

Robust Value at Risk Prediction

Loriano Mancini, Fabio Trojani

September 2007 Discussion Paper no. 36

Editor:

Prof. Jörg Baumberger
University of St. Gallen
Department of Economics
Bodanstr. 1
CH-9000 St. Gallen
Phone +41 71 224 22 41
Fax +41 71 224 28 85
Email joerg.baumberger@unisg.ch

Publisher:

Department of Economics
University of St. Gallen
Bodanstrasse 8
CH-9000 St. Gallen
Phone +41 71 224 23 25
Fax +41 71 224 22 98

Electronic Publication:

<http://www.vwa.unisg.ch>

Robust Value at Risk Prediction*

Loriano Mancini, Fabio Trojani

Author's address:

Prof. Dr. Fabio Trojani
Institut für Banken und Finanzen
Rosenbergstrasse 52
9000 St. Gallen
Tel. +41 71 2247074
Fax +41 71 2247088
Email fabio.trojani@unisg.ch
Website www.sbf.unisg.ch

* Correspondence Information: Loriano Mancini, Swiss Banking Institute, University of Zurich, Plattenstrasse 14, 8032 Zurich, Switzerland, phone: +41(0)44 634 2942, fax: +41(0)44 634 4997, E-mail: mancini@isb.uzh.ch. E-mail address for Fabio Trojani: fabio.trojani@unisg.ch. We gratefully acknowledge the financial support of the Swiss National Science Foundation (NCCR-FinRisk and Grants Nr. 101312-103781/1 and 100012-105745/1) and the University Research Priority Program "Finance and Financial Markets" University of Zurich. We thank Claudio Ortelli and Claudia Ravanelli for many valuable suggestions on an earlier draft. Much of this paper was written while Mancini visited the Department of Operations Research and Financial Engineering, Princeton University, whose hospitality is gratefully acknowledged. Finally, we would like to remember our Professor and friend Pietro Balestra, who had taught and transmitted us with incomparable magistry and enthusiasm the beauty of Econometrics.

Abstract

We propose a general robust semiparametric bootstrap method to estimate conditional predictive distributions of GARCH-type models. Our approach is based on a robust estimator for the parameters in GARCH-type models and a robustified resampling method for standardized GARCH residuals, which controls the bootstrap instability due to influential observations in the tails of standardized GARCH residuals. Monte Carlo simulation shows that our method consistently provides lower VaR forecast errors, often to a large extent, and in contrast to classical methods never fails validation tests at usual significance levels. We test extensively our approach in the context of real data applications to VaR prediction for market risk, and find that only our robust procedure passes all validation tests at usual confidence levels. Moreover, the smaller tail estimation risk of robust VaR forecasts implies VaR prediction intervals that can be nearly 20% narrower and 50% less volatile over time. This is a further desirable property of our method, which allows to adapt risky positions to VaR limits more smoothly and thus more efficiently.

Keywords

Backtesting, M-estimator, Extreme Value Theory, Breakdown Point.

JEL Classification

C14, C15, C23, C59.

1 Introduction

Large portfolios of traded assets held by many financial institutions have made the measurement of market risk, i.e. the risk of losses on the trading book due to adverse market movements, a primary concern for regulators and internal risk managers. The Basel Committee (1996) requires that financial institutions hold a certain amount of capital against market risk. This capital is called Value at Risk (VaR) and must be sufficient to cover losses on the trading book over a ten days holding period 99% of the times. In practice, VaR measures are computed for several holding periods and confidence levels. For internal purposes, for instance, most banks use VaR at a 95% confidence level and a horizon of one day. From a statistical viewpoint, VaR is the quantile of the profit and loss (P&L) distribution of a portfolio over a certain holding period. Hence, a key issue in implementing VaR and related risk measures is to obtain accurate estimates for the tails of the conditional P&L distribution at the relevant horizons. In the financial literature two main directions have been followed to estimate P&L conditional distributions for market risk management: fully nonparametric historical simulation methods and semiparametric/nonparametric bootstrap methods based on dynamic models for asset returns; see Duffie and Pan (1997) for an overview. Fully nonparametric methods are easy to implement, but unfortunately do not provide accurate VaR predictions. Semiparametric methods have been found to work well and are commonly called Filtered Historical Simulation (FHS) methods; see, among others, Pritsker (1997), Hull and White (1998), Diebold, Schuermann, and Stroughair (1998), Barone-Adesi, Giannopoulos, and Vosper (1999), McNeil and Frey (2000), Pritsker (2001), and Kuester, Mittnik, and Paolella (2006). In these papers, parametric GARCH-type models for returns are fitted using pseudo maximum likelihood (PML). Then standardized GARCH residuals are resampled using either (i) a fully nonparametric bootstrap or (ii) a semiparametric bootstrap based on PML estimations of generalized Pareto distributions for the tails of GARCH residuals. Therefore, semiparametric methods allow for time varying conditional moments of returns (via parametric GARCH-type models) and non-

parametric structures in conditional distribution of returns (as GARCH residual distributions are estimated nonparametrically). The last feature is crucial in applications and avoids too simplistic assumptions on return conditional distributions, such as normality; see for instance JP Morgan’s RiskMetrics (1995). In contrast to a fully nonparametric method, the semiparametric approach resamples approximately iid standardized GARCH residuals without relying on resampling procedures for non iid data and the necessary assumptions for those to hold; see for instance Künsch (1989).¹

In this paper we propose a class of robust semiparametric bootstrap methods to estimate conditional predictive distributions of asset returns in GARCH-type volatility models. This approach can account for general parametric specifications of time varying conditional mean and variance of asset returns. As an application, we use the proposed robust method to compute Value at Risk predictions over different forecasting horizons. Our robust approach combines a robust estimator for parametric GARCH-type volatility dynamics and a robustified resampling method for standardized GARCH residuals. Hence, robustness is achieved in two steps. In the first step, we estimate a parametric GARCH-type model using the optimal bounded influence estimator in Mancini, Ronchetti, and Trojani (2005). In the second step, we fit the robust generalized Pareto density estimator in Dupuis (1999) and Juárez and Schucany (2004) to the tails of the GARCH residuals distribution and we resample from this distribution. In order to ensure robustness of the whole bootstrap procedure, both robustification steps are necessary. It is well-known that outliers

¹A potentially alternative way to estimate risk measures for P&L distributions is to apply directly statistical tools designed to estimate regression quantiles and conditional distributions; see for instance Koenker and Bassett (1978), Foresi and Peracchi (1995), Peracchi (2002), and Engle and Manganelli (2004). From a robustness perspective, drawbacks of regression quantiles are their behavior under heteroscedasticity and the non robustness to “bad” leverage points; see Koenker and Bassett (1982). Moreover, regression quantile methods are naturally applied to one day ahead VaR predictions, as illustrated in Engle and Manganelli (2004). Their efficient extension to the estimation of risk measures for longer horizons is, however, not straightforward.

or influential points in the data can strongly bias and highly inflate PML parameter estimates of GARCH-type models; see for instance Sakata and White (1998) and Mancini et al. (2005). In our semiparametric bootstrap context, this feature can induce inaccurately estimated GARCH residuals and unreliable bootstrap distributions. In addition, standard bootstrap procedures suffer from an intrinsic robustness problem, especially when bootstrapping fat tailed distributions. A small number of outliers in the data can cause the breakdown of quantile estimates based on nonparametric bootstrap distributions of residuals; see for instance Singh (1998), Davidson and Flachaire (2005), Gagliardini, Trojani, and Urga (2005) and Camponovo, Scaillet, and Trojani (2006). The robust extreme value estimator in our approach controls the bootstrap instability deriving from outliers or influential points in the tails of GARCH residual distributions. We show that the standard nonparametric bootstrap procedure has a low breakdown point, meaning that VaR forecasts can be heavily affected by a few influential points, especially when long forecast horizons are considered. Robustness can be enhanced by fitting a generalized Pareto distribution to the tails of the residual distribution and sampling tail residuals from this density. Sampling from a parametric density can circumvent the robustness problem of nonparametric bootstrap procedures. However, to ensure a sufficiently large breakdown point for the estimator of the generalized Pareto tails, a robust extreme value estimator is needed. Such an estimator is not subject to the instability problem of PML estimators for generalized Pareto distributions as noted, among others, in Cowell and Victoria-Feser (1996).

Extreme or outlying observations—due for instance to abnormal liquidity conditions or market crashes—are a fundamental component of data generating processes of financial returns and, in particular, of the riskiness we aim to measure. Therefore, they cannot be simply disregarded when estimating measures of market risk. However, such observations are typically caused by relatively rare market conditions that are very difficult to anticipate using historical return information and almost impossible to model conveniently in a simple parametric model for asset returns, such

as a GARCH-type model. To avoid biased GARCH estimates that are not representative for the typical volatility dynamics in the data, we estimate a parametric GARCH model with the bounded influence estimator proposed in Mancini et al. (2005). To ensure the stability of the GARCH residual bootstrap procedure we fit the robust generalized Pareto estimator in Dupuis (1999) and Juárez and Schucany (2004) to the tails of the GARCH residual distribution.

We study the accuracy of VaR predictions at 5% and 1% confidence levels for one day and ten days ahead forecast horizons by performing Monte Carlo simulation and in real data applications. Monte Carlo simulation shows that compared to the non robust counterparts our robust procedure suffers very moderate efficiency losses, when estimating a GARCH model with conditionally Gaussian innovations. At the same time, it offers substantial improvements in accuracy under several forms of departure from conditional normality, which are likely to be present in financial data. Procedures based on classical methods imply less reliable VaR predictions under conditionally non Gaussian data, especially for several days ahead horizons. In particular, in the presence of outlying observations, robust VaR predictions at 1% confidence level have mean square prediction errors several times smaller than those of classical procedures. In nearly all Monte Carlo experiments and for all VaR confidence levels and forecast horizons, our robust bootstrap procedure has the lowest mean square prediction errors, often by a large extent. Moreover, in contrast to classical methods our procedure never fails validation tests based on VaR violations at 10% level.

The simulation evidence is confirmed by the real data application. We backtest VaR prediction methods using about twenty years of S&P 500, Dollar-Yen, Microsoft and Boeing historical returns. Overall, robust VaR predictions are found to be more accurate than classical ones: in our backtesting exercises the robust VaR forecasting method is the only one that passes all validation tests at a 10% significance level, both for all VaR confidence levels and all forecast horizons. Moreover, the robust procedure implies a smaller tail estimation risk, which is reflected by robust VaR prediction intervals that are more accurate and more stable over time. For instance, in the

case of S&P 500 and Boeing, robust VaR prediction intervals are nearly 20% narrower and 50% less volatile than classical ones. In risk management, the stability over time of VaR predictions is a desirable feature, because financial institutions cannot adjust rapidly the capital base. At the same time, stable VaR profiles over time allow to adapt outstanding portfolio risk exposures to VaR limits more smoothly and thus more efficiently. Therefore, the time stability of VaR predictions estimated by robust semiparametric bootstrap methods is a further desirable property of our approach.

Section 2 introduces semiparametric bootstrap and extreme value estimation methods for VaR predictions, along with their robust versions. Section 3 presents Monte Carlo simulation comparing the performance of classical and robust, nonparametric and semiparametric, bootstrap VaR prediction methods, under different forms of conditional non normality of returns. Section 4 presents the real data application to VaR prediction and backtesting for four financial time series. Section 5 concludes.

2 Setting

This section introduces the different semiparametric bootstrap procedures for GARCH-type models relevant for our analysis.

2.1 Return Dynamics and Measures of Market Risk

Let $\mathcal{Y} := (Y_t)_{t \in \mathbb{Z}}$ be a strictly stationary time series process on probability space $(\mathbb{R}^\infty, \mathcal{F}, \mathbb{P}_*)$, modelling the daily rate of return on a financial asset with price P_t at time t , i.e. $Y_t := P_t/P_{t-1} - 1$. We assume that \mathbb{P}_* can be “approximated” by some parametric model $\mathcal{P} := \{\mathbb{P}_\theta, \theta \in \Theta \subseteq \mathbb{R}^p\}$. Precisely, $\mathbb{P}_* \in \mathcal{U}(\mathbb{P}_{\theta_0})$ for some $\theta_0 \in \Theta$, where $\mathcal{U}(\mathbb{G})$ denotes a nonparametric neighborhood of

distribution \mathbb{G} . Under \mathbb{P}_{θ_0} , process \mathcal{Y} satisfies the dynamic model

$$Y_t = \mu_t(\theta_0) + \sigma_t(\theta_0) Z_t, \quad (1)$$

where $\mu_t(\theta_0)$ and $\sigma_t^2(\theta_0)$ parameterize the conditional mean and conditional variance of Y_t , given information \mathcal{F}_{t-1} up to time $t - 1$. Under probability \mathbb{P}_{θ_0} , innovations Z_t 's are a strong white noise, i.e. $Z_t \sim IIN(0, 1)$. We denote by F the distribution function of Z_t under \mathbb{P}_* . Given $Y_1^m := (Y_1, \dots, Y_m)$, denote by \mathbb{P}_*^m ($\mathbb{P}_{\theta_0}^m$) the m -dimensional marginal distribution of Y_1^m under \mathbb{P}_* (\mathbb{P}_{θ_0}). $F_{t,t+h}$ is the conditional distribution function of h days returns $Y_{t,t+h} := P_{t+h}/P_t - 1$ under \mathbb{P}_* , given information \mathcal{F}_t . For $0 < \alpha < 1$ and horizon h days, lower quantile $y_{t,t+h}^\alpha$ of $F_{t,t+h}$ is defined by

$$y_{t,t+h}^\alpha := \inf\{y \in \mathbb{R} : F_{t,t+h}(y) \geq \alpha\}.$$

The Value at Risk (VaR) at time t of an institution investing in a portfolio of assets with market price P_t is $VaR_{t,t+h}^\alpha = P_t + R_{t,t+h}^\alpha$, where $R_{t,t+h}^\alpha$ is the reserve amount such that the probability of a loss over the next h days is not above some level α :²

$$\alpha = \mathbb{P}_*(P_{t+h} + R_{t,t+h}^\alpha < 0 | \mathcal{F}_t) = \mathbb{P}_*(P_{t+h} - P_t < -VaR_{t,t+h}^\alpha | \mathcal{F}_t). \quad (2)$$

In other words, $-VaR_{t,t+h}^\alpha = P_t y_{t,t+h}^\alpha$ is the α -quantile of the conditional profit and loss (P&L) distribution under \mathbb{P}_* over the next h days, given \mathcal{F}_t . In contrast to historical simulation techniques, $VaR_{t,t+h}^\alpha$ reflects current available information such as high or low volatility periods, which is important to achieve accurate VaR forecasts.³ Another measure of market risk is the Expected Shortfall (ES) $S_{t,t+h}^\alpha$ (Artzner, Delbaen, Eber, and Heath (1999)), defined by

$$S_{t,t+h}^\alpha := E_*[Y_{t,t+h} | Y_{t,t+h} < y_{t,t+h}^\alpha, \mathcal{F}_t],$$

where $E_*[\cdot]$ denotes expectation with respect to \mathbb{P}_* .⁴ For horizon $h = 1$ day,

$$y_{t,t+1}^\alpha = \mu_{t+1}(\theta_0) + \sigma_{t+1}(\theta_0) z_\alpha, \quad S_{t,t+1}^\alpha = \mu_{t+1}(\theta_0) + \sigma_{t+1}(\theta_0) E_*[Z | Z < z_\alpha],$$

²To simplify the exposition, equation (2) is based on a continuous profit and loss distribution under \mathbb{P}_* .

