CENTER FOR TAX AND ACCOUNTING RESEARCH

Center for Risk Management

TAXATION, ACCOUNTING, AND FINANCE

TAF WORKING PAPER

No. 105 / May 2026

Application of Novel Exponential (Semi-)Parametric Short and Long Memory GARCH Models under Regulatory Requirements of Basel III

Hanke, Dominik / Uhde, André / Feng, Yuanhua

PADERBORN UNIVERSITY

www.taf-wps.cetar.org.

Application of Novel Exponential (Semi-)Parametric Short and Long Memory GARCH Models under Regulatory Requirements of Basel III

Dominik Hanke, André Uhde and Yuanhua Feng

Faculty of Business Administration and Economics

Paderborn University

Abstract

This paper evaluates the forecasting performance of an expanded class of (semi-)parametric GARCH models belonging to the EGARCH family (EGF), including recently introduced long and short memory specifications and their semiparametric extensions. The semiparametric variants employ a multiplicative volatility decomposition into conditional and slowly varying unconditional components, where the latter is estimated via a data-driven local polynomial smoother to accommodate non-stationarities commonly observed in financial time series. Based on the revised Basel Committee framework for market-risk assessment, all models are capable of producing rolling one-day-ahead forecasts for Value at Risk (VaR) and Expected Shortfall (ES) under a wide range of symmetric and skewed innovation distributions. Their forecasting accuracy is examined using the regulatory traffic light tests for VaR and the recently developed ES-specific traffic light procedure, complemented by the regulatory loss function. In addition, model selection incorporates both a recently proposed corrected firm-oriented loss function that accounts for opportunity costs and the Weighted Absolute Deviation (WAD) criterion. The empirical comparison demonstrates that (semiparametric) long memory GARCH models - particularly those combining fractional dynamics with nonparametric scale adjustments - can serve as valuable alternatives to traditional parametric short memory models, offering more stable volatility estimates and improved tail-risk forecasts for practical risk management applications.

Keywords: semiparametric GARCH extension, data-driven local polynomial smoother, long memory, GARCH models, Value at Risk, Expected Shortfall, traffic light test, backtesting, Basel III, market risk

1 Introduction

The progressive liberalization of financial systems and the concomitant global integration of heterogeneous capital markets have fundamentally transformed the architecture of modern financial systems, fostering more intricate information transmission mechanisms and amplifying systemic interdependencies. As a consequence, the global financial landscape has become more susceptible to novel forms of speculative dynamics and contagion-driven crises, distinct from traditional macroeconomic instability patterns. Prompted by the 2007 global financial crisis, the Basel Committee on Banking Supervision (BCBS) introduced new guidelines for measuring market risk in banks' trading books through the "Fundamental Review of the Trading Book (FRTB)" framework (BCBS, 2013). The review was essential as the financial crisis had highlighted limitations in the Value at Risk (VaR) method (Morgan, 1996), widely used by banks to quantify market risk. In response, the BCBS recommended replacing VaR with Expected Shortfall (ES) (Acerbi & Tasche, 2002), a more coherent risk measure, under the FRTB. The BCBS also suggested using the established "traffic light tests" to backtest the quality of ES models (BCBS, 1996; 2016). However, ES backtesting is more challenging because it considers not only the frequency but also the magnitude of unexpected losses. Given their ability to model time-varying volatility with remarkable precision, GARCH-type models have emerged as a cornerstone in empirical finance for capturing the dynamic nature of conditional variance processes. Building on the seminal work of Engle (1982) and Bollerslev (1986), these models have been extensively employed across diverse domains of financial econometrics. Applications include the valuation of primary financial assets under stochastic volatility (Merton, 1980), the pricing of derivative securities (Wiggins, 1987), the empirical measurement of information flows and market efficiency (Conrad et al., 1991), and the formulation of optimal hedging strategies (Baillie & Myers, 1989). Their flexibility - reflected in asymmetric and multivariate extensions - has further consolidated their status as indispensable tools for both theoretical inquiry and applied financial research. More recently, modeling long-range dependence in volatility has become a key topic in financial econometrics. Empirical studies suggest that long memory GARCH (LM-GARCH) models are highly effective in forecasting the conditional volatility of asset returns, often surpassing short memory GARCH models (see, e.g., Giot & Laurent, 2003; Degiannakis, 2004; Tang & Shieh, 2006; Grané & Veiga, 2008; Härdle & Mungo, 2008; Baillie & Morana, 2009; Demiralay & Ulusoy, 2014; Aloui & Ben Hamida, 2015; Royer, 2022; Hanke et al., 2025). Ding et al. (1993), Ding and Granger (1996), Andersen and Bollerslev (1997), Andersen et al. (1999), and Cotter (2005), among others, provided evidence for long memory in the autocorrelations of absolute or squared observations in financial time series. Building on the introduction of the fractional ARIMA (FARIMA) model by Granger and Joyeux (1980), Baillie et al. (1996) introduced the fractionally integrated GARCH (FIGARCH) model, which effectively captures long-term volatility patterns in various financial data sets (e.g., Bollerslev & Mikkelsen, 1996; Tse, 1998; Beine et al., 2002; Baillie & Morana, 2009). In contrast, some suggest that these long-term dynamics may arise from deterministic structural shifts in unconditional variance (e.g., Lamoureux & Lastrapes, 1990; Mikosch & Stărică, 2004). Most empirical studies employing GARCH-type models rely on the assumption of covariance stationarity, implying time-invariant first- and second-order parameters. However, growing evidence suggests that the statistical properties of financial return series evolve over time, often driven by macroeconomic conditions, structural breaks, and regime shifts. Consequently, conventional parametric GARCH frameworks are extended to account for non-stationary dynamics that better capture the evolving nature of financial markets. This paper, therefore, centers on applying new semiparametric (long memory) GARCH (Semi-GARCH) models, which fit within a broader category of non-stationary volatility models, as discussed in Sucarrat (2019). For example, Beran and Ocker (2001) found evidence of a non-constant deterministic scale function in volatility series using the semiparametric fractional autoregressive (SEMIFAR) model (Beran & Ocker, 1999). Similarly, Feng (2004) observed that conditional heteroskedasticity often coincides with volatility shifts in financial return series. While conditional heteroskedasticity usually does not violate covariance stationarity, volatility-level changes often indicate at best local stationarity. A non-stationary process can be transformed into weak stationarity by removing its deterministic component, as shown by Feng (2004) and Van Bellegem and Von Sachs (2004), where the time-varying unconditional variance is estimated through kernel smoothing of the squared residuals, assuming volatility decomposes multiplicatively into conditional and unconditional components that change gradually over time. Engle and Rangel (2008) and Brownlees and Gallo (2010) adopted similar multiplicative decompositions using exponential quadratic and penalized B-splines, respectively. Mazur and Pipień (2012) developed the almost periodically correlated (APC-) GARCH model, using the Flexible Fourier Form by Gallant (1981, 1984) to parameterize the scaling function. More recently, Amado and Teräsvirta (2014) introduced the time-varying GARCH model under this framework (see also Amado & Teräsvirta, 2008; 2013; 2017), highlighting the empirical and practical significance of accounting for deterministic shifts in the unconditional variance of financial returns.

This paper introduces and applies several semiparametric short and long memory GARCH models under a variety of different innovation distributions. Similar to Letmathe et al. (2022), an adapted SEMIFAR algorithm proposed by Beran and Feng (2002a) to estimate the time-varying unconditional variance is used, employing a local polynomial estimator. After removing the deterministic component from the data, parametric GARCH models are fitted to the residuals, now approximately stationary. We demonstrate practical applications of this approach by analyzing daily returns of the 20 most popular and biggest US stocks. Additionally, our models are used in quantitative risk measurement, as illustrated by one-step-ahead, out-of-sample VaR and ES forecasts with a forecast horizon of one year, aligning with recent BCBS recommendations (BCBS, 2019). Comparative studies show that Semi-GARCH models are a compelling alternative to traditional parametric GARCH models.

The paper is organized as follows: Section 2 reviews established short and long memory GARCH models, including the fractionally integrated Log-GARCH (Geweke, 1986; Pantula, 1986; Milhøj, 1987; Feng et al., 2020) and EGARCH. Section 3 presents the novel specifications of the EGARCH family (EGF), namely the modulus-transformed Log-GARCH model (Feng et al., 2023) and the modified EGARCH (Hanke et al., 2025; Feng et al., forthcoming). In Section 4, the semiparametric estimation process is briefly outlined and the practical application for rolling forecasts is discussed. Section 5 details the application of our models to VaR and ES calculations. Empirical findings are provided in Section 6, and conclusions are drawn in Section 7.

2 Parametric Volatility Modeling

2.1 Established Models

Let $r_t^*, t = 1, \dots, n$ denote logarithmic returns with $E(r_t^*) = \mu_{r^*}$. Then the centralized returns are defined as $r_t = r_t^* - \mu_{r^*}$. Following Engle (1982) and Bollerslev (1986), a standard GARCH(p, q)-process specification for the time-varying conditional variance σ_t^2 is given by:

$$r_t = \sigma_t \varepsilon_t \quad \varepsilon_t \sim i.i.d.(0,1) \tag{1}$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \phi_i r_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{2}$$

where ε_t is a series of i.i.d. innovations with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t^2 - 1) = 0$. Let $\mathcal{F}_t = \sigma(\{r_t : s < t\})$ be a sigma algebra representing the history of the return process up to time t, so that $(\mathcal{F}_t)_{t \in \mathbb{Z}}$ is a natural filtration (McNeil et al., 2015). Then the random variable $var(r_t | \mathcal{F}_{t-1}) = \sigma_t^2$ is the conditional variance of r_t. Different parametric volatility models arise from alternative specifications of σ_t^2. Here, the non-negativity conditions $\phi_i \geq 0$ for $\forall i \in$

$\{1, \dots, p\}$, and $\beta_j \geq 0$ for $\forall j \in \{1, \dots, q\}$, with $p \geq 0$ and $q \geq 0$, are required. When $q = 0$ the model reduces to ARCH(p), and when $p = q = 0$ to homoscedastic white noise (Bollerslev, 1986). For strict stationarity of GARCH processes see e.g. Nelson (1990), Bougerol and Picard (1992), and Ling and McAleer (2002).

Ding et al. (1993) first documented long-range dependence in the volatility of S&P 500 daily returns, challenging the earlier view that volatility dynamics exhibit purely exponential decay. Subsequent studies confirmed similar persistence patterns across a wide range of financial assets, including intraday and high-frequency returns (e.g., Ding & Granger, 1996; Andersen & Bollerslev, 1997; Andersen et al., 1999; Cotter, 2005). This motivated the development of LM-GARCH specifications. Baillie et al. (1996) proposed the FIGARCH model, which introduces the fractional differencing operator $(1-B)^d$ to the AR polynomial, yielding hyperbolic decay in the autocorrelation of squared innovations - rather than the exponential decay implied by standard GARCH - allowing the model to capture the pronounced persistence commonly observed in financial volatility. According to Hosking (1981b) and Granger and Joyeux (1980), the fractional differencing operator $(1 - B)^d$ can be expressed as

$$(1 - B)^d = \sum_{k=0}^{\infty} \theta(d) B^k \tag{3}$$

where

$$\theta(d) = (-1)^k \frac{\Gamma(d + 1)}{\Gamma(k + 1)\Gamma(d - k + 1)} \tag{4}$$

$d \in (-0.5, 0.5)$ and $\Gamma(\cdot)$ denotes the Gamma function. This expansion generalizes the conventional differencing operator to fractional orders, allowing for intermediate degrees of persistence between stationarity and non-stationarity. The FIGARCH(p, d, q) is defined as:

$$\sigma_t^2 = \widetilde{\omega} + \left(1 - \beta^{-1}(B)\phi(B)(1 - B)^d\right) r_t^2 \tag{5}$$

with $d \in (0,1)$, $\widetilde{\omega} = \omega\beta^{-1}(B), \phi(B) = 1 - \sum_{j=1}^{p} \phi_i B^i$ AR(p), $\beta(B) = 1 - \sum_{j=1}^{q} \beta_j B^j$ MA(q) polynomials, and the other parameters are defined as in the GARCH model. It is assumed that both $\phi(B)$ and $\beta(B)$ have all roots lying outside the unit circle, ensuring the invertibility of the lag polynomials. The FIGARCH encompasses the GARCH model as a special case when $d = 0$ and the IGARCH model when $d = 1$. Although the FIGARCH process is not covariance stationary, Baillie et al. (1996) argue that it remains strictly stationary and ergodic for $d \in [0,1]$. More recently, Giraitis et al. (2018) established necessary and sufficient conditions for the existence of a covariance-stationary solution to the FIGARCH($0, d, 0$) model (with zero intercept) when $d \in (0, 0.5)$. Furthermore, the FIGARCH specification is closely related to the covariance-stationary LMGARCH process proposed by Karanasos et al. (2004).

The APARCH(p, q) model, introduced by Ding et al. (1993), which imposes a Box-Cox power transformation to both the volatility process and the asymmetric absolute residuals to deal with the problem of asymmetric responses to positive and negative shocks (leverage effect discovered by Black (1976)). The APARCH(p, q) is defined as:

$$\sigma_t^\delta = \omega + \sum_{i=1}^{p} \phi_i (|r_{t-i}| - \kappa r_{t-i})^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta \tag{6}$$

where $\delta \geq 0$ measures the degree of Box-Cox power transformation, $|\kappa_i| < 1$ is the asymmetry parameter, which takes a positive value if the leverage effect exists, and the other parameters align with the GARCH definition. The APARCH model constitutes a comprehensive framework nesting the ARCH model (Engle, 1982), the GARCH model (Bollerslev, 1986), the Log-ARCH model (Geweke, 1986; Pantula, 1986), the GJR-GARCH model (Glosten et al., 1993), and the Threshold GARCH (TGARCH) model (Zakoïan, 1994), among others. The long memory extension was introduced by Tse (1998) and is defined as:

$$\sigma_t^\delta = \widetilde{\omega} + \left(1 - \beta^{-1}(\mathrm{B})\phi(B)(1-B)^d\right)(|r_t| - \kappa r_t)^\delta \tag{7}$$

where all parameters are defined as in the APARCH model and d is the long memory parameter. When $d = 0$, the APARCH model admits an ARMA-type representation in terms of σ_t^δ. The FIAPARCH model thus generalizes several existing volatility frameworks: it nests the FIGARCH model when $\delta = 2$ and $\kappa = 0$, and it further encompasses additional long memory volatility specifications as special cases, two of which are defined in Ayensu et al. (forthcoming) as the FITGARCH and the FI-GJR-GARCH.

As a baseline model of the overarching EGF (e.g., Feng et al., forthcoming; Hanke et al., 2025), the EGARCH introduced by Nelson (1991) also accounts for leverage effects and is defined as:

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^{p} \phi_i g_E(\varepsilon_{t-i}) + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2 \tag{8}$$

where $g_E(\varepsilon_t) = \kappa\varepsilon_t + \gamma[|\varepsilon_t| - E(|\varepsilon_t|)]$ is a non-degenerate, strict and weakly stationary, ergodic i.i.d. zero-mean error process with finite variance (if κ and γ both do not equal zero), called the news impact function (see, e.g., Nelson, 1991; Bollerslev & Mikkelsen, 1996) that allows the volatility process $\{\sigma_t^2\}$ to respond asymmetrically to shocks, and ω, ϕ_i and β_j are real-valued coefficients. For practical implementation and convenience, the ARMA representation of the EGARCH is defined as:

$$\ln\sigma_t^2 = \omega + \phi^{-1}(B)\psi(B)g_E(\varepsilon_{t-1}) = \omega + \sum_{k=0}^{\infty}\beta_k g_E(\varepsilon_{t-k-1}) \tag{9}$$

where $\omega = E(ln\,\sigma_t^2)$ is a real number, the infinite moving average coefficients $\beta_k = \phi^{-1}(B)\psi(B)$ satisfy $\sum_{k=0}^{\infty}|\beta_k| < \infty$, and $E(g_E(\varepsilon_{t-k-1})) = 0$ and $0 < var[g_E(\varepsilon_t)] = E[g_E{}^2(\varepsilon_t)] < \infty$ so that $\ln\sigma_t^2$ is a stationary process. Hence, $\ln r_t^2$ can be represented as a linear signal-plus-noise model, as outlined in Equation (3) of Zaffaroni (2009). The autoregressive and moving-average lag polynomials $\phi(B)$ and $\psi(B)$ have to satisfy the standard root conditions required for the existence of a stationary solution. The stationarity and ergodicity of the process $\{\ln\,\sigma_t^2\}$ are established under the conditions outlined in Theorem 2.1 of Nelson (1991). Moreover, under the assumptions of Theorem 2.2 therein, both $\{r_t\}$ and $\{\sigma_t^2\}$ possess finite moments of all orders when the innovations ε_t follow a Generalized Error Distribution (GED) with tail-thickness parameter $v > 1$, a class that includes the normal distribution ($v = 2$). Conversely, for heavy-tailed distributions such as the Student-t distribution and GED with $v < 1$, the existence of finite moments generally fails. As a remedy, Feng et al. (forthcoming) introduced the Asymmetric Laplace (AL) distribution - parameterized by a density order parameter $P \geq 0$ - for the innovations of models within the EGF. According to Proposition 1 in their study, for $P > 0$, the existence of finite moments $E[\sigma_t^m] < \infty$ and $E[r_t^{2m}] < \infty$ is guaranteed if the moment generating function (mgf) of ε_1 exists for some $\epsilon \in (m \cdot b_{\max} \cdot \pi_1, m \cdot b_{\max} \cdot \pi_2)$, where $m \geq 1$, $b_{\max} = \sup\{\beta_k\}$, $\pi_1 = \min(\kappa - \gamma, \kappa + \gamma)$, and $\pi_2 = \max(\kappa - \gamma, \kappa + \gamma)$. For $P = 1, \ldots, 8$ the mgf of ϵ exists within neighborhoods $(-2,2), \ldots, (-3\sqrt{2}, 3\sqrt{2})$, with the range increasing by $\sqrt{2(P+1)}$. In the special case $P = 0$, the AL distribution simplifies to the standard Laplace distribution, which corresponds to a GED with $v = 1$ and is a member of the Sargan family of distributions (Missiakoulis, 1983). In this case, the EGARCH process may exhibit at most weak (wide-sense) stationarity, as the mgf exists only in the narrower domain $(-\sqrt{2}, \sqrt{2})$. The findings of Feng et al. (forthcoming) considerably expand the applicability of the EGARCH model by providing distributional flexibility for the innovations. For further insights into the autocovariance structure and higher-order moment properties of EGARCH models, see Breidt et al. (1998), Surgailis and Viano (2002), and He et al. (2002). Bollerslev and Mikkelsen (1996) introduced the long memory extension by defining the FIEGARCH(p, d, q) as

$$\ln\sigma_t^2 = \omega + \phi^{-1}(B)(1-B)^{-d}\psi(B)g_E(\varepsilon_{t-1}) = \omega + \sum_{k=0}^{\infty}\beta_k g_E(\varepsilon_{t-k-1}) \tag{10}$$

Here, $-0.5 < d < 0.5$ (ensuring covariance stationarity and invertibility of $\{\ln\ \sigma_t^2\}$) and $\beta_k = \phi^{-1}(B)(1-B)^{-d}\psi(B)$ are defined such that $\sum_{k=0}^{\infty}|\beta_k| = \infty$ but $\sum_{k=0}^{\infty}\beta_k^2 < \infty$, ensuring long memory dependence with finite variance. Here, the polynomial $\phi(z) = 0$ has all roots located outside the unit circle, guaranteeing model invertibility. When $d = 0$, the fractional differencing term vanishes, recovering the conventional EGARCH structure. The autocovariance properties and long memory behavior of the FIEGARCH model have been rigorously investigated in Surgailis and Viano (2002), who provide a comprehensive characterization of its dependence structure and asymptotic properties.

The Log-GARCH model - originally introduced by Geweke (1986), Pantula (1986), and Milhøj (1987), and later refined by Francq et al. (2013) - has gained significant attention in the literature as the first attempt to model σ_t^2 as log-linear that can be modeled by AR or ARMA processes. In particular, Francq and Sucarrat (2018) define the Log-GARCH(p, q) process through

$$\ln\sigma_t^2 = \omega + \sum_{i=1}^{p}\phi_i \ln r_{t-i}^2 + \sum_{j=1}^{q}\beta_j \ln\sigma_{t-j}^2 \tag{11}$$

which aligns with the EGARCH formulation by relaxing non-negativity constraints on model parameters, yet differs in assuming log-linearity for both $\{r_t^2\}$ and $\{\sigma_t^2\}$, whereas the EGARCH assumes log-linearity only for $\{\sigma_t^2\}$. An asymmetric extension - the asymmetric Log-GARCH of Francq et al. (2013) - incorporates leverage effects. However, the basic Log-GARCH model can perform poorly in the presence of zero or near-zero returns. Several remedies have been proposed, including the use of demeaned returns, replacing zero values with optimal estimates (Sucarrat & Escribano, 2018), or applying a log-modulus transformation to the innovations (Feng et al., 2023). Under the moment condition $E[(\ln \varepsilon_1^2)^2] < \infty$, the process $\{\ln\ \sigma_t^2\}$ satisfies an ARMA-type representation:

$$\ln\sigma_t^2 = \alpha + (\phi^{-1}(B)\psi(B) - 1)\eta_t = \alpha + \sum_{k=1}^{\infty}\beta_k\,\eta_{t-k} \tag{12}$$

where $\eta_t = \ln\ \varepsilon_t^2 - E(\ln\ \varepsilon_1^2)$ is an i.i.d. sequence with $E[\eta_t] = 0$ and $E[\eta_t^2] = 1$. Stationarity and invertibility are guaranteed under the usual root conditions for these polynomials (Francq & Sucarrat, 2018). These representations confirm earlier findings (Nelson, 1991; Francq et al., 2017; Feng et al., 2020) showing that the Log-GARCH is a special case of the EGARCH when $g_E(\varepsilon_t) = \eta_t$. Furthermore, under Theorem 2 of Feng et al. (2020), the processes $\{\sigma_t^2\}$ and $\{r_t\}$ are strictly stationary and ergodic. Let $\beta_{\max} = \max(\beta_k) \geq 1$ and $\beta_{\min} = \inf(\beta_k) \leq 0$, and assume $E[\varepsilon_1^m] < \infty$ for some $m > 0$. According to Feng et al. (2020), covariance stationarity of $\{r_t\}$ holds if $m > 2\beta_{\max}$ and $\beta_{\min} > -0.5$, while finite fourth moments require $m > 4\beta_{\max}$ and $\beta_{\min} > -0.25$. Consequently, $\{r_t\}$ can exhibit finite moments of any order provided that ε_1 admits finite moments of all orders and $\beta_k \geq 0$ for all k. Interestingly, the moment conditions for the Log-GARCH model are less restrictive than those for the EGARCH, as they do not require the existence of the mgf. However, an additional constraint arises on negative coefficients, since $E(\varepsilon_t^m) = \infty$ for typical innovation distributions when $m \leq -1$. In the simple Log-GARCH(1,1) case, finite fourth moments $E(r_t^4) < \infty$ exist if $\phi_1 > -0.25$ and $E(|\varepsilon_1|^m) < \infty$ for some $m \in [4,8)$ (Feng et al., 2020, Proposition 1). The Fractionally Integrated Log-GARCH (FILog-GARCH)(p, d, q) model proposed by Feng et al. (2020) extends the Log-GARCH framework to capture long memory behavior in volatility. The model represents a special case of the FIAPARCH for $\delta \to 0$ and is defined by:

$$\ln \sigma_t^2 = \omega + [\phi^{-1}(B)(1-B)^{-d}\psi(B) - 1]\,\eta_t = \omega + \sum_{k=1}^{\infty} \beta_k \eta_{t-k} \tag{13}$$

for $d \in [0,0.5)$, where $\beta_k = [\phi^{-1}(B)(1-B)^{-d}\psi(B) - 1]$ and the lag polynomials $\phi(z)$ and $\psi(z)$ share no common factors and have all roots outside the unit circle, ensuring stationarity and invertibility. All other parameters parallel those in the Log-GARCH model. In this model, $\{\ln r_t^2\}$ is assumed to follow a linear FARIMA (Fractional Autoregressive Integrated Moving Average) process, as originally introduced by Hosking (1981a) and Granger and Joyeux (1980). When $d = 0$, the FILog-GARCH reduces to the standard Log-GARCH specification. Theorem 2 of Feng et al. (2020) establishes the existence of strict stationary solutions to the FILog-GARCH process, while Corollary 1 therein provides necessary and sufficient conditions for covariance stationarity and the existence of higher-order moments. Specifically, the process possesses finite moments of any order if and only if the innovations ε_t admit the corresponding finite moments and $\beta_k \geq 0$ for all k. A key distinction between the Log-GARCH and the EGARCH models lies in their moment requirements. Whereas the EGARCH model relies on the existence of the mgf of ε_t, the Log-GARCH requires only the existence of finite moments of ε_t up to a given order, offering greater flexibility when modeling heavy-tailed innovation distributions. However, the Log-GARCH does not capture leverage effects, imposes constraints on negative coefficients in its MA(∞) representation, and may exhibit inefficiencies when dealing with zero or near-zero returns.

2.2 Recent members of the EGARCH family

To address certain limitations of the Log-GARCH framework, Feng et al. (2023) recently substituted the standard logarithmic transformation with the modulus–log transformation $\zeta_t =$

sign $(\varepsilon_t)\ln(|\varepsilon_t|+1)$, originally introduced by John and Draper (1980) as the Box–Cox transformation with power parameter zero (Box and Cox (1964) on general shifted log-transformations with the power parameter set to zero), improving robustness in the presence of zero or near-zero returns while allowing for potential leverage effects. The asymmetric and magnitude terms are modeled using centered absolute log-transformations, $|\zeta_t| - E(|\zeta_t|)$, ensuring $E(\zeta_t) = 0$ under symmetric innovation distributions. This framework generalizes to a new class of log-modulus volatility models called modulus-adjusted Log-GARCH (MLog-GARCH) defined by

$$\ln \sigma_t^2 = \omega + \phi^{-1}(B)\psi(B) g_{MLog}(\varepsilon_{t-1}) = \omega + \sum_{k=0}^{\infty} \beta_k g_{MLog}(\varepsilon_{t-k-1}) \tag{14}$$

where $g_{MLog}(\varepsilon_t) = \kappa[\zeta_t(\varepsilon_t) - E(\zeta_t)] + \gamma[|\zeta_t(\varepsilon_t)| - E(|\zeta_t|)]$ is an i.i.d. zero-mean process with $0 < var(g_{MLog}(\varepsilon_t)) < \infty$. As in standard log-volatility models, $\omega = E[\ln(\sigma_t^2)]$ holds by construction. Unlike the Log-GARCH, $g_{MLog}(\varepsilon_t)$ is derived from the EGARCH structure and expressed directly in terms of the innovations rather than observed returns. Thus, the model forms another asymmetric exponential specification. Feng et al. (2023) highlight the distinctive limiting behavior of $\ln(1+|x|)$, which approaches $|x|$ as $x \to 0$ and $\ln |x|$ as $|x| \to \pm\infty$. Consequently, MLog-GARCH retains EGARCH-like flexibility near zero while exhibiting tail dynamics akin to Log-GARCH models. Since both sides of the volatility equation are on a logarithmic scale, the existence conditions for finite moments mirror those of Log-GARCH models and are considerably weaker than those of EGARCH. In particular, $E[\exp(g_{MLog}(\varepsilon_{t-k-1}))]^m < \infty$ whenever $E(|\varepsilon_1|^m) < \infty$ for any $m > 0$. Furthermore, the model removes the non-negativity restriction on Log-GARCH coefficients, as $(|\varepsilon_t| + 1)^m \in (0,1)$ for $m < 0$. As near-zero innovations are handled in a nearly linear way, the transformation also ensures numerical stability for near-zero innovations, since $\zeta_t = 0$ when $\varepsilon_t = 0$ is well-defined. Thus, the MLog-GARCH model integrates the robustness of Log-GARCH with the asymmetry-capturing capability of EGARCH. Assuming κ and γ are not simultaneously zero, $g_{MLog}(\varepsilon_t)$ forms an i.i.d. zero-mean process with finite variance, implying - by Theorem 2.1 of Nelson (1991) - that $\ln(\sigma_t^2)$ is strictly stationary and ergodic. Unlike EGARCH, however, $\exp[g_{MLog}(\varepsilon_{t-k-1})]$ appears in a power rather than an exponential form, providing enhanced moment stability and analytical tractability. As shown by Feng et al. (2023), the MLog-GARCH finite-moments requirements parallel those of the Log-GARCH model and holds, for instance, when innovations follow a Student-t distribution with sufficiently high degrees of freedom. In contrast, EGARCH models generally lack this moment-preserving characteristic, granting the

MLog-GARCH a distinct theoretical advantage in ensuring well-behaved higher-order dynamics. Moreover, the MLog-GARCH framework circumvents the numerical instability associated with zero or near-zero observations and permits negative coefficient values without violating model admissibility, enhancing both flexibility and robustness in empirical applications. The long memory extension of the MLog-GARCH (FIMLog-GARCH) was recently introduced by Feng et al. (2023) and is defined as

$$\ln \sigma_t^2 = \omega + (1-B)^{-d}\phi^{-1}(B)\psi(B)g_{MLog}(\varepsilon_{t-1}) = \omega + \sum_{k=0}^{\infty} \beta_k g_{MLog}(\varepsilon_{t-k-1}) \tag{15}$$

where $g_{MLog}(\varepsilon_t)$ is as previously defined, and $d \in (0, 0.5)$ denotes the fractional integration parameter governing long memory. Under the regularity conditions established by Feng et al. (2023), the FIMLog-GARCH process admits finite moments of any desired order, provided that the innovation distribution possesses moments of the corresponding order - a property shared with the FILog-GARCH but not with the FIEGARCH. However, as in the FIEGARCH framework, the invertibility of $\ln(\sigma_t^2)$ in the (FI-)MLog-GARCH remains an open theoretical issue. Its verification requires analyzing the convergence of a highly intricate recursive expansion in terms of r_t, involving an infinite product of infinite products (Feng et al., forthcoming), which lies beyond the scope of the present study.

