İSTANBUL UNIVERSITY PRESS