³See for instance Gouriéroux, Laurent, and Scaillet (2000) for a general discussion on conditional VaR, and

Aït-Sahalia and Lo (2000) for an economic interpretation of VaR in a general equilibrium model.

⁴See for instance Scaillet (2004) for nonparametric estimation of Expected Shortfall.

where z_α is the α -quantile of the distribution of Z_t . Hence estimation of VaR or ES can be obtained by (i) estimating model (1) for daily conditional moment dynamics and (ii) estimating the tail distribution of daily standardized residual Z_t . For longer horizons $h \geq 2$, estimated daily dynamics are only the starting point to estimate $F_{t,t+h}$. Joint conditional distributions of $(Z_{t+1}, \dots, Z_{t+h})$, $(\mu_{t+1}, \dots, \mu_{t+h})$ and $(\sigma_{t+1}, \dots, \sigma_{t+h})$ have to be estimated, a task which is considerably more difficult. Filtering Historical Simulation (FHS) methods estimate $F_{t,t+h}$ by applying a semiparametric bootstrap to dynamic model (1) over horizon $[t, t+h]$.

Our goal is to develop robust semiparametric bootstrap methods for estimating $F_{t,t+h}$. In this setup, underlying distribution \mathbb{P}_* is unknown and belongs to some nonparametric neighborhood of parametric reference model \mathbb{P}_{θ_0} . Hence \mathbb{P}_{θ_0} has to be regarded as an “approximate” description of the true data generating process \mathbb{P}_* .

2.2 Estimations of GARCH-type Dynamics

We use pseudo maximum likelihood (PML, Gouriéroux, Monfort, and Trognon (1984)) estimators for the parameters of dynamic model (1) and bounded influence, conditionally unbiased, optimal versions of such PML estimators (Mancini, Ronchetti, and Trojani (2005)).

2.2.1 Pseudo Maximum Likelihood Estimation

Typically, model (1) is estimated by a PML approach, under the nominal assumption of Gaussian innovations. The functional PML estimator (PMLE) $a(\cdot)$ is implicitly defined by the asymptotic estimating equation

$$E_*[s(Y_1^m; a(\mathbb{P}_*^m))] = 0,$$

where

$$s(Y_1^m; \theta) = \frac{1}{\sigma_m^2(\theta)} \frac{\partial \mu_m(\theta)}{\partial \theta} \varepsilon_m(\theta) + \frac{1}{2\sigma_m^2(\theta)} \frac{\partial \sigma_m^2(\theta)}{\partial \theta} \left(\frac{\varepsilon_m^2(\theta)}{\sigma_m^2(\theta)} - 1 \right) \quad (3)$$

and $\varepsilon_m(\theta) := \sigma_m(\theta)Z_m$. Under model \mathbb{P}_{θ_0} , PMLE and MLE coincide, implying an asymptotically optimal covariance matrix

$$V(s; \theta_0) = E_{\theta_0}[s(Y_1^m; \theta_0) s(Y_1^m; \theta_0)^\top]^{-1} =: I(\theta_0)^{-1},$$

where $I(\theta_0)$ is the information matrix. However, if \mathbb{P}_{θ_0} is slightly different from the true data generating process for \mathcal{Y} , PML estimators can induce highly biased and inefficient inference results on θ_0 ; see for instance Sakata and White (1998) and Mancini et al. (2005). In Section 3, we investigate by Monte Carlo simulation how such estimates affect Value at Risk forecasts under several realistic forms of local departures from conditional normality in model (1).

2.2.2 Optimal Robust Estimation

As a robust estimator for conditional moment dynamics in model (1) we use the optimal conditionally unbiased bounded-influence estimator for GARCH-type models developed in Mancini et al. (2005). Such a robust estimator is efficient and computationally feasible for highly nonlinear models, and compares therefore favorably with other robust estimators, such as robust GMM estimators (Ronchetti and Trojani (2001)) or robust EMM estimators (Ortelli and Trojani (2005)) for time series. It is defined as follows. Let

$$\psi_c(s(Y_1^m; \theta)) := A(\theta) (s(Y_1^m; \theta) - \tau(Y_1^{m-1}; \theta)) w(Y_1^m; \theta),$$

where $s(Y_1^m; \theta)$ is the PML score function defined in (3) and the weighting function

$$w(Y_1^m; \theta) := \min(1, c \|A(\theta) (s(Y_1^m; \theta) - \tau(Y_1^{m-1}; \theta))\|^{-1}).$$

Define a robust functional M -estimator $\bar{a}(\cdot)$ of θ implicitly by

$$E_*[\psi_c(s(Y_1^m; \bar{a}(\mathbb{P}_*^m)))] = 0, \tag{4}$$

where non singular matrix $A(\theta) \in \mathbb{R}^p \times \mathbb{R}^p$ and \mathcal{F}_{m-1} -measurable random vector $\tau(Y_1^{m-1}; \theta) \in \mathbb{R}^p$ are determined by the implicit equations

$$E_{\theta_0}[\psi_c(s(Y_1^m; \theta_0)) \psi_c(s(Y_1^m; \theta_0))^\top] = I, \quad (5)$$

$$E_{\theta_0}[\psi_c(s(Y_1^m; \theta_0)) | \mathcal{F}_{m-1}] = 0. \quad (6)$$

The estimating function ψ_c is a truncated version of PML score (3) and can be interpreted as a weighted PML score because by construction $\|\psi_c(s(Y_1^m; \theta))\| \leq c$. Constant $c \geq \sqrt{p}$ is selected by the researcher and controls the degree of robustness of the estimator \bar{a} . It can be chosen according to an estimator's efficiency criterion (as in Mancini et al. (2005)) or based on a more inference oriented principle (see Ronchetti and Trojani (2001)). As in this paper we are interested in efficient estimates of GARCH-type models (the preliminary step in our robust semiparametric bootstrap method), we apply an efficiency criterion in Section 2.4 to select c . Note that when $c = \infty$, \bar{a} is the PMLE of θ_0 under \mathbb{P}_* .

Formally, optimality results in Mancini et al. (2005) hold for Markovian estimating functions, which encompass ARCH-type, but not GARCH-type processes. As in Sakata and White (1998), however, we can realistically expect that our robust estimator performs satisfactorily also under well-behaved GARCH models with sufficient memory decay. Moreover, we can use a Markovian estimating function implied by an approximating ARCH-type process with a sufficient number of lags in the corresponding conditional variance function. For instance, in the standard GARCH(1,1) model we can rewrite the conditional variance $\sigma_t^2(\theta_0)$ as an invertible ARMA process,

$$\begin{aligned} \sigma_t^2(\theta_0) &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2(\theta_0) + \alpha_2 \sigma_{t-1}^2(\theta_0) \\ &= \sum_{j=0}^{+\infty} \alpha_2^j (\alpha_0 + \alpha_1 \varepsilon_{t-1-j}^2(\theta_0)) \\ &= \sum_{j=0}^{l-1} \alpha_2^j (\alpha_0 + \alpha_1 \varepsilon_{t-1-j}^2(\theta_0)) + \alpha_2^l \sigma_{t-l}^2(\theta_0) =: \sigma_t^2(\theta_0)_{\text{trunc}} + \alpha_2^l \sigma_{t-l}^2(\theta_0) \\ &\approx \sigma_t^2(\theta_0)_{\text{trunc}}, \end{aligned}$$

and use the approximation in the last row, as for a sufficiently large lag l we can expect $\alpha_2^l \sigma_{t-l}^2(\theta_0) \approx$

0. Of course, the bounded-influence estimator \bar{a}_{trunc} based on $\sigma_t^2(\theta_0)_{\text{trunc}}$ is slightly biased, but such a bias can be expected to be negligible for a sufficiently large l . In our Monte Carlo simulation (Section 3), we compare robust procedures based on both non Markovian GARCH-type estimating functions and Markovian ARCH-type estimating functions based on a truncated GARCH volatility function. We find that for the lag $l = 30$ estimation results under these two methods are virtually identical. These results are discussed in Section 3.1.

To implement a semiparametric bootstrap for VaR estimation, conditional mean and conditional volatility dynamics of model (1) have to be specified. A substantial amount of empirical evidence suggests that financial asset return volatility is stochastic and mean reverting, return innovations are non-normal, and equity return volatility responds asymmetrically to positive and negative returns; see for instance Ghysels, Harvey, and Renault (1996). In discrete time settings, the stochastic volatility is often described using GARCH model introduced by Engle (1982) and Bollerslev (1986). Comprehensive surveys of the GARCH and related models are Bollerslev, Chou, and Kroner (1992) and Bollerslev, Engle, and Nelson (1994). Several different GARCH-type models for asset returns have been proposed in the financial literature. Our robust bootstrap method can accommodate quite general parametric specifications for conditional mean and conditional variance of asset returns. The different specifications imply different estimating functions, but do not change the overall procedure. In our simulations and empirical applications, we adopt a fairly flexible semiparametric model for asset returns. We assume an AR(1) model for the conditional mean $\mu_t(\theta_0)$ and an asymmetric GARCH(1,1) model for the conditional variance $\sigma_t(\theta_0)$ (Glosten, Jagannathan, and Runkle (1993)),

$$\mu_t(\theta_0) = \rho_0 + \rho_1 Y_{t-1}, \quad (7)$$

$$\sigma_t^2(\theta_0) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2(\theta_0) + \alpha_2 \sigma_{t-1}^2(\theta_0) + \alpha_3 \varepsilon_{t-1}^2(\theta_0) I_{t-1}(\theta_0), \quad (8)$$

where $\alpha_0, \alpha_1, \alpha_2 > 0$, $|\rho_1| < 1$, $\alpha_1 + \alpha_2 + \alpha_3/2 < 1$, $I_{t-1}(\theta_0) = 1$ when $\varepsilon_{t-1}(\theta_0) < 0$, and $I_{t-1}(\theta_0) = 0$ otherwise. The AR(1) model for μ_t captures potential autocorrelations in daily

returns, for instance due to nonsynchronous transactions in different index components. $\alpha_3 > 0$ accounts for the leverage effect,⁵ that is the stronger impact of “bad news” ($\varepsilon_{t-1}(\theta_0) < 0$) than good “good news” ($\varepsilon_{t-1}(\theta_0) \geq 0$) on conditional variance $\sigma_t^2(\theta_0)$. Compared to symmetric GARCH models, asymmetric GARCH models are better able to fit volatility dynamics of equity and index returns; see for instance Engle and Ng (1993) and Engle and Rosenberg (2002). To our knowledge, however, robust estimators of asymmetric GARCH processes have not yet been applied in the statistics and econometrics literature.

2.3 Robust Semiparametric Bootstrap of GARCH-type Processes

We study semiparametric bootstrap methods for GARCH-type processes consisting of two steps:

1. A preliminary estimation of the AR(1), asymmetric GARCH(1,1) model (7)–(8).
2. A bootstrap procedure for estimated scaled residuals in model (7)–(8).

In the first step, we use the PML and the optimal robust estimators in Section 2.2. Given observed daily rate of returns y_1, \dots, y_T , we denote by $\hat{\theta} = (\hat{\rho}_0 \ \hat{\rho}_1 \ \hat{\alpha}_0 \ \hat{\alpha}_1 \ \hat{\alpha}_2 \ \hat{\alpha}_3)^\top$ the estimated parameter vector in model (7)–(8), obtained either by PML or optimal robust estimator. In the second step, we apply bootstrap estimators of $F_{T,T+h}$, based on different bootstrap procedures for estimated scaled residuals in model (7)–(8). Estimated scaled residuals from the asymmetric GARCH(1,1) model are defined by

$$\hat{z}_t = \frac{y_t - \mu_t(\hat{\theta})}{\sigma_t(\hat{\theta})}, \quad t = 1, \dots, T,$$

⁵The name leverage effect was introduced by Black (1976) who suggested that a large negative return increases the financial and operating leverage, and rises equity return volatility; see also Christie (1982). Campbell and Hentschel (1992) suggested an alternative explanation based on the market risk premium and volatility feedback effects; see also the more recent discussion by Bekaert and Wu (2000). We shall use the name leverage effect as it is commonly used by researchers when referring to the asymmetric reaction of volatility to positive and negative return innovations.

where conditional means and conditional volatilities $\{\mu_t(\hat{\theta}), \sigma_t(\hat{\theta})\}_{t=1}^T$ are computed recursively from (7)–(8) after having substituted sensible starting values. We denote by \mathbb{P}_T the empirical distribution of estimated scaled residuals $\hat{z}_1, \dots, \hat{z}_T$. Engle and Gonzalez-Rivera (1991) introduce and investigate the theoretical properties of semiparametric GARCH models driven by empirical innovations \hat{z}_t . Applying the robust estimator (4) extends the estimation of such GARCH models to the robust setting.

Accurate parameter estimates $\hat{\theta}$ are of critical importance as they determine the dynamics of conditional means and conditional volatilities as well as the estimated scaled residuals \hat{z}_t . Optimal robust estimator $\bar{a}(\cdot)$ can provide accurate parameter estimates of model (7)–(8) under conditional non normality of general form and in the presence of outliers or other influential points. In contrast, classical PML estimates can be highly biased and inefficient in such a setting. Monte Carlo simulation (Section 3) investigates these issues.

2.3.1 Nonparametric Residual Bootstrap and VaR Estimation

Nonparametric residual bootstrap relies on the nonparametric empirical estimator \mathbb{P}_T as an estimator of the innovation distribution F . Estimation of one day ahead VaR forecast $\hat{y}_{T,T+1}^\alpha$ for horizon $h = 1$ day is easily obtained by means of the empirical quantile \hat{z}_α of \mathbb{P}_T , yielding $\hat{y}_{T,T+1}^\alpha = \mu_{T+1}(\hat{\theta}) + \sigma_{T+1}(\hat{\theta}) \hat{z}_\alpha$.