A natural generalization of the previously discussed exponential volatility frameworks leads to the EGF (Feng et al., forthcoming). This subclass modifies the conventional EGARCH specification by replacing the leverage term with a modulus-based transformation. It is important to note that in both cases, the MA filters are retained as in the standard EGARCH formulation, ensuring consistency in the underlying volatility dynamics. Extending the conventional EGARCH specification by introducing distinct power parameters for the asymmetric and magnitude terms yields a Double-Power Modulus EGARCH class, defined as

$$\ln \sigma_t^2 = \omega + \phi^{-1}(B)\psi(B)g_{DP}(\varepsilon_{t-1}) = \omega + \sum_{k=0}^{\infty} \beta_k g_{DP}(\varepsilon_{t-k-1}) \tag{16}$$

where $g_{DP}(\varepsilon_t)$ is a measurable function of ε_t satisfying $E[g_{DP}(\varepsilon_t)] = 0$ as well as $0 < var\ [g_{DP}(\varepsilon_t)] < \infty$. Hence, $\{g_{DP}(\varepsilon_t)\}$ forms an i.i.d. zero-mean sequence with finite variance, implying that $\{\ln\ \sigma_t^2\}$ is stationary and that $\{\ln\ r_t^2\}$ corresponds to a linear signal-plus-noise representation (Zaffaroni, 2009). The general framework nests several models, including the MEGARCH and MLog-GARCH (Feng et al., 2023), and constitutes a specific subclass of the broader exponential volatility models discussed by Zaffaroni (2009) as well as those discussed by Surgailis and Viano (2002). The function $g_{DP}(\varepsilon_t)$ is specified as

$$g_{DP}(\varepsilon_t) = \kappa\{g_a(\varepsilon_t) - E[g_a(\varepsilon_t)]\} + \gamma\{g_m(\varepsilon_t) - E[g_m(\varepsilon_t)]\} \tag{17}$$

with real-valued coefficients κ and γ. The novelty of the MEGARCH model lies in its use of distinct power parameters for the leverage and magnitude components, thereby overcoming the inherent limitations of traditional specifications that rely on a single power parameter. Setting $g_a(\varepsilon_t) = \varepsilon_t$ and $g_m(\varepsilon_t) = |\varepsilon_t|$ recovers the standard EGARCH model of Nelson (1991). While any suitable transformation of ε_t can be applied, suggested transformations are:

$$\begin{aligned} g_{1,a}(\varepsilon_t) &= \text{sign}\,(\varepsilon_t)\,|\,\varepsilon_t\,|^{p_a}/p_a, \qquad g_{2,a}(\varepsilon_t) = \text{sign}\,(\varepsilon_t)((|\,\varepsilon_t\,| + 1)^{p_a} - 1)/p_a, \\ g_{1,m}(\varepsilon_t) &= |\,\varepsilon_t\,|^{p_m}/p_m, \qquad g_{2,m}(\varepsilon_t) = ((|\,\varepsilon_t\,| + 1)^{p_m} - 1)/p_m, \end{aligned} \tag{18}$$

where p_a and p_m are real-valued power parameters. The error transformation $g_{1,\cdot}$ performs a power transformation and $g_{2,\cdot}$ performs both a modulus and a power transformation simultaneously. When $p_i \to 0$, $i \in \{a, m\}$, these transformations reduce to logarithmic forms, $g_{1,a}(\varepsilon_t) = sign(\varepsilon_t)\ln(|\varepsilon_t|)$ and $g_{2,a}(\varepsilon_t) = sign(\varepsilon_t)\ln(|\varepsilon_t| + 1)$. The Log-GARCH model arises as a special case when $\kappa = 0$, $g_{1,m}(\varepsilon_t)$ is used and $p_m = 0$.

Although introducing two distinct power parameters increases flexibility, empirical evidence suggests that model selection criteria rarely favor such added complexity (Feng et al., forthcoming). Thus, we propose fixing $p_a = 0$ for the sign effect and $p_m = 1$ for the magnitude term, yielding a Modified EGARCH (MEGARCH) model that combines the EGARCH (Nelson, 1991) and MLog-GARCH formulations:

$$\ln \sigma_t^2 = \omega + \phi^{-1}(B)\psi(B)g_{ME}(\varepsilon_{t-1}) = \omega + \sum_{k=0}^{\infty} \beta_k g_{ME}(\varepsilon_{t-k-1}) \tag{19}$$

where $g_{ME}(\varepsilon_t) = \kappa\,[\zeta_t(\varepsilon_t) - E(\zeta_t(\varepsilon_t))] + \gamma[(|\varepsilon_t|) - E(|\varepsilon_t|)]$. The existence conditions for finite moments coincide with those of the EGARCH model since the tail behavior of $g_{ME}(\varepsilon_t)$ is dominated by the magnitude term. Empirical tests across diverse financial return series confirm that this double-power modulus structure performs competitively in practice (Hanke et al., 2025). Moreover, the modulus transformation is well-defined for all real powers, including negative values, enhancing model flexibility. Allowing fractional exponents $p_s < 1$ can further improve empirical fit, though the interpretation of negative power parameters remains an open theoretical issue (Feng et al., forthcoming). Introducing fractional differencing into the AR polynomial yields the FIMEGARCH(*p, d, q*) model:

$$\ln \sigma_t^2 = \omega + (1 - B)^{-d}\phi^{-1}(B)\psi(B)g_{ME}(\varepsilon_{t-1}) = \omega + \sum_{k=0}^{\infty} \beta_k g_{ME}(\varepsilon_{t-k-1}) \tag{20}$$

which combines the long memory properties of the FIEGARCH (Bollerslev & Mikkelsen, 1996) with the modulus-log dynamics of the FIMLog-GARCH (Feng et al., 2023). Assuming $d < 0.5$ and regularity conditions consistent with Theorem 2.1 of Nelson (1991), $\{\ln\ \sigma_t^2\}$ is strictly stationary and ergodic, with stationary solutions depicted in Hanke et al. (2025) and Feng et al. (forthcoming). Because $1 \leq (|\ \varepsilon_t\ | + 1) \leq e^{|\varepsilon_t|}$, the conditions for finite moments mirror those of the (FI-)EGARCH, allowing negative coefficients. Thus, Theorem 2.1 of Nelson (1991) ensures finite moments of arbitrary order under standard regularity assumptions. However, unlike in the MLog-GARCH framework, the existence of finite moments in the MEGARCH model do not exist for t-distributed or GED innovations with degrees of freedom $\nu < 1$ (Feng et al., forthcoming). In the limiting case of $\nu = 1$, corresponding to the Laplace distribution, Feng et al. (forthcoming) show that the MEGARCH model - and, by extension, its long memory variant - may exhibit covariance stationarity under certain parameter constellations. As in the (FI)EGARCH literature, the invertibility of $\ln\ \sigma_t^2$ remains unresolved and lies beyond the scope of this study.

3 Semiparametric Extension of GARCH Models

Empirical evidence shows that observed volatility often reflects both conditional heteroskedasticity and a slowly varying unconditional component. Standard GARCH models capture the former but assume a constant unconditional variance, an assumption frequently violated when structural changes or persistent volatility shifts occur. Semiparametric extensions address this limitation by decomposing $\{\sigma_t^2\}$ into a dynamic conditional part and a smoothly evolving unconditional scale (Feng, 2004; Van Bellegem & von Sachs, 2004). This places them within the broader class of non-stationary volatility models discussed by Sucarrat (2019). Beran and Ocker (2001) demonstrated the presence of deterministic scale functions in a variety of time series settings, providing empirical evidence of smoothly varying volatility structures. Building on this idea, Feng (2004) observed that financial return series frequently exhibit both conditional heteroskedasticity and time-varying scale behavior and introduced the Semi-GARCH model. This framework was later extended to the long memory context by Letmathe et al. (2022), enabling the modeling of persistent volatility dynamics under non-stationary scale variation. In this paper, the unconditional component is estimated via local polynomial smoothing - rather than kernel smoothing proposed by Feng (2004) and Van Bellegem and von Sachs (2004) - allowing for more flexible and a lower-bias approximation of gradual volatility dynamics.

3.1 Semiparametric EGF models

The basic idea is to multiplicatively decompose the volatility into a non-stationary and a stationary part by inserting a non-negative smoothing scale function $s(\tau_{t,n})$. The parametric GARCH models can then be applied to the scale-removed process. Let

$$r_t = s(\tau_{t,n})\sigma_t\varepsilon_t \tag{21}$$

Here, $\tau_{t,n} = (t - 0.5)/n$ describes the rescaled time.[1] The stationary part can then be defined as $\xi_t^2 = \frac{r_t^2}{s^2(\tau_{t,n})} = \sigma_t^2\varepsilon_t^2$. Assuming $\{\xi_t\}$ follows an EGARCH-type formulation:

$$\xi_t = \sigma_t\varepsilon_t \tag{22}$$

$$\ln\sigma_t^2 = \omega + \sum_{k=0}^{\infty}\beta_k g(\varepsilon_{t-k-1}) \tag{23}$$

r_t follows a Semi-GARCH model from the EGF. With different parametric formulations of σ_t^2 presented in Section 2, various Semi-GARCH models emerge under the condition that strictly stationary solutions exist for these models. Hence, within this semiparametric framework, the total volatility can be decomposed into three components: (i) a long-run, smoothly varying local scale function $s(\tau_{t,n})$; (ii) a short-run, conditionally heteroskedastic component σ_t; and (iii) the baseline volatility represented by the standard deviation of the innovation process ε_t.

Under the assumption that the coefficients β_k are square-summable, this semiparametric EGARCH framework naturally accommodates long memory behavior in the logarithmic conditional variance. This feature was first noted by Nelson (1991) within a stationary context. Accordingly, the proposed model generalizes and unifies the EGF of Feng et al. (forthcoming) within a semiparametric setting. When the coefficient sequence $\{\beta_k\}$ is absolutely summable (i.e., $\sum_{k=0}^{\infty}|\beta_k| < \infty$), thus ruling out integrated or long memory dynamics, the framework provides semiparametric EGARCH-type models with short memory.

To ensure that the models are well-defined, in the following it is assumed that $\{\xi_t^2\}$ is a log-linear process and $var(\xi_t) = 1$ and $E(ln\,\xi_t^2)$ exists, which implies that $\xi_t \neq 0$ almost surely.[2] Furthermore, let $y_t = \ln r_t^2$, $g(\tau_t) = \ln s^2(\tau_{t,n}) + \mu_{l\xi^2}$ and $z_t = \ln\xi_t^2 - \mu_{l\xi^2} = \ln\sigma_t^2 + \eta_t -$

[1] It should be noted that this specification of $\tau_{t,n}$ is essential for deriving the asymptotic properties of the local polynomial estimator of $s(\tau_{t,n})$ (see, e.g., Beran & Feng, 2002b). The rescaled time has been widely adopted (see, e.g., Robinson, 1989; Beran & Ocker, 2001; Beran & Feng, 2002a,b,c; Feng, 2004; Amado & Teräsvirta, 2013; Letmathe et al., 2022).

[2] That implies that the intercept term ω in non-stationary GARCH and APARCH models is fixed. Specifically, $\omega = 1 - \sum_{i=1}^{p}\alpha_i - \sum_{i=1}^{q}\beta_j$, and $\omega = 1 - \sum_{i=1}^{p}\alpha_i EZ_i - \sum_{j=1}^{q}\beta_j$, where $EZ_i = E[(|\,\varepsilon_t\,| - \kappa_i\varepsilon_t)^{\delta}]$, respectively. For non-stationary EGARCH-type models, it typically holds that $E[\ln(\sigma_t^2)] < 0$ by Jensen's inequality, which ensures the existence of $E[\ln\ \xi_t^2]$.

$(\mu_{l\xi^2} - \mu_{l\varepsilon^2})$, where $\mu_{l\xi^2} = E(\ln \xi_t^2)$ and $\eta_t = \ln \varepsilon_t^2 - \mu_{l\varepsilon^2}$, where $\mu_{l\varepsilon^2} = E(\ln \varepsilon_t^2)$. Due to the log-transformation of r_t^2, y_t can be represented by an additive regression model with a deterministic, non-parametric functional form:

$$y_t = g(\tau_{t,n}) + z_t \tag{24}$$

where $g(\tau_{t,n})$ is a nonnegative smooth trend function and $\{z_t\}$ is a zero-mean stationary process with autocovariance function $\gamma_z(l) = cov(z_t, z_t + l)$ that satisfies either (i) $\sum_{l=-\infty}^{\infty} |\gamma_z(l)| < \infty$, corresponding to a short memory process; or (ii) $\sum_{l=-\infty}^{\infty} |\gamma_z(l)| = \infty$, corresponding to a long memory process of $\{z_t\}$. The additive decomposition isolates the deterministic trend component from the stationary process, thereby allowing for a nonparametric estimation of $g(\tau_{t,n})$. As shown in Letmathe et al. (2022), z_t follows a FARIMA(p, q) process given by:

$$\phi(B)(1 - B)^d z_t = \psi(B)\, \eta_t \tag{25}$$

Subsequently, well-developed SEMIFAR algorithms are applicable for estimating $g(\tau_{t,n})$ and z_t (Beran & Feng, 2002a; Beran & Feng, 2002b; Beran & Feng, 2002c; Beran et al., 2015). One can see that now an estimator for the variance is given by $\hat{s}^2(\tau_{t,n}) = \hat{C}_s\, exp[\hat{g}(\tau_{t,n})]$ and $\hat{\sigma}_t^2 = \hat{C}_\sigma \exp[\hat{z}_t]$, where $C_s = \exp(-\mu_{l\xi^2})$ and $C_\sigma = \exp(\mu_{l\xi^2} - \mu_{l\varepsilon^2})$ can be estimated consistently by the scale-adjusted returns and by standardizing the innovations under the assumption that $E(\xi_t^4) < \infty$ and $var(\xi_t) = 1$.

In the Semi-FI-Log-GARCH model, the long memory parameter d is not affected by a power- or log-transformation (Surgailis & Viano, 2002; Feng et al., 2020). This necessary condition is not yet proven for the remaining conventional LM-GARCH models. However, the estimation procedure assumes that $\{\xi_t^2\}$ follows a log-linear structure and that the fractional integration parameter d remains invariant under the logarithmic transformation.

3.2 Related Approaches

Several prior studies have adopted similar methodologies to decompose the conditional variance σ_t^2 and to estimate its deterministic component $s(\tau_{t,n})$. For early work using the kernel estimator, see Altman (1990) as well as Hall and Hart (1990). Subsequent work introduced more structured semiparametric GARCH frameworks, notably the Semi-GARCH model of Feng (2004) and the time-modulated GARCH of Van Bellegem and von Sachs (2004). More recently, Jiang et al. (2021) also make use of the kernel estimator. Related approaches estimate the deterministic scale via multiplicative decompositions, such as the exponential-quadratic specification of Engle and Rangel (2008), penalized B-splines in Brownlees and Gallo (2010), the Flexible Fourier Form (Gallant, 1981; 1984) of Mazur and Pipień (2012), and the logistic transition functions used in the time-varying GARCH models of Amado and Teräsvirta (2008, 2013, 2014, 2017). Additional extensions include the GARCH-MIDAS model of Engle et al. (2013), the Box–Cox Semi-GARCH of Zhang et al. (2017), and penalized-spline estimators as in Krivobokova and Kauermann (2007) and Feng and Härdle (2025). Most recently, Feng and Peitz (2025) introduced object-driven smoothing to solve the non-stationarity problem. Collectively, these methods relax the assumption of time-invariant parameters and are designed to accommodate evolving volatility structures. Moreover, local polynomial regression provides an effective remedy for the boundary bias problem commonly encountered in traditional kernel-based estimation methods. Local polynomial regression inherently performs automatic kernel adaptation (Hastie & Loader, 1993), allowing most established results from kernel regression theory to be directly extended to this framework.

Building on this literature, the present paper - together with Letmathe et al. (2022) - advances the semiparametric modeling of volatility. Our framework explicitly incorporates long memory dynamics in the stochastic component, a feature central to financial volatility but largely ignored in earlier semiparametric specifications. Additionally, we estimate the deterministic scale function $s(\tau_{t,n})$ using a local polynomial smoother embedded within an adapted SEMIFAR algorithm (Beran & Feng, 2002a; Letmathe et al., 2023).

4 Estimation and practical implementation

The SEMIFAR model of Beran and Feng (2002c) provides a unified framework for disentangling a deterministic scale component from both short- and long-range dependence in the stochastic part of a time series. Estimation proceeds in two stages: a nonparametric step that recovers the slowly varying scale function, followed by a parametric step that estimates the fractional differencing and dependence parameters governing the stationary component. Beran et al. (2015) subsequently introduced the exponential SEMIFAR (ESEMIFAR) model, which extends this framework under the assumption that $\{\xi_t\}$ is log-linear, thereby enabling semiparametric modeling of volatility on a multiplicative scale.

4.1 Local Polynomial Smoothing

The estimation of $g^{(v)}(\tau_{t,n})$, the v-th derivative, is performed by a local polynomial estimator (see e.g., Beran & Feng, 2002a; Beran & Feng, 2002b; Beran & Feng, 2002c; Beran et al., 2013). Given the assumption that g is continuously differentiable at least $(l + 1)$ times in $\tau_{0,n}$,

$g(\tau_{t,n})$ can be approximated by a local polynomial of order l for $\tau_{t,n}$ in a neighborhood of $\tau_{0,n}$ by a simple application of the Taylor expansion (Beran & Feng, 2002b):

$$g(\tau_{t,n}) = \sum_{v=0}^{l} \frac{m^{(v)}(\tau_{0,n})}{v!} (\tau_{t,n} - \tau_{0,n})^{v} + R_l(\tau_{t,n}) \tag{26}$$

where $R_l(\tau_{t,n})$ represents the remaining approximation error. In the following, a kernel weighting function of order two with compact support on $[-1,1]$ is assumed (Gasser & Müller, 1979), which can be represented in polynomial form as follows:

$$K(u) = \begin{cases} \sum_{i=0}^{r} a_i u^{2i} & for\ |u| \leq 1 \\ 0 & for\ |u| > 1 \end{cases} \tag{27}$$

with $r \in \{0, 1, 2, \ldots\}$ including the uniform, Epanechnikov, bisquare, and triweight kernels and weights a_i such that $\int_{-1}^{1} K(u)du = 1$ holds. By solving the locally weighted least squares problem, an estimate $\hat{g}^{(v)}(\tau_{t,n})$ $(v \leq l)$ is obtained by:

$$Q = \sum_{i=1}^{n} \left[y_t - \sum_{j=1}^{l} \lambda_j (\tau_{t,n} - \tau_{0,n})^j \right]^2 K\left(\frac{\tau_{t,n} - \tau_{0,n}}{b}\right) \rightarrow min \tag{28}$$

where b indicates the bandwidth, which ensures through the Kernel function $K(\tau_{t,n} - \tau_{0,n}/b)$ that only observations in the neighborhood of $\tau_{0,n}$ are used in the minimization problem. Considering the stated formulas, it is easy to see that $\hat{g}^{(v)}(\tau_{0,n}) = v!\,\hat{\lambda}_v$, with $\hat{\lambda} = (\hat{\lambda}_0, \hat{\lambda}_1, \ldots, \hat{\lambda}_l)$ and $v = 1, 2, \ldots, l$. It is important to note that Equation (2.7) in Beran and Feng (2002c) demonstrates that $\hat{g}^{(v)}(\tau_{0,n})$ is a linear smoother with weight vector $w^{(v)}(\tau_{0,n}) = (w_1^{(v)}, \ldots, w_n^{(v)})^{\top}$, where $w_t^{(v)} = 0$ for $|\ \tau_{t,n} - \tau_{0,n}\ |> b$. This formulation underscores the local averaging property of the estimator, which remains valid irrespective of the short or long memory dependence structure of $\{z_t\}$. Define $m = l + 1$. Under Assumption 3 (i) below, $l - v \geq 1$, so $m \geq v + 2$ with $m - v$ being even (see e.g., Fan & Gijbels, 1996; Beran & Feng, 2002c). This condition guarantees that $\hat{g}^{(v)}$ is asymptotically equivalent to some kernel regression of order m and possesses automatic boundary correction, with the bias uniformly of order $O(b^{m-v})$. Assumption 3 (ii) further imposes a bandwidth–sample size trade-off ensuring the consistency of $\hat{g}^{(v)}$ even in the presence of long-range dependence. It also ensures that the variance of $\hat{g}^{(v)}$ converges to zero asymptotically, thereby establishing mean integrated squared error (MISE) consistency of the estimator over the entire support [0,1]. Under $E(\xi_t^4) < \infty$ the asymptotic variance of $\hat{g}^{(v)}$ is well defined and $\hat{C}_s$ can be estimated consistently.

To prevent boundary bias, we differentiate between interior and boundary points, where a point $\tau_{t,n}$ is said to be an interior point if it lies within $[b, 1-b]$, a left boundary point if it satisfies $\tau_{t,n} \in [0, b)$, and a right boundary point if $\tau_{t,n} \in (1-b, 1]$. Similar to Beran and Feng (2002b), a left boundary point can be expressed as $\tau_{t,n} = cb$ with $c \in [0,1)$, while for each interior point c is automatically set to 1. Beran and Feng (2002b) make four assumptions regarding the asymptotic properties of the proposed procedure:

Assumption 1. The weighting function $K(u)$ is a second-order kernel, i.e. a symmetric density, with compact support on [-1, 1].

Assumption 2. g is a function that can be continuously differentiated at least m times on [0, 1] with $m \geq v+2$ and $m-v$ is even.

Assumption 3. (i) $l-v$ is odd; (ii) the bandwidth fulfills: $h \to 0, (nh)^{1-2d}h^{2v} \to \infty, when\ n \to \infty$, with $d \in [0, 0.5)$.

Assumption 4. A local polynomial adjustment of order $p = k-1$ is used.

Under Assumptions 1 to 4, it can be shown that $\hat{g}^{(v)}$ converges to $g^{(v)}$ in the interior as well as at boundary points and asymptotic formulas for the bias, variance, and MISE of $\hat{g}$ can be derived. The asymptotic bias and variance of $\hat{g}^{(v)}$ (Beran & Feng, 2002b) are given by:

$$E\left(\hat{g}^{(v)} - g^{(v)}\right) = b^{m-v}\frac{g^{(m)}(\tau)\lambda_{(v,m,c)}}{m!}o\left(b^{(m-v)}\right) \quad (29)$$

and

$$(nb)^{1-2d}b^{2v}var\left(\hat{g}^{(v)}\right) = V(c) + o(1) \quad (30)$$

Here, $V(c) \in (0, \infty)$ and $\lambda_{(v,m,c)} \coloneqq \int_{-c}^{1} u^m K_{(v,m,c)}(u)du$ is the kernel constant of the equivalent kernel $K_{(v,m,c)}$, which acts as a m-th order kernel in the interior ($c=1$) and as a boundary kernel ($c<1$) for the estimation of $g^{(v)}$. Beran and Feng (2002b) also provide explicit formulas for calculating $V(c)$, which is necessary for selecting the bandwidth using the iterative-plug-in method; the expression for $V(1)$ at an interior point is given by $V(1) = 2\pi c_f \int_{-1}^{1} K_{(v,m,1)}^2(x)\,dx$ for $d=0$ and $V(1) = 2\pi c_f \Gamma(1-2d)\sin(\pi d)\int_{-1}^{1}\int_{-1}^{1} K_{(v,m,1)}(x)K_{(v,m,1)}(y)|x-y|^{2d-1}\,dx\,dy$ for $d>0$, where c_f represents the spectral density of the ARMA part of the respective Semi-GARCH at frequency zero. It is worth noting that, similar to $K(u)$, the equivalent kernel $K_{(v,m,c)}(u)$ is polynomial in form and has been employed in earlier studies, including Gasser et al. (1985) and Feng et al. (2020). For $v=0$ and $l=1$, the equivalent kernel $K_{(0,2,1)}(u)$ reduces to the weight function $K(u)$, whereas for $\nu = 0$ and $p = 3$, it yields a fourth-order equivalent kernel. It is well established that the asymptotic properties of the local polynomial estimator coincide with those of its (asymptotically) equivalent kernel representation (see, e.g. Ruppert & Wand, 1994; Beran & Feng, 2002c). To reduce complexity and to save space, the case of antipersistence (where $d < 0$) is omitted, as only the two cases above are relevant for the applications following below.

The asymptotic MISE (AMISE) is defined in Theorem 2 of Beran and Feng (2002b). The asymptotic optimal bandwidth that minimizes the AMISE is given by:

$$b_{opt} = C_{opt} n^{(2d-1)/(2m+1-2d)} \tag{31}$$

where C_{opt} is defined as follows:

$$C_{opt} = \left(\frac{[m!]^2}{2(m-\nu)} \frac{(2\nu + 1 - 2d)}{\lambda^2_{(\nu,m,1)}} \frac{(d_b - c_b)V(1)}{I[g^{(m)}]} \right)^{1/(2m+1-2d)} \tag{32}$$

where $I\left[g^{(m)}\right] = \int_{c_b}^{d_b} (g^m(\tau))^2 \, d\tau > 0$ where both c_b and d_b are constants lying in the interval [0,1] where $c_b < d_b$ to control for the boundary effect and $\lambda = \int_{-1}^{1} x^m K(x)\, dx$. For estimating the short memory version, the smooth function is used, setting $d = 0$ by default (see Feng et al., 2020; Feng et al., 2022). Adopting a bandwidth of order $b = O(b_{opt})$ leads to the asymptotically optimal rate of MISE convergence, given by MISE $= O(n^{2(2d-1)(m-\nu)/(2m+1-2d)})$. This rate corresponds to the minimax efficiency bound for estimating the ν-th derivative of an m-times continuously differentiable function g when long-range dependence is present. In the absence of long memory - i.e., under short-range dependence (see Feng et al., 2020; 2022) - the expression simplifies to MISE $= O(n^{-2(m-\nu)/(2m+1)})$ (Ayensu et al., 2025). Based on these results, an iterative plug-in algorithm to estimate the optimal bandwidth will be presented in the next section. An adapted version of Algorithm B by Beran and Feng (2002b) is used (Letmathe et al., 2022).

4.2 Iterative-Plug-In Algorithm for SEMIFARIMA models

A key practical challenge in nonparametric regression lies in the data-driven selection of the bandwidth parameter. Under the assumption of short memory, several well-known approaches have been proposed, including those by Gasser et al. (1991), Ruppert et al. (1995), Beran and Feng (2002a), and Feng et al. (2020). For processes incorporating long memory, relevant contributions include Ray and Tsay (1997), Beran and Feng (2002c), and Letmathe et al. (2022). In the present study, we adopt the iterative plug-in (IPI) approach proposed by Letmathe et al. (2022) for semiparametric FARIMA (SEMIFARIMA) models, which extends Algorithm B of

Beran and Feng (2002a). The latter algorithm, originally based on the IPI idea of Gasser et al. (1991), employs an exponential inflation step to accelerate the relative convergence rate of the estimated bandwidth. The IPI bandwidth selection procedure for local linear ($m = 2$) and local cubic ($m = 4$) estimators $\hat{g}$ can be summarized as follows:

1) An initial bandwidth b_0 is pre-specified to initiate the iterative procedure. Furthermore, the autoregressive (p) and moving-average (q) orders are determined by the BIC.
2) In iteration j, the function g is estimated by minimizing the locally weighted least squares criterion for y_t, using the bandwidth b_{j-1} from the previous iteration. Subsequently, the residuals $\{\tilde{z}_t\} = y_t - \hat{g}(\tau_{t,n})$ are computed, and the parameters $\hat{d}$ and $\hat{V}$ are then obtained by fitting a FARIMA model to the residual series $\{\tilde{z}_t\}$.
3) The bandwidth is then updated using the exponential inflation method (Beran and Feng, 2002a), defined as follows:

$$b_{d,j} = b_{j-1}^{\alpha}, \; with\; \alpha = \begin{cases} \frac{5-2\hat{d}}{7-2\hat{d}}(optimal), or \frac{5-2\hat{d}}{9-2\hat{d}}(naive), if l = 1 \\ \frac{9-2\hat{d}}{11-2\hat{d}}(optimal), or \frac{9-2\hat{d}}{13-2\hat{d}}(naive), if l = 3 \\ 0.5\,(stable), for\; l \in \{1,3\}^{3} \end{cases} \tag{33}$$

$I[g^{(m)}]$ is estimated using $b_{d,j}$ and a local polynomial of order $l^* = l + 2$. Compute

$$b_j = \left(\frac{[m!]^2}{2m}\frac{(1-2\hat{d})}{\beta^2}\frac{(d_b - c_b)\hat{V}(1)}{I[\hat{g}^{(m)}]}\right)^{1/(2m+1-2\hat{d})} * n^{(2\hat{d}-1)/(2m+1-2\hat{d})} \tag{34}$$

4) Steps 2) and 3) are repeated until convergence or a given number of iterations is reached. Set $\hat{b}_{opt} = b_j$ as the bandwidth selected by the algorithm.

