Journal of Empirical Finance 36 (2016) 86–99

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

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

Marco Bee [a,*], Debbie J. Dupuis [b,1], Luca Trapin [c]

[a] *Department of Economics and Management, University of Trento, via Inama, 5-38122 Trento, Italy*
[b] *Department of Decision Sciences, HEC Montreal, 3000, chemin de la Côte-Sainte-Catherine, Montréal, Québec H3T 2A7, Canada*
[c] *IMT Institute for Advanced Studies Lucca, Italy*

ARTICLE INFO

Article history:
Received 17 May 2015
Received in revised form 30 November 2015
Accepted 7 January 2016
Available online 14 January 2016

JEL classification:
C4
C5
G1

Keywords:
Realized volatility
High-frequency data
Extreme Value Theory
Value-at-Risk
Expected Shortfall

ABSTRACT

This article applies realized volatility forecasting to Extreme Value Theory (EVT). We propose a two-step approach where returns are first pre-whitened with a high-frequency based volatility model, and then an EVT based model is fitted to the tails of the standardized residuals. This realized EVT approach is compared to the conditional EVT of McNeil & Frey (2000). We assess both approaches' ability to filter the dependence in the extremes and to produce stable out-of-sample VaR and ES estimates for one-day and ten-day time horizons. The main finding is that GARCH-type models perform well in filtering the dependence, while the realized EVT approach seems preferable in forecasting, especially at longer time horizons.

1. Introduction

Accurate assessment of the tail behavior 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 behavior 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.

* Corresponding author. Tel.: +39 0461 282 296; fax: +39 0461 282 241.
E-mail addresses: marco.bee@unitn.it (M. Bee), debbie.dupuis@hec.ca (D.J. Dupuis), luca.trapin@imtlucca.it (L. Trapin).
[1] Tel.: +1 514 340 6952; fax: +1 514 340 5634.

http://dx.doi.org/10.1016/j.jempfin.2016.01.006
0927-5398/

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.[2] This period is relevant from the regulatory perspective as the risk capital of a bank 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 (2016). 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 (2009), 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 paper is organized as follows: in Section 2, we present the C-EVT of McNeil and Frey (2000); in Section 3 we introduce the RV-EVT; in Section 4 we compare the two approaches, looking separately at the filtering and forecasting components; in Section 5 we perform robustness checks aimed at consolidating the evidence from the main analysis; in Section 6 we give the concluding remarks. Some technical details appear in the Appendix and further data analyses can be found in the Supplementary material.

2. The conditional EVT approach

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

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

where μ_t and σ_t^2, are respectively the conditional mean and variance, functions of the information sets $\mathcal{F}_{t-1}$ and $\mathcal{H}_{t-1}$, and ϵ_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 σ_t^2 as a function of its past values and the past values of ϵ_t, assuming ϵ_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 $\mathbb{E}(\epsilon_t^4) < \infty$. Formally, they consider an AR(1)-GARCH(1,1),

$$r_t = \mu + \phi_1 r_{t-1} + \sigma_t \epsilon_t \tag{2}$$

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{3}$$

where $\epsilon_t \sim F$ with zero mean and unit variance, and $\alpha_1 + \beta_1 < 1$ to guarantee stationarity. Suppose[3] 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 ξ and scale parameter $\nu > 0$. That is, $\Pr(\epsilon_t - u \leq x | \epsilon_t > u)$ goes to

$$G(x; \xi, \nu) = \begin{cases} 1 - \{1 + \xi x/\nu\}^{-\frac{1}{\xi}} & \text{for} \quad \xi \neq 0 \\ 1 - \exp\{-x/\nu\} & \text{for} \quad \xi = 0 \end{cases} \tag{4}$$

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

[2] 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.

[3] The following argument equally applies to ϵ_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*.