Estimation of VaR measures for horizons $h \geq 2$ days is more involved and obtained as follows. Select randomly with replacement scaled innovations z_1^*, \dots, z_h^* from \mathbb{P}_T . Compute recursively bootstrap daily returns $y_{T+1}^*, \dots, y_{T+h}^*$ and portfolio prices $p_{T+1}^*, \dots, p_{T+h}^*$ as follows.

- For y_{T+1}^* and p_{T+1}^* :

$$y_{T+1}^* = \hat{\rho}_0 + \hat{\rho}_1 y_T + \sigma_{T+1} z_1^*, \quad p_{T+1}^* = p_T(1 + y_{T+1}^*).$$

- For y_{T+j}^* and p_{T+j}^* , $j = 2, \dots, h$:

$$y_{T+j}^* = \hat{\rho}_0 + \hat{\rho}_1 y_{T+j-1}^* + \sigma_{T+j}^* z_j^*, \quad p_{T+j}^* = p_{T+j-1}^* (1 + y_{T+j}^*),$$

where

$$\sigma_{T+j}^* = (\hat{\alpha}_0 + \hat{\alpha}_1 \sigma_{T+j-1}^{*2} z_{T+j-1}^{*2} + \hat{\alpha}_2 \sigma_{T+j-1}^{*2} + \hat{\alpha}_3 \sigma_{T+j-1}^{*2} z_{T+j-1}^{*2} I_{T+j-1}^*)^{1/2},$$

with $I_{T+j-1}^* = 1$ if $z_{T+j-1}^* < 0$, and $I_{T+j-1}^* = 0$ otherwise.

Using $B = 10,000$, say, bootstrap samples $\{z_1^{*(b)}, \dots, z_h^{*(b)}\}_{b=1}^B$ compute an estimate of $F_{T,T+h}$ as the bootstrap distribution $F_{T,T+h}^*$ of h days ahead simulated returns $y_{T,T+h}^* := p_{T+h}^*/p_T - 1$. The h days ahead VaR forecast at 1% level, say, is given by the 1% empirical quantile of distribution $F_{T,T+h}^*$. The bootstrap method provides an estimate of the entire conditional predictive distribution $F_{T,T+h}$ and other risk measures, such as Expected Shortfall, can be readily computed.

2.3.2 Breakdown Analysis of Nonparametric Residual Bootstrap

Bootstrap estimate $F_{T,T+h}^*$ is obtained by nonparametric bootstrap of estimated scaled innovations $\hat{z}_1, \dots, \hat{z}_T$. Such innovations can well include a certain amount of outliers or influential points. Via the resampling procedure, such observations can heavily affect the estimated quantiles of $F_{T,T+h}^*$ and hence the corresponding VaR forecasts. To investigate this issue, we compute the bootstrap breakdown point of quantile estimates implied by $F_{T,T+h}^*$; see also Singh (1998). Let b_α denote the breakdown point of VaR forecast $\hat{y}_{T,T+h}^\alpha$ implied by $F_{T,T+h}^*$. That is, Tb_α is the smallest number of scaled innovations in the original sample that need to go to $-\infty$ in order to force $\hat{y}_{T,T+h}^\alpha$ to go to $-\infty$. In practice, scaled innovations do not go to $-\infty$ and $\hat{y}_{T,T+h}^\alpha$ could be unreliable well before reaching $-\infty$. The asymptotic definition of the breakdown point is a convenient tool to assess the robustness property of an estimator. By definition, Tb_α is an integer between 1 and T . Without loss of generality, we assume the outliers causing the breakdown of $\hat{y}_{T,T+h}^\alpha$ to be in the lower tail of the empirical distribution \mathbb{P}_T .

Proposition 1

$$b_\alpha = 1 - (1 - \alpha)^{1/h}.$$

Proof. See Appendix A.

Hence the breakdown point $b_\alpha \rightarrow 0$ when $h \rightarrow +\infty$. That is, for a longer horizon h fewer outliers are sufficient to carry $\hat{y}_{T,T+h}^\alpha$ to $-\infty$. For $h = 1$ day, $b_\alpha = \alpha$ as the α -quantile is estimated by the corresponding empirical quantile of \mathbb{P}_T . Table 5 presents numerical values of b_α for different horizons h . Given the very low breakdown point prevailing for longer horizons, such as ten days, we can expect a quite bad performance of bootstrap estimates $\hat{y}_{T,T+h}^\alpha$ implied by nonparametric residual bootstrap, already under a moderate number of outliers or influential points in the data. Our Monte Carlo simulation in Section 3.4 confirms this conjecture.

2.3.3 Semiparametric Residual Bootstrap and VaR Estimation

Semiparametric residual bootstrap with Extreme Value Theory (EVT) relies on a different estimator of the innovation distribution F . Instead of using nonparametric estimator \mathbb{P}_T , the basic idea is to approximate the lower and upper tail of F by means of a parametric generalized Pareto distribution (GPD) whose distribution function $G_{\xi,\beta}$ is

$$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi x/\beta)^{-1/\xi}, & \xi \neq 0, \\ 1 - \exp(-x/\beta), & \xi = 0, \end{cases}$$

where the support of $G_{\xi,\beta}$ is $[0, +\infty)$ for $\xi \geq 0$, and $[0, -\beta/\xi]$ for $\xi < 0$. To estimate the upper tail of F with a GPD, fix a “high” threshold \bar{u} (such as the 90th percentile of $\{\hat{z}_j\}_{j=1}^T$). Then for any $k > \bar{u}$

$$\mathbb{P}_*(Z_t > k) = \mathbb{P}_*(Z_t > \bar{u}) \mathbb{P}_*(Z_t > k | Z_t > \bar{u}) = \mathbb{P}_*(Z_t > \bar{u}) \mathbb{P}_*(Z_t - \bar{u} > k - \bar{u} | Z_t > \bar{u}). \quad (9)$$

In (9), $\mathbb{P}_*(Z_t > \bar{u})$ is easily estimated nonparametrically by estimator $\sum_{j=1}^T \mathbf{1}\{\hat{z}_j > \bar{u}\}$, where $\mathbf{1}\{z > \bar{u}\} = 1$ when $z > \bar{u}$. Excess distribution $F_{\bar{u}}(k - \bar{u}) := 1 - \mathbb{P}_*(Z_t - \bar{u} > k - \bar{u} | Z_t > \bar{u})$ above

threshold \bar{u} can be approximated with a generalized Pareto distribution using the limit result (see Balkema and de Haan (1974) and Pickands (1975))

$$\lim_{\bar{u} \rightarrow \bar{x}} \sup_{0 \leq x < \bar{x} - \bar{u}} |F_{\bar{u}}(x) - G_{\xi, \beta(\bar{u})}(x)| = 0,$$

where \bar{x} is the (finite or infinite) right endpoint of F and $\beta(\bar{u})$ is a positive measurable function. To estimate lower tail of F with a GPD, fix a “low” threshold $\underline{u} < 0$ (such as the 10th percentile of $\{\hat{z}_j\}_{j=1}^T$) and apply for every $k < \underline{u}$ the above procedure to excess losses $x = -(k - \underline{u})$ using a GPD with distribution function $G_{\xi, \beta(-\underline{u})}$.

Given GPD parameter estimates $\hat{\xi}^{(1)}, \hat{\beta}^{(1)}$ and $\hat{\xi}^{(2)}, \hat{\beta}^{(2)}$ for the lower and the upper tail of F , scaled innovations z_1^*, \dots, z_h^* are sampled from \mathbb{P}_T and adjusted as follows. For $j = 1, \dots, h$:

- a) If $z_j^* < \underline{u}$, sample a $\text{GPD}(\hat{\xi}^{(1)}, \hat{\beta}^{(1)})$ distributed excess loss x_1 and return $\underline{u} - x_1$.
- b) If $z_j^* > \bar{u}$, sample a $\text{GPD}(\hat{\xi}^{(2)}, \hat{\beta}^{(2)})$ distributed excess gain x_2 and return $\bar{u} + x_2$.
- c) If $\underline{u} \leq z_j^* \leq \bar{u}$, return scaled residual z_j^* itself.

Using adjusted scaled residuals, semiparametric bootstrap methods apply the same procedure as in Section 2.3.1 for nonparametric residual bootstrap in order to estimate the conditional predictive density $F_{T, T+h}$.

2.3.4 Pseudo Maximum Likelihood and Robust Estimations of Generalized Pareto Distribution

Sampling scaled innovations from a parametric GPD tail can circumvent the robustness problem of standard nonparametric residual bootstrap procedures. However, GPD parameters $\zeta := (\xi, \beta)^\top$ have to be estimated and also it can be important to ensure the robustness of the $\text{GPD}(\hat{\xi}^{(1)}, \hat{\beta}^{(1)})$ and $\text{GPD}(\hat{\xi}^{(2)}, \hat{\beta}^{(2)})$ tails. In our Monte Carlo simulation and empirical application, we use the PML and the robust estimators of GPD parameters $\zeta := (\xi, \beta)^\top$.

Let $q(\cdot)$ denote the PML estimator for the GPD parameters and \mathbb{G}_* the true tail distribution of threshold exceedances X , $q(\cdot)$ is defined by the asymptotic estimating equation

$$E_{\mathbb{G}_*}[s_{\text{gpd}}(X; q(\mathbb{G}_*))] = 0,$$

where $s_{\text{gpd}}(x; \zeta)$ is the GPD score function

$$s_{\text{gpd}}(x; \zeta) = \begin{bmatrix} s_\xi(x; \zeta) = \xi^{-2} \log(1 + \xi x / \beta) - (1 + 1/\xi)(1 + \xi x / \beta)^{-1} x / \beta \\ s_\beta(x; \zeta) = -\beta^{-1} + (1 + 1/\xi)\beta^{-1}(1 + \xi x / \beta)^{-1} \xi x / \beta \end{bmatrix}. \quad (10)$$

As the score function $s_{\text{gpd}}(x; \zeta)$ is unbounded in x , this estimator is not robust. Given positive constant $c_{\text{gpd}} \geq \sqrt{2}$, optimal bounded influence estimator of GPD parameters, noted \bar{q} , is implicitly defined by the estimating equation (see Dupuis (1999))

$$E_{\mathbb{G}_*}[\psi_c(s_{\text{gpd}}(X; \bar{q}(\mathbb{G}_*)))] = 0, \quad (11)$$

where $s_{\text{gpd}}(x; \zeta)$ is the GPD score function (10) and

$$\begin{aligned} \psi_c(s_{\text{gpd}}(X; \zeta)) &:= A(\zeta) (s_{\text{gpd}}(X; \zeta) - \tau(\zeta)) w(X; \zeta), \\ w(X; \zeta) &:= \min(1, c_{\text{gpd}} \|A(\zeta) (s_{\text{gpd}}(X; \zeta) - \tau(\zeta))\|^{-1}). \end{aligned}$$

Matrix $A(\zeta)$ and vector $\tau(\zeta)$ are solutions of the equations

$$\begin{aligned} E_{\zeta_0} [\psi_c(s_{\text{gpd}}(X; \zeta_0)) \psi_c(s_{\text{gpd}}(X; \zeta_0))^\top] &= I \\ E_{\zeta_0} [\psi_c(s_{\text{gpd}}(X; \zeta_0))] &= 0. \end{aligned}$$

The GPD is the natural limiting distribution of a large class of tail distributions; see Embrechts, Klüppelberg, and Mikosch (1997) for details. Hence it is reasonable to assume that the true tail distribution of the data, \mathbb{G}_* , can be approximated by a GPD distribution \mathbb{G}_{ζ_0} , say. However, several authors have emphasized the instability of PML estimator for GPD distribution when a moderate number of influential points is present in the sample; see for instance Cowell and Victoria-Feser (1996) and references therein. Optimal robust estimator \bar{q} in (11) can provide

accurate GPD parameter estimates also in such problematic settings. Furthermore, as shown in Juárez and Schucany (2004), estimator \bar{q} compares favorably also with respect to more involved robust extreme value estimators when the true distribution has support $[0, +\infty)$, which is the relevant case for our application.

2.4 Choices of Threshold Levels and Robustness Tuning Constants

The choice of thresholds \underline{u} and \bar{u} in EVT applications determines the trade-off between bias and variance of the estimator for GPD parameters. Based on Monte Carlo evidence, McNeil and Frey (2000) suggest to use the empirical 10th and 90th quantiles of $\{z_j\}_{j=1}^T$, i.e. $\underline{u} = z_{0.10}$ and $\bar{u} = z_{0.90}$, for estimating lower and higher quantiles in risk management applications.⁶ In our Monte Carlo experiments and empirical applications we select such thresholds for PML and robust estimators of GPD.

An important choice in the application of robust estimators for GARCH models and for GPD is the tuning constants c and c_{gpd} in (4) and (11), respectively. Such constants control the degree of robustness of estimators. Following Mancini et al. (2005), we set the levels of c and c_{gpd} to achieve a given asymptotic efficiency under the reference models \mathbb{P}_{θ_0} and \mathbb{G}_{ζ_0} . To our knowledge, this approach has not yet been applied in models with highly nonlinear dynamics as those studied here. Figure 5 illustrates the relative efficiency of estimators \bar{a} and \bar{q} with respect to the corresponding ML estimators under the reference models \mathbb{P}_{θ_0} and \mathbb{G}_{ζ_0} , as a function of constant c and c_{gpd} . For example the relative efficiency of \bar{a} is measured as $\text{trace}(V(s; \hat{\theta}_n))/\text{trace}(V(\psi_c; \hat{\theta}_n))$, where asymptotic covariance matrices V are estimated using $n = 20,000$ observations simulated under the reference model \mathbb{P}_{θ_0} , s is the ML score function in (3), $\hat{\theta}_n$ the ML estimate, ψ_c the estimating function of the robust estimator in (4) and $\hat{\theta}_n$ the optimal robust estimate. In our Monte Carlo

⁶McNeil and Frey (2000) find that the resulting quantile estimates are rather insensitive to threshold levels for wide ranges of \underline{u} and \bar{u} ; see also Gonzalo and Olmo (2004) for a related study.

simulation, the choice $c = 11$ for robust estimator \bar{a} implies approximately 98% asymptotic relative efficiency (see the left plot in Figure 5). The relative efficiency for \bar{q} is computed analogously and the choice $c_{\text{gpd}} = 8$ implies an asymptotic relative efficiency of about 98% (see the right plot in Figure 5).

3 Monte Carlo Simulation

We compare the performance of semiparametric bootstrap for GARCH processes based on (i) PML or robust estimators for GARCH-type models, introduced in Section 2.2, and (ii) nonparametric or semiparametric residual bootstrap methods, introduced in Section 2.3.