The natural extension to short memory models is given by setting $d = 0$ (Feng et al., 2020). As a result of estimating g and b_{opt}, the detrended or adjusted residuals $\tilde{z}_t$ or $\hat{\xi}_t$ can be investigated with one of the GARCH models presented above. It should be noted that the findings in Beran and Feng (2002a,b,c), Beran and Ocker (1999, 2001), and Beran et al. (2013) continue to apply to the IPI for SEMIFARIMA models. For a more detailed explanation of the procedure, including modifications and adaptations of the IPI algorithm, readers are referred to Feng et al. (2021) and Letmathe et al. (2023).

4.3 Estimation of the GARCH parameter vector

To estimate the parameter vector θ of the GARCH models discussed in Sections 2, we rely on maximum likelihood estimation (MLE), following the approach formalized by White (1982).

[3] For complementary results under the assumption of short memory see Feng et al. (2020; 2022).

Specifically, six alternative conditional distributions for the model innovations are considered: the Student-t, the GED, the AL distribution (Feng, forthcoming), and their respective skewed variants following the idea of Fernández and Steel (1998). Exploring this set of distributions provides the necessary flexibility to accommodate key empirical regularities of financial return series, including excess kurtosis and asymmetric dynamics (see e.g. Mandelbrot, 1963). For conditionally t-distributed innovations, the QMLE is defined by $\hat{L}_n(\theta) = \frac{1}{n}\sum_{t=1}^{n} \hat{l}_t(\theta)$ and

$$\hat{l}_t(\theta) = c_0 - \frac{1}{2}\ln\left(\hat{\sigma}_t^2(F_{t-1};\theta)\right) - \frac{k+1}{2}\ln\left(1 + \frac{r_t^2}{\hat{\sigma}_t^2(F_{t-1};\theta)(k-2)}\right) \tag{35}$$

Here, $\ln\left(\sigma_t^2(F_{t-1};\theta)\right)$ represents the conditional variance specification implied by the particular GARCH model under consideration, given the information set F_{t-1}. The return series $\{r_t\}$ is assumed to be observable and serves as the basis for likelihood-based inference. Additionally, $c_0 = \ln\left(\Gamma\left(\frac{k+1}{2}\right) / \left[\sqrt{\pi(k-2)}\Gamma\left(\frac{k}{2}\right)\right]\right)$ and k denotes the degrees of freedom. As an additional conditional distribution, the GED accommodates both heavy and light tails through its shape parameter, allowing for a better fit to varying market conditions, whereas the t-distribution imposes consistently heavy tails. In GARCH-type models, the GED improves volatility estimation by reducing sensitivity to extreme observations while maintaining efficiency in parameter estimation. Following Nelson (1991) (2.4) and (2.5), a random variable x is standardized GED-distributed, if its density function follows:

$$f(x) = v\left(\Gamma\left(\frac{3}{v}\right)\right)^{\frac{1}{2}}\left(2\left(\Gamma\left(\frac{1}{v}\right)\right)^{\frac{3}{2}}\right)^{-1} e^{-|x|^v}\left(\frac{\Gamma(3/v)}{\Gamma(1/v)}\right)^{\frac{v}{2}} \tag{36}$$

where v describes the thickness of the tails, which allows the innovation process to exhibit thin ($v > 1$) or fat ($v < 1$) tails and thus provides a more versatile alternative to the standard normal distribution (which postulates $v = 2$) and the t-distribution. Assuming conditionally standardized GED innovations, the log-likelihood function of the process $\{r_t\}$ can be expressed as

$$\hat{l}_t(\theta) = c_1 - \frac{1}{2}\ln\hat{\sigma}_t^2(F_{t-1};\theta) - \left(\frac{\Gamma(3/v)}{\Gamma(1/v)}\right)^{\frac{v}{2}} \frac{|r_t|^v}{\hat{\sigma}_t^v(F_{t-1};\theta)} \tag{37}$$

where c_1 is defined as $\ln\left(\frac{v}{2}\right) + 1/2\ln\left(\Gamma\left(\frac{3}{v}\right)\right) - 3/2\ln(\Gamma\left(\frac{1}{v}\right))$ (Nelson, 1991). Under the aforementioned standard innovation distributions, the (M)EGARCH models exhibit either no finite moments or finite moments of all orders (see Theorem 2.2 in Nelson, 1991). The sole exception arises when the innovations follow the standard symmetric Laplace distribution (also referred to as the double exponential distribution), corresponding to the GED with $\nu = 1$. In this case, (M)EGARCH processes may be weakly stationary, yet their fourth moments can never be finite. To address these limitations, alternative conditional distributions have been proposed. A particularly appealing class is the Sargan family of distributions introduced by Missiakoulis (1983) and suggested by Feng et al. (forthcoming). Let $\{Y_i\}_{i=0}^{P} \sim L(b)$ denote i.i.d. zero-mean Laplace random variables with scale parameter $b > 0$. Missiakoulis (1983) showed that, for fixed P, the distribution of the average $X = \bar{Y}$ defines a one-parameter Sargan distribution, which we denote as the Average Laplace distribution $AL(\alpha, P)$. If $P = 0$, it reduces to the Laplace distribution $L(b)$, itself a member of the GED family with $\nu = 1$. The Laplace distribution is parameterized such that $\mathrm{E}[Y_0] = 0$ and $Var(Y_0) = 2b^2 = 2(P+1)^2/\alpha^2$. Hence, given α, the scale parameter b is no longer free. The density of the $AL(P)$ distribution for x is

$$f_{AL}(x; \alpha, P) = \frac{\alpha_P K}{2\sigma_t} e^{-\alpha|x|} \sum_{j=0}^{P} c_j {\alpha_P}^j |x|^j \tag{38}$$

where $a = \sqrt{2(P+1)}$, and the coefficients $\{c_j\}$ are defined recursively as $c_0 = c_1 = 1$ and $c_j = \frac{2(P-j+1)}{j(2P-j+1)} c_{j-1}$ for $\forall j \; = \; 2, 3, \ldots, P$ as well as $K = 2^{-2P}\binom{2P}{P}$ for $P \geq 0$. As shown by Tse (1987), these distributions are symmetric and unimodal. By construction, their variance is

$$Var(x) = \frac{Var(y)}{P+1} = \frac{2(P+1)}{{\alpha_P}^2} \tag{39}$$

Therefore, in the standardized AL distributions with unit variance, the scale parameter a is fixed at values depending only on P. By the central limit theorem, as $P \to \infty$, the distribution $AL(P)$ converges in distribution to $N(0,1)$. Thus, the AL family provides a one-parameter approximation to the normal distribution, distinct from the approach of the Student-t family. For a numerical comparison of $AL(P)$ for $P = 0,1, \ldots, 5$, see Tse (1987), who also contrasts this distribution with an alternative Sargan distribution proposed by Goldfeld and Quandt (1981). The log-likelihood function, assuming $AL(P)$ distributed innovations, is defined as before and

$$\hat{l}_t(\theta) = c_2 - ln\, 2\hat{\sigma}_t(F_{t-1}; \theta) - \alpha_P \frac{|r_t|}{\hat{\sigma}_t(F_{t-1}; \theta)} + ln \left(\sum_{j=0}^{P} c_j {\alpha_P}^j \left| \frac{r_t}{\hat{\sigma}_t(F_{t-1}; \theta)} \right|^j \right) \tag{40}$$

where $c_2 = \ln(\alpha_P K)$. It is worth noting that the $AL(P)$ distribution broadens the applicability of the EGF by allowing these models to attain wide-sense stationarity under certain conditions.

Additionally, we use the skewed variants of the symmetric unimodal distributions introduced earlier, following the construction of Fernandez and Steel (1998). This flexible class of skewed distributions is indexed by a skewness parameter $\vartheta \in \mathbb{R}^+$ and is defined as

$$f(x \mid \vartheta) = \frac{2}{\vartheta + \vartheta^{-1}} [f(\vartheta^{-1}x)H(x) + f(\vartheta x)H(-x)] \tag{41}$$

where $H(\cdot)$ denotes the Heaviside step function, and $f(\cdot)$ is the density of the underlying symmetric distribution. The first two moments of this distribution are given by $E(X) = M_1(\vartheta - \vartheta^{-1})$, $Var(X) = (M_2 - M_1^2)(\vartheta^2 + \vartheta^{-2}) + 2M_1^2 - M_2$ where M_1 and M_2 denote the first and second absolute moments of the baseline symmetric distribution, respectively (see Eq. (5) in Fernández & Steel, 1998). The key idea underlying this construction is the use of inverse scale factors on the positive and negative half-lines of the real axis to induce skewness while preserving unimodality. Alternative approaches for generating skewed distributions from symmetric ones include hidden truncation (Azzalini, 1985), order-statistics transformations (Jones, 2004), and inverse probability integral transformations (Ferreira & Steel, 2006). The parameter ϑ directly governs asymmetry: for $\vartheta = 1$, the distribution reduces to its symmetric form, i.e., $f(x \mid \vartheta = 1) = f(x)$. Values $\vartheta < 1$ imply left-skewness, while $\vartheta > 1$ imply right-skewness. A standardized version with zero mean and unit variance has the density function

$$f_s(z \mid \vartheta) = \sqrt{Var(X)} f\left(z\sqrt{Var(X)} + E(X) \mid \vartheta\right) \tag{42}$$

For standardized skewed innovations, the QMLE is obtained by maximizing

$$\hat{l}_t(\theta) = \frac{1}{2} ln\,(Var(X)) - \frac{1}{2} ln\, \hat{\sigma}_t^2(F_{t-1};\theta) + ln \left(f\left(\frac{r_t\sqrt{Var(X)}}{\hat{\sigma}_t(F_{t-1};\theta)} + E(X) \mid \vartheta \right) \right) \tag{43}$$

The QMLE $\hat{\theta}$ is obtained by solving the respective log-likelihood function of the conditional distributions and is given by a measurable solution of

$$\hat{\theta}_n = \arg\min_{\theta} \hat{L}_n(\theta) \, ^4 \tag{44}$$

All models are fitted using the `fEGarch` R package form Schulz et al. (2025). The estimation procedure is implemented by nonlinear optimization using the augmented Lagrange algorithm of Ye (1997), as implemented in the R package `Rsolnp` (Ghalanos & Theussl, 2011). The corresponding standard errors are derived from the variance-covariance matrix, obtained by inverting the Hessian matrix of the augmented Lagrangian at the optimum. The standard errors, denoted $s.e.(\hat{\theta})$, correspond to the square roots of the diagonal elements of this matrix.

[4] It is well known that under conventional regularity assumptions, the QMLE preserves both consistency and asymptotic normality, even when the innovations ε_t deviate from conditional normality.

4.3 Rolling one-step ahead forecasts

The forecasts of semiparametric LM-GARCH models can be estimated utilizing the `fracdiff` package embedded in the `fEGarch` package for estimating FARIMA processes in R. For this purpose, a truncated AR(∞) representation of $\tilde{z}_t$ can be used to calculate the conditional expected values from which the volatility is derived (Letmathe et al., 2022):

$$\hat{z}_t = \sum_{i=1}^{L} \hat{\lambda}_i \tilde{z}_{t-i}, \quad \tilde{z}_t = 0 \; for \; t < i \tag{45}$$

Here L represents the number of lags and $\hat{\lambda}_i$ denote the coefficients of $\hat{\lambda}(B) = \hat{\beta}^{-1}(B)(1-B)^{\hat{d}}\hat{\phi}(B) = 1 - \sum_{i=1}^{L} \hat{\lambda}_i B^i$ with $i = 1, \dots, L$ and $\tilde{z}_t = y_t - \hat{g}(\tau_{t,n})$. In the short memory case, d is set to zero and $\tilde{z}_{t-i}$ is exchanged by $\tilde{y}_{t-i} = y_{t-i} - \bar{y}$. As proposed by Sucarrat (2019), a three-step estimation procedure to obtain the total volatility is followed. We follow the methodology of Letmathe et al. (2022) and estimate the deterministic component of y_t, $g(\tau_{t,n})$ non-parametrically by a data-driven local polynomial smoother described above.

1) Obtain an estimate of $g(\tau_{t,n})$ by applying the IPI for SEMIFARIMA processes and consequently determine the residuals $\tilde{z}_t = y_t - \hat{g}(\tau_{t,n})$.
2) Fit a FARIMA(p^*, q) model to the calculated $\tilde{z}_t$ and determine $\hat{z}_t$ by means of the truncated AR(∞) approach described above.
3) The total volatilities are then consequently defined by

$$\hat{\zeta}_t = \sqrt{\hat{C}_s \, exp[\hat{g}(\tau_{t,n})/2]} \sqrt{\hat{C}_\sigma \, exp[\hat{z}_t/2]} = \hat{s}(\tau_{t,n})\sigma_t \tag{46}$$

The parameters $\hat{C}_s$ and $\hat{C}_\sigma$ are estimated similarly to Sucarrat (2019) (see also Sucarrat et al. (2016) and Escribano and Sucarrat (2018)). We obtain $\hat{C}_s = var[exp(\tilde{z}_t/2)] = exp(-\hat{\mu}_{l\xi^2})$ and $\hat{C}_\sigma = var[\hat{\xi}_t / exp(\hat{z}_t/2)] = exp(\hat{\mu}_{l\xi^2} - \hat{\mu}_{l\varepsilon^2})$, where $\hat{\mu}_{l\xi^2} = -\,ln[\frac{1}{n}\sum_{i=1}^{n} exp(\tilde{z}_t)]$ and $\hat{\mu}_{l\varepsilon^2} = -\,ln[\frac{1}{n}\sum_{i=1}^{n} exp(\hat{\eta}_t)]$, which are smearing estimators (Duan, 1983) of $E(\ln \xi_t^2)$ and $E(\ln \varepsilon_t^2)$. Consequently, the total volatility calculated in step 3) only requires an estimate of $\mu_{l\varepsilon^2}$ since $\hat{C}_s\hat{C}_\sigma = \exp(-\hat{\mu}_{l\varepsilon^2})$. For the purely parametric specifications, the estimation procedure remains unchanged in principle but becomes substantially simpler due to the omission of the smoothing component. The rolling one-day forecasts are therefore given by

$$\hat{z}_{n+k} = \sum_{i=1}^{L} \hat{\lambda}_i \, \tilde{z}_{n+k-i} \tag{47}$$

where $k = 1, \ldots, K$ and K is the number of out-of-sample observations. This paper follows the methodology of extrapolating the last in-sample estimate of the deterministic component $g(\tau_{t,n})$ as a prediction for the unconditional volatility for the out-of-sample period. Hence, $\tilde{z}_{n+k-i} = y_t - \hat{g}(\tau_t)$ if $n + k - i \leq n$ and $\tilde{z}_{n+k-i} = y_{n+k-i} - \hat{g}(\tau_n)$ otherwise. The rolling one-day-ahead forecasts for the total volatility can be determined by inserting $\hat{z}_{n+k}$ and $\hat{s}(\tau_{t,n})$ into (46).

5 Application to the regulatory requirements of banks' quantitative risk management

In financial economics, risk refers to the potential deviation of a financial variable from its expected value under uncertainty, with large negative realizations posing serious threats to firms, especially to banks. Adequate capitalization and efficient capital allocation are therefore central to both institutional stability and shareholder value, particularly in a highly interconnected banking system. The global financial crisis of 2007 highlighted the necessity of rigorous methods for identifying and quantifying such risks. Quantitative risk management addresses this need through statistical and econometric tools, among which VaR and ES are the most widely adopted. Previous research has shown that removing deterministic components from financial time series can enhance the accuracy of quantitative risk measure estimation (see, e.g., Peitz, 2016; Letmathe et al., 2022). Semi-LM-GARCH models, which combine semiparametric flexibility with long memory dynamics, are therefore promising candidates for enhanced tail-risk forecasting. In the empirical application, the proposed models are fitted on an in-sample segment of $n_{\mathrm{in}} = n - K$ trading days, with $K = 250$ corresponding to approximately one year of observations. Rolling one-step-ahead VaR and ES forecasts at the regulatory confidence levels of 99% and 97.5% are then produced for the K-day test window, in line with Basel Committee recommendations (BCBS, 2019). For a short overview of the regulatory framework, VaR, and ES, see Appendices A.1 and A.3 respectively.

5.1 One-day-ahead forecast of VaR and ES

Let $n_{in} = n - K$ and $-r_t$ be the linearized losses. A one-day rolling forecast for the VaR with a confidence level α using a parametric GARCH model is given by:

$$\widehat{VaR}_{n_{in}+k}(\alpha) = -\bar{r} + \hat{\sigma}_{n_{in}+k} F^{-1}_{dist(\theta)}(\alpha) \tag{48}$$

where $k = 1, \dots, K$, $F_{dist(\theta)}^{-1}$ denotes the cumulative distribution function of the employed (skewed) distribution with parameter vector θ, and $\bar{r}$ is the in-sample average return. Please note that $F_{dist(\theta)}^{-1}$ is an adjusted version of the CDF, which arises from the standardization of the respective innovations to ensure unit variance.[5] Similarly, the VaR can be calculated using semiparametric approaches. For tractability, the final in-sample estimate of $s(\tau_{t,n})$ is employed as a proxy for the unconditional volatility during the out-of-sample phase, such that $\hat{\zeta}_{n_{\text{in}}+k} = \hat{s}(\tau_{t,n})\,\hat{\sigma}_{n_{in}+k}$. The corresponding rolling one-step-ahead forecast of the VaR at confidence level α using a semiparametric GARCH model is then defined as

$$\widehat{VaR}_{n_{in}+k}(\alpha) = -\bar{r} + \hat{\zeta}_{n_{in}+k} F_{dist(\theta)}^{-1}(\alpha) \tag{49}$$

For the ES, a one-day rolling forecast using a parametric GARCH model is obtained as follows:

$$\widehat{ES}_{n_{in}+k}(\alpha) = -\bar{r} + \hat{\sigma}_{n_{in}+k} ES_{\varepsilon,\alpha} \tag{50}$$

For a semiparametric GARCH model as

$$\widehat{ES}_{n_{in}+k}(\alpha) = -\bar{r} + \hat{\zeta}_{n_{in}+k} ES_{\varepsilon,\alpha} \tag{51}$$

For different (skewed) distributions, $ES_{\varepsilon,\alpha}$ is defined as $ES_{\varepsilon,\alpha} = (1-\alpha)^{-1} \int_{\alpha}^{1} F_{dist(\theta)}^{-1}(\alpha)$.[6]

5.2 Backtesting VaR and ES

We backtest VaR and ES forecasts over a test window of $K = 250$ trading days. Model accuracy is assessed by comparing predicted risk forecasts with realized losses. For VaR, we apply the BCBS (2016) traffic light test based on exceedance counts; for ES, we use the Costanzino-Curran (2018) traffic light procedure at $\alpha = 97.5\%$. Define the empirical hit sequence:

$$I_{n_{in}+k}(\alpha) = \mathbf{1}\{-r_{n_{in}+k} > \widehat{VaR}_{n_{in}+k}(\alpha)\} \tag{52}$$

[5] In GARCH-type models, the innovation process is typically assumed to have mean zero and unit variance. Because quantiles scale linearly with multiplicative constants, the standardized quantile at level α is calculated by multiplying the quantile of the unstandardized distribution by the inverse of the respective distribution standard deviation to correctly adjust for standardization. For the Student-$t(v)$ distribution, the scaling term is equal $\sqrt{v-2/v}$, for $GED(v)$ $\sqrt{\Gamma(1/v)/\Gamma(3/v)}$, and for $AL(P)$ $a_P/\sqrt{2(P+1)}$.

[6] For the t-distribution with degrees of freedom v, $\widehat{ES}_{\varepsilon}(\alpha)$ is the ES of a standardized conditional t-distribution with unit variance. For $v > 2$, the result is $ES_{\varepsilon}(\alpha) = \frac{f_v[F_v^{-1}(\alpha)]}{1-\alpha} \frac{v+[F_v^{-1}(\alpha)]^2}{v-1} \sqrt{\frac{v-2}{v}}$ with f_v as the density function of a conditional t-distribution (McNeil et al., 2015). For a symmetric GED with shape $v > 0$ and scale $\lambda = \left(\frac{\Gamma(1/v)}{\Gamma(3/v)}\right)^{1/2}$, so the variance is unit, let q_α be the right-tail α-quantile (i.e., $F(q_\alpha) = \alpha$). Define $w = (\frac{q_\alpha}{\lambda})^{v}$. Then the ES of the standardized GED is $ES_{\varepsilon,\alpha} = \frac{\int_{q_\alpha}^{\infty} x f(x)\, dx}{\int_{q_\alpha}^{\infty} f(x)\, dx} = \lambda \frac{\Gamma(\frac{2}{v}, w)}{\Gamma(\frac{1}{v}, w)}$. For a symmetric $AL(P)$ innovation with order $P \geq 0$ and unit variance obtained by setting $a = \sqrt{2(P+1)}$, let q_α be the right-tail α-quantile. Define $z = a\, q_\alpha$. With coefficients $\{c_j\}_{j=0}^{P}$ from the *AL(P)* density (Missiakoulis, 1983; $c_0 = 1$, $c_{j+1} = \frac{2(P-j)}{(j+1)(2P-j)} c_j$), the ES of the standardized *AL(P)* is $ES_{\eta,\alpha} = \frac{\int_{q_\alpha}^{\infty} x f(x)\, dx}{\int_{q_\alpha}^{\infty} f(x)\, dx} = \frac{1}{a} \frac{\sum_{j=0}^{P} c_j \Gamma(j+2, z)}{\sum_{j=0}^{P} c_j \Gamma(j+1, z)}$, where $\Gamma(\cdot,\cdot)$ is the upper incomplete gamma function.

so that $\sum_{t=1}^{K} I_t = N_\alpha$ indicates the number of VaR violations. Because I_t is Bernoulli with success probability $1-\alpha$, the exceedance count follows $N_\alpha \sim \text{Binomial}(K, 1-\alpha)$. A two-sided binomial test is therefore used to evaluate correct calibration. Under the BCBS criteria, the green zones are $N_1 \leq 10$ for $\alpha = 97.5\%$ and $N_2 \leq 4$ for $\alpha = 99\%$. The corresponding expected numbers of violations are $E[N_1] = 6.25$ and $E[N_2] = 2.5$. Well-specified VaR models should produce violation counts that do not deviate substantially from these benchmarks, since large deviations indicate either systematic underestimation or overestimation of tail risk.[7]

More recently, Costanzino and Curran (2018) provide a traffic light test approach for backtesting the ES based on the severity of breaches. Define a violation residual as

$$\hat{\varepsilon}^*_{n_{in}+k} = -\frac{r_{n_{in}+k} - \bar{r}}{\hat{\sigma}_{n_{in}+k}} \; or \; \hat{\varepsilon}^*_{n_{in}+k} = -\frac{r_{n_{in}+k} - \bar{r}}{\hat{\zeta}_{n_{in}+k}} \tag{52}$$

when using a parametric or a semiparametric GARCH model, respectively. These residuals have to be standardized to unit variance with the aforementioned scale parameters, if not standardized. Note that $\varepsilon^*_{n_{in}+k}$ behaves like an i.i.d. random variable with continuous density, therefore the conditions in Equation (14) in Costanzino and Curran (2018) are fulfilled. Define

$$\omega_{n_{in}+k} = \begin{cases} 1 - \dfrac{1 - F_\theta\left(\hat{\varepsilon}^*_{n_{in}+k}\right)}{1-\alpha}, & when\; I_{n_{in}+k} = 1 \\ 0, & otherwise \end{cases} \tag{53}$$

where F_θ denotes the fitted cumulative distribution function of the standardized innovations. The test statistic for the ES is then the weighted sum of risk measure exceedances:

$$T_{ES} = \sum_{k=1}^{K} \omega_{n_{in}+k} \tag{54}$$

Costanzino and Curran (2015; 2018) also show that $(T_{ES} - \mu_{ES})/\sqrt{K} \xrightarrow{asymt} N(\mu_{ES}, h_{ES})$ with $\mu_{ES} = (1-\alpha)K/2$ and $h_{ES} = (1-\alpha)(1+3\alpha)/12$. For a backtest at level $\alpha = 97.5\%$ with $K = 250$ observations, the T_{ES} statistic has expected value $\mu_T = 3.125$ and an approximate standard deviation $\sqrt{K}\sigma_T \approx 1.43$. Asymptotic and finite-sample critical values for the ES traffic light statistic T_{ES} are also calculated, resulting in green-zone boundaries at 5.48 and 5.70, respectively. The latter corrects distortions that arise in regulatory settings with limited sample sizes. These findings are particularly relevant given the BCBS shift from VaR to ES under the FRTB, making the T_{ES} traffic light test a coherent and informative tool for evaluating tail-risk forecasts. In this study, a model is deemed adequate when N_1, N_2, and T_{ES} all lie within their

[7] The allocation for the violations of the VaR series according to the Basel parameters and for the ES suggested by Costanzino and Curran (2018) based on 250 observations can be seen in the tables of Appendix A.4.

respective green zones, ensuring consistency across coverage and tail-severity dimensions, enhancing statistical rigor while aligning with supervisory requirements. Among the models that pass these tests, selection criteria are used to select the best-performing specification.

5.3 Selection Criteria

Based on the estimated market risk, financial institutions are required by regulation to hold an appropriate amount of capital as a buffer against potential losses. If a model successfully meets all regulatory requirements as well as coverage and independence criteria, selection is often guided by loss functions that balance statistical accuracy with opportunity costs associated with capital allocation.[8] Their objective is to identify the model that minimizes the expected loss. In the context of regulatory practice, the BCBS has emphasized the use of regulatory loss functions. Let RM_t be a risk measure at time t (consequently a VaR_t or an ES_t). For the regulatory loss function, predictive accuracy is measured by

$$LF_i = \sum_{t=1}^{n} l_{t,i} \; where \; l_{t,1} = \left(\widehat{RM}_t - r_t\right)^2 \tag{55}$$

for $r_t < RM_t$ and $t \in \{T-250, \ldots, T\}$. In the regulatory framework $l_{t,i}$ is defined by 0 if $r_t > RM_t$.[9] Using this loss function for quantitative risk measures, not only aligns with the capital adequacy framework under the FRTB, but also provides an operationally meaningful

[8] We conduct several backtests as side conditions, namely the independence, conditional and unconditional coverage test (Christoffersen, 1998; Kupiec, 1995): Take the (empirical) hit sequence $I_{n_{in}+k}(\alpha) = I_t$. Under the null hypothesis I_t is a Bernoulli process with probability p. Let K_0 and K_1 denote the number of zeros and ones in the hit sequence, respectively. Under the null hypothesis of unconditional coverage ($H_{0,\mathrm{uc}}$), exceedances are assumed to occur independently with a fixed probability $p = 1-\alpha$. Given this, the likelihood of the observed sequence is $L(I_t, p) = p^{K_1}(1-p)^{K_0}$. To evaluate whether the model's predicted violation probability matches the empirical one, an alternative likelihood is constructed using the maximum likelihood estimate of the violation probability, $\widehat{w} = K_1/(K_0+K_1)$, which yields $L(I_t, \widehat{w}) = \widehat{w}^{K_1}(1-\widehat{w})^{K_0}$. The likelihood ratio statistic for testing $H_{0,\mathrm{uc}}$ is then $LR_{\mathrm{uc}} = -2\ln\left(\frac{L(I_t,p)}{L(I_t,\widehat{w})}\right)$, which asymptotically follows a chi-square distribution with one degree of freedom. A rejection of $H_{0,\mathrm{uc}}$ indicates that the observed frequency of violations is inconsistent with the model's predicted coverage level. If $K_1 = 0$, the unconditional coverage test cannot be carried out, since the maximum likelihood estimate $\widehat{w}$ is not well defined in this case. In such situations, we simply regard the test as not applicable. Additionally, we carry out the conditional coverage test introduced by Christoffersen (1998), which combines the conditional coverage test by Kupiec (1995) with an independence test. Let w_{ij} denote the probability that an outcome $i \in I_t$ on day $t-1$ is followed by an outcome $j \in I_t$ on day t. The null hypothesis for the conditional coverage test is ($H_{0,\mathrm{cc}}$): $w_{01} = w_{11} = p$, meaning that the probability of observing a violation is independent of the previous day's state. Let K_{ij}be the number of transitions from state i to state j. The likelihood under the unrestricted transition probabilities is $L(I_t, \widehat{w}_{01}, \widehat{w}_{11}) = (1-\widehat{w}_{01})^{K_0-K_{01}}\,\widehat{w}_{01}^{K_{01}}(1-\widehat{w}_{11})^{K_1-K_{11}}\,\widehat{w}_{11}^{K_{11}}$, where $\widehat{w}_{01} = K_{01}/K_0$ and $\widehat{w}_{11} = K_{11}/K_1$. The conditional coverage test statistic is defined as $LR_{\mathrm{cc}} = -2\ln\left[\frac{L(I_t,p)}{L(I_t,\widehat{w}_{01},\widehat{w}_{11})}\right]$ which, under the null hypothesis, follows an asymptotic χ^2 distribution with two degrees of freedom. For finite samples, it may occur that $K_{11} = 0$. In such cases, following Christoffersen and Pelletier (2004), the likelihood $L(I_t, \widehat{w}_{01}, \widehat{w}_{11})$ is computed using only the first two terms of the expression. At the confidence level $\alpha = 0.99$, it is common to encounter models with $K_1 \leq 1$. When $K_1 = 0$, the conditional coverage test is deemed not applicable. In fact, the independence - or conditional coverage - test is only informative when $K_1 \geq 2$. Consequently, if $K_1 = 1$, the test outcome does not provide meaningful evidence and can likewise be disregarded.

