 15 | 14 |
| (26) | UC | 1 | 0.84 | 0.84 | 1 | 0.84 | 0.55 | (23) | UC | 0.83 | **0.04** | **0.04** | 0.07 | 0.07 | **0.04** |
| | IN | 0.09 | 0.69 | 0.69 | 0.79 | 0.12 | 0.56 | | IN | 0.14 | 0.35 | 0.35 | 0.31 | 0.31 | 0.35 |
| | CC | 0.23 | 0.86 | 0.86 | 0.92 | 0.27 | 0.67 | | CC | 0.31 | 0.07 | 0.07 | 0.11 | 0.11 | 0.07 |
| | DQ | 0.80 | 0.68 | 0.68 | 0.46 | 0.48 | 0.89 | | DQ | 0.35 | 0.55 | 0.56 | 0.66 | 0.66 | 0.56 |
| DJ | NV | 24 | 25 | 26 | 25 | 26 | 26 | RUS | NV | 21 | 22 | 22 | 23 | 21 | 25 |
| (25) | UC | 0.84 | 1 | 0.84 | 1 | 0.84 | 0.84 | (25) | UC | 0.40 | 0.53 | 0.53 | 0.68 | 0.40 | 1 |
| | IN | 0.13 | 0.81 | 0.58 | 0.51 | 0.58 | 0.74 | | IN | 0.17 | 0.15 | 0.15 | 0.14 | 0.17 | 0.10 |
| | CC | **0.01** | 0.92 | 0.80 | 0.77 | 0.88 | 0.80 | | CC | 0.27 | 0.28 | 0.28 | 0.29 | 0.27 | 0.25 |
| | DQ | 0.61 | 0.94 | 0.80 | 0.84 | 0.90 | 0.98 | | DQ | 0.72 | 0.78 | 0.78 | 0.85 | 0.67 | 0.19 |
| ESX | NV | 29 | 23 | 23 | 23 | 23 | 24 | SPX | NV | 21 | 20 | 21 | 22 | 22 | 21 |
| (25) | UC | 0.42 | 0.68 | 0.68 | 0.68 | 0.68 | 0.84 | (25) | UC | 0.40 | 0.29 | 0.40 | 0.53 | 0.53 | 0.40 |
| | IN | 0.55 | 0.14 | 0.14 | 0.14 | 0.14 | 0.12 | | IN | 0.17 | 0.82 | 0.90 | 0.97 | 0.97 | 0.90 |
| | CC | 0.57 | 0.30 | 0.30 | 0.30 | 0.30 | 0.28 | | CC | 0.27 | 0.53 | 0.67 | 0.78 | 0.78 | 0.67 |
| | DQ | 0.08 | 0.32 | 0.32 | 0.32 | 0.31 | 0.20 | | DQ | 0.73 | 0.89 | 0.88 | 0.92 | 0.92 | 0.93 |
| MIB | NV | 30 | 30 | 30 | 29 | 30 | 32 | SMI | NV | 26 | 28 | 28 | 26 | 27 | 28 |
| (25) | UC | 0.32 | 0.32 | 0.32 | 0.42 | 0.32 | 0.17 | (25) | UC | 0.84 | 0.55 | 0.55 | 0.84 | 0.69 | 0.55 |
| | IN | 0.05 | 0.05 | 0.05 | 0.06 | 0.05 | 0.40 | | IN | 0.10 | 0.64 | 0.64 | 0.77 | 0.71 | 0.64 |
| | CC | 0.09 | 0.09 | 0.09 | 0.12 | 0.09 | 0.25 | | CC | 0.24 | 0.71 | 0.71 | 0.89 | 0.81 | 0.71 |
| | DQ | **0.02** | **0.01** | **0.01** | **0.01** | **0.01** | **0.01** | | DQ | 0.10 | **0.01** | **0.01** | 0.20 | **0.04** | **0.01** |
| FT | NV | 25 | 26 | 25 | 25 | 25 | 28 | | | | | | | | |
| (25) | UC | 1 | 0.84 | 1 | 1 | 1 | 0.54 | | | | | | | | |
| | IN | 0.85 | 0.78 | 0.85 | 0.85 | 0.85 | 0.68 | | | | | | | | |
| | CC | 0.93 | 0.89 | 0.93 | 0.93 | 0.93 | 0.72 | | | | | | | | |
| | DQ | 0.38 | 0.11 | 0.62 | 0.11 | 0.11 | 0.13 | | | | | | | | |

**Table 21: Performance measures for the one-day-ahead VaR forecast during a period of turmoil with type-III link function**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | PM | (a) | (b) | ( c) | (d) | (e) | (f) | Index | PM | (a) | (b) | ( c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX | NV | 28 | 20 | 20 | 18 | 18 | 21 | IBX* | NV | 29 | 29 | 29 | 29 | 30 | 31 |
| (26) | UC | 0.69 | 0.21 | 0.21 | 0.09 | 0.09 | 0.30 | (26) | UC | 0.55 | 0.55 | 0.55 | 0.55 | 0.43 | 0.33 |
| | IN | 0.08 | 0.21 | 0.21 | 0.26 | 0.26 | 0.18 | | IN | 0.77 | 0.77 | 0.77 | 0.77 | 0.84 | 0.91 |
| | CC | 0.18 | 0.20 | 0.20 | 0.12 | 0.12 | 0.23 | | CC | 0.76 | 0.76 | 0.76 | 0.76 | 0.68 | 0.58 |
| | DQ | 0.64 | 0.57 | 0.57 | 0.68 | 0.67 | 0.73 | | DQ | 0.39 | 0.15 | 0.15 | 0.72 | 0.33 | 0.28 |
| AOI | NV | 17 | 21 | 21 | 21 | 21 | 20 | IPC | NV | 22 | 18 | 18 | 17 | 17 | 19 |
| (25) | UC | 0.08 | 0.40 | 0.40 | 0.40 | 0.40 | 0.29 | (25) | UC | 0.53 | 0.13 | 0.13 | 0.08 | 0.08 | 0.20 |
| | IN | 0.27 | 0.19 | 0.19 | 0.19 | 0.19 | 0.21 | | IN | 0.97 | 0.25 | 0.25 | 0.27 | 0.27 | 0.22 |
| | CC | 0.12 | 0.28 | 0.28 | 0.28 | 0.28 | 0.25 | | CC | 0.78 | 0.16 | 0.16 | 0.12 | 0.12 | 0.20 |
| | DQ | 0.42 | 0.57 | 0.57 | 0.57 | 0.56 | 0.49 | | DQ | 0.68 | 0.49 | 0.48 | 0.59 | 0.59 | 0.37 |
| BVP | NV | 19 | 19 | 19 | 19 | 19 | 19 | KCI* | NV | 18 | 21 | 21 | 23 | 22 | 21 |
| (24) | UC | 0.28 | 0.28 | 0.28 | 0.28 | 0.28 | 0.28 | (25) | UC | 0.13 | 0.40 | 0.40 | 0.68 | 0.53 | 0.40 |
| | IN | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | 0.77 | | IN | 0.68 | 0.28 | 0.28 | 0.10 | 0.33 | 0.28 |
| | CC | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | | CC | 0.28 | 0.38 | 0.38 | 0.23 | 0.49 | 0.37 |
| | DQ | 0.12 | 0.68 | 0.68 | 0.85 | 0.86 | 0.68 | | DQ | 0.47 | 0.18 | 0.17 | 0.03 | 0.30 | 0.21 |
| CAC* | NV | 24 | 24 | 23 | 22 | 23 | 22 | NSQ | NV | 20 | 21 | 22 | 22 | 22 | 23 |
| (26) | UC | 0.68 | 0.68 | 0.54 | 0.41 | 0.54 | 0.41 | (25) | UC | 0.29 | 0.40 | 0.53 | 0.53 | 0.53 | 0.68 |
| | IN | 0.13 | 0.92 | 0.99 | 0.94 | 0.99 | 0.94 | | IN | 0.20 | 0.89 | 0.97 | 0.16 | 0.16 | 0.95 |
| | CC | 0.28 | 0.87 | 0.79 | 0.68 | 0.79 | 0.68 | | CC | 0.24 | 0.67 | 0.78 | 0.29 | 0.29 | 0.87 |
| | DQ | 0.28 | 0.27 | 0.25 | 0.28 | 0.35 | 0.26 | | DQ | 0.67 | 0.20 | 0.33 | 0.27 | 0.27 | 0.37 |
| DAX | NV | 26 | 27 | 27 | 26 | 25 | 29 | NK | NV | 22 | 14 | 14 | 15 | 15 | 14 |
| (26) | UC | 1 | 0.84 | 0.84 | 1.00 | 0.84 | 0.55 | (23) | UC | 0.83 | **0.04** | **0.04** | 0.07 | 0.07 | **0.04** |
| | IN | 0.09 | 0.69 | 0.69 | 0.79 | 0.86 | 0.56 | | IN | 0.14 | 0.35 | 0.35 | 0.31 | 0.31 | 0.35 |
| | CC | 0.23 | 0.86 | 0.86 | 0.92 | 0.92 | 0.67 | | CC | 0.31 | 0.07 | 0.07 | 0.11 | 0.11 | 0.07 |
| | DQ | 0.80 | 0.68 | 0.68 | 0.47 | 0.76 | 0.89 | | DQ | 0.35 | 0.55 | 0.56 | 0.66 | 0.66 | 0.56 |
| DJ | NV | 24 | 25 | 26 | 25 | 26 | 26 | RUS | NV | 21 | 22 | 22 | 23 | 21 | 25 |
| (25) | UC | 0.84 | 1 | 0.84 | 1 | 0.84 | 0.84 | (25) | UC | 0.40 | 0.53 | 0.53 | 0.68 | 0.40 | 1 |
| | IN | 0.13 | 0.81 | 0.58 | 0.51 | 0.58 | 0.74 | | IN | 0.17 | 0.15 | 0.15 | 0.14 | 0.17 | 0.10 |
| | CC | **0.01** | 0.92 | 0.80 | 0.77 | 0.88 | 0.80 | | CC | 0.27 | 0.28 | 0.28 | 0.29 | 0.27 | 0.25 |
| | DQ | 0.61 | 0.94 | 0.80 | 0.84 | 0.90 | 0.98 | | DQ | 0.72 | 0.78 | 0.78 | 0.85 | 0.67 | 0.19 |
| ESX | NV | 29 | 23 | 23 | 23 | 23 | 24 | SPX | NV | 21 | 20 | 21 | 22 | 22 | 21 |
| (25) | UC | 0.42 | 0.68 | 0.68 | 0.68 | 0.68 | 0.84 | (25) | UC | 0.40 | 0.29 | 0.40 | 0.53 | 0.53 | 0.40 |
| | IN | 0.55 | 0.14 | 0.14 | 0.14 | 0.14 | 0.12 | | IN | 0.17 | 0.82 | 0.90 | 0.97 | 0.97 | 0.89 |
| | CC | 0.57 | 0.30 | 0.30 | 0.30 | 0.30 | 0.28 | | CC | 0.27 | 0.53 | 0.67 | 0.78 | 0.78 | 0.66 |
| | DQ | 0.08 | 0.32 | 0.32 | 0.32 | 0.31 | 0.20 | | DQ | 0.73 | 0.89 | 0.88 | 0.92 | 0.92 | 0.93 |
| MIB | NV | 30 | 30 | 30 | 29 | 30 | 32 | SMI | NV | 26 | 28 | 28 | 26 | 27 | 28 |
| (25) | UC | 0.32 | 0.32 | 0.32 | 0.42 | 0.32 | 0.17 | (25) | UC | 0.84 | 0.55 | 0.55 | 0.84 | 0.69 | 0.55 |
| | IN | 0.05 | 0.05 | 0.05 | 0.06 | 0.05 | 0.41 | | IN | 0.10 | 0.64 | 0.64 | 0.77 | 0.71 | 0.64 |
| | CC | 0.09 | 0.09 | 0.09 | 0.12 | 0.09 | 0.25 | | CC | 0.24 | 0.71 | 0.71 | 0.89 | 0.81 | 0.71 |
| | DQ | **0.02** | **0.01** | **0.01** | **0.01** | **0.01** | **0.01** | | DQ | 0.10 | **0.01** | **0.01** | 0.20 | **0.04** | **0.01** |
| FT | NV | 25 | 26 | 25 | 25 | 25 | 28 | | | | | | | | |
| (25) | UC | 1 | 0.84 | 1 | 1 | 1 | 0.55 | | | | | | | | |
| | IN | 0.85 | 0.78 | 0.85 | 0.85 | 0.85 | 0.68 | | | | | | | | |
| | CC | 0.93 | 0.89 | 0.93 | 0.93 | 0.93 | 0.72 | | | | | | | | |
| | DQ | 0.38 | 0.11 | 0.62 | 0.11 | 0.11 | 0.13 | | | | | | | | |

**Table 22: Performance of one-day-ahead ES forecast during a period of turmoil with type-I link function**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.748 | 0.721 | 0.815 | 0.772 | 0.625 | 0.565 |
| AOI | 0.214 | 0.734 | 0.537 | 0.695 | 0.832 | 0.812 |
| BVP | 0.385 | 0.394 | 0.639 | 0.576 | 0.545 | 0.868 |
| CAC | **0.013** | 0.063 | 0.069 | 0.151 | 0.23 | 0.442 |
| DAX | 0.633 | 0.546 | 0.703 | 0.681 | 0.644 | 0.919 |
| DJ | 0.154 | 0.586 | 0.716 | 0.947 | 0.998 | 0.758 |
| ESX | 0.760 | 0.939 | 0.87 | 0.801 | 0.93 | 0.975 |
| MIB | 0.234 | 0.400 | 0.381 | 0.220 | 0.326 | 0.505 |
| FT | 0.186 | 0.253 | 0.180 | 0.330 | 0.371 | 0.432 |
| IBX | 0.126 | 0.428 | 0.421 | 0.590 | 0.559 | 0.283 |
| IPC | 0.048 | 0.632 | 0.565 | 0.315 | 0.267 | 0.651 |
| KCI | **0.001** | 0.096 | 0.081 | 0.048 | 0.068 | **0.02** |
| NSQ | 0.409 | 0.998 | 0.999 | 0.999 | 0.994 | 0.994 |
| NK | 0.852 | 0.524 | 0.765 | 0.85 | 0.839 | 0.181 |
| RUS | 0.348 | 0.933 | 0.905 | 0.924 | 0.924 | 0.736 |
| SPX | 0.075 | 0.158 | 0.279 | 0.695 | 0.832 | 0.624 |
| SMI | 0.447 | 0.052 | 0.085 | 0.253 | 0.184 | 0.142 |

**Table 23: Performance of one-day-ahead ES forecast during a period of turmoil with type-II link function**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.706 | 0.616 | 0.724 | 0.652 | 0.634 | 0.511 |
| AOI | 0.223 | 0.650 | 0.717 | 0.643 | 0.797 | 0.175 |
| BVP | 0.372 | 0.670 | 0.655 | 0.565 | 0.545 | 0.672 |
| CAC* | **0.009** | 0.109 | 0.073 | 0.149 | 0.244 | **0.032** |
| DAX | 0.642 | 0.470 | 0.410 | 0.434 | 0.441 | 0.640 |
| DJ | 0.156 | 0.743 | 0.843 | 0.963 | 0.988 | 0.885 |
| ESX | 0.750 | 0.932 | 0.904 | 0.918 | 0.944 | 0.678 |
| MIB | 0.272 | 0.214 | 0.189 | 0.134 | 0.242 | 0.461 |
| FT | 0.223 | 0.274 | 0.191 | 0.313 | 0.381 | 0.354 |
| IBX* | 0.131 | 0.452 | 0.421 | 0.574 | 0.560 | 0.354 |
| IPC | 0.067 | 0.656 | 0.555 | 0.293 | 0.264 | 0.315 |
| KCI* | **0.001** | 0.081 | 0.058 | 0.051 | 0.081 | 0.078 |
| NSQ | 0.421 | 0.991 | 0.992 | 0.996 | 0.998 | 0.993 |
| NK | 0.879 | 0.625 | 0.614 | 0.710 | 0.667 | 0.580 |
| RUS | 0.342 | 0.874 | 0.915 | 0.865 | 0.910 | 0.832 |
| SPX | 0.072 | 0.062 | 0.135 | 0.563 | 0.680 | 0.287 |
| SMI | 0.488 | 0.186 | 0.201 | 0.095 | 0.155 | 0.390 |

**Table 24: Performance of one-day-ahead ES forecast during a period of turmoil with type-III link function**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.724 | 0.617 | 0.724 | 0.677 | 0.615 | 0.519 |
| AOI | 0.235 | 0.644 | 0.697 | 0.651 | 0.793 | 0.155 |
| BVP | 0.352 | 0.672 | 0.633 | 0.565 | 0.508 | 0.655 |
| CAC* | **0.020** | 0.115 | 0.058 | 0.103 | 0.274 | **0.026** |
| DAX | 0.654 | 0.493 | 0.418 | 0.412 | 0.468 | 0.628 |
| DJ | 0.165 | 0.761 | 0.838 | 0.971 | 0.993 | 0.896 |
| ESX | 0.767 | 0.928 | 0.908 | 0.916 | 0.944 | 0.743 |
| MIB | 0.257 | 0.192 | 0.197 | 0.160 | 0.258 | 0.069 |
| FT | 0.199 | 0.258 | 0.183 | 0.294 | 0.362 | 0.475 |
| IBX* | 0.125 | 0.458 | 0.418 | 0.548 | 0.539 | 0.348 |
| IPC | 0.050 | 0.619 | 0.548 | 0.298 | 0.273 | 0.333 |
| KCI* | **0.003** | 0.076 | 0.082 | 0.051 | 0.063 | 0.066 |
| NSQ | 0.481 | 0.994 | 0.996 | 0.997 | 0.994 | 0.989 |
| NK | 0.860 | 0.567 | 0.643 | 0.713 | 0.696 | 0.580 |
| RUS | 0.323 | 0.881 | 0.905 | 0.877 | 0.905 | 0.809 |
| SPX | 0.075 | 0.070 | 0.128 | 0.587 | 0.670 | 0.299 |
| SMI | 0.454 | 0.191 | 0.186 | 0.102 | 0.140 | 0.230 |

**Table 25: Performance measures for the ten-day-ahead VaR forecast during a period of turmoil with type-I link function**. Actual and expected number of violations of the VaR. Cases in which the null hypothesis of correct unconditional coverage is rejected at level $\alpha = 10\%$ are in bold. In *italics*, the best performer. The * indicates series for which a higher threshold was required for appropriate GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 25 | *26 (0.63)* | **9 (0.03)** | **9 (0.03)** | **9 (0.03)** | **9 (0.03)** | **12 (0.10)** |
| AOI | 25 | *22 (0.69)* | *18 (0.57)* | *18 (0.57)* | *18 (0.57)* | *18 (0.57)* | 17 (0.46) |
| BVP | 24 | **10 (0.07)** | *14 (0.31)* | *14 (0.31)* | *14 (0.31)* | *14 (0.31)* | 13 (0.24) |
| CAC* | 25 | 33 (0.58) | *19 (0.50)* | *19 (0.50)* | *19 (0.50)* | *19 (0.50)* | *19 (0.54)* |
| DAX | 25 | *32 (0.81)* | **14 (0.02)** | **13 (0.01)** | **14 (0.02)** | **14 (0.02)** | 14 (0.27) |
| DJ | 25 | **12 (0.06)** | 19 (0.61) | 19 (0.61) | 19 (0.61) | 19 (0.61) | *20 (0.64)* |
| ESX | 25 | *21 (0.78)* | 12 (0.22) | 12 (0.22) | 12 (0.22) | 12 (0.22) | 10 (0.14) |
| MIB | 25 | 39 (0.64) | *32 (0.96)* | *32 (0.96)* | *32 (0.96)* | *32 (0.96)* | 34 (0.48) |
| FT | 25 | 43 (0.40) | 17 (0.47) | 17 (0.47) | *18 (0.49)* | *18 (0.49)* | 17 (0.49) |
| IBX | 25 | 37 (0.38) | 26 (0.93) | 26 (0.93) | 26 (0.93) | 26 (0.93) | *25 (0.97)* |
| IPC | 25 | *29 (0.99)* | **12 (0.02)** | **12 (0.02)** | **13 (0.02)** | **13 (0.02)** | **13 (0.07)** |
| KCI | 24 | **11 (0.07)** | 17 (0.41) | *18 (0.50)* | *18 (0.50)* | *18 (0.50)* | *18 (0.47)* |
| NSQ | 25 | *23 (0.83)* | 21 (0.77) | 21 (0.77) | 21 (0.77) | 20 (0.77) | 21 (0.71) |
| NK | 23 | **5 (0.003)** | **2 (0.001)** | **2 (0.001)** | **2 (0.001)** | **2 (0.001)** | **2 (0.001)** |
| RUS | 25 | *19 (0.47)* | 13 (0.14) | 13 (0.14) | 14 (0.19) | 14 (0.19) | 13 (0.14) |
| SPX | 25 | 15 (0.14) | 18 (0.52) | 18 (0.52) | *19 (0.61)* | *19 (0.61)* | *19 (0.55)* |
| SMI* | 25 | *27 (0.88)* | 22 (0.63) | 22 (0.63) | 21 (0.53) | 21 (0.53) | 20 (0.64) |

**Table 26: Performance measures for the ten-day-ahead VaR forecast during a period of turmoil with type-II link function**. Actual and expected number of violations of the VaR. Cases in which the null hypothesis of correct unconditional coverage is rejected at level $\alpha = 10\%$ are in bold. In *italics*, the best performer. The * indicates series for which a higher threshold was required for appropriate GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 25 | *26 (0.63)* | **9 (0.021)** | **9 (0.02)** | **9 (0.02)** | **10 (0.04)** | **12 (0.10)** |
| AOI | 25 | *22 (0.69)* | 12 (0.17) | 12 (0.17) | 12 (0.17) | 13 (0.23) | 14 (0.25) |
| BVP* | 24 | **10 (0.06)** | 9 (0.15) | **8 (0.09)** | **8 (0.09)** | **8 (0.09)** | *10 (0.14)* |
| CAC* | 25 | 33 (0.58) | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* |
| DAX | 25 | *32 (0.81)* | **14 (0.02)** | **13 (0.02)** | **13 (0.02)** | **14 (0.03)** | 14 (0.27) |
| DJ | 25 | **12 (0.06)** | 21 (0.73) | 21 (0.73) | *24 (0.95)* | 23 (0.88) | 22 (0.81) |
| ESX | 25 | *21 (0.78)* | 12 (0.21) | 12 (0.21) | 12 (0.21) | 12 (0.21) | 10 (0.14) |
| MIB | 25 | 39 (0.64) | 32 (0.59) | 32 (0.59) | 32 (0.59) | *31 (0.63)* | 34 (0.48) |
| FT | 25 | 43 (0.40) | *17 (0.47)* | *17 (0.47)* | 15 (0.40) | *17 (0.49)* | 16 (0.42) |
| IBX | 25 | 37 (0.38) | 26 (0.96) | 26 (0.96) | 26 (0.96) | 26 (0.96) | *25 (0.97)* |
| IPC | 25 | *29 (0.99)* | **12 (0.07)** | **12 (0.07)** | **12 (0.07)** | **13 (0.09)** | **12 (0.07)** |
| KCI | 24 | **11 (0.07)** | 17 (0.39) | *18 (0.47)* | *18 (0.47)* | *18 (0.47)* | *18 (0.47)* |
| NSQ | 25 | *23 (0.83)* | 19 (0.61) | 19 (0.61) | 19 (0.61) | 20 (0.66) | 20 (0.63) |
| NK | 22 | **5 (0.003)** | **2 (0.001)** | **2 (0.001)** | **3 (0.001)** | **2 (0.001)** | **3 (0.001)** |
| RUS | 25 | *19 (0.47)* | 14 (0.18) | 13 (0.13) | 15 (0.24) | 15 (0.24) | 13 (0.15) |
| SPX* | 25 | 15 (0.14) | *23 (0.87)* | 22 (0.78) | 22 (0.78) | 22 (0.78) | 22 (0.78) |
| SMI* | 25 | *27 (0.88)* | 22 (0.83) | 22 (0.83) | 22 (0.81) | 21 (0.73) | 20 (0.65) |

**Table 27: Performance measures for the ten-day-ahead VaR forecast during a period of turmoil with type-III link function**. Actual and expected number of violations of the VaR. Cases in which the null hypothesis of correct unconditional coverage is rejected at level $\alpha = 10\%$ are in bold. In *italics*, the best performer. The * indicates series for which a higher threshold was required for appropriate GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 26 | *31 (0.63)* | **9 (0.02)** | **9 (0.02)** | **9 (0.02)** | **10 (0.04)** | **11 (0.09)** |
| AOI | 25 | *22 (0.69)* | 13(0.21) | 13 (0.21) | 12 (0.17) | 13 (0.23) | 14 (0.25) |
| BVP* | 24 | **10 (0.06)** | 8 (0.16) | **8 (0.09)** | **8 (0.09)** | **8 (0.09)** | 10 (0.14) |
| CAC* | 25 | 33 (0.58) | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* |
| DAX | 25 | *32 (0.81)* | 14 (0.27) | 13 (0.25) | 13 (0.25) | 14 (0.27) | 14 (0.27) |
| DJ | 25 | **12 (0.06)** | 21 (0.73) | 21 (0.73) | *24 (0.95)* | 23 (0.88) | 22 (0.81) |
| ESX | 25 | *21 (0.78)* | 12 (0.21) | 12 (0.21) | 12 (0.21) | 12 (0.21) | 10 (0.14) |
| MIB | 25 | 39 (0.64) | 32 (0.59) | 32 (0.59) | 32 (0.59) | *31 (0.63)* | 34 (0.48) |
| FT | 25 | 43 (0.40) | *17 (0.47)* | *17 (0.47)* | 15 (0.40) | *17 (0.49)* | 16 (0.43) |
| IBX | 25 | 37 (0.38) | 26 (0.96) | 26 (0.96) | 26 (0.96) | 26 (0.96) | *25 (0.97)* |
| IPC | 25 | *29 (0.99)* | **12 (0.07)** | **12 (0.07)** | **12 (0.07)** | **13 (0.09)** | **12 (0.07)** |
| KCI | 24 | **11 (0.07)** | 17 (0.39) | *18 (0.47)* | *18 (0.47)* | *18 (0.47)* | *18 (0.47)* |
| NSQ | 25 | *23 (0.83)* | 19 (0.61) | 19 (0.61) | 19 (0.61) | 20 (0.66) | 20 (0.64) |
| NK | 22 | **5 (0.003)** | **2 (0.001)** | **2 (0.001)** | **3 (0.001)** | **2 (0.001)** | **3 (0.001)** |
| RUS | 25 | *19 (0.47)* | 14 (0.18) | 13 (0.13) | 14 (0.18) | 15 (0.24) | 13 (0.45) |
| SPX* | 25 | 15 (0.14) | *23 (0.87)* | 22 (0.78) | 22 (0.78) | 22 (0.78) | 22 (0.78) |
| SMI* | 25 | *27 (0.88)* | 22 (0.83) | 22 (0.83) | 22 (0.81) | 21 (0.73) | 20 (0.65) |

**Table 28: Performance of ten-day-ahead ES forecast during a period of turmoil with type-I link function**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.954 | 0.999 | 1.000 | 1.000 | 0.992 | 0.814 |
| AOI* | 1.000 | 0.192 | 0.175 | 0.227 | 0.180 | 0.239 |
| BVP | 0.887 | 0.938 | 0.951 | 0.959 | 0.955 | 0.924 |
| CAC | 0.385 | 0.275 | 0.305 | 0.309 | 0.318 | 0.225 |
| DAX | 0.623 | 0.937 | 0.892 | 0.927 | 0.939 | 0.753 |
| DJ | 0.971 | 0.599 | 0.572 | 0.506 | 0.503 | 0.877 |
| ESX | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.997 |
| MIB | **0.001** | **0.001** | **0.002** | **0.002** | **0.002** | **0.017** |
| FT | 0.177 | **0.032** | **0.031** | **0.044** | **0.044** | **0.02** |
| IBX | 0.564 | 0.349 | 0.327 | 0.331 | 0.318 | 0.317 |
| IPC | 0.567 | 0.593 | 0.594 | 0.600 | 0.697 | 0.546 |
| KCI* | 1.000 | 0.964 | 0.976 | 0.971 | 0.976 | 0.965 |
| NSQ | 0.820 | 0.495 | 0.427 | 0.476 | 0.476 | 0.377 |
| NK | 0.924 | 1.000 | 1.000 | 1.000 | 1.000 | 0.778 |
| RUS | 0.997 | 0.888 | 0.901 | 0.882 | 0.890 | 0.881 |
| SPX | 0.996 | 0.553 | 0.540 | 0.769 | 0.745 | 0.751 |
| SMI | 0.054 | 0.078 | 0.086 | 0.073 | 0.128 | 0.199 |

**Table 29: Performance of ten-day-ahead ES forecast during a period of turmoil with type-II link function**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.954 | 1.000 | 1.000 | 0.988 | 1.000 | 0.832 |
| AOI* | 1.000 | 0.139 | 0.128 | 0.149 | 0.149 | 0.362 |
| BVP | 0.887 | 0.860 | 0.881 | 0.874 | 0.874 | 0.832 |
| CAC | 0.385 | 0.277 | 0.295 | 0.273 | 0.288 | 0.207 |
| DAX | 0.623 | 0.845 | 0.746 | 0.741 | 0.826 | 0.720 |
| DJ | 0.971 | 0.707 | 0.623 | 0.856 | 0.778 | 0.757 |
| ESX | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| MIB | **0.001** | **0.001** | **0.001** | **0.001** | **0.001** | **0.028** |
| FT | 0.177 | **0.034** | **0.029** | **0.019** | **0.039** | **0.008** |
| IBX | 0.564 | 0.314 | 0.348 | 0.358 | 0.388 | 0.314 |
| IPC | 0.567 | 0.421 | 0.384 | 0.363 | 0.543 | 0.500 |
| KCI* | 1.000 | 0.959 | 0.983 | 0.980 | 0.971 | 0.969 |
| NSQ | 0.820 | 0.555 | 0.516 | 0.554 | 0.768 | 0.353 |
| NK | 0.924 | 1.000 | 1.000 | 0.896 | 1.000 | 0.773 |
| RUS | 0.997 | 0.920 | 0.932 | 0.922 | 0.928 | 0.929 |
| SPX | 0.996 | 0.962 | 0.941 | 0.905 | 0.923 | 0.915 |
| SMI | 0.054 | 0.066 | 0.076 | 0.104 | 0.080 | 0.126 |

**Table 30: Performance of ten-day-ahead ES forecast during a period of turmoil with type-III link function**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.954 | 1.000 | 1.000 | 0.986 | 1.000 | 0.818 |
| AOI* | 1.000 | 0.153 | 0.173 | 0.184 | 0.084 | 0.357 |
| BVP | 0.887 | 0.856 | 0.901 | 0.863 | 0.873 | 0.835 |
| CAC | 0.385 | 0.281 | 0.300 | 0.261 | 0.276 | 0.203 |
| DAX | 0.623 | 0.823 | 0.743 | 0.760 | 0.828 | 0.717 |
| DJ | 0.971 | 0.682 | 0.603 | 0.834 | 0.777 | 0.782 |
| ESX | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 |
| MIB | **0.001** | **0.001** | **0.001** | **0.001** | **0.001** | **0.016** |
| FT | 0.177 | **0.026** | **0.027** | **0.027** | **0.030** | **0.013** |
| IBX | 0.564 | 0.328 | 0.317 | 0.379 | 0.402 | 0.317 |
| IPC | 0.567 | 0.406 | 0.407 | 0.377 | 0.542 | 0.510 |
| KCI* | 1.000 | 0.968 | 0.975 | 0.978 | 0.964 | 0.970 |
| NSQ | 0.820 | 0.561 | 0.509 | 0.539 | 0.752 | 0.392 |
| NK | 0.924 | 1.000 | 1.000 | 0.877 | 1.000 | 0.744 |
| RUS | 0.997 | 0.961 | 0.955 | 0.929 | 0.965 | 0.943 |
| SPX | 0.996 | 0.947 | 0.932 | 0.919 | 0.912 | 0.930 |
| SMI | 0.054 | 0.074 | 0.075 | 0.100 | 0.113 | 0.107 |

## B.5 Microstructure noise

In this section, we present the results obtained with the sub-sampled realized measures and the realized kernel. Given results in Appendixes B.3 and B.4, we report only the results for the type-I link function. Figures 40-43 show extremal index estimates when using sub-sampled realized measures and the realized kernel for both the full sample and the period of turmoil only. We find slightly better filtering for some series, slightly worse for others. Results are presented for one-day-ahead VaR (Tables 31-34), one-day-ahead ES (Tables 35-38), ten-day-ahead VaR (Tables 39-42) and ten-day-ahead ES (Tables 43-46). In summary, the results from both the full and restricted samples are in line with those previously obtained. Therefore, according to this analysis we conclude that the microstructure noise does not seem to affect our results.

**Figure 40: Extremal index estimates with sub-sampled realized measures**. Values of $\theta$ estimated on the standardized residuals obtained from the different windows of length $S = 2000$.

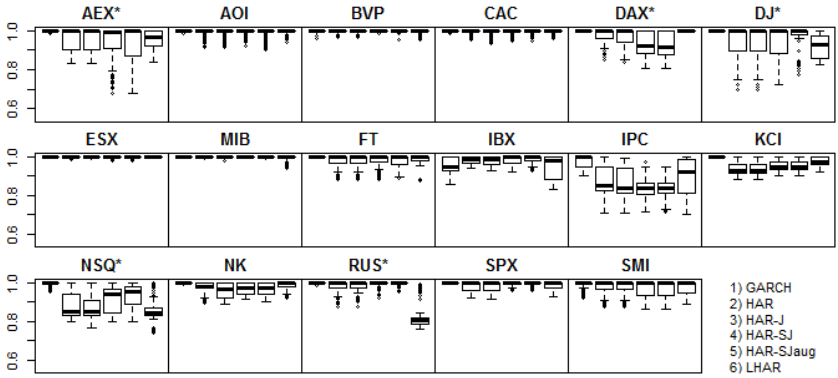


**Figure 41: Extremal index estimates with realized kernel**. Values of $\theta$ estimated on the standardized residuals obtained from the different windows of length $S = 2000$.

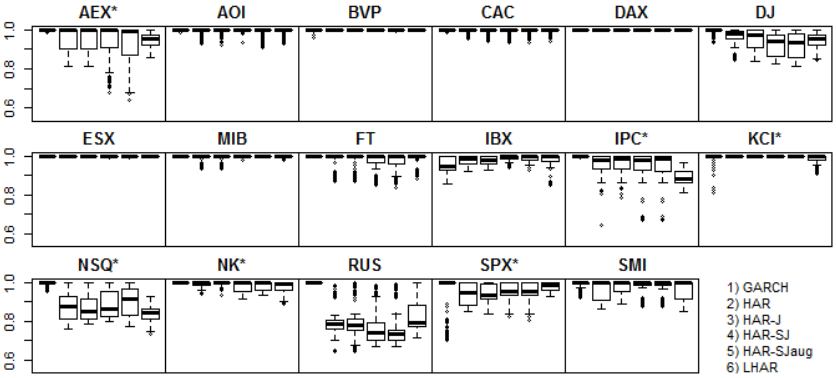



**Figure 42: Extremal index estimates during a period of turmoil with subsampled realized measures**. Values of $\theta$ estimated on the standardized residuals obtained from the different windows of length $S = 500$.

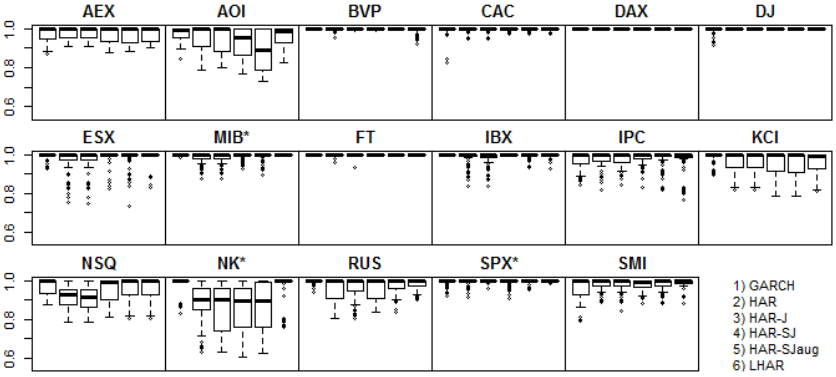


**Figure 43: Extremal index estimates during a period of turmoil with realized kernel**. Values of $\theta$ estimated on the standardized residuals obtained from the different windows of length $S = 500$.