	PML GARCH	Robust GARCH
Empirical dist.	fhs	fhs rob
PML GPD	evt	—
Robust GPD	—	evt rob

For brevity, we call the VaR prediction method based on nonparametric residual bootstrap fhs; see the above panel. When GARCH dynamics are estimated with the robust estimator (4) we call this method fhs rob; evt rob uses robust estimators both for GARCH dynamics and for GPD tail estimations, while evt uses standard PML estimators at both stages. Comparing fhs and fhs rob VaR predictions allows to assess the potential improvement of VaR forecasts due to robust instead of PML estimation of the GARCH model. In Section 3.4 we compare VaR forecasts of previous methods based on the true GARCH parameters. In that context comparing evt and evt rob allows to assess the potential improvement of VaR forecasts due to robust instead of PML estimation of tail distributions. Hence the simulation design allows to evaluate the contribution of each robustification step to the accuracy of VaR predictions.⁷

⁷Considering the two hybrid cases PML GARCH/Robust GPD and Robust GARCH/PML GPD in the above panel is therefore unnecessary.

To study the accuracy of the different methods, we compute out-of-sample VaR forecasts at 1% and 5% confidence levels for horizons $h = 1$ day and $h = 10$ days, under an AR(1), asymmetric GARCH(1,1) model for daily returns. Precisely, we simulate the following dynamics for \mathcal{Y} .

1. *Student t_5 innovation model.* In this experiment, scaled innovation in model (1) is given by

$$Z_t = ((\nu - 2)/\nu)^{1/2} T_\nu, \quad (12)$$

where random variable T_ν has a Student- t distribution with $\nu = 5$ degrees of freedom. Hence $Z_t \sim IID(0, 1)$ and model (1) is dynamically correctly specified.

2. *Laplace innovation model.* Scaled innovation in (1) is given by

$$Z_t = 2^{-1/2} L, \quad (13)$$

where random variable L has a Laplace (or Double exponential) distribution. Such a distribution has a symmetric convex density and displays fatter tails than the t_5 distribution. Also in this experiment $Z_t \sim IID(0, 1)$ and model (1) is dynamically correctly specified.

3. *Replace-innovative model.* Under such a model observed process $\mathcal{Y} := (Y_t)_{t \in \mathbb{Z}}$ is generated according to the data generating process

$$Y_t = \begin{cases} \rho_0 + \rho_1 Y_{t-1} + \varepsilon_t, & \text{with probability } 1 - \kappa, \\ \check{Y}_t, & \text{with probability } \kappa, \end{cases} \quad (14)$$

where $\check{Y}_t \sim N(0, \varrho^2)$, $\varepsilon_t \sim N(0, \sigma_t^2)$ and σ_t^2 is given by (8). Hence at time t there is a probability κ that observation Y_t is not generated by the GARCH dynamics. The possible “shock” \check{Y}_t will affect future realizations of the process, mainly by “inflating” the conditional variance on subsequent days. In this experiment, model (1) is “slightly” misspecified as the dynamic equations (7)–(8) are not satisfied for every t . We set $\kappa = 0.2\%$ and $\varrho = 10$. The probability of contamination, κ , is very low and implies (on average) 4 contaminated observations out of 2,000 observations. The choice for ϱ allows us to compare the accuracy

of the different VaR estimators under very infrequent, but dramatic, (symmetric) shocks. Such shocks could occur over short time periods in real data, as for instance in daily equity returns.

We simulate an AR(1), asymmetric GARCH(1,1) model for the following parameter choice $\rho_0 = \rho_1 = 0.01$, $\alpha_0 = 0.03$, $\alpha_1 = 0.02$, $\alpha_2 = 0.8$, and $\alpha_3 = 0.2$, under the above distributions for Y_t and for a sample size $T = 2,000$. This parameter choice reflects somehow parameter estimates typically obtained for daily percentage index or exchange rate returns; see for instance Bollerslev et al. (1994). At the reference model \mathbb{P}_{θ_0} , annualized volatility of Y_t is about 12%. The tuning constants for the robust GARCH estimators \bar{a} and \bar{a}_{trunc} are $c = 11$. The one for the robust GPD estimator \bar{q} is $c_{\text{gpd}} = 8$. Each model is simulated 1,000 times. For each simulated sample path, we estimate the model (7)–(8) using classical and robust estimators and we apply the four VaR prediction methods (fls, fls rob, evt, evt rob) to compute conditional VaR forecasts (as a percentage of p_T) at confidence levels $\alpha = 5\%$, 1% and horizons $h = 1$ day, 10 days ahead. In the financial industry, virtually only out-of-sample VaR forecasts are required and in-sample measurements of VaR are by far less important. In our simulations and empirical applications, all VaR forecasts are out-of-sample ones.

3.1 GARCH Dynamics Estimation

Table 1 presents bias and mean square error (MSE) (in percentage) of PML, robust and “truncated” robust parameter estimates, \bar{a} and \bar{a}_{trunc} , for the AR(1), asymmetric GARCH(1,1) model (7)–(8); see also Figures 1–4. Under reference model \mathbb{P}_{θ_0} , PMLE (which is indeed MLE) is only slightly more efficient than the robust estimators of GARCH models. In all other experiments, robust estimators always outperform classical PML estimator in terms of mean square errors. In particular, under the contaminated replace-innovative model, both robust estimators largely outperform the PML estimator. Under all considered models, the overall performances of the two

robust estimators \bar{a} and \bar{a}_{trunc} are very close, although \bar{a} has somewhat lower mean square errors. We interpret the last finding in support of the application of \bar{a} to dynamic models with estimating functions depending on an infinite number of process coordinates, such as GARCH models.

3.2 VaR Violations

Table 2 shows the number of violations of fhs, fhs rob, evt and evt rob VaR forecasts for horizons $h = 1$ day, 10 days, and confidence levels $\alpha = 5\%$, 1% . In the i -th simulation, a violation occurs when the actual loss is larger than the predicted VaR, i.e. $I(i) := \mathbf{1}\{y_{T,T+h}(i) < \hat{y}_{T,T+h}^\alpha(i)\} = 1$ and zero otherwise. Under the null hypothesis that the proposed method estimates the VaR correctly, the test statistic $\sum_{i=1}^{1000} I(i)$ is binomially distributed, $\text{Bin}(1000, \alpha)$, as the 1,000 simulations are independently drawn. Hence for $\alpha = 0.05$ and 0.01 the expected number of violations are 50 and 10, and two-side confidence intervals at 95% level are $[37, 64]$ and $[4, 17]$, respectively. For all horizons and VaR confidence levels, Table 2 shows that all methods exhibit numbers of violations within such intervals. Hence apparently all methods are rather accurate in forecasting VaR, when relying only on violation tests. Indeed, it is known that FHS-type VaR forecasts tend to display the proper number of violations; see for instance Kuuster, Mittnik, and Paoletta (2006) for a recent comparison of these methods. However, Table 2 also hints some differences among VaR prediction methods. In the first two Monte Carlo experiments (Student t_5 and Laplace innovation models), only evt rob never exhibits test p -values below 0.10, even though estimated GARCH models are correctly specified. These results suggest that evt rob can outperform other approaches even in setups relatively favorable to classical methods, but this aspect is not clearly detected by violation tests. In the next section we study the precision of VaR forecasts, which, of course, plays a key role in measuring market risk.

In our simulation study VaR violations both at horizons $h = 1$ day and 10 days are all mutually independent as the different sample paths are independently drawn. Hence this approach allows

to test for VaR violations accurately. In empirical applications, VaR violations are not necessarily independent and it is standard practice to apply the independence test proposed by Christoffersen (1998).⁸ In the empirical application (Section 4) we apply also the independence and other tests.

3.3 VaR Prediction

Left panel in Table 3 shows bias and MSE of fhs, fhs rob, evt and evt rob in predicting VaR at $h = 1$ day ahead horizon and $\alpha = 5\%$, 1% confidence levels. The true VaR is computed under the true data generating process. In all experiments the robust versions of FHS and EVT methods outperform the corresponding classical versions in terms of MSE. The reduction in VaR forecast MSE is small in the Laplace innovation model, but reaches about 80% in the contaminated replace-innovative model. In particular, fhs rob provides more accurate VaR forecasts than fhs because relies on more accurate estimates of volatility dynamics; see Table 1. In all but one case, evt rob has the lowest VaR forecast MSE.

Right panel in Table 3 shows the accuracy of VaR predictions at $h = 10$ days horizon. The true VaR is obtained by simulating 100,000 times true dynamics for \mathcal{Y} over the ten days horizon $[T, T + h]$. In the first two experiments the dynamic model (1) is correctly specified and all VaR prediction methods tend to perform similarly in predicting VaR at 5% level, although evt rob outperforms all other methods in predicting VaR at 1% level. In the third experiment the dynamic model (1) is misspecified and both FHS methods perform very poorly, with fhs rob having the largest MSE for VaR prediction at the 1% level. At first sight, the last finding might appear puzzling given the higher accuracy of robust GARCH estimates documented in Table 1. The result is explained by the very low breakdown point of VaR predictions based on nonparametric residuals bootstrap and the larger absolute residuals produced by robust estimation in this case.⁹ This

⁸The independence test would be trivially verified in the present simulation setup.

⁹To raise the breakdown point of nonparametric bootstrap quantiles, Singh (1998) suggests to winsorize the data before bootstrapping. We winsorized scaled innovations at 0.5% and 1% levels, respectively, and then we

point is discussed in more detail in Section 3.4 below. In terms of MSE, robust EVT predictions largely outperforms all other methods, often by several orders of magnitude. These findings put in perspective the results based on violation tests discussed in the previous section. Even though the estimated GARCH model is misspecified and `evt rob` provides more accurate VaR predictions, violation tests have a quite low power in discriminate among the different VaR prediction methods; see Table 2. Measuring the precision of VaR forecasts reveals that `evt rob` largely outperforms all other methods. In our empirical application (Section 4), we report common violation test results and measure, in addition, the precision of VaR forecasts by investigating model estimation risk.

3.4 Residual Bootstrap Breakdown Point and Quantile Estimates

Accuracy

The accuracy of above bootstrap procedures for GARCH-type processes depends on the accuracy of (i) GARCH parameter estimates and (ii) quantile estimates implied by GARCH residuals bootstrap. We showed above that robust GARCH parameter estimators have higher accuracy under different forms of conditional non normality of returns. To deepen the analysis of the relative accuracy between residual nonparametric bootstrap and residual semiparametric bootstrap using EVT, we repeat the previous Monte Carlo simulation using true GARCH parameters in the estimation of scaled residuals $\hat{z}_1, \dots, \hat{z}_T$. In this way, we eliminate estimation errors due to GARCH parameter estimation and we can disentangle the contribution of nonparametric, classical and robust EVT methods to VaR predictions. We also investigate the theoretical predictions of Proposition 1 on breakdown points of bootstrap quantiles.

We estimate 5% and 1% quantiles (VaR) of ten days ahead return distribution. As GARCH parameters are not estimated classical and robust FHS methods coincide and we call them simply computed VaR predictions using classical and robust FHS over ten days horizon. MSE's of the winsorized VaR predictions did decrease, but only by a small amount and the results are not reported.

“resampling” in this section. We also consider different ways of implementing semiparametric bootstrap methods using EVT. We make an additional distinction depending on whether the quantile of simulated ten days ahead distribution is estimated empirically or using a GPD (PML or robust) estimator. This distinction highlights the additional contribution of parametric GPD over nonparametric tail estimations in producing accurate VaR forecasts.¹⁰ Summarizing, we compute VaR predictions using the following five approaches:

1. Resampling (FHS),
2. EVT applied to both daily returns and to simulated ten days ahead returns,
3. Robust EVT applied to both daily returns and to simulated ten days ahead returns,
4. EVT applied to daily returns, and empirical quantile estimation applied to ten days ahead returns,
5. Robust EVT applied to daily returns, and empirical quantile estimation applied to ten days ahead returns.

Table 4 presents simulation results for the five bootstrap methods under the previous three Monte Carlo experiments. As expected, all MSE’s of VaR forecasts in Table 4 are lower than the corresponding ones in Table 3, as variability deriving from estimation of GARCH parameters is absent now. In the first two experiments (Student t_5 and Laplace innovations), the data generating processes do not produce “outliers”. Resampling procedures and robust EVT perform satisfactorily for quantiles at the 5% level. Classical EVT method is the least precise in this case, by a large amount for the Laplace innovations case. For 1% quantiles, robust EVT methods have a uniformly higher accuracy.

¹⁰In all previous simulations, classical and robust EVT estimations were always applied to both daily and ten days ahead return distributions.

Under the replace-innovative model, nonparametric resampling method breaks down in the estimation of the 1% quantile in the presence of 0.20% outliers in the data, whereas it produces accurate results in estimating 5% quantiles. From Table 5, the breakdown point of VaR at 5% level corresponds to 0.51% contamination by outliers, whereas the breakdown point of VaR at 1% level is 0.10%. Hence as predicted by Proposition 1, 0.20% of outliers breaks down VaR predictions at 1%, but not VaR predictions at 5% obtained by nonparametric residual bootstrap.¹¹ In predicting VaR at 1%, robust EVT method applied both to daily and ten days ahead returns has the lowest MSE across all simulation experiments. Overall, applying robust EVT estimators to both daily and ten days ahead returns seems to be particularly important when forecasting VaR at low confidence level 1% and/or data are contaminated by outliers. Given the very small percentage of outliers generated in our simulations (0.20%) and the difficulties documented for the other methods, VaR forecasts based on robust EVT and robust GARCH estimations can be expected to be the most reliable ones in many realistic applications.