[9] The authors are aware of more recent approaches that consider opportunity costs for excess held regulatory capital. Let $\Omega \geq 0$ be a penalty term that weights the positive deviations from the risk measure, then Sarma et al (2003) set $l_{t,3} = \Omega|RM_t|$, Abad et al. (2015) use $l_{t,4} = \Omega|RM_t - r_t|$, and Feng and Peitz (2025) propose $l_{t,2} = \Omega \min(|RM_t - r_t|, |RM_t|)$. As Feng and Peitz (2025)'s proposal combines those made by the former, we restrict our analysis to using $l_{t,1}$ and $l_{t,2}$.

benchmark for assessing whether candidate models produce risk forecasts that are both conservative and robust in the tail.

While the regulatory loss function reflects supervisory priorities by rewarding conservative forecasts, it does not necessarily align with the economic objectives of financial institutions. As noted by Sarma et al. (2003), firms evaluate VaR and ES models not only by their ability to control tail losses but also by their opportunity cost of holding capital. Therefore, once models have passed regulatory admissibility filters, model selection should incorporate a loss function consistent with firms' economic objectives. Sarma et al. (2003) proposed such a VaR-based loss function, penalizing exceedances quadratically and imposing an opportunity-cost charge when losses are covered. However, their criterion is discontinuous at the VaR threshold and may overweight small exceedances. Abad et al. (2015) improved continuity but unintentionally impose opportunity-cost penalties even when realized returns are positive. To resolve these inconsistencies, Feng and Peitz (2025) introduce a corrected loss function that preserves continuity and applies opportunity costs only when economically justified. Let $\Omega \geq 0$ be a penalty term that introduces opportunity costs for positive deviations from the risk measure, then

$$l_{t,2} = \begin{cases} (\widehat{RM}_t - r_t)^2, & r_t < \widehat{\mathrm{VaR}}_t(\alpha), \\ \Omega \min[|\widehat{RM}_t - r_t|, |\widehat{RM}_t|], & \text{otherwise,} \end{cases} \tag{56}$$

which unifies the desirable features of both earlier approaches while correcting their deficiencies. When the return is negative, the new function behaves like the continuous specification of Abad et al. (2015), penalizing uncovered losses more heavily while applying a proportional opportunity cost only to the relevant portion of the covered loss. When the return is positive, the loss reduces to the structure proposed by Sarma et al. (2003), ensuring that no opportunity cost is charged on realized gains. The resulting loss function is continuous at the VaR boundary and economically consistent across all return realizations. Following Feng and Peitz (2025), the parameter Ω in the corrected loss function is set to one basis point. This corrected firm-oriented loss measure therefore provides (i) coherent incentives for capital allocation, (ii) a stable and interpretable metric for comparing candidate risk models, and (iii) a more realistic representation of a firm's profit–risk trade-off.

While loss functions are well suited for compliance-driven evaluation, they may not fully capture the nuanced trade-off between statistical efficiency and tail accuracy required for econometric model comparison. To address this limitation, Feng et al. (2020) proposed the Weighted Absolute Deviation (WAD) criterion, which jointly backtests the VaR and ES and emphasizing deviations in the tail region by assigning greater weights to extreme observations, thereby offering a more sensitive diagnostic of model performance under stress. Combining regulatory consistency with advanced econometric methodology ensures that the selected models are both practically acceptable and statistically sound.

$$WAD = \frac{|N_1 - \mu_1|}{\mu_1} + \frac{|N_2 - \mu_2|}{\mu_2} + \frac{|T_{ES} - \mu_T|}{\mu_T} \tag{56}$$

An important advantage of employing the WAD criterion lies in its consistency with backtesting outcomes. The WAD score of a model is negatively associated with the magnitude of the test statistics N_1, N_2, and T_{ES}: the poorer the model performance in the tail, the larger the resulting WAD value. This monotonicity ensures that model rankings derived from WAD do not contradict the qualitative assessment provided by traffic light style backtests.

6 Empirical Results

The practical relevance of extending the models to long memory and semiparametric volatility models as well as incorporating skewed innovation distributions is demonstrated by comparing the accuracy of VaR and ES rolling forecasts and by contrasting the parameter estimates of the selected GARCH models. Selected through popularity, media coverage, and trading volume, the models are fitted to daily return series of the 20 largest S&P 500 companies by market capitalization as of December 31, 2024 - namely: Apple (AAPL), Nvidia (NVDA), Microsoft (MSFT), Amazon (AMZN), Alphabet Class A (ABEA), Alphabet Class C (ABEC), Meta (META), Tesla (TSLA), Broadcom (AVGO), Berkshire Hathaway Class B (BRK-B), Eli Lilly (LLY), Walmart (WMT), JPMorgan Chase (JPM), Visa (V), Mastercard (MA), Exxon Mobil (XOM), Oracle (ORCL), UnitedHealth Group (UNH), Costco (COST), and Procter & Gamble (PG). Where possible, the sample period for each stock spans January 1, 2010, to December 31, 2024, covering multiple market regimes - including the COVID-19 crisis - and thus captures potential long-term volatility dynamics. For most of the selected stocks, their IPOs were before 2010, thus 15 years of daily data are available, justifying both long memory modeling and semiparametric smoothing approaches for VaR and ES forecasting. The return series are retrieved from Yahoo Finance via a Python-based data crawler. For model evaluation, each time series is divided into a training sample and a test sample, where the final $K = 250$ trading days (in line with BCBS (2019)) are reserved for backtesting, and the remaining $N - K$ observations are used for model fitting. For semiparametric specifications, the local polynomial trend is first estimated using the IPI procedure depicted in section 4.3, followed by fitting the parametric components to the detrended returns. All estimations are conducted using the `fEGarch` package developed by Schulz et al. (2025). For each time series, 24 models - covering parametric and semiparametric, short and long memory variants - are estimated under six innovation distributions. Although EGARCH and MEGARCH models can, in principle, be estimated under t and skewed t innovations, the empirical results show these are not favorable. In consensus with the empirical results, such specifications are not suggested to ensure that all fitted models possess finite fourth moments, allowing for reliable volatility and risk estimation. For comparability, all models are estimated with orders $(1,1)$ or $(1, d, 1)$. The models are subsequently evaluated at the 97.5% and 99% confidence levels, as prescribed by BCBS (2019).

The descriptive statistics in Tables 1 and 2 strongly support the use of GARCH-type models for the analyzed return series. The pronounced excess kurtosis and skewness across assets indicate clear deviations from normality, reflecting the heavy-tailed and asymmetric nature typical of financial returns. This skewed and leptokurtic behavior is characteristic of financial time series and implies that models assuming normal innovations are likely to underestimate tail risk. Consequently, non-Gaussian innovations - such as the Student-t, GED, or $AL(P)$ distributions - that accommodate heavy tails and skewness are more appropriate for capturing these dynamics. The large and statistically significant Jarque-Bera statistics (Jarque & Bera, 1980) further confirm the rejection of normality for all series, reinforcing the need for non-Gaussian conditional specifications and robust quasi-maximum likelihood estimation techniques. The ARCH-LM statistics (Engle, 1982) confirm conditional heteroskedasticity, validating the presence of time-varying volatility. Both the elevated ARCH-LM and Ljung-Box statistics (Ljung & Box, 1978) for squared returns indicate persistent volatility clustering, especially for technology stocks such as Nvidia, Microsoft, and Meta, which may warrant the use of extended or long memory specifications. This pattern reflects Mandelbrot's (1963) observation that information in financial markets arrives in bursts, generating volatility that mean-reverts only gradually. The combined evidence of fat tails, asymmetry, and volatility persistence justifies employing parametric and semiparametric GARCH frameworks with flexible, skewed, and heavy-tailed error distributions to better capture conditional risk and enhance VaR and ES forecast accuracy.

Stock	AAPL	NVDA	MSFT	AMZN	ABEA	ABEC	META	TSLA	AVGO	BRK-B
Skewness	-0.2303734	0.2362566	-0.2060224	0.009349564	0.1163553	0.03482222	-0.4850172	0.00310525	0.01418615	0.02056311
Kurtosis	8.437999	9.733539	10.98501	9.009085	10.43061	9.19243	25.61192	7.776031	11.32286	12.59232
Jacque-Bera	4682.3	7163	10050	5676.7	8688.6	4328.9	67744	3470	10890	14465
ARCH-LM	386.45	145.35	673.64	90.564	121.48	148.21	17.136	163.74	426.87	1150.5
Ljung-Box (m=5)	495.67	161.09	1016	84.431	135.97	191.73	14.139	216.16	668.7	2347.4
Ljung-Box (m=10)	883.9	226.29	1400.4	126.85	185.86	256.83	22.048	258.05	799.9	3180.4
Ljung-Box (m=20)	1116.6	331.46	1634.6	160.24	242.99	327.39	25.016	367.64	837.64	3475.9

Table 1 - Skewness, kurtosis, Jacque-Bera test statistic (Jarque & Bera, 1980), ARCH-Lagrange Multiplier test statistic (Engle, 1982), and Ljung-Box test statistic (squared returns) (Ljung & Box, 1978) for the selected stocks (part 1).

Index	LLY	WMT	JPM	V	MA	XOM	ORCL	UNH	COST	PG
Skewness	0.6455045	-0.2079431	-0.07334083	-0.1078255	-0.005758725	-0.1572038	0.02884843	-0.3582289	-0.4663827	-0.06593537
Kurtosis	13.72568	19.03588	12.62558	13.0844	11.93714	10.12647	17.50793	13.82978	12.21778	14.2345
Jacque-Bera	18347	40453	14569	15995	12557	7999.6	33090	18519	13494	19845
ARCH-LM	179.99	261.07	911.51	485.71	651.13	881.35	316.4	1009.8	259.96	1173
Ljung-Box (m=5)	216.51	319.45	1841.5	564.76	898.7	1438.8	400.35	1302.5	203.54	2520.9
Ljung-Box (m=10)	323.56	387.73	2678.7	1099.2	1697.3	2498.6	514.72	2344.8	434.04	3717.5
Ljung-Box (m=20)	358.61	428.99	3314.2	1401.2	2141.3	3513.9	526.88	2885.2	544.94	4281

Table 2 - Skewness, kurtosis, Jacque-Bera test statistic (Jarque & Bera, 1980), ARCH-Lagrange Multiplier test statistic (Engle, 1982), and Ljung-Box test statistic (squared returns) (Ljung & Box, 1978) for the selected stocks (part 2).

6.1 Backtesting Results

Tables 3 to 10 present the test statistics of the loss functions, N_1, N_2, T_{ES}, and the WAD for all parametric and semiparametric specifications, respectively. A model is regarded as adequately specified when N_1, N_2, and T_{ES} fall within their designated green zones. Among these, the model with the lowest selection criterion value is considered the most precise in forecasting the VaR and ES. The additional measure N_{ES} in Tables 7 to 10 denotes the number of cases in which realized losses exceeded the ES forecast; it serves only as supplementary information and is not used for model classification. The empirical results show $N_{ES} \leq N_2$. Due to space constraints, only the models under the distribution associated with the overall best model are shown. Most selected models satisfy the traffic light criteria, with the exception of the return series of JPM and UNH for which no satisfactory model could be identified. The best-selected models for those series are positioned in the yellow zone accordingly.[10] Similar to Letmathe et al. (2022), it is empirically observed that $T_{ES} < 5.70$ generally holds whenever both N_1 and N_2 lie within their respective green zones. In some instances, however, N_2 falls into the yellow zone while the condition $T_{ES} < 5.70$ remains satisfied. When both N_1 and N_2 are clearly positioned in the yellow zone, T_{ES} also shifts accordingly, confirming that T_{ES} serves as a reliable and informative statistic for evaluating ES forecasts within the backtesting framework. Surprisingly, and in contrast to the findings of Letmathe et al. (2022), the selected models are predominantly parametric. However, in three cases, the best models are estimated semiparametrically, which indicates that the estimation procedure can be enhanced by the semiparametric extension. While in general accounting for a potentially underlying deterministic scale function could be crucial for passing the traffic light test (see e.g., Letmathe et al., 2022) and the overall in-sample fit of parametric models is usually enhanced by incorporating a long-term risk smoothing function (Ayensu et al., 2025), semiparametric models are by construction inherently prone to the risk of not adapting to regime shifts fast enough, especially in forecasting settings when the estimate for long term risk is held constant. We will see in the next section, however, that they are nevertheless useful and important alternatives to their parametric counterparts, as they produce more stable and stationary GARCH estimates, which are able to identify and account for spurious long memory behavior. The results provide a clear and, in some respects, unexpected assessment of the relative forecasting performance of the competing volatility models. Models selected as optimal according to each criterion are highlighted in bold. Across both the

[10] Detailed analysis of those two time series reveals that the volatility regime changes drastically in 2024. While these time series are heavily affected by the Covid-19-induced uncertainty and the models are trained on data prior 2024, this volatility behavior cannot be appropriately controlled for with backward looking GARCH models.

regulatory and corrected loss functions, members of the EGF usually outperform their peer GARCH models. Depending on the risk measure and loss function considered, especially the (adjusted) Log-GARCH models (incorporating long memory) emerge as the most effective specifications, as up to 17 of the selected 20 models are of this origin. When focusing on the corrected loss function and the ES at the 97.5% level, 13 of the 20 selected models correspond to the standard Log-GARCH class (6 short memory and 7 long memory variants). An additional three specifications belong to the modulus-adjusted Log-GARCH family (1 short and 2 long memory models), while only two selected models are short memory EGARCH models. For the AVGO and LLY series, the selected models are a long memory MEGARCH and a long memory standard GARCH, respectively. These findings suggest that Log-GARCH-type structures (incorporating a long memory and modulus transformation) are particularly well suited for market-risk forecasting, in line with the conclusions of Letmathe et al. (2022). While the selection on in-sample criteria reveals stronger preference for long memory volatility models (see, e.g., Hanke et al., 2025), in total, 11 of the 20 preferred models feature long memory behavior, underscoring the relevance of persistent volatility dynamics for improving tail-risk forecasts. Somewhat unexpectedly, more than half of the selected models do not incorporate sign-dependent asymmetry, implying that additional nonlinear complexity (e.g., leverage effects) does not necessarily translate into better risk forecasts. This may also explain why comparatively parsimonious models outperform more elaborate recently developed structures, despite the latter's attractive statistical properties. Another noteworthy observation is that for certain assets - such as Microsoft - only models from the Log-GARCH family are capable of passing the traffic light test, highlighting their robustness in regulatory settings. For Walmart, standard GARCH-type models substantially overestimate risk, as reflected in the regulatory loss function. An advantage of the corrected loss function is its ability to discriminate among models even in periods without exceedances, by incorporating opportunity costs into the evaluation criterion. As shown, the selected models and distributions usually coincide between the respective loss functions (especially when holding the risk measures constant), which underscores the consistency of the respective selection criteria. However, the results from the WAD criterion usually deviate substantially from those based on the loss functions, yielding alternative GARCH specifications with regulatorily suitable forecasts. For both, JPM and UNH no suitable models across all GARCH variants and innovation distributions could be identified.

Furthermore, the results show that models employing skewed innovation distributions consistently outperform their symmetric counterparts, leading to markedly improved tail-risk forecasts. This finding suggests that practitioners should incorporate asymmetric error distributions when modeling volatility, as doing so better captures empirical return asymmetries, reduces the likelihood of severe unforeseen losses, and lowers unnecessary capital charges. In turn, this frees regulatory capital that can be allocated more efficiently to enhance profitability.

Selection Criterion	AAPL				NVDA				MSFT				AMZN				ABEA			
	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES
G	0.205889	0.184846	1.064113	1.041177	1.148126	0.771209	2.996951	2.781872	1.304416	1.263195	2.085092	2.07224	1.875406	1.792333	3.100837	2.921747	1.850582	1.567764	2.877902	2.675129
S-G	0.618468	0.581907	1.306336	1.179634	3.087961	2.234547	4.513737	3.794791	1.299056	1.195466	2.062891	1.996832	2.303195	2.054834	3.269221	3.073909	2.543008	2.155293	3.398059	3.073613
APA	0.176954	0.158401	1.020358	1.000142	0.701852	0.482312	2.497282	2.420667	1.508734	1.442757	2.302603	2.284946	1.428671	1.189984	2.467065	2.318964	1.925827	1.686142	2.944132	2.781535
S-APA	0.737933	0.697434	1.373363	1.220104	2.083890	1.468292	3.423453	2.922284	1.570700	1.513358	2.338354	2.321621	1.907069	1.648234	2.840488	2.656570	2.625135	2.323597	3.461471	3.221280
EG	0.136284	0.119355	0.993177	**0.983547**	0.772861	0.510670	2.565326	2.448128	1.446402	1.385117	2.253521	2.241783	1.557708	1.324134	2.610147	2.452686	1.953154	1.71376	2.986560	2.825672
S-EG	0.694682	0.655797	1.335571	1.201093	2.372367	1.725240	3.683133	3.140633	1.526602	1.457341	2.306087	2.280477	2.016967	1.722677	2.924647	2.735819	2.636265	2.333357	3.476660	3.235664
MEG	0.160146	0.142125	1.016714	1.000397	0.735696	0.462861	2.514967	2.390378	1.491501	1.424899	2.285275	2.269753	1.504353	1.249429	2.534623	2.378096	1.996116	1.755471	3.018597	2.856816
S-MEG	0.762957	0.723125	1.400546	1.260502	2.470521	1.769153	3.772511	3.177123	1.572222	1.510379	2.337918	2.318219	1.962922	1.658678	2.858760	2.67045	2.685495	2.374629	3.514236	3.265396
MLogG	0.171432	0.153221	1.018285	1.002362	0.61256	0.383601	2.415685	2.324988	1.526687	1.454137	2.324355	2.301671	1.404432	**1.162197**	**2.441715**	**2.295442**	1.844587	1.609436	2.874935	2.717516
S-MLogG	0.784707	0.743725	1.41114	1.268593	2.286638	1.626411	3.579702	3.021399	1.599319	1.532967	2.367001	2.341935	1.798172	1.525152	2.720930	2.548387	2.516815	2.221972	3.349945	3.116566
LogG	0.341187	0.322502	1.156754	1.117533	0.718237	0.527599	2.565896	2.510199	1.295936	1.263217	2.127718	2.132403	1.686481	1.364375	2.734494	2.574726	1.300297	1.039057	2.396955	2.219638
S-LogG	0.597359	0.562371	1.250991	1.142129	2.940195	2.206573	4.214646	3.582501	1.265233	**1.149168**	2.070788	**2.005717**	1.996597	1.720200	2.988678	2.813624	2.171663	1.835575	3.043304	2.771545
FIG	0.209390	0.187760	1.068229	1.051806	1.052808	0.870257	2.921514	2.821241	**1.235814**	1.215940	**2.034666**	2.036934	1.892868	1.713334	3.035879	2.863579	1.513318	1.266405	2.577843	2.410404
S-FIG	0.413779	0.385239	1.162255	1.085455	2.007042	1.332636	3.566719	3.05504	1.363835	1.216428	2.125787	2.032867	2.193952	1.831989	3.110139	2.920210	2.232257	1.890945	3.162034	2.893292
FIAPA	0.222586	0.202381	1.066365	1.042479	0.790583	0.749758	2.568845	2.607587	1.378913	1.340661	2.187938	2.186849	1.541318	1.303895	2.593622	2.444961	1.852752	1.651425	2.887703	2.759846
S-FIAPA	0.580292	0.546950	1.271032	1.165222	1.714641	1.190567	3.18634	2.747471	1.525981	1.477618	2.296716	2.287452	1.947066	1.724404	2.939400	2.760531	2.508831	2.240344	3.401542	3.199885
FIEG	0.174135	0.155385	1.029610	1.018830	0.679971	0.458677	2.528944	2.425904	1.411327	1.362720	2.218522	2.204932	1.709937	1.464874	2.772992	2.614216	1.718731	1.497317	2.791098	2.648549
S-FIEG	0.696920	0.657990	1.337258	1.203635	2.380889	1.737470	3.690428	3.150316	1.547011	1.459849	2.309937	2.282294	2.017486	1.722746	2.924498	2.735832	2.634117	2.331235	3.475042	3.234165
FIMEG	0.202056	0.182270	1.055883	1.035566	0.671150	0.430918	2.507110	2.390097	1.487256	1.431795	2.274561	2.255441	1.666690	1.407869	2.710976	2.552756	1.799369	1.575496	2.855553	2.711105
S-FIMEG	0.765432	0.725540	1.402424	1.263157	2.479190	1.781279	3.779969	3.186796	1.620567	1.513227	2.364510	2.320342	1.963503	1.659093	2.858834	2.670694	2.685331	2.374526	3.514281	3.265522
FIMLogG	0.199517	0.180019	1.047402	1.034604	0.477280	0.300947	**2.373248**	**2.309867**	1.450378	1.459144	2.251283	2.305010	1.516625	1.298891	2.602872	2.453521	1.667041	1.450112	2.731327	2.594426
S-FIMLogG	0.787029	0.745988	1.412916	1.271162	2.293616	1.637702	3.585652	3.030365	1.601196	1.536028	2.368433	2.344144	1.799091	1.52646	2.721720	2.549320	2.516412	2.221650	3.349768	3.116491
FILogG	**0.032563**	**0.024761**	**0.984366**	1.099247	**0.424257**	**0.252012**	2.387039	2.324653	1.369228	1.208695	2.099148	2.088932	**1.365224**	1.374964	2.773313	2.613233	**1.083272**	**0.839074**	**2.220509**	**2.062938**
S-FILogG	0.479865	0.450029	1.172697	1.109436	2.136824	1.611387	3.539232	3.112379	1.252121	1.111975	2.062494	1.966626	1.951983	1.704463	2.983132	2.806931	1.934491	1.622611	2.869272	2.627154
Distr.	GED	GED	GED	s-t	GED	s-t	GED	s-t	GED	s-t	GED	s-t	GED	s-t	s-t	s-t	s-t	s-t	s-t	s-t

Table 3 – Model selection criteria based on forecasts of the 99% VaR and the 97.5% ES, evaluated using the **regulatory loss function** (**RLF**) and the **corrected loss function** (**CLF**). The considered volatility specifications include GARCH (G), APARCH (APA), EGARCH (EG), MEGARCH (MEG), Log-GARCH (LogG), and MLog-GARCH (MLogG). Models prefixed by S denote semiparametric specifications, while FI indicates the inclusion of a fractional integration term and hence long-memory dynamics. For each return series, the preferred model under the respective criterion is highlighted in bold, and the corresponding innovation distribution is reported in the Distr. row. All values are multiplied by 1000 for convenience. The prefix "s" before a distribution indicates the skewed variant.

Selection Criterion	ABEC				META				TSLA				AVGO				BRK-B			
	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES
G	1.956268	1.875784	2.954718	2.894292	2.584410	2.454310	3.939486	3.856663	4.934351	3.694930	7.100418	6.111558	1.937190	1.520830	3.517829	3.191540	0.111579	0.085565	0.655028	0.652111
S-G	1.638008	1.555835	2.602733	2.538504	2.698188	2.583415	4.059123	3.823001	7.734469	5.521621	9.518726	7.594026	2.140364	1.722425	3.721334	3.389705	0.286658	0.240191	0.750011	0.736367
APA	2.010166	1.93763	2.982137	2.929002	2.708807	2.585360	3.92743	3.731624	4.681693	3.348541	6.847601	5.760786	1.416983	1.002418	2.911961	2.580185	0.125032	0.100235	0.661831	0.659901
S-APA	1.716490	1.646263	2.674332	2.622081	2.524117	2.411497	3.837859	3.600909	7.266208	4.908130	9.018873	6.933628	1.430695	1.022681	2.926069	2.598501	0.310304	0.266878	0.760216	0.747101
EG	1.982731	1.912837	2.970189	2.920053	2.811053	2.686977	4.04388	3.845405	4.577900	3.326384	6.756881	5.750502	1.428334	1.024449	2.924599	2.604146	0.113459	0.088923	0.652397	0.651067
S-EG	1.740996	1.672026	2.696157	2.645271	2.501384	2.388316	3.812809	3.610447	7.283616	4.978460	9.024008	6.987478	1.407482	1.016303	2.909163	2.599097	0.286519	0.244422	0.734215	0.722195
MEG	1.993156	1.920955	2.971467	2.918883	2.600883	2.477451	3.846622	3.649498	4.435142	3.140167	6.619901	5.572321	1.293674	0.891484	2.816898	2.499464	0.120374	0.095090	0.655295	0.653806
S-MEG	1.741741	1.671766	2.691355	2.639397	2.463436	2.349985	3.777654	3.551884	7.160791	4.798988	8.905817	6.812589	1.326322	0.928703	2.844821	2.529047	0.287488	0.245096	0.731883	0.719891
MLogG	1.890939	1.817994	2.867878	2.814323	2.501738	2.379577	3.723116	3.503674	4.378198	3.091816	6.565542	5.534060	1.492452	1.055888	2.975937	2.622222	0.092096	0.069817	0.638705	0.638705
S-MLogG	1.693575	1.622797	2.64147	2.588576	2.445436	2.332403	3.745085	3.503377	7.184817	4.886761	8.908456	6.875454	1.454635	1.033120	2.945439	2.604820	0.258763	0.216491	0.709091	0.697748
LogG	1.505706	1.430972	2.553529	2.499468	**2.119889**	**1.999750**	**3.469239**	3.556173	4.107814	2.770543	6.365033	5.330418	1.545522	1.108782	3.030326	2.683298	**0.014550**	**0.005278**	**0.585116**	**0.590181**
S-LogG	1.626073	1.547944	2.607909	2.548346	2.829443	2.713887	4.153713	3.865029	9.044398	6.874266	10.69136	8.763123	1.546578	1.134244	3.061809	2.738660	0.179146	0.130005	0.665927	0.654033
FIG	1.615192	1.538984	2.643276	2.587145	2.384772	2.257305	3.790119	3.686167	4.299293	3.325382	6.49844	5.732494	1.438543	1.054734	3.100587	2.815391	0.112406	0.084847	0.646926	0.643098
S-FIG	1.531338	1.456278	2.531439	2.475315	2.711465	2.596684	4.069207	3.646570	7.266113	5.429166	9.111833	7.534853	1.374697	0.993267	3.060377	2.777757	0.203951	0.162880	0.691138	0.680598
FIAPA	1.872275	1.805431	2.874766	2.827544	2.361431	2.237600	3.661702	3.552814	4.199451	3.216055	6.393906	5.600160	1.229946	0.870002	2.770714	2.496426	0.103871	0.081297	0.634075	0.633158
S-FIAPA	1.747350	1.679246	2.741102	2.692029	2.550805	2.436887	3.893323	3.450287	7.011952	5.012655	8.844240	7.081432	1.061977	0.725179	2.631475	2.381019	0.216401	0.179220	0.679127	0.669419
FIEG	1.745468	1.676805	2.768825	2.720344	2.641278	2.518877	3.908089	3.543435	4.266438	3.176695	6.495264	5.643664	1.143762	0.782563	2.693263	2.418436	0.125438	0.099748	0.635737	0.633479
S-FIEG	1.738419	1.669435	2.693985	2.643093	2.496470	2.383378	3.808832	3.602547	7.234647	5.163874	8.979032	7.270561	1.159784	0.808073	2.700934	2.433852	0.193321	0.159100	0.659099	0.651321
FIMEG	1.784736	1.713582	2.792392	2.741186	2.463754	2.341427	3.740390	3.398423	4.103234	2.974971	6.338443	5.453089	0.991305	0.638026	2.574675	2.309753	0.128704	0.102419	0.636603	0.634126
S-FIMEG	1.738970	1.668991	2.689044	2.637092	2.452359	2.338892	3.768656	3.545388	7.105646	4.958207	8.855206	7.074095	**1.335149**	**0.937126**	**2.851674**	**2.535109**	0.195165	0.160334	0.658922	0.651002
FIMLogG	1.689990	1.618616	2.698963	2.647538	2.392979	2.273551	3.646401	**3.375876**	4.110985	2.921608	6.357491	5.421237	1.533228	1.093096	3.009696	2.651513	0.096156	0.073745	0.618968	0.618876
S-FIMLogG	1.691010	1.620229	2.639324	2.586436	2.440086	2.327055	3.740098	3.496828	7.133995	4.773650	8.861781	6.776825	1.463783	1.041862	2.952748	2.611400	0.172761	0.139126	0.639601	0.632850
FILogG	**0.566963**	**0.535016**	**1.821450**	**1.807419**	3.037305	2.899829	4.335827	3.615517	**2.863988**	**2.144400**	**5.272571**	**4.775483**	2.148026	1.606164	3.545324	3.089294	0.035338	0.019657	0.677680	0.674402
S-FILogG	1.581721	1.50463	2.592380	2.534695	2.341748	2.219468	3.669140	3.692592	9.202845	7.230790	10.87205	9.141077	1.322546	0.933002	2.867460	2.570102	0.128246	0.088427	0.622910	0.614277
Distr.	GED	GED	GED	GED	s-GED	s-GED	s-GED	s-t	s-GED	s-t	s-GED	s-t	s-t	s-t	s-t	s-t	t	t	s-AL	s-AL

Table 4 - Model selection criteria based on forecasts of the 99% VaR and the 97.5% ES, evaluated using the **regulatory loss function** (**RLF**) and the **corrected loss function** (**CLF**). The considered volatility specifications include GARCH (G), APARCH (APA), EGARCH (EG), MEGARCH (MEG), Log-GARCH (LogG), and MLog-GARCH (MLogG). Models prefixed by S denote semiparametric specifications, while FI indicates the inclusion of a fractional integration term and hence long-memory dynamics. For each return series, the preferred model under the respective criterion is highlighted in bold, and the corresponding innovation distribution is reported in the Distr. row. All values are multiplied by 1000 for convenience. The prefix "s" before a distribution indicates the skewed variant.