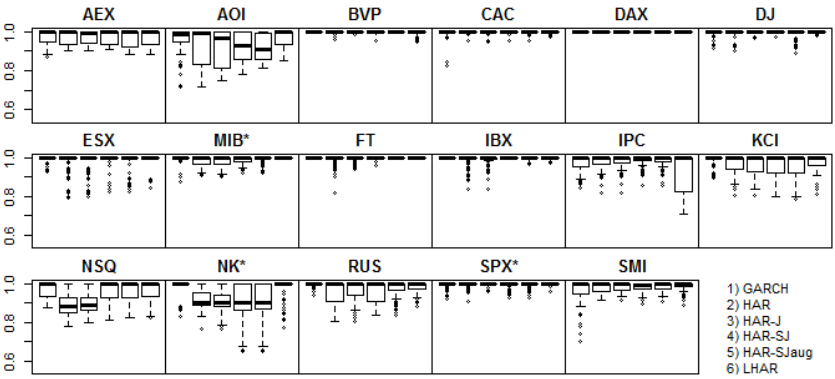


**Table 31: Performance measures for the one-day-ahead VaR forecast with sub-sampled realized measures**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | PM | (a) | (b) | (c) | (d) | (e) | (f) | Index | PM | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX* | NV | 22 | 18 | 18 | 18 | 20 | 18 | IBX | NV | 26 | 19 | 19 | 20 | 22 | 20 |
| (18) | UC | 0.35 | 1 | 1 | 1 | 0.64 | 1 | (18) | UC | 0.07 | 0.81 | 0.81 | 0.64 | 0.36 | 0.64 |
| | IN | 0.46 | 0.55 | 0.55 | 0.55 | 0.51 | 0.55 | | IN | 0.38 | 0.20 | 0.20 | 0.22 | 0.27 | 0.50 |
| | CC | 0.50 | 0.83 | 0.83 | 0.83 | 0.71 | 0.83 | | CC | 0.14 | 0.42 | 0.42 | 0.42 | 0.36 | 0.71 |
| | DQ | 0.65 | 0.72 | 0.73 | 0.73 | 0.17 | 0.72 | | DQ | **0.003** | 0.12 | 0.13 | **0.001** | **0.001** | 0.55 |
| AOI | NV | 16 | 15 | 16 | 16 | 17 | 11 | IPC | NV | 13 | 10 | 11 | 10 | 11 | 9 |
| (17) | UC | 0.80 | 0.62 | 0.81 | 0.81 | 1 | 0.12 | (17) | UC | 0.31 | 0.06 | 0.12 | 0.06 | 0.12 | **0.03** |
| | IN | 0.58 | 0.61 | 0.58 | 0.59 | 0.56 | 0.71 | | IN | 0.66 | 0.73 | 0.71 | 0.73 | 0.71 | 0.76 |
| | CC | 0.82 | 0.77 | 0.83 | 0.83 | 0.84 | 0.27 | | CC | 0.54 | 0.17 | 0.27 | 0.17 | 0.27 | 0.09 |
| | DQ | 0.99 | 0.53 | 0.001 | 0.61 | **0.002** | 0.12 | | DQ | 0.27 | 0.84 | 0.92 | 0.85 | 0.92 | 0.73 |
| BVP | NV | 13 | 11 | 11 | 12 | 13 | 13 | KCI | NV | 13 | 21 | 21 | 19 | 20 | 16 |
| (17) | UC | 0.31 | 0.12 | 0.12 | 0.20 | 0.31 | 0.31 | (17) | UC | 0.31 | 0.35 | 0.35 | 0.63 | 0.48 | 0.80 |
| | IN | 0.65 | 0.70 | 0.70 | 0.68 | 0.65 | 0.65 | | IN | 0.65 | 0.47 | 0.47 | 0.21 | 0.49 | 0.58 |
| | CC | 0.53 | 0.27 | 0.27 | 0.40 | 0.53 | 0.53 | | CC | 0.53 | 0.49 | 0.49 | 0.40 | 0.60 | 0.82 |
| | DQ | 0.25 | 0.17 | 0.17 | 0.27 | 0.05 | 0.28 | | DQ | 0.97 | 0.69 | 0.69 | 0.14 | 0.73 | 0.93 |
| CAC | NV | 29 | 27 | 27 | 26 | 26 | 24 | NSQ* | NV | 19 | 17 | 17 | 18 | 19 | 20 |
| (18) | UC | **0.01** | 0.08 | 0.05 | 0.08 | 0.08 | 0.18 | (17) | UC | 0.63 | 1 | 1 | 0.81 | 0.63 | 0.48 |
| | IN | 0.33 | 0.37 | 0.37 | 0.38 | 0.38 | 0.42 | | IN | 0.52 | 0.56 | 0.56 | 0.54 | 0.52 | 0.50 |
| | CC | **0.03** | 0.09 | 0.09 | 0.14 | 0.14 | 0.29 | | CC | 0.72 | 0.84 | 0.84 | 0.80 | 0.71 | 0.61 |
| | DQ | **0.02** | 0.07 | 0.07 | 0.09 | 0.09 | 0.10 | | DQ | 0.15 | 0.15 | 0.15 | 0.16 | 0.14 | 0.18 |
| DAX* | NV | 19 | 17 | 17 | 13 | 15 | 13 | NK | NV | 15 | 16 | 16 | 16 | 16 | 14 |
| (18) | UC | 0.81 | 0.81 | 0.81 | 0.21 | 0.46 | 0.21 | (16) | UC | 0.80 | 1 | 1 | 1 | 1 | 0.61 |
| | IN | 0.52 | 0.57 | 0.57 | 0.66 | 0.61 | 0.66 | | IN | 0.13 | 0.15 | 0.15 | 0.15 | 0.15 | 0.11 |
| | CC | 0.78 | 0.82 | 0.81 | 0.41 | 0.67 | 0.41 | | CC | 0.30 | 0.34 | 0.34 | 0.34 | 0.35 | 0.24 |
| | DQ | 0.16 | 0.14 | 0.14 | 0.21 | 0.37 | 0.80 | | DQ | 0.50 | **0.008** | **0.009** | **0.01** | **0.009** | **0.003** |
| DJ* | NV | 21 | 20 | 20 | 19 | 21 | 17 | RUS* | NV | 23 | 20 | 20 | 22 | 22 | 15 |
| (17) | UC | 0.35 | 0.48 | 0.48 | 0.63 | 0.34 | 1 | (17) | UC | 0.16 | 0.47 | 0.47 | 0.24 | 0.24 | 0.62 |
| | IN | 0.47 | 0.50 | 0.50 | 0.52 | 0.47 | 0.56 | | IN | 0.43 | 0.49 | 0.50 | 0.45 | 0.45 | 0.61 |
| | CC | 0.49 | 0.61 | 0.61 | 0.71 | 0.49 | 0.83 | | CC | 0.28 | 0.61 | 0.61 | 0.38 | 0.38 | 0.77 |
| | DQ | 0.52 | 0.70 | 0.70 | 0.98 | 0.93 | 0.99 | | DQ | **0.005** | 0.15 | 0.13 | 0.15 | 0.65 | 0.35 |
| ESX | NV | 30 | 16 | 16 | 15 | 15 | 13 | SPX | NV | 24 | 23 | 22 | 24 | 23 | 21 |
| (18) | UC | **0.009** | 0.63 | 0.63 | 0.46 | 0.46 | 0.21 | (17) | UC | 0.11 | 0.16 | 0.24 | 0.11 | 0.16 | 0.35 |
| | IN | 0.31 | 0.59 | 0.59 | 0.61 | 0.61 | 0.66 | | IN | 0.41 | 0.43 | 0.45 | 0.41 | 0.43 | 0.47 |
| | CC | 0.02 | 0.76 | 0.76 | 0.67 | 0.67 | 0.42 | | CC | 0.19 | 0.28 | 0.38 | 0.19 | 0.28 | 0.49 |
| | DQ | **0.001** | 0.26 | 0.26 | 0.38 | 0.33 | 0.20 | | DQ | 0.16 | 0.36 | 0.37 | 0.32 | 0.33 | 0.93 |
| MIB | NV | 31 | 23 | 23 | 24 | 24 | 20 | SMI | NV | 17 | 19 | 18 | 20 | 20 | 17 |
| (18) | UC | **0.005** | 0.26 | 0.26 | 0.17 | 0.17 | 0.64 | (17) | UC | 1 | 0.63 | 0.81 | 0.48 | 0.48 | 1 |
| | IN | 0.29 | 0.44 | 0.44 | 0.42 | 0.42 | 0.50 | | IN | 0.56 | 0.51 | 0.54 | 0.50 | 0.50 | 0.56 |
| | CC | **0.01** | 0.38 | 0.38 | 0.29 | 0.29 | 0.71 | | CC | 0.84 | 0.72 | 0.80 | 0.61 | 0.61 | 0.84 |
| | DQ | **0.03** | **0.02** | **0.02** | **0.04** | 0.24 | 0.06 | | DQ | 0.68 | 0.35 | 0.39 | 0.16 | 0.17 | 0.41 |
| FT | NV | 20 | 23 | 23 | 24 | 24 | 12 | | | | | | | | |
| (18) | UC | 0.64 | 0.26 | 0.26 | 0.18 | 0.18 | 0.13 | | | | | | | | |
| | IN | 0.50 | 0.44 | 0.44 | 0.42 | 0.42 | 0.69 | | | | | | | | |
| | CC | 0.70 | 0.38 | 0.38 | 0.28 | 0.28 | 0.29 | | | | | | | | |
| | DQ | 0.71 | 0.05 | 0.05 | **0.04** | 0.24 | 0.91 | | | | | | | | |

**Table 32: Performance measures for the one-day-ahead VaR forecast with realized kernel**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | PM | (a) | (b) | (c) | (d) | (e) | (f) | Index | PM | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX* | NV | 22 | 18 | 18 | 18 | 18 | 18 | IBX | NV | 26 | 19 | 19 | 20 | 21 | 19 |
| (18) | UC | 0.35 | 1 | 1 | 1 | 1 | 1 | (18) | UC | 0.07 | 0.81 | 0.81 | 0.64 | 0.49 | 0.81 |
| | IN | 0.46 | 0.55 | 0.55 | 0.55 | 0.55 | 0.55 | | IN | 0.38 | 0.20 | 0.20 | 0.22 | 0.25 | 0.52 |
| | CC | 0.50 | 0.83 | 0.83 | 0.83 | 0.83 | 0.83 | | CC | 0.14 | 0.42 | 0.42 | 0.42 | 0.39 | 0.78 |
| | DQ | 0.65 | 0.72 | 0.72 | 0.72 | 0.72 | 0.72 | | DQ | **0.003** | 0.13 | 0.10 | **0.001** | **0.001** | 0.56 |
| AOI | NV | 16 | 13 | 15 | 16 | 19 | 11 | IPC* | NV | 16 | 8 | 8 | 9 | 9 | 7 |
| (17) | UC | 0.80 | 0.31 | 0.61 | 0.81 | 0.63 | 0.12 | (17) | UC | 0.81 | **0.01** | **0.01** | **0.03** | **0.03** | **0.005** |
| | IN | 0.58 | 0.66 | 0.61 | 0.59 | 0.52 | 0.71 | | IN | 0.59 | 0.78 | 0.79 | 0.76 | 0.76 | 0.81 |
| | CC | 0.82 | 0.54 | 0.77 | 0.83 | 0.72 | 0.27 | | CC | 0.83 | 0.05 | 0.05 | 0.09 | 0.09 | **0.02** |
| | DQ | 0.99 | 0.34 | 0.006 | 0.61 | **0.001** | 0.12 | | DQ | 0.47 | 0.59 | 0.59 | 0.73 | 0.74 | 0.47 |
| BVP | NV | 13 | 13 | 13 | 13 | 14 | 13 | KCI* | NV | 13 | 22 | 23 | 19 | 20 | 17 |
| (17) | UC | 0.31 | 0.31 | 0.31 | 0.31 | 0.45 | 0.31 | (17) | UC | 0.31 | 0.24 | 0.16 | 0.63 | 0.48 | 1 |
| | IN | 0.65 | 0.65 | 0.65 | 0.65 | 0.62 | 0.65 | | IN | 0.65 | 0.28 | 0.32 | 0.21 | 0.23 | 0.56 |
| | CC | 0.53 | 0.53 | 0.53 | 0.53 | 0.66 | 0.53 | | CC | 0.53 | 0.28 | 0.23 | 0.40 | 0.38 | 0.83 |
| | DQ | 0.25 | 0.05 | 0.05 | **0.04** | 0.06 | 0.29 | | DQ | 0.98 | 0.35 | 0.32 | 0.28 | 0.16 | 0.36 |
| CAC | NV | 29 | 25 | 25 | 26 | 26 | 22 | NSQ* | NV | 19 | 19 | 19 | 19 | 20 | 21 |
| (18) | UC | **0.01** | 0.12 | 0.12 | 0.08 | 0.08 | 0.36 | (17) | UC | 0.63 | 0.63 | 0.63 | 0.63 | 0.48 | 0.35 |
| | IN | 0.33 | 0.40 | 0.40 | 0.38 | 0.38 | 0.46 | | IN | 0.51 | 0.51 | 0.52 | 0.51 | 0.50 | 0.47 |
| | CC | **0.03** | 0.20 | 0.20 | 0.14 | 0.14 | 0.50 | | CC | 0.71 | 0.71 | 0.71 | 0.71 | 0.61 | 0.49 |
| | DQ | **0.02** | 0.07 | 0.07 | 0.09 | 0.10 | 0.09 | | DQ | 0.15 | 0.18 | 0.18 | 0.19 | 0.16 | 0.16 |
| DAX | NV | 19 | 14 | 14 | 15 | 16 | 16 | NK* | NV | 17 | 15 | 15 | 15 | 14 | 13 |
| (18) | UC | 0.81 | 0.32 | 0.32 | 0.46 | 0.63 | 0.63 | (16) | UC | 0.80 | 0.80 | 0.80 | 0.80 | 0.61 | 0.43 |
| | IN | 0.52 | 0.64 | 0.64 | 0.61 | 0.59 | 0.59 | | IN | 0.17 | 0.13 | 0.13 | 0.11 | 0.11 | 0.09 |
| | CC | 0.78 | 0.55 | 0.55 | 0.67 | 0.76 | 0.76 | | CC | 0.37 | 0.29 | 0.30 | 0.24 | 0.24 | 0.17 |
| | DQ | 0.16 | 0.28 | 0.27 | 0.32 | 0.43 | 0.53 | | DQ | 0.58 | **0.003** | **0.005** | **0.003** | **0.004** | **0.03** |
| DJ | NV | 19 | 15 | 17 | 19 | 20 | 16 | RUS | NV | 23 | 29 | 29 | 26 | 26 | 24 |
| (17) | UC | 0.63 | 0.62 | 1 | 0.63 | 0.48 | 0.81 | (17) | UC | 0.16 | **0.008** | **0.008** | 0.0 | 0.04 | 0.11 |
| | IN | 0.52 | 0.61 | 0.56 | 0.52 | 0.49 | 0.59 | | IN | 0.43 | 0.32 | 0.32 | 0.37 | 0.37 | 0.41 |
| | CC | 0.72 | 0.77 | 0.84 | 0.71 | 0.61 | 0.83 | | CC | 0.28 | 0.02 | 0.02 | 0.08 | 0.08 | 0.19 |
| | DQ | 0.57 | 0.99 | 0.99 | 0.98 | 0.97 | 0.99 | | DQ | **0.005** | **0.001** | **0.002** | **0.008** | 0.02 | 0.13 |
| ESX | NV | 30 | 18 | 19 | 16 | 16 | 15 | SPX* | NV | 27 | 26 | 24 | 24 | 23 | 22 |
| (18) | UC | **0.009** | 1 | 0.81 | 0.63 | 0.63 | 0.46 | (17) | UC | 0.02 | 0.04 | 0.11 | 0.11 | 0.17 | 0.24 |
| | IN | 0.31 | 0.55 | 0.52 | 0.59 | 0.59 | 0.61 | | IN | 0.35 | 0.37 | 0.41 | 0.41 | 0.43 | 0.45 |
| | CC | 0.02 | 0.82 | 0.78 | 0.76 | 0.76 | 0.67 | | CC | 0.05 | 0.08 | 0.19 | 0.19 | 0.28 | 0.38 |
| | DQ | 0.001 | 0.35 | 0.09 | 0.27 | 0.38 | 0.39 | | DQ | 0.11 | 0.23 | 0.32 | 0.32 | 0.36 | 0.86 |
| MIB | NV | 31 | 24 | 24 | 25 | 24 | 18 | SMI | NV | 17 | 17 | 17 | 19 | 19 | 15 |
| (18) | UC | **0.005** | 0.18 | 0.18 | 0.12 | 0.18 | 1 | (17) | UC | 1 | 1 | 1 | 0.63 | 0.63 | 0.61 |
| | IN | 0.29 | 0.42 | 0.42 | 0.40 | 0.42 | 0.54 | | IN | 0.56 | 0.56 | 0.54 | 0.52 | 0.52 | 0.61 |
| | CC | 0.01 | 0.28 | 0.28 | 0.20 | 0.28 | 0.82 | | CC | 0.84 | 0.84 | 0.80 | 0.72 | 0.72 | 0.77 |
| | DQ | **0.003** | **0.004** | **0.004** | **0.02** | 0.38 | 0.06 | | DQ | 0.35 | 0.35 | 0.39 | 0.37 | 0.37 | 0.82 |
| FT | NV | 20 | 15 | 14 | 13 | 14 | 14 | | | | | | | | |
| (18) | UC | 0.64 | 0.46 | 0.32 | 0.21 | 0.32 | 0.32 | | | | | | | | |
| | IN | 0.50 | 0.61 | 0.63 | 0.66 | 0.64 | 0.63 | | | | | | | | |
| | CC | 0.70 | 0.67 | 0.54 | 0.41 | 0.54 | 0.54 | | | | | | | | |
| | DQ | 0.71 | 0.95 | 0.96 | 0.92 | 0.96 | 0.95 | | | | | | | | |

**Table 33: Performance measures for the one-day-ahead VaR forecast during a period of turmoil with sub-sampled realized measures**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | PM | (a) | (b) | (c) | (d) | (e) | (f) | Index | PM | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX | NV | 28 | 22 | 21 | 22 | 19 | 18 | IBX | NV | 29 | 27 | 26 | 29 | 30 | 25 |
| (26) | UC | 0.69 | 0.41 | 0.30 | 0.41 | 0.14 | 0.10 | (26) | UC | 0.55 | 0.84 | 1 | 0.55 | 0.43 | 0.84 |
| | IN | 0.07 | 0.93 | 0.86 | 0.93 | 0.23 | 0.26 | | IN | 0.76 | 0.62 | 0.55 | 0.76 | 0.84 | 0.49 |
| | CC | 0.18 | 0.68 | 0.55 | 0.68 | 0.16 | 0.12 | | CC | 0.76 | 0.82 | 0.79 | 0.76 | 0.68 | 0.73 |
| | DQ | 0.64 | 0.53 | 0.41 | 0.53 | 0.77 | 0.07 | | DQ | 0.39 | 0.11 | 0.61 | 0.71 | 0.35 | 0.92 |
| AOI | NV | 17 | 22 | 21 | 21 | 21 | 19 | IPC | NV | 22 | 19 | 19 | 18 | 18 | 19 |
| (25) | UC | 0.08 | 0.53 | 0.40 | 0.40 | 0.40 | 0.20 | (25) | UC | 0.53 | 0.20 | 0.20 | 0.13 | 0.13 | 0.20 |
| | IN | 0.27 | 0.16 | 0.19 | 0.19 | 0.19 | 0.23 | | IN | 0.97 | 0.22 | 0.22 | 0.25 | 0.25 | 0.22 |
| | CC | 0.12 | 0.30 | 0.28 | 0.28 | 0.28 | 0.21 | | CC | 0.78 | 0.20 | 0.20 | 0.16 | 0.16 | 0.20 |
| | DQ | 0.42 | 0.62 | 0.57 | 0.57 | 0.57 | 0.40 | | DQ | 0.68 | 0.80 | 0.80 | 0.73 | 0.73 | 0.79 |
| BVP | NV | 19 | 21 | 21 | 20 | 20 | 19 | KCI | NV | 18 | 24 | 25 | 24 | 25 | 20 |
| (24) | UC | 0.28 | 0.52 | 0.52 | 0.39 | 0.39 | 0.28 | (25) | UC | 0.13 | 0.83 | 1 | 0.84 | 1 | 0.29 |
| | IN | 0.77 | 0.93 | 0.93 | 0.85 | 0.85 | 0.77 | | IN | 0.67 | 0.45 | 0.15 | 0.13 | 0.16 | 0.23 |
| | CC | 0.51 | 0.77 | 0.77 | 0.65 | 0.65 | 0.51 | | CC | 0.28 | 0.70 | 0.35 | 0.29 | 0.35 | 0.27 |
| | DQ | 0.12 | 0.70 | 0.70 | 0.68 | 0.68 | 0.85 | | DQ | 0.47 | 0.23 | 0.05 | **0.03** | **0.04** | 0.12 |
| CAC | NV | 24 | 24 | 25 | 23 | 23 | 22 | NSQ | NV | 20 | 22 | 22 | 22 | 22 | 19 |
| (26) | UC | 0.68 | 0.68 | 0.84 | 0.54 | 0.54 | 0.41 | (25) | UC | 0.28 | 0.53 | 0.53 | 0.53 | 0.53 | 0.20 |
| | IN | 0.13 | 0.41 | 0.48 | 0.99 | 0.99 | 0.16 | | IN | 0.20 | 0.97 | 0.97 | 0.97 | 0.16 | 0.22 |
| | CC | 0.28 | 0.63 | 0.72 | 0.79 | 0.79 | 0.26 | | CC | 0.24 | 0.78 | 0.78 | 0.78 | 0.29 | 0.20 |
| | DQ | 0.28 | 0.08 | 0.06 | 0.34 | 0.33 | 0.25 | | DQ | 0.67 | 0.12 | 0.34 | 0.33 | 0.22 | 0.51 |
| DAX | NV | 26 | 27 | 27 | 27 | 27 | 27 | NK* | NV | 22 | 15 | 15 | 16 | 17 | 15 |
| (26) | UC | 1 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | (23) | UC | 0.83 | 0.07 | 0.07 | 0.11 | 0.18 | 0.07 |
| | IN | 0.09 | 0.72 | 0.72 | 0.72 | 0.72 | 0.72 | | IN | 0.14 | 0.33 | 0.33 | 0.30 | 0.27 | 0.51 |
| | CC | 0.23 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | | CC | 0.30 | 0.11 | 0.11 | 0.16 | 0.21 | 0.14 |
| | DQ | 0.79 | 0.41 | 0.41 | 0.42 | 0.41 | 0.38 | | DQ | 0.35 | 0.59 | 0.60 | 0.69 | 0.74 | 0.22 |
| DJ | NV | 24 | 25 | 25 | 25 | 24 | 23 | RUS | NV | 21 | 20 | 20 | 20 | 20 | 19 |
| (25) | UC | 0.84 | 1 | 1 | 1 | 0.83 | 0.68 | (25) | UC | 0.39 | 0.29 | 0.29 | 0.29 | 0.29 | 0.20 |
| | IN | 0.13 | 0.51 | 0.51 | 0.51 | 0.88 | 0.95 | | IN | 0.17 | 0.20 | 0.20 | 0.20 | 0.20 | 0.22 |
| | CC | 0.10 | 0.77 | 0.77 | 0.77 | 0.92 | 0.87 | | CC | 0.27 | 0.24 | 0.24 | 0.24 | 0.24 | 0.20 |
| | DQ | 0.61 | 0.40 | 0.40 | 0.44 | 0.81 | 0.73 | | DQ | 0.72 | 0.77 | 0.77 | 0.55 | 0.56 | 0.66 |
| ESX | NV | 29 | 20 | 20 | 21 | 23 | 23 | SPX* | NV | 21 | 22 | 22 | 21 | 20 | 20 |
| (25) | UC | 0.68 | 0.29 | 0.29 | 0.40 | 0.68 | 0.68 | (25) | UC | 0.40 | 0.53 | 0.53 | 0.40 | 0.29 | 0.29 |
| | IN | 0.14 | 0.20 | 0.20 | 0.17 | 0.14 | 0.14 | | IN | 0.17 | 0.97 | 0.97 | 0.89 | 0.20 | 0.82 |
| | CC | 0.29 | 0.24 | 0.24 | 0.27 | 0.29 | 0.29 | | CC | 0.27 | 0.79 | 0.78 | 0.66 | 0.24 | 0.53 |
| | DQ | 0.16 | 0.68 | 0.51 | 0.66 | 0.32 | 0.38 | | DQ | 0.73 | 0.64 | 0.64 | 0.89 | 0.76 | 0.82 |
| MIB* | NV | 31 | 29 | 29 | 29 | 30 | 24 | SMI | NV | 26 | 29 | 30 | 31 | 31 | 30 |
| (25) | UC | 0.23 | 0.42 | 0.42 | 0.42 | 0.31 | 0.84 | (25) | UC | 0.83 | 0.42 | 0.31 | 0.23 | 0.23 | 0.32 |
| | IN | **0.04** | 0.06 | 0.06 | 0.06 | 0.08 | 0.13 | | IN | 0.10 | 0.76 | 0.84 | 0.47 | 0.46 | 0.52 |
| | CC | 0.06 | 0.12 | 0.12 | 0.12 | 0.08 | 0.28 | | CC | 0.24 | 0.65 | 0.56 | 0.35 | 0.35 | 0.46 |
| | DQ | **0.002** | 0.11 | 0.11 | 0.12 | **0.009** | 0.11 | | DQ | 0.10 | **0.004** | **0.01** | **0.006** | **0.006** | **0.01** |
| FT | NV | 25 | 28 | 28 | 28 | 28 | 24 | | | | | | | | |
| (25) | UC | 1 | 0.54 | 0.54 | 0.54 | 0.54 | 0.84 | | | | | | | | |
| | IN | 0.85 | 0.64 | 0.65 | 0.65 | 0.65 | 0.92 | | | | | | | | |
| | CC | 0.93 | 0.71 | 0.71 | 0.71 | 0.71 | 0.93 | | | | | | | | |
| | DQ | 0.38 | 0.06 | 0.06 | 0.07 | 0.07 | 0.29 | | | | | | | | |

**Table 34: Performance measures for the one-day-ahead VaR forecast during a period of turmoil with realized kernel**. For each model (GARCH (*a*), HAR (*b*), HAR-J (*c*), HAR-SJ (*d*), HAR-SJ*aug* (*e*), LHAR (*f*)), we report: the actual number of violations (NV); the p-values for the unconditional coverage (UC), the independence assumption (IN), the conditional coverage (CC), and the DQ test (DQ). Expected number of violations are in parentheses. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| Index | PM | (a) | (b) | (c) | (d) | (e) | (f) | Index | PM | (a) | (b) | (c) | (d) | (e) | (f) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AEX | NV | 28 | 21 | 21 | 22 | 20 | 18 | IBX | NV | 29 | 27 | 26 | 29 | 30 | 25 |
| (26) | UC | 0.69 | 0.30 | 0.30 | 0.41 | 0.21 | 0.09 | (26) | UC | 0.55 | 0.84 | 1 | 0.55 | 0.43 | 0.84 |
| | IN | 0.07 | 0.86 | 0.86 | 0.93 | 0.79 | 0.26 | | IN | 0.76 | 0.62 | 0.55 | 0.76 | 0.84 | 0.49 |
| | CC | 0.18 | 0.55 | 0.55 | 0.68 | 0.42 | 0.12 | | CC | 0.76 | 0.82 | 0.80 | 0.76 | 0.68 | 0.73 |
| | DQ | 0.64 | 0.41 | 0.42 | 0.53 | 0.37 | 0.66 | | DQ | 0.39 | 0.12 | 0.60 | 0.72 | 0.36 | 0.93 |
| AOI* | NV | 17 | 22 | 22 | 22 | 22 | 20 | IPC | NV | 22 | 18 | 18 | 18 | 18 | 17 |
| (25) | UC | 0.08 | 0.53 | 0.53 | 0.53 | 0.53 | 0.29 | (25) | UC | 0.53 | 0.13 | 0.13 | 0.13 | 0.13 | 0.08 |
| | IN | 0.27 | 0.16 | 0.16 | 0.16 | 0.16 | 0.21 | | IN | 0.97 | 0.25 | 0.25 | 0.25 | 0.25 | 0.27 |
| | CC | 0.12 | 0.30 | 0.30 | 0.30 | 0.30 | 0.24 | | CC | 0.78 | 0.16 | 0.16 | 0.16 | 0.16 | 0.12 |
| | DQ | 0.42 | 0.62 | 0.62 | 0.75 | 0.76 | 0.47 | | DQ | 0.68 | 0.74 | 0.74 | 0.74 | 0.74 | 0.61 |
| BVP | NV | 19 | 22 | 22 | 20 | 20 | 17 | KCI | NV | 18 | 22 | 24 | 22 | 22 | 22 |
| (24) | UC | 0.28 | 0.67 | 0.67 | 0.39 | 0.39 | 0.12 | (25) | UC | 0.13 | 0.53 | 0.84 | 0.53 | 0.53 | 0.53 |
| | IN | 0.77 | 0.99 | 0.99 | 0.85 | 0.85 | 0.62 | | IN | 0.67 | 0.07 | 0.13 | 0.07 | 0.08 | 0.97 |
| | CC | 0.51 | 0.87 | 0.87 | 0.65 | 0.65 | 0.26 | | CC | 0.28 | 0.17 | 0.29 | 0.16 | 0.16 | 0.78 |
| | DQ | 0.12 | 0.83 | 0.83 | 0.67 | 0.67 | 0.78 | | DQ | 0.47 | **0.02** | **0.04** | **0.02** | **0.02** | 0.53 |
| CAC | NV | 24 | 25 | 26 | 24 | 24 | 23 | NSQ | NV | 20 | 22 | 22 | 22 | 21 | 22 |
| (26) | UC | 0.68 | 0.84 | 1 | 0.68 | 0.68 | 0.54 | (25) | UC | 0.28 | 0.53 | 0.53 | 0.53 | 0.40 | 0.53 |
| | IN | 0.13 | 0.47 | 0.54 | 0.92 | 0.92 | 0.15 | | IN | 0.20 | 0.97 | 0.97 | 0.97 | 0.17 | 0.97 |
| | CC | 0.28 | 0.72 | 0.79 | 0.87 | 0.87 | 0.27 | | CC | 0.24 | 0.78 | 0.78 | 0.78 | 0.27 | 0.78 |
| | DQ | 0.28 | 0.07 | 0.03 | 0.30 | 0.30 | 0.22 | | DQ | 0.67 | 0.33 | 0.33 | 0.33 | 0.23 | 0.31 |
| DAX | NV | 26 | 27 | 26 | 27 | 27 | 26 | NK* | NV | 22 | 15 | 16 | 17 | 17 | 15 |
| (26) | UC | 1 | 0.84 | 1 | 0.84 | 0.84 | 1 | (23) | UC | 0.83 | 0.07 | 0.11 | 0.18 | 0.18 | 0.07 |
| | IN | 0.09 | 0.72 | 0.79 | 0.72 | 0.72 | 0.75 | | IN | 0.14 | 0.32 | 0.30 | 0.27 | 0.27 | 0.46 |
| | CC | 0.23 | 0.87 | 0.91 | 0.87 | 0.87 | 0.90 | | CC | 0.30 | 0.11 | 0.16 | 0.21 | 0.21 | 0.12 |
| | DQ | 0.79 | 0.41 | 0.23 | 0.42 | 0.42 | 0.74 | | DQ | 0.35 | 0.63 | 0.61 | 0.73 | 0.74 | 0.29 |
| DJ | NV | 24 | 23 | 22 | 25 | 25 | 20 | RUS | NV | 21 | 21 | 22 | 20 | 20 | 18 |
| (25) | UC | 0.84 | 0.68 | 0.53 | 1 | 1 | 0.29 | (25) | UC | 0.39 | 0.39 | 0.53 | 0.29 | 0.29 | 0.13 |
| | IN | 0.13 | 0.39 | 0.33 | 0.51 | 0.51 | 0.82 | | IN | 0.17 | 0.17 | 0.15 | 0.20 | 0.20 | 0.25 |
| | CC | **0.01** | 0.60 | 0.49 | 0.76 | 0.76 | 0.53 | | CC | 0.27 | 0.27 | 0.28 | 0.24 | 0.24 | 0.16 |
| | DQ | 0.61 | 0.66 | 0.40 | 0.41 | 0.43 | 0.80 | | DQ | 0.72 | 0.68 | 0.78 | 0.56 | 0.56 | 0.54 |
| ESX | NV | 29 | 23 | 24 | 23 | 23 | 24 | SPX* | NV | 21 | 23 | 23 | 22 | 22 | 21 |
| (25) | UC | 0.68 | 0.84 | 0.53 | 0.68 | 0.68 | 0.83 | (25) | UC | 0.40 | 0.67 | 0.68 | 0.54 | 0.53 | 0.40 |
| | IN | 0.14 | 0.13 | 0.16 | 0.14 | 0.14 | 0.12 | | IN | 0.17 | 0.95 | 0.95 | 0.97 | 0.97 | 0.90 |
| | CC | 0.29 | 0.29 | 0.30 | 0.30 | 0.29 | 0.28 | | CC | 0.27 | 0.87 | 0.87 | 0.78 | 0.78 | 0.67 |
| | DQ | 0.16 | 0.14 | 0.26 | 0.57 | 0.16 | 0.20 | | DQ | 0.73 | 0.71 | 0.72 | 0.63 | 0.63 | 0.89 |
| MIB | NV | 30 | 29 | 29 | 29 | 30 | 24 | SMI* | NV | 26 | 29 | 30 | 31 | 31 | 27 |
| (25) | UC | 0.32 | 0.42 | 0.42 | 0.42 | 0.32 | 0.84 | (25) | UC | 0.83 | 0.42 | 0.32 | 0.23 | 0.23 | 0.69 |
| | IN | 0.05 | 0.06 | 0.06 | 0.06 | 0.05 | 0.12 | | IN | 0.10 | 0.76 | 0.83 | 0.46 | 0.46 | 0.71 |
| | CC | 0.09 | 0.12 | 0.12 | 0.12 | 0.09 | 0.28 | | CC | 0.24 | 0.65 | 0.56 | 0.35 | 0.35 | 0.81 |
| | DQ | **0.01** | 0.14 | 0.14 | 0.13 | **0.01** | 0.12 | | DQ | 0.10 | **0.003** | **0.01** | **0.006** | **0.006** | **0.03** |
| FT | NV | 25 | 28 | 27 | 25 | 25 | 26 | | | | | | | | |
| (25) | UC | 1 | 0.54 | 0.68 | 1 | 1 | 0.84 | | | | | | | | |
| | IN | 0.85 | 0.65 | 0.71 | 0.85 | 0.85 | 0.78 | | | | | | | | |
| | CC | 0.93 | 0.71 | 0.81 | 0.93 | 0.93 | 0.89 | | | | | | | | |
| | DQ | 0.38 | 0.10 | 0.49 | 0.64 | 0.64 | **0.04** | | | | | | | | |

**Table 35: Performance of one-day-ahead ES forecast with sub-sampled realized measures**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|
| AEX* | **0.019** | 0.165 | 0.170 | 0.074 | **0.034** | 0.315 |
| AOI | 0.927 | 0.659 | 0.636 | 0.756 | 0.79 | 0.943 |
| BVP | 0.401 | 1.000 | 0.997 | 0.998 | 1.000 | 1.000 |
| CAC | 0.101 | 0.675 | 0.676 | 0.55 | 0.464 | 0.707 |
| DAX* | 0.303 | 0.477 | 0.522 | 0.165 | 0.388 | 0.262 |
| DJ* | 0.996 | 1.000 | 1.000 | 0.988 | 0.780 | 1.000 |
| ESX | 0.694 | 0.461 | 0.456 | 0.301 | 0.284 | 0.265 |
| MIB | 0.8 | 0.838 | 0.852 | 0.921 | 0.848 | 0.875 |
| FT | 0.396 | 0.544 | 0.291 | 0.192 | 0.161 | 0.692 |
| IBX | 0.315 | 0.26 | 0.264 | 0.289 | 0.244 | 0.700 |
| IPC | 0.811 | 0.992 | 0.991 | 0.898 | 0.903 | 0.989 |
| KCI | 0.073 | 0.255 | 0.219 | 0.223 | 0.266 | 0.175 |
| NSQ* | 0.425 | 0.240 | 0.279 | 0.300 | 0.08 | 0.805 |
| NK | **0.019** | **0.027** | **0.034** | **0.04** | 0.071 | 0.110 |
| RUS* | 0.961 | 0.601 | 0.519 | 0.780 | 0.702 | 0.436 |
| SPX | 0.983 | 0.989 | 0.966 | 0.956 | 0.9 | 0.995 |
| SMI | 0.189 | 0.874 | 0.825 | 0.861 | 0.887 | 0.783 |