4 Real Data Estimation and Backtesting

We backtest VaR prediction methods on four historical series of daily rate of returns: the S&P 500 index from December 1988 to July 2003, the Dollar-Yen exchange rate from January 1986 to January 2005, Microsoft share price from March 1986 to January 2005, and Boeing share prices from January 1980 to January 2005. The data are downloaded from Datastream. To backtest the four VaR prediction methods on a historical series y_1, \dots, y_N , where $N \gg n$, we compute the out-of-sample VaR forecast $\hat{y}_{T, T+h}^\alpha$ for each $T \in \mathcal{T} = \{n, n+1, \dots, N-h\}$ using a time window of n historical daily returns for each estimation. We set $n = 2,000$ and hence about eight years of data were used for each estimation. For each day $T \in \mathcal{T}$, we predict VaR's at horizons $h = 1$ day, 10 days, and confidence levels $\alpha = 1\%$, 5%, using fhs, fhs rob, evt and evt rob. AR(1),

¹¹This finding is confirmed by the results in the right panel of Table 3.

asymmetric GARCH(1,1) model is estimated using the PML estimator (3) and the optimal robust estimator (4) with tuning constant $c = 8$, the estimates are updated every 500 days and the tuning constant of the optimal robust GPD estimator is $c_{\text{gpd}} = 6$.¹²

4.1 Data and GARCH Estimation

Table 6 shows summary statistics for the daily rate of returns. As expected, different assets have different characteristics. For example, Dollar-Yen exchange rates have relatively high skewness and Microsoft returns high kurtosis unconditionally. These different characteristics make the backtesting exercise particularly interesting. Table 7 shows the AR(1), asymmetric GARCH(1,1) model (7)–(8) estimates given by classical PML and robust \bar{a} estimators. In several occasions and especially for the volatility parameters the two estimates are rather different and next sections show how they induce different VaR forecasts. Interestingly, the asymmetry parameter a_3 turns out to be (very close to) zero for the Dollar-Yen exchange rates. This parameter controls the asymmetric impact of positive and negative shocks on conditional variance. As negative shocks or depreciations of one currency correspond to positive shocks or appreciations of the other, asymmetric effects can be absent in exchange rates. For example, using options data Xu and Taylor (1994) document symmetric, instead of skewed, implied volatility smiles for exchange rates, confirming no asymmetric impact of shock returns on volatility.

4.2 Backtesting VaR Prediction

To assess the forecasting performance of the different VaR prediction methods we adopt the testing framework proposed by Christoffersen (1998). Although this framework has become a

¹²Considering the “noisier” nature of real data, as opposed to simulated data, and the different characteristics of financial time series (indexes, stocks and exchange rates) used in backtesting, we take a somewhat more conservative viewpoint setting the robustness tuning constants c and c_{gpd} to lower levels than in the Monte Carlo study.

standard theoretical framework for evaluating out-of-sample forecasts, for completeness and to setup the notation we briefly recall it here as well; see also Christoffersen (2003). Denote by $I_t := \mathbf{1}\{y_{t+1} < \hat{y}_{t,t+1}^\alpha\}$, then $I_t = 1$ when the actual return is lower than the predicted VaR (i.e. a violation occurs) and zero otherwise. If the VaR prediction method efficiently uses all the available information, then

$$E[I_t | \mathcal{F}_t] = \alpha \quad (15)$$

and I_t is uncorrelated with any function in the information set available at time t . If (15) holds, then VaR violations will occur with the correct unconditional and conditional probability, and neither the forecast for VaR nor that for I_t could be improved.

The test of unconditional coverage is based on the unconditional expectation of (15), that is

$$H_0 : E[I_t] = \alpha \quad \text{vs.} \quad H_A : E[I_t] \neq \alpha.$$

Under the null hypothesis, the likelihood-ratio (LR) test statistic

$$LR_{uc} = -2 \log[\mathcal{L}(\alpha) / \mathcal{L}(\hat{\alpha})] \stackrel{\text{asy}}{\sim} \chi^2(1),$$

where $\mathcal{L}(\alpha)$ is the binomial likelihood and $\hat{\alpha} = n_1 / (n_0 + n_1)$ is the MLE of α , that is the ratio of the number of violations n_1 to the total number of observations $(n_0 + n_1)$.

The test of independence aims at verifying possible clusterings of violations over time. This phenomenon occurs for instance when the VaR prediction method does not account for changing volatilities properly. Under the null hypothesis, a violation today has no influence on the probability of a violation tomorrow. Under the alternative hypothesis $\{I_t\}$ is a binary first-order Markov chain with transition probability $\pi_{ij} = P(I_t = j | I_{t-1} = i)$ and the likelihood is

$$\mathcal{L}(\Pi) = (1 - \pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1 - \pi_{11})^{n_{10}} \pi_{11}^{n_{11}},$$

where n_{ij} is the number of observations with value i followed by value j . The ML estimators are

$$\hat{\pi}_{01} = \frac{n_{01}}{n_{00} + n_{01}}, \quad \hat{\pi}_{11} = \frac{n_{11}}{n_{10} + n_{11}}.$$

Under the null hypothesis of independence, $\pi_{01} = \pi_{11} =: \pi_0$, the likelihood is

$$\mathcal{L}(\pi_0) = (1 - \pi_0)^{n_{00}+n_{10}} \pi_0^{n_{01}+n_{11}}$$

and the ML estimate is $\hat{\pi}_0 = (n_{01} + n_{11}) / (n_{00} + n_{10} + n_{01} + n_{11})$. The LR test is

$$LR_{\text{ind}} = -2 \log[\mathcal{L}(\hat{\pi}_0) / \mathcal{L}(\hat{\Pi})] \stackrel{\text{asy}}{\sim} \chi^2(1).$$

The test of independence cannot account for the exact conditional coverage (15) as π_0 is estimated.

The test of conditional coverage imposes $\pi_0 = \alpha$ and LR test is

$$LR_{\text{cc}} = -2 \log[\mathcal{L}(\alpha) / \mathcal{L}(\hat{\Pi})] \stackrel{\text{asy}}{\sim} \chi^2(2).$$

The LR test statistic $LR_{\text{cc}} = LR_{\text{uc}} + LR_{\text{ind}}$ allows to check in which respect the violation series $\{I_t\}$ does not satisfy the correct conditional coverage (15).

Tables 8 and 9 show, respectively, number of violations and p -values of unconditional, independence and conditional coverage tests for one day ahead VaR forecasts. In all tests, backtested assets and VaR confidence levels, only evt rob never displays p -values below 0.10. All the other methods have some difficulties for example with the conditional coverage test in the S&P 500 backtesting, where evt has a p -value of 0.012—while evt rob has a p -value of 0.121. These empirical findings confirm the previous simulation results, where only evt rob never displays p -values below 0.10; see for instance Kuester, Mittnik, and Paolella (2006) and references therein for related empirical studies. Left panels in Tables 11–14 report summary statistics for one day ahead out-of-sample VaR predictions. In these tables “VaR” denotes the average VaR forecasts, Δ the average daily changes in VaR predictions $\{\hat{y}_{T+1,T+1+h}^\alpha - \hat{y}_{T,T+h}^\alpha\}_{T \in \mathcal{T}}$, Δ^2 the corresponding empirical second moment, and $|\Delta|\%$ the average absolute changes divided by the corresponding VaR forecasts in percentage. The last three statistics describe the changes over time of VaR forecasts. In nearly all backtested time series and VaR confidence levels, evt rob has the lowest values of Δ^2 and $|\Delta|\%$. For instance, in the Dollar-Yen and Boeing backtesting and VaR forecasts at 1% level, $|\Delta|\%$ for

evt is about 8.5% and 15.7% larger than the ones for evt rob, respectively. Overall, evt rob tends to produce the most stable VaR predictions over time. Given the correct conditional coverage, the stability over time of VaR predictions is a desirable feature. Using evt rob VaR predictions, financial institutions can adjust outstanding portfolio risk exposures to VaR limits more smoothly and thus more efficiently.

All the previous empirical findings are largely confirmed by the corresponding empirical analysis of ten days ahead VaR predictions. Table 10 shows the number of violations of ten days ahead VaR forecasts and robust two-side p -values for the null hypothesis that the given method correctly predicts VaR.¹³ The lowest p -value for evt rob is 0.28. All the other methods appear to be too conservative in predicting VaR at 5% level in Boeing backtesting, with p -values below 0.07. Right panels in Tables 11–14 show summary statistics for ten days ahead VaR forecasts. In particular, average levels of VaR predictions at 1% level are quite different across different methods. As in the previous case of one day ahead VaR forecasts, in nearly all backtested time series and VaR confidence levels, evt rob VaR forecasts are the most stable over time in terms of squared and absolute relative changes, Δ^2 and $|\Delta|\%$. As an example, Figure 6 shows ten days ahead VaR forecasts and 1% confidence level reserve amounts¹⁴ for the S&P 500 index. In several occasions the four VaR prediction methods provide different VaR forecasts. For instance at beginning of April and June '03, reserve amounts estimated using evt are more than 15% higher than the ones estimated using evt rob. A financial institutions which relies on evt would incur in substantial opportunity costs, as interest rates earned on reserve amounts are virtually zero. These costs could be avoided using evt rob.¹⁵

Altogether, the previous empirical findings show that evt rob provides more accurate (in terms

¹³Robust standard errors are computed using the Newey and West (1987) covariance matrix with $h - 1$ lags.

¹⁴Reserve amounts refer to a \$100 long position in the S&P 500 index at the beginning of the backtesting period.

¹⁵Additional summary statistics and plots of VaR forecasts at different horizons and confidence levels for the four VaR prediction methods are available from the authors upon request.

of violations) and stable over time VaR predictions than the other methods. Robust VaR predictions rely on robust estimations of both GARCH models and tail distributions, which can be quite different (and more accurate) than classical ones according to our simulation study. Hence the different temporal profiles of VaR forecasts are certainly due to different estimates of volatility dynamics and tail distributions. However, an additional source of variability in VaR forecasts can be model estimation risk. In particular, tail distributions are re-estimated every day and this procedure can affect time profiles of VaR forecasts. Different prediction methods can cope with estimation risk differently and the next section investigates this issue.

4.3 Tail Estimation Risk

The previous Monte Carlo study points out that the VaR prediction methods behave very differently in terms of precision of VaR forecasts—as demonstrated by the widely different MSE’s in Table 3. Accuracy of VaR forecasts cannot be measured in empirical applications as true VaR’s are unknown. However, it is possible to quantify the precision of VaR forecasts by providing prediction intervals for the VaR forecast itself. When a VaR prediction method behaves properly in terms of VaR violations, the narrower the prediction interval for VaR forecasts, the more accurate the VaR predictions.¹⁶ Christoffersen and Gonçalves (2005) suggest a resampling technique to address this issue and we follow their methodology here. As many other resampling techniques, this procedure can be very computationally demanding. To keep the computational burden feasible, we limit the analysis to fhs and evt rob ten days ahead VaR forecasts. For both methods and for each day $T \in \mathcal{T}$, we obtain $S = 199$ conditional VaR predictions $\{\hat{y}_{T,T+h}^{\alpha(s)}\}_{s=1}^S$ at confidence levels $\alpha = 5\%$ and 1% , i.e. we repeat S times the forecasting procedures in Sections 2.3.1 and 2.3.3 for classical FHS and robust EVT methods, respectively. The robust EVT method is particularly demanding

¹⁶Precisely, narrower prediction intervals imply more accurate VaR forecasts when the nominal coverage of the prediction interval is attained and VaR forecasts are unbiased.

because on each day and for each one of the 199 random samples GPD distributions are fitted using the robust estimator to both tails of the bootstrap GARCH residual distribution and to the left tail of simulated ten days ahead return distribution. Then for each day T we compute the prediction interval at 80% confidence level for the VaR forecast $\hat{y}_{T,T+h}^\alpha$,

$$\left[Q_{0.1} \left(\{\hat{y}_{T,T+h}^{\alpha(s)}\}_{s=1}^S \right), Q_{0.9} \left(\{\hat{y}_{T,T+h}^{\alpha(s)}\}_{s=1}^S \right) \right],$$

where $Q_x(\cdot)$ is the x -quantile of the empirical distribution of $\{\hat{y}_{T,T+h}^{\alpha(s)}\}_{s=1}^S$. Other confidence levels are certainly conceivable but the results based on the 80% confidence level are likely to be representative of the findings based on other confidence levels. Given the relatively low number of random samples, $S = 199$, the 80% confidence level seems to be a reasonable choice.

Table 15 shows absolute and relative average widths of ten days ahead VaR prediction intervals in the backtesting period. For all backtested assets and VaR confidence levels, evt rob has narrower prediction intervals than fhs, both in absolute and relative terms. Therefore, evt rob provides more accurate and reliable VaR predictions than fhs. For example, in S&P 500 and Boeing backtesting and for VaR forecasts at 1% level, classical FHS relative prediction intervals are about 14% and 21% larger than the ones for robust EVT.

In Table 15 the columns headed var% report average variances of daily changes in prediction intervals during the backtesting period. In all but one case, daily changes of evt rob prediction intervals have smaller variances than daily changes of fhs prediction intervals. For example, in S&P 500 and Boeing backtesting and for VaR forecasts at 1% level, such variances for evt rob are nearly 50% the ones for fhs. These results largely confirm the previous empirical findings that evt rob provides VaR predictions more stable over time than other methods. Overall, our robust procedure appears to control estimation risk in a better way than classical procedures and this induces more stable VaR profiles over time.

5 Conclusion

We propose a general approach to estimate conditional predictive distributions of asset returns based on robust semiparametric bootstrap methods for GARCH-type processes. Our approach adopts a robust estimator for parametric GARCH-type models and a robustified resampling method for standardized GARCH residuals. In the latter, a robust extreme value estimator controls the bootstrap instability deriving from influential observations in the tails of GARCH residual distributions. We show theoretically and by Monte Carlo simulation that to ensure robustness of the whole bootstrap procedure both robustification steps are necessary. We apply this procedure to Value at Risk (VaR) predictions over different forecasting horizons. Monte Carlo simulation shows that our robust bootstrap procedure offers large improvements in accuracy of VaR predictions, especially for several days ahead horizons and in the presence of outlying or influential observations. In nearly all Monte Carlo experiments, our robust bootstrap procedure has lower mean square prediction errors, often by a large extent, and in contrast to classical methods never fails validation tests at usual significance levels. Theoretical predictions of bootstrap breakdown points are confirmed by our simulations and non robust bootstrap procedures break down approximately at the calculated breakdown point. The simulation evidence is confirmed by the real data application. We backtest VaR prediction methods using S&P 500, Dollar-Yen, Microsoft and Boeing historical returns. For all VaR confidence levels and forecast horizons, only our robust bootstrap procedure never fails all validation tests at usual significance levels. When accounting for tail estimation error, our robust procedure provides more accurate and more stable over time VaR prediction intervals than other methods. Indeed, robust VaR prediction intervals can be almost 20% narrower and 50% less volatile than classical ones. Hence robust procedures allow to adapt outstanding risky positions to VaR limits more smoothly and thus more efficiently.

Robust semiparametric bootstrap methods have applications beyond risk management. Future applications cover, for instance, robust bootstrap procedures for parametric estimators of GARCH-

type models or estimation of fund performance measures using bootstrap methods.

A Proof of Proposition 1

Let η denote the fraction of outliers in the original sample of scaled innovations $\hat{z}_1, \dots, \hat{z}_T$. Then it follows

$$\mathbb{P}_T(y_{T,T+h}^* \text{ has at least one outlier }) = 1 - (1 - \eta)^h.$$

By definition, $\hat{y}_{T,T+h}^\alpha$ is the α -quantile of $F_{T,T+h}^*$. Therefore, $\hat{y}_{T,T+h}^\alpha$ breaks down when a sufficiently large proportion of simulated returns $y_{T,T+h}^*$ is corrupted, i.e. when

$$\mathbb{P}_T(y_{T,T+h}^* \text{ has at least one outlier }) \geq \alpha.$$

The probability on the left hand side gives the fraction of corrupted $y_{T,T+h}^*$ in the simulation.