	LLY				WMT				JPM				V				MA			
Selection Criterion	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES
G	1.612841	1.302607	2.675945	2.453376	**0.000000**	**0.000000**	0.609848	0.621099	2.148002	1.953088	2.97834	2.821624	0.785414	0.658109	1.500125	1.419285	0.057108	0.031182	0.769694	0.76491
S-G	1.681743	1.390167	2.615828	2.398047	0.011089	0.002051	0.531842	0.535998	3.243053	3.018668	3.929208	3.736658	1.410385	1.189518	1.986142	1.805246	0.362293	0.256140	0.998334	0.970208
APA	1.53018	1.234985	2.579128	2.368361	0.001830	**0.000000**	0.593786	0.603027	2.271878	2.082327	2.990038	2.830244	0.810419	0.704053	1.508249	1.444559	0.124651	0.075090	0.888840	0.878297
S-APA	1.610562	1.325729	2.538469	2.326317	0.021344	0.011411	0.531372	0.536670	3.713127	3.456840	4.272893	4.040660	1.541262	1.346336	2.080059	1.919599	0.742432	0.633794	1.353034	1.322403
EG	1.505900	1.212737	2.544458	2.335155	0.000708	**0.000000**	0.592700	0.602340	2.246903	2.059602	2.987996	2.831781	0.830008	0.741655	1.541713	1.494453	0.148914	0.094382	0.910582	0.899285
S-EG	1.634728	1.348906	2.542802	2.328391	0.018909	0.009567	**0.527160**	**0.532460**	3.616097	3.360198	4.194103	3.963026	1.481000	1.298059	2.040495	1.891591	0.746116	0.627850	1.359601	1.329549
MEG	1.522893	1.227762	2.559125	2.347518	0.000637	**0.000000**	0.596893	0.606733	2.351630	2.179021	3.073956	2.931342	0.841455	0.743644	1.552711	1.497167	0.130600	0.079837	0.888023	0.877588
S-MEG	1.672348	1.383963	2.578989	2.361805	0.018245	0.008976	0.529882	0.535146	3.550301	3.325734	4.112197	3.911612	1.546447	1.357893	2.097187	1.942184	0.655806	0.536681	1.265256	1.235435
MLogG	1.486886	1.201468	2.517545	2.316471	0.003261	0.000138	0.600539	0.609618	2.343488	2.148644	3.039678	2.874640	0.820007	0.713966	1.512793	1.448915	0.108247	0.062348	0.858908	0.849645
S-MLogG	1.663424	1.378205	2.563712	2.350236	0.023090	0.012851	0.54025	0.545998	3.827897	3.558902	4.353892	4.108239	1.518118	1.334336	2.052157	1.900862	0.670001	0.552008	1.271416	1.242552
LogG	1.746493	1.420442	2.711345	2.468197	**0.000000**	**0.000000**	0.624207	0.633607	**1.654052**	**1.423087**	**2.420281**	**2.234104**	**0.530824**	0.552320	**1.212000**	1.233900	**0.009490**	0.006506	**0.705549**	**0.706620**
S-LogG	1.946581	1.604636	2.802061	2.531439	0.014540	0.005690	0.552374	0.558600	3.596135	3.273499	4.13885	3.846788	1.120046	0.919522	1.657577	1.496978	0.378478	0.260081	1.012471	0.985481
FIG	**0.996836**	**0.760495**	**2.076462**	**1.923897**	**0.000000**	**0.000000**	0.630337	0.641930	2.167528	1.980895	2.983292	2.834952	0.872987	0.732611	1.564182	1.467475	0.064673	0.033913	0.769970	0.762231
S-FIG	1.089887	0.822394	2.129092	1.949424	0.003767	0.000094	0.534650	0.540786	2.828124	2.619240	3.553559	3.380169	1.179915	0.928819	1.798508	1.601064	0.226125	0.133863	0.900490	0.878400
FIAPA	1.099332	0.866854	2.128073	1.972323	0.000784	**0.000000**	0.599215	0.608401	2.278575	2.102582	3.012203	2.866092	0.739150	0.619551	1.445622	1.366500	0.187285	0.140966	0.964312	0.951202
S-FIAPA	1.190307	0.931334	2.185766	2.007158	0.014542	0.006328	0.554195	0.560824	3.423875	3.189426	4.006978	3.797744	1.406810	1.202761	1.964002	1.795715	0.649918	0.565620	1.285310	1.258155
FIEG	1.053284	0.811514	2.127456	1.969047	0.000024	**0.000000**	0.592786	0.602198	2.546779	2.360113	3.247868	3.091515	0.865107	0.727976	1.540082	1.442754	0.249306	0.184349	0.986017	0.970722
S-FIEG	1.633983	1.348178	2.542186	2.327811	0.004121	0.000346	0.551355	0.558891	3.251084	3.028475	3.855447	3.658792	1.291305	1.084874	1.87104	1.702031	0.747835	0.630588	1.360985	1.330895
FIMEG	1.074308	0.829181	2.147825	1.985979	0.000178	**0.000000**	0.593486	0.602862	2.645712	2.472608	3.319087	3.174912	0.89466	0.752717	1.563285	1.461023	0.217830	0.155183	0.954233	0.939254
S-FIMEG	1.671883	1.383540	2.578587	2.361447	0.004534	0.000487	0.553446	0.560999	3.268915	3.065259	3.852938	3.674304	1.370634	1.155735	1.938873	1.761913	0.657436	0.539370	1.266619	1.236751
FIMLogG	1.029546	0.791852	2.102097	1.950889	0.002768	0.000039	0.587997	0.595962	2.592942	2.396887	3.241047	3.073399	0.849476	0.714811	1.505936	1.409597	0.203028	0.141939	0.941203	0.926618
S-FIMLogG	1.663666	1.378528	2.563883	2.350457	0.008369	0.002302	0.560366	0.567042	3.821929	3.552995	4.348408	4.102847	1.521262	1.340096	2.054873	1.905682	0.671496	0.554438	1.272808	1.243898
FILogG	1.516797	1.234803	2.501259	2.302528	0.000782	**0.000000**	0.611429	0.619623	1.747524	1.534572	2.504953	2.336112	0.554089	**0.456432**	1.222441	**1.167104**	0.025503	**0.001585**	0.727869	0.722894
S-FILogG	1.570595	1.254241	2.452557	2.210336	0.009905	0.002693	0.566519	0.573289	3.198384	2.896921	3.771372	3.502796	1.035276	0.881257	1.589856	1.471811	0.323039	0.227992	0.985134	0.960300
Distr.	s-t	s-t	s-t	s-t	t	t	AL	AL	t	t	t	t	s-GED	s-t	s-GED	s-t	s-GED	s-t	AL	AL

Table 5 - Model selection criteria based on forecasts of the 99% VaR and the 97.5% ES, evaluated using the **regulatory loss function** (**RLF**) and the **corrected loss function** (**CLF**). The considered volatility specifications include GARCH (G), APARCH (APA), EGARCH (EG), MEGARCH (MEG), Log-GARCH (LogG), and MLog-GARCH (MLogG). Models prefixed by S denote semiparametric specifications, while FI indicates the inclusion of a fractional integration term and hence long-memory dynamics. For each return series, the preferred model under the respective criterion is highlighted in bold, and the corresponding innovation distribution is reported in the Distr. row. All values are multiplied by 1000 for convenience. The prefix “s” before a distribution indicates the skewed variant.

	XOM				ORCL				UNH				COST				PG			
Selection Criterion	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES	RLF VaR	RLF ES	CLF VaR	CLF ES
G	0.067249	0.048001	0.760996	0.764834	0.762116	0.691892	1.887334	1.761153	4.650485	4.537619	5.552595	5.455315	2.835238	2.772558	3.542879	3.493285	0.690707	0.593867	1.269539	1.208609
S-G	0.123718	0.093465	0.765721	0.761332	0.806066	0.734836	1.870442	1.720548	6.534571	6.393463	7.236426	7.107140	1.433523	1.379292	2.401730	2.364540	0.754071	0.656113	1.307386	1.243332
APA	0.092072	0.074625	0.795932	0.800582	0.583277	0.526564	1.635197	1.549621	4.712466	4.610660	5.609019	5.521941	3.384172	3.326398	4.082739	4.037467	0.625802	0.528938	1.194896	1.130872
S-APA	0.142477	0.115258	0.776568	0.773257	0.682957	0.618706	1.689584	1.581186	6.602200	6.469183	7.308403	7.186781	2.244957	2.188083	3.157678	3.116466	0.755487	0.651783	1.287565	1.214184
EG	0.082158	0.065309	0.790040	0.794898	0.666017	0.603557	1.732442	1.627254	4.642878	4.542447	5.564929	5.479571	3.396516	3.338482	4.109431	4.064220	0.636806	0.543250	1.215704	1.215704
S-EG	0.135640	0.109325	0.773810	0.770870	0.763719	0.696212	1.775848	1.653709	6.481451	6.348857	7.195931	7.074874	2.356948	2.299639	3.268683	3.227089	0.756291	0.654710	1.299562	1.229234
MEG	0.087966	0.071222	0.796354	0.800831	0.655120	0.593080	1.736912	1.639565	4.613184	4.514010	5.519799	5.435543	3.364009	3.306010	4.070583	4.025321	0.632115	0.537969	1.207949	1.147353
S-MEG	0.140385	0.113584	0.778721	0.775646	0.758407	0.690899	1.782076	1.671183	6.491386	6.349589	7.191041	7.060617	2.427142	2.370709	3.331811	3.290956	0.760800	0.659049	1.299066	1.228209
MLogG	0.085197	0.068163	0.796279	0.800883	0.573813	0.518258	1.632416	1.557507	4.695522	4.589029	5.567703	5.475619	3.32308	3.264903	4.031494	3.986082	0.571872	0.474718	1.142755	1.078930
S-MLogG	0.137922	0.111728	0.775493	0.772784	0.656415	0.594169	1.666553	1.567860	6.694983	6.542958	7.364507	7.223508	2.413243	2.356694	3.316498	3.275511	0.709640	0.605638	1.242817	1.169416
LogG	0.061530	0.044982	0.763675	0.783566	**0.174341**	**0.145974**	**1.400586**	1.481476	5.061827	4.922339	5.869135	5.744211	2.509504	2.444191	3.274763	3.224077	0.434858	0.329520	**0.992724**	**0.924974**
S-LogG	0.107133	0.083144	**0.744915**	**0.744516**	0.514506	0.460830	1.529749	1.429299	7.180169	6.996168	7.793243	7.620346	1.497363	1.441637	2.476966	2.439045	0.519604	0.401301	1.072506	0.992334
FIG	0.046986	0.030458	0.769949	0.774033	0.421868	0.368753	1.668128	1.630669	4.400773	4.284758	5.322282	5.222351	2.380285	2.315738	3.133066	3.082318	0.628904	0.534350	1.233015	1.177169
S-FIG	0.132185	0.099707	0.793513	0.787999	0.796005	0.725236	1.863556	1.576657	6.417294	6.277641	7.125736	6.998001	1.507819	1.452043	2.465684	2.426815	0.676256	0.571348	1.264616	1.199937
FIAPA	**0.027571**	**0.016669**	0.779984	0.786493	0.509490	0.456095	1.583120	1.512880	4.700403	4.596317	5.622066	5.533257	3.110939	3.051965	3.830352	3.784076	0.632132	0.543558	1.208770	1.152344
S-FIAPA	0.144420	0.117314	0.797489	0.794927	0.577414	0.517925	1.619905	1.527057	6.065162	5.943133	6.816828	6.707015	2.21293	2.154757	3.112323	3.069627	0.708499	0.606983	1.255341	1.185757
FIEG	0.079626	0.063007	0.805647	0.810525	0.525458	0.469830	1.630391	1.531016	4.791862	4.689745	5.708716	5.621718	3.119037	3.058455	3.837256	3.789628	0.672883	0.582138	1.246743	1.188772
S-FIEG	0.136097	0.111881	0.785311	0.783457	0.764897	0.697370	1.776631	1.654097	6.480452	6.347866	7.194902	7.073849	2.361031	2.303736	3.272101	3.230503	0.753636	0.651952	1.297740	1.227429
FIMEG	0.086014	0.069126	0.812046	0.816062	0.517670	0.462274	1.632475	1.536871	4.801117	4.697747	5.699659	5.611103	3.114998	3.054435	3.823951	3.776208	0.672685	0.581459	1.241788	1.183047
S-FIMEG	0.140396	0.113603	0.778834	0.775754	0.759753	0.692223	1.782976	1.671784	6.491804	6.350027	7.191353	7.060943	2.431178	2.374759	3.335170	3.294311	0.759047	0.657228	1.297917	1.227071
FIMLogG	0.082286	0.065545	0.817331	0.821660	0.577192	0.521459	1.634402	1.464960	4.891679	4.779405	5.759255	5.661429	3.039351	2.978026	3.754491	3.706162	0.587100	0.491697	1.155338	1.093224
S-FIMLogG	0.138040	0.111842	0.775641	0.772925	0.657638	0.595355	1.667363	1.497209	6.694993	6.543006	7.364466	7.223501	2.417309	2.360770	3.319922	3.278928	0.708193	0.604172	1.241891	1.168524
FILogG	0.034406	0.021046	0.777563	0.805560	0.393984	0.349555	1.482162	**1.411517**	**3.085962**	**2.963059**	**4.047124**	**3.941686**	**1.078934**	**1.016779**	**2.093341**	**2.051123**	**0.431718**	**0.324518**	1.002491	0.934366
S-FILogG	0.129826	0.104229	0.775622	0.774965	0.398243	0.350284	1.448828	1.367462	6.810954	6.627272	7.437525	7.265271	1.39484	1.337281	2.360041	2.320211	0.507161	0.388573	1.065976	0.986734
Distr.	s-t	s-t	s-t	t	s-GED	s-GED	s-GED	s-t	s-GED	s-GED	s-GED	s-GED	s-GED	s-GED	s-GED	s-GED	s-t	s-t	s-t	s-t

Table 6 - Model selection criteria based on forecasts of the 99% VaR and the 97.5% ES, evaluated using the **regulatory loss function** (**RLF**) and the **corrected loss function** (**CLF**). The considered volatility specifications include GARCH (G), APARCH (APA), EGARCH (EG), MEGARCH (MEG), Log-GARCH (LogG), and MLog-GARCH (MLogG). Models prefixed by S denote semiparametric specifications, while FI indicates the inclusion of a fractional integration term and hence long-memory dynamics. For each return series, the preferred model under the respective criterion is highlighted in bold, and the corresponding innovation distribution is reported in the Distr. row. All values are multiplied by 1000 for convenience. The prefix "s" before a distribution indicates the skewed variant.

Selection Criterion	**AAPL** N_1	N_2	ES	T_{ES}	WAD	**NVDA** N_1	N_2	ES	T_{ES}	WAD	**MSFT** N_1	N_2	ES	T_{ES}	WAD	**AMZN** N_1	N_2	ES	T_{ES}	WAD	**ABEA** N_1	N_2	ES	T_{ES}	WAD
G	6	2	2	2.96057	0.29264	7	2	2	3.12606	0.32034	10	6	4	5.35158	2.71251	4	1	1	2.14943	1.27218	8	4	4	4.89829	4.89829
S-G	11	5	5	6.37082	2.79866	12	7	7	7.93762	4.26004	12	5	4	5.45910	2.66691	8	2	2	3.19250	0.50160	10	7	7	7.00206	7.00206
APA	6	2	2	3.40964	0.33108	10	1	1	2.90108	1.27165	8	5	5	5.18908	1.94050	5	1	1	2.42564	1.02380	8	3	3	3.93822	3.93822
S-APA	13	8	8	8.77642	5.08846	12	10	9	8.67570	5.69622	8	6	5	5.41774	2.41368	5	3	2	3.24999	0.44000	12	5	5	7.32186	7.32186
EG	6	2	2	3.22631	0.27242	9	2	1	2.95147	0.69553	7	6	4	4.96254	2.10801	4	2	1	2.38379	0.79719	8	4	3	3.91896	3.91896
S-EG	13	7	7	8.46518	4.58886	12	12	12	8.99915	6.59973	9	6	5	5.15135	2.48843	6	2	2	3.15793	0.25054	12	5	5	7.50993	7.50993
MEG	6	2	2	3.38717	0.32390	10	2	1	3.02661	0.83148	8	5	4	5.21344	1.94830	4	2	1	2.44020	0.77914	8	4	3	4.00680	4.00680
S-MEG	13	8	8	8.66113	5.05156	12	12	12	9.07571	6.62423	8	6	5	5.48065	2.43381	6	3	2	3.17616	0.25637	11	7	5	7.14097	7.14097
MLogG	6	2	2	3.61739	0.39756	9	1	1	2.76564	1.15500	8	5	5	5.16692	1.93342	5	1	1	2.38511	1.03677	8	3	3	3.72650	3.72650
S-MLogG	14	8	8	8.81081	5.25946	13	11	11	9.04034	6.37291	8	5	5	5.41092	2.01150	5	3	2	3.14131	0.40522	12	6	4	7.16920	7.16920
LogG	7	3	3	3.98160	0.59411	7	1	1	2.02982	1.07046	**6**	**3**	**2**	**3.68463**	**0.41908**	5	1	1	1.32908	1.37470	9	4	3	3.65833	3.65833
S-LogG	10	5	5	6.06738	2.54156	17	8	8	9.86251	6.07600	8	5	3	3.97582	1.55226	6	1	1	2.30890	0.90115	12	6	6	6.98968	6.98968
FIG	5	2	2	2.76339	0.51571	**6**	**2**	**2**	**2.91531**	**0.30710**	11	5	4	5.21389	2.42845	4	2	1	2.19212	0.85852	7	4	4	4.37760	4.37760
S-FIG	9	3	3	4.44196	1.06143	11	5	5	6.26905	2.76610	12	6	4	5.63102	3.12193	4	2	1	2.36080	0.80455	12	7	7	5.93817	5.93817
FIAPA	6	2	2	3.29073	0.29303	10	2	2	2.91812	0.86620	8	6	5	5.16189	2.33181	5	2	1	2.50806	0.59742	8	3	3	3.43678	3.43678
S-FIAPA	13	6	6	6.76287	3.64412	12	6	5	7.04735	3.57515	8	6	6	5.51854	2.44593	5	2	2	3.11485	0.40325	11	4	4	5.60153	5.60153
FIEG	**6**	**2**	**2**	**3.16709**	**0.25347**	9	1	1	2.64579	1.19335	8	6	6	5.07960	2.30547	4	2	1	2.33821	0.81177	7	2	2	3.43609	3.43609
S-FIEG	13	7	7	8.48001	4.59360	12	12	12	9.00761	6.60243	9	6	5	5.15860	2.49075	**6**	**2**	**2**	**3.15721**	**0.25031**	12	5	5	7.33981	7.33981
FIMEG	6	2	2	3.34191	0.30941	9	1	1	2.78773	1.14793	8	6	6	5.43599	2.41952	4	2	1	2.41760	0.78637	6	3	3	3.61798	3.61798
S-FIMEG	13	8	8	8.67627	5.05641	12	12	12	9.08319	6.62662	8	6	5	5.48789	2.43613	6	3	2	3.17618	0.25638	10	7	5	7.13641	7.13641
FIMLogG	6	2	2	3.56664	0.38133	7	1	1	2.25370	0.99882	8	5	5	5.18325	1.93864	5	1	1	2.42581	1.02374	7	3	3	3.22415	3.22415
S-FIMLogG	14	8	8	8.82570	5.26423	13	11	11	9.05222	6.37671	8	5	5	5.41929	2.01417	5	3	2	3.14232	0.40554	12	6	4	7.15153	7.15153
FILogG	5	2	1	2.50148	0.59953	7	2	2	2.36666	0.56267	11	4	3	4.64964	1.84788	3	1	1	1.32107	1.69726	**6**	**2**	**2**	**3.15224**	**3.15224**
S-FILogG	10	4	4	5.04907	1.81570	15	7	7	7.94108	4.74115	9	5	4	4.14949	1.76784	8	1	1	2.27758	1.15117	10	4	4	5.74400	5.74400
Distr.	s-GED					s-AL					t					s-t					s-GED				

Table 7 - Model selection criteria based on forecasts of the 97.5% and 99% VaR as well as the 97.5% ES, evaluated using **Weighted Absolute Deviation (WAD)** . The considered volatility specifications include GARCH (G), APARCH (APA), EGARCH (EG), MEGARCH (MEG), Log-GARCH (LogG), and MLog-GARCH (MLogG). Models prefixed by S denote semiparametric specifications, while FI indicates the inclusion of a fractional integration term and hence long-memory dynamics. For each return series, the preferred model under the respective criterion is highlighted in bold, and the corresponding innovation distribution is reported in the Distr. row. The prefix "s" before a distribution indicates the skewed variant.

	ABEC					META					TSLA					AVGO					BRK-B				
Selection Criterion	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD
G	8	5	2	4.75944	1.80302	6	2	1	2.90766	0.30955	7	3	2	3.97138	0.59084	5	3	3	3.49220	0.51750	6	4	4	3.99875	0.91960
S-G	7	5	3	4.76417	1.64453	3	1	1	2.03714	1.46811	8	6	4	5.34219	2.38950	5	3	3	3.56879	0.54201	8	6	6	5.31230	2.37993
APA	10	3	2	3.88517	1.04325	9	1	1	3.61166	1.19573	7	3	2	4.01678	0.60537	9	4	4	4.76208	1.56386	6	4	3	3.92072	0.89463
S-APA	11	3	2	4.66225	1.45192	5	1	1	1.79843	1.22450	7	6	5	5.31260	2.22003	9	4	4	4.79351	1.57392	9	6	6	5.83421	2.70695
EG	9	2	2	3.74708	0.83906	8	2	1	3.39560	0.56659	7	2	2	3.93961	0.58068	9	4	3	4.66558	1.53299	6	4	3	3.80699	0.85824
S-EG	11	3	2	4.39721	1.36711	5	1	1	1.81094	1.22050	7	7	5	5.32669	2.62454	9	4	3	4.65339	1.52909	9	6	6	5.41291	2.57213
MEG	10	3	2	3.89894	1.04766	9	1	1	3.22630	1.07242	7	2	2	4.05403	0.61729	8	4	4	4.23173	1.23416	6	4	3	3.78125	0.85000
S-MEG	10	3	3	4.22702	1.15265	5	1	1	1.79139	1.22675	7	7	5	5.39280	2.64569	8	4	4	4.21174	1.22776	9	6	6	5.29949	2.53584
MLogG	10	3	3	3.68491	0.97917	8	1	1	3.61157	1.03570	7	3	2	3.75664	0.52213	9	4	4	5.00851	1.64272	**6**	**3**	**3**	**3.60212**	**0.39268**
S-MLogG	10	3	3	4.14163	1.12532	5	1	1	1.91519	1.18714	7	5	5	5.30431	1.81738	9	4	4	4.95081	1.62426	9	6	6	5.41661	2.57332
LogG	7	3	3	3.73563	0.51540	7	2	1	3.17389	0.33565	5	2	2	3.16028	0.41129	9	4	3	4.66327	1.53225	5	3	3	3.22420	0.43174
S-LogG	7	3	3	4.07503	0.62401	4	2	1	2.35095	0.80769	8	6	5	5.38499	2.40320	8	4	3	4.30949	1.25904	10	5	5	5.39917	2.32773
FIG	7	5	3	4.57268	1.58326	6	1	1	2.17038	0.94548	7	3	2	4.14543	0.64654	5	3	3	3.19097	0.42111	6	4	4	4.02015	0.92645
S-FIG	7	3	2	4.37544	0.72014	3	1	1	2.04559	1.46541	8	5	4	5.14253	1.92561	**5**	**3**	**3**	**3.17773**	**0.41687**	7	5	5	4.98207	1.71426
FIAPA	9	2	2	3.56312	0.78020	7	1	1	2.10555	1.04623	7	3	2	4.38526	0.72328	8	3	3	4.14335	1.20587	6	3	3	3.87156	0.47890
S-FIAPA	8	3	2	3.74239	0.67756	5	1	1	1.82170	1.21706	7	5	5	5.21414	1.78852	7	4	3	3.91565	0.97301	9	6	6	5.16525	2.49288
FIEG	8	2	2	3.57780	0.62489	7	1	1	2.88621	0.79641	7	3	2	3.84815	0.55141	7	4	3	4.00996	1.00319	6	4	3	4.10131	0.95242
S-FIEG	11	3	2	4.38291	1.36253	5	1	1	1.80426	1.22264	7	5	5	4.94834	1.70347	7	4	3	4.04804	1.01537	7	5	5	4.75792	1.64254
FIMEG	8	3	2	3.67354	0.65553	7	1	1	2.68586	0.86052	7	4	2	3.94922	0.98375	7	4	4	3.70848	0.90671	6	4	3	4.11195	0.95583
S-FIMEG	10	3	3	4.21090	1.14749	5	1	1	1.77637	1.23156	7	5	5	5.01582	1.72506	8	4	4	4.22520	1.23207	7	5	5	4.73587	1.63548
FIMLogG	8	3	2	3.49553	0.59857	8	1	1	2.95807	0.93342	6	3	2	3.52838	0.36908	9	4	4	5.06210	1.65987	6	3	3	3.84925	0.47176
S-FIMLogG	10	3	3	4.12731	1.12074	5	1	1	1.91490	1.18723	7	5	5	5.27216	1.80709	9	4	4	4.96688	1.62940	6	6	6	4.66908	1.93410
FILogG	**6**	**2**	**2**	**2.75423**	**0.35865**	**6**	**3**	**2**	**3.11565**	**0.24299**	**6**	**2**	**2**	**2.93793**	**0.29986**	10	5	4	4.98859	2.19635	6	4	3	3.76326	0.84424
S-FILogG	7	3	2	3.89055	0.56498	3	2	2	2.10680	1.04582	8	5	4	5.24314	1.95780	8	4	3	3.88672	1.12375	7	5	5	4.59869	1.59158
Distr.	s-t					s-GED					t					s-t					s-AL				

Table 8 - Model selection criteria based on forecasts of the 97.5% and 99% VaR as well as the 97.5% ES, evaluated using **Weighted Absolute Deviation (WAD)** . The considered volatility specifications include GARCH (G), APARCH (APA), EGARCH (EG), MEGARCH (MEG), Log-GARCH (LogG), and MLog-GARCH (MLogG). Models prefixed by S denote semiparametric specifications, while FI indicates the inclusion of a fractional integration term and hence long-memory dynamics. For each return series, the preferred model under the respective criterion is highlighted in bold, and the corresponding innovation distribution is reported in the Distr. row. The prefix "s" before a distribution indicates the skewed variant.