**Table 36: Performance of one-day-ahead ES forecast with realized kernel**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|
| AEX* | **0.016** | 0.274 | 0.273 | 0.112 | 0.05 | 0.414 |
| AOI | 0.939 | 0.554 | 0.632 | 0.784 | 0.836 | 0.950 |
| BVP | 0.408 | 1.000 | 0.998 | 0.999 | 1.000 | 1.000 |
| CAC | 0.100 | 0.278 | 0.296 | 0.542 | 0.471 | 0.483 |
| DAX | 0.271 | 0.073 | 0.058 | 0.251 | 0.372 | 0.443 |
| DJ | 0.995 | 0.983 | 0.989 | 1.000 | 0.999 | 1.000 |
| ESX | 0.684 | 0.475 | 0.521 | 0.332 | 0.261 | 0.418 |
| MIB | 0.825 | 0.840 | 0.832 | 0.924 | 0.889 | 0.768 |
| FT | 0.392 | 0.143 | 0.146 | 0.338 | 0.526 | 0.582 |
| IBX | 0.332 | 0.434 | 0.390 | 0.391 | 0.338 | 0.805 |
| IPC* | 0.992 | 0.936 | 0.967 | 0.900 | 0.830 | 0.994 |
| KCI* | **0.045** | 0.209 | 0.199 | **0.041** | 0.163 | 0.144 |
| NSQ* | 0.469 | 0.136 | 0.183 | 0.242 | 0.074 | 0.712 |
| NK* | **0.016** | **0.058** | **0.012** | **0.023** | **0.021** | 0.141 |
| RUS | 0.946 | 0.611 | 0.650 | 0.466 | 0.434 | 0.623 |
| SPX* | 0.997 | 0.995 | 0.956 | 0.984 | 0.981 | 0.998 |
| SMI | 0.188 | 0.735 | 0.744 | 0.875 | 0.858 | 0.536 |

**Table 37: Performance of one-day-ahead ES forecast during a period of turmoil with sub-sampled realized measures**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.713 | 0.808 | 0.742 | 0.79 | 0.729 | 0.552 |
| AOI | 0.284 | 0.847 | 0.722 | 0.664 | 0.737 | 0.695 |
| BVP | 0.376 | 0.607 | 0.554 | 0.454 | 0.450 | 0.763 |
| CAC | **0.018** | 0.061 | 0.094 | 0.183 | 0.152 | 0.267 |
| DAX | 0.634 | 0.801 | 0.794 | 0.903 | 0.869 | 0.982 |
| DJ | 0.157 | 0.756 | 0.832 | 0.889 | 0.922 | 0.914 |
| ESX | 0.744 | 0.745 | 0.585 | 0.824 | 0.956 | 0.979 |
| MIB* | 0.454 | 0.571 | 0.604 | 0.581 | 0.678 | 0.378 |
| FT | 0.195 | 0.374 | 0.288 | 0.312 | 0.292 | 0.439 |
| IBX | 0.112 | 0.496 | 0.425 | 0.550 | 0.557 | 0.375 |
| IPC | 0.061 | 0.408 | 0.396 | 0.537 | 0.450 | 0.728 |
| KCI | **0.001** | 0.110 | 0.138 | 0.052 | 0.062 | 1.000 |
| NSQ | 0.438 | 0.991 | 0.995 | 0.993 | 0.993 | 0.975 |
| NK* | 0.870 | 0.825 | 0.804 | 0.888 | 0.930 | 0.221 |
| RUS | 0.353 | 0.907 | 0.880 | 0.902 | 0.906 | 0.829 |
| SPX* | 0.083 | 0.616 | 0.727 | 0.724 | 0.625 | 0.904 |
| SMI | 0.474 | 0.106 | 0.101 | 0.104 | 0.135 | 0.240 |

**Table 38: Performance of one-day-ahead ES forecast during a period of turmoil with realized kernel**. P-values from the one-sided upper-tail bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJaug | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.731 | 0.733 | 0.780 | 0.799 | 0.657 | 0.561 |
| AOI* | 0.240 | 0.812 | 0.790 | 0.706 | 0.810 | 0.869 |
| BVP | 0.374 | 0.81 | 0.758 | 0.520 | 0.492 | 0.801 |
| CAC | **0.02** | 0.141 | 0.225 | 0.212 | 0.174 | 0.555 |
| DAX | 0.639 | 0.677 | 0.370 | 0.825 | 0.792 | 0.869 |
| DJ | 0.157 | 0.563 | 0.430 | 0.924 | 0.931 | 0.694 |
| ESX | 0.756 | 0.953 | 0.949 | 0.917 | 0.964 | 0.988 |
| MIB | 0.242 | 0.454 | 0.506 | 0.409 | 0.54 | 0.382 |
| FT | 0.201 | 0.252 | 0.290 | 0.354 | 0.341 | 0.439 |
| IBX | 0.147 | 0.485 | 0.416 | 0.579 | 0.571 | 0.415 |
| IPC | 0.058 | 0.721 | 0.676 | 0.570 | 0.643 | 0.592 |
| KCI | **0.001** | 0.092 | 0.173 | 0.069 | 0.056 | 0.083 |
| NSQ | 0.437 | 0.967 | 0.962 | 0.989 | 0.992 | 0.980 |
| NK* | 0.851 | 0.610 | 0.810 | 0.939 | 0.930 | 0.367 |
| RUS | 0.375 | 0.877 | 0.859 | 0.900 | 0.896 | 0.803 |
| SPX* | 0.092 | 0.512 | 0.671 | 0.799 | 0.791 | 0.880 |
| SMI* | 0.511 | **0.032** | 0.096 | 0.110 | 0.115 | 0.101 |

**Table 39: Performance measures for the ten-day-ahead VaR forecast with sub-sampled realized measures**. Actual and expected number of violations of the VaR. Cases in which the null hypothesis of correct unconditional coverage is rejected at level $\alpha = 10\%$ are in bold. The best performer is in *italics*. The * indicates series for which a higher threshold was required for appropriate GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 18 | **34 (0.08)** | *18 (0.99)* | 17 (0.88) | *18 (0.99)* | *18 (0.99)* | *18 (0.99)* |
| AOI* | 17 | 25 (0.43) | 23 (0.30) | 23 (0.35) | *22 (0.34)* | *22 (0.37)* | *22 (0.52)* |
| BVP | 17 | 12 (0.49) | 22 (0.54) | 22 (0.54) | 22 (0.54) | 23 (0.49) | *20 (0.68)* |
| CAC* | 18 | 29 (0.23) | *26 (0.36)* | *26 (0.36)* | *26 (0.39)* | *26 (0.39)* | 28 (0.32) |
| DAX | 18 | 27 (0.33) | 22 (0.66) | 22 (0.66) | *20 (0.82)* | 23 (0.58) | 21 (0.75) |
| DJ | 17 | 12 (0.51)) | *18 (0.93)* | *18 (0.93)* | 20 (0.73) | 20 (0.73) | 19 (0.82) |
| ESX | 18 | 32 (0.12) | 19 (0.90) | 19 (0.90) | *18 (0.99)* | 20 (0.81) | 21 (0.71) |
| MIB | 18 | **50 (0.02)** | 25 (0.36) | 25 (0.36) | 25 (0.36) | 25 (0.36) | *23 (0.49)* |
| FT | 18 | 33 (0.14) | 16 (0.82) | 16 (0.82) | 15 (0.68) | 16 (0.81) | *18 (0.95)* |
| IBX | 18 | **39 (0.07)** | 25 (0.39) | 24 (0.44) | *22 (0.58)* | 25 (0.40) | 25 (0.36) |
| IPC | 17 | *14 (0.68)* | 9 (0.30) | 8 (0.23) | 9 (0.30) | 10 (0.32) | 10 (0.32) |
| KCI* | 17 | *11 (0.45)* | 29 (0.50) | 29 (0.50) | 29 (0.50) | 29 (0.50) | 30 (0.24) |
| NSQ | 17 | 20 (0.73) | 16 (0.84) | 16 (0.84) | 15 (0.72) | 16 (0.84) | *17 (0.95)* |
| NK | 16 | *24 (0.57)* | 26 (0.40) | 26 (0.50) | 27 (0.37) | 27 (0.37) | 25 (0.48) |
| RUS | 17 | 18 (0.94) | 16 (0.94) | *17 (0.87)* | 14 (0.69) | 14 (0.69) | 16 (0.87) |
| SPX | 17 | 10 (0.38) | 18 (0.93) | 18 (0.93) | *17 (0.96)* | *17 (0.96)* | *17 (0.96)* |
| SMI | 17 | 24 (0.41) | *13 (0.43)* | *13 (0.43)* | *13 (0.43)* | *13 (0.43)* | 12 (0.43) |

**Table 40: Performance measures for the ten-day-ahead VaR forecast with realized kernel**. Actual and expected number of violations of the VaR. Cases in which the null hypothesis of correct unconditional coverage is rejected at level $\alpha = 10\%$ are in bold. The best performer is in *italics*. The * indicates series for which a higher threshold was required for appropriate GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 18 | **34 (0.08)** | 16 (0.76) | 16 (0.76) | 19 (0.89) | *18 (0.99)* | 19 (0.90) |
| AOI* | 17 | 25 (0.43) | 24 (0.30) | *23 (0.29)* | *23 (0.29)* | *23 (0.16)* | *23 (0.44)* |
| BVP | 17 | 12 (0.49) | 23 (0.46) | 23 (0.46) | 22 (0.54) | 23 (0.48) | *19 (0.77)* |
| CAC* | 18 | 29 (0.23) | 26 (0.36) | 26 (0.36) | *25 (0.39)* | 26 (0.39) | 27 (0.36) |
| DAX | 18 | 27 (0.33) | 21 (0.74) | *20 (0.82)* | *20 (0.82)* | 22 (0.66) | 22 (0.66) |
| DJ | 17 | 12 (0.51) | *18 (0.93)* | *18 (0.93)* | 19 (0.82) | 19 (0.82) | 19 (0.82) |
| ESX | 18 | 32 (0.12) | 21 (0.73) | 21 (0.73) | *20 (0.82)* | 21 (0.73) | 21 (0.75) |
| MIB | 18 | **50 (0.02)** | 26 (0.32) | 26 (0.32) | 25 (0.36) | 25 (0.36) | *23 (0.49)* |
| FT | 18 | 33 (0.14) | 19 (0.84) | *18 (0.95)* | *18 (0.95)* | *18 (0.95)* | 23 (0.48) |
| IBX | 18 | **39 (0.07)** | *24 (0.42)* | *24 (0.42)* | *24 (0.42)* | 25 (0.38) | 25 (0.36) |
| IPC | 17 | *14 (0.68)* | **12 (0.05)** | **11 (0.07)** | **13 (0.02)** | **13 (0.02)** | 8 (0.21) |
| KCI* | 17 | *11 (0.45)* | 29 (0.23) | 29 (0.23) | 29 (0.23) | 29 (0.25) | 30 (0.24) |
| NSQ | 17 | *20 (0.73)* | 21 (0.65) | 21 (0.65) | *20 (0.73)* | *20 (0.73)* | *20 (0.72)* |
| NK | 16 | *24 (0.57)* | 25 (0.44) | 27 (0.35) | 25 (0.41) | 25 (0.41) | 26 (0.40) |
| RUS | 17 | *18 (0.94)* | 19 (0.85) | 19 (0.85) | *18 (0.94)* | *18 (0.94)* | 19 (0.85) |
| SPX | 17 | 10 (0.38) | *18 (0.93)* | 19 (0.82) | 16 (0.87) | 16 (0.87) | 20 (0.72) |
| SMI* | 17 | 24 (0.41) | *12 (0.43)* | *12 (0.43)* | *12 (0.43)* | *12 (0.43)* | *12 (0.43)* |

**Table 41: Performance measures for the ten-day-ahead VaR forecast during a period of turmoil with sub-sampled realized measures**. Actual and expected number of violations of the VaR. Cases in which the null hypothesis of correct unconditional coverage is rejected at level $\alpha = 10\%$ are in bold. The best performer is in *italics*. The * indicates series for which a higher threshold was required for appropriate GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 25 | *26 (0.63)* | **9 (0.02)** | **9 (0.02)** | **9 (0.02)** | **9 (0.02)** | 12 (0.10) |
| AOI | 25 | *22 (0.69)* | 19 (0.61) | 19 (0.61) | 19 (0.61) | 19 (0.61) | 18 (0.52) |
| BVP | 24 | **10 (0.06)** | *15 ( 0.34)* | *15 (0.33)* | 14 (0.27) | *15 (0.35)* | 14 (0.29) |
| CAC* | 25 | 33 (0.58) | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* | 18 (0.46) |
| DAX | 25 | *32 (0.81)* | 13 (0.25) | 13 (0.25) | 14 (0.27) | 14 (0.27) | 14 (0.27) |
| DJ | 25 | **12 (0.06)** | *20 (0.66)* | *20 (0.66)* | *20 (0.66)* | *20 (0.66)* | 19 (0.60) |
| ESX* | 25 | *21 (0.78)* | 11 (0.15) | 11 (0.15) | **10 (0.09)** | **10 (0.09)** | 9 (0.08) |
| MIB | 25 | 39 (0.64) | 32 (0.59) | *31 (0.63)* | *31 (0.63)* | *31 (0.63)* | 32 (0.59) |
| FT* | 25 | 43 (0.40) | *19 (0.59)* | *19 (0.59)* | 18 (0.53) | 18 (0.53) | 17 (0.49) |
| IBX | 25 | 37 (0.38) | 26 (0.96) | 27 (0.89) | *25 (0.98)* | 26 (0.96) | 24 (0.90) |
| IPC | 25 | *29 (0.99)* | **12 (0.07)** | **12 (0.07)** | **13 (0.09)** | **13 (0.09)** | 13 (0.09) |
| KCI | 24 | **11 (0.07)** | *24 (0.39)* | 23 (0.39) | 23 (0.47) | 23 (0.47) | 17 (0.42) |
| NSQ | 25 | 23 (0.83) | *26 (0.95)* | 21 (0.87) | 21 (0.87) | 21 (0.87) | 22 (0.78) |
| NK | 23 | **5 (0.01)** | **1 (0.003)** | **2 (0.006)** | **2 (0.006)** | **2 (0.006)** | **2 (0.007)** |
| RUS | 25 | *19 (0.47)* | 15 (0.24) | 16 (0.32) | 15 (0.24) | 15 (0.24) | 13 (0.15) |
| SPX | 25 | 15 (0.14) | 18 (0.50) | 19 (0.57) | 18 (0.50) | 18 (0.50) | *21 (0.71)* |
| SMI | 25 | *27 (0.88)* | 21 (0.75) | 21 (0.75) | 20 (0.67) | 20 (0.67) | 20 (0.67) |

**Table 42: Performance measures for the ten-day-ahead VaR forecast during a period of turmoil with realized kernel**. Actual and expected number of violations of the VaR. Cases in which the null hypothesis of correct unconditional coverage is rejected at level $\alpha = 10\%$ are in bold. The best performer is in *italics*. The * indicates series for which a higher threshold was required for appropriate GPD behavior.

| Index | Expected | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|---|
| AEX | 25 | *26 (0.63)* | **10 (0.02)** | **10 (0.02)** | **9 (0.02)** | **10 (0.02)** | 12 (0.09) |
| AOI | 25 | *22 (0.69)* | 19 (0.61) | 19 (0.61) | 19 (0.61) | 18 (0.61) | 16 (0.37) |
| BVP | 24 | **10 (0.06)** | *15 (0.33)* | *15 (0.33)* | *15 (0.27)* | *15 (0.33)* | 14 (0.29) |
| CAC* | 25 | 33 (0.58) | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* | *19 (0.54)* | 18 (0.46) |
| DAX | 25 | *32 (0.81)* | 14 (0.27) | 12 (0.25) | 14 (0.27) | 14 (0.27) | 14 (0.27) |
| DJ | 25 | **12 (0.06)** | 18 (0.65) | 18 (0.65) | *19 (0.65)* | *19 (0.65)* | 18 (0.54) |
| ESX | 25 | *21 (0.78)* | **11 (0.08)** | **11 (0.08)** | **11 (0.08)** | **11 (0.08)** | 10 (0.14) |
| MIB | 25 | 39 (0.64) | 32 (0.60) | 33 (0.62) | *31 (0.60)* | 32 (0.62) | 33 (0.53) |
| FT* | 25 | 43 (0.40) | *20 (0.59)* | *20 (0.59)* | *20 (0.59)* | 19 (0.53) | 19 (0.59) |
| IBX | 25 | 37 (0.38) | 27 (0.96) | 27 (0.96) | 26 (0.89) | 27 (0.96) | *24 (0.90)* |
| IPC | 25 | *29 (0.99)* | **12 (0.07)** | **11 (0.07)** | **13 (0.09)** | **13 (0.09)** | 14 (0.13) |
| KCI | 24 | **11 (0.07)** | 18 (0.39) | 18 (0.39) | *20 (0.47)* | *20 (0.47)* | 17 (0.42) |
| NSQ | 25 | 23 (0.83) | *25 (0.95)* | 24 (0.87) | 24 (0.87) | 24 (0.87) | 24 (0.95) |
| NK | 23 | **5 (0.01)** | **2 (0.006)** | **2 (0.006)** | **2 (0.006)** | **2 (0.006)** | **1 (0.002)** |
| RUS | 25 | *19 (0.47)* | 15 (0.24) | 17 (0.31) | 15 (0.24) | 15 (0.24) | 13 (0.15) |
| SPX | 25 | 15 (0.14) | 18 (0.50) | 19 (0.57) | 18 (0.50) | 18 (0.50) | *20 (0.64)* |
| SMI* | 25 | *27 (0.88)* | 22 (0.75) | 22 (0.75) | 21 (0.67) | 21 (0.67) | 19 (0.57) |

**Table 43: Performance of ten-day-ahead ES forecast with sub-sampled realized measures**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.114 | 0.298 | 0.231 | 0.275 | 0.414 | 0.437 |
| AOI* | 0.946 | 0.163 | 0.095 | 0.268 | 0.165 | 0.056 |
| BVP | 0.673 | 0.081 | 0.068 | 0.081 | **0.038** | 0.09 |
| CAC | 0.283 | 0.408 | 0.420 | 0.367 | 0.367 | 0.732 |
| DAX | **0.045** | 0.068 | 0.077 | **0.024** | 0.108 | **0.026** |
| DJ | 0.345 | 0.215 | 0.215 | 0.158 | 0.133 | 0.109 |
| ESX | 0.153 | 0.117 | 0.127 | 0.058 | **0.059** | 0.292 |
| MIB | 0.288 | 0.925 | 0.934 | 0.942 | 0.932 | 0.843 |
| FT | 0.119 | 0.319 | 0.318 | 0.216 | 0.288 | 0.249 |
| IBX | 0.738 | 0.340 | 0.303 | 0.141 | **0.045** | 0.400 |
| IPC | 0.220 | 0.811 | 0.274 | 0.749 | 0.858 | 0.829 |
| KCI* | **0.027** | 0.350 | 0.312 | 0.356 | 0.362 | 0.276 |
| NSQ | 0.062 | **0.010** | **0.015** | **0.005** | **0.004** | **0.012** |
| NK | **0.007** | **0.013** | **0.007** | **0.004** | **0.011** | **0.021** |
| RUS | 0.089 | **0.076** | **0.145** | **0.061** | **0.049** | 0.110 |
| SPX | 0.282 | **0.040** | **0.039** | **0.019** | **0.020** | **0.021** |
| SMI | 0.479 | 0.475 | 0.474 | 0.430 | 0.430 | 0.398 |

**Table 44: Performance of ten-day-ahead ES forecast with realized kernel**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.114 | 0.249 | 0.277 | 0.291 | 0.304 | 0.479 |
| AOI* | 0.946 | 0.296 | 0.087 | 0.299 | 0.295 | 0.068 |
| BVP | 0.673 | 0.082 | 0.104 | 0.097 | **0.033** | 0.071 |
| CAC | 0.283 | 0.460 | 0.490 | 0.343 | 0.463 | 0.684 |
| DAX | **0.045** | 0.0.057 | **0.018** | **0.019** | 0.053 | 0.063 |
| DJ | 0.345 | 0.135 | 0.110 | 0.278 | 0.283 | 0.218 |
| ESX | 0.153 | 0.240 | 0.219 | 0.173 | 0.059 | 0.158 |
| MIB | 0.288 | 0.919 | 0.918 | 0.903 | 0.916 | 0.762 |
| FT | 0.119 | 0.276 | 0.158 | 0.233 | 0.222 | 0.266 |
| IBX | 0.738 | 0.422 | 0.564 | 0.450 | 0.454 | 0.623 |
| IPC | 0.220 | 0.633 | 0.614 | 0.737 | 0.788 | 1.000 |
| KCI* | **0.027** | 0.265 | 0.233 | 0.284 | 0.323 | 0.246 |
| NSQ | 0.062 | 0.066 | 0.087 | 0.079 | **0.023** | **0.009** |
| NK | **0.007** | **0.004** | **0.006** | **0.002** | **0.001** | **0.018** |
| RUS | 0.089 | 0.093 | 0.086 | 0.054 | **0.044** | 0.084 |
| SPX | 0.282 | **0.034** | **0.039** | **0.018** | **0.022** | 0.098 |
| SMI | 0.479 | 0.544 | 0.541 | 0.527 | 0.531 | 0.491 |

**Table 45: Performance of ten-day-ahead ES forecast during a period of turmoil with sub-sampled realized measures**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.954 | 0.982 | 0.992 | 0.992 | 0.992 | 0.795 |
| AOI* | 1.000 | 0.263 | 0.306 | 0.277 | 0.320 | 0.326 |
| BVP | 0.887 | 0.945 | 0.929 | 0.909 | 0.939 | 0.946 |
| CAC | 0.385 | 0.296 | 0.264 | 0.276 | 0.238 | 0.187 |
| DAX | 0.623 | 0.884 | 0.903 | 0.845 | 0.878 | 0.740 |
| DJ | 0.971 | 0.713 | 0.669 | 0.629 | 0.645 | 0.715 |
| ESX | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| MIB | **0.001** | **0.007** | **0.001** | **0.001** | **0.002** | **0.030** |
| FT | 0.177 | 0.061 | 0.068 | **0.026** | **0.026** | **0.026** |
| IBX | 0.564 | 0.377 | 0.426 | 0.334 | 0.182 | 0.243 |
| IPC | 0.567 | 0.582 | 0.592 | 0.672 | 0.706 | 0.714 |
| KCI* | 1.000 | 0.960 | 0.976 | 0.985 | 0.983 | 0.945 |
| NSQ | 0.820 | 0.632 | 0.571 | 0.559 | 0.636 | 0.355 |
| NK | 0.924 | 1.000 | 1.000 | 0.877 | 1.000 | 1.000 |
| RUS | 0.997 | 0.953 | 0.963 | 0.950 | 0.943 | 0.908 |
| SPX | 0.996 | 0.592 | 0.632 | 0.532 | 0.476 | 0.885 |
| SMI | 0.054 | **0.007** | **0.016** | **0.024** | **0.012** | 0.083 |

**Table 46: Performance of ten-day-ahead ES forecast during a period of turmoil with realized kernel**. P-values from the one-sided upper-tail block-bootstrap test for the null of mean zero difference between actual returns and ES when VaR is exceeded. Rejection at level $\alpha = 5\%$ is in bold. The * denotes series for which a higher threshold was required for appropriate GPD behavior.

| | GARCH | HAR | HAR-J | HAR-SJ | HAR-SJ-AUG | LHAR |
|---|---|---|---|---|---|---|
| AEX | 0.954 | 1.000 | 1.000 | 0.999 | 1.000 | 0.832 |
| AOI* | 1.000 | 0.345 | 0.269 | 0.345 | 0.262 | 0.256 |
| BVP | 0.887 | 0.949 | 0.909 | 0.934 | 0.949 | 0.946 |
| CAC | 0.385 | 0.321 | 0.316 | 0.356 | 0.339 | 0.156 |
| DAX | 0.623 | 0.868 | 0.735 | 0.824 | 0.826 | 0.601 |
| DJ | 0.971 | 0.672 | 0.601 | 0.750 | 0.707 | 0.748 |
| ESX | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| MIB | **0.001** | **0.002** | **0.001** | **0.001** | **0.001** | **0.005** |
| FT | 0.177 | 0.091 | 0.103 | 0.064 | **0.016** | **0.041** |
| IBX | 0.564 | 0.357 | 0.378 | 0.325 | 0.168 | 0.190 |
| IPC | 0.567 | 0.671 | 0.592 | 0.763 | 0.755 | 0.819 |
| KCI* | 1.000 | 0.960 | 0.967 | 0.968 | 0.956 | 0.920 |
| NSQ | 0.820 | 0.347 | 0.391 | 0.392 | 0.451 | 0.254 |
| NK | 0.924 | 1.000 | 1.000 | 0.877 | 1.000 | 1.000 |
| RUS | 0.993 | 0.974 | 0.949 | 0.944 | 0.943 | 0.915 |
| SPX | 0.996 | 0.705 | 0.461 | 0.647 | 0.791 | 0.836 |
| SMI | 0.054 | 0.160 | 0.131 | 0.212 | 0.224 | 0.162 |

# Chapter 4

# Extremal Behaviour of Financial Returns and Models

## 4.1 Introduction

A large body of the financial economics literature has been devoted to study the autocorrelation of asset returns, particularly for its relevance in connection with the Efficient Market Hypothesis (Conrad and Kaul, 1988; Lo and MacKinlay, 1988). The empirical evidence suggests that daily returns present a mild and often not significant positive autocorrelation that varies according to other variables, such as the volume of trades Campbell et al. (1993) and the volatility of asset returns (LeBaron, 1992). In contrast, the volatility of asset returns is a highly predictable quantity (Bollerslev, 1986; Engle, 1982) which exhibits a strong degree of persistence.

While the serial correlation of stock returns and volatility has attracted a lot of attention, serial dependence in extreme returns has been unexpectedly overlooked. Understanding the behaviour of extreme observations is crucial, as they are informative of *tail risk*. Are daily extreme observations in financial returns isolated events or do they tend to occur

together? Does the occurrence of an extreme event increase the probability of observing another event of similar magnitude? Are financial econometrics models capable of accurately capturing the observed behaviour in the extremes? These questions are of interest for both regulatory purposes and financial risk management, and this paper attempts to provide some answers.

The left panel of Figure 44 reports the price path recorded by the S&P500 index in the last part of 2008, in the aftermath of the Lehman Brothers collapse. In red, the days from September 26 to October 10 on which the index plummeted from 1200 points to 900, losing almost one fourth of its capitalization. The figure emphasizes that this crash was not due to one single observation but to a sequence of extreme negative observations. In particular, five out of ten days experienced a daily return below three standard deviations from the mean. The right panel of Figure 44 reports a similar scenario for the Euro Stoxx 50, an index of Eurozone blue-chips stocks, in the fall of 2011, during the European sovereign debt crisis. In this case, the index lost $20\%$ of its value in two weeks and five out of twelve days registered negative values below two standard deviations from the mean. These two examples give the feeling of the importance of studying the dependence in extreme returns and understanding which models are eventually able to capture it appropriately.

As extreme returns happen near the boundaries of the support of the return distribution, a definition of *extremal dependence* entailing an asymptotic behaviour seems proper. Let $\{X_t\}$ be a stationary sequence and denote the conditional tail probability of the random vector $(X_t, X_{t+h})$ at level $x$ as

$$c(x,h) = \Pr(X_{t+h} > x | X_t > x), \tag{4.1}$$

then $(X_t, X_{t+h})$ are asymptotically independent if $c(x,h) \to 0$ as $x \to \infty$, while they are asymptotically dependent, and thus present extremal dependence, if $c(x,h) \to C > 0$ as $x \to \infty$. The extremal behaviour of traditional classes of models for the daily returns, such as the Generalized AutoRegressive Conditional Heteroskedastic (GARCH) of Bollerslev (1986) and the Stochastic Volatility (SV) pioneered by Clark (1973), have been

widely studied, amongst others, by Mikosch and Starica (2000), Davis and Mikosch (2009b) and Mikosch and Rezapur (2013). In particular, while GARCH processes exhibit extremal dependence regardless of the assumptions on the error term, SV processes present extremal dependence only if the volatility component satisfies certain conditions (Mikosch and Rezapur, 2013).

Empirical assessment of the extremal dependence in financial returns is restrained to few small examples (Davis et al., 2009; Liu and Tawn, 2013). Our first contribution is to provide a comprehensive empirical analysis of the extremal behaviour of daily returns. Exploiting different statistical procedures, we assess the existence and the persistence of the extremal dependence in both positive and negative extreme returns and in the extremes of the return variance for 17 international equity indexes. In particular, we rely on three different approaches: the intervals estimator of the extremal index (Ferro and Segers, 2003), a quantity emerging from Extreme Value Theory and characterizing the extremal behaviour of a dependent sequence; the formal test of extremal dependence developed in Liu and Tawn (2013) and based on the coefficient of tail dependence of Ledford and Tawn (2003); the extremogram of Davis et al. (2009) providing an analogue of the correlogram to estimate the serial dependence in the extremes. Our results provide evidence of strong and persistent dependence in both the upper and lower tails of the return distribution and in the extremes of the return variance.

The compelling evidence of extremal dependence in real data leads to favour processes for the daily returns in the GARCH class or SV processes satisfying the conditions of Mikosch and Rezapur (2013). However, recent contributions in the high-frequency (HF) financial econometrics literature have led to new classes of models that will likely become the new standard in the near future. It is relevant to understand whether processes for the daily returns and volatility based on HF data can account for the extremal dependence observed in the data.

With the availability of HF data, research on the volatility of returns has taken new avenues. Intra-daily returns are used to construct nonparametric estimators of the daily asset price variation, termed realized

measures (Barndorff-Nielsen and Shephard, 2002). Most of the research on this topic has been devoted to the development of models that use the realized measures to improve the accuracy of volatility forecasts. Hansen and Lunde (2011) classify the existing approaches in two classes. The *reduced-form* class provides time series models for the realized measures. Within this class, the Heterogeneous AutoRegressive (HAR) model of Corsi (2009) has probably been the most successful. The *model-based* class jointly models the returns and the realized measures of volatility. Notable examples within this class are the Multiplicative Error Model (MEM) introduced by Engle and Gallo (2006), the HEAVY model of Shephard and Sheppard (2010), and the Realized GARCH of Hansen et al. (2012). In this paper, we consider models that belong to this second class and refer to them as to *HF-based volatility processes*. We establish a simple framework to obtain HF-based volatility processes and provide results for the existence of processes that exhibit extremal dependence. Furthermore, we use Monte Carlo simulations to explore the degree and persistence of the extremal dependence of such processes and assess whether it is consistent with that exhibited by real time series. We find that modelling the persistence in the realized measure and the leverage effect (Glosten et al., 1993) is crucial to explain the observed pattern of extreme dependence in both tails of financial returns.

The empirical findings on the extremal dependence are also relevant because of the important implications they bear on continuous-time models. Given that these models are commonly used in practice, for example in option pricing, it is important for them to account for the observed pattern of dependence in the extremes. Extremal properties of continuous-time models have been discussed, amongst others, in Fasen (2009), but the issue we want to address is how to make the extremal behaviour a central aspect of the estimation strategy. Estimation methods for continuous-time volatility models typically optimize the fit to the body of the return distribution, but when the interest is on the tails, they should be the focus in the estimation strategy. We consider the Simulated Method of Moments (SMM) of Corradi and Distaso (2006) and include extreme moment conditions that allow to identify both the dependence

in the tails and the leverage effect. In this sense, we propose an alternative approach to that of Corsi and Renò (2012) and Harvey and Shephard (1996) to estimate asymmetric continuous-time volatility models. We use this strategy to estimate a two-factor GARCH Diffusion process (Nelson, 1990), and find that it perfectly captures the empirical pattern of extremal dependence.

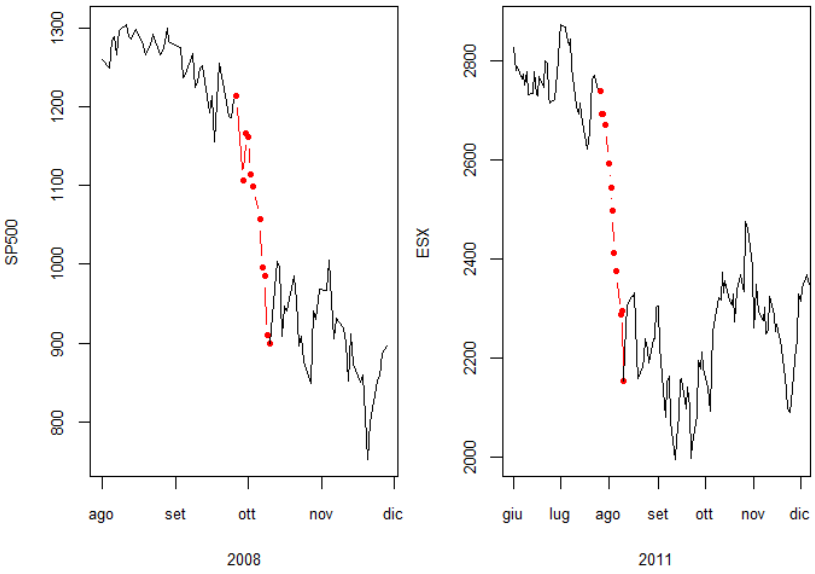
Summarizing, this article gives three contributions: first, it provides evidence of strong and persistent dependence in both positive and negative extreme returns and in the extremes of the return volatility; second, it studies the extremal properties of various HF-based volatility processes, and finds that modelling the persistence in the realized measure and the leverage effect is important to capture the empirical dependence in the extremes; third, it addresses how to make the extremal dependence a central aspect in the estimation of a continuous-time volatility model.

The remainder of the chapter is organized as follows: Section 4.2 describes three different estimators of the extremal dependence; Section 4.3 provides evidence of dependence in the extremes of 17 equity indexes; Section 4.4 presents results concerning the extremal behaviour of different HF-based volatility processes; Section 4.5 contains estimation of continuous-time models; Section 4.6 concludes. A supplementary appendix contains additional details and results.

## 4.2 Measuring the extremal dependence

Measuring and estimating the extremal dependence in a time series is a rather challenging problem. Since financial time series are not Gaussian processes, the autocorrelation function is not well-suited for describing the dependence structure in the extremes. In what follows, we present three different but related measures of extremal dependence and their estimators. Note that the definition in Equation (4.1) characterizes the asymptotic behaviour of the conditional tail probability at lag $h$, but reality allows us to deal only with finite samples. To circumvent this prob-

**Figure 44: Stock market crashes**. The left panel shows the S&P500 (SP500), while the right panel shows the EuroStoxx50 (ESX).



lem, we define a threshold $u$ as a high empirical quantile[1] of the sample $\{X_t\}_{t=1}^{T}$ and study the behaviour of the exceedances of such threshold, i.e. $X_t > u$.

### 4.2.1 Extremal index

The extremal index is a quantity arising from the Extreme Value Theory of weakly-dependent data, and characterizes both the frequency with which extreme events occur and the clustering features of an extreme event, once such an event occurs. Let $\{X_t\}$ be a strictly stationary sequence with marginal distribution function $F$, finite or infinite right end point $\omega = \sup\{x : F(x) < 1\}$ and survival function $\overline{F} = 1 - F$. Let $X_1, \ldots, X_n$ be a random sample from $F$ and denote the sample maxima as $M_n = \max\{X_1, \ldots, X_n\}$. The process $\{X_t\}$ is said to have extremal index $\theta \in [0, 1]$ if for integer $n \geq 1$ and for each $\tau > 0$ there exists a sequence of real numbers $\{u_n\}$ such that as $n \to \infty$ the following are equivalent,

$$n\overline{F}(u_n) \to \tau \tag{4.2}$$

$$\Pr(M_n \leq u_n) \to e^{-\theta\tau} \tag{4.3}$$

If $\theta = 1$ then exceedances of a high threshold are independent, and the standard Extreme Value Theory results on independent sequences hold (Embrechts et al., 1997). If $\theta < 1$ then exceedances tend to cluster in the limit and therefore they are dependent (Leadbetter et al., 1983). Ledford and Tawn (2003) establish a bound relationship between the conditional probability in Equation (4.1) and the extremal index. In particular, they show that when $\{X_t\}$ is asymptotically dependent for at least one lag then $\theta < 1$. Conversely, if the process is asymptotically independent at all lags, theoretical results suggest that $\theta = 1$. We refer to Ledford and Tawn (2003) for a detailed discussion.