When this fraction is larger than α , $\hat{y}_{T,T+h}^\alpha$ breaks down. Therefore,

$$b_\alpha = \arg \min_{\eta} \{ 1 - (1 - \eta)^h \geq \alpha \},$$

implying the result. □

Gaussian innovations

Estimator	$\rho_0 = 0.01$		$\rho_1 = 0.01$		$\alpha_0 = 0.03$		$\alpha_1 = 0.02$		$\alpha_2 = 0.8$		$\alpha_3 = 0.2$	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
pml	0.0105	0.0163	0.0253	0.0523	0.0631	0.0042	-0.1589	0.0227	-0.0529	0.0883	0.0198	0.1115
rob \bar{a}	0.0066	0.0165	0.0212	0.0561	0.0597	0.0043	-0.0242	0.0237	-0.0809	0.0894	-0.0305	0.1130
rob \bar{a}_{trunc}	0.0083	0.0170	0.0033	0.0572	0.0671	0.0043	0.0079	0.0241	-0.0979	0.0918	-0.0154	0.1150

Student t_5 innovations

pml	-0.0250	0.0146	-0.1059	0.0696	0.2127	0.0123	0.1634	0.0547	-0.9415	0.2370	0.2669	0.3264
rob \bar{a}	-0.0341	0.0134	-0.1426	0.0583	0.0834	0.0067	-0.3421	0.0232	-0.7573	0.1602	-0.6577	0.2069
rob \bar{a}_{trunc}	-0.0349	0.0135	-0.1427	0.0585	0.0801	0.0070	-0.3158	0.0238	-0.6968	0.1771	-0.6828	0.2203

Laplace innovations

pml	-0.0263	0.0141	-0.1163	0.0681	0.1609	0.0076	0.0121	0.0308	-0.6708	0.1706	0.2565	0.2452
rob \bar{a}	-0.0337	0.0135	-0.1433	0.0589	0.1287	0.0071	-0.4353	0.0252	-0.6674	0.1673	-0.2996	0.2257
rob \bar{a}_{trunc}	-0.0373	0.0136	-0.1459	0.0597	0.1278	0.0066	-0.3937	0.0252	-0.6457	0.1690	-0.2861	0.2350

Replace-innovative model

pml	0.4225	0.0937	-0.4564	0.4612	3.9489	0.6484	4.9875	2.1803	-4.6313	2.7812	3.8014	5.8839
rob \bar{a}	0.0015	0.0184	0.0578	0.0610	1.0472	0.0489	1.2385	0.2152	-1.1921	0.3871	0.8978	0.8934
rob \bar{a}_{trunc}	0.0033	0.0186	0.0447	0.0619	1.0640	0.0516	1.2970	0.2327	-1.1735	0.4297	0.9111	0.9284

Table 1: Percentage bias and MSE of estimated parameters of the AR(1), asymmetric GARCH(1,1) model (7)–(8).

	$h = 1$ day ahead VaR forecasts			$h = 10$ days ahead VaR forecasts		
	Stud. t_5	Laplace	rep.-innov.	Stud. t_5	Laplace	rep.-innov.
fhs 5%	42 (0.25)	41 (0.19)	45 (0.47)	62 (0.08)	62 (0.08)	44 (0.38)
fhs rob 5%	44 (0.38)	43 (0.31)	47 (0.66)	61 (0.11)	61 (0.11)	43 (0.31)
evt 5%	41 (0.19)	43 (0.31)	47 (0.66)	62 (0.08)	62 (0.08)	44 (0.38)
evt rob 5%	43 (0.31)	44 (0.38)	50 (1.00)	60 (0.15)	61 (0.11)	42 (0.25)
fhs 1%	14 (0.20)	15 (0.11)	10 (1.00)	15 (0.11)	13 (0.34)	8 (0.53)
fhs rob 1%	15 (0.11)	15 (0.11)	7 (0.34)	16 (0.06)	14 (0.20)	5 (0.11)
evt 1%	14 (0.20)	15 (0.11)	8 (0.53)	13 (0.34)	13 (0.34)	8 (0.53)
evt rob 1%	13 (0.34)	14 (0.20)	7 (0.34)	12 (0.53)	13 (0.34)	7 (0.34)

Table 2: Number of violations under the different simulation models. In parentheses, two-side p -values for null hypothesis H_0 : number of violations equals to the expected number of violations (i.e. 50 and 10 for 5% and 1% confidence levels, respectively).

$h = 1$ day ahead VaR forecasts					$h = 10$ days ahead VaR forecasts				
Student t_5 innovations									
	VaR 5%		VaR 1%		VaR 5%		VaR 1%		
	bias	MSE	bias	MSE	bias	MSE	bias	MSE	
fhs	−0.0023	0.0043	−0.0041	0.0165	−0.0044	0.0831	−0.0134	0.3574	
fhs rob	−0.0038	0.0038	−0.0075	0.0144	0.0148	0.0831	0.0454	0.3950	
evt	−0.0042	0.0040	0.0046	0.0136	−0.0139	0.0835	0.0222	0.3368	
evt rob	−0.0067	0.0035	0.0014	0.0119	−0.0123	0.0821	−0.0235	0.2877	
Laplace innovations									
fhs	−0.0036	0.0047	−0.0079	0.0161	−0.0132	0.0809	−0.0265	0.3357	
fhs rob	−0.0046	0.0046	−0.0110	0.0159	−0.0005	0.0805	0.0079	0.3463	
evt	−0.0088	0.0041	−0.0078	0.0128	−0.0135	0.0812	0.0476	0.3187	
evt rob	−0.0137	0.0040	−0.0187	0.0126	−0.0257	0.0827	−0.0347	0.2864	
Replace-innovative model									
fhs	0.0053	0.1283	0.0483	0.2481	0.0208	1.4550	0.2790	6.1700	
fhs rob	0.0032	0.0241	0.0119	0.0535	0.0928	1.4371	0.5971	14.9516	
evt	−0.0189	0.1306	0.1612	0.2786	0.0310	1.4466	−0.1933	3.9456	
evt rob	−0.0231	0.0240	0.0818	0.0710	−0.0629	0.4249	−0.6181	1.5884	

Table 3: Bias and MSE of classical and robust, FHS and EVT VaR prediction methods for $h = 1$ day, 10 days ahead and confidence levels $\alpha = 5\%$, 1% under the different simulation models.

Student t_5 innovations				
	quantile 5%		quantile 1%	
	bias	MSE	bias	MSE
resampling	0.0025	0.0566	0.0032	0.2223
daily evt, 10 days evt	−0.0145	0.0733	0.0193	0.2784
daily evt rob, 10 days evt rob	0.0135	0.0599	0.0300	0.1913
daily evt, 10 days emp. quant.	−0.0047	0.0696	−0.0276	0.2706
daily evt rob, 10 days emp. quant.	0.0096	0.0554	0.0181	0.2209
Laplace innovations				
resampling	0.0012	0.0594	−0.0024	0.2007
daily evt, 10 days evt	−0.0354	0.1143	−0.0015	0.4107
daily evt rob, 10 days evt rob	0.0043	0.0624	0.0355	0.1889
daily evt, 10 days emp. quant.	−0.0280	0.1100	−0.0305	0.4169
daily evt rob, 10 days emp. quant.	−0.0012	0.0603	0.0293	0.2002
Replace-innovative model				
resampling	−0.0073	0.1283	0.3037	7.3803
daily evt, 10 days evt	−0.0143	0.1633	−0.2485	1.2285
daily evt rob, 10 days evt rob	−0.0044	0.1547	−0.3853	0.8931
daily evt, 10 days emp. quant.	0.0052	0.1588	−0.3086	1.1703
daily evt rob, 10 days emp. quant.	−0.0167	0.1329	−0.3985	0.9217

Table 4: Bias and MSE of quantile estimates for resampling (first row), EVT applied to daily and 10 days ahead returns (second row), robust EVT applied to daily and 10 days ahead returns (third row), EVT applied to daily returns and empirical quantile estimation applied to 10 days ahead returns (fourth row), robust EVT applied to daily returns and empirical quantile estimation applied to 10 days ahead returns (fifth row), under the different simulation models.

h	1	2	3	4	5	6	7	8	9	10
$\alpha = 5\%$	5.00	2.53	1.70	1.27	1.02	0.85	0.73	0.64	0.57	0.51
$\alpha = 1\%$	1.00	0.50	0.33	0.25	0.20	0.17	0.14	0.13	0.11	0.10

Table 5: For different time horizons (h in days) each entry represents the minimal percentage of outliers in estimated scaled innovations $\hat{z}_1, \dots, \hat{z}_T$ that is sufficient to cause breakdown of VaR estimates based on FHS method.

	Sample size	Mean	Std.	Skew.	Kurt.	Min	Max
S&P 500	3,799	0.034	1.029	-0.163	7.137	-7.113	5.573
Dollar-Yen	4,969	0.016	0.688	0.496	7.596	-3.505	6.795
Microsoft	4,918	0.146	2.565	-0.228	12.223	-31.111	19.552
Boeing	6,535	0.055	1.938	-0.052	8.388	-17.625	15.347

Table 6: Summary statistics for the daily rate of returns in percentage.

S&P 500

Estimator	ρ_0		ρ_1		α_0		α_1		α_2		α_3		Persistency		Ann. Vol.	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
pml	0.042	0.005	0.054	0.019	0.009	0.005	0.008	0.009	0.936	0.030	0.092	0.064	0.990	0.005	14.987	3.650
rob \bar{a}	0.047	0.003	0.050	0.011	0.008	0.003	0.006	0.007	0.936	0.023	0.089	0.049	0.987	0.005	12.437	2.335

Dollar-Yen

pml	-0.007	0.016	0.024	0.013	0.013	0.006	0.064	0.012	0.910	0.022	0.000	0.000	0.974	0.015	11.487	0.877
rob \bar{a}	-0.004	0.015	0.012	0.006	0.010	0.007	0.051	0.013	0.921	0.029	0.000	0.000	0.972	0.017	9.745	0.212

Microsoft

pml	0.165	0.030	-0.011	0.015	0.286	0.170	0.043	0.016	0.869	0.062	0.065	0.020	0.944	0.046	38.556	5.558
rob \bar{a}	0.162	0.034	-0.014	0.019	0.278	0.151	0.043	0.015	0.867	0.056	0.058	0.014	0.940	0.042	35.420	3.570

Boeing

pml	0.054	0.018	0.030	0.035	0.371	0.519	0.044	0.034	0.804	0.240	0.072	0.070	0.885	0.174	31.815	6.022
rob \bar{a}	0.053	0.015	0.039	0.034	0.194	0.192	0.040	0.038	0.876	0.106	0.040	0.019	0.935	0.063	26.171	4.645

Table 7: Mean and standard deviation (Std.) of parameter estimates, persistency (i.e., $\alpha_1 + \alpha_2 + \alpha_3/2$), and annualized volatility in percentage (Ann. Vol.) of the AR(1), asymmetric GARCH(1,1) model (7)–(8). Asset returns in percentage.

	fhs 5%	fhs rob 5%	evt 5%	evt rob 5%	Expected 5%	fhs 1%	fhs rob 1%	evt 1%	evt rob 1%	Expected 1%
S&P 500	108	106	109	104	90	19	21	18	19	18
Dollar-Yen	147	149	152	153	149	32	33	35	33	30
Microsoft	155	153	156	152	146	39	38	39	38	29
Boeing	211	234	217	244	227	47	49	41	46	45

Table 8: Number of violations of classical and robust, FHS and EVT, $h = 1$ day ahead VaR predictions.

Unconditional coverage, LR_{uc}									
	fhs 5%	fhs rob 5%	evt 5%	evt rob 5%	fhs 1%	fhs rob 1%	evt 1%	evt rob 1%	
S&P 500	0.059	0.092	0.046	0.139	0.814	0.489	1.000	0.814	
Dollar-Yen	0.899	0.966	0.769	0.706	0.675	0.550	0.342	0.550	
Microsoft	0.447	0.552	0.398	0.610	0.083	0.117	0.083	0.117	
Boeing	0.276	0.305	0.501	0.120	0.808	0.592	0.508	0.838	
Independence, LR_{ind}									
S&P 500	0.108	0.129	0.028	0.153	0.524	0.481	0.546	0.524	
Dollar-Yen	0.781	0.569	0.420	0.681	0.404	0.389	0.433	0.389	
Microsoft	0.781	0.993	0.900	0.726	0.550	0.522	0.550	0.522	
Boeing	0.531	0.059	0.420	0.148	0.515	0.560	0.388	0.450	
Conditional coverage, LR_{cc}									
S&P 500	0.046	0.076	0.012	0.121	0.794	0.614	0.834	0.794	
Dollar-Yen	0.955	0.849	0.692	0.856	0.646	0.577	0.468	0.577	
Microsoft	0.720	0.838	0.695	0.826	0.185	0.239	0.185	0.239	
Boeing	0.454	0.099	0.576	0.105	0.785	0.731	0.553	0.736	

Table 9: Likelihood ratio p -values for violations of classical and robust, FHS and EVT, $h = 1$ day ahead VaR predictions.

	fhs 5%	fhs rob 5%	evt 5%	evt rob 5%	Expt. 5%	fhs 1%	fhs rob 1%	evt 1%	evt rob 1%	Expt. 1%
S&P 500	75 (0.42)	75 (0.42)	76 (0.46)	75 (0.42)	90	14 (0.49)	14 (0.51)	13 (0.38)	16 (0.75)	18
Dollar-Yen	135 (0.64)	147 (0.97)	148 (1.00)	157 (0.76)	148	25 (0.67)	26 (0.74)	20 (0.38)	24 (0.61)	30
Microsoft	151 (0.84)	150 (0.87)	152 (0.82)	155 (0.74)	146	27 (0.84)	25 (0.69)	29 (0.99)	30 (0.93)	29
Boeing	164 (0.06)	161 (0.05)	166 (0.07)	190 (0.28)	226	57 (0.59)	58 (0.57)	59 (0.53)	63 (0.43)	45

Table 10: Number of violations of classical and robust, FHS and EVT, $h = 10$ days ahead VaR predictions. In parentheses, robust two-side p -values for null hypothesis H_0 : number of violations equals to the expected number of violations (Expt.).