	LLY					WMT					JPM					V					MA				
Selection Criterion	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD
G	7	3	3	3.88227	0.56233	6	0	0	1.67725	1.50328	7	5	5	4.78478	1.65113	7	4	4	4.50691	1.16221	6	3	3	3.30305	0.29698
S-G	11	5	4	5.90372	2.64919	12	4	3	4.91958	2.09427	10	6	6	7.66211	3.45188	8	7	7	6.62234	3.19915	12	7	6	7.19310	4.02179
APA	8	3	3	4.01149	0.76368	5	1	1	1.67041	1.26547	9	7	7	6.36295	3.27614	7	5	5	4.75399	1.64128	4	3	3	2.57021	0.73753
S-APA	10	3	3	5.63246	1.60239	10	3	2	4.39217	1.20549	14	10	10	10.39402	6.56609	7	7	7	6.40236	2.96875	11	7	7	6.36705	3.59746
EG	8	3	3	3.99468	0.75830	5	1	1	1.44953	1.33615	9	7	7	6.23042	3.23373	7	5	5	4.68037	1.61772	4	3	3	2.74658	0.68109
S-EG	12	3	3	5.78898	1.97247	10	3	2	4.39109	1.20515	14	9	9	10.13033	6.08171	7	7	7	6.33995	2.94879	9	7	7	6.05698	3.17823
MEG	8	3	3	4.05519	0.77766	4	1	1	1.41450	1.50736	9	5	5	6.27889	2.44925	7	6	5	4.79847	2.05551	4	3	3	2.69916	0.69627
S-MEG	11	3	3	5.81891	1.82205	10	3	3	4.36366	1.19637	14	10	8	10.17681	6.49658	7	7	7	6.42123	2.97479	11	7	7	6.23530	3.55530
MLogG	10	3	3	4.35485	1.19355	5	1	1	1.65506	1.27038	9	6	6	6.59210	2.94947	7	5	4	4.73896	1.63647	4	3	3	2.57529	0.73591
S-MLogG	11	3	3	6.23875	1.95640	9	2	2	3.89901	0.88768	14	10	10	10.85290	6.71293	8	7	7	6.49671	3.15895	12	7	6	6.34117	3.74917
LogG	12	5	5	5.99750	2.83920	5	1	1	1.76976	1.23368	7	5	5	4.99915	1.71973	6	5	5	4.39305	1.44578	6	3	3	2.79318	0.34618
S-LogG	12	7	5	7.40444	4.08942	6	1	1	3.28681	0.69178	13	9	9	9.95107	5.86434	8	6	6	6.08340	2.62669	13	7	7	6.83294	4.06654
FIG	7	3	3	3.66836	0.49388	3	0	0	0.86176	2.24424	7	5	5	4.63958	1.60467	7	6	4	4.95481	2.10554	6	3	3	3.79011	0.45283
S-FIG	9	3	3	4.50582	1.08186	10	2	2	3.98045	1.07374	10	5	5	6.32583	2.62427	8	7	7	6.41700	3.13344	10	6	6	5.83317	2.86661
FIAPA	9	3	3	3.50656	0.76210	4	1	1	1.59536	1.44948	10	7	6	6.07614	3.34436	7	4	6	4.78339	1.25069	5	3	3	2.75107	0.51966
S-FIAPA	9	3	3	4.32459	1.02387	8	1	1	3.50561	1.00179	14	10	10	9.31470	6.22070	7	7	7	6.35705	2.95426	9	6	6	5.63135	2.64203
FIEG	8	3	3	3.35460	0.55347	4	1	1	1.56335	1.45973	10	7	7	7.00370	3.64119	7	6	6	5.40025	2.24808	6	3	3	3.37643	0.32046
S-FIEG	12	3	3	5.78651	1.97168	8	1	1	3.07952	0.89456	14	10	10	9.21272	6.18807	7	7	7	6.29827	2.93545	9	7	7	6.06539	3.18092
FIMEG	**7**	**3**	**3**	**3.40539**	**0.40972**	7	1	1	1.67384	1.18437	10	8	8	6.97503	4.03201	7	6	5	5.51810	2.28579	6	3	3	3.39086	0.32508
S-FIMEG	11	3	3	5.81641	1.82125	9	1	1	3.21767	1.06965	14	10	8	10.16726	6.49352	7	7	7	6.36950	2.95824	12	7	7	6.24851	3.71952
FIMLogG	8	3	3	3.12715	0.48069	6	1	1	2.32744	0.89522	10	8	7	7.41989	4.17436	7	5	4	5.37213	1.83908	**6**	**3**	**3**	**3.12867**	**0.24117**
S-FIMLogG	11	3	3	6.23744	1.95598	9	2	2	3.30043	0.69614	14	10	10	10.84754	6.71121	8	7	6	6.50788	3.16252	12	7	6	6.35531	3.75370
FILogG	10	4	3	4.67971	1.69751	5	2	2	2.12105	0.72126	**6**	**5**	**5**	**4.69828**	**1.54345**	**7**	**4**	**4**	**4.20039**	**1.06412**	7	3	3	3.09743	0.32882
S-FILogG	13	5	4	6.44207	3.14146	**6**	**2**	**1**	**3.03567**	**0.26859**	13	8	7	8.69393	5.06206	8	6	6	5.97506	2.59202	15	7	7	7.83980	4.70874
Distr.	s-GED					AL					GED					s-GED					s-AL				

Table 9 - Model selection criteria based on forecasts of the 97.5% and 99% VaR as well as the 97.5% ES, evaluated using **Weighted Absolute Deviation (WAD)** . The considered volatility specifications include GARCH (G), APARCH (APA), EGARCH (EG), MEGARCH (MEG), Log-GARCH (LogG), and MLog-GARCH (MLogG). Models prefixed by S denote semiparametric specifications, while FI indicates the inclusion of a fractional integration term and hence long-memory dynamics. For each return series, the preferred model under the respective criterion is highlighted in bold, and the corresponding innovation distribution is reported in the Distr. row. The prefix "s" before a distribution indicates the skewed variant.

	XOM					ORCL					UNH					COST					PG				
Selection Criterion	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD	N_1	N_2	ES	T_{ES}	WAD
G	5	4	3	3.12521	0.80007	8	3	3	3.93075	0.73784	10	6	5	6.54290	3.09373	4	2	2	2.41854	0.78607	5	3	3	3.70946	0.58703
S-G	9	4	4	4.02554	1.32817	8	4	3	4.15284	1.20891	13	10	9	9.11461	5.99668	2	2	2	1.77480	1.31206	5	4	3	3.89072	1.04503
APA	6	3	3	3.14297	0.24575	7	3	3	4.03749	0.61200	11	5	4	6.74705	2.91906	4	2	2	2.55803	0.74143	6	5	5	4.13638	1.36364
S-APA	8	5	5	4.62575	1.76024	7	5	3	4.22731	1.47274	12	11	9	9.55253	6.37681	2	2	2	1.88502	1.27679	6	5	5	4.41970	1.45430
EG	6	3	2	3.08014	0.25436	7	3	3	3.64391	0.48605	11	5	4	6.47532	2.83210	4	2	2	2.42995	0.78242	6	5	4	4.12419	1.35974
S-EG	8	5	4	4.47072	1.71063	8	4	3	3.96701	1.14944	12	11	9	9.48907	6.35650	2	2	2	1.89753	1.27279	6	5	5	4.41972	1.45431
MEG	6	3	2	3.17174	0.25496	7	4	3	3.64198	0.88543	11	4	4	6.67659	2.49651	4	2	2	2.50671	0.75785	6	5	5	4.18537	1.37932
S-MEG	7	5	5	4.58185	1.58619	7	4	3	3.98954	0.99665	13	11	10	9.54283	6.53371	2	2	2	1.88992	1.27523	6	5	5	4.46758	1.46962
MLogG	5	3	3	3.06075	0.42056	6	3	3	3.90652	0.49009	11	6	4	7.16292	3.45213	4	2	2	2.60238	0.72724	6	5	5	4.18340	1.37869
S-MLogG	8	5	4	4.48182	1.71418	7	3	3	4.08421	0.62695	14	11	11	9.93685	6.81979	2	2	2	1.89111	1.27484	7	5	5	4.46576	1.54904
LogG	5	3	3	2.80344	0.50290	6	3	2	3.48561	0.35540	15	10	9	9.83907	6.54850	3	2	2	2.14541	1.03347	6	4	4	3.99308	0.91778
S-LogG	8	4	3	3.73314	1.07460	6	2	2	3.60760	0.39443	17	12	10	12.21185	8.42779	2	2	2	1.77520	1.31194	8	4	4	4.15074	1.20824
FIG	5	4	4	3.28601	0.85152	4	3	3	2.89243	0.63442	10	6	5	5.85640	2.87405	4	2	2	2.47205	0.76894	5	3	3	3.46824	0.50984
S-FIG	8	4	4	3.99750	1.15920	4	3	2	2.68945	0.69937	13	10	7	8.81743	5.90158	2	2	2	1.78629	1.30839	5	3	3	3.61609	0.55715
FIAPA	5	3	3	2.95206	0.45534	6	5	3	4.19920	1.38374	11	4	4	6.23473	2.35511	5	2	2	2.60016	0.56795	5	5	5	4.02436	1.48780
S-FIAPA	8	5	4	4.53279	1.73049	7	5	3	4.11288	1.43612	12	10	6	8.87506	5.76002	2	2	2	1.90456	1.27054	6	5	5	4.29684	1.41499
FIEG	6	3	2	3.27283	0.28731	6	3	2	3.14505	0.24642	11	4	4	6.48384	2.43483	5	2	2	2.69626	0.53720	6	5	4	4.11384	1.35643
S-FIEG	8	5	4	4.35749	1.67440	8	4	3	3.96907	1.15010	12	11	9	9.48375	6.35480	2	2	2	1.89806	1.27262	6	5	5	4.41217	1.45189
FIMEG	6	2	2	3.36750	0.31760	6	3	2	3.29113	0.29316	11	4	4	6.85332	2.55306	5	2	2	2.84124	0.49080	6	5	5	4.17213	1.37508
S-FIMEG	7	5	5	4.58226	1.58632	7	4	3	3.99288	0.99772	12	11	9	9.53687	6.37180	2	2	2	1.89044	1.27506	6	5	5	4.46192	1.46781
FIMLogG	5	3	2	3.19429	0.42217	6	3	3	3.77084	0.44667	11	5	5	7.41422	3.13255	5	2	2	2.90165	0.47147	6	5	5	4.14289	1.36572
S-FIMLogG	8	5	4	4.48329	1.71465	7	3	3	4.08743	0.62798	14	11	11	9.93089	6.81788	2	2	2	1.89157	1.27470	7	5	5	4.45949	1.54704
FILogG	6	3	3	2.59528	0.40951	6	2	2	3.24432	0.27818	12	7	6	7.79189	4.21340	4	2	2	2.68123	0.70201	6	4	4	3.99975	0.91992
S-FILogG	9	4	3	4.01098	1.32351	6	2	2	3.39214	0.32548	17	12	10	11.68638	8.25964	2	2	2	1.81336	1.29973	7	4	4	4.11935	1.03819
Distr.	s-t					t					s-t					AL					s-t				

Table 10 - Model selection criteria based on forecasts of the 97.5% and 99% VaR as well as the 97.5% ES, evaluated using **Weighted Absolute Deviation (WAD)** . The considered volatility specifications include GARCH (G), APARCH (APA), EGARCH (EG), MEGARCH (MEG), Log-GARCH (LogG), and MLog-GARCH (MLogG). Models prefixed by S denote semiparametric specifications, while FI indicates the inclusion of a fractional integration term and hence long-memory dynamics. For each return series, the preferred model under the respective criterion is highlighted in bold, and the corresponding innovation distribution is reported in the Distr. row. The prefix "s" before a distribution indicates the skewed variant.

Furthermore, the results indicate that the WAD model selection criterion aligns more closely with the regulatory traffic light assessment than conventional loss functions. When N_1, N_2, or T_{ES} lie outside the regulatorily appropriate green zones, the WAD becomes large by construction, which in turn makes this model the least favorable for selection. This is illustrated by the Microsoft series: the loss-function-optimal model fails to fall within the green zone of the traffic light test. Only the second-best model based on loss functions - yet the best model among those satisfying regulatory constraints - yields a model consistent with regulatory requirements. The WAD never selects a model that is not assigned to the best possible zone. This underscores the usefulness of WAD as a selection device that prioritizes tail-risk adequacy and regulatory compliance rather than purely focusing on the severity of the breaches.

An overview of the WAD values for all models based on the distribution of best selected model across all series is given in Tables 7 to 10. Members of the EGF are selected in most cases, and (adjusted) Log-GARCH variants - particularly their long memory extensions - are chosen for more than half of the series, consistent with the findings based on loss functions. While models most closely related to the classical Log-GARCH appear less frequently, the broader EGF class outperforms its benchmark competitors in three out of four cases. This provides strong evidence that these specifications offer credible and robust alternatives to conventional short memory models. Notably, 17 of the 20 selected models incorporate long memory dynamics, highlighting the empirical relevance of accounting for persistence when using GARCH models in applied risk management.

Figures 1 and 2 present the one-step-ahead rolling forecasts of the 97.5% VaR and ES for the ABEA series across all GARCH specifications and their semiparametric counterparts. The 99% VaR is omitted for readability, as it typically lies very close to the 97.5% ES. The dark and light blue lines depict the VaR and ES forecasts, while exceedances are marked by red circles and crosses. The figure shows that the semiparametric models occasionally underestimate tail risk at the 97.5% level, as evidenced by a higher-than-expected frequency of breaches, placing them in the yellow zone of the regulatory traffic light framework. The displayed risk measures clearly show that semiparametric and parametric models behave differently when confronted with the same volatility environment. The semiparametric specifications tend to track short-term fluctuations in the scale-adjusted data more closely, which occasionally results in insufficient tail protection. By contrast, the parametric specifications generally yield fewer violations, more closely satisfy the supervisory thresholds, and produce smoother and more conservative risk forecasts, resulting in fewer exceedances. This is particularly noticeable for the Log-GARCH variants, whose conditional variance dynamics yield wider safety buffers and fewer violations.

The difference illustrates how semiparametric detrending sharpens short-run volatility signals but may also reduce the margin of conservativeness embedded in fully parametric long memory models. Additionally, models that incorporate fractional integration display a more persistent and pronounced volatility response, especially during sustained high-volatility episodes. Their semiparametric counterparts typically exhibit attenuated persistence because the slowly changing scale component is removed prior to estimating the conditional dynamics.

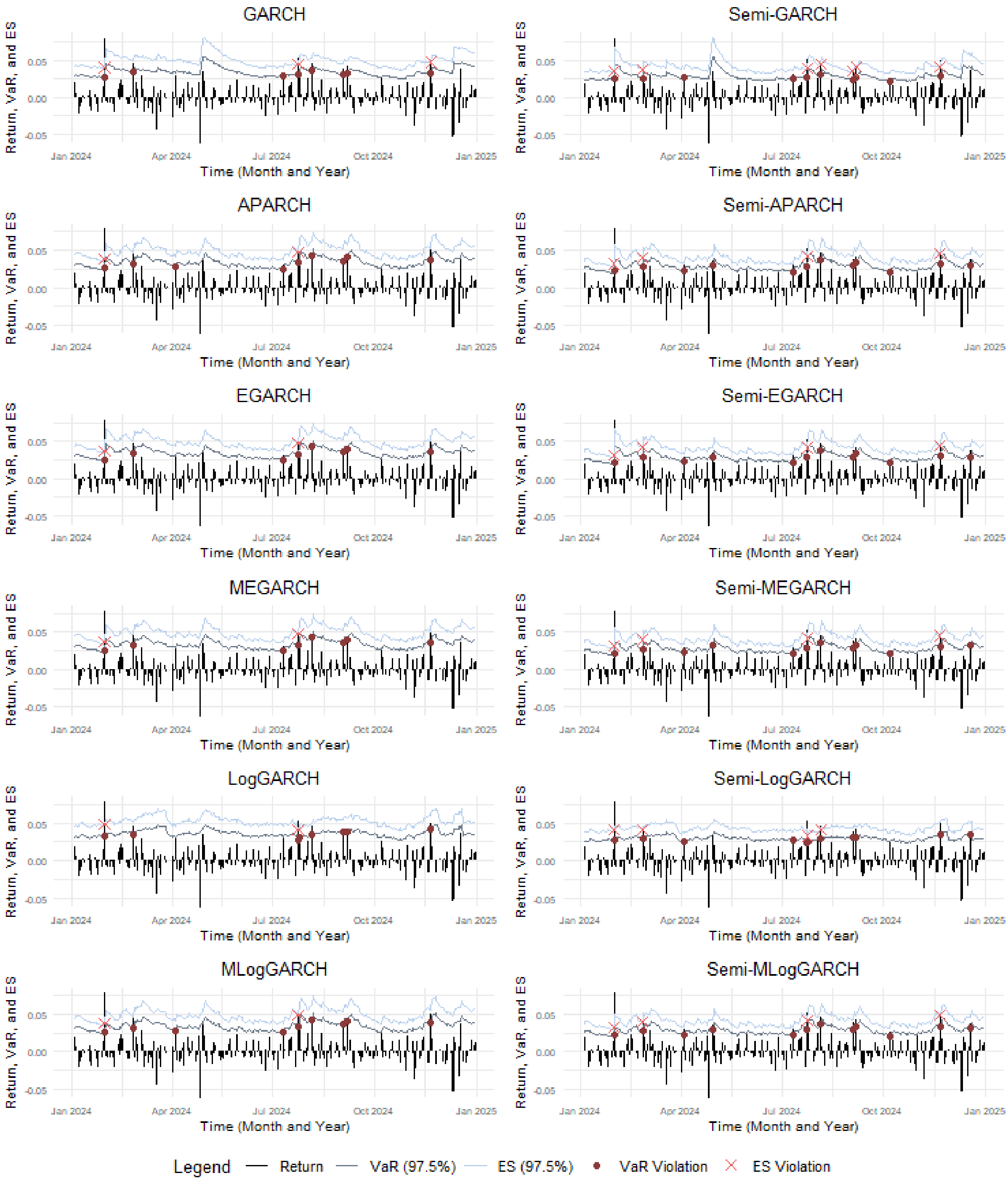

Figure 1 - One-step-ahead rolling forecasts of the 97.5% VaR and ES for the Alphabet Class A series across the full set of short memory GARCH specifications.

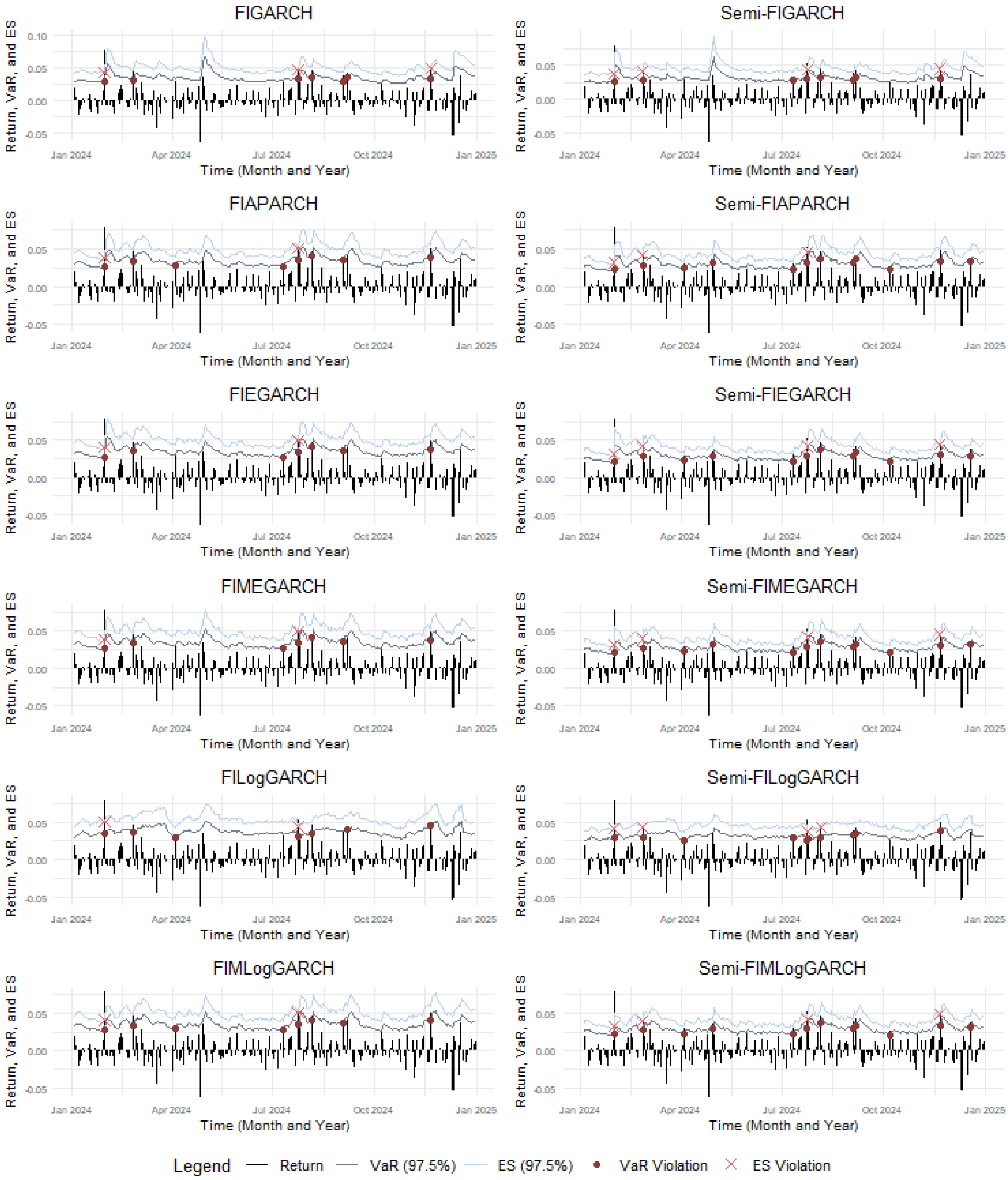

Figure 2 - One-step-ahead rolling forecasts of the 97.5% VaR and ES for the Alphabet Class A series across the full set of long memory GARCH specifications.

This underscores the role of long-range dependence in shaping tail-risk forecasts: ignoring deterministic scale variation may artificially inflate the estimated memory parameter, whereas isolating it yields cleaner - but sometimes less conservative - risk forecasts.

6.2 Fitted model parameters

Due to space constraints and given the growing regulatory prominence of ES, Tables 11 to 14 report only the parameter estimates in accordance with the 97.5% ES corrected loss function and the WAD criterion; full results are available upon request. The tables present the estimated coefficients $\hat{\omega}$, $\hat{\phi}$, $\hat{\psi}$, the fractional integration parameter $\hat{d}$, and the degrees of freedom $\hat{\nu}$ of the selected models. Tables 11 and 13 compare all specifications, while Tables 12 and 14 focus exclusively on the semiparametric variants. For EGARCH models incorporating leverage, the sign and magnitude effects are denoted by $\hat{\kappa}$ and $\hat{\gamma}$; in the APARCH family, $\hat{\kappa}$ and $\hat{\delta}$ denote the leverage and transformation parameters, respectively. Across all models, all estimated coefficients fall within admissible parameter regions and are generally highly significant. Intercepts in the EGARCH class are negative by construction, whereas standard GARCH and APARCH models exhibit strictly positive intercepts due to parameter restrictions. Short memory models tend to exhibit larger $\hat{\phi}$ estimates, reflecting their reliance on autoregressive persistence in the absence of fractional differencing. Estimated asymmetry parameters vary across stocks but consistently display the expected signs and strong statistical significance.

The estimated degrees of freedom underscore the inadequacy of Gaussian innovations, confirming heavy-tailed return behavior and validating the distributional insights from Table 1. Likewise, the estimated skewness parameters, though often only moderately deviating from unity, indicate economically meaningful asymmetry and coincide with the early descriptive analysis. Even small amounts of skewness can materially alter VaR and ES forecasts by reallocating probability mass toward one tail. The frequent selection of skewed distributions reveals their practical relevance: symmetric distributions generally fail to capture the asymmetric downside risk characteristic of equity returns, particularly in periods of market stress. Assets with pronounced left skewness exhibit larger downside risk, while right-skewed series reflect return patterns dominated by upward jumps or recovery episodes.

For the fitted parametric FIGARCH models, the estimate $\hat{\omega} \approx 0$ implies that the conditions of the adjusted FIGARCH specification in Karanasos et al. (2004) are effectively met. Moreover, their Condition (6) is satisfied, so assuming finite fourth moments for return series is empirically justified. When long memory models are selected, parametric estimates of $\hat{d}$ are usually moderate but may sometimes fall outside the stationary range - most prominently for NVDA, META, and AVGO in recently introduced EGF-type specifications (LLY, ORCL, COST for WAD selected models). These violations likely reflect spurious long memory arising from slowly evolving unconditional variance. The NVDA series will be subject to further investigation. The semiparametric counterparts yield substantially smaller $\hat{d}$ values, always within the stationarity region. This systematic reduction arises because the semiparametric approach removes the deterministic scale component prior to parametric estimation, thereby preventing structural shifts from being misinterpreted as stochastic long memory. Consequently, semiparametric specifications provide more stable volatility estimates and more reliable risk forecasts, particularly from a regulatory standpoint. Usually, there exists a semiparametric counterpart which fulfills all regulatory requirements. They mitigate spurious long memory effects, improve distributional calibration, and yield innovation structures that better reflect empirical equity-return dynamics. Tables 12 and 14 show that the best semiparametric GARCH specifications are predominantly long memory variants, underscoring the empirical relevance of persistent volatility dynamics in return series. Notably, the semiparametric

Members of the EGARCH Family

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\psi}$	$\hat{\kappa}$	$\hat{\gamma}$	$\hat{d}$	$\hat{\nu}$ or $\hat{P}$	Dist.	Skew.
AAPL	EG	-8.2594 (0.0932)	0.9576 (0.0086)		-0.0980 (0.0152)	0.1878 (0.0219)		5.1507 (0.4406)	t	
NVDA	FIMLogG	-7.3562 (0.3912)	0.4317 (0.1591)		-0.1021 (0.0315)	0.4043 (0.0751)	0.6359 (0.0651)	4.4812 (0.3298)	s-t	0.9948 (0.0230)
MSFT	S-LogG	-0.0473 (0.0650)	0.9384 (0.0124)	-0.8848 (0.0171)				4.3710 (0.3283)	s-t	0.9948 (0.0199)
AMZN	MLogG	-7.9402 (0.1146)	0.9741 (0.0064)		-0.1292 (0.0243)	0.3295 (0.0432)		4.6058 (0.3430)	s-t	0.9784 (0.0230)
ABEA	FILogG	-8.1075 (0.2000)	0.2697 (0.0424)	-0.5287 (0.0547)			0.3011 (0.0325)	3.9758 (0.2743)	s-t	0.9508 (0.0189)
ABEC	FILogG	-8.4120 (0.1759)	0.2117 (0.0753)	-0.4054 (0.0887)			0.2444 (0.0305)	1.2592 (0.0439)	GED	
META	FIMLogG	-6.8783 (0.4258)	0.5198 (0.1458)		-0.1364 (0.0339)	0.2884 (0.0646)	0.6503 (0.0663)	4.0921 (0.3136)	s-t	0.9657 (0.0245)
TSLA	FILogG	-6.4488 (0.1582)	0.3587 (0.0480)	-0.5823 (0.0553)			0.2572 (0.0307)	3.6766 (0.2475)	s-t	0.9606 (0.0182)
AVGO	FIMEG	-7.7241 (0.3269)	0.4174 (0.1999)		-0.1721 (0.0385)	0.1910 (0.0406)	0.5923 (0.0858)	4.8882 (0.3871)	s-t	0.9575 (0.0221)
BRK-B	LogG	-9.0468 (0.0786)	0.9569 (0.0095)	-0.8860 (0.0151)				1.0000 (NA)	s-AL	1.0107 (0.0196)
WMT	S-EG	-0.2556 (0.0541)	0.9171 (0.0233)		-0.0366 (0.0148)	0.1824 (0.0298)		1.0000 (NA)	AL	
JPM	LogG	-8.3569 (0.1006)	0.9738 (0.0058)	-0.9150 (0.0113)				4.6516 (0.3645)	t	
V	FILogG	-8.6466 (0.1850)	0.2607 (0.0616)	-0.4774 (0.0762)			0.2837 (0.0303)	4.6700 (0.3615)	s-t	0.9203 (0.0199)
MA	LogG	-8.3997 (0.0921)	0.9656 (0.0067)	-0.8934 (0.0117)				1.0000 (NA)	AL	
XOM	S-LogG	-0.0756 (0.0627)	0.9581 (0.0103)	-0.9092 (0.0145)				6.6407 (0.7158)	t	
ORCL	FILogG	-8.3940 (0.1627)	0.2207 (0.0783)	-0.3940 (0.0914)			0.2470 (0.0275)	3.8150 (0.2471)	s-t	0.9246 (0.0189)
UNH	FILogG	-8.3543 (0.1662)	0.3999 (0.0473)	-0.6108 (0.0575)			0.2723 (0.0376)	1.1569 (0.0339)	s-GED	0.9013 (0.0088)
COST	FILogG	-8.4716 (0.0376)	0.4055 (0.0146)	-0.5635 (0.0140)			0.2064 (0.0117)	1.0816 (0.0323)	s-GED	0.8359 (0.0067)
PG	LogG	-9.2215 (0.0774)	0.9588 (0.0083)	-0.9030 (0.0130)				4.2977 (0.3141)	s-t	0.9741 (0.0202)

Standard GARCH

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\beta}$	$\hat{\kappa}$	$\hat{\gamma}$	$\hat{d}$	$\hat{\nu}$ or $\hat{P}$	Dist.	Skew.
LLY	FIG	0.0000 (0.0000)	0.0839 (0.2545)	0.2281 (0.2774)			0.2557 (0.0506)	3.9395 (0.2542)	s-t	0.9717 (0.0218)

Table 11 – **Overall best model** selected by the **corrected loss function for the ES** at the 97.5% confidence level. Standard errors are reported in parentheses below the coefficient estimates. A preceding "s" denotes the corresponding skewed version of the innovation distribution.