Exploiting point process theory, Hsing et al. (1988) show that the rescaled times of exceedances of a sequence $u_n$ for a stationary process

[1] We will always consider exceedances over a high quantile and change the underlying sample depending on what we want to study. For instance, we will consider the negated returns to study the dependence in the negative extreme returns.

converge weakly to a compound Poisson process with mean cluster size $\theta^{-1}$. Ferro and Segers (2003) propose an *intervals estimator* for the extremal index $\theta$, based on a limiting characterization of the interexceedance times. In particular, they show that properly normalized times between consecutive exceedances converge to a random variable which is zero with probability $1-\theta$ and exponential with mean $\theta^{-1}$ with probability $\theta$. Given this characterization, they propose a moment-based estimator of $\theta$.

Suppose a random sample $X_1, \ldots, X_n$ from $F$ and a high threshold $u$. Let $N(u) = \sum_{t=1}^{n} I(X_t > u)$ be the number of observations exceeding $u$, and let $1 \leq S_1 < \cdots < S_n \leq n$ be the exceedance times. The observed interexceedance times are $E_i = S_{i+1} - S_i$ for $i = 1, \ldots, n-1$. The intervals estimator is a consistent estimator of the extremal index and is defined as

$$\tilde{\theta}(u) = \begin{cases} 1 \wedge \widehat{\theta}(u) & \text{if } \max\{E_i : 1 \leq i \leq n-1\} \leq 2 \\ 1 \wedge \widehat{\theta}^*(u) & \text{if } \max\{E_i : 1 \leq i \leq n-1\} > 2 \end{cases} \tag{4.4}$$

where

$$\widehat{\theta}(u) = \frac{2(\sum_{i=1}^{n-1} E_i)^2}{(n-1)\sum_{i-1}^{n-1} E_i^2} \qquad \widehat{\theta}^*(u) = \frac{2\{\sum_{i-1}^{n-1}(E_i - 1)\}^2}{(n-1)\sum_{i-1}^{n-1}(E_i-1)(E_i-2)}.$$

See Ferro and Segers (2003) for further details.

### 4.2.2 Tail dependence coefficient

Ledford and Tawn (2003) suggest estimating the extremal dependence of a stationary sequence $\{X_t\}$ with the coefficient of tail dependence. Let $\{X_t\}$ be a stationary time series with unit-Fréchet margin, i.e.

$$F(x) = \exp(-x^{-1}), \qquad x > 0$$

then under weak regularity conditions, we have that as $x \to \infty$,

$$\Pr(X_{t+h} > x | X_t > x) = \frac{\mathcal{L}_h(x)}{x^{1/\eta_h - 1}} \tag{4.5}$$

where $\eta_h \in (0,1)$ and $\mathcal{L}_h(x)$ is such that $\lim_{x\to\infty} \mathcal{L}_h(tx)/\mathcal{L}_h(x) \to 1$. The coefficient of tail dependence between $X_{t+h}$ and $X_t$ is defined as $\Lambda_h = 2\eta_h - 1$ where $-1 < \Lambda_h \leq 1$.

The limiting conditional tail probability in Equation (4.5) is positive if $\eta_h = 1$ and $\mathcal{L}_h(x) \not\to 0$, otherwise it is equal to 0 if $\eta_h < 1$ or if $\eta_h = 1$ and $\mathcal{L}_h(x) \to 0$ as $x \to \infty$. Consequently, $\Lambda_h$ provides a measure of serial dependence between extreme values $h$ lags apart. In particular, if $\mathcal{L}_h(x) \not\to 0$ then $\Lambda_h = 1$ implies asymptotic dependence, meaning that conditional on a large observation at time $t$ there is a positive probability that another similar large observation occurs at time $t + h$. On the other hand, $\Lambda_h < 1$ implies asymptotic independence.

An estimator of the tail dependence coefficient can be obtained as follows. Given a stationary series $\{X_t\}$, apply the probability integral transform and obtain a stationary unit-Fréchet series $\{Y_t\}$,

$$Y_t = -\frac{1}{\log[\widehat{F}_x(X_t)]}$$

where $\widehat{F}_x$ is the empirical distribution function. Define the sequence of pairwise-minima of observations $h$ lags apart as $\{V_t^{(h)}\}$, where $V_t^{(h)} = \min(Y_t, Y_{t+h})$. As a consequence of $Y_t$ being Fréchet, we have that $V^{(h)}$ is regularly varying,

$$\Pr(V^{(h)} > x) = \frac{\mathcal{L}_h(x)}{x^{1/\eta_h}}$$

and the tail index $\eta_h$ can easily be estimated with the Hill (1975) estimator

$$\widehat{\eta}_h = \left\{ \frac{1}{N_u} \sum_{i=1}^{n} \log\left(\frac{V_i^{(h)}}{u}\right) I\{V_i^{(h)} > u\} \right\}^{-1}$$

where $u$ is a high threshold and $N_u$ the number of observations exceeding $u$. The transformation $\widehat{\Lambda}_h = 2\widehat{\eta}_h - 1$ gives an estimate of the tail dependence coefficient for observations $h$ lags apart.

### 4.2.3 Extremogram

Davis et al. (2009) introduce another tool for measuring the extremal dependence in a strictly stationary series: the *extremogram*. In many respects, one can view the extremogram as the extreme-value analogue of

the autocorrelation function of a stationary process. For a strictly stationary $\mathbb{R}^d$-valued time series $X_t$, the extremogram is defined as a limiting sequence given by

$$\gamma_{AB}(h) = \lim_{n\to\infty} n\text{cov}(I_{[a_n^{-1}X_t\in A]}, I_{[a_n^{-1}X_{t+h}\in B]}) \tag{4.6}$$

where $I_{[\cdot]}$ is the indicator function, $a_n$ is a suitably chosen normalization sequence and $A, B$ are two fixed sets bounded away from zero. Davis et al. (2009) use the theory of regular variation to provide sufficient conditions for the existence of the limits $\gamma_{AB}(h)$ in (4.6). Note that setting $d = 1$ and $A = B = (1, \infty)$ one recovers the tail dependence coefficient introduced in Section 4.2.2 for the vector $(X_t, X_{t+h})$. Indeed, if $a_n$ is such that $nP(X > a_n) \sim 1$ as $n \to \infty$ then,

$$\begin{aligned} n\text{cov}(I_{[X_t>a_n]}, I_{[X_{t+h}>a_n]}) &\sim \frac{\Pr(X_t > a_n, X_{t+h} > a_n) - (\Pr(X_t > a_n))^2}{\Pr(X_t > a_n)} \\ &\sim P(X_{t+h} > a_n | X_t > a_n) \end{aligned}$$

A natural estimator for the extremogram can be obtained by replacing the limiting sequence $a_n$ in Equation (4.6) with a high quantile of the process. Defining $a_m$ as the $m$th upper order statistics of the sample $X_1, \ldots, X_n$, the *sample extremogram* is given by

$$\widehat{\gamma}_{AB}(h) = \frac{\sum_{t=1}^{n-h} I_{[a_m^{-1}X_t\in A, a_m^{-1}X_{t+h}\in B]}}{\sum_{t=1}^{n} I_{[a_m^{-1}X_t\in A]}}$$

In order to have a consistent result, $m = m_n \to \infty$ with $m/n \to 0$ as $n \to \infty$. Under suitable mixing conditions and other distributional assumptions that ensure the limit in Equation (4.6) exists, Davis and Mikosch (2009a) show that $\widehat{\gamma}_{AB}$ is asymptotically normal,

$$\sqrt{n/m}(\widehat{\gamma}_{AB} - \gamma_{AB:m}(h)) \xrightarrow{d} N(0, \sigma^2_{AB}(h)) \tag{4.7}$$

where

$$\gamma_{AB:m}(h) = \Pr(a_m^{-1}X_{t+h} \in B | a_m^{-1}X_t \in A) \tag{4.8}$$

Equation (4.8) is usually referred to as the pre-asymptotic extremogram. This asymptotic result requires some clarifications:

(i) $\sigma^2_{AB}(h)$ is based on an infinite sum of unknown quantities and is often difficult to estimate. For this reason Davis et al. (2012) develop a bootstrap procedure to approximate the distribution of $(\widehat{\gamma}_{AB} - \gamma_{AB:m})$ and construct asymptotically correct confidence bands for the PA-extremogram.

(ii) $\gamma_{AB:m}(h)$ can be considered a finite approximation of its asymptotic limit $\gamma_{AB}(h)$, and cannot always be replaced by this latter in Equation (4.7), see Davis et al. (2013) for details.

In the rest of the paper we will only consider the extremogram for univariate time series, i.e. $d = 1$. Setting the threshold $u$ at a high quantile of the unconditional distribution and fixing the regions $A = B = [1, +\infty]$, the sample extremogram at lag $h$ simply becomes

$$\widehat{\gamma}(h) = \frac{\sum_{t=1}^{n-h} I_{[X_t > u, X_{t+h} > u]}}{\sum_{t=1}^{n} I_{[X_t > u]}}.$$

## 4.3 Do financial returns exhibit extremal dependence?

Although the extremal dependence is an important aspect of financial returns, empirical investigations of this character have not received the deserved attention. Small assessment of such behaviour can be found in the empirical sections of Davis et al. (2009), Davis et al. (2012) and Liu and Tawn (2013), but to the best of our knowledge, a sound comprehensive analysis on a large dataset has never been performed. The purpose of this section is to fill this gap.

We investigate whether the upper and lower tails of the daily returns exhibit extremal dependence. Furthermore, given the relevance of the volatility in the financial econometrics literature, we think that insights on its extremal behaviour are important from a modelling perspective. To this end, we inspect the serial dependence in the extreme observations of the squared returns (R2) and the realized variance (RV). These are both non-parametric estimators of the variation of the price path of

an asset, but while R2 is built upon close-to-close daily observations, RV is obtained as the sum of squares of intra-day returns recorded during the opening hours of the stock exchange. Formally, let $p_t$ denote the log-price of an asset at time $t$ and $r_{t,\Delta} = p_t - p_{t-\Delta}$ the discretely sampled $\Delta$-period return. The RV on day $t$ is then defined as

$$RV_t = \sum_{j=1}^{N} r^2_{t-1+(j\cdot\Delta)} \tag{4.9}$$

### 4.3.1 Data description

The empirical analysis is based on the Oxford-Man Institute "Realized Library" version 0.2 (Heber et al., 2009). We consider 17 different stock indices from the beginning of 2000 to the end of 2014, see Table 47. For each asset, we consider the daily returns and the 5-min RV.

**Table 47: Data description**. Time series considered (*Asset*), the abbreviation (*Abbr.*) that we use throughout and their length (*T*). The starting date and the ending dates of the samples are respectively, 2 January 2000 and 31 December 2014. Differences in the number of observations arise from the different closures of the stock exchanges.

| Asset | Abbr. | T | Asset | Abbr. | T |
|---|---|---|---|---|---|
| Amsterdam Exchange Index | AEX | 3816 | IBEX35 | IBX | 3782 |
| All Ordinaries Index | AOI | 3743 | IPC Mexico | IPC | 3748 |
| Bovespa Index | BVP | 3664 | Korea Composite Index | KCI | 3690 |
| CAC40 | CAC | 3817 | Nasdaq 100 | NSQ | 3747 |
| DAX30 | DAX | 3795 | Nikkei 225 | NK | 3630 |
| Dow Jones Industrial | DJ | 3746 | Russel 2000 Index | RUS | 3745 |
| Euro Stoxx 50 | ESX | 3794 | SP500 | SPX | 3744 |
| FTSE MIB | MIB | 3778 | Swiss Market Index | SMI | 3749 |
| FTSE100 | FT | 3764 | | | |

### 4.3.2 Extremal index

We start by considering the extremal index, as it is the most natural parameter to describe the degree of dependence in the extremes of a time series. For each asset in Table 47, we estimate the extremal index on

the upper tail of the following four time series: the daily log-returns $r_t$, the negated returns $l_t = -r_t$, and the two variance proxies, $r_t^2$ and $RV_t$. We use the intervals estimator described in Section 4.2.1, setting the threshold $u$ at the $95$th empirical quantile of each series, and compute the $95\%$-confidence bounds with the bootstrap procedure described in Ferro and Segers (2003). Recall that an extremal index $\theta < 1$ implies extremal dependence. The estimates $\widehat{\theta}$ reported in Table 48 are well below one for each series of extremes across the different assets. Furthermore, the bootstrap confidence intervals do not contain the value of one in any case, suggesting that both tails and the variance exhibit extremal dependence.

**Table 48: Extremal index estimates**. Estimated values (*est.*) of $\widehat{\theta}$ at the threshold corresponding to the $95$th quantile, and 95%-bootstrap confidence bounds ($q_{0.025}$, $q_{0.975}$).

| | **Upper tail** | | | **Lower tail** | | | **Squared return** | | | **Realized variance** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *est.* | $q_{0.025}$ | $q_{0.975}$ | *est.* | $q_{0.025}$ | $q_{0.975}$ | *est.* | $q_{0.025}$ | $q_{0.975}$ | *est.* | $q_{0.025}$ | $q_{0.975}$ |
| AEX | 0.29 | 0.23 | 0.50 | 0.23 | 0.16 | 0.41 | 0.18 | 0.11 | 0.40 | 0.06 | 0.04 | 0.17 |
| AOI | 0.29 | 0.20 | 0.61 | 0.28 | 0.19 | 0.64 | 0.20 | 0.13 | 0.46 | 0.12 | 0.08 | 0.42 |
| BVP | 0.54 | 0.42 | 0.82 | 0.46 | 0.36 | 0.66 | 0.35 | 0.25 | 0.56 | 0.24 | 0.14 | 0.44 |
| CAC | 0.27 | 0.22 | 0.46 | 0.26 | 0.18 | 0.60 | 0.19 | 0.13 | 0.49 | 0.09 | 0.06 | 0.23 |
| DAX | 0.29 | 0.23 | 0.45 | 0.27 | 0.18 | 0.59 | 0.19 | 0.13 | 0.34 | 0.08 | 0.04 | 0.29 |
| DJ | 0.31 | 0.25 | 0.45 | 0.26 | 0.18 | 0.47 | 0.08 | 0.05 | 0.47 | 0.08 | 0.04 | 0.25 |
| ESX | 0.35 | 0.28 | 0.53 | 0.29 | 0.20 | 0.65 | 0.16 | 0.11 | 0.41 | 0.10 | 0.06 | 0.22 |
| MIB | 0.17 | 0.12 | 0.64 | 0.37 | 0.31 | 0.51 | 0.14 | 0.10 | 0.45 | 0.12 | 0.07 | 0.28 |
| FT | 0.26 | 0.20 | 0.46 | 0.22 | 0.16 | 0.42 | 0.12 | 0.08 | 0.32 | 0.10 | 0.06 | 0.27 |
| IBX | 0.31 | 0.25 | 0.59 | 0.18 | 0.13 | 0.69 | 0.15 | 0.11 | 0.43 | 0.15 | 0.11 | 0.33 |
| IPC | 0.38 | 0.29 | 0.57 | 0.32 | 0.25 | 0.47 | 0.22 | 0.16 | 0.39 | 0.12 | 0.07 | 0.38 |
| KCI | 0.43 | 0.35 | 0.57 | 0.31 | 0.25 | 0.47 | 0.39 | 0.31 | 0.51 | 0.09 | 0.06 | 0.34 |
| NSQ | 0.22 | 0.17 | 0.35 | 0.14 | 0.10 | 0.43 | 0.09 | 0.06 | 0.29 | 0.05 | 0.03 | 0.17 |
| NK | 0.47 | 0.36 | 0.77 | 0.60 | 0.51 | 0.73 | 0.32 | 0.25 | 0.45 | 0.20 | 0.13 | 0.42 |
| RUS | 0.26 | 0.20 | 0.44 | 0.36 | 0.28 | 0.51 | 0.13 | 0.08 | 0.46 | 0.09 | 0.05 | 0.33 |
| SPX | 0.25 | 0.20 | 0.39 | 0.12 | 0.09 | 0.56 | 0.09 | 0.06 | 0.41 | 0.07 | 0.04 | 0.23 |
| SMI | 0.30 | 0.23 | 0.51 | 0.28 | 0.20 | 0.54 | 0.20 | 0.14 | 0.42 | 0.06 | 0.04 | 0.19 |

### 4.3.3 Tail dependence coefficient

Liu and Tawn (2013) propose a statistical test based on the tail dependence coefficient described in Section 4.2.2 to draw insights on the extremal dependence of the upper and lower tails of the asset return distri-

bution[2]. They propose to test whether the tail behaviour of the returns are better represented by a GARCH model, which implies extremal dependence, or a SV model with Gaussian innovations, which entails extremal independence. More formally, given that for adjacent observations $(h = 1)$ a GARCH process presents a coefficient of tail dependence $\Lambda_h = 1$, while a Gaussian SV process have $\Lambda_h = 0$, they substantially test the following hypothesis,

$$\mathcal{H}_0 : \Lambda_1 = 1 \qquad \text{vs} \qquad \mathcal{H}_1 : \Lambda_1 = 0$$

Once an estimate of the tail dependence coefficient $\widehat{\Lambda}_1$ is obtained with the procedure outlined in Section 4.2.2, the decision of rejecting $\mathcal{H}_0$ is based on an optimal cut-off value $\lambda^* \in [0, 1]$. If $\widehat{\Lambda}_1 \geq \lambda^*$ then the series is better fitted by a GARCH, hence it presents extremal dependence, otherwise the SV model is favoured and the extremes are independent. The value of $\lambda^*$ is obtained with a Monte Carlo procedure detailed in Appendix C.1.2. For a time series of $4000$ observations and a threshold level at the $95$th empirical quantile we find that $\lambda^* = 0.18$ for both tails (due to symmetry arguments).

Table 49 reports the values of $\widehat{\Lambda}_1$ for the upper and lower tails of the 17 assets considered. In bold, the time series for which the null hypothesis of GARCH tails is rejected favouring a SV process. The results highlight that $\mathcal{H}_0$ is rejected only on one occasion, thus the series present extremal dependence both in the upper and lower tails, confirming the results obtained with the extremal index.

### 4.3.4 Sample extremogram

The preceding analysis aimed to ascertain the presence of extremal dependence. We now turn our attention to the magnitude and the persistence of such dependence. By means of the sample extremogram introduced in Section 4.2.3, we check the degree of dependence in the extremes up to $100$ lags.

[2]This test can be only applied to the return time series, therefore R2 and RV are not considered here.

**Table 49: Tail dependence coefficient estimates**. Values of $\widehat{\Lambda}_1$ for the upper and lower tails of the different assets. In bold the time series for which $\mathcal{H}_0$ is rejected.

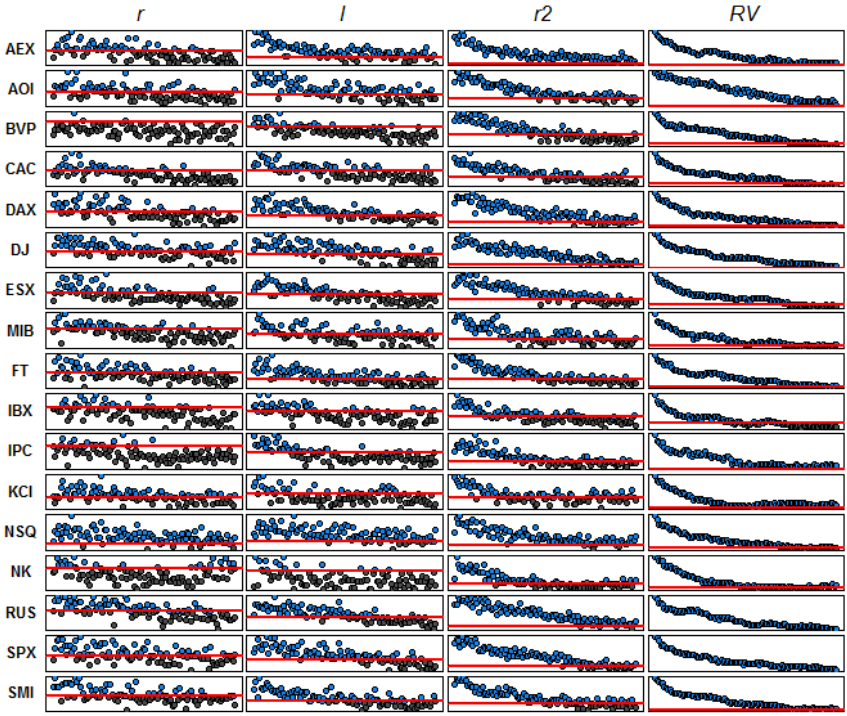
| Asset | Upper tail | Lower Tail | Asset | Upper tail | Lower tail |
|---|---|---|---|---|---|
| AEX | 0.23 | 0.48 | IBX | 0.25 | 0.27 |
| AOI | 0.39 | 0.42 | IPC | 0.34 | 0.46 |
| BVP | **0.11** | 0.45 | KCI | 0.39 | 0.40 |
| CAC | 0.37 | 0.37 | NSQ | 0.41 | 0.51 |
| DAX | 0.32 | 0.43 | NK | 0.35 | 0.21 |
| DJ | 0.38 | 0.42 | RUS | 0.27 | 0.28 |
| ESX | 0.25 | 0.35 | SPX | 0.33 | 0.43 |
| MIB | 0.31 | 0.33 | SMI | 0.53 | 0.53 |
| FT | 0.19 | 0.38 | | | |

The first two columns of Figure 45 report the sample extremograms for the upper and lower tails of the 17 assets, obtained at the threshold level $u$ corresponding to the *95*th quantile of $r_t$ and $l_t$ respectively. We also report the 99%-confidence bound obtained under the assumption of independence with the permutation procedure described in Davis et al. (2012). Values of the sample extremogram extending beyond this bound support evidence of extremal dependence at the corresponding lag. Overall, most of the series present a degree of extremal dependence that persists for more than one lag. Exceptions are BVP, which does not exhibit any dependence in the upper tail, consistent with the result obtained with the tail dependence coefficient, and AEX, IBX, and IPC which present a mild degree of dependence. The lower tail tends to exhibit stronger tail dependence, with a higher magnitude at the early lags and a higher degree of persistence.

The third and the fourth columns of Figure 45 show the sample extremograms at the *95*th quantile for R2 and RV. Dependence in the extremes of the second moment is even stronger than that observed at the return level. All the series exhibit a high degree of dependence that lasts for more than $50$ lags. This evidence on the return variance agrees with what we observe in the returns. Indeed, as put forth in Davis and

Mikosch (2009a) and Mikosch and Rezapur (2013), extremal dependence in the volatility process is crucial to generate extremal dependence in the return process.

**Figure 45: Sample Extremograms**. The points show the values of the sample extremogram at the 95th quantile up to 100 lags. The 99%-confidence bound in red is equal to 0.09 in each panel. Blue points are those exceeding the confidence bound and highlight the significance of the dependence at the corresponding lag. Grey points correspond to not significant lags.




### 4.3.5 Summary

Overall, this analysis indicates that extremal dependence is strong and persistent in both tails of financial returns, and in both R2 and RV. Since

some degree of arbitrariness is implied by the choice of the *95*th quantile to run the extreme value analysis, a sensitivity analysis is opportune. We repeat the study at the *97*th quantile, and the results, reported in Appendix C.1, confirm the conclusions above.

## 4.4 Extremal behaviour of HF-based volatility models

The empirical analysis in Section 4.3 shows that financial returns present extremal dependence, therefore econometric models able to capture this behaviour are required. Extremal properties of standard econometric models, such as GARCH and SV, have been widely studied. Recent models based on HF data prove to outperform the standard GARCH and SV models in terms of forecast (Brownlees and Gallo, 2010; Engle and Gallo, 2006; Hansen et al., 2012; Shephard and Sheppard, 2010), however nothing is known about their extremal behaviour. This section fills the gap.

To study the extremal properties of a class of HF-based volatility processes, we set up an appropriate environment where a general realized measure characterizes the conditional returns distribution. In Section 4.3 we consider the realized variance, but there exist many other measures that can be used instead (Barndorff-Nielsen and Shephard, 2007). Let $r_t$ be the return and $x_t$ the realized measure at time $t$, we will consider the following general framework,

$$
\begin{aligned}
r_t &= \sigma_t Z_t && (4.10)\\
\sigma_t^2 &= \omega + \kappa x_{t-1}, \qquad \omega, \kappa \geq 0 && (4.11)\\
x_t &= m(\mathcal{F}_{t-1}; \beta; \eta_t) && (4.12)
\end{aligned}
$$

where $Z_t \sim iid(0,1)$, $m(\cdot)$ is a general function defined on the information set available at time $t-1$, $\mathcal{F}_{t-1}$, and depending on the vector of parameters $\beta$ and the error term $\eta_t \sim iid(0, \sigma_\eta^2)$. We will refer throughout to these three equations as the *return equation*, the *variance equation*, and the *realized equation*. Equation (4.11) nests inside the HEAVY structure of Shephard and Sheppard (2010), as they define $\sigma_t^2 = \omega + \kappa RV_{t-1} + \delta\sigma_{t-1}^2$.

The processes defined by Equations (4.10)-(4.12) can be considered latent volatility processes, but the fact that we rely on the observable $x_t$ represents a crucial difference with the standard class of SV models.

The extremal behaviour of such a general class of processes strongly depends on the specification of the function $m(\cdot)$ and the specification of the error term $\eta_t$. We provide theoretical arguments for the existence of HF-based volatility processes that are asymptotically independent and HF-based volatility processes that instead present extremal dependence. We then use Monte Carlo simulations to obtain a simulated version of the pre-asymptotic extremogram of these processes and compare it with the sample extremogram obtained from the real data[3]. Indeed, proving asymptotic dependence between consecutive observations is not enough for empirical purposes, as one also need to know whether the degree and the persistence of the extremal dependence implied by a process is consistent with that observed in the real data.

In what follows, we start considering a class of models that are asymptotically independent and then show how extreme dependent HF-based volatility processes can be constructed following Mikosch and Rezapur (2013). In order to be coherent with the empirical analysis of Section 4.3, we will only consider the RV as realized measure, but this does not prevent the use of other measures.

### 4.4.1 Asymptotically independent HF-based volatility processes

Let $r_t$ and $RV_t$ be the return and the realized variance at time $t$ respectively. The first specification we consider for the *realized equation* is a logarithmic HAR model (Andersen et al., 2007; Corsi, 2009) with Gaussian innovations $\eta_t \sim N(0, \sigma_\eta^2)$. This simple three-factor model has attracted lot of attention in the financial econometrics literature because of its ability to capture the long-range pattern in the volatility decay. Casting this

[3]Note that we refer to the simulated pre-asymptotic extremogram of the process and not to the extremogram of the process because this would require deriving a specific asymptotic limit. In contrast, the simulated pre-asymptotic extremogram is obtained by applying the sample extremogram to a sequence of random observations generated by the process, and it gives insights on the extremal behaviour only at a finite level.

model in the environment defined by Equations (4.10)-(4.12), we obtain the **HAR** process,

$$\begin{aligned} r_t &= \sigma_t Z_t, \\ \sigma_t^2 &= \omega + \kappa RV_{t-1}, \\ \log RV_t &= \beta_0 + \beta_d \log RV_{t-1} + \beta_w \log RV_{t-1}^{(5)} + \beta_m \log RV_{t-1}^{(22)} + \eta_t, \end{aligned} \tag{4.13}$$

where $Z_t \sim N(0,1)$, $\eta_t \sim N(0, \sigma_\eta^2)$, $\log RV_{t-1}^{(h)} = \frac{1}{h}\sum_{j=1}^{h} \log RV_{t-j}$ with $\log RV_{t-1}^{(1)} \equiv \log RV_{t-1}$, and $(\beta_d + \beta_w + \beta_m) < 1$ to guarantee stationarity. A logarithmic specification for $RV_t$ is used because it does not require positivity constraints on the parameters, and it is preferable from an econometric perspective as $\log RV_t$ exhibits a bell-shaped distribution. The extremal behaviour of such process is defined in Proposition 1.

**Proposition 1.** *Consider the stationary process in (4.13), then the sequences* $\{RV_t\}$, $\{\sigma_t\}$ *and* $\{|r_t|\}$ *are all asymptotically independent.*

*Proof.* Note first that the realized equation of the process in (4.13) can be written as an autoregressive (AR) process of order 22,

$$\begin{aligned} \log RV_t &= \beta_0 + \beta_d \log RV_{t-1} + \beta_w \underbrace{\log RV_{t-1}^{(5)}}_{\frac{1}{5}\log RV_{t-1} + \cdots + \frac{1}{5}\log RV_{t-5}} \\ &\quad + \beta_m \underbrace{\log RV_{t-1}^{(22)}}_{\frac{1}{22}\log RV_{t-1} + \cdots + \frac{1}{22}\log RV_{t-22}} + \eta_t \\ &= \beta_0 + (\beta_d + \tfrac{1}{5}\beta_w + \tfrac{1}{22}\beta_m)\log RV_{t-1} + (\tfrac{1}{5}\beta_w + \tfrac{1}{22}\beta_m)\log RV_{t-2} \\ &\quad + \cdots + \tfrac{1}{22}\beta_m \log RV_{t-22} + \eta_t \\ &= \widetilde{\beta}_0 + \widetilde{\beta}_1 \log RV_{t-1} + \cdots + \widetilde{\beta}_{22} \log RV_{t-22} + \eta_t \end{aligned}$$

where $\widetilde{\beta}_0 = \beta_0$, $\widetilde{\beta}_1 = \beta_d + \frac{1}{5}\beta_w + \frac{1}{22}\beta_m$ and so on. Therefore, $\log RV_t$ is a Gaussian linear process. Note that $\log \sigma_t^2$ inherits the behaviour from $\log RV_{t-1}$ up to a shift in scale and location. Without loss of generality, we can set the constants $(\omega, \kappa) = (0, 1)$ so that $\log \sigma_t^2 = \log RV_{t-1}$ and the asymptotic independence of $RV_t$, $\sigma_t$ and $|r_t|$ follows from Theorem 1 of Breidt and Davis (1998). □

There is wide evidence of an asymmetric effect between the returns and the volatility, which goes under the name of *leverage* effect (Glosten et al., 1993). We consider also an asymmetric model as in Maheu and McCurdy (2011), and obtain the *lev***HAR** process,

$$\begin{aligned} r_t &= \sigma_t Z_t, \\ \sigma_t^2 &= \omega + \kappa RV_{t-1}, \\ \log RV_t &= \beta_0 + \beta_d \log RV_{t-1} + \beta_w \log RV_{t-1}^{(5)} + \beta_m \log RV_{t-1}^{(22)} + \gamma Z_t + \eta_t, \end{aligned} \tag{4.14}$$

where $Z_t \sim N(0,1)$, $\eta_t \sim N(0, \sigma_\eta^2)$, $(\beta_d + \beta_w + \beta_m) < 1$ to guarantee stationarity, and $\gamma$ parameter accounts for the asymmetric feedback between the innovations and the volatility. This process presents the same extremal behaviour of the process in (4.13).

Although these HF-based volatility processes are asymptotically independent, knowledge of the dependence they generate above a high quantile is still relevant in practical applications such as risk management and regulation. We compute the simulated pre-asymptotic extremogram implied by these processes and compare it with the sample extremogram of a real time series. We consider only the time series of the S&P500 which has $N = 3743$ observations. We set the parameters of each process at values corresponding to the Maximum Likelihood (ML) estimates obtained by fitting the models to the S&P500, and generate $B = 100$ samples of $N$ observations for the $r_t$ and $RV_t$ series. For each sample $b \in \{1, \dots, B\}$, we estimate the sample extremogram of both the upper and lower tails of the returns, R2 and RV. This allows us to obtain the simulated pre-asymptotic extremograms of the series $r_t$, $l_t$, $r_t^2$ and $RV_t$, and compute the confidence interval associated to them. We can thus compare the sample extremograms obtained on the S&P500 with those implied by the random processes.

In the second and third rows of Figure 46, we report the simulated pre-asymptotic extremograms computed at the *95*th quantile for the HAR and the *lev*HAR processes with the corresponding 95%-confidence intervals. As can be seen from the grey-shaded area nicely covering the sample extremogram of RV, both models appropriately account for the

slow decay of the dependence in the extremes of RV. However, they fail to capture the dependence in both the tails of the returns and in the extremes of R2. For comparison purposes, the first row of Figure 46 reports the simulated pre-asymptotic extremograms of a GJR-GARCH(1,1) process (Glosten et al., 1993). The grey-shaded areas perfectly covering the sample extremograms obtained on the S&P500 suggest that the GARCH process perform much better than the HAR and *lev*HAR processes, particularly in the lower tail and R2. This result highlights the importance of a model endowed with the appropriate extremal properties to characterize the behaviour of financial returns. Finally, comparing the HAR with the *lev*HAR, it seems that adding the leverage parameter improves the results. In particular, not only it allows to capture the asymmetric degree of dependence in the two tails, but it produces a stronger dependence in the extremes of the volatility.

### 4.4.2 Asymptotically dependent HF-based volatility processes

We have seen that the HAR and *lev*HAR processes are asymptotically independent. We consider two different approaches to obtain extremal dependence in a HF-based volatility process. The first one requires changing the structure of the *realized equation* with respect to the one used in Section 4.4.1. The second one involves the modelling of the variance of the variance, $\sigma_\eta^2$.

Mikosch and Rezapur (2013) show that a stochastic volatility process with recurrent structure (Kesten, 1973) exhibits extremal dependence. We consider the stationary **MEM** process which takes the following form

$$\begin{array}{rcl} r_t & = & \sigma_t Z_t, \\ \sigma_t^2 & = & \omega + \kappa RV_{t-1} + \delta\sigma_{t-1}^2, \\ RV_t & = & \mu_t \eta_t^2, \\ \mu_t & = & \beta_0 + \beta_1 RV_{t-1} + \beta_2 \mu_{t-1}, \end{array} \tag{4.15}$$

where $Z_t \sim N(0,1)$, $\eta_t \sim N(0,1)$, $\omega, \kappa, \delta \geq 0$, $\delta \in (0,1]$, $\beta_0, \beta_1, \beta_2 \geq 0$, and $\beta_1 + \beta_2 \in [0,1)$. The *realized equation* is defined as a MEM (Brownlees and Gallo, 2010; Engle, 2002; Engle and Gallo, 2006)with the mean fol-

lowing a recurrent equation. This structure was proposed in Shephard and Sheppard (2010) that also add the autoregressive term $\sigma_{t-1}^2$ in the *variance equation*, so we include it to be coherent with their framework though it is not relevant to prove the extremal dependence property in Proposition 2.

**Proposition 2.** *Consider the stationary process in (4.15), then the sequences* $\{RV_t\}$, $\{\sigma_t\}$ *and* $\{|r_t|\}$ *present extremal dependence, i.e. the extremal indexes* $\theta_{RV}$, $\theta_\sigma$ *and* $\theta_{|r|}$ *are strictly lower than one.*

*Proof.* Note first that $\mu_t$ can be represented as a stochastic recurrent equations Kesten (1973),

$$\mu_t = \beta_0 + \mu_{t-1}\left(\beta_1 \eta_{t-1}^2 + \beta_2\right).$$

Assume, without lost of generality that $(\omega, \kappa, \delta) = (0, 1, 0)$ so that $\sigma_t^2 = RV_{t-1}$. Extremal dependence of $RV_t$, $\sigma_t$ and $|r_t|$ directly follows from Theorem 4.4 of Mikosch and Rezapur (2013). □

The fourth row of Figure 46 shows the simulated pre-asymptotic extremograms computed at the *95*th quantile for an asymmetric MEM process and the corresponding 95%-confidence intervals. Note that the leverage effect is obtained with the realized semi-variance as suggested in Shephard and Sheppard (2010), rather than resorting on the lagged sign of the innovations as in (4.14). This process perfectly captures the dependence in both tails of the returns and R2, outperforming also the GJR-GARCH process.

Corsi et al. (2008) find strong evidence of time-variation and dependence in the variance of the variance and propose to model this behaviour. Following their intuition, we relax the assumption that $\eta_t \sim iid(0, \sigma_\eta^2)$ and propose the **HAR-G** process,

$$\begin{aligned} r_t &= \sigma_t Z_t, \\ \sigma_t^2 &= \omega + \kappa RV_{t-1}, \\ \log RV_t &= \beta_0 + \beta_d \log RV_{t-1} + \beta_w \log RV_{t-1}^{(5)} + \beta_m \log RV_{t-1}^{(22)} + \eta_t \\ \sigma_{\eta,t}^2 &= \psi_0 + \psi_1 \eta_{t-1}^2 + \psi_2 \sigma_{\eta,t-1}^2 \end{aligned} \tag{4.16}$$

with $Z_t \sim N(0,1)$ and $\eta_t \sim N(0, \sigma_{\eta,t}^2)$. To guarantee stationarity of the process we require $\beta_d + \beta_w + \beta_m < 1$ and $\psi_1 + \psi_2 < 1$. We do not present a

formal proof for the extremal behaviour of this process. However, based on the result of Borkovec (2000) showing that an AR(1)-ARCH(1) process exhibits extremal dependence, we conjecture that the extremes of $\log RV_t$ are asymptotically dependent and that this property is inherited by $\sigma_t^2$ and $|r_t|$ as a consequence of Theorem 4.4 of Mikosch and Rezapur (2013).