	$h = 1$ day ahead VaR forecasts				$h = 10$ days ahead VaR forecasts			
	VaR	$\Delta \times 10^4$	$\Delta^2 \times 100$	$ \Delta \%$	VaR	$\Delta \times 10^4$	$\Delta^2 \times 100$	$ \Delta \%$
fhs 5%	1.921	1.625	4.712	6.575	6.327	12.787	38.255	5.658
fhs rob 5%	1.925	1.335	4.328	6.435	6.455	12.709	36.673	5.562
evt 5%	1.921	1.535	4.717	6.564	6.338	13.735	37.270	5.450
evt rob 5%	1.928	1.295	4.350	6.436	6.288	10.379	33.848	5.362
fhs 1%	3.124	0.458	8.649	5.293	10.664	21.861	100.833	5.629
fhs rob 1%	3.108	0.699	8.349	5.289	10.934	22.565	99.066	5.594
evt 1%	3.285	1.389	9.572	5.190	10.725	24.678	93.707	5.162
evt rob 1%	3.261	1.057	9.097	5.184	10.470	16.365	85.725	5.098

Table 11: Averages of VaR forecasts and daily changes in VaR forecasts, Δ , for the S&P 500 index from 12/88 to 07/03. $|\Delta|\%$ are Δ in absolute values, divided by VaR forecasts and in percentage.

	$h = 1$ day ahead VaR forecasts				$h = 10$ days ahead VaR forecasts			
	VaR	$\Delta \times 10^4$	$\Delta^2 \times 100$	$ \Delta \%$	VaR	$\Delta \times 10^4$	$\Delta^2 \times 100$	$ \Delta \%$
fhs 5%	1.037	-0.556	0.463	3.518	3.364	-0.861	3.979	3.291
fhs rob 5%	1.032	-0.694	0.431	3.144	3.300	-0.835	3.595	3.106
evt 5%	1.024	-0.498	0.450	3.541	3.305	-1.088	3.825	3.186
evt rob 5%	1.021	-0.554	0.413	3.128	3.262	-1.163	3.306	2.935
fhs 1%	1.653	-0.953	1.103	3.302	4.992	-2.263	9.882	3.616
fhs rob 1%	1.650	-1.189	1.075	3.065	4.880	-2.516	9.023	3.458
evt 1%	1.637	-1.351	1.111	3.320	5.012	-2.646	9.379	3.448
evt rob 1%	1.642	-1.390	1.082	3.059	4.960	-3.271	8.146	3.204

Table 12: Averages of VaR forecasts and daily changes in VaR forecasts, Δ , for the Dollar-Yen exchange rate from 01/86 to 01/05. $|\Delta|\%$ are Δ in absolute values, divided by VaR forecasts and in percentage.

	$h = 1$ day ahead VaR forecasts				$h = 10$ days ahead VaR forecasts			
	VaR	$\Delta \times 10^4$	$\Delta^2 \times 100$	$ \Delta \%$	VaR	$\Delta \times 10^4$	$\Delta^2 \times 100$	$ \Delta \%$
fhs 5%	3.209	-2.359	9.994	4.299	9.594	-3.788	56.051	4.020
fhs rob 5%	3.210	-2.451	9.301	4.297	9.669	-3.391	53.757	4.013
evt 5%	3.216	-2.266	9.968	4.287	9.479	-3.497	54.895	3.965
evt rob 5%	3.223	-2.373	9.280	4.286	9.429	-3.822	50.359	3.905
fhs 1%	5.274	-4.253	25.895	4.151	15.269	-8.646	143.345	4.347
fhs rob 1%	5.287	-4.196	24.015	4.124	15.445	-7.746	138.954	4.330
evt 1%	5.227	-4.716	25.514	4.129	15.413	-6.925	133.829	4.050
evt rob 1%	5.259	-4.751	23.633	4.101	15.229	-9.727	123.972	4.085

Table 13: Averages of VaR forecasts and daily changes in VaR forecasts, Δ , for the Microsoft stock price from 03/86 to 01/05. $|\Delta|\%$ are Δ in absolute values, divided by VaR forecasts and in percentage.

	$h = 1$ day ahead VaR forecasts				$h = 10$ days ahead VaR forecasts			
	VaR	$\Delta \times 10^4$	$\Delta^2 \times 100$	$ \Delta \%$	VaR	$\Delta \times 10^4$	$\Delta^2 \times 100$	$ \Delta \%$
fhs 5%	2.780	-1.007	7.437	4.990	9.360	-2.803	27.446	3.067
fhs rob 5%	2.729	-1.396	5.264	4.822	9.605	-3.750	28.781	3.207
evt 5%	2.761	-1.037	7.385	5.003	9.258	-2.924	26.211	2.987
evt rob 5%	2.696	-1.413	5.212	4.866	8.968	-3.738	26.135	3.160
fhs 1%	4.509	-0.902	16.496	4.357	14.715	-1.070	73.233	3.571
fhs rob 1%	4.474	-1.494	11.079	3.863	15.544	-2.283	84.840	3.769
evt 1%	4.680	-0.984	17.388	4.315	14.701	-1.947	67.628	3.279
evt rob 1%	4.601	-1.582	11.654	3.816	14.118	-2.599	67.078	3.432

Table 14: Averages of VaR forecasts and daily changes in VaR forecasts, Δ , for the Boeing stock price from 01/80 to 01/05. $|\Delta|\%$ are Δ in absolute values, divided by VaR forecasts and in percentage.

Prediction interval VaR 5%			Prediction interval VaR 1%		
Absolute	Relative%	var%	Absolute	Relative%	var%
S&P 500					
fhs	0.283	4.474	0.143	0.692	6.550
evt rob	0.257	4.115	0.097	0.589	5.727
Dollar-Yen					
fhs	0.121	3.548	0.076	0.250	4.909
evt rob	0.112	3.344	0.064	0.240	4.708
Microsoft					
fhs	0.393	4.056	0.101	0.807	5.387
evt rob	0.359	3.766	0.079	0.753	4.938
Boeing					
fhs	0.346	3.932	0.103	0.832	5.969
evt rob	0.308	3.550	0.082	0.668	4.934

Table 15: Prediction intervals that account for tail estimation risk in ten days ahead VaR forecasts. For each day T in the backtesting period, we obtain $S = 199$ ten days ahead conditional VaR forecasts using a resampling technique applied to classical FHS and robust EVT methods. Then we compute prediction intervals at 80% confidence level for ten days ahead VaR forecasts. Absolute is the average width of the 80% prediction intervals (90% minus 10% percentile) of VaR forecasts; Relative% is the average width of the 80% prediction intervals as a percentage of the VaR forecasts; var% is the average variance of daily changes of Relative%.

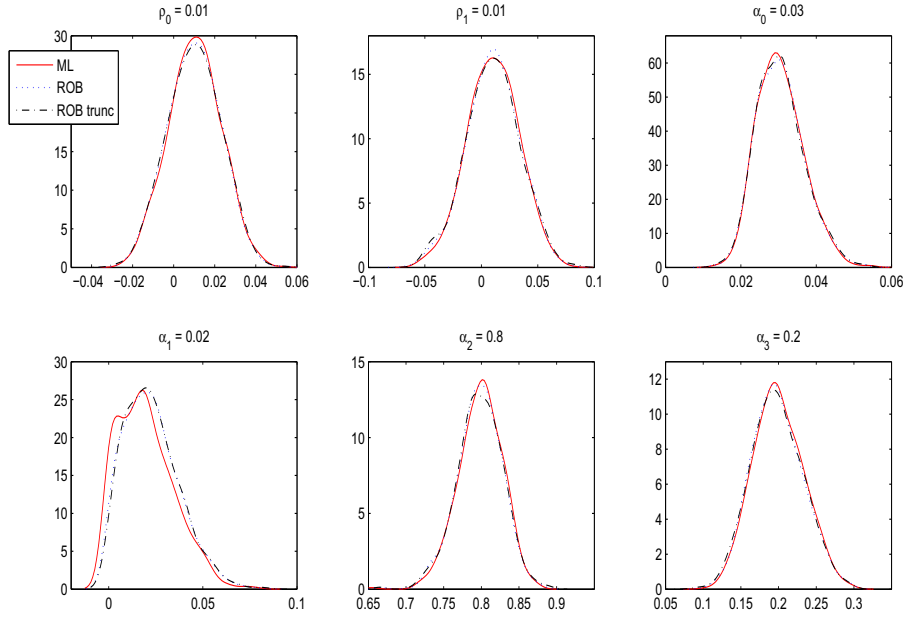


Figure 1: Estimated densities of $\hat{\theta}$ for the AR(1), asymmetric GARCH(1,1) model (7)–(8) under Gaussian innovations.

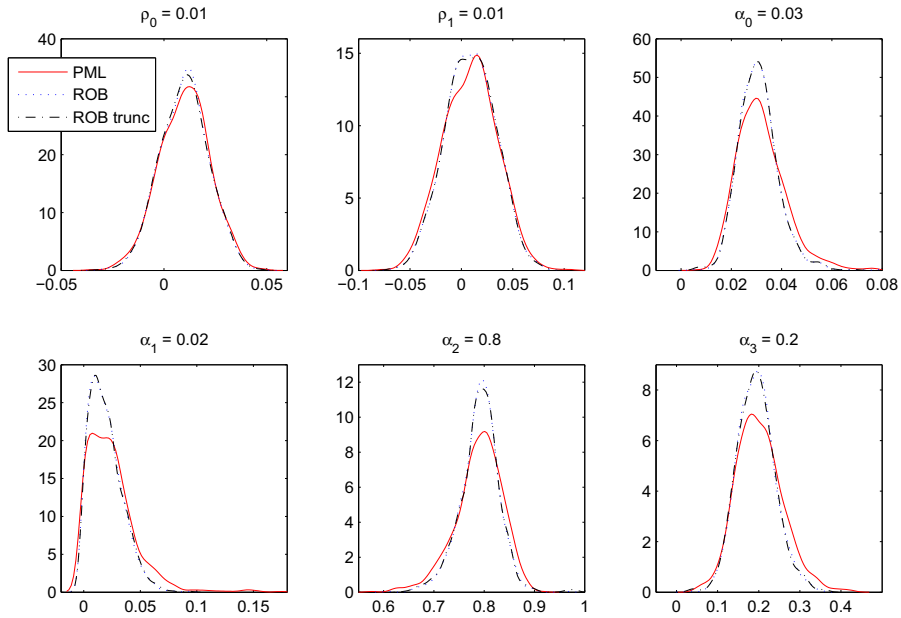


Figure 2: Estimated densities of $\hat{\theta}$ for the AR(1), asymmetric GARCH(1,1) model (7)–(8) under scaled Student t_5 innovations.

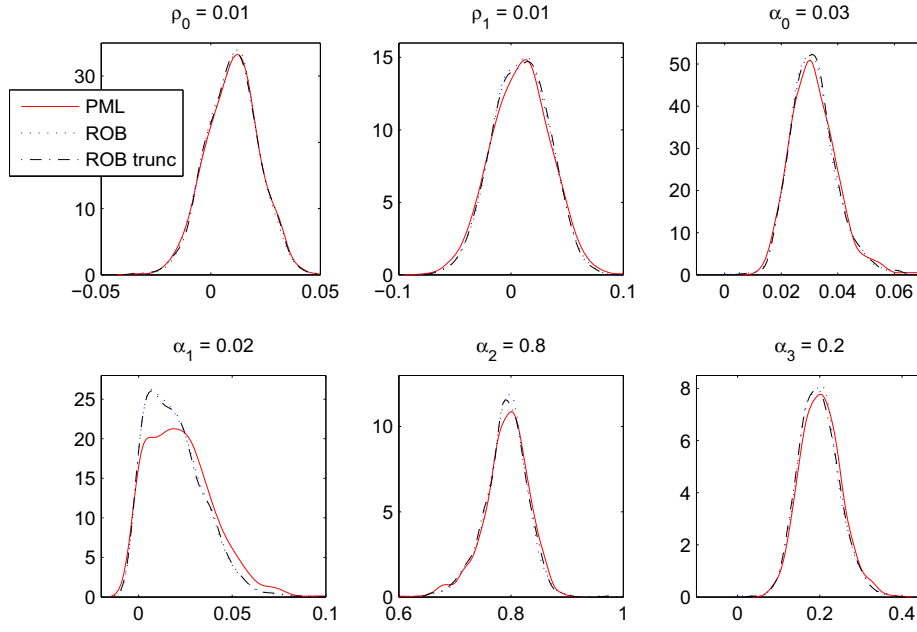


Figure 3: Estimated densities of $\hat{\theta}$ for the AR(1), asymmetric GARCH(1,1) model (7)–(8) under scaled Laplace innovations.

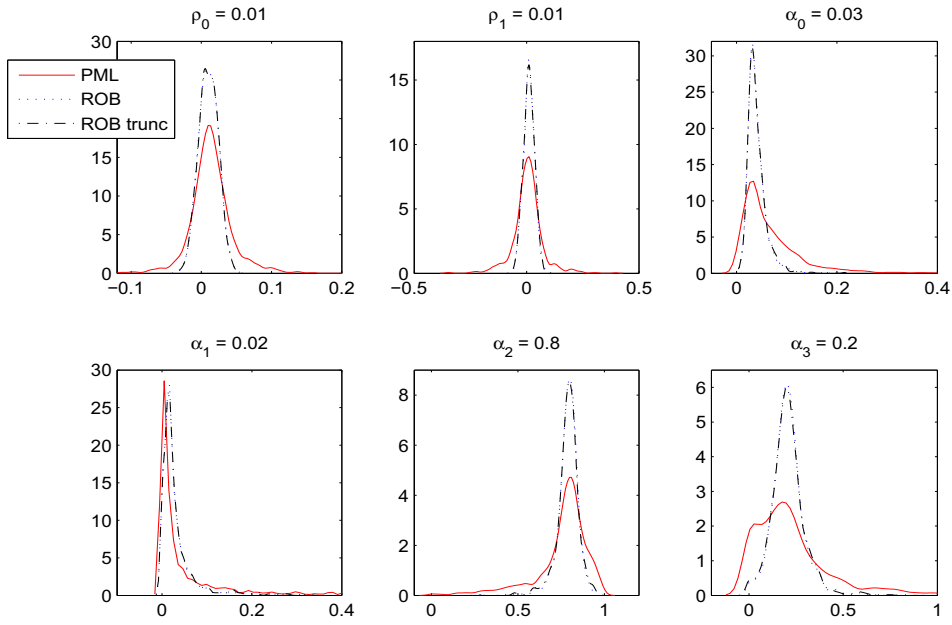


Figure 4: Estimated densities of $\hat{\theta}$ for the AR(1), asymmetric GARCH(1,1) model (7)–(8) under replace-innovative model.