Members of the EGARCH Family

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\psi}$	$\hat{\kappa}$	$\hat{\gamma}$	$\hat{d}$	$\hat{\nu}$ *or* $\hat{P}$	Dist.	Skew.
MSFT	S-LogG	-0.0473 (0.0650)	0.9384 (0.0124)	-0.8848 (0.0171)				4.3710 (0.3283)	s-t	0.9948 (0.0199)
AMZN	S-MLogG	-0.1412 (0.0773)	0.9532 (0.0112)		-0.1376 (0.0267)	0.3504 (0.0466)		4.6280 (0.3465)	s-t	0.9852 (0.0212)
ABEA	S-FILogG	-0.0031 (0.1212)	0.2528 (0.0632)	-0.4260 (0.0773)			0.2105 (0.0323)	4.0602 (0.2846)	s-t	0.9538 (0.0194)
ABEC	S-LogG	-0.0196 (0.0718)	0.9108 (0.0243)	-0.8578 (0.0296)				4.2207 (0.3741)	s-t	0.9450 (0.0227)
TSLA	S-FIMLogG	0.0360 (0.0781)	0.9519 (0.0240)		-0.0604 (0.0299)	0.2632 (0.0705)	0.0000 (0.1455)	3.9006 (0.2792)	s-t	0.9922 (0.0194)
BRK-B	S-FILogG	-0.0477 (0.1366)	0.1914 (0.0868)	-0.3441 (0.1016)			0.2221 (0.0276)	4.9696 (0.4165)	s-t	1.0179 (0.0217)
WMT	S-EG	-0.2556 (0.0541)	0.9171 (0.0233)		-0.0366 (0.0148)	0.1824 (0.0298)		1.0000 (NA)	AL	
V	S-FILogG	-0.3131 (0.1443)	0.2223 (0.0861)	-0.3909 (0.1022)			0.2319 (0.0294)	4.6930 (0.3599)	s-t	0.9193 (0.0198)
XOM	S-LogG	-0.0756 (0.0627)	0.9581 (0.0103)	-0.9092 (0.0145)				6.6407 (0.7158)	t	
ORCL	S-FILogG	-0.2133 (0.1254)	0.1250 (0.1203)	-0.2522 (0.1361)			0.1959 (0.0274)	3.8014 (0.2445)	s-t	0.9279 (0.0190)
COST	S-FILogG	-0.0626 (0.0729)	0.4499 (0.1715)	-0.5027 (0.1866)			0.1093 (0.0399)	4.1689 (0.2926)	s-t	0.9782 (0.0203)
PG	S-FILogG	-0.0781 (0.1308)	0.2670 (0.0764)	-0.4405 (0.0937)			0.2220 (0.0327)	4.1714 (0.2977)	s-t	0.9541 (0.0196)

Standard GARCH

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\beta}$	$\hat{\kappa}$	$\hat{\gamma}$	$\hat{d}$	$\hat{\nu}$ *or* $\hat{P}$	Dist.	Skew.
AAPL	S-FIG	0.1856 (0.0418)	0.3046 (0.1054)	0.4963 (0.1343)			0.3242 (0.0833)	4.5416 (0.3684)	s-t	0.9664 (0.0206)
NVDA	S-FIG	0.2483 (0.0646)	0.0620 (0.2018)	0.1830 (0.2250)			0.1908 (0.0514)	1.1501 (0.0329)	s-GED	0.9938 (0.0062)
LLY	S-FIG	0.2809 (0.0503)	0.0000 (0.0013)	0.1212 (0.0554)			0.2341 (0.0508)	3.7555 (0.2604)	s-t	0.9732 (0.0191)
JPM	S-FIG	0.1415 (0.0312)	0.1254 (0.1880)	0.3365 (0.2263)			0.3357 (0.0686)	5.2154 (0.4658)	s-t	0.9841 (0.0214)

APARCH

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\beta}$	$\hat{\kappa}$	$\hat{\delta}$	$\hat{d}$	$\hat{\nu}$ *or* $\hat{P}$	Dist.	Skew.
META	S-FIAPA	0.3854 (0.0470)	0.1684 (0.1135)	0.3826 (0.1534)	0.5708 (0.1511)	1.0324 (0.1657)	0.2767 (0.0739)	4.2114 (0.3275)	s-t	0.9740 (0.0222)
AVGO	S-FIAPA	0.3068 (0.0290)	0.2452 (0.0651)	0.5550 (0.1107)	0.5472 (0.1229)	1.0007 (0.1623)	0.4133 (0.0815)	4.9497 (0.4031)	s-t	0.9660 (0.0207)
MA	S-FIAPA	0.2731 (0.0330)	0.0237 (0.0868)	0.2884 (0.1141)	0.6782 (0.1105)	1.1268 (0.1246)	0.3700 (0.0537)	5.0717 (0.4270)	s-t	0.9080 (0.0196)
UNH	S-FIAPA	0.3374 (0.0438)	0.2034 (0.0735)	0.4516 (0.1167)	0.8066 (0.1923)	1.0613 (0.1673)	0.3151 (0.0680)	4.7973 (0.3979)	s-t	0.9781 (0.0210)

Table 12 – **Best semiparametric model** selected by the **corrected loss function for the ES** at the 97.5% confidence level. Standard errors are reported in parentheses below the coefficient estimates. A preceding "s" denotes the corresponding skewed version of the innovation distribution.

estimates of persistence are more stable and internally consistent, indicating that our approach provides a robust framework for capturing genuine long-range dependence while avoiding spurious long memory induced by non-stationary scale variation. This aligns with earlier evidence from Ding et al. (1993) and Lobato and Savin (1998), who documented true long memory behavior in squared returns that cannot be attributed solely to structural breaks or deterministic trends.

Members of the EGARCH Family

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\psi}$	$\hat{\kappa}$	$\hat{\gamma}$	$\hat{d}$	$\hat{\nu}$ or $\hat{P}$	Dist.	Skew.
AAPL	FIEG	-8.2516 (0.1552)	0.8050 (0.1148)		-0.0953 (0.0162)	0.1660 (0.0242)	0.3125 (0.1441)	1.2652 (0.0385)	s-GED	0.9960 (0.0130)
MSFT	LogG	-8.2987 (0.0775)	0.9572 (0.0085)	-0.8998 (0.0134)				4.3564 (0.3297)	t	
AMZN	S-FIEG	-0.1367 (1.6215)	0.9480 (0.0135)		-0.0843 (0.0008)	0.1610 (0.0380)	0.0000 (0.0000)	4.5995 (0.0654)	s-t	0.9855 (0.9837)
ABEA	FILogG	-8.5131 (0.1878)	0.2818 (0.0475)	-0.4955 (0.0589)			0.2681 (0.0313)	1.1038 (0.0310)	s-GED	0.9810 (0.0259)
ABEC	FILogG	-8.0132 (0.2205)	0.2401 (0.0519)	-0.4882 (0.0644)			0.2971 (0.0351)	3.8699 (0.3244)	s-t	0.9270 (0.0214)
META	FILogG	-7.6365 (0.1779)	0.2518 (0.0589)	-0.4550 (0.0682)			0.2546 (0.0264)	1.0417 (0.0309)	s-GED	1.0180 (0.0062)
TSLA	FILogG	-6.4847 (0.1637)	0.3509 (0.0538)	-0.5733 (0.0669)			0.2597 (0.0354)	3.6497 (0.2471)	t	
BRK-B	MLogG	-9.0766 (0.0927)	0.9550 (0.0093)		-0.1490 (0.0257)	0.4869 (0.0559)		1.0000 (NA)	s-AL	1.0048 (0.0216)
LLY	FIMEG	-8.8401 (0.3178)	0.3364 (0.2456)		-0.0630 (0.0309)	0.2041 (0.0514)	0.6009 (0.0878)	1.0783 (0.2568)	s-GED	0.9837 (0.0215)
WMT	S-FILogG	-0.0697 (0.1288)	0.3952 (0.0521)	-0.5541 (0.0644)			0.2233 (0.0381)	1.0000 (NA)	AL	
JPM	FILogG	-8.5905 (0.1776)	0.2338 (0.0743)	-0.4085 (0.0867)			0.2552 (0.0254)	1.2294 (0.0360)	GED	
V	FILogG	-8.8621 (0.1865)	0.2885 (0.0399)	-0.5148 (0.0438)			0.2926 (0.0218)	1.2088 (0.0344)	s-GED	0.9190 (0.0090)
MA	FIMLogG	-8.1550 (0.2388)	0.6426 (0.0954)		-0.2181 (0.0333)	0.4539 (0.0589)	0.4543 (0.0687)	1.0000 (NA)	s-AL	0.9075 (0.0176)
ORCL	FIEG	-8.3882 (0.3218)	0.4362 (0.1570)		-0.0733 (0.0218)	0.2146 (0.0365)	0.5647 (0.0648)	3.7194 (0.2403)	t	
COST	FIMLogG	-9.1230 (0.3098)	0.1482 (0.1685)		-0.1769 (0.0364)	0.5597 (0.0811)	0.5815 (0.0632)	1.0000 (NA)	AL	

Standard GARCH

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\beta}$	$\hat{\kappa}$	$\hat{\gamma}$	$\hat{d}$	$\hat{\nu}$ or $\hat{P}$	Dist.	Skew.
NVDA	FIG	0.0001 (0.0000)	0.2265 (0.0854)	0.5039 (0.1008)			0.3462 (0.0494)	1.0000 (NA)	s-AL	1.0106 (0.0209)
AVGO	S-FIG	0.2325 (1.5487)	0.0003 (0.0001)	0.0971 (0.0273)			0.2263 (0.0262)	4.7008 (0.3698)	s-t	0.9765 (0.0168)
PG	FIG	0.0000 (0.0000)	0.0000 (0.1906)	0.1589 (0.2134)			0.2808 (0.0587)	4.4189 (0.3347)	s-t	0.9564 (0.0223)

APARCH

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\beta}$	$\hat{\kappa}$	$\hat{\delta}$	$\hat{d}$	$\hat{\nu}$ or $\hat{P}$	Dist.	Skew.
XOM	APA	0.0008 (0.0002)	0.0835 (0.0112)	0.9251 (0.0095)	0.3820 (0.0897)	0.6353 (0.2261)		7.1042 (0.8123)	s-t	0.9745 (0.0236)
UNH	FIAPA	0.0019 (0.0003)	0.2295 (0.0431)	0.6168 (0.0916)	0.6706 (0.1334)	1.1866 (0.1584)	0.4761 (0.0931)	4.6764 (0.3737)	s-t	0.9560 (0.0225)

Table 13 - Overall best model selected by the **WAD**. Standard errors are reported in parentheses below the coefficient estimates. A preceding "s" denotes the corresponding skewed version of the innovation distribution.

The relevance of semiparametric GARCH specifications becomes evident when examining cases in which the parametric long memory parameter is markedly inflated and lies outside the stationary region. For example, for the NVDA series (Figure 3), the estimated fractional parameter in Table 11 exceeds 0.5, a value incompatible with covariance stationarity. The standardized returns exhibit clear signs of non-stationary scale variation, consistent with earlier findings that deterministic shifts or structural breaks may masquerade as long memory (Francq & Zakoïan, 2012; Jiang et al., 2021). Panels (b) and (d) reveal that once the smooth scale function $\hat{s}(\tau_{t,n})$ is estimated and removed, the resulting series becomes visibly more homogeneous, substantially improving the suitability of parametric volatility models and their associated forecasts. The estimated scale function in panel (c) effectively captures slow-moving changes in long-run risk. A closer inspection indicates that Nvidia's apparent non-stationarity stems not from genuine fractional persistence but from structural breaks in the unconditional

Members of the EGARCH Family

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\psi}$	$\hat{\kappa}$	$\hat{\gamma}$	$\hat{d}$	$\hat{\nu}$ *or* $\hat{P}$	Dist.	Skew.
MSFT	S-LogG	-0.1563	0.9300	-0.8757				1.1326	GED	
		(0.0568)	(0.0157)	(0.0213)				(0.0332)		
AMZN	S-FIE	-0.1367	0.9480		-0.0843	0.1610	0.0000	4.5995	s-t	0.9855
		(1.6215)	(0.0135)		(0.0008)	(0.0380)	(0.0000)	(0.0654)		(0.0136)
ABEA	S-FILogG	-0.0031	0.2528	-0.4260			0.2105	4.0602	s-t	0.9538
		(0.1212)	(0.0632)	(0.0773)			(0.0323)	(0.2846)		(0.0194)
META	S-FILogG	-0.0756	-0.2045	0.1321			0.0924	4.1284	t	
		(0.0738)	(0.3843)	(0.4106)			(0.0284)	(0.3136)		
TSLA	S-FIEG	0.0080	0.4167		-0.0372	0.1970	0.4926	3.9436	s-t	0.9909
		(0.2012)	(0.2505)		(0.0199)	(0.0471)	(0.1295)	(0.2780)		(0.0193)
BRK-B	S-FILogG	-0.0477	0.1927	-0.3457			0.2225	4.9754	t	
		(0.1371)	(0.0867)	(0.1016)			(0.0277)	(0.4172)		
WMT	S-FILogG	-0.0697	0.3952	-0.5541			0.2233	1.0000	AL	
		(0.1288)	(0.0521)	(0.0644)			(0.0381)	(NA)		
V	S-FILogG	-0.3131	0.2223	-0.3909			0.2319	4.6930	s-t	0.9193
		(0.1443)	(0.0861)	(0.1022)			(0.0294)	(0.3599)		(0.0198)
XOM	S-LogG	-0.1309	0.9514	-0.8988				1.4229	s-GED	0.9799
		(0.0616)	(0.0114)	(0.0156)				(0.0455)		(0.0196)
ORCL	S-LogG	-0.2745	0.9126	-0.8395				1.0609	s-GED	0.9373
		(0.0784)	(0.0206)	(0.0264)				(0.0286)		(0.0049)

Standard GARCH

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\beta}$	$\hat{\kappa}$	$\hat{\gamma}$	$\hat{d}$	$\hat{\nu}$ *or* $\hat{P}$	Dist.	Skew.
AAPL	S-FIG	0.1856	0.3046	0.4963			0.3242	4.5416	s-t	0.9664
		(0.0418)	(0.1054)	(0.1343)			(0.0833)	(0.3684)		(0.0206)
NVDA	S-FIG	0.2483	0.0620	0.1830			0.1908	1.1501	s-GED	0.9938
		(0.0646)	(0.2018)	(0.2250)			(0.0514)	(0.0329)		(0.0062)
AVGO	S-FIG	0.2325	0.0003	0.0971			0.2263	4.7008	s-t	0.9765
		(0.0406)	(0.0064)	(0.0096)			(0.0228)	(0.0367)		(0.0129)
JPM	S-FIG	0.1443	0.0000	0.1485			0.2695	1.2918	s-GED	0.9906
		(0.0327)	(0.3391)	(0.3710)			(0.0532)	(0.0400)		(0.0134)
PG	S-FIG	0.1851	0.0000	0.1865			0.3149	4.2454	s-t	0.9532
		(0.0379)	(0.1901)	(0.2296)			(0.0758)	(0.3294)		(0.0200)

APARCH

Stock	Model	$\hat{\omega}$	$\hat{\phi}$	$\hat{\beta}$	$\hat{\kappa}$	$\hat{\delta}$	$\hat{d}$	$\hat{\nu}$ *or* $\hat{P}$	Dist.	Skew.
ABEC	S-FIAPA	0.3453	0.1137	0.3273	0.5562	1.1935	0.2756	1.1413	s-GED	0.9609
		(0.0283)	(0.0577)	(0.0658)	(0.0828)	(0.0969)	(0.0272)	(0.0380)		(0.0100)
LLY	S-FIAPA	0.3245	0.3120	0.5175	0.2543	1.2273	0.3187	1.0918	s-GED	0.9778
		(0.0397)	(0.1012)	(0.1559)	(0.1823)	(0.2374)	(0.0901)	(0.0312)		(0.0026)
MA	S-FIAPA	0.2731	0.0237	0.2884	0.6782	1.1268	0.3700	5.0717	s-t	0.9080
		(0.0330)	(0.0868)	(0.1141)	(0.1105)	(0.1246)	(0.0537)	(0.4270)		(0.0196)
UNH	S-FIAPA	0.3374	0.2034	0.4516	0.8066	1.0613	0.3151	4.7973	s-t	0.9781
		(0.0438)	(0.0735)	(0.1167)	(0.1923)	(0.1673)	(0.0680)	(0.3979)		(0.0210)
COST	S-FIAPA	0.3559	0.3880	0.5256	0.6529	1.0253	0.2521	1.0000	AL	
		(0.0460)	(0.1153)	(0.1418)	(0.1490)	(0.1446)	(0.0657)	(NA)		

Table 14 – Best semiparametric model selected by the **WAD**. Standard errors are reported in parentheses below the coefficient estimates. A preceding "s" denotes the corresponding skewed version of the innovation distribution.

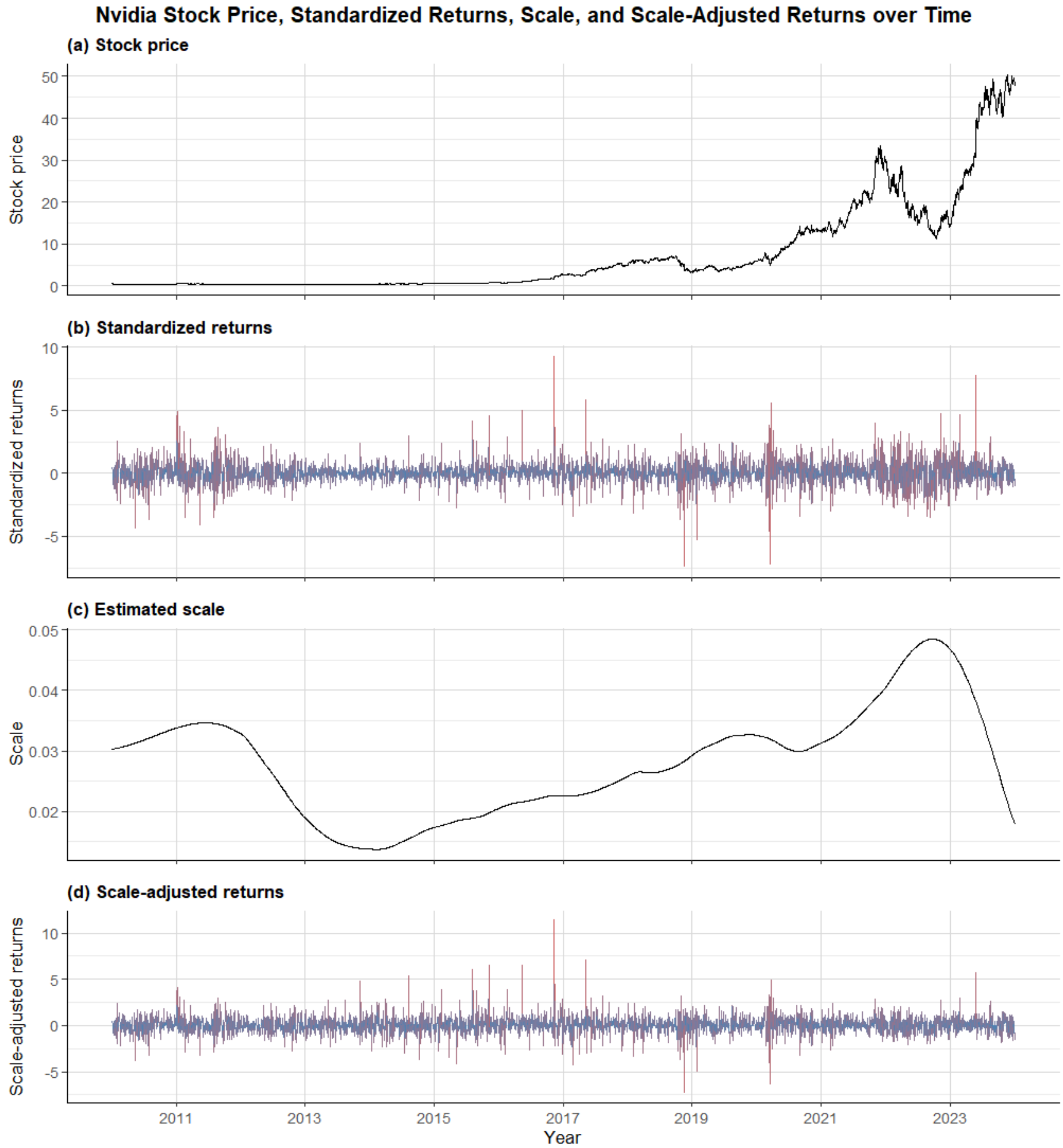

Figure 3 – Slowly evolving unconditional variance and non-stationarity. Nvidia stock price, standardized returns, scale function and scale-adjusted returns over time.

variance - particularly during 2010–2012, associated with post-crisis market conditions and uncertainty in the GPU sector, and again from 2020 onward amid COVID-19-induced regime changes and the firm's strategic pivot toward AI, which is not caused by a stochastic non-stationarity factor. These breaks are evident in Panels (a) and (b), where the unconditional variance shifts sharply across regimes. As highlighted in Sibbertsen (2004), deterministic trends and structural breaks can induce spurious long memory, and vice versa. Consistent with this observation, Table 12 shows that the estimated fractional parameter returns to the stationary range once the semiparametric scale adjustment is applied, demonstrating that the original non-stationary estimate reflected spurious long memory and an artifact of neglected structural variation. Failing to account for such deterministic scale shifts can lead to misinterpreting non-stationarity as spurious persistence, producing volatility forecasts that are unstable, biased, and potentially misleading for risk measurement. Since the consistency of QMLE hinges on strict stationarity (Francq & Zakoïan, 2012), these findings underscore the need to test for and remove non-stationary components prior to GARCH estimation and the construction of confidence or forecast intervals.

Overall, the findings demonstrate that Semi-GARCH models represent a robust and effective alternative to conventional parametric approaches in market risk modeling and that recognizing and modeling scale functions and long memory behavior yields substantial economic and financial benefits. Especially in cases where parametric models exhibit weak performance (due to overestimation of spurious persistence), adopting a semiparametric framework can yield more accurate and stable results in forecasting volatility and associated risk measures. The presence of long or dual long memory in many economic and financial time series is unsurprising, as market adjustments often occur gradually - allowing the effects of initial shocks to persist over extended periods. Such delayed adjustments, driven by factors like information frictions, transaction costs, and slow structural responses, can induce persistence in volatility and deviations from linearity in the data-generating process. From a forecasting perspective, modeling these features enhances both point and interval predictions. In financial markets, accounting for long memory, dual long memory, or smooth trend effects enhances the accuracy of volatility and risk forecasts, reduces inefficiencies, and improves portfolio and risk management strategies. Hence, integrating these persistent dynamics within parametric or semiparametric frameworks represents a valuable step toward a more refined and realistic analysis of market processes.

7 Conclusion

In the financial sector, quantitative risk management plays a crucial role in identifying, measuring, and mitigating risks that could threaten the solvency of financial institutions. In particular, the adequate capitalization of banks enhances their resilience to crises and contributes to the stability of the overall financial system, where institutions are closely interconnected. The financial crises of recent decades have underscored the necessity of dedicated regulatory frameworks for banks and have emphasized the importance of risk awareness and transparency in a sector that exerts a profound influence on global economic development.

This study contributes to the modeling of market risk by focusing on two key risk measures widely used in practice: the VaR and the ES. Their application follows the Basel regulatory standards, ensuring a strong alignment with real-world risk management practices. Beyond standard GARCH specifications, the paper applies different classes of semiparametric short and long memory models, which are capable of simultaneously capturing conditional heteroskedasticity and slowly changing persistent dependencies in financial time series. The underlying deterministic scale function approximating the unconditional variance is estimated nonparametrically via a SEMIFARIMA model using local polynomial smoothing, with the optimal bandwidth determined through an iterative plug-in algorithm (Letmathe et al., 2023; Letmathe et al., 2022). This approach removes non-stationary trend components from the return series, after which the parametric GARCH-type models are fitted to the detrended data.

We utilize the proposed (Semi-LM-)GARCH specifications to obtain out-of-sample forecasts of the VaR and ES measures. The resulting VaR and ES forecasts are evaluated over the test period using the traffic light backtesting framework, and overall model performance is assessed through different loss functions and the WAD criterion recently introduced by Feng et al. (2020), which jointly accounts for all backtesting results. We complement the regulatory loss function with the recently proposed corrected loss function of Feng and Peitz (2025), which augments the evaluation criterion by incorporating the opportunity costs associated with holding excessive regulatory capital. Moreover, we further strengthen the practical relevance of GARCH-based risk forecasting by incorporating asymmetric, skewed innovation distributions, including the recently introduced Average Laplace (AL) distribution of Feng et al. (forthcoming), which provides additional flexibility for capturing tail asymmetries in financial returns especially through maintaining desirable statistical properties of specific members of the EGF. While parametric specifications are often selected as optimal under regulatory constraints and model-selection criteria, they can occasionally generate spurious long memory behavior - an issue effectively mitigated by the semiparametric extensions. Overall, the EGF family, particularly the (adjusted) Log-GARCH variants, exhibits notably strong performance within the regulatory framework for VaR and ES forecasting. Generally (specifically under the WAD criterion) long memory specifications were identified as the best-performing models. Statistical properties of the parametric processes were enhanced by introducing a scale function for long-term unconditional heteroskedasticity. Although the semiparametric extensions were selected only in rare cases, the results confirm that long memory dynamics and the estimation of scale functions remain economically and statistically relevant. Importantly, the semiparametric extension helps offset over- or underestimation biases inherent in purely parametric models, yielding more stable and robust volatility forecasts. Therefore, the empirical study demonstrates that semiparametric long memory models offer a valuable complement to traditional parametric frameworks. When standard models prove inefficient or overly restrictive, semiparametric approaches provide a flexible and stable alternative for improved inference and prediction. Additionally, the results provide strong evidence for the superior performance of skewed innovation distributions, underscoring their practical relevance in market risk assessment. In practice, it is advisable to consider and regularly re-evaluate multiple model specifications, as model performance may vary with changing market conditions. Against this backdrop, the proposed models and empirical findings may assist both banking supervisors and financial institutions in refining ES estimation and improving corresponding backtesting procedures.

Future research should build upon this work by extending the univariate GARCH framework to multivariate settings, enabling a comprehensive assessment of portfolio-wide risk dynamics and interdependencies among assets. Moreover, promising directions for further development include hybrid modeling approaches that combine traditional econometric frameworks with deep learning architectures, thereby uniting the interpretability and theoretical rigor of classical models with the flexibility and adaptive learning capabilities of modern machine learning techniques. In particular, forecasting the deterministic scale function in semiparametric specifications using recurrent neural networks (RNNs), rather than fixing it at its last in-sample estimate, seems to be an interesting topic for future research. This would allow the unconditional component to evolve over the test period and could substantially improve VaR and ES forecast accuracy.

References

Abad, P., Muela, S. B., & Martín, C. L. (2015). The role of the loss function in value-at-risk comparisons. *The Journal of Risk Model Validation, 9*(1): 1-19.

Acerbi, C., & Tasche, D. (2002). Expected shortfall: a natural coherent alternative to value at risk. *Economic Notes*, *31*(2), 379-388.

Adrian, T., & Shin, H. S. (2014). Procyclical leverage and value-at-risk. *The Review of Financial Studies*, *27*(2), 373-403.

Aloui, C., & Ben Hamida, H. (2015). Estimation and Performance Assessment of Value-at-Risk and Expected Shortfall Based on Long-Memory GARCH-Class Models. *Finance a Uver: Czech Journal of Economics & Finance*, *65*(1).

Altman, N. S. (1990). Kernel smoothing of data with correlated errors. *Journal of the American Statistical Association*, *85*(411), 749-759.

Amado, C., & Terasvirta, T. (2008). Modelling conditional and unconditional heteroskedasticity with smoothly time-varying structure. *CREATES Research Paper*.

Amado, C., & Teräsvirta, T. (2013). Modelling volatility by variance decomposition. *Journal of Econometrics*, *175*(2), 142-153.

Amado, C., & Teräsvirta, T. (2014). Modelling changes in the unconditional variance of long stock return series. *Journal of Empirical Finance*, *25*, 15-35.

Amado, C., & Teräsvirta, T. (2017). Specification and testing of multiplicative time-varying GARCH models with applications. *Econometric Reviews*, *36*(4), 421-446.

Andersen, T. G., & Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, *4*(2-3), 115-158.

Andersen, T. G., Bollerslev, T., & Lange, S. (1999). Forecasting financial market volatility: Sample frequency vis-à-vis forecast horizon. *Journal of Empirical Finance*, *6*(5), 457-477.

Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*(3), 203-228.

Ayensu, O. K., Feng, Y., & Schulz, D. (2025). Semiparametric extensions of EGARCH-type models and their implementation in R, *Center for International Economics, Paderborn University*.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 171-178.

Baillie, R. T., & Morana, C. (2009). Modelling long memory and structural breaks in conditional variances: An adaptive FIGARCH approach. *Journal of Economic Dynamics and Control*, *33*(8), 1577-1592.

Baillie, R. T., & Myers, R. J. (1991). Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics*, *6*(2), 109-124.

Bardet, J. M., & Wintenberger, O. (2009). Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *Annals of Statistics, 37*(5), 2730-2759.

Basel Committee on Banking Supervision (1996). Supervisory Framework for the Use of Backtesting in Conjunction With The Internal Models Approach to Market Risk Capital Requirements. *Bank for International Settlements Basel*.

Basel Committee on Banking Supervision (2013). Fundamental review of the trading book: A revised market risk framework. *Bank for International Settlements Basel*.

Basel Committee on Banking Supervision (2016). Minimum capital requirements for market risk. *Bank for International Settlements Basel*.

Basel Committee on Banking Supervision (2017). Finalizing post-crisis reforms. *Bank for International Settlements Basel*.

Basel Committee on Banking Supervision (2019). Minimum capital requirements for market risk (revised). *Bank for International Settlements Basel*.

Beine, M., Laurent, S., & Lecourt, C. (2002). Accounting for conditional leptokurtosis and closing days effects in FIGARCH models of daily exchange rates. *Applied Financial Economics*, *12*(8), 589-600.

Beran, J., & Feng, Y. (2002a). Iterative plug-in algorithms for SEMIFAR models—definition, convergence, and asymptotic properties. *Journal of Computational and Graphical Statistics*, *11*(3), 690-713.

Beran, J., & Feng, Y. (2002b). Local polynomial fitting with long-memory, short-memory and antipersistent errors. *Annals of the Institute of Statistical Mathematics*, *54*(2), 291-311.

Beran, J., & Feng, Y. (2002c). SEMIFAR models—A semiparametric approach to modelling trends, long-range dependence and nonstationarity. *Computational Statistics & Data Analysis*, *40*(2), 393-419.