In the fifth and sixth rows of Figure 46, we report the simulated pre-asymptotic extremograms computed at the *95*th quantile for the HAR-G and the *lev*HAR-G processes and the corresponding 95%-confidence intervals. The latter process simply corresponds to the *lev*HAR process, but we also include GARCH dynamics to model the variance of $\log RV_t$. First, we note that both processes imply a higher degree of dependence across the four series compared to the baseline HAR and *lev*HAR processes. However, the HAR-G process is still unable to account for the whole dependence exhibited by the sample extremograms of the S&P500. This suggests that though this process might be asymptotically dependent, the degree of dependence it entails would not be sufficient to appropriately model financial returns. In contrast, the *lev*HAR-G process generates a high degree of dependence in the extremes, highlighting the importance of including a leverage parameter.

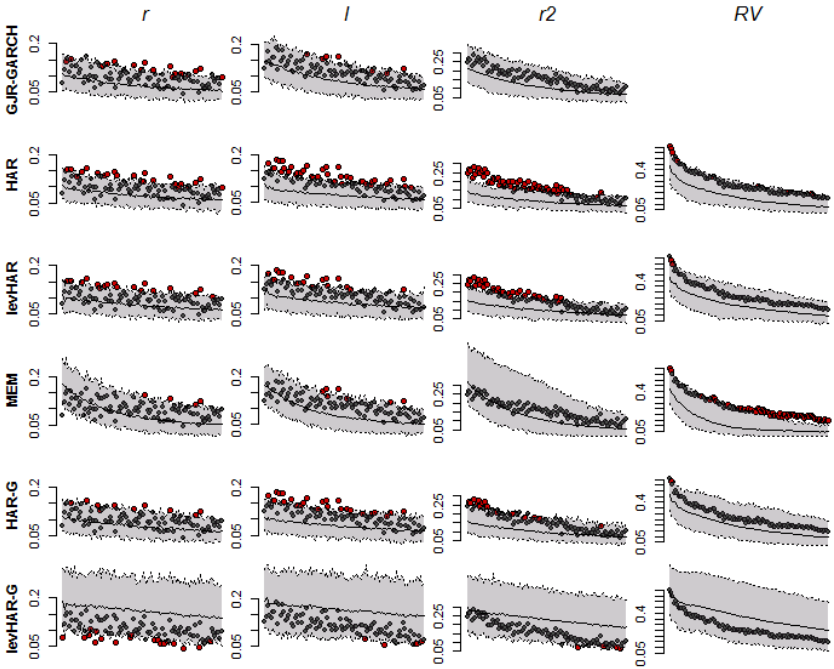
### 4.4.3 Summary

In this section, we have shown that HF-based volatility processes can exhibit extremal dependence under appropriate specifications. We have seen that the asymmetric MEM and the *lev*HAR-G processes successfully replicate the dependence in the extremes of the S&P500. To strengthen this evidence, we run the same Monte Carlo analysis at the *97*th quantile. Results reported in Appendix C.2 confirm our findings.

## 4.5 Continuous-time models

Continuous-time volatility models are widely used in practice for derivative pricing, portfolio choice and risk management, and it is thus important that their extremal behaviour is consistent with that observed in real

**Figure 46: HF-based processes**. The grey-shaded areas correspond to the 95%-confidence interval of the simulated pre-asymptotic extremograms. The points represent the first 100 lags of the sample extremograms computed at the 95th quantile of the series $r_t$, $l_t$, $r_t^2$ and $RV_t$ of the S&P500. Red points identify the lags at which the dependence exceeds the confidence bounds.



data. Extremal dependence of continuous-time models has been investigated for several classes of models (Fasen, 2009). In this section, we do not want to examine further the theoretical properties of such models, but we want to show how to make the extremal dependence a central character of the estimation strategy.

Estimation of continuous-time volatility models is complicated by the fact that the volatility is a latent process. To overcome this issue, recent works have tried to use the daily realized measures to make the latent volatility observable at discrete points in time, see Bollerslev and Zhou (2002) and Corsi and Renò (2012) amongst others. We propose to use the Simulated Method of Moments (SMM) defined in Corradi and Distaso (2006) along with extreme moment conditions to match the extremal dependence that characterizes daily financial returns with that implied by the continuous-time model.

Let $\mathcal{S}$ be a sample of observations of length $T$ and $\theta \in \Theta$ be the vector of parameters governing the dynamics of the continuous-time model to be estimated. Let $\bar{g}_T^*$ be a vector of empirical moments obtained from the sample $\mathcal{S}$ and $\tilde{g}_S(\tilde{\theta})$ be a vector of simulated moments obtained from $S$ random samples generated from the continuous-time model with parameters $\tilde{\theta}$. The SMM estimator is obtained as the minimizer of the quadratic form

$$\widehat{\theta} = \arg\min_{\theta \in \Theta} (\bar{g}_T^* - \tilde{g}_S(\theta)) W_T^{-1} (\bar{g}_T^* - \tilde{g}_S(\theta))$$

where $W_T$ is a properly chosen positive semi-definite matrix.

To account for the persistence in the volatility, we estimate a two-factor model as in Corsi and Renò (2012). Fasen (2009) shows that the GARCH diffusion of Nelson (1990) presents extremal dependence, therefore we consider this process to model the volatility dynamics. Let $p_t$ be the logarithmic price at time $t$, the model specification is

$$\begin{array}{rcl} dp_t & = & \sqrt{V_t^1} dW_t^1 + \sqrt{V_t^2} dW_t^2, \\ dV_t^1 & = & \kappa_1 \left(\omega_1 - V_t^1\right) dt + \eta_1 V_t^1 dB_t^1, \\ dV_t^2 & = & \kappa_2 \left(\omega_2 - V_t^2\right) dt + \eta_2 V_t^2 dB_t^2, \end{array} \tag{4.17}$$

where $B_t^1 \perp B_t^2$ and $W_t^1 \perp W_t^2$ are Brownian motions with $\operatorname{corr}(W_t^i, B_t^i) = \rho_i$, for $i = 1, 2$.

**Table 50: Continuous-time model estimates**. SMM estimates for the parameters of the model in Equation (4.17) obtained with the annualized RV of the S&P500.

| **Parameters** | $\kappa_1$ | $\kappa_2$ | $\omega_1$ | $\omega_2$ | $\eta_1$ | $\eta_2$ | $\rho_1$ | $\rho_2$ |
|---|---|---|---|---|---|---|---|---|
| **SMM estimates** | 0.516 | 0.005 | 0.005 | 0.034 | 1.598 | 0.190 | 0.396 | -0.943 |

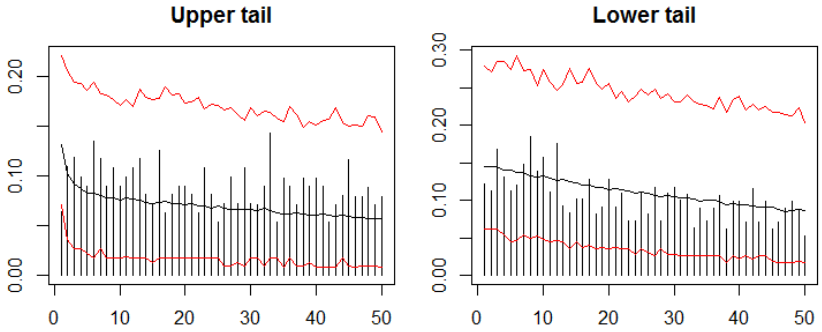
We use the following moment conditions to estimate the model: mean and variance of the Quadratic Variation (QV); the average autocorrelation of QV for lags 1-10, 11-20, 21-30 and 31-40; the average extremogram of both the upper and lower tails of the return distribution for lags 1-10. Since QV is unobservable, we replace the empirical moments for QV with those of RV. Overall we have eight moment conditions that exactly match the number of parameters of the model in (4.17).

Distinguishing between the extreme moment conditions of the upper and lower tails, one can identify the parameters $\rho_1$ and $\rho_2$. From this perspective, we add to the work of Harvey and Shephard (1996) and Corsi and Renò (2012) which try to estimate asymmetric continuous-time models.

Table 50 reports the SMM estimates for the parameters of the model in Equation (4.17) obtained on the S&P500. The results are compatible with those typically encountered in the literature (Corsi and Renò, 2012). The sum of the long-term parameters, $\omega_1 + \omega_2$, is fairly close to the unconditional mean of the annualized RV, equal to 0.033. The values of $\kappa_1$ and $\rho_1$ imply a fast-mean reverting factor positively correlated with the returns, while the values of $\kappa_2$ and $\rho_2$ indicate a slow mean-reverting factor presenting a strong negative correlation with the returns.

Figure 47 reports the simulated pre-asymptotic extremograms at the 95th quantile obtained with the parameters in Table 50 and the sample extremograms of the S&P500 for both the return tails. We can see that the sample extremograms are within the 95%-confidence bounds of the simulated extremograms implied by the two-factor GARCH diffusion, therefore we conclude that this model perfectly captures the degree of dependence in the extremes of the daily returns.

**Figure 47: Continuous-time model extremogram**. The bars show the sample extremogram for the S&P500. The black line and the red lines correspond respectively to the mean and the 95% confidence bands of the simulated extremogram of the model in Equation (4.17).




## 4.6 Conclusions

Although empirical investigations of autocorrelations in returns and volatility have long history in financial economics and econometrics, dependence in the extremes has been underrated. Understanding the extremal properties of financial time series and related models is of central importance to gauge financial risk, and this paper gives three contributions in this direction.

First, we add new evidence on the extremal behaviour of daily stock returns, finding that dependence in the extremes is strong and persistent in both tails of the return distribution, and also in the volatility process.

Second, we study the extremal behaviour of several HF-based volatility processes and show how to obtain a process which exhibits extremal dependence. Furthermore, we find that modelling the persistence in RV and the leverage effect is crucial to account for the degree of dependence observed in the extremes of financial time series.

Finally, we show how extreme moment conditions can be used within a SMM estimator to estimate asymmetric continuous-time volatility mod-

els. We find that a two-factor GARCH diffusion process perfectly model the dependence in the tails of the S&P500.

# Appendix C

## C.1 Do Financial returns exhibit extremal dependence?

This section of the Appendix reports additional analysis concerning the empirical part of the paper on the 17 equity stock indices.

### C.1.1 Extremal index

The interval estimator of the extremal index, described in Section 4.2.1, requires setting a threshold level that identifies extreme observations. In Section 4.3.2, we set this threshold at the *95*th quantile of each variable for which the extremal index is estimated. Table 51 reports the results obtained at the *97*th quantile.

### C.1.2 Tail dependence coefficient

**The cut-off value $\lambda^*$**

To test the null hypothesis $\mathcal{H}_0$ in Section 4.3.2, we consider the *parsimonious method* of Liu and Tawn (2013). This consists in fixing a value $\lambda^* \in [0, 1]$ and deciding that the series is better fitted by a GARCH model when $\widehat{\Lambda} > \lambda^*$ and preferring a SV model with Gaussian innovations otherwise.

Assume that the return $r_t$ follows a multiplicative error process $r_t = \sigma_t Z_t$ with $Z_t \sim N(0, 1)$. The volatility dynamics of the GARCH(1,1) pro-

**Table 51: Extremal index estimates**. Estimated values (*est.*) of $\widehat{\theta}$ at the threshold level corresponding to the *97*th quantile and bootstrap confidence bounds ($q_{0.025}$, $q_{0.975}$) at the 0.05 level.

| | Upper tail | | | Lower tail | | | Squared return | | | Realized variance | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *est.* | $q_{0.025}$ | $q_{0.975}$ | *est.* | $q_{0.025}$ | $q_{0.975}$ | *est.* | $q_{0.025}$ | $q_{0.975}$ | *est.* | $q_{0.025}$ | $q_{0.975}$ |
| AEX | 0.21 | 0.15 | 0.63 | 0.25 | 0.19 | 0.42 | 0.18 | 0.11 | 0.50 | 0.10 | 0.06 | 0.24 |
| AOI | 0.22 | 0.16 | 0.65 | 0.29 | 0.19 | 0.58 | 0.25 | 0.17 | 0.49 | 0.11 | 0.06 | 0.42 |
| BVP | 0.59 | 0.47 | 0.78 | 0.42 | 0.30 | 0.68 | 0.31 | 0.20 | 0.53 | 0.28 | 0.14 | 0.68 |
| CAC | 0.26 | 0.20 | 0.56 | 0.25 | 0.18 | 0.63 | 0.20 | 0.14 | 0.43 | 0.10 | 0.06 | 0.32 |
| DAX | 0.21 | 0.15 | 0.54 | 0.30 | 0.22 | 0.52 | 0.17 | 0.12 | 0.32 | 0.07 | 0.05 | 0.21 |
| DJ | 0.16 | 0.12 | 0.61 | 0.16 | 0.12 | 0.78 | 0.12 | 0.08 | 0.49 | 0.07 | 0.04 | 0.45 |
| ESX | 0.18 | 0.13 | 0.64 | 0.29 | 0.21 | 0.51 | 0.12 | 0.08 | 0.41 | 0.10 | 0.06 | 0.28 |
| MIB | 0.26 | 0.19 | 0.66 | 0.22 | 0.15 | 0.49 | 0.22 | 0.15 | 0.49 | 0.11 | 0.07 | 0.33 |
| FT | 0.24 | 0.18 | 0.46 | 0.19 | 0.14 | 0.45 | 0.16 | 0.10 | 0.34 | 0.16 | 0.10 | 0.48 |
| IBX | 0.21 | 0.17 | 0.69 | 0.25 | 0.19 | 0.59 | 0.23 | 0.16 | 0.56 | 0.13 | 0.09 | 0.44 |
| IPC | 0.33 | 0.25 | 0.53 | 0.29 | 0.22 | 0.40 | 0.22 | 0.15 | 0.37 | 0.17 | 0.10 | 0.46 |
| KCI | 0.41 | 0.34 | 0.56 | 0.50 | 0.42 | 0.65 | 0.30 | 0.23 | 0.53 | 0.14 | 0.09 | 0.52 |
| NSQ | 0.09 | 0.07 | 0.49 | 0.14 | 0.10 | 0.49 | 0.08 | 0.05 | 0.37 | 0.07 | 0.05 | 0.24 |
| NK | 0.44 | 0.33 | 0.69 | 0.48 | 0.36 | 0.67 | 0.26 | 0.18 | 0.48 | 0.22 | 0.13 | 0.47 |
| RUS | 0.18 | 0.12 | 0.41 | 0.33 | 0.23 | 0.60 | 0.13 | 0.07 | 0.55 | 0.13 | 0.07 | 0.60 |
| SPX | 0.13 | 0.10 | 0.59 | 0.15 | 0.11 | 0.62 | 0.12 | 0.07 | 0.47 | 0.08 | 0.04 | 0.50 |
| SMI | 0.28 | 0.21 | 0.54 | 0.33 | 0.24 | 0.59 | 0.20 | 0.13 | 0.44 | 0.08 | 0.05 | 0.26 |

cess and the first-order SV process of Taylor (1982) are respectively defined by the following two equations,

$$\sigma_t^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{C.1}$$

$$\log(\sigma_t^2) = \omega + \psi_1 \log(\sigma_{t-1}^2) + \nu_t \tag{C.2}$$

where $\nu_t \sim N(0, \sigma_\nu)$. To set $\lambda^*$, Liu and Tawn (2013) propose to calculate by Monte Carlo simulations the Type-I and Type-II errors of the hypothesis test

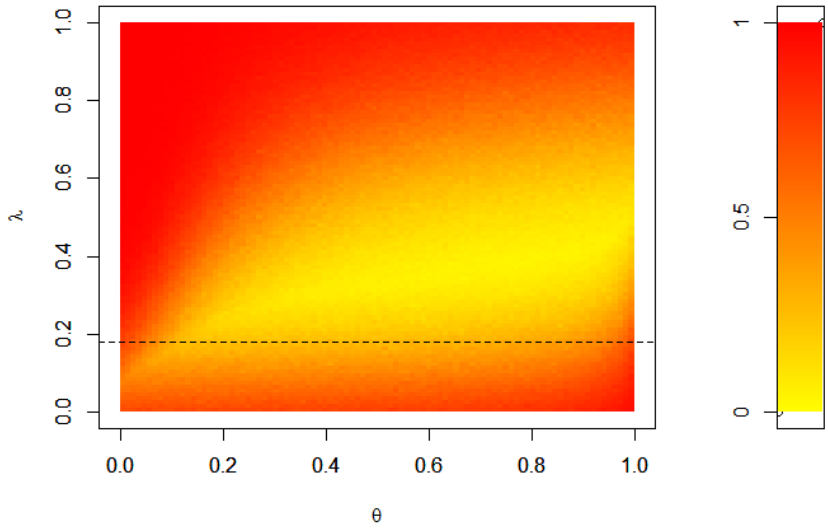
$$\mathcal{H}_0 : \text{the series is GARCH} \qquad \text{vs} \qquad \mathcal{H}_1 : \text{the series is SV}$$

for different parameters of the GARCH and SV models in Equations (C.1)-(C.2), and different cut-off points $\lambda^*$. Let $P_1(\alpha_1, \lambda)$ be the probability that a GARCH series with parameter $\alpha_1$ is identified as a SV series (Type-I error) for the cut-off point $\lambda$ and $P_2(\psi_1, \lambda)$ be the probability that a SV series with parameter $\psi_1$ is identified as GARCH (Type-II error). The optimal cut-off point $\lambda^*$ is then given by,

$$\lambda^* = \arg \min_{\lambda \in (0,1)} \left\{ \max_{\alpha_1, \psi_1 \in (0,1)} \{\max\{P_1(\alpha_1, \lambda), P_2(\psi_1, \lambda)\}\} \right\}.$$

We fix the length of each random series to $n = 4000$, consistent with the length of the time series we considered. We set the threshold level to calculate the coefficient of tail dependence at the *95*th quantile. We allow the GARCH parameter $\alpha_1$ to vary in the interval $\alpha \in [0.01, 0.98]$, with $\alpha_1 + \beta_1 = 0.98$ and the SV parameter $\psi_1$ to range within $[0.01, 0.98]$. We find that for these specific settings the decision rule cut-off is $\lambda^* = 0.18$. Figure 48 reports the values of $\max\{P_1(\alpha_1, \lambda), P_2(\psi_1, \lambda)\}$ for the different candidate cut-off points obtained using $100$ replications for each parameter.

**Figure 48:** $\max\{P_1(\alpha_1, \lambda), P_2(\psi_1, \lambda)\}$ for $\lambda^* \in (0, 1)$ with $\theta = \alpha_1 = \psi_1$. The dotted line represent the optimal cut-off value $\lambda^* = 0.18$.



**Robustness checks**

In Section 4.3.3, we estimate the coefficient of tail dependence on both tails at threshold level corresponding to the *95*th quantile. Table 52 reports the estimated coefficients at the threshold level corresponding to the *97*th quantile. With the Monte Carlo procedure described above, we have that $\lambda^* = 0.16$ in this case.

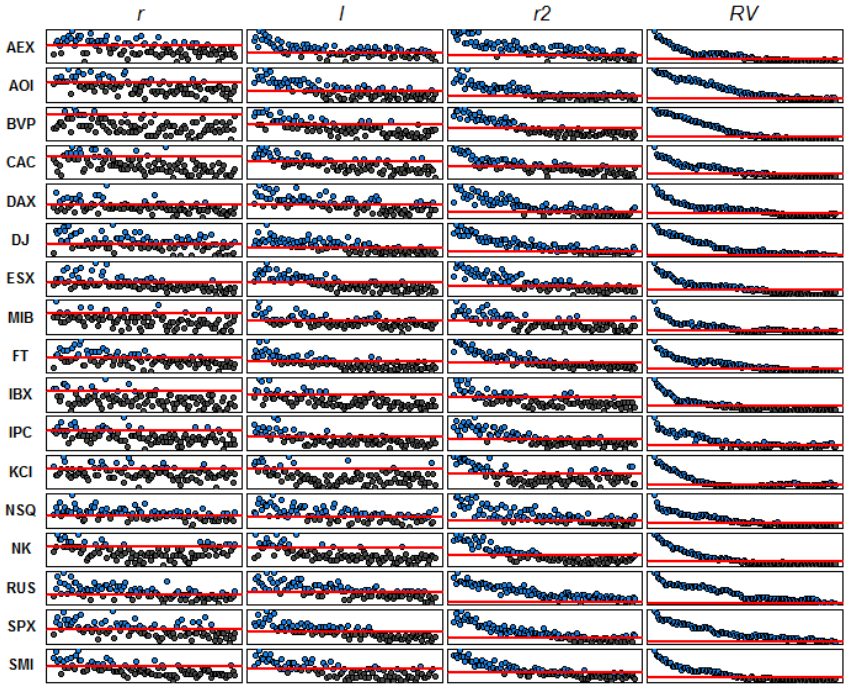
**Table 52: Tail dependence coefficient estimates**. Values of $\widehat{\Lambda}_1$ for the upper and lower tails of the different assets. In bold the time series for which $\mathcal{H}_0$ is rejected.

| Asset | Upper tail | Lower Tail | Asset | Upper tail | Lower tail |
|---|---|---|---|---|---|
| AEX | 0.36 | 0.54 | IBX | 0.37 | 0.37 |
| AOI | 0.40 | 0.44 | IPC | 0.33 | 0.55 |
| BVP | **0.15** | 0.56 | KCI | 0.37 | 0.35 |
| CAC | 0.23 | 0.34 | NSQ | 0.37 | 0.49 |
| DAX | 0.21 | 0.29 | NK | 0.45 | **0.15** |
| DJ | 0.43 | 0.49 | RUS | 0.41 | 0.36 |
| ESX | 0.43 | 0.39 | SPX | 0.40 | 0.51 |
| MIB | 0.32 | 0.27 | SMI | 0.57 | 0.52 |
| FT | 0.18 | 0.31 | | | |

## C.1.3 Extremogram

Figure 45 reports the sample extremograms obtained at the *97*th quantile for the series $r_t$, $l_t$, $RV_t$ and $r_t^2$. We also report the 99%-confidence bound obtained under the assumption of independence with the permutation procedure described in Davis et al. (2012). All the series present a degree of extremal dependence that persists for more than one lag, particularly in R2 and RV. These results are coherent with those obtained in Figure 45 at the *95*th quantile.

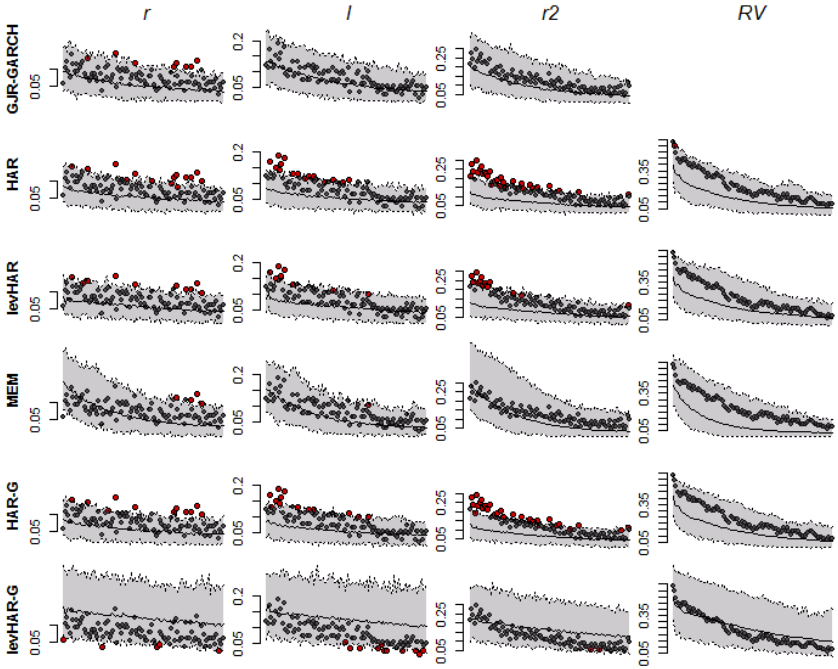
**Figure 49: Sample Extremograms**. The points show the values of the sample extremogram at the 97th quantile up to 100 lags. The 99%-confidence bound in red is equal to 0.09 in each panel. Blue points exceed the confidence bound and highlight the significance of the dependence at the corresponding lag. Grey points correspond to not significant lags.



## C.2 Extremal behaviour of HF-based volatility models

In Section 4.4, we consider exceedances over the *95*th quantile to approximate the limiting extremal dependence. To strengthen our conclusion, the simulated extremogram at the *97*th quantile are reported in Figure 50.

**Figure 50: HF-based processes**. The grey-shaded areas correspond to the 95%-confidence interval of the simulated pre-asymptotic extremograms. The points represent the first 100 lags of the sample extremograms computed at the *97*th quantile of the series $r_t$, $l_t$, $r_t^2$ and $RV_t$ of the S&P500. Red points identify the lags at which the dependence exceeds the confidence bounds.



# Chapter 5

# Realized Peaks over Threshold: a High-Frequency Extreme Value Approach for Financial Time Series

## 5.1 Introduction

*Tail-risk* has been at the heart of discussions among economists, bankers and world leaders in the aftermath of the 2008 stock market crash. Although being an elusive notion, tail risk tends to be associated to large negative events which have a positive but rather small probability of occurrence. Appropriate management of this kind of risk is of the utmost importance from both policy and regulatory perspectives and for the internal risk control of financial institutions. For this purpose, several risk measures have been defined which require forecasting quantiles deep in the lower tail of the financial returns distribution.

Traditional parametric methods based on estimation of entire densities are mostly ill-suited for the assessment of extreme quantiles. These

parametric methods strive to produce a good fit in regions where most of the data fall, potentially at the expense of a good fit in the tails. Similarly, it is well-known that non-parametric methods of density estimation such as kernel smoothing perform poorly in the tails.

Extreme Value Theory (EVT) is a branch of probability theory which focusses on extreme outcomes and provides models for them. In particular, instead of forcing a single distribution on the entire sample, this theory allows for the investigation of only the tails of the sample distribution using limit laws. Estimates of probabilities associated with quantiles even higher than the most extreme observations are then obtained by extrapolation.

Use of EVT in financial applications has become more and more common over the last fifteen years. Danielsson and de Vries (1997) and Longin (2000) use EVT to model the *unconditional* return distribution and emphasize its accuracy in predicting tail-risk. In a critical discussion of the use of EVT in risk management, Diebold et al. (2000) put forth both the opportunities and pitfalls of such applications. Their main criticism regards the time dependence that characterize financial returns. Specifically, while the probabilistic results underlying the theory hold for *iid* observations, time series in economics and finance usually do not satisfy this requirement. Despite that, they support the approach and foster its application to the tails of the *conditional* return distribution.

To model the tails of the time-varying conditional return distribution, two different paths have been taken. One consists in specifying a model for the conditional mean and variance, and then applying an EVT-based model to the tails of the standardized residuals (McNeil and Frey, 2000). If the model for the first two conditional moments completely characterize the dependence structure, then the standardized residuals should be approximately *iid*. A second strategy is to fit a non-stationary extremal model to account for the dependence in the original data (Chavez-Demoulin et al., 2005, 2014). The benefit of the first strategy is that any non-stationarities are estimated much more precisely than based on the extremal data alone. However, if the extremes of the residuals present any form of heterogeneity, it will be necessary to model them directly.

The claim that a model for the conditional mean and variance can produce *iid* residuals implies that higher conditional moments are constant over time, but the evidence seems to argue otherwise (Hansen, 1994; Harvey and Siddique, 1999). Furthermore, the presence of switching-regimes, as argued in Mikosch and Starica (2004), makes the task of pre-whitening the return time series even more difficult. These considerations support direct modelling of the extremes of the original sample, and the current paper proposes a novel dynamic extreme value approach to do so.

Dynamic EVT modelling in finance requires finding an economically sound source of information that can be used to explain the time-varying behaviour of the extremes, and finding an appropriate way to use this information within a suitable model. Current approaches exploit information in past daily exceedances. Chavez-Demoulin et al. (2005) suggest using a self-exciting process to model the probability of exceeding a high threshold of the negated return distribution (*loss distribution*), and use a time-varying Generalized Pareto (GP) distribution with the past exceedances as covariates to model the size of the exceedances. Alternatively, Chavez-Demoulin et al. (2014) model the intensity parameter of the non-homogeneous Poisson process describing the exceedance rate and the time-varying scale parameter of the GP with non-parametric Bayesian smoothers.

In this paper, we consider a completely different perspective and build models for the extremes based on measures obtained from high-frequency (HF) data, i.e. intra-daily returns. We show that HF data hold information about the tail beyond that contained in the daily exceedances. There is a growing literature which attempts to exploit HF data to enrich models for lower frequency data. Inclusion of HF data to model the conditional second moment of the return distribution has been proposed by Shephard and Sheppard (2010), Noureldin et al. (2012), Hansen et al. (2012), and Hansen et al. (2014). De Lira Salvatierra and Patton (2015) propose incorporating HF data into models for the time-varying dependence in a copula function, while Oh and Patton (2015) use a HF-based measure of correlation to disentangle the linear from the non-linear de-

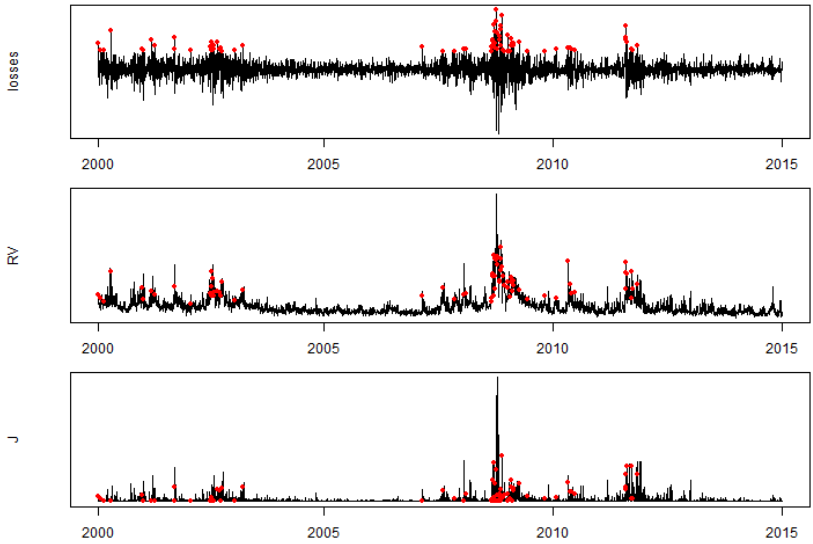
pendence in a portfolio of stocks and then model the non-linear dependence with joint-symmetric copulas. Similarly, we propose to model the tails of the conditional returns distribution with a class of EVT models that incorporate HF information. The intuition underlying the proposed approach is that HF-based measures, such as realized variation, are likely to contain useful information regarding the extreme behaviour. If we define extreme occurrences as observations exceeding a high quantile of the loss distribution, then this high quantile is more likely to be exceeded when the realized variance is high. Similarly, the magnitude of such events may also be related to the realized measures. This is well illustrated in Figure 51 where losses for the S&P 500 are shown for the 2000-2014 period. Red points indicate the days on which the *98*th quantile of the loss distribution is exceeded. They tend to be concentrated in periods of high realized variance (RV) and high jump activity (J). These two quantities will be formally defined in Section 5.3.

From a methodological perspective, we use the Peaks-over-threshold (POT) method of Davison and Smith (1990) and propose a *realized* POT approach (RPOT). Borrowing techniques from the non-stationary EVT literature, we consider different models for the time-varying probability of exceeding a high threshold, exploring Generalized Linear Models and non-homogeneous Poisson processes with different realized measures as covariates. The size of the exceedances are modelled using a GP distribution with time-varying parameters that are functions of the realized measure.

The remainder of the chapter is organized as follows. Section 5.2 reviews the standard and the conditional POT approaches. Section 5.3 describes the realized measures and how they are employed in the RPOT approach. In-sample results showing that HF data convey information on the behaviour of the extremes beyond that provided in low frequency data appear in Section 5.4. Section 5.5 shows that out-of-sample forecasts of standard risk measures obtained with the RPOT approach outperform those from standard models. Section 5.6 contains some robustness checks to consolidate the evidence from the main analysis. Section 5.7 concludes. A supplemental appendix contains additional details and

results.

**Figure 51: S&P 500**. Daily negative returns (*losses*), Realized Variance (RV) and Jump component (J) from the beginning of 2000 to the end of 2014. Red dots indicate days for which a threshold set at the 98th quantile of the loss distribution is exceeded.




## 5.2 Extreme Value Theory

### 5.2.1 The Peaks Over Threshold approach

Let $\{Y_t\}_{t=1}^T$ be a sequence of *iid* random variables from a distribution function $F$ with upper end point $v_F := \sup\{Y_t : F(Y_t) < 1\}$. Define the extremes of $\{Y_t\}_{t=1}^T$ to be the exceedances of a high threshold $u$, $u < v_F$. As $u \to v_F$, Pickands (1975) shows that the distribution of the excesses $(Y_t - u)_+$ converges to a GP distribution G with shape parameter $\xi$ and

scale parameter $\nu > 0$. That is, $\Pr(Y - u \leq y | Y > u)$ goes to

$$G(y; \xi, \nu) = \begin{cases} 1 - \{1 + \xi y/\nu\}^{-\frac{1}{\xi}} & \text{for} \quad \xi \neq 0 \\ 1 - \exp\{-y/\nu\} & \text{for} \quad \xi = 0 \end{cases} \tag{5.1}$$

When $\xi > 0$, $F$ has Pareto-type upper tail with tail index $1/\xi$. For a given threshold $u$, the POT approach is based on the decomposition of the tail of F as

$$1 - F(y) = (1 - F(u))(1 - F_u(y)) \tag{5.2}$$

where $\phi \equiv (1 - F(u)) = \Pr(Y > u)$ and $F_u(y) = \Pr(Y - u \leq y | Y > u)$. Letting $\mathcal{E} = \{t \in \{1, \ldots, T\} | Y_t > u\}$ be the set containing the times at which an exceedance occurs, the number of exceedances $N_u(T) = card\{\mathcal{E}\}$ can be modelled as a homogeneous Poisson process with rate $\lambda$ and the size of excesses $\{W_j = Y_j - u | j \in \mathcal{E}\}$ with the limiting GP distribution (Coles, 2001). An estimate of the tail probability in Equation (5.2) can thus be obtained as

$$\widehat{\bar{F}}(y) = \widehat{\phi} \left(1 + \widehat{\xi} \frac{y - \widehat{u}}{\widehat{\nu}}\right)^{-\frac{1}{\widehat{\xi}}} \tag{5.3}$$

where $\bar{F} = 1 - F$, $\widehat{u}$ is an appropriately chosen threshold, $\widehat{\xi}$ and $\widehat{\nu}$ are estimates of the GP parameters, and $\widehat{\phi} = 1 - \exp(-\widehat{\lambda})$ is an estimate of the exceedance probability obtained from the Poisson process with estimated intensity $\widehat{\lambda}$. In particular, letting $g$ be the GP density, the joint likelihood function for the POT approach is

$$\begin{aligned} \mathcal{L}(\lambda, \xi, \nu) &= \Pr(N_u = n) \prod_{j=1}^{n} g(y_j - u) \\ &= \tfrac{\lambda^n}{n!} \exp^{-\lambda} \prod_{j=1}^{n} \tfrac{1}{\nu} \left[1 + \tfrac{\xi(y_j - u)}{\nu}\right]^{-(1+1/\xi)} \\ &= \mathcal{L}(\lambda)\mathcal{L}(\xi, \nu) \end{aligned} \tag{5.4}$$

where $\mathcal{L}(\lambda)$ and $\mathcal{L}(\xi, \nu)$ are the Poisson and GP likelihoods, respectively[1]. Separate maximization leads to the Maximum Likelihood (ML) estimators $\widehat{\lambda}$, $\widehat{\xi}$, and $\widehat{\nu}$. We have the closed-form estimator $\widehat{\lambda} = n/T$ and so $\widehat{\phi} = 1 - \exp(-n/T)$. Contrarily, $\widehat{\xi}$ and $\widehat{\nu}$ are found by numerical optimization (Hosking and Wallis, 1987) and their asymptotic normality is established for $\xi > -0.5$ (Smith, 1985).