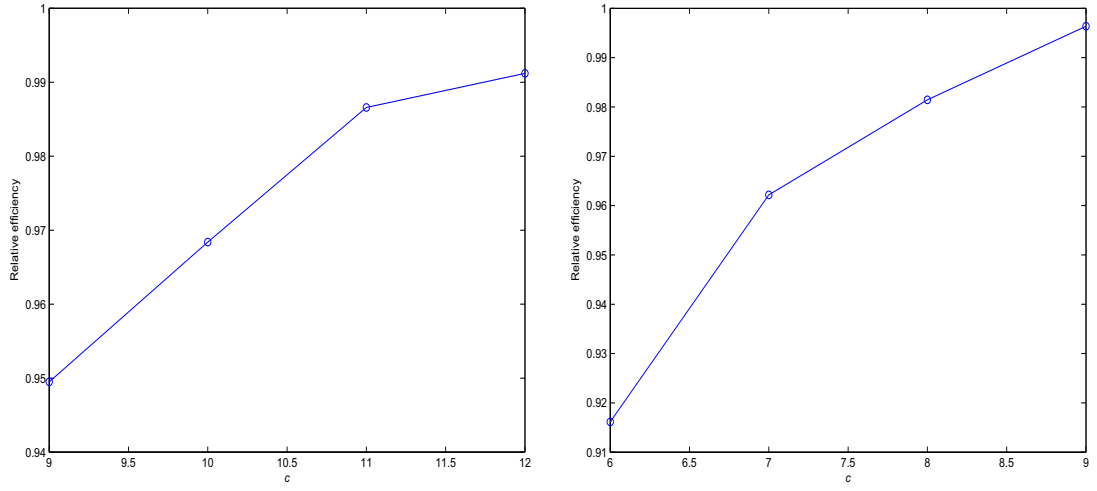


Figure 5: Relative efficiency of the optimal robust estimator \bar{a} in equation (4) with respect to the MLE under the reference model \mathbb{P}_{θ_0} in the AR(1), asymmetric GARCH(1,1) model (left plot); relative efficiency of the optimal robust estimator \bar{q} in equation (11) with respect to the MLE under the reference model \mathbb{G}_{ζ_0} in the GPD model (right plot).

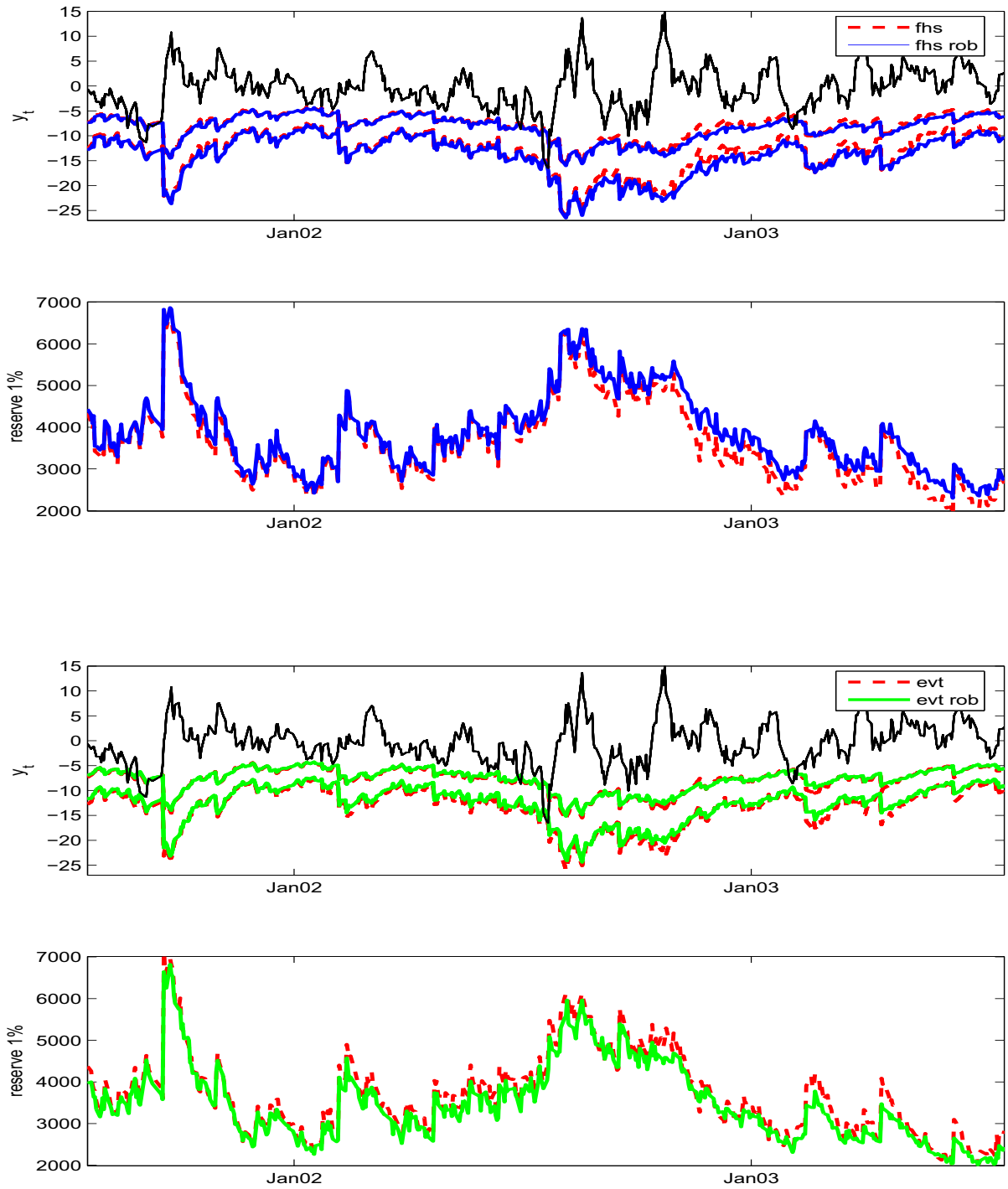


Figure 6: Backtesting on the S&P 500 index from 07/01 to 07/03. Ten days ahead out-of-sample forecasts of VaR's at 5%, 1% confidence levels (superimposed on the ten days rate of returns in percentage) and reserve amounts at 1% level using the classical and robust FHS methods (first two plots) and the classical and robust EVT methods (last two plots).

References

- Aït-Sahalia, Y., and A. Lo, 2000, “Nonparametric Risk Management and Implied Risk Aversion,” *Journal of Econometrics*, 94, 9–51.
- Artzner, P., F. Delbaen, J. Eber, and D. Heath, 1999, “Coherent measures of risk,” *Mathematical Finance*, 9, 203–228.
- Balkema, A., and L. de Haan, 1974, “Residual life time at great age,” *Annals of Probability*, 2, 792–804.
- Barone-Adesi, G., K. Giannopoulos, and L. Vosper, 1999, “VaR without correlations for nonlinear portfolios,” *Journal of Futures Markets*, 19, 583–602.
- Basel Committee, 1996, *Overview of the amendment to the capital accord to incorporate market risks*. Basle Committee on Banking Supervision.
- Bekaert, G., and G. Wu, 2000, “Asymmetric volatility and risk in equity markets,” *Review of Financial Studies*, 13, 1–42.
- Black, F., 1976, “Studies of stock market volatility changes,” in *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economic Statistic Section*, pp. 177–181.
- Bollerslev, T. P., 1986, “Generalized autoregressive conditional heteroskedasticity,” *Journal of Econometrics*, 31, 307–327.
- Bollerslev, T. P., R. Y. Chou, and K. F. Kroner, 1992, “ARCH modelling in finance: A review of the theory and empirical evidence,” *Journal of Econometrics*, 52, 5–59.
- Bollerslev, T. P., R. F. Engle, and D. B. Nelson, 1994, “ARCH models,” in *Handbook of Econometrics*, ed. by R. F. Engle, and D. McFadden. Vol. IV. North-Holland, Amsterdam, pp. 2959–3038.

- Campbell, J. Y., and L. Hentschel, 1992, “No news is good news: An asymmetric model of changing volatility in stock returns,” *Journal of Financial Economics*, 31, 281–318.
- Camponovo, L., O. Scaillet, and F. Trojani, 2006, “Robust subsampling,” working paper, University of Lugano.
- Christie, A., 1982, “The stochastic behavior of common stock variances: Value, leverage and interest rate effects,” *Journal of Financial Economics*, 10, 407–432.
- Christoffersen, P. F., 1998, “Evaluating interval forecasts,” *International Economic Review*, 39, 841–862.
- , 2003, *Elements of financial risk management*. Academic Press, San Diego.
- Christoffersen, P. F., and S. Gonçalves, 2005, “Estimation risk in financial risk management,” *Journal of Risk*, pp. 1–29.
- Cowell, F. A., and M.-P. Victoria-Feser, 1996, “Robustness properties of inequality measures,” *Econometrica*, 64, 77–101.
- Davidson, R., and E. Flachaire, 2005, “Asymptotic and bootstrap inference for inequality and poverty measures,” working paper, McGill University.
- Diebold, F. X., T. Schuermann, and J. Stroughair, 1998, “Pitfalls and opportunities in the use of Extreme Value Theory in risk management,” in *Advances in Computational Finance*, ed. by A.-P. N. Refenes, J. D. Moody, and A. N. Burgess. Kluwer Academic Publishers, Amsterdam, pp. 3–12.
- Duffie, D., and J. Pan, 1997, “An overview of Value at Risk,” *Journal of Derivatives*, 4, 7–49.
- Dupuis, D. J., 1999, “Exceedances over high thresholds: A guide to threshold selection,” *Extremes*, 1, 251–261.

- Embrechts, P., C. Klüppelberg, and T. Mikosch, 1997, *Modelling extremal events for insurance and finance*. Springer, Berlin.
- Engle, R. F., 1982, “Autoregressive conditional heteroskedasticity with estimate of the variance of United Kingdom inflation,” *Econometrica*, 50, 987–1007.
- Engle, R. F., and G. Gonzalez-Rivera, 1991, “Semiparametric ARCH models,” *Journal of Business and Economic Statistics*, 9, 345–359.
- Engle, R. F., and S. Manganelli, 2004, “CAViaR: Conditional autoregressive Value at Risk by regression quantiles,” *Journal of Business and Economic Statistics*, 22, 367–381.
- Engle, R. F., and V. Ng, 1993, “Measuring and testing the impact of news on volatility,” *Journal of Finance*, 48, 1749–1778.
- Engle, R. F., and J. V. Rosenberg, 2002, “Empirical pricing kernels,” *Journal of Financial Economics*, 64, 341–372.
- Foresi, S., and F. Peracchi, 1995, “The conditional distribution of excess returns: An empirical analysis,” *Journal of the American Statistical Association*, 90, 451–466.
- Gagliardini, P., F. Trojani, and G. Urga, 2005, “Robust GMM tests for structural breaks,” *Journal of Econometrics*, 129, 139–182.
- Ghysels, E., A. C. Harvey, and E. Renault, 1996, “Stochastic volatility,” in *Handbook of Statistics*, ed. by G. S. Maddala, and C. R. Rao. North-Holland, Amsterdam, pp. 119–191.
- Glosten, L. R., R. Jagannathan, and D. E. Runkle, 1993, “On the relation between the expected value and the volatility of the nominal excess return on stocks,” *Journal of Finance*, 48, 1779–1801.
- Gonzalo, J., and J. Olmo, 2004, “Which extreme values are really extreme?,” *Journal of Financial Econometrics*, 2, 349–369.

- Gouriéroux, C., J. P. Laurent, and O. Scaillet, 2000, “Sensitivity analysis of Values at Risk,” *Journal of Empirical Finance*, 7, 225–245.
- Gouriéroux, C., A. Monfort, and A. Trognon, 1984, “Pseudo maximum likelihood methods: Theory,” *Econometrica*, 52, 681–700.
- Hull, J., and A. White, 1998, “Incorporating volatility updating into the historical simulation method for VaR,” *Journal of Risk*, 1, 5–19.
- Juárez, S. F., and W. R. Schucany, 2004, “Robust and efficient estimation for the generalized Pareto distribution,” *Extremes*, 7, 237–251.
- Koenker, R. W., and G. Bassett, 1978, “Regression quantiles,” *Econometrica*, 46, 33–50.
- , 1982, “Robust tests for heteroskedasticity based on regression quantiles,” *Econometrica*, 50, 43–61.
- Kuester, K., S. Mittnik, and M. S. Paolella, 2006, “Value-at-Risk prediction: A comparison of alternative strategies,” *Journal of Financial Econometrics*, 4, 53–89.
- Künsch, H., 1989, “The jackknife and the bootstrap for general stationary observations,” *Annals of Statistics*, 17, 1217–1241.
- Mancini, L., E. Ronchetti, and F. Trojani, 2005, “Optimal conditionally unbiased bounded-influence inference in dynamic location and scale models,” *Journal of the American Statistical Association*, 100, 628–641.
- McNeil, A. J., and R. Frey, 2000, “Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach,” *Journal of Empirical Finance*, 7, 271–300.
- Newey, W. K., and K. D. West, 1987, “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 55, 703–708.

- Ortelli, C., and F. Trojani, 2005, “Robust efficient method of moments,” *Journal of Econometrics*, 128, 69–97.
- Peracchi, F., 2002, “On estimating conditional quantiles and distribution functions,” *Computational Statistics and Data Analysis*, 38, 433–447.
- Pickands, J., 1975, “Statistical inference using extreme order statistics,” *Annals of Statistics*, 3, 119–131.
- Pritsker, M., 1997, “Towards assessing the magnitude of Value at Risk errors due to errors in the correlation matrix,” *Financial Engineering News*, 2, 14–16.
- , 2001, “The hidden dangers of historical simulation,” working paper, Finance and Economics discussion series 2001-27. Washington: Board of Governors of the Federal Reserve System.
- RiskMetrics, 1995, *RiskMetrics technical document, 3rd edition*. J.P. Morgan.
- Ronchetti, E., and F. Trojani, 2001, “Robust inference with GMM estimators,” *Journal of Econometrics*, 101, 37–69.
- Sakata, S., and H. White, 1998, “High breakdown point conditional dispersion estimation with application to S&P 500 daily returns volatility,” *Econometrica*, 66, 529–567.
- Scaillet, O., 2004, “Nonparametric estimation and sensitivity analysis of expected shortfall,” *Mathematical Finance*, 14, 115–129.
- Singh, K., 1998, “Breakdown theory for bootstrap quantiles,” *Annals of Statistics*, 26, 1719–1732.
- Xu, X., and S. J. Taylor, 1994, “The term structure of volatility implied by foreign exchange options,” *Journal of Financial and Quantitative Analysis*, 29, 57–74.