Beran, J., & Ocker, D. (1999). SEMIFAR forecasts, with applications to foreign exchange rates. *Journal of Statistical Planning and Inference*, *80*(1-2), 137-153.

Beran, J., & Ocker, D. (2001). Volatility of stock-market indexes—an analysis based on SEMIFAR models. *Journal of Business & Economic Statistics*, *19*(1), 103-116.

Beran, J., Feng, Y., Ghosh, S., and Kulik, R. (2013). Long-Memory Processes. Springer, Berlin.

Beran, J., Feng, Y., & Ghosh, S. (2015). Modelling long-range dependence and trends in duration series: an approach based on EFARIMA and ESEMIFAR models. *Statistical Papers*, *56*(2), 431-451.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177-181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307-327.

Bollerslev, T., & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, *73*(1), 151-184.

Bougerol, P., & Picard, N. (1992). Strict stationarity of generalized autoregressive processes. *The Annals of Probability*, *20*(4), 1714-1730.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *26*(2), 211-243.

Breidt, F. J., Crato, N., & De Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, *83*(1-2), 325-348.

Brownlees, C. T., & Gallo, G. M. (2010). Comparison of volatility measures: a risk management perspective. *Journal of Financial Econometrics*, *8*(1), 29-56.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 841-862.

Conrad, J., Gultekin, M. N., & Kaul, G. (1991). Asymmetric predictability of conditional variances. *The Review of Financial Studies*, *4*(4), 597-622.

Costanzino, N., & Curran, M. (2015). Backtesting general spectral risk measures with application to expected shortfall. *The Journal of Risk Model Validation, 9*(1), 21.

Costanzino, N., & Curran, M. (2018). A simple traffic light approach to backtesting expected shortfall. *Risks*, *6*(1), 2.

Cotter, J. (2005). Uncovering long memory in high frequency UK futures. *The European Journal of Finance*, *11*(4), 325-337.

Degiannakis, S. (2004). Volatility forecasting: evidence from a fractional integrated asymmetric power ARCH skewed-t model. *Applied Financial Economics*, *14*(18), 1333-1342.

Demiralay, S., & Ulusoy, V. (2014). Value-at-risk predictions of precious metals with long memory volatility models. *MPRA Paper No. 53229*.

Diebold, F. X. (1986). Modeling the persistence of conditional variances: A comment. *Econometric Reviews*, *5*(1), 51-56.

Ding, Z., & Granger, C. W. (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of Econometrics*, *73*(1), 185-215.

Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1*(1), 83-106.

Douc, R., Roueff, F., & Soulier, P. (2008). On the existence of some ARCH(∞) processes. *Stochastic Processes and their Applications*, *118*(5), 755-761.

Duan, N. (1983). Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association*, *78*(383), 605-610.

El Ghourabi, M., Francq, C., & Telmoudi, F. (2016). Consistent estimation of the value at risk when the error distribution of the volatility model is misspecified. *Journal of Time Series Analysis*, *37*(1), 46-76.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), 1-50.

Engle, R. F., & Rangel, J. G. (2008). The spline-GARCH model for low-frequency volatility and its global macroeconomic causes. *The Review of Financial Studies*, *21*(3), 1187-1222.

Engle, R. F., Ghysels, E., & Sohn, B. (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics*, *95*(3), 776-797.

Escribano, A., & Sucarrat, G. (2018). Equation-by-equation estimation of multivariate periodic electricity price volatility. *Energy Economics*, *74*, 287-298.

Fan, J., & Gijbels, I. (1996). Local Polynomial Modelling and Its Applications: Monographs on Statistics and Applied Probability 66 (1st ed.). *Routledge*.

Feng, Y. (2004). Simultaneously modeling conditional heteroskedasticity and scale change. *Econometric Theory*, *20*(3), 563-596.

Feng, Y., & Karl Härdle, W. (2025). A data-driven P-spline smoother and the P-Spline-GARCH models. *Journal of Nonparametric Statistics*, 1-22.

Feng, Y., & Peitz, C. (2025). Semiparametric GARCH models for value-at-risk and expected shortfall: an object-driven procedure. *Journal of Risk, 27*(5).

Feng, Y., Beran, J., & Letmathe, S. (2021). Smoothing long memory time series. *R package version 1.0.1*.

Feng, Y., Beran, J., Letmathe, S., & Ghosh, S. (2020). Fractionally integrated Log-GARCH with application to value at risk and expected shortfall. *Center for International Economics, Paderborn University.*

Feng, Y., Beran, J., Letmathe, S., & Schulz, D. (2024). esemifar: Smoothing Long Memory Time Series. *R package version 2.0.1.*

Feng, Y., Gries, T., & Letmathe, S. (2023). FIEGARCH, modulus asymmetric FILog-GARCH and trend-stationary dual long memory time series. *Center for International Economics, Paderborn University.*

Feng, Y., Letmathe, S., & Schulz, D. (2023). smoots: Nonparametric Estimation of the Trend and Its Derivatives in TS. *R package version 1.1.4*.

Feng, Y., Peitz, C., & Siddiqui, S. (*forthcoming*). A few useful members of the EGARCH family with short- or long-memory in volatility. *Center for International Economics, Paderborn University.*

Fernández, C., & Steel, M. F. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, *93*(441), 359-371.

Ferreira, J. T. S., & Steel, M. F. J. (2006). A constructive representation of univariate skewed distributions. *Journal of the American Statistical Association*, *101*(474), 823-829.

Föllmer, H., & Schied, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, *6*(4), 429-447.

Francq, C., & Sucarrat, G. (2018). An exponential chi-squared QMLE for log-GARCH models via the ARMA representation. *Journal of Financial Econometrics*, *16*(1), 129-154.

Francq, C., & Zakoïan, J. M. (2012). Strict stationarity testing and estimation of explosive and stationary generalized autoregressive conditional heteroscedasticity models. *Econometrica*, *80*(2), 821-861.

Francq, C., & Zakoïan, J. M. (2015). Risk-parameter estimation in volatility models. *Journal of Econometrics*, *184*(1), 158-173.

Francq, C., Wintenberger, O., & Zakoïan, J. M. (2013). GARCH models without positivity constraints: Exponential or log GARCH?. *Journal of Econometrics*, *177*(1), 34-46.

Francq, C., Wintenberger, O., & Zakoïan, J. M. (2018). Goodness-of-fit tests for Log-GARCH and EGARCH models. *Test*, *27*(1), 27-51.

Gallant, A. R. (1981). On the bias in flexible functional forms and an essentially unbiased form: the Fourier flexible form. *Journal of Econometrics*, *15*(2), 211-245.

Gallant, A. R. (1984). The Fourier flexible form. *American Journal of Agricultural Economics*, *66*(2), 204-208.

Gasser, T., & Müller, H. G. (1979). Kernel estimation of regression functions. Smoothing techniques for curve estimation. *Springer*, 23-68.

Gasser, T., Kneip, A., & Köhler, W. (1991). A flexible and fast method for automatic smoothing. *Journal of the American Statistical Association*, *86*(415), 643-652.

Gasser, T., Müller, H. G., & Mammitzsch, V. (1985). Kernels for nonparametric curve estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 238-252.

Geweke, J. (1986). Modeling the persistence of conditional variances: A comment. *Econometric Reviews*, *5*(1), 51-56.

Ghalanos, A., & Theussl, S. (2011). Rsolnp: General Non-linear Optimization Using Augmented Lagrange Multiplier Method. *R package version 1.11.*

Giot, P., & Laurent, S. (2003). Value-at-risk for long and short trading positions. *Journal of Applied Econometrics*, *18*(6), 641-663.

Giraitis, L., Surgailis, D., & Škarnulis, A. (2018). Stationary integrated ARCH(∞) and AR(∞) processes with finite variance. *Econometric Theory*, *34*(6), 1159-1179.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779-1801.

Goldfeld, S. M., & Quandt, R. E. (1981). Econometric modelling with non-normal disturbances. *Journal of Econometrics*, *17*(2), 141-155.

Grané, A., & Veiga, H. (2008). Accurate minimum capital risk requirements: A comparison of several approaches. *Journal of Banking & Finance*, *32*(11), 2482-2492.

Granger, C. W. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, *14*(2), 227-238.

Granger, C. W., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis*, *1*(1), 15-29.

Hall, P., & Hart, J. D. (1990). Nonparametric regression with long-range dependence. *Stochastic Processes and Their Applications*, *36*(2), 339-351.

Härdle, W. K., & Mungo, J. (2008). Value-at-risk and expected shortfall when there is long range dependence. *Discussion Papers from Humboldt University*.

Hastie, T., & Loader, C. (1993). Local regression: Automatic kernel carpentry. *Statistical Science*, 120-129.

He, C., Teräsvirta, T., & Malmsten, H. (2002). Moment structure of a family of first-order exponential GARCH models. *Econometric Theory*, *18*(4), 868-885.

Hosking, J. R. (1981a). Fractional differencing. *Biometrika 68*, 165-176.

Hosking, J. R. (1981b). Lagrange-multiplier tests of multivariate time-series models. *Journal of the Royal Statistical Society: Series B (Methodological)*, *43*(2), 219-230.

Iqbal, R., Sorwar, G., Baker, R., & Choudhry, T. (2020). Multiday expected shortfall under generalized t distributions: evidence from global stock market. *Review of Quantitative Finance and Accounting*, *55*(3), 803-825.

J. P. Morgan (1996). Risk Metrics technology document. Morgan Guaranty Trust Company of New York, *New York*, 35-65.

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, *6*(3), 255-259.

Jiang, F., Li, D., & Zhu, K. (2021). Adaptive inference for a semiparametric generalized autoregressive conditional heteroskedasticity model. *Journal of Econometrics*, *224*(2), 306-329.

John, J. A., & Draper, N. R. (1980). An alternative family of transformations. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *29*(2), 190-197.

Jones, M. C. (2004). Families of distributions arising from distributions of order statistics. *Test*, *13*(1), 1-43.

Karanasos, M., Psaradakis, Z., & Sola, M. (2004). On the Autocorrelation Properties of Long-Memory GARCH Processes. *Journal of Time Series Analysis*, *25*(2), 265-282.

Kazakevičius, V., & Leipus, R. (2002). On stationarity in the ARCH (∞) model. *Econometric Theory*, *18*(1), 1-16.

Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives, 3*(2).

Lamoureux, C. G., & Lastrapes, W. D. (1990). Persistence in variance, structural change, and the GARCH model. *Journal of Business & Economic Statistics*, *8*(2), 225-234.

Letmathe, S., Beran, J., & Feng, Y. (2023). An extended exponential SEMIFAR model with application in R. *Communications in Statistics-Theory and Methods*, *53*(22), 7914-7926.

Letmathe, S., Feng, Y., & Uhde, A. (2022). Semiparametric GARCH models with long memory applied to value-at-risk and expected shortfall. *Journal of Risk, 25*(2).

Ling, S., & McAleer, M. (2002). Necessary and sufficient moment conditions for the GARCH (r, s) and asymmetric power GARCH (r, s) models. *Econometric Theory*, *18*(3), 722-729.

Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297-303.

Lobato, I. N., & Savin, N. E. (1998). Real and spurious long-memory properties of stock-market data. *Journal of Business & Economic Statistics*, *16*(3), 261-268.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, *36*(4), 394.

Mazur, B., & Pipień, M. (2012). On the empirical importance of periodicity in the volatility of financial returns-time varying GARCH as a second order APC (2) process. *Central European Journal of Economic Modelling and Econometrics*, 95-116.

McNeil, A. J., Frey, R., & Embrechts, P. (2015). Quantitative risk management: Concepts, Techniques and Tools: Revised edition. *Princeton University Press*.

Mikosch, T., & Stărică, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *Review of Economics and Statistics*, *86*(1), 378-390.

Milhøj, A. (1987). A conditional variance model for daily deviations of an exchange rate. *Journal of Business & Economic Statistics*, *5*(1), 99-103.

Missiakoulis, S. (1983). Sargan densities: which one? *Journal of Econometrics*, *23*(2), 223-233.

Nelson, D. B. (1990). Stationarity and persistence in the GARCH (1, 1) model. *Econometric Theory*, *6*(3), 318-334.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.

Peitz, C. (2016). Die parametrische und semiparametrische Analyse von Finanzzeitreihen: neue Methoden, Modelle und Anwendungsmöglichkeiten. *Springer-Verlag*.

Ray, B. K., & Tsay, R. S. (1997). Bandwidth selection for kernel regression with long-range dependent errors. *Biometrika*, *84*(4), 791-802.

Royer, J. (2023). Conditional asymmetry in Power ARCH (∞) models. *Journal of Econometrics*, *234*(1), 178-204.

Ruppert, D., & Wand, M. P. (1994). Multivariate locally weighted least squares regression. *The Annals of Statistics*, 1346-1370.

Ruppert, D., Sheather, S. J., & Wand, M. P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association*, *90*(432), 1257-1270.

Sarma, M., Thomas, S., & Shah, A. (2003). Selection of Value-at-Risk models. *Journal of Forecasting*, *22*(4), 337-358.

Schulz, D., Feng, Y., Peitz, C. & Ayensu, O. K. (2025). fEGarch: SM/LM EGARCH & GARCH, VaR/ES Backtesting & Dual LM Extensions. *R package version 1.0.1*.

Sibbertsen, P. (2004). Long memory in volatilities of German stock returns. *Empirical Economics*, *29*(3), 477-488.

Sucarrat, G. (2019). The log-GARCH model via ARMA representations. *Financial Mathematics, Volatility and Covariance Modelling, 2*, 336.

Sucarrat, G., & Escribano, A. (2018). Estimation of log-GARCH models in the presence of zero returns. *The European Journal of Finance*, *24*(10), 809-827.

Sucarrat, G., Grønneberg, S., & Escribano, A. (2016). Estimation and inference in univariate and multivariate log-GARCH-X models when the conditional density is unknown. *Computational Statistics & Data Analysis*, *100*, 582-594.

Surgailis, D., & Viano, M. C. (2002). Long memory properties and covariance structure of the EGARCH model. *ESAIM: Probability and Statistics*, *6*, 311-329.

Tang, T. L., & Shieh, S. J. (2006). Long memory in stock index futures markets: A value-at-risk approach. *Physica A: Statistical Mechanics and its Applications*, *366*, 437-448.

Tse, Y. K. (1987). A note on Sargan densities. *Journal of Econometrics*, *34*(3), 349-354.

Tse, Y. K. (1998). The conditional heteroscedasticity of the yen–dollar exchange rate. *Journal of Applied Econometrics*, *13*(1), 49-55.

Van Bellegem, S., & Von Sachs, R. (2004). Forecasting economic time series with unconditional time-varying variance. *International Journal of Forecasting*, *20*(4), 611-627.

Zaffaroni, P. (2009). Whittle estimation of EGARCH and other exponential volatility models. *Journal of Econometrics*, *151*(2), 190-200.

Zakoïan, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control*, *18*(5), 931-955.

Zhang, X., Feng, Y., & Peitz, C. (2017). A general class of SemiGARCH models based on the Box-Cox transformation, *Center for International Economics, Paderborn University*.

Appendix

A.1 Regulatory framework: Treatment of market risk under Basel III

The BCBS serves as the primary international body responsible for developing and promoting uniform global standards for banking regulation, with the overarching goal of safeguarding financial system stability. Although its resolutions are not legally binding, they function as key reference points for national supervisory authorities. Within the European Union, the BCBS recommendations are transposed into binding law through the Capital Requirements Regulation (CRR) and the Capital Requirements Directive (CRD). Over time, the BCBS framework has undergone several major revisions, culminating in three successive accords: Basel I (1988), Basel II (2004), and Basel III (2010). The latter introduced substantially stricter capital requirements in response to the 2007–2008 global financial crisis, aiming to enhance the resilience of the banking sector at both the micro- and macroprudential levels. In 2012, the BCBS initiated the *Fundamental Review of the Trading Book (FRTB)*, revising the treatment of market risk. Following two consultation papers, the final document "Minimum Capital Requirements for Market Risk" was published in 2016 (BCBS, 2016). The corresponding amendments were largely implemented in the EU through CRR II.

The FRTB framework delineates the trading book as well as the banking book and specifies two approaches for measuring market risk: the Standardized Approach (SA) and the Internal Model Approach (IMA). Both were considerably strengthened. The standardized approach was made more risk-sensitive, while the internal model approach now faces stricter validation and approval requirements. Under the new framework, banks must perform daily measurements of market risk in their trading books. Traditionally, most banks have relied on internal VaR models, calibrated at a 99% confidence level, to quantify market risk. However, the FRTB replaces VaR with the ES as the new risk metric.[11] The ES must be computed daily at a 97.5% confidence level, with expanded data and calibration requirements. To ensure model reliability, regular backtesting is mandated and the use of internal models remains subject to prior supervisory approval. Under current rules, the 99% VaR is evaluated over the previous trading year (typically 250 observations) using a traffic light system, where supervisory actions depend on the number of VaR exceedances. As of now, the BCBS has not yet released an official backtesting framework for ES.

[11] The ES has replaced the VaR because, unlike VaR, it is a coherent risk measure that accounts for the magnitude of losses beyond the chosen confidence level, providing a more comprehensive assessment of tail risk. For a short introduction to the axioms of risk measures, see Appendix A.2.

A.2 Requirements for risk measures

In the academic literature, several sets of axioms have been proposed to characterize the desirable properties of risk measures. Artzner et al. (1999) introduced the foundational framework for coherent risk measures, while Föllmer and Schied (2002) extended this concept to convex risk measures. Consistent with the formulation adopted in this paper, McNeil et al. (2015) define a risk measure as a real-valued functional on a linear space of random variables $\mathcal{M}$, which is assumed to include all constant elements. This definition admits two common interpretations for the elements of $\mathcal{M}$. First, the elements may represent future net asset values of a position or portfolio, denoted by V, with V_0 indicating the corresponding current net asset value. Alternatively, the elements may describe future losses, defined as $L = V_0 - V$. Under these conventions, a risk measure can be interpreted either as the amount of capital that must be added to a position to make it acceptable - expressed as $\tilde{\varrho}(V)$ - or as the total capital required to cover potential losses, denoted by $\varrho(L)$. The relationship between the two notations is described by

$$\varrho(L) = V_0 + \tilde{\varrho}(V) \tag{A2.1}$$

as the total capital is made up of available capital plus additional capital. A risk measure is described in context as $\varrho: L \rightarrow \varrho(L)$. The axioms can be defined on this basis.

Monotony*: For L_1, L_2 $\in \mathcal{M}$ such that $L_1 < L_2$, then $\varrho(L_1) \leq \varrho(L_2)$ holds.*

That is easy to understand from an economic point of view, as higher losses lead to more risk capital. On the other hand, positions with $\varrho(L) < 0$ do not require equity backing.

Translation Invariance*: For $L \in \mathcal{M}$ and for each $l \in \mathbb{R}$, $\varrho(L + l) = \varrho(L) + l$.*

If a high-risk position is expanded or reduced by the deterministic value l, the equity backing is adjusted by this value. Alternatively, $\tilde{\varrho}(V + k) = \tilde{\varrho}(V) - k$ for $k \in \mathbb{R}$. This results in $\tilde{\varrho}\left(V + \tilde{\varrho}(V)\right) = 0$. A future position with the value $V + \tilde{\varrho}(V)$ therefore requires no further equity backing.

Subadditivity*: For L_1, L_2 $\in \mathcal{M}$, $\varrho(L_1 + L_2) \leq \varrho(L_1) + \varrho(L_2)$ applies.*

Subadditivity takes into account that risk can be reduced through diversification. For example, an institution whose capital adequacy is determined on the basis of a non-subadditive risk measure would have an incentive to legally split into different subsidiaries in order to reduce the regulatory capital requirements.

Positive Homogeneity*: For $L \in \mathcal{M}$ and for every $\lambda > 0$, $\varrho(\lambda L) = \lambda \varrho(L)$ applies.*

With the assumption that subadditivity is given, positive homogeneity can also be justified, since subadditivity for $n \in \mathbb{N}$ implies:

$$\varrho(nL) = \varrho(L + \cdots + L) \leq n\varrho(L)$$

By means of these properties, one can establish a definition for a coherent risk measure.

***Coherent Risk Measure**: A risk measure ϱ whose domain includes the convex cone $\mathcal{M}$ is said to be coherent (on $\mathcal{M}$) if the previous axioms apply.*

***Convexity**: For all $L_1, L_2 \in \mathcal{M}$ and all $\lambda \in [0,1]$, $\varrho(\lambda L_1 + (1-\lambda)L_2) \leq \lambda\varrho(L_1) + (1-\lambda)\varrho(L_2)$ applies.*

In economic terms, the convexity axiom is again based on the fact that diversification reduces risks. Convexity can be used to establish a definition for a convex risk measure.

***Convex Risk Measures**: A risk measure ϱ on $\mathcal{M}$ is called convex (on $\mathcal{M}$) if the axioms of monotony, translation invariance and convexity are fulfilled.*

Every coherent risk measure is also convex, whereas the reverse is not true (McNeil et al., 2015).

A.3 Value at Risk and Expected Shortfall

The VaR metric was introduced in 1996 by the major US bank JPMorgan Chase in the fourth version of "RiskMetrics" (J.P. Morgan, 1996). It arose from the basic idea of determining an upper limit for the loss from a financial position or a portfolio. The practical relevance of this risk measure is demonstrated, among other things, by the fact that VaR was included in the second Basel Accord published by the BCBS and in the Solvency II Directive.[12] The VaR can be defined as the maximum loss L of a risky position for a certain time horizon t that will not be exceeded with a given confidence level of $1 - \alpha$ or, alternatively, the minimum extreme loss that occurs with a probability of α in a given period. The distribution function of loss L is defined as $F_L(l) = P(L < l)$. According to Equation (2.5) in McNeil et al. (2015), the negative returns, $-r_t$, can be interpreted as first-order approximations of relative losses. This approximation is employed throughout this paper. Statistically, the VaR represents the α -percent quantile of a loss distribution and can be defined as follows:

$$VaR_\alpha = \inf\{l \in \mathbb{R}: P(L > l) \leq 1 - \alpha\} = \inf\{l \in \mathbb{R}: F_L(l) \geq \alpha\}$$

with $0 < \alpha < 1$. The usual confidence level for VaR in banking regulatory practice is 99% with a time horizon of one or ten days. However, the VaR does not contain any information about the residual risk when the VaR is exceeded and a loss of an unknown amount is incurred.

There are several established approaches for estimating the VaR. The most common include the variance-covariance method, the Monte Carlo simulation, and the historical simulation. The first two are parametric approaches, typically based on time-invariant estimates of volatility, whereas the historical simulation relies solely on empirical return data without assuming a specific distribution. As discussed earlier, financial return series frequently exhibit heteroskedasticity, making GARCH-type models particularly suitable for capturing the time-varying nature of volatility. In this context, the VaR evolves dynamically over time in accordance with the conditional variance forecasts. To compute the VaR, assumptions regarding the loss distribution function F_L are required. Given a fitted model the VaR and ES can be computed from the conditional loss distribution implied by the assumed innovation distribution. Let $\hat{\mu}_t$ denote the estimated or forecasted conditional mean and $\hat{\sigma}_t$ the corresponding conditional standard deviation at time t. Define $\mathrm{VaR}_{\varepsilon,\alpha}$ and $\mathrm{ES}_{\varepsilon,\alpha}$ as the time-invariant VaR and ES, respectively, of a sequence of i.i.d. standardized innovations $\{\varepsilon_t\}$ with mean zero and unit variance. Assuming the return process satisfies $r_t = \mu_t + \sigma_t \varepsilon_t$ where μ_t and σ_t represent the

[12] VaR was originally developed for the quantification of market price risks. However, the concept was also adopted and adapted for other risks, e.g. credit risk (credit value at risk) and liquidity risk (liquidity-at-risk).

true conditional mean and conditional standard deviation, the conditional VaR of r_t can be expressed as follows:

$$\widehat{\mathrm{VaR}}_{t,\alpha} = \hat{\mu}_t + \hat{\sigma}_t \, \mathrm{VaR}_{\varepsilon,\alpha},$$

Here, $\mathrm{VaR}_{\varepsilon,\alpha}$ depend solely on the assumed conditional distribution of the standardized innovations ε_t. Define

$$\mathrm{VaR}_{\varepsilon,\alpha} = F_\varepsilon^{-1}(\alpha)$$

where $F_\varepsilon(\cdot)$ denotes the CDF of the innovation process ε_t. Hence, $F_\varepsilon^{-1}(\alpha)$ yields the quantile of the innovation distribution at level α. For many innovation distributions, these quantities must be numerically estimated, particularly when the distribution depends on unknown parameters that are themselves estimated from the data.[13] Introducing skewness via the Fernández and Steel (1998) transformation does not change the VaR estimation mechanics, but it changes the innovation quantile that is plugged into the VaR as now ε_t follows the FS-skewed version of the baseline symmetric distribution, as described above.

Although the VaR provides information about the probability of losses exceeding a given threshold, it offers no insight into the magnitude of losses once this threshold is breached. This limitation became particularly evident in the aftermath of the 2008 global financial crisis, when it was revealed that many financial institutions had underestimated tail risks and that extreme losses had been systematically downplayed. Moreover, VaR is designed to capture losses under normal market conditions, and thus fails to adequately reflect the effects of rare or extreme events, such as financial crises, which exert a profound influence on market dynamics (Adrian & Shin, 2014). In addition, Artzner et al. (1999) demonstrated that VaR does not qualify as a coherent risk measure, since it violates the principle of subadditivity.[14] In addition to the criticism that VaR is not coherent, Artzner et al. (1999) presented an alternative, the ES as a coherent and improved risk measure. The latter is calculated as the arithmetic average of the extreme losses that exceed the VaR. For a loss L with $E(|L|) < \infty$ and a density function F_L, the ES with a confidence level $\alpha \in (0, 1)$ is defined as:

[13] For instance, under the standard normal distribution, these quantities are obtained as $\mathrm{VaR}_{\varepsilon,\alpha} = \Phi^{-1}(\alpha)$ where $\Phi^{-1}(\cdot)$ and $\phi(\cdot)$ denote the quantile function and probability density function of the standard normal distribution, respectively. When the innovations follow a Student-t distribution with ν degrees of freedom, the expression becomes $\mathrm{VaR}_{\varepsilon,\alpha} = t_\nu^{-1}(\alpha)$ where $t_\nu^{-1}(\cdot)$ and $f_{t_\nu}(\cdot)$ represent the quantile function and probability density function of the standardized Student-t distribution, respectively. Similarly, the quantile function of the standardized GED with shape parameter ν or of the $AL(P)$ can be used to calculate the VaR under respective distributions.

[14] Consider two portfolios with loss distributions $F_{L_1}$and F_{L_2}, and let the aggregated portfolio loss be $L = L_1 + L_2$ with corresponding distribution F_L. Then the VaR of the combined portfolio is not necessarily equal to the sum of the individual VaRs; that is, $q_\alpha(F_L) = q_\alpha(F_{L_1}) + q_\alpha(F_{L_2})$ does not necessarily apply. As a result, VaR fails to account for diversification benefits, a fundamental property expected of a sound and coherent risk measure.

$$ES_\alpha = (1-\alpha)^{-1} \int_\alpha^1 q_x(F_L)\, dx$$

With $q_u(F_L)$ as the quantile function of F_L. The proximity to VaR is shown by:

$$ES_{\varepsilon,\alpha} = (1-\alpha)^{-1} \int_\alpha^1 VaR_{\varepsilon,x}\, dx$$

In terms of the conditional standard deviation estimated by a GARCH process, the estimated and forecasted ES of r_t can be simply estimated by

$$\widehat{\mathrm{ES}}_{t,\alpha} = \hat{\mu}_t + \hat{\sigma}_t\, \mathrm{ES}_{\varepsilon,\alpha}$$

under varying estimations of $\mathrm{ES}_{\varepsilon,\alpha}$ conditional on the assumed distribution. Both VaR and ES estimation can be extended to skewed distributions as defined above.

A.4 Traffic Light Test Tables

The tables below report the Basel Committee traffic light test thresholds for backtesting the 97.5% and 99% VaR, as well as the ES-based traffic light procedure proposed by Costanzino and Curran (2018). The tables summarize the regulatory classification into green, yellow, and red zones based on exceedance counts or test statistics and their associated cumulative probabilities, providing the benchmark used throughout the empirical analysis.

Backtesting zone	Violations	Cumulative probability
Green	$N \leq 10$	$< 95\%$
Yellow	$9 \leq N \leq 16$	$< 99\%$
Red	$17 \leq N$	$\geq 99.99\%$

Table A4.1 – Critical values for the Basel traffic light test for 97.5% VaR (BCBS, 2016).

Backtesting zone	Violations	Cumulative probability
Green	$N \leq 4$	$< 95\%$
Yellow	$5 \leq N \leq 9$	$< 99\%$
Red	$10 \leq N$	$\geq 99.99\%$

Table A4.2 – Critical values for the Basel traffic light test for 99% VaR (BCBS, 2016).

Backtesting zone	Violations	Cumulative probability
Green	0	0.18%
	1.3929	10%
	2.1131	25%
	3.0276	50%
	4.0520	75%
	5.0622	90%
	5.7049	95%
Yellow	6.9844	99%
	8.5258	99.9%
	9.8833	99.99%
Red	> 9.8833	$\geq 99.99\%$

Table A4.3 – Critical values for a traffic light-test approach for backtesting the 97.5% ES based on the severity of breaches, following Costanzino and Curran (2018).