### 5.2.2 The *Conditional* Peaks Over Threshold approach

Daily financial returns violate the *iid* assumption underlying the POT approach and this hinders the applicability to the tails of the unconditional return distribution. Suppose that the process of interest $\{Y_t\}_{t=1}^{T}$ is stationary with marginal distribution $F$ with upper end point $v_F := \sup\{Y_t : F(Y_t) < 1\}$ and satisfies the mixing condition,

$$\begin{aligned} &|\Pr(Y_{t_1} \leq u_n, \ldots, Y_{t_p} \leq u_n, Y_{s_1} \leq u_n, \ldots, Y_{s_q} \leq u_n) \\ &- \Pr(Y_{t_1} \leq u_n, \ldots, Y_{t_p} \leq u_n) \Pr(Y_{s_1} \leq u_n, \ldots, Y_{s_q} \leq u_n)| \leq \alpha(n, l) \end{aligned} \tag{5.5}$$

for $\alpha(n, l_n) \to 0$ for some sequence $l_n = o(n)$ and $u_n \to v_F$ as $n \to \infty$, with $t_1 < \cdots < t_p < s_1 < \cdots < s_q$ and $s_1 - t_p > l$. Under this assumption on the dependence structure[2], the sequence of exceedances is no longer independent, invalidating the homogeneous Poisson model for counts, but the convergence result to the GP distribution still holds and $\widehat{\xi}$ and $\widehat{\nu}$ are still asymptotically normal (Drees, 2000). This argument keeps EVT appealing to estimate the tails of the conditional return distribution. In financial applications, the POT approach is typically applied to the time series of residuals from a filtering model. In particular, McNeil and Frey

[1] An alternative characterization of the likelihood function can be obtained with the binomial approximation of the Poisson process. In this case, $\phi$ enters the likelihood equation

$$\mathcal{L}(\nu, \xi, \phi) = \prod_{t=1}^{T} (1-\phi)^{1-I_t} \left( \frac{\phi}{\nu} \left[ 1 + \frac{\xi(y_t - u)}{\nu} \right]_{+}^{-1/\xi - 1} \right)^{I_t}$$

where $I_t$ takes value 1 if $t \in \mathcal{E}$ and zero otherwise. This representation will be useful in Section 3.1 where we model $\phi$ directly.

[2] The Condition in Equation (5.5) is referred to as the $D(u_n)$ condition and it is satisfied for large classes of Gaussian stationary sequences, including many Gaussian linear processes, e.g. Gaussian ARMA and fractional ARIMA processes.

(2000) suggest a two-step procedure that combines a GARCH model to account for the volatility and a POT approach to estimate the tails of the standardized residuals from the GARCH model. This approach works reasonably well, however if dependence remains in the extremes of the residuals, then tail probabilities are poorly estimated. Harvey and Siddique (1999) provide evidence of time variation in higher moments of the conditional return distribution, implying that GARCH models are not able to fully capture dependence in the returns. Furthermore, difficulties may also arise from the possible presence of non-stationarities. In particular, Mikosch and Starica (2004) explain how regime switches could be an explanation for the long-range dependence patterns that characterize financial returns.

An alternative way to estimate the tails of the conditional distribution consists in modelling them directly, borrowing techniques from the non-stationary EVT literature. Suppose that $\{Y_t\}_{t=1}^{T}$ is a random process which can be either stationary or non-stationary, and satisfy or not the condition in Equation (5.5). Following Davison and Smith (1990), the tail of the conditional distribution of $Y_t$ can be decomposed following

$$\Pr(Y_t > y|\mathcal{F}_{t-1}) = \Pr(Y_t > u|\mathcal{F}_{t-1}) \Pr(Y_t - u > y|Y_t > u, \mathcal{F}_{t-1}) \quad (5.6)$$

where $\mathcal{F}_{t-1}$ is the information set of the process up to time $t-1$. An estimate of the conditional tail probability at time $t$ can be obtained combining a dynamic model for $\phi_t = \Pr(Y_t > u|\mathcal{F}_{t-1})$ such as a Generalized Linear Model (Davison and Smith, 1990) or a non-homogeneous Poisson process for the counts $N_u(t)$ (Chavez-Demoulin et al., 2005, 2014) and a GP distribution with parameters depending on covariates for $\Pr(Y_t - u > y|Y_t > u, \mathcal{F}_{t-1})$, see Coles (2001) and references therein.

With a slight misuse of notation, the joint likelihood of the conditional tail can be written as[3],

$$\mathcal{L}(\phi, \xi, \nu) = f_{\phi_1,\xi_1,\nu_1}(\phi_1, \xi_1, \nu_1) \prod_{t=2}^{T} f_{\phi_t,\xi_t,\nu_t|\mathcal{F}_{t-1}}(\phi_t, \xi_t, \nu_t|\mathcal{F}_{t-1}) \quad (5.7)$$

[3]We write the conditional likelihood as a function of the parameter of interest $\phi$, instead of $\lambda$ as in (5.4).

where $f_{\phi_t,\xi_t,\nu_t|\mathcal{F}_{t-1}}(\phi_t, \xi_t, \nu_t|\mathcal{F}_{t-1})$ is the joint density of the model at time $t$ conditioned on time $t-1$. Assuming conditional independence between the rate and the magnitude of the exceedances, one obtains

$$f_{\phi_t,\xi_t,\nu_t|\mathcal{F}_{t-1}}(\phi_t, \xi_t, \nu_t|\mathcal{F}_{t-1}) = f_{\phi_t|\mathcal{F}_{t-1}}(\phi_t|\mathcal{F}_{t-1}) f_{\xi_t,\nu_t|\mathcal{F}_{t-1}}(\xi_t, \nu_t|\mathcal{F}_{t-1}) \tag{5.8}$$

where $f_{\phi_t|\mathcal{F}_{t-1}}(\phi_t|\mathcal{F}_{t-1})$ and $f_{\xi_t,\nu_t|\mathcal{F}_{t-1}}(\xi_t, \nu_t|\mathcal{F}_{t-1})$ are respectively the density of the models for the exceedance rate and the density of the GP distribution. The joint likelihood can thus be separated into two components that can be maximized separately,

$$\mathcal{L}(\phi, \xi, \nu) = \left\{ f_{\phi_1}(\phi_1) \prod_{t=2}^{T} f_{\phi_t|\mathcal{F}_{t-1}}(\phi_t|\mathcal{F}_{t-1}) \right\} \left\{ f_{\xi_1,\nu_1}(\xi_1, \nu_1) \prod_{t=2}^{T} f_{\xi_t,\nu_t|\mathcal{F}_{t-1}}(\xi_t, \nu_t|\mathcal{F}_{t-1}) \right\} \tag{5.9}$$

leading to the sequence of ML estimators $\widehat{\phi}$, $\widehat{\xi}$, and $\widehat{\nu}$.

Limiting arguments concerning the validity of this approach and the asymptotic properties of the ML estimators strongly depend on the assumptions on the $Y_t$'s. Assuming $\{Y_t\}$ to be stationary, Hall and Tajvidi (2000) and Beirlant and Goegebeur (2004) establish the asymptotic normality for several semi-parametric classes of estimators $\widehat{\xi}$ and $\widehat{\nu}$ both when the true conditional distribution is GP and when the conditional distribution of excesses converges to the GP. Similar arguments can be used to prove asymptotic properties of fully parametric estimators, provided that the functions used to describe the dynamic parameters satisfy standard conditions on differentiability (Wang and Tsai, 2009). When the series $\{Y_t\}$ is non-stationary, the theory of Hüsler (1986) proves the existence of an extreme value limit for the sample maxima of $Y_t$, and we use this result to support the time-varying POT approach. Asymptotic results for the statistical procedures strongly depend on the knowledge of the form of non-stationarity. We simply assume that the specification of the parameter accounts for the possible non-stationarities, and use the standard asymptotic results.

In what follows, we consider the conditional decomposition in Equation (5.6), and show how to incorporate HF data into the models for the exceedance rate and the GP. In both cases, we will rely on parametric specifications of the dynamic parameters and will base the inference on

standard asymptotic arguments from ML theory.

## 5.3 The *Realized* Peaks Over Threshold approach

In order to model the tail of the conditional return distribution, Chavez-Demoulin et al. (2005) and Chavez-Demoulin et al. (2014) develop respectively fully parametric and non-parametric extensions of the POT approach, considering $\mathcal{F}_t$ to be the information set generated by the daily price path. We propose to augment the available information set with HF data and model daily extremes with measures built upon these observations. We denote this new information set by $\mathcal{H}_t$ where $\mathcal{F}_t \subset \mathcal{H}_t$.

Realized measures are non-parametric estimators of the variation of the price path of an asset. They ignore the variation of prices overnight and sometimes the variation in the first few minutes of the trading day when recorded prices may contain large errors. A good background for realized measures can be found in survey articles by Barndorff-Nielsen and Shephard (2007) and Andersen et al. (2009).

Let $r_{t,\Delta} = p_t - p_{t-\Delta}$ be the discretely sampled $\Delta$-period return on day $t$. The most common realized measure is the realized variance (RV),

$$RV_t = \sum_{j=1}^{N} r^2_{t-1+j\cdot\Delta,\Delta} \tag{5.10}$$

where $N = 1/\Delta$. Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002) show that, if prices are observed without noise then, as $\Delta \to 0$, this measure consistently estimates the quadratic variation of the price process $p_t$ on day $t$. When the price process includes a jump component, the quadratic variation contains the contributions of both continuous and discontinuous sample paths. As these two components may contain different sources of information, Barndorff-Nielsen and Shephard (2004) define the bipower variation (BV)

$$BV_t = \frac{\pi}{2} \sum_{j=2}^{N} |r_{t-1+j\cdot\Delta,\Delta}||r_{t-1+(j-1)\cdot\Delta,\Delta}| \tag{5.11}$$

which consistently estimates the variation due to the continuous sample path. The jump variation (J) can then be estimated by taking the difference between RV and BV,

$$J_t = \max(0, RV_t - BV_t). \tag{5.12}$$

As our approach includes realized measures as covariates in the models for the exceedances, throughout we will refer to it as *Realized Peaks over Threshold* (RPOT). In what follows, we discuss the two components of the conditional likelihood in Equation (5.9) individually and show how realized measures can be incorporated in each. First, we present models for the exceedance rate, then we treat the modelling of excess size.

### 5.3.1 Modelling the exceedance rate

We start following Davison and Smith (1990) who propose combining the approach for stationary data with regression modelling so that the parameters are modelled as linear functions of covariates. Treating the occurrence of an exceedance as a dichotomous variable, the probability of an event can thus be modelled with a Logit function,

$$\phi_t = \frac{1}{1 + \exp(\varphi_0 + \varphi_1 RM_{t-1})} \tag{5.13}$$

where $RM_{t-1}$ is a generic realized measure. More generally, several lags of realized measures can be included in a length $p+1$ row-vector $\mathbf{RM}_t'$ of regressors at time $t$. Parameters $\boldsymbol{\varphi} = [\varphi_0, \varphi_1, \ldots, \varphi_p]$ are easily estimated through maximization of the likelihood function,

$$\mathcal{L}(\boldsymbol{\varphi}) = \prod_{t=(l+1)}^{T} \left(\exp\left(\mathbf{RM}_t'\boldsymbol{\varphi}\right)\right)^{I_t} \frac{1}{1 + \exp\left(\mathbf{RM}_t'\boldsymbol{\varphi}\right)}$$

where $l$ is the lag at which a given realized measure becomes available[4], and $I_t$ takes value one if $t \in \mathcal{E} = \{t \in \{1, \ldots, T\} | Y_t > u\}$ and zero oth-

[4]In the simplest case, $\mathbf{RM}_t = [1, RM_{t-1}]$ as in Equation (5.13). When the regressors are functions of past observations, it may require several lags before all are available. For example, for a sequence of observations $y_1, \ldots, y_T$, if we use a 30-day moving average of the $y_t$'s as a regressor, then the first observation will be available at $l = 30$.

erwise. ML estimates of the parameters are obtained through numerical techniques.

As an alternative approach to obtain an estimate of the time-varying exceedance rate, Chavez-Demoulin et al. (2005) propose to model the number of exceedances $N_u(t)$ with a non-homogeneous Poisson process. In this case $\phi = 1 - \exp\left(-\int_0^t \lambda_u du\right)$ where $\lambda_t$ is the intensity parameter of the Poisson process at time $t$. They consider a self-exciting process depending on past exceedances to determine the intensity parameter $\lambda_t$. We rather allow $\lambda_t$ to change as a function of the realized measure. In particular, considering a log-linear specification as in Coles (2001) to guarantee the positivity of $\lambda_t$, we let

$$\lambda_t(\zeta_0, \zeta_1) = \exp\left(\zeta_0 + \zeta_1 RM_{t-1}\right) \tag{5.14}$$

where $RM_{t-1}$ is a generic realized measure. More generally, the likelihood function for the Poisson process is expressed as

$$\mathcal{L}(\boldsymbol{\zeta}) = \exp\left(-\int_{T-l} \exp\left(\mathbf{R}\mathbf{M}_t'\boldsymbol{\zeta}\right) dt\right) \prod_{j \in \mathcal{E}} \exp\left(\mathbf{R}\mathbf{M}_j'\boldsymbol{\zeta}\right)$$

and ML estimates of the parameters $\boldsymbol{\zeta} = [\zeta_0, \zeta_1, \ldots, \zeta_p]$ are obtained through numerical techniques.

### 5.3.2 Modelling the excess size

The natural model for the excesses of a high threshold is the GP distribution. In the case of a non-stationary process, Davison and Smith (1990) suggest adding linear covariates in the scale $\nu_t$ and shape $\xi_t$ parameters. In our specific case, $\nu_t$ varies according to a log-link function of the realized measures, e.g. with one realized measure

$$\nu_t(\kappa_0, \kappa_1) = \exp(\kappa_0 + \kappa_1 RM_{t-1}). \tag{5.15}$$

The estimation of a time-varying shape parameter $\xi_t$ introduces a lot of uncertainty. To gain stability, we keep it constant, $\xi_t = \xi$, similarly to Chavez-Demoulin et al. (2014). More generally, the likelihood function

for the size of the excesses assumes the form

$$\mathcal{L}(\boldsymbol{\kappa}, \xi) = \prod_{t=(l+1)}^{T} \left( \frac{1}{\exp\left(\mathbf{RM}_t'\boldsymbol{\kappa}\right)} \left[ 1 + \frac{\xi(y_t - u)}{\exp\left(\mathbf{RM}_t'\boldsymbol{\kappa}\right)} \right]_+^{-1/\xi - 1} \right)^{I_t} \tag{5.16}$$

and $\boldsymbol{\kappa} = [\kappa_0, \kappa_1, \ldots, \kappa_p]$ and $\xi$ are the parameters that must be estimated, $I_t$ takes value one when an exceedance occurs, i.e. $t \in \mathcal{E}$, and zero otherwise. ML estimates of the parameters are obtained through numerical techniques.

### 5.3.3 Estimation of the conditional risk measures

Quantile-based risk measures such as the Value-at-Risk (VaR) and the Expected Shortfall (ES) are central tools for quantitative risk management in the financial industry. Denote by $L_t(x) = -F_t(y)$ the *loss distribution* at time $t$. The one-day-ahead VaR at level $\alpha$ is defined[5] as the $(1-\alpha)$-quantile of the loss distribution,

$$VaR_{t|t-1}^{\alpha} = \inf\{x \in \mathbb{R} : L_{t|t-1}(x) \geq 1 - \alpha\}$$

where $L_{t|t-1}(x)$ denotes the loss distribution conditional on the information at time $t-1$. To obtain a VaR forecast with the RPOT approach, it is simply necessary to invert Equation (5.3) and plug-in the forecast of the parameters obtained for the loss distribution, i.e.

$$\widehat{VaR}_{t|t-1}^{\alpha} = \widehat{u} + \frac{\widehat{\nu}_{t|t-1}}{\widehat{\xi}} \left[ \left( \frac{\widehat{\phi}_{t|t-1}}{1-\alpha} \right)^{\widehat{\xi}} - 1 \right] \tag{5.17}$$

where $\widehat{\nu}_{t|t-1}$ is the parameter forecast from the dynamic GP distribution and $\widehat{\phi}_{t|t-1}$ is the forecast from the threshold exceedance model. In particular, when we model the probability of an exceedance by a Logit function, then $\widehat{\phi}_{t|t-1} = 1/(1 + \exp(\widehat{\varphi_0} + \widehat{\varphi_1} RM_{t-1}))$ for the model in (5.13) while for the non-homogeneous Poisson process we have $\widehat{\phi}_{t|t-1} = 1 - \exp(-\exp(\widehat{\zeta_0} + \widehat{\zeta_1} RM_{t-1}))$ for the model in (5.14).

[5] In Section 5, we compute the VaR at level $\alpha = 0.01$. The latter is often called the 99%-VaR in the literature.

The one-day-ahead conditional ES at level $\alpha$ is defined as

$$\widehat{ES}^{\alpha}_{t|t-1} = \frac{\widehat{VaR}^{\alpha}_{t|t-1}}{1-\widehat{\xi}} + \frac{\widehat{\nu}_{t|t-1} - \widehat{\xi}\widehat{u}}{1-\widehat{\xi}}. \tag{5.18}$$

$\widehat{VaR}^{\alpha}_{t|t-1}$ and $\widehat{ES}^{\alpha}_{t|t-1}$ are point forecasts. To obtain the confidence interval estimates, we use a post-blackened bootstrap (Davison and Hinkley, 1997). First, we fit a RPOT model to the observations and obtain the residuals $R_j = 1/\widehat{\xi}\log\left\{1+\widehat{\xi}\left((Y_j-\widehat{u})/\widehat{\nu}_j\right)\right\}$ for each exceedance $Y_j$, $j \in \mathcal{E}$. Then, we obtain $B$ bootstrap samples of the residuals $R_j$ and apply the RPOT model to each sample to obtain a percentile based confidence interval of the VaR and ES.

## 5.4 Empirical analysis

### 5.4.1 Data description

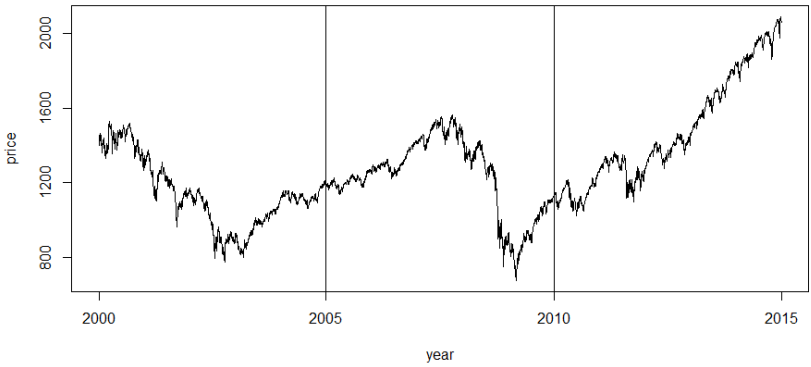
The empirical analysis is based on the S&P500 index from January 1, 2000 to December 31, 2014. We obtain the time series of returns, 5-min realized variance and 5-min bipower variation from the Oxford-Man Institute "Realised Library" version 0.2 (Heber et al., 2009).

The sequence of extremes is obtained by fixing a threshold level $u$. The choice of the threshold is important, implying a balance between bias and variance. On the one hand, smaller $u$ means more observations used for inference. On the other hand, probability theory suggests choosing a high $u$ for limiting results to apply. In this paper, we consider a threshold level corresponding to the *90*th quantile of the loss distribution. Graphical analysis and tests assessing the merits of this choice are performed.

In what follows, we consider three five-year sub-samples: 2000-2004, 2005-2009, 2010-2014. This allows us to discuss the fitted models in three different regimes of the stock market. As can be seen in Figure 52, the first sub-sample contains the downward trend consequent to the dot-com bubble explosion, and part of the subsequent rebound. The second sub-sample contains the 2008 crash following the default of several banks

and insurance companies. The last sub-sample contains the recovery and the bullish trend of the most recent years.

**Figure 52: S&P500 index data.**




### 5.4.2 Model estimation

This section is devoted to the estimation of different specifications of the models presented in Section 5.3. For each model, we consider four different specifications with increasing sets of covariates. The set of covariates at time $t$ are obtained as measures of the information set $\mathcal{H}_t$ and are as follows:

I. $\{\log(RV_{t-1})\}$

II. $\{\log(BV_{t-1}), \log(1+J_{t-1})\}$

III. $\{\log(RV_{t-1}), \log(\overline{RV}^{W}_{t-1}), \log(\overline{RV}^{M}_{t-1})\}$

IV. $\{\log(BV_{t-1}), \log(1+J_{t-1}), \log(\overline{RV}^{W}_{t-1}), \log(\overline{RV}^{M}_{t-1})\}$

where $\overline{RV}^{W}_{t-1} = \frac{1}{4}\sum_{i=2}^{5} RV_{t-i}$ and $\overline{RV}^{M}_{t-1} = \frac{1}{17}\sum_{i=6}^{22} RV_{t-i}$. The choice of these sets requires a bit of discussion. Set I contains the simplest re-

alized measure: the realized variance. In the introduction, we argue that the intuition behind the RPOT approach is a possible relationship between excesses and the realized variance. Using the lagged realized variance as a covariate is thus a natural specification. Set II contains the bipower variation and the jump component. This choice follows the argument of Andersen et al. (2007) which suggests that distinguishing between the information coming from the continuous and discontinuous sample paths could be valuable. Sets III and IV can be considered extensions where we add the information coming from the average weekly and monthly realized variance. This gives a heterogeneous autoregressive (HAR) structure (Corsi, 2009) that allows us to see whether realized measures at different time-horizons are useful to predict the behaviour of the exceedances. We consider the logarithmic transformation of these realized measures as it is preferable from a modelling perspective[6].

Table 53 reports the estimated parameters for the Logit model in Equation (5.13) and more elaborate specifications. Results for specification I show that the coefficient on RV is strongly significant across the three windows. This strong statistical evidence confirms the usefulness of HF data for modelling the exceedance rate. Decomposing the RV in the contribution from the continuous and the discontinuous sample paths in specification II allows us to see that the jump component has a negligible role on the exceedance rate. In particular, while the coefficient on the BV is strongly significant across the windows, the jump coefficient is never significant. Finally, information on the variance at further lags in a HAR fashion does not seem to add significantly to the prediction of the exceedance rate. Indeed, in both specifications III and IV, the coefficients on $\overline{RV}^M$ are never significant and those on $\overline{RV}^W$ are significant on only one occasion.

In Table 54, we report the results obtained with the non-homogeneous Poisson process for the intensity function in Equation (5.14) and more elaborate specifications. They substantially confirm what we found with

[6]In Section 5.4, we scale the jump component in order to have covariates of similar magnitude. Indeed, while $\log(RV_t)$ is much greater than $RV_t$, $log(1 + J_t) \approx J_t$, with the consequence that the parameter associated to the latter would be huge.

the Logit model: the coefficient on RV and BV are strongly significant while coefficients on the jump component and longer past averaged realized volatilities are not significant.

Table 55 shows the estimated parameters of the dynamic GP distribution with constant $\xi$ and $\nu_t$ parameter as in Equation (5.15) and more elaborate specifications. Also in this case, the inclusion of HF data contributes significantly to explain the size of the excesses. The coefficients on RV and BV are strongly significant even when realized variances at longer lags are added. As with the exceedance rate, the coefficients on $\overline{RV}^W$ and $\overline{RV}^M$ are generally not significant, suggesting that the inclusion of additional lags of realized variation does not much improve the fit of the model. The coefficient on the jump component $J$ is significant only in one case.

As for the $\xi$ parameter in Table 55, it is very close to zero or not significantly different from zero in the sample periods 2000-2004 and 2005-2009, while it is negative and strongly significant in the last five-year sub-sample considered. These results imply that the right tail of the *loss distribution* has either an exponential decay or even lighter during the most recent years. As we are trying to estimate the tail of the conditional return distribution, we do not expect the $\xi$ parameter to be strongly positive as it is usually observed for the unconditional return distribution. Furthermore, the results of Table 55 are consistent with the regimes of the different sub-samples. Indeed, from Figure 52 we can clearly see that while the first two sub-periods are characterized by extreme negative returns, the last sub-sample shows mainly an upward trend.

Overall, it appears that HF data apport a meaningful contribution to understanding the behaviour of the excesses. While the one-day lagged realized variation is useful for modelling both the exceedance rate and the size of excesses, the contributions of jumps and past volatilities are limited.

**Table 53: Fitted Logit models** for specifications:

I. $\text{logit}(\phi_t) = \varphi_0 + \varphi_1 \log(RV_{t-1})$

II. $\text{logit}(\phi_t) = \varphi_0 + \varphi_2 \log(BV_{t-1}) + \varphi_3 \log(1 + J_{t-1})$

III. $\text{logit}(\phi_t) = \varphi_0 + \varphi_1 \log(RV_{t-1}) + \varphi_4 \log(\overline{RV}^W_{t-1}) + \varphi_5 \log(\overline{RV}^M_{t-1})$

IV. $\text{logit}(\phi_t) = \varphi_0 + \varphi_2 \log(BV_{t-1}) + \varphi_3 \log(1 + J_{t-1}) + \varphi_4 \log(\overline{RV}^W_{t-1}) + \varphi_5 \log(\overline{RV}^M_{t-1})$

on S&P500 returns. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\varphi_0$ | $\varphi_1$ | $\varphi_2$ | $\varphi_3$ | $\varphi_4$ | $\varphi_5$ |
|---|---|---|---|---|---|---|
| | | | 2000-2004 | | | |
| I. | 5.46*** | 0.84*** | | | | |
| | (1.02) | (0.11) | | | | |
| II. | 6.78*** | | 0.97*** | -0.07 | | |
| | (1.12) | | (0.12) | (0.08) | | |
| III. | 8.15*** | 0.38* | | | 0.52* | 0.25 |
| | (1.34) | (0.18) | | | (0.24) | (0.23) |
| IV. | 9.07*** | | 0.62*** | -0.10 | 0.37 | 0.24 |
| | (1.44) | | (0.18) | (0.09) | (0.25) | (0.23) |
| | | | 2005-2009 | | | |
| I. | 5.02*** | 0.79*** | | | | |
| | (0.68) | (0.08) | | | | |
| II. | 5.13*** | | 0.78*** | 0.00 | | |
| | (0.78) | | (0.08) | (0.07) | | |
| III. | 5.85*** | 0.30 | | | 0.30 | 0.29 |
| | (0.74) | (0.17) | | | (0.22) | (0.17) |
| IV. | 5.86*** | | 0.27 | 0.00 | 0.31 | 0.29 |
| | (0.81) | | (0.18) | (0.07) | (0.22) | (0.17) |
| | | | 2010-2014 | | | |
| I. | 3.77*** | 0.61*** | | | | |
| | (0.87) | (0.09) | | | | |
| II. | 4.69*** | | 0.68*** | -0.03 | | |
| | (1.05) | | (0.11) | (0.08) | | |
| III. | 4.80*** | 0.45*** | | | 0.11 | 0.16 |
| | (1.10) | (0.14) | | | (0.18) | (0.17) |
| IV. | 5.46*** | | 0.54*** | -0.03 | 0.05 | 0.18 |
| | (1.24) | | (0.15) | (0.08) | (0.19) | (0.17) |

**Table 54: Fitted non-homogeneous Poisson models** for specifications:

I. $\lambda_t = \exp\left(\zeta_0 + \zeta_1 \log(RV_{t-1})\right)$

II. $\lambda_t = \exp\left(\zeta_0 + \zeta_2 \log(BV_{t-1}) + \zeta_3 \log(1 + J_{t-1})\right)$

III. $\lambda_t = \exp\left(\zeta_0 + \zeta_1 \log(RV_{t-1}) + \zeta_4 \log(\overline{RV}^W_{t-1}) + \zeta_5 \log(\overline{RV}^M_{t-1})\right)$

IV. $\lambda_t = \exp\left(\zeta_0 + \zeta_2 \log(BV_{t-1}) + \zeta_3 \log(1 + J_{t-1}) + \zeta_4 \log(\overline{RV}^W_{t-1}) + \zeta_5 \log(\overline{RV}^M_{t-1})\right)$

on S&P500 returns. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\zeta_0$ | $\zeta_1$ | $\zeta_2$ | $\zeta_3$ | $\zeta_4$ | $\zeta_5$ |
|---|---|---|---|---|---|---|
| | | | 2000-2004 | | | |
| I. | 4.15*** | 0.71*** | | | | |
| | (0.87) | (0.10) | | | | |
| II. | 5.01*** | | 0.79*** | -0.05 | | |
| | (0.90) | | (0.10) | (0.07) | | |
| III. | 6.53*** | 0.31* | | | 0.43 | 0.24 |
| | (1.18) | (0.16) | | | (0.22) | (0.21) |
| IV. | 7.16*** | | 0.50** | -0.07 | 0.30 | 0.24 |
| | (1.24) | | (0.16) | (0.08) | (0.22) | (0.21) |
| | | | 2005-2009 | | | |
| I. | 3.34*** | 0.62*** | | | | |
| | (0.52) | (0.06) | | | | |
| II. | 3.49*** | | 0.62*** | -0.02 | | |
| | (0.59) | | (0.07) | (0.05) | | |
| III. | 4.02*** | 0.23 | | | 0.23 | 0.24 |
| | (0.57) | (0.15) | | | (0.19) | (0.15) |
| IV. | 4.07*** | | 0.21 | -0.01 | 0.25 | 0.24 |
| | (0.62) | | (0.15) | (0.05) | (0.19) | (0.15) |
| | | | 2010-2014 | | | |
| I. | 2.74*** | 0.52*** | | | | |
| | (0.75) | (0.08) | | | | |
| II. | 3.48*** | | 0.58*** | -0.03 | | |
| | (0.90) | | (0.09) | (0.07) | | |
| III. | 3.65*** | 0.39*** | | | 0.09 | 0.14 |
| | (0.97) | (0.12) | | | (0.16) | (0.15) |
| IV. | 4.23*** | | 0.46*** | -0.03 | 0.04 | 0.16 |
| | (1.09) | | (0.14) | (0.07) | (0.17) | (0.15) |

**Table 55: Fitted dynamic Generalized Pareto models**. The $\xi$ parameter is constant, while $\nu_t$ is allowed to vary according to

I. $\nu_t = \exp\left(\kappa_0 + \kappa_1 \log(RV_{t-1})\right)$

II. $\nu_t = \exp\left(\kappa_0 + \kappa_2 \log(BV_{t-1}) + \kappa_3 \log(1 + J_{t-1})\right)$

III. $\nu_t = \exp\left(\kappa_0 + \kappa_1 \log(RV_{t-1}) + \kappa_4 \log(\overline{RV}^W_{t-1}) + \kappa_5 \log(\overline{RV}^M_{t-1})\right)$

IV. $\nu_t = \exp\left(\kappa_0 + \kappa_2 \log(BV_{t-1}) + \kappa_3 \log(1 + J_{t-1}) + \kappa_4 \log(\overline{RV}^W_{t-1}) + \kappa_5 \log(\overline{RV}^M_{t-1})\right)$

*,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\kappa_0$ | $\kappa_1$ | $\kappa_2$ | $\kappa_3$ | $\kappa_4$ | $\kappa_5$ | $\xi$ |
|---|---|---|---|---|---|---|---|
| | | | | 2000-2004 | | | |
| I. | -2.27* | 0.31** | | | | | 0.02 |
| | (0.98) | (0.11) | | | | | (0.09) |
| II. | -1.84 | | 0.35** | -0.04 | | | 0.01 |
| | (1.06) | | (0.12) | (0.08) | | | (0.09) |
| III. | -0.37 | 0.05 | | | 0.33 | 0.15 | 0.00 |
| | (1.46) | (0.15) | | | (0.20) | (0.23) | (0.09) |
| IV. | 0.82 | | 0.09 | -0.06 | 0.38 | 0.19 | 0.00 |
| | (1.54) | | (0.15) | (0.08) | (0.20) | (0.23) | (0.09) |
| | | | | 2005-2009 | | | |
| I. | -0.95 | 0.42*** | | | | | 0.00 |
| | (0.61) | (0.07) | | | | | (0.07) |
| II. | -0.83 | | 0.42*** | -0.01 | | | -0.02 |
| | (0.67) | | (0.08) | (0.05) | | | (0.07) |
| III. | -0.51 | 0.06 | | | 0.48** | -0.07 | -0.10 |
| | (0.56) | (0.12) | | | (0.18) | (0.14) | (0.07) |
| IV. | 0.03 | | 0.11 | 0.22 | 0.19 | 0.07 | -0.86 |
| | (0.46) | | (0.20) | (1.58) | (0.36) | (0.51) | (0.73) |
| | | | | 2010-2014 | | | |
| I. | -0.19 | 0.47*** | | | | | -0.29*** |
| | (0.69) | (0.07) | | | | | (0.08) |
| II. | 0.86 | | 0.56** | -0.12* | | | -0.42*** |
| | (0.59) | | (0.06) | (0.06) | | | (0.08) |
| III. | 0.37 | 0.30*** | | | 0.32** | -0.10 | -0.42*** |
| | (0.67) | (0.09) | | | (0.11) | (0.10) | (0.09) |
| IV. | 0.91 | | 0.43** | -0.09 | 0.17 | -0.04 | -0.50*** |
| | (0.51) | | (0.08) | (0.05) | (0.12) | (0.09) | (0.10) |

### 5.4.3 Model diagnostics

Use of appropriate model diagnostics is crucial to determine whether the assumed model describes the behaviour of the underlying observations. In this section, we make use of graphical methods and formal testing procedures to assess the validity of the models fitted in Section 5.4.2.

First, we perform a deviance $\chi^2$ test on the residuals of the Logit model, a standard goodness-of-fit test in Logit regression literature (Hosmer and Lemeshow, 2004). Given a model $\mathcal{M}_0$ and a sequence of observations $y$, the deviance is defined as

$$D(y, \mathcal{M}_0) = -2\left(\log \Pr(y|\widehat{\theta}_0) - \log \Pr(y|\widehat{\theta}_s)\right)$$

where $\widehat{\theta}_0$ denotes the fitted values of the parameters in the model $\mathcal{M}_0$ and $\widehat{\theta}_s$ denotes the fitted parameters for the saturated model, i.e. a model with a parameter for every observation. Two measures of deviance are particularly important in a Logit model: the *null deviance* which represents the deviance for a model with only the intercept, and the *model deviance* representing the deviance of the fitted model. To assess the contribution of the predictors, one can subtract the model deviance from the null deviance, i.e.

$$D_{null} - D_{fitted} = -2\left(\log \Pr(y|\widehat{\theta}_n) - \log \Pr(y|\widehat{\theta}_0)\right)$$

and assess the difference on a $\chi^2$ distribution with degrees of freedom equal to the difference in the number of parameters estimated. If the model deviance is significantly smaller than the null deviance then one can conclude that the predictors significantly improve the model fit.

Another commonly used goodness-of-fit test in a Logit regression is the Hosmer-Lemeshow test (Hosmer and Lemeshow, 2004). In our specific case, it tests the null hypothesis of equality between the observed frequency of exceedances and that expected from the fitted model. For each observation $y_i$ in the sample, the predicted probability $\pi_i$ of exceeding the threshold is computed. Then, the $y_i$'s are split in $G$ groups of size $N_g$ according to the rank of their predicted probabilities, with

$g \in \{1, \ldots, G\}$. Finally, for each group $g$, the average predicted probability $\pi_g = (N_g)^{-1} \sum_{i_g=1}^{N_g} y_{i_g}$ and the expected number of exceedances $E_g = N_g \pi_g (1 - \pi_g)$ are computed. The test statistic then becomes a Pearson $\chi^2$ statistic with the following form,

$$H = \sum_{g=1}^{G} \frac{(O_g - E_g)^2}{E_g}$$

where $O_g$ is the number of observed exceedances in the $g$th group. The test statistic follows asymptotically a $\chi^2$ distribution with $G - 2$ degrees of freedom. As there are no specific rules to set the number of groups, we set $G = 10$ as it is usual in this literature, confident that our large sample size leads to enough observations in every decile.

Table 56 reports the p-values for both the deviance $\chi^2$ and the Hosmer-Lemeshow tests. In the first test, the null hypothesis of equal explanatory power between the null and the fitted model is rejected for all the specifications, across the three windows. At the same time, the null hypothesis of the Hosmer-Lemeshow is rejected on only one occasion, suggesting that in general the observed exceedance rate and that implied by the model do not differ significantly. These results indicate a good fit of the Logit model across the different windows.

The usual approach to validate a non-homogeneous Poisson process with intensity $\lambda_t$ is to transform it into a unit-rate homogeneous Poisson process (Ogata, 1988), using the transformation

$$t_i^H = \int_0^{t_i^{NH}} \lambda_t dt$$

where $t_i^{NH}$ and $t_i^H$ are the times of the occurrence $i$ for the non-homogeneous and corresponding unit-rate homogeneous Poisson processes, respectively. Then, the validation analysis consists in analyzing the inter-arrival times of the homogeneous Poisson process, $d_i = t_i^H - t_{i-1}^H$, that under the null hypothesis must be an *iid* exponential sample. Equivalently, the validation can be based on the transformed distances $\exp(-d_i)$ that must be an *iid* uniform sample and Table 57 reports the p-values of the

Kolmogorov-Smirnov test on this assumption. To account for the uncertainty associated to the estimated parameters, we compute the distribution of the test statistic by means of Monte Carlo simulations[7]. Non rejection of the null across the different windows suggests that the non-homogeneous Poisson process presents a good fit of the observed exceedance rate.

To assess the goodness of fit of the dynamic GP distribution as a model for the size of the excesses, we perform a graphical validation as in Coles (2001). When data are assumed to be identically distributed, goodness of fit can be performed by means of a qq-plot or formal testing as in Choulakian and Stephens (2001). However, in the non-stationary case, the lack of homogeneity among observations means that some modifications are required. Diagnostic procedures are applied to a standardized version of the data, conditional on the fitted parameter values. Consider a threshold $\widehat{u}$ leading to $k$ threshold exceedances $y_1, \ldots, y_k$ where the estimated model is

$$Y_j - \widehat{u} \sim GP(\widehat{\nu}_j, \widehat{\xi}_j)$$

with $j \in 1, \ldots, k$ and where $\widehat{\nu}_j$ and $\widehat{\xi}_j$ are respectively the estimated scale and shape parameters of the GP distribution at the time when the exceedance occurs. Transforming the observations $Y_j - \widehat{u}$ to a standard exponential distribution

$$\widetilde{Y}_j = \frac{1}{\widehat{\xi}_j} \log\left\{1 + \widehat{\xi}_j \left(\frac{Y_j - \widehat{u}}{\widehat{\sigma}_j}\right)\right\} \tag{5.19}$$

and denoting the ordered values of the observed $\widetilde{Y}_j$'s as $\tilde{y}_{(1)}, \ldots, \tilde{y}_{(k)}$, a quantile plot is obtained using the pairs

$$\left\{(\tilde{y}_{(i)}, -\log(1 - i/(k+1)));\ i = 1, \ldots, k\right\}.$$

The latter plot appears in Figure 53 for our dynamic GP distribution. The fits are generally satisfactory for all the specifications across the three windows considered. These results confirm that the dynamic GP assumption for the excesses over the *90*th quantile is representative of the underlying data and confirm the adequacy of the chosen threshold.

[7]Note that using the theoretical limiting distribution of the test statistic does not change

**Table 56: Diagnostics for Logit Model**. P-values for the $\chi^2$ test on the difference between the deviance residuals and the null residuals and for the Hosmer-Lemeshow test on the null of equality between the expected and observed frequency of exceedances. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | *$\chi^2$ test* | | | *Hosmer-Lemeshow test* | | |
|---|---|---|---|---|---|---|
| *Model* | $2000-2004$ | $2005-2009$ | $2010-2014$ | $2000-2004$ | $2005-2009$ | $2010-2014$ |
| I. | 0.000*** | 0.000*** | 0.000*** | 0.319 | 0.627 | 0.552 |
| II. | 0.000*** | 0.000*** | 0.000*** | 0.009** | 0.485 | 0.987 |
| III. | 0.000*** | 0.000*** | 0.000*** | 0.391 | 0.764 | 0.468 |
| IV. | 0.000*** | 0.000*** | 0.000*** | 0.777 | 0.618 | 0.823 |

**Table 57: Diagnostics for non-homogeneous Poisson models.** P-values for Kolmogorov-Smirnov test on the uniform assumption for the transformed residuals of the process. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| Model | 2000-2004 | 2005-2009 | 2010-2014 |
|---|---|---|---|
| I. | 0.852 | 0.122 | 0.401 |
| II. | 0.918 | 0.138 | 0.192 |
| III. | 0.284 | 0.150 | 0.344 |
| IV. | 0.696 | 0.124 | 0.286 |

**Figure 53: QQ-plot for dynamic GP distribution**. From top to bottom, time intervals 2000-2004, 2005-2009 and 2010-2014, respectively. From left to right, specifications I, II, III, IV, respectively. Each panel has the theoretical quantiles of a unit-rate exponential distribution on the $x$-axis, and the empirical quantiles obtained with transformation (5.19) on the $y$-axis.

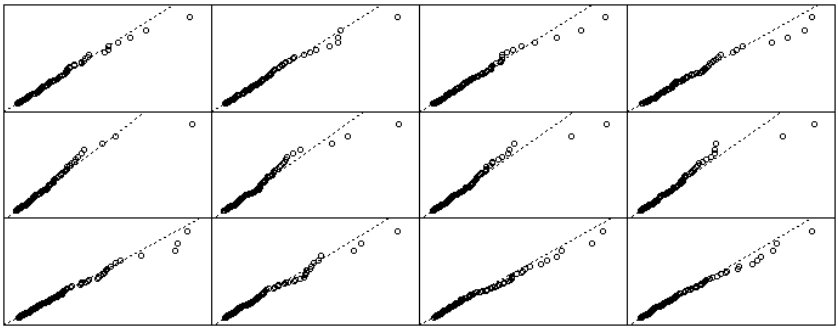


**Table 58: Model selection**. BIC of Logit models, non-homogeneous Poisson process and dynamic GP distribution for each window. The specification which presents the lowest BIC appears in bold.

| | Logit model | | | | non-homogeneous Poisson model | | | | Dynamic GP | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | I. | II. | III. | IV. | I. | II. | III. | IV. | I. | II. | III. | IV. |
| 2000-2004 | 761.53 | 755.37 | 744.54 | **743.73** | 781.78 | 778.19 | **765.84** | 766.57 | -960.87 | **-962.37** | -951.92 | -953.08 |
| 2005-2009 | **716.70** | 724.42 | 717.45 | 725.13 | **748.38** | 756.08 | 750.95 | 758.56 | -862.92 | -863.79 | **-875.97** | -875.84 |
| 2010-2014 | 786.52 | 790.50 | **773.63** | 779.23 | 805.51 | 810.11 | **791.75** | 797.59 | -981.67 | **-995.02** | -946.44 | -957.36 |

### 5.4.4 Model selection

In Section 5.4.3 we show that several specifications of both the exceedance rate models and the dynamic GP distribution adequately fit the data. As a consequence, selecting the simplest model that explains as much of the variation in the data as possible becomes a relevant issue.

Table 58 reports the values for the BIC (Schwarz, 1978) obtained for both the exceedance rate models and the GP distribution. The table emphasizes how the best specification differs across the three samples, coherently with the different regimes. Indeed, different behaviour in the extremes may require more or less structured models. The results suggest that a specification for the exceedance rate which contains longer past volatilities is preferred in 2000-2004 and 2010-2014, while the one-day lag volatility is sufficient in the middle sample. An insight on these results comes from Figure 51. It is evident that, in the second sub-sample, most of the exceedances occur over a short time, reflecting the higher risk perception after the defaults in the banking and insurance sectors in 2008. On the contrary, exceedances tend to occur more uniformly in the other sub-samples. The fast increase in the exceedance rate during the period 2005-2009 explains why a model depending only on the most recent volatility may be preferred. Contrarily, the opposite situation seems to occur on the specification for the size of the exceedance. Specifically, in this case a more structured model seems to be preferred in the sub-sample 2005-2009.

---

the results.

### 5.4.5 Do HF data provide additional valuable information?

In the previous sections, we have seen that RPOT models fit the data reasonably well, confirming that HF based measures are informative with respect to threshold exceedances. An interesting question is whether HF data add information beyond that carried by low frequency (LF) observations. As previously indicated in Section 1, past works on non-stationary EVT in finance exploit only information on past exceedances to learn the behaviour of future exceedances. We now consider models for the exceedance rate and the size of excesses using both past exceedances and the realized variation as covariates. Our intention here is to verify whether the HF-based measures add information to that conveyed by past exceedances.

Let $I_t$ be an indicator function taking value 1 when an exceedance occurs, i.e. $t \in \mathcal{E}$ and 0 otherwise. Let $W_t$ be excess size at time $t \in \mathcal{E}$. In Table 59, we report the results for the Logit model with the lagged variable $I_{t-1}$ as a covariate. In the first row, we see that the coefficient $\varphi_{LF}$ is significant in the first two periods. This result is consistent with that obtained in Chavez-Demoulin et al. (2005) with a self-exiting process. In the second row, we see that adding RV can reduce the explanatory power of $I_{t-1}$. The coefficient $\varphi_{HF}$ is strongly significant and similar in magnitude to that observed in Table 53. Similar conclusions are derived from the parameter estimates of the non-homogeneous Poisson models in Table 60.

As to the excess size, we proceed as in Chavez-Demoulin et al. (2005) and assume that exceedances are Markov with $W_t|W_{t-1}$ distributed as a GP distribution with $\nu_t$ depending on $W_{t-1}$. As shown in Table 61, the LF coefficient $\kappa_{LF}$ tends to be significant, but adding RV strongly reduces its explanatory power and the HF coefficient $\kappa_{HF}$ is significant. Particularly noteworthy is the significant additional contribution of the HF data during the 2005-2009 period which contains the crisis: $\varphi_{LF}$ and $\varphi_{HF}$ in Table 59, and $\kappa_{LF}$ and $\kappa_{HF}$ in Table 61, are significant for that period.

Overall, this analysis shows that HF data convey information beyond that provided by LF data on the behaviour of the extremes, confirming the merits of the RPOT approach.

**Table 59: Fitted Logit models with HF and LF covariates**.

*LF*. $\text{logit}(\phi_t) = \varphi + \varphi_{LF} I_{t-1}$

*HF*. $\text{logit}(\phi_t) = \varphi + \varphi_{LF} I_{t-1} + \varphi_{HF} \log(RV_{t-1})$

on S&P500 returns. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | 2000-2004 | | | 2005-2009 | | | 2010-2014 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\varphi$ | $\varphi_{LF}$ | $\varphi_{HF}$ | $\varphi$ | $\varphi_{LF}$ | $\varphi_{HF}$ | $\varphi$ | $\varphi_{LF}$ | $\varphi_{HF}$ |
| *LF.* | -2.26*** | 0.61* | | -2.26*** | 0.59* | | -2.22*** | 0.31 | |
| | (0.10) | (0.26) | | (0.10) | (0.26) | | (0.10) | (0.28) | |
| *HF.* | 5.61*** | -0.11 | 0.86*** | 5.70*** | -0.62* | 0.85*** | 4.34*** | -0.46 | 0.66*** |
| | (1.09) | (0.28) | (0.12) | (0.76) | (0.30) | (0.08) | (0.94) | (0.31) | (0.10) |

**Table 60: Fitted non-homogeneous Poisson models with HF and LF covariates**.

*LF*. $\lambda_t = \exp\left(\zeta + \zeta_{LF} I_{t-1}\right)$

*HF*. $\lambda_t = \exp\left(\zeta + \zeta_{LF} I_{t-1} + \zeta_{HF} \log(RV_{t-1})\right)$

on S&P500 returns. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | 2000-2004 | | | 2005-2009 | | | 2010-2014 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\zeta$ | $\zeta_{LF}$ | $\zeta_{HF}$ | $\zeta$ | $\zeta_{LF}$ | $\zeta_{HF}$ | $\zeta$ | $\zeta_{LF}$ | $\zeta_{HF}$ |
| *LF.* | -2.37*** | 0.54* | | -2.36*** | 0.52* | | -2.33*** | 0.27 | |
| | (0.10) | (0.24) | | (0.10) | (0.24) | | (0.09) | (0.26) | |
| *HF.* | 4.26*** | -0.09 | 0.72*** | 3.70*** | -0.42 | 0.65*** | 3.21*** | -0.39 | 0.56*** |
| | (0.93) | (0.25) | (0.10) | (0.55) | (0.25) | (0.06) | (0.80) | (0.28) | (0.08) |

## 5.5 Out-of-sample forecast

In Section 5.4, we show that HF data contribute significantly toward explaining the behaviour of daily extreme returns. However, in-sample fit does not guarantee a satisfactory out-of-sample forecast performance. In

**Table 61: Fitted dynamic Generalized Pareto models with HF and LF covariates**. The $\xi$ parameter is constant while the $\nu_t$ parameter is allowed to vary according to

$$LF.\ \ \nu_t = \exp\left(\kappa + \kappa_{LF} W_{t-1}\right)$$

$$HF.\ \ \nu_t = \exp\left(\kappa + \kappa_{LF} W_{t-1} + \kappa_{HF} \log(RV_{t-1})\right)$$

*,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | 2000-2004 | | | 2005-2009 | | | 2010-2014 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\kappa$ | $\kappa_{LF}$ | $\kappa_{HF}$ | $\kappa$ | $\kappa_{LF}$ | $\kappa_{HF}$ | $\kappa$ | $\kappa_{LF}$ | $\kappa_{HF}$ |
| *LF.* | -4.94*** | 3.29 | | -4.75*** | 26.39*** | | -5.01*** | 40.89*** | |
| | (0.17) | (12.92) | | (0.10) | (6.27) | | (0.13) | (8.63) | |
| *HF.* | -2.26* | -0.11 | 0.31*** | -1.87*** | 15.19** | 0.32*** | -1.44 | 16.61 | 0.36*** |
| | (1.01) | (12.93) | (0.11) | (0.59) | (5.24) | (0.06) | (0.98) | (9.95) | (0.09) |

this section, we investigate whether HF data also lead to good out-of-sample forecasts. To this end, we perform an out-of-sample analysis of the one-day-ahead VaR and ES forecasts defined respectively in Equations (5.17)-(5.18). As a benchmark, we also report the results for the commonly used conditional EVT (C-EVT) model of McNeil and Frey (2000). For the latter, we filter the daily returns with an ARMA(1,1)-GJR-GARCH(1,1) and then model the upper tail of the negated estimated residuals with the POT method.

We apply a rolling-window scheme to obtain a time series of VaR and ES predictions at level $\alpha = 0.01$. Let $n$ be the size of the available sample and $s$ be the length of the window. We have two sequences of forecasts: $\{VaR_t^\alpha\}_{t=s+1}^n$ and $\{ES_t^\alpha\}_{t=s+1}^n$ of length $m = n - s$, where each prediction is obtained considering the observations $l_{t-s}, \ldots, l_{t-1}$. We investigate the performance of the RPOT and C-EVT approaches during the *Global Financial Crisis* and the *European sovereign debt crises*[8]. We consider the S&P500 from the beginning of 2000 to the end of 2012, resulting in a time series of $n = 3240$ observations. This allows us to produce $m = 1240$ predictions, when considering a window of size $s = 2000$.

[8]The Global Financial Crisis refers to the period of turmoil that affected the stock market in the period 2007-2009, following the failure of several banks and insurance companies. The European sovereign debt crises indicates the period of distress that affected the stock market in 2010-2012, due to the deterioration in the creditworthiness of several European countries.

The threshold level is fixed at the *90*th quantile of the unconditional loss distribution obtained from the $s$ observations. Figure 54 shows one-day-ahead VaR and ES forecasts and their 95% confidence intervals for the simplest RPOT specification.

We evaluate the performance of the RPOT approach by performing a battery of tests. We consider the binary indicator of VaR failure

$$H_{t+1} = I\left(l_{t+1} > \widehat{VaR}^{\alpha}_{t+1|t}\right),$$

where $I(\cdot)$ is the indicator function. Commonly used tests are the Unconditional Coverage (UC), Independence (IND), and Conditional Coverage (CC) suggested by Christoffersen (1998) and the Dynamic Quantile (DQ) test suggested by Engle and Manganelli (2004). To evaluate the performance for the ES we test the hypothesis that conditional upon exceeding the *99*th quantile of the loss distribution, the difference between the actual return and the predicted ES has mean zero. In particular, we perform a one-sided test with the alternative that the mean is greater than zero using a bootstrap that makes no assumption about the distribution of the differences (McNeil and Frey, 2000).

In Table 62, we report the p-values of the different tests. These results show that the RPOT models outperform the C-EVT. In particular, while all specifications of the RPOT approach yield large p-values, the null of UC is rejected for the C-EVT at the 5% level. Furthermore, the RPOT approach does not exhibit any problem with respect to the independence requirement, while we reject the null hypothesis of the DQ test for C-EVT. Finally, the superior performance of the RPOT is also confirmed by the CC test where the RPOT performs reasonably well and the C-EVT fails the test at the 10% level.

## 5.6 Further analysis

### 5.6.1 Sensitivity to the threshold choice

An important question when modelling the extremes using any form of POT is whether the results are robust to the choice of the threshold. Al-

**Figure 54: Risk measures**. In green, estimated conditional VaR (left panel) and ES (right panel) at level $\alpha = 0.01$ with the RPOT approach. RV is the only covariate in the dynamic GP distribution and the Logit model used for the exceedance rate. In blue, 95% confidence intervals obtained with the post-blackened bootstrap described in Section 5.3.3.

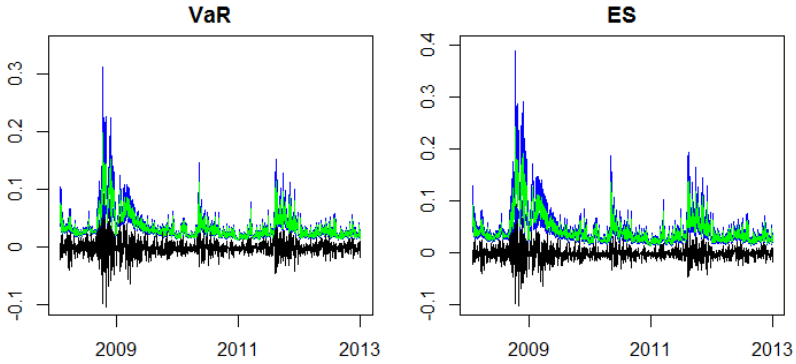



**Table 62: VaR measures**. For each model (C-EVT, Logit (L), Poisson(Nhp)) and specifications (I-IV), we report: the percentage of violations (Violation); the p-values for the unconditional coverage (UC), the independence assumption (IND), the conditional coverage (CC), and the DQ tests, and the bootstrap test for the ES (BOOT). Rejection at level of significance 0.05 appears in bold.

| | **Violation** | **UC** | **IND** | **CC** | **DQ** | **BOOT** |
|---|---|---|---|---|---|---|
| *C-EVT* | 1.61 | **0.046** | 0.418 | 0.097 | **0.022** | 0.939 |
| *L.I* | 1.29 | 0.325 | 0.518 | 0.494 | 0.867 | 0.348 |
| *L.II* | 1.29 | 0.325 | 0.518 | 0.494 | 0.910 | 0.459 |
| *L.III* | 1.29 | 0.325 | 0.199 | 0.267 | 0.040 | 0.714 |
| *L.IV* | 1.29 | 0.325 | 0.518 | 0.494 | 0.120 | 0.638 |
| *Nhp.I* | 1.29 | 0.325 | 0.518 | 0.494 | 0.866 | 0.273 |
| *Nhp.II* | 1.37 | 0.214 | 0.492 | 0.360 | 0.866 | 0.415 |
| *Nhp.III* | 1.45 | 0.134 | 0.260 | 0.170 | 0.055 | 0.753 |
| *Nhp.IV* | 1.29 | 0.325 | 0.518 | 0.494 | 0.120 | 0.565 |

though we have seen that the quality of the fit at the *90*th quantile is remarkably good, we also perform an analysis at a threshold level corresponding to the *95*th quantile. Results reported in Appendix D.1 confirm the findings of the previous sections. While the one-day lagged RV and BV have a highly significant impact in both the models for the exceedance rate and the GP distribution, the jump component and the past volatilities are significant on very few occasions. In this case as well, the diagnostic tools support the good fit of the different specifications. Finally, an analysis of the out-of-sample risk measures forecasts confirms that the RPOT approach outperforms the C-EVT.

### 5.6.2 An investigation over a shorter sample

We investigate whether good forecasts of the risk measures are maintained even when the models are estimated over shorter training periods. In particular, while in Section 5.5 we use a rolling window of $2000$ observations to fit the models, here we reduce this interval to $1000$ observations. The intention behind this exercise is to verify how the performance of the models are affected by a higher degree of uncertainty in parameter estimates. Results reported in Appendix D.2 show that the good quality of the RPOT is maintained.

### 5.6.3 Impact of negative jumps

In Section 5.4, we considered the jump variation as a possible covariate and ascertained that it does not play a relevant role in forecasting the extremes. However, as we focussed only on the negative tail, it could be the case that only the negative jumps are relevant. Patton and Sheppard (2015) find that distinguishing between positive and negative jumps is extremely relevant when estimating the impact of the jumps on future volatilities. Here we use the signed jump variation, defined as

$$\Delta J_t = RS_t^+ - RS_t^-$$

where $RS_t^+$ and $RS_t^-$ are respectively the positive and negative realized semi-variances. This measure converges in probability to the daily dif-

ference of jump variation induced by jumps of opposite signs. A further decomposition allows us to obtain the positive and negative jump variations which are defined respectively as,

$$\begin{aligned} \Delta J_t^+ &= \Delta J_t \cdot I\left(RS_t^+ - RS_t^- > 0\right) \\ \Delta J_t^- &= \Delta J_t \cdot I\left(RS_t^+ - RS_t^- < 0\right) \end{aligned}$$

where $I(\cdot)$ is the indicator function. Assuming that jumps are rare, these quantities broadly capture the variation of positive and negative jumps.

We substitute the jump variation $J$ with the negative jump variation $\Delta J^-$ in models for the exceedance rate and the dynamic GP distribution. Results reported in Appendix D.3 suggest that the negative jump component also does not have a substantial impact on the extremes.

## 5.7 Conclusions

The availability of high-frequency data has lead to breakthroughs in the financial econometrics literature, and models that exploit this source of information are superseding standard econometric models. In this paper, we propose a novel high-frequency extreme value approach where realized measures are used to model the time-varying behaviour of extreme returns. In-sample fit of these models shows that high-frequency data add information on the extremes, beyond that conveyed by low-frequency data. Moreover, out-of-sample forecasts of standard risk measures outperform those of the standard conditional EVT approach.

Our proposed approach works well and we intend to build upon it. We are working on refinements and extensions of the RPOT with the aim of establishing a complete framework where high-frequency data are used within extreme value models. For example, adding parametric or non-parametric smoothing components could enhance the stability of the extreme value models. Furthermore, the intuition of this paper can be used to extend the time-varying threshold model of Wang et al. (2012) to financial returns. Finally, the development of multivariate models where the realized covariance is used as a source of information would allow for the modelling of the joint occurrence of extreme events. Such a frame-

work could be used to provide a much needed extreme value perspective measure of contagion effects among assets.

# Appendix D

## D.1 Sensitivity to the threshold choice

Throughout the paper, we consider as extremes the observations exceeding the *90*th quantile of the unconditional distribution. In order to assess that the results obtained are not a mere consequence of the threshold choice, we proceed with a sensitivity analysis. In this section we consider a threshold corresponding to the *95*th quantile of the unconditional distribution. In Tables 63-65, we report the in-sample results for both the models for the exceedance rate and for the GP distribution. We can see that also at this threshold level, RV and BV are strongly significant, while the jump component and past lagged variances are not particularly relevant. The results reported in Tables 66-67 and Figure 55 confirm that the models adequately fit the data. Table 69 reports the results of the out-of-sample analysis, and confirms that the RPOT approach still outperforms the C-EVT.

**Table 63: Fitted Logit models** for specifications:

I. $\text{logit}(\phi_t) = \varphi_0 + \varphi_1 \log(RV_{t-1})$

II. $\text{logit}(\phi_t) = \varphi_0 + \varphi_2 \log(BV_{t-1}) + \varphi_3 \log(1 + J_{t-1})$

III. $\text{logit}(\phi_t) = \varphi_0 + \varphi_1 \log(RV_{t-1}) + \varphi_4 \log(\overline{RV}^W_{t-1}) + \varphi_5 \log(\overline{RV}^M_{t-1})$

IV. $\text{logit}(\phi_t) = \varphi_0 + \varphi_2 \log(BV_{t-1}) + \varphi_3 \log(1 + J_{t-1}) + \varphi_4 \log(\overline{RV}^W_{t-1}) + \varphi_5 \log(\overline{RV}^M_{t-1})$

on S&P500 returns. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\varphi_0$ | $\varphi_1$ | $\varphi_2$ | $\varphi_3$ | $\varphi_4$ | $\varphi_5$ |
|---|---|---|---|---|---|---|
| | | | 2000-2004 | | | |
| I. | 6.40*** | 1.04*** | | | | |
| | (1.35) | (0.15) | | | | |
| II. | 7.62*** | | 1.15*** | -0.06 | | |
| | (1.44) | | (0.16) | (0.1) | | |
| III. | 11.07*** | 0.39 | | | 0.69* | 0.49 |
| | (1.93) | (0.24) | | | (0.32) | (0.31) |
| IV. | 12.14*** | | 0.63** | -0.11 | 0.56 | 0.49 |
| | (2.06) | | (0.24) | (0.11) | (0.33) | (0.31) |
| | | | 2005-2009 | | | |
| I. | 5.69*** | 0.97*** | | | | |
| | (0.87) | (0.1) | | | | |
| II. | 5.84*** | | 0.96*** | -0.01 | | |
| | (0.99) | | (0.11) | (0.07) | | |
| III. | 6.75*** | 0.26 | | | 0.60* | 0.24 |
| | (0.96) | (0.23) | | | (0.29) | (0.23) |
| IV. | 6.76*** | | 0.23 | 0.00 | 0.62* | 0.23 |
| | (1.04) | | (0.24) | (0.07) | (0.3) | (0.23) |
| | | | 2010-2014 | | | |
| I. | 5.01*** | 0.83*** | | | | |
| | (1.12) | (0.12) | | | | |
| II. | 6.60*** | | 0.96*** | -0.11 | | |
| | (1.4) | | (0.14) | (0.11) | | |
| III. | 7.05*** | 0.57** | | | 0.06 | 0.41 |
| | (1.41) | (0.19) | | | (0.25) | (0.22) |
| IV. | 8.48*** | | 0.74*** | -0.12 | -0.03 | 0.46* |
| | (1.6) | | (0.21) | (0.11) | (0.26) | (0.22) |

**Table 64: Fitted non-homogeneous Poisson models** for specifications:

I. $\lambda_t = \exp\left(\zeta_0 + \zeta_1 \log(RV_{t-1})\right)$

II. $\lambda_t = \exp\left(\zeta_0 + \zeta_2 \log(BV_{t-1}) + \zeta_3 \log(1 + J_{t-1})\right)$

III. $\lambda_t = \exp\left(\zeta_0 + \zeta_1 \log(RV_{t-1}) + \zeta_4 \log(\overline{RV}^W_{t-1}) + \zeta_5 \log(\overline{RV}^M_{t-1})\right)$

IV. $\lambda_t = \exp\left(\zeta_0 + \zeta_2 \log(BV_{t-1}) + \zeta_3 \log(1 + J_{t-1}) + \zeta_4 \log(\overline{RV}^W_{t-1}) + \zeta_5 \log(\overline{RV}^M_{t-1})\right)$

on S&P500 returns. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\zeta_0$ | $\zeta_1$ | $\zeta_2$ | $\zeta_3$ | $\zeta_4$ | $\zeta_5$ |
|---|---|---|---|---|---|---|
| | | | 2000-2004 | | | |
| I. | 5.39*** | 0.93*** | | | | |
| | (1.19) | (0.14) | | | | |
| II. | 6.22*** | | 1.01*** | -0.05 | | |
| | (1.21) | | (0.14) | (0.1) | | |
| III. | 9.68*** | 0.35 | | | 0.61* | 0.47 |
| | (1.74) | (0.22) | | | (0.3) | (0.29) |
| IV. | 10.46*** | | 0.54* | -0.09 | 0.49 | 0.48 |
| | (1.83) | | (0.22) | (0.1) | (0.31) | (0.29) |
| | | | 2005-2009 | | | |
| I. | 4.27*** | 0.82*** | | | | |
| | (0.69) | (0.08) | | | | |
| II. | 4.46*** | | 0.82*** | -0.02 | | |
| | (0.78) | | (0.09) | (0.05) | | |
| III. | 5.25*** | 0.22 | | | 0.49 | 0.22 |
| | (0.79) | (0.21) | | | (0.26) | (0.21) |
| IV. | 5.29*** | | 0.2 | -0.01 | 0.52* | 0.22 |
| | (0.85) | | (0.21) | (0.06) | (0.26) | (0.21) |
| | | | 2010-2014 | | | |
| I. | 4.09*** | 0.74*** | | | | |
| | (0.98) | (0.11) | | | | |
| II. | 5.34*** | | 0.84*** | -0.10 | | |
| | (1.17) | | (0.12) | (0.1) | | |
| III. | 6.02*** | 0.51** | | | 0.06 | 0.37 |
| | (1.29) | (0.17) | | | (0.23) | (0.21) |
| IV. | 7.24*** | | 0.66*** | -0.10 | -0.03 | 0.42* |
| | (1.43) | | (0.2) | (0.1) | (0.25) | (0.2) |

**Table 65: Fitted dynamic Generalized Pareto models**. The $\xi$ parameter is constant, while $\nu_t$ is allowed to vary according to

I. $\nu_t = \exp\left(\kappa_0 + \kappa_1 \log(RV_{t-1})\right)$

II. $\nu_t = \exp\left(\kappa_0 + \kappa_2 \log(BV_{t-1}) + \kappa_3 \log(1 + J_{t-1})\right)$

III. $\nu_t = \exp\left(\kappa_0 + \kappa_1 \log(RV_{t-1}) + \kappa_4 \log(\overline{RV}^W_{t-1}) + \kappa_5 \log(\overline{RV}^M_{t-1})\right)$

IV. $\nu_t = \exp\left(\kappa_0 + \kappa_2 \log(BV_{t-1}) + \kappa_3 \log(1 + J_{t-1}) + \kappa_4 \log(\overline{RV}^W_{t-1}) + \kappa_5 \log(\overline{RV}^M_{t-1})\right)$

*,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\kappa_0$ | $\kappa_1$ | $\kappa_2$ | $\kappa_3$ | $\kappa_4$ | $\kappa_5$ | $\xi$ |
|---|---|---|---|---|---|---|---|
| | | | | 2000-2004 | | | |
| I. | -3.49** | 0.16 | | | | | 0.00 |
| | (1.3) | (0.15) | | | | | (0.13) |
| II. | -2.96* | | 0.22 | -0.09 | | | -0.02 |
| | (1.34) | | (0.15) | (0.1) | | | (0.13) |
| III. | -2.77 | 0.14 | | | 0.25 | -0.13 | 0.03 |
| | (2.21) | (0.22) | | | (0.28) | (0.35) | (0.13) |
| IV. | -2.49 | | 0.21 | -0.09 | 0.25 | -0.17 | 0.01 |
| | (2.24) | | (0.23) | (0.11) | (0.28) | (0.35) | (0.14) |
| | | | | 2005-2009 | | | |
| I. | -0.74 | 0.48*** | | | | | 0.13 |
| | (0.92) | (0.12) | | | | | (0.14) |
| II. | -0.58 | | 0.48*** | -0.01 | | | 0.10 |
| | (1.03) | | (0.12) | (0.07) | | | (0.14) |
| III. | -0.18 | 0.11 | | | 0.48 | -0.05 | -0.03 |
| | (0.85) | (0.16) | | | (0.26) | (0.2) | (0.15) |
| IV. | -0.41 | | 0.08 | 0.00 | 0.49 | -0.06 | -0.03 |
| | (0.9) | | (0.18) | (0.08) | (0.27) | (0.21) | (0.15) |
| | | | | 2010-2014 | | | |
| I. | 0.42 | 0.56*** | | | | | -0.27* |
| | (1.08) | (0.12) | | | | | (0.12) |
| II. | 2.42* | | 0.74* | -0.19* | | | -0.43*** |
| | (0.98) | | (0.1) | (0.09) | | | (0.11) |
| III. | 0.79 | 0.27 | | | 0.6*** | -0.29* | -0.57** |
| | (0.91) | (0.14) | | | (0.14) | (0.12) | (0.18) |
| IV. | 1.97*** | | 0.44*** | -0.1* | 0.4*** | -0.16* | -0.74* |
| | (0.27) | | (0.04) | (0.05) | (0.07) | (0.07) | (0.26) |

**Table 66: Diagnostics for Logit Model**. P-values for the $\chi^2$ test on the difference between the deviance residuals and the null residuals and for the Hosmer-Lemeshow test on the null of equality between the expected and observed frequency of exceedances. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\chi^2 test$ | | | *Hosmer-Lemeshow test* | | |
|---|---|---|---|---|---|---|
| Model | 200-2004 | 2005-2009 | 2010-2014 | 2000-2004 | 2005-2009 | 2010-2014 |
| I. | 0.000*** | 0.000*** | 0.000*** | 0.776 | 0.551 | 0.178 |
| II. | 0.000*** | 0.000*** | 0.000*** | 0.118 | 0.456 | 0.937 |
| III. | 0.000*** | 0.000*** | 0.000*** | 0.137 | 0.527 | 0.847 |
| IV. | 0.000*** | 0.000*** | 0.000*** | 0.656 | 0.365 | 0.537 |

**Table 67: Diagnostics for non-homogeneous Poisson models.** P-values for Kolmogorov-Smirnov test on the uniform assumption for the transformed residuals of the process. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| Model | 2000-2004 | 2005-2009 | 2010-2014 |
|---|---|---|---|
| I. | 0.905 | 0.058 | 0.183 |
| II. | 0.911 | 0.025* | 0.170 |
| III. | 0.922 | 0.181 | 0.662 |
| IV. | 0.822 | 0.184 | 0.819 |

**Figure 55: QQ-plot for dynamic GP distribution**. From top to bottom, time intervals 2000-2004, 2005-2009 and 2010-2014, respectively. From left to right, specifications I, II, III, IV, respectively. Each panel has the theoretical quantiles of a unit-rate exponential distribution on the $x$-axis, and the empirical quantiles obtained with the exponential transformation on the $y$-axis.

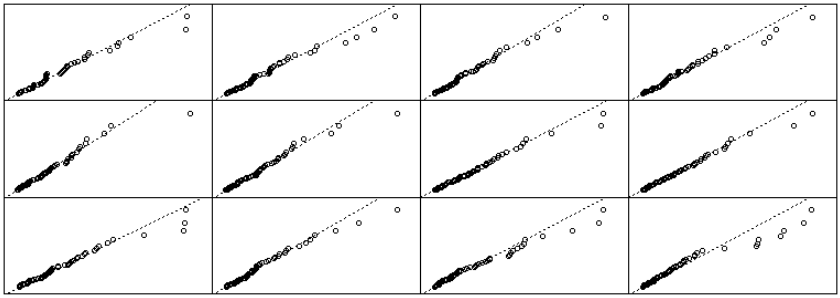


**Table 68: Model selection**. BIC of Logit models, non-homogeneous Poisson process and dynamic GP distribution for each window. The specification which presents the lowest BIC appears in bold.

| | Logit model | | | | non-homogeneous Poisson model | | | | Dynamic GP | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | I. | II. | III. | IV. | I. | II. | III. | IV. | I. | II. | III. | IV. |
| 2000-2004 | 458.24 | 456.72 | **436.17** | 438.58 | 464.96 | 464.71 | **443.67** | 446.77 | -471.35 | **-472.58** | -451.80 | -452.75 |
| 2005-2009 | **415.84** | 422.94 | 417.59 | 424.97 | **429.22** | 436.42 | 431.92 | 439.26 | -420.36 | -420.76 | **-426.49** | -426.19 |
| 2010-2014 | 466.80 | 470.67 | **462.76** | 466.83 | 473.74 | 478.24 | **469.72** | 474.12 | -490.19 | **-504.15** | -483.56 | -493.75 |

**Table 69: VaR measures**. For each model (C-EVT, Logit (L), Poisson(Nhp)) and specifications (I-IV), we report: the percentage of violations (Violation); the p-values for the unconditional coverage (UC), the independence assumption (IND), the conditional coverage (CC), and the DQ tests, and the bootstrap test for the ES (BOOT). Rejection at level of significance 0.05 appears in bold.

| | **Violation** | **UC** | **IND** | **CC** | **DQ** | **BOOT** |
|---|---|---|---|---|---|---|
| *C-EVT* | 0.016 | **0.046** | 0.418 | 0.097 | **0.022** | 0.938 |
| *L.I* | 0.012 | 0.473 | 0.544 | 0.635 | 0.952 | 0.245 |
| *L.II* | 0.011 | 0.655 | 0.572 | 0.762 | 0.991 | 0.284 |
| *L.III* | 0.011 | 0.655 | 0.572 | 0.762 | 0.023 | 0.229 |
| *L.IV* | 0.012 | 0.473 | 0.544 | 0.635 | 0.000 | 0.311 |
| *Nhp.I* | 0.010 | 0.865 | 0.600 | 0.850 | 0.997 | 0.158 |
| *Nhp.II* | 0.010 | 0.865 | 0.600 | 0.850 | 0.997 | 0.235 |
| *Nhp.III* | 0.012 | 0.473 | 0.544 | 0.635 | 0.035 | 0.303 |
| *Nhp.IV* | 0.013 | 0.325 | 0.518 | 0.494 | **0.000** | 0.360 |

## D.2 An investigation over a shorter sample

We now assess whether the RPOT approach keeps providing good forecasts of the risk measures when the window size is 1000 instead of 2000. Results reported in Table 70 show that in this case the C-EVT approach does not fail the tests of coverage. However, looking at the violations, it still seems that the RPOT does a better job, thus confirming the previous findings.

**Table 70: VaR measures**. For each model (C-EVT, Logit (L), Poisson(Nhp)) and specifications (I-IV), we report: the percentage of violations (Violation); the p-values for the unconditional coverage (UC), the independence assumption (IND), the conditional coverage (CC), and the DQ tests, and the bootstrap test for the ES (BOOT). Rejection at level of significance 0.05 appears in bold.

| | **Violation** | **UC** | **IND** | **CC** | **DQ** | **BOOT** |
|---|---|---|---|---|---|---|
| *C-EVT* | 0.014 | 0.146 | 0.469 | 0.263 | **0.039** | 0.978 |
| *L.I* | 0.009 | 0.655 | 0.659 | 0.814 | 0.996 | 0.289 |
| *L.II* | 0.008 | 0.455 | 0.688 | 0.692 | 0.996 | 0.347 |
| *L.III* | 0.006 | 0.168 | 0.748 | 0.364 | **0.060** | 0.211 |
| *L.IV* | 0.008 | 0.455 | 0.688 | 0.692 | **0.000** | 0.560 |
| *Nhp.I* | 0.010 | 0.897 | 0.602 | 0.856 | 0.999 | 0.368 |
| *Nhp.II* | 0.011 | 0.684 | 0.574 | 0.777 | 0.999 | 0.477 |
| *Nhp.III* | 0.008 | 0.455 | 0.688 | 0.692 | 0.209 | 0.298 |
| *Nhp.IV* | 0.010 | 0.877 | 0.630 | 0.871 | **0.005** | 0.607 |

## D.3 Adding the signed jump variation as covariate

In this section, we investigate the impact of the negative jumps, substituting the jump component $J$ with the negative jump variation $\Delta J_t^-$. In particular, as $\Delta J_t^-$ is either negative or zero, we use its negated version for practical purposes. Consequently, if the coefficient associated to $\Delta J_t^-$ has positive sign, then the higher the negative jump, the higher

the impact on the dependent variable. Tables 71-73 report the results for the threshold level at the *90*th quantile. They show that the negative jump has a negative and statistically significant impact on the exceedance rate models in the period 2005-2009, while it has a negative and statistically significant impact on the GP distribution in the period 2010-2014. These results are no longer present when we shift the threshold at the *95*th quantile, see Tables 74-76. Overall, we can claim that the negative jumps have no forecasting power on the daily extremes.

**Table 71: Fitted logit models with negative jump variation**. Parameter estimates for models:

II. $\text{logit}(\phi_t) = \varphi_0 + \varphi_2 \log(BV_{t-1}) + \varphi_3 \log(1 + \Delta J^-_{t-1})$

IV. $\text{logit}(\phi_t) = \varphi_0 + \varphi_2 \log(BV_{t-1}) + \varphi_3 \log(1 + \Delta J^-_{t-1}) + \varphi_4 \log(\overline{RV}^W_{t-1}) + \varphi_5 \log(\overline{RV}^M_{t-1})$

on S&P500 returns. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\varphi_0$ | $\varphi_2$ | $\varphi_3$ | $\varphi_4$ | $\varphi_5$ |
|---|---|---|---|---|---|
| | | | 2000-2004 | | |
| II. | 6.64*** | 0.95*** | 0.03 | | |
| | (1.16) | (0.13) | (0.08) | | |
| IV. | 8.52*** | 0.60*** | 0.00 | 0.35 | 0.23 |
| | (1.41) | (0.19) | (0.08) | (0.24) | (0.23) |
| | | | 2005-2009 | | |
| II. | 6.11*** | 0.88*** | -0.21* | | |
| | (0.80) | (0.09) | (0.09) | | |
| IV. | 6.80*** | 0.38* | -0.20* | 0.30 | 0.30 |
| | (0.83) | (0.18) | (0.09) | (0.23) | (0.17) |
| | | | 2010-2014 | | |
| II. | 4.66*** | 0.68*** | -0.02 | | |
| | (1.05) | (0.11) | (0.07) | | |
| IV. | 5.36*** | 0.54*** | -0.02 | 0.05 | 0.18 |
| | (1.20) | (0.16) | (0.07) | (0.19) | (0.16) |

**Table 72: Fitted non-homogeneous Poisson models with negative jump variation**. Parameter estimates for models:

II. $\lambda_t = \exp\left(\zeta_0 + \zeta_2 \log(BV_{t-1}) + \zeta_3 \log(1 + \Delta J^-_{t-1})\right)$

IV. $\lambda_t = \exp\left(\zeta_0 + \zeta_2 \log(BV_{t-1}) + \zeta_3 \log(1 + \Delta J^-_{t-1}) + \zeta_4 \log(\overline{RV}^W_{t-1}) + \zeta_5 \log(\overline{RV}^M_{t-1})\right)$

on the S&P500. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\zeta_0$ | $\zeta_2$ | $\zeta_3$ | $\zeta_4$ | $\zeta_5$ |
|---|---|---|---|---|---|
| | | | 2000-2004 | | |
| II. | 4.93*** | 0.78*** | -0.02 | | |
| | (0.95) | (0.10) | (0.06) | | |
| IV. | 6.76*** | 0.48*** | -0.00 | 0.28 | 0.22 |
| | (1.22) | (0.16) | (0.06) | (0.22) | (0.20) |
| | | | 2005-2009 | | |
| II. | 3.91*** | 0.66*** | -0.14* | | |
| | (0.56) | (0.06) | (0.07) | | |
| IV. | 4.58*** | 0.26 | -0.14* | 0.24 | 0.24 |
| | (0.61) | (0.15) | (0.07) | (0.19) | (0.15) |
| | | | 2010-2014 | | |
| II. | 3.47*** | 0.57*** | -0.02 | | |
| | (0.89) | (0.09) | (0.06) | | |
| IV. | 4.14*** | 0.45*** | -0.01 | 0.03 | 0.16 |
| | (1.05) | (0.14) | (0.06) | (0.17) | (0.14) |

**Table 73: Dynamic Generalized Pareto with negative Jump variation**. The $\xi$ parameter is kept constant, while the $\nu$ parameter is allowed to vary according to the following models:

II. $\nu_t = \exp\left(\kappa_0 + \kappa_2 \log(BV_{t-1}) + \kappa_3 \log(1 + \Delta J^-_{t-1})\right)$

IV. $\nu_t = \exp\left(\kappa_0 + \kappa_2 \log(BV_{t-1}) + \kappa_3 \log(1 + \Delta J^-_{t-1}) + \kappa_4 \log(\overline{RV}^W_{t-1}) + \kappa_5 \log(\overline{RV}^M_{t-1})\right)$

*,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\kappa_0$ | $\kappa_2$ | $\kappa_3$ | $\kappa_4$ | $\kappa_5$ |
|---|---|---|---|---|---|
| | | | 2000-2004 | | |
| II. | -1.78 | 0.36** | -0.03 | | |
| | (1.14) | (0.13) | (0.08) | | |
| IV. | 0.76 | 0.10 | -0.04 | 0.37 | 0.18 |
| | (1.57) | (0.16) | (0.08) | (0.20) | (0.23) |
| | | | 2005-2009 | | |
| II. | -0.98 | 0.40*** | 0.02 | | |
| | (0.65) | (0.07) | (0.07) | | |
| IV. | -0.72 | 0.02 | 0.05 | 0.52 | -0.10 |
| | (0.63) | (0.13) | (0.08) | (0.19) | (0.14) |
| | | | 2010-2014 | | |
| II. | 0.79 | 0.55*** | -0.13*** | | |
| | (0.62) | (0.06) | (0.04) | | |
| IV. | 1.20** | 0.43*** | -0.15*** | 0.18** | -0.04 |
| | (0.42) | (0.06) | (0.04) | (0.06) | (0.07) |

**Table 74: Fitted logit models with negative jump variation**. Parameter estimates for models:

II. $\text{logit}(\phi_t) = \varphi_0 + \varphi_2 \log(BV_{t-1}) + \varphi_3 \log(1 + \Delta J^-_{t-1})$

IV. $\text{logit}(\phi_t) = \varphi_0 + \varphi_2 \log(BV_{t-1}) + \varphi_3 \log(1 + \Delta J^-_{t-1}) + \varphi_4 \log(\overline{RV}^W_{t-1}) + \varphi_5 \log(\overline{RV}^M_{t-1})$

on S&P500 returns. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\varphi_0$ | $\varphi_2$ | $\varphi_3$ | $\varphi_4$ | $\varphi_5$ |
|---|---|---|---|---|---|
| | | | 2000-2004 | | |
| II. | 8.13*** | 1.20*** | -0.12 | | |
| | (1.51) | (0.17) | (0.10) | | |
| IV. | 11.89*** | 0.67** | -0.08 | 0.52 | 0.45 |
| | (2.02) | (0.25) | (0.10) | (0.32) | (0.31) |
| | | | 2005-2009 | | |
| II. | 6.40*** | 1.02*** | -0.10 | | |
| | (1.01) | (0.12) | (0.08) | | |
| IV. | 7.29*** | 0.29) | -0.09 | 0.61* | 0.24 |
| | (1.06) | (0.24) | (0.08) | (0.29) | (0.23) |
| | | | 2010-2014 | | |
| II. | 6.91*** | 0.99*** | -0.15 | | |
| | (1.37) | (0.14) | (0.11) | | |
| IV. | 8.52*** | 0.81*** | -0.15 | 0.09 | 0.45* |
| | (1.54) | (0.22) | (0.11) | (0.26) | (0.22) |

**Table 75: Fitted non-homogeneous Poisson models with negative jump variation**. Parameter estimates for models:

II. $\lambda_t = \exp\left(\zeta_0 + \zeta_2 \log(BV_{t-1}) + \zeta_3 \log(1 + \Delta J^-_{t-1})\right)$

IV. $\lambda_t = \exp\left(\zeta_0 + \zeta_2 \log(BV_{t-1}) + \zeta_3 \log(1 + \Delta J^-_{t-1}) + \zeta_4 \log(\overline{RV}^W_{t-1}) + \zeta_5 \log(\overline{RV}^M_{t-1})\right)$

on the S&P500. *,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\zeta_0$ | $\zeta_2$ | $\zeta_3$ | $\zeta_4$ | $\zeta_5$ |
|---|---|---|---|---|---|
| | | | 2000-2004 | | |
| II. | 6.60*** | 1.04*** | -0.09 | | |
| | (1.25) | (0.14) | (0.09) | | |
| IV. | 10.22*** | 0.57** | -0.06 | 0.45 | 0.45 |
| | (1.78) | (0.22) | (0.09) | (0.30) | (0.29) |
| | | | 2005-2009 | | |
| II. | 4.63*** | 0.83*** | -0.06 | | |
| | (0.74) | (0.08) | (0.06) | | |
| IV. | 5.62*** | 0.23 | -0.07 | 0.51* | 0.23 |
| | (0.85) | (0.20) | (0.07) | (0.26) | (0.21) |
| | | | 2010-2014 | | |
| II. | 5.54*** | 0.85*** | -0.12 | | |
| | (1.31) | (0.12) | (0.09) | | |
| IV. | 7.23*** | 0.71*** | -0.13 | -0.08 | 0.41 |
| | (1.37) | (0.21) | (0.10) | (0.25) | (0.21) |

**Table 76: Dynamic Generalized Pareto with negative Jump variation**. The $\xi$ parameter is kept constant, while the $\nu$ parameter is allowed to vary according to the following models:

II. $\nu_t = \exp\left(\kappa_0 + \kappa_2 \log(BV_{t-1}) + \kappa_3 \log(1 + \Delta J^-_{t-1})\right)$

IV. $\nu_t = \exp\left(\kappa_0 + \kappa_2 \log(BV_{t-1}) + \kappa_3 \log(1 + \Delta J^-_{t-1}) + \kappa_4 \log(\overline{RV}^W_{t-1}) + \kappa_5 \log(\overline{RV}^M_{t-1})\right)$

*,**,*** indicate significance at the 5%, 1% and 0.1% levels, respectively.

| | $\kappa_0$ | $\kappa_2$ | $\kappa_3$ | $\kappa_4$ | $\kappa_5$ |
|---|---|---|---|---|---|
| | | | 2000-2004 | | |
| II. | -3.84* | 0.12 | 0.13 | | |
| | (1.42) | (0.16) | (0.12) | | |
| IV. | -3.49 | 0.13 | 0.12 | 0.21 | -0.17 |
| | (2.26) | (0.23) | (0.12) | (0.27) | (0.34) |
| | | | 2005-2009 | | |
| II. | -0.55 | 0.48*** | -0.02 | | |
| | (1.01) | (0.12) | (0.08) | | |
| IV. | -0.19 | 0.09 | 0.001 | 0.49 | -0.05 |
| | (0.96) | (0.18) | (0.08) | (0.28) | (0.21) |
| | | | 2010-2014 | | |
| II. | 2.01* | 0.70*** | -0.17* | | |
| | (1.07) | (0.11) | (0.08) | | |
| IV. | 0.27 | 0.27*** | -0.19*** | 0.55** | -0.32 |
| | (0.51) | (0.08) | (0.05) | (0.11) | (0.07) |

# References


Agrawal, A. and Tandon, K. Anomalies or illusions? evidence from stock markets in eighteen countries. *Journal of International Money and Finance*, 13(1): 83–106, 1994. 6, 9

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453):42–55, 2001. 70, 180

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625, 2003. 72

Andersen, T. G., Bollerslev, T., and Diebold, F. X. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics*, 89(4):701–720, 2007. 73, 93, 154, 186

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. Parametric and nonparametric volatility measurement. *Handbook of Financial Econometrics*, 1:67–138, 2009. 180

Andersen, T. G., Bollerslev, T., and Huang, X. A reduced form framework for modeling volatility of speculative prices based on realized variation measures. *Journal of Econometrics*, 160(1):176–189, 2011. 71, 72

Ariel, R. A. High stock returns before holidays: Existence and evidence on possible causes. *The Journal of Finance*, 45(5):1611–1626, 1990. 5

Axtell, R. L. Zipf distribution of US firm sizes. *Science*, 293(5536):1818–1820, 2001. 1

Baillie, R. T. and Bollerslev, T. The message in daily exchange rates: A conditional-variance tale. *Journal of Business & Economic Statistics*, 20(1):60–68, 2002. 6

Bandi, F. M. and Renò, R. Price and volatility co-jumps. *Journal of Financial Economics (Forthcoming)*, 2015. 67, 73, 83

Barndorff-Nielsen, O. E. and Shephard, N. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B*, 64(2):253–280, 2002. 70, 140, 180

Barndorff-Nielsen, O. E. and Shephard, N. Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2(1):1–37, 2004. 73, 180

Barndorff-Nielsen, O. E. and Shephard, N. *Variation, jumps and high frequency data in financial econometrics*. Econometric Society Monograph. Cambridge University Press: Cambridge, UK, 2007. 153, 180

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536, 2008. 75, 91

Barro, R. J. Rare disasters, Asset prices, and Welfare costs. *American Economic Review*, 99(1):243–264, 2006. 1

Basrak, B., Davis, R. A., and Mikosch, T. Regular variation of GARCH processes. *Stochastic Processes and their Applications*, 99(1):95–115, 2002. 14, 78

Beirlant, J. and Goegebeur, Y. Local polynomial maximum likelihood estimation for pareto-type distributions. *Journal of Multivariate Analysis*, 89(1):97–118, 2004. 179

Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327, 1986. 8, 68, 82, 137, 138

Bollerslev, T. and Ghysels, E. Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics*, 14(2):139–151, 1996. 6, 61

Bollerslev, T. and Todorov, V. Tails, fears, and risk premia. *The Journal of Finance*, 66(6):2165–2211, 2011. 7, 61

Bollerslev, T. and Zhou, H. Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*, 109(1):33–65, 2002. 161

Bollerslev, T., Kretschmer, U., Pigorsch, C., and Tauchen, G. A discrete-time model for daily S&P500 returns and realized variations: Jumps and leverage effects. *Journal of Econometrics*, 150(2):151–166, 2009. 72

Borkovec, M. Extremal behavior of the autoregressive process with ARCH(1) errors. *Stochastic Processes and their applications*, 85(2):189–207, 2000. 159

Bottazzi, G. and Secchi, A. A new class of asymmetric exponential power densities with applications to economics and finance. *Industrial and Corporate Change*, 20:991–1030, 2011. 16

Bouman, S. and Jacobsen, B. The halloween indicator, "sell in may and go away": another puzzle. *American Economic Review*, 92(5):1618–1635, 2002. 6

Breidt, F. J. and Davis, R. A. Extremes of stochastic volatility models. *Annals of Applied Probability*, pages 664–675, 1998. 155

Brownlees, C. T. and Gallo, G. M. Comparison of volatility measures: A risk management perspective. *Journal of Financial Econometrics*, 8(1):29–56, 2010. 67, 71, 76, 85, 153, 157

Brownlees, C. T., Engle, R. F., and Kelly, B. T. A practical guide to volatility forecasting through calm and storm. *Journal of Risk*, 14(2):1–20, 2011. 82

Campbell, J. Y., Grossman, S. J., and Wang, J. Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics*, 108:905–939, 1993. 137

Chan, L. K., Karceski, J., and Lakonishok, J. On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies*, 12(5): 937–974, 1999. 26

Chavez-Demoulin, V., Davison, A. C., and McNeil, A. J. Estimating value-at-risk: A point process approach. *Quantitative Finance*, 5(2):227–234, 2005. 172, 173, 178, 180, 182, 196

Chavez-Demoulin, V., Embrechts, P., and Sardy, S. Extreme-quantile tracking for financial time series. *Journal of Econometrics*, 181(1):44–52, 2014. 172, 173, 178, 180, 182

Choulakian, V. and Stephens, M. Goodness-of-fit tests for the generalized pareto distribution. *Technometrics*, 43(4):478–484, 2001. 29, 76, 193

Christoffersen, P., Heston, S., and Jacobs, K. Option valuation with conditional skewness. *Journal of Econometrics*, 131(1):253–284, 2006. 61

Christoffersen, P. F. Evaluating interval forecasts. *International Economic Review*, 39(4):841–862, 1998. 76, 85, 199

Claessens, S., Dasgupta, S., and Glen, J. Return behavior in emerging stock markets. *The World Bank Economic Review*, 9(1):131–151, 1995. 5, 9

Clark, P. K. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155, 1973. 138

Clements, M. P., Galvão, A. B., and Kim, J. H. Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. *Journal of Empirical Finance*, 15 (4):729–750, 2008. 67, 71

Coles, S. *An Introduction to Statistical Modeling of Extreme Values*. Springer, 2001. 176, 178, 182, 193

Conrad, J. and Kaul, G. Time-variation in expected returns. *Journal of Business*, 61(4):409–425, 1988. 137

Corhay, A., Hawawini, G., and Michel, P. Seasonality in the risk-return relationship: Some international evidence. *The Journal of Finance*, 42(1):49–68, 1987. 9

Corradi, V. and Distaso, W. Semi-parametric comparison of stochastic volatility models using realized measures. *The Review of Economic Studies*, 73(3):635–667, 2006. 140, 161

Corsi, F. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196, 2009. 66, 72, 78, 140, 154, 186

Corsi, F. and Renò, R. Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics*, 30(3):368–380, 2012. 74, 141, 161, 162

Corsi, F., Mittnik, S., Pigorsch, C., and Pigorsch, U. The volatility of realized volatility. *Econometric Reviews*, 27(1-3):46–78, 2008. 158

Corsi, F., Pirino, D., and Reno, R. Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, 159(2):276–288, 2010. 73

Cross, F. The behavior of stock prices on Fridays and Mondays. *Financial Analysts Journal*, 29(6):67–69, 1973. 5

Danielsson, J. and de Vries, C. G. Tail index and quantile estimation with very high frequency data. *Journal of Empirical Finance*, 4(2):241–257, 1997. 65, 172

Davis, R. A. and Mikosch, T. Extreme value theory for GARCH processes. In *Handbook of Financial Time Series*, pages 187–200. Springer, 2009a. 14, 15, 146, 151

Davis, R. A. and Mikosch, T. Extremes of stochastic volatility models. In *Handbook of Financial Time Series*, pages 355–364. Springer, 2009b. 78, 139

Davis, R. A., Mikosch, T., et al. The extremogram: A correlogram for extreme events. *Bernoulli*, 15(4):977–1009, 2009. 139, 145, 146, 147

Davis, R. A., Mikosch, T., and Cribben, I. Towards estimating extremal serial dependence via the bootstrapped extremogram. *Journal of Econometrics*, 170 (1):142–152, 2012. 147, 151, 168

Davis, R. A., Mikosch, T., and Zhao, Y. Measures of serial extremal dependence and their estimation. *Stochastic Processes and their Applications*, 123(7):2575–2602, 2013. 147

Davison, A. C. and Hinkley, D. V. *Bootstrap methods and their application*. Cambridge University Press, 1997. 184

Davison, A. C. and Smith, R. L. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B*, 52(3):393–442, 1990. 3, 65, 174, 178, 181, 182

De Lira Salvatierra, I. A. and Patton, A. J. Dynamic copula models and high frequency data. *Journal of Empirical Finance*, 30(1):120–135, 2015. 173

Diebold, F. X., Schuermann, T., and Stroughair, J. D. Pitfalls and opportunities in the use of extreme value theory in risk management. *The Journal of Risk Finance*, 1(2):30–35, 2000. 172

Dittmar, R. F. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *The Journal of Finance*, 57(1):369–403, 2002. 56

Drees, H. Weighted approximations of tail processes for $\beta$-mixing random variables. *Annals of Applied Probability*, 10(4):1274–1301, 2000. 177

Dupuis, D. J. Modeling waves of extreme temperature: the changing tails of four cities. *Journal of the American Statistical Association*, 107(497):24–39, 2012. 2

Dupuis, D. J., Sun, Y., and Wang, H. J. Detecting change-points in extremes. *Statistics and Its Interface*, 8(1):19–31, 2015. 12, 13, 15

Efron, B. and Tibshirani, R. J. *An introduction to the bootstrap*. CRC Press, 1994. 85

Embrechts, P., Klüppelberg, C., and Mikosch, T. *Modelling extremal events: for insurance and finance*. Springer, 1997. 1, 2, 6, 10, 15, 143

Embrechts, P., Resnick, S. I., and Samorodnitsky, G. Extreme value theory as a risk management tool. *North American Actuarial Journal*, 3(2):30–41, 1999. 6

Engle, R. New frontiers for arch models. *Journal of Applied Econometrics*, 17(5): 425–446, 2002. 157

Engle, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007, 1982. 137

Engle, R. F. and Gallo, G. M. A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, 131(1):3–27, 2006. 140, 153, 157

Engle, R. F. and Manganelli, S. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381, 2004. 76, 85, 199

Fama, E. F. The behavior of stock-market prices. *Journal of Business*, 38(1):34–105, 1965. 1

Fama, E. F. Efficient capital markets: II. *The Journal of Finance*, 46(5):1575–1617, 1991. 6

Fasen, V. Extremes of continuoustime processes. In Mikosch, T., Krei, J.-P., Davis, R. A., and Andersen, T. G., editors, *Handbook of Financial Time Series*, pages 653–667. Springer Berlin Heidelberg, 2009. 140, 161

Fasen, V., Klppelberg, C., and Lindner, A. Extremal behavior of stochastic volatility models. In Shiryaev, A., Grossinho, M., Oliveira, P., and Esquvel, M., editors, *Stochastic Finance*, pages 107–155. Springer US, 2006. 78

Ferro, C. A. and Segers, J. Inference for clusters of extreme values. *Journal of the Royal Statistical Society: Series B*, 65(2):545–556, 2003. 15, 29, 77, 139, 144, 149

Fisher, R. A. and Tippett, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge Univ Press, 1928. 2

French, K. R. Stock returns and the weekend effect. *Journal of Financial Economics*, 8(1):55–69, 1980. 5

French, K. R. and Roll, R. Stock return variances: The arrival of information and the reaction of traders. *Journal of Financial Economics*, 17(1):5–26, 1986. 6

Gabaix, X. Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, 114(3):739–767, 1999. 1

Gabaix, X. Power laws in economics and finance. *Annual Review of Economics*, 1: 255–293, 2009. 1

Gallant, A. R., Rossi, P. E., and Tauchen, G. Stock prices and volume. *Review of Financial Studies*, 5(2):199–242, 1992. 6, 9

Giot, P. and Laurent, S. Modelling daily value-at-risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, 11(3):379–398, 2004. 67, 71, 85

Glosten, L. R., Jagannathan, R., and Runkle, D. E. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801, 1993. 6, 8, 140, 156, 157

Gultekin, M. N. and Gultekin, N. B. Stock market seasonality: International evidence. *Journal of Financial Economics*, 12(4):469–481, 1983. 5, 9

Hall, P. and Tajvidi, N. Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, 15(2):153–167, 2000. 179

Hansen, B. E. Autoregressive conditional density estimation. *International Economic Review*, 35(3):705–730, 1994. 173

Hansen, P. and Lunde, A. *Forecasting volatility using high frequency data*. Oxford University Press, 2011. 140

Hansen, P. R. and Lunde, A. A forecast comparison of volatility models: Does anything beat a GARCH (1, 1)? *Journal of Applied Econometrics*, 20(7):873–889, 2005. 27

Hansen, P. R., Huang, Z., and Shek, H. H. Realized garch: a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27 (6):877–906, 2012. 140, 153, 173

Hansen, P. R., Lunde, A., and Voev, V. Realized beta garch: A multivariate garch model with realized measures of volatility. *Journal of Applied Econometrics*, 29 (5):774–799, 2014. 173

Harvey, A. C. and Shephard, N. Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, 14(4):429–434, 1996. 141, 162

Harvey, C. R. and Siddique, A. Autoregressive conditional skewness. *Journal of financial and quantitative analysis*, 34(04):465–487, 1999. 56, 173, 178

Harvey, C. R. and Siddique, A. Conditional skewness in asset pricing tests. *The Journal of Finance*, 55(3):1263–1295, 2000. 56

Hawawini, G. and Keim, D. B. On the predictability of common stock returns: World-wide evidence. *Handbooks in Operations Research and Management Science*, 9:497–544, 1995. 6

Heber, G., Lunde, A., Shephard, N., and Sheppard, K. Oxford-Man Institute realized library, version 0.2. Oxford-Man Institute, University of Oxford, 2009. 75, 148, 184

Heston, S. L. and Sadka, R. Seasonality in the cross-section of stock returns. *Journal of Financial Economics*, 87(2):418–445, 2008. 6

Heston, S. L. and Sadka, R. Seasonality in the cross section of stock returns: The international evidence. *Journal of Financial and Quantitative Analysis*, 45:1133–1160, 2010. 6

Hill, B. M. A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3(5):1163–1174, 1975. 12, 145

Hosking, J. R. and Wallis, J. R. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3):339–349, 1987. 177

Hosmer, D. W. and Lemeshow, S. *Applied logistic regression*. John Wiley & Sons, 2004. 191

Hsing, T., Hüsler, J., and Leadbetter, M. On the exceedance point process for a stationary sequence. *Probability Theory and Related Fields*, 78(1):97–112, 1988. 143

Huang, X. and Tauchen, G. The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3(4):456–499, 2005. 83

Hüsler, J. Extreme values of non-stationary random sequences. *Journal of Applied Probability*, 23(4):937–950, 1986. 3, 179

Jaffe, J. and Westerfield, R. The week-end effect in common stock returns: The international evidence. *The Journal of Finance*, 40(2):433–454, 1985. 5

Jalal, A. and Rockinger, M. Predicting tail-related risk measures: The consequences of using GARCH filters for non-GARCH data. *Journal of Empirical Finance*, 15(5):868–877, 2008. xii, 6, 66, 79, 82, 84

Jondeau, E. and Rockinger, M. Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance*, 10(5):559–581, 2003. 6

Kamstra, M. J., Kramer, L. A., and Levi, M. D. A careful re-examination of seasonality in international stock markets: Comment on sentiment and stock returns. *Journal of Banking & Finance*, 36(4):934–956, 2012. 6

Keim, D. B. Size-related anomalies and stock return seasonality: Further empirical evidence. *Journal of Financial Economics*, 12(1):13–32, 1983. 5, 9

Kelly, B. and Jiang, H. Tail risk and asset prices. *Review of Financial Studies*, 10 (27):2841–2871, 2014. 8

Kesten, H. Random difference equations and renewal theory for products of random matrices. *Acta Mathematica*, 131(1):207–248, 1973. 14, 157, 158

Kim, C.-W. and Park, J. Holiday effects and stock returns: Further evidence. *Journal of Financial and Quantitative Analysis*, 29(01):145–157, 1994. 5

Kim, M. and Lee, S. Test for tail index change in stationary time series with Pareto-type marginal distribution. *Bernoulli*, 15(2):325–356, 2009. 7, 10, 12

Lakonishok, J. and Maberly, E. The weekend effect: Trading patterns of individual and institutional investors. *The Journal of Finance*, 45(1):231–243, 1990. 5

Leadbetter, M. R., Lindgren, G., and Rootzén, H. *Extremes and related properties of random sequences and processes*. Springer, New York, 1983. 3, 143

Leadbetter, M. Extremes and local dependence in stationary sequences. *Probability Theory and Related Fields*, 65(2):291–306, 1983. 15, 76

LeBaron, B. Some relations between volatility and serial correlations in stock market returns. *Journal of Business*, (2):199–219, 1992. 137

Ledford, A. W. and Tawn, J. A. Diagnostics for dependence within time series extremes. *Journal of the Royal Statistical Society: Series B*, 65(2):521–543, 2003. 139, 143, 144

Ling, S. and McAleer, M. Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics*, 106(1):109–117, 2002. 9

Liu, L. Y., Patton, A. J., and Sheppard, K. Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. *Journal of Econometrics*, 187(1):293–311, 2015. 75

Liu, Y. and Tawn, J. A. Volatility model selection for extremes of financial time series. *Journal of Statistical Planning and Inference*, 143(3):520–530, 2013. 78, 139, 147, 149, 165, 166

Lo, A. W. and MacKinlay, A. C. Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1):41–66, 1988. 137

Longin, F. M. From value at risk to stress testing: The extreme value approach. *Journal of Banking & Finance*, 24(7):1097–1130, 2000. 172

Maheu, J. M. and McCurdy, T. H. Do high-frequency measures of volatility improve forecasts of return distributions? *Journal of Econometrics*, 160(1):69–76, 2011. 156

Maheu, J. M., McCurdy, T. H., and Zhao, X. Do jumps contribute to the dynamics of the equity premium? *Journal of Financial Economics*, 110(2):457–477, 2013. 61

Mandelbrot, B. The variation of certain speculative prices. *Journal of Business*, 36 (4):394–419, 1963. 1

McFarland, J. W., Pettit, R. R., and Sung, S. K. The distribution of foreign exchange price changes: Trading day effects and risk measurement. *The Journal of Finance*, 37(3):693–715, 1982. 1

McNeil, A. J. Extreme value theory for risk managers. In *Internal Modeling and CAD II*, pages 93–113. RiskBooks, 1999. 6

McNeil, A. J. and Frey, R. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, 7(3):271–300, 2000. 2, 3, 4, 6, 65, 67, 68, 69, 76, 82, 172, 177, 198, 199

McNeil, A. J., Frey, R., and Embrechts, P. *Quantitative risk management: concepts, techniques, and tools,*. Princeton University Press, 2005. 7

Meitz, M. and Saikkonen, P. Parameter estimation in nonlinear AR–GARCH models. *Econometric Theory*, 27(6):1236–1278, 2011. 11

Merton, R. C. On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics*, 8(4):323–361, 1980. 26

Mikosch, T. and Rezapur, M. Stochastic volatility models with possible extremal clustering. *Bernoulli*, 19(5A):1688–1713, 2013. 139, 152, 154, 157, 158, 159

Mikosch, T. and Starica, C. Limit theory for the sample autocorrelations and extremes of a GARCH (1, 1) process. *Annals of Statistics*, 28(5):1427–1451, 2000. 14, 139

Mikosch, T. and Starica, C. Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *Review of Economics and Statistics*, 86(1):378–390, 2004. 26, 173, 178

Nelson, D. B. Arch models as diffusion approximations. *Journal of econometrics*, 45(1):7–38, 1990. 141, 161

Nelson, D. B. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 59(2):347–370, 1991. 11, 73, 82

Nelson, D. B. Filtering and forecasting with misspecified arch models i: Getting the right variance with the wrong model. *Journal of Econometrics*, 52(1):61–90, 1992. 72

Noureldin, D., Shephard, N., and Sheppard, K. Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 27(6):907–933, 2012. 173

Officer, R. R. Seasonality in Australian capital markets: Market efficiency and empirical issues. *Journal of Financial Economics*, 2(1):29–51, 1975. 5

Ogata, Y. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401):9–27, 1988. 192

Oh, D. H. and Patton, A. J. High dimension copula-based distributions with mixed frequency data. *Journal of Econometrics (Forthcoming)*, 2015. 173

Pan, J. Stochastic volatility with reset at jumps. *Available at SSRN 135048*, 1997. 82

Patton, A. J. and Sheppard, K. Good volatility, bad volatility: Signed jumps and the persistence of volatility. *Review of Economics and Statistics*, 97(3):683–697, 2015. 73, 74, 93, 94, 201

Pickands, J. Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1):119–131, 1975. 10, 29, 68, 76, 175

Roll, R. Vas ist das? *The Journal of Portfolio Management*, 9(2):18–28, 1983. 5

Rozeff, M. S. and Kinney, W. J. Capital market seasonality: The case of stock returns. *Journal of Financial Economics*, 3(4):379–402, 1976. 5, 7, 9

Scarrott, C. and MacDonald, A. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT*, 10(1):33–60, 2012. 29

Schaller, H. and Norden, S. V. Regime switching in stock market returns. *Applied Financial Economics*, 7(2):177–191, 1997. 82

Schwarz, G. Estimating the dimension of a model. *Annals of Statistics*, 6(2):461–464, 1978. 11, 195

Shephard, N. and Sheppard, K. Realising the future: Forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 25(2):197–231, 2010. 66, 71, 140, 153, 158, 173

Smith, R. L. Maximum likelihood estimation in a class of non regular cases. *Biometrika*, 72(1):67–90, 1985. 177

Smith, R. L. Extreme value theory based on the r largest annual events. *Journal of Hydrology*, 86(1):27–43, 1986. 2

Smith, R. L. Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statistical Science*, pages 367–377, 1989. 2

Taylor, S. J. Financial returns modelled by the product of two stochastic processes. A study of daily sugar prices 1961-79. In *Time Series Analysis: Theory and Practice I*, pages 203–226. New-Holland, 1982. 166

Todorov, V. and Tauchen, G. Volatility jumps. *Journal of Business & Economic Statistics*, 29(3):356–371, 2011. 73

Wang, H. and Tsai, C.-L. Tail index regression. *Journal of the American Statistical Association*, 104(487):1233–1240, 2009. 179

Wang, H. J., Li, D., and He, X. Estimation of high conditional quantiles for heavy-tailed distributions. *Journal of the American Statistical Association*, 107 (500):1453–1464, 2012. 202

Wood, R. A., McInish, T. H., and Ord, J. K. An investigation of transactions data for NYSE stocks. *The Journal of Finance*, 40(3):723–739, 1985. 6

Zakoian, J.-M. Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, 18(5):931–955, 1994. 11

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. A tale of two time scales. *Journal of the American Statistical Association*, 100(472):1394–1411, 2005. 75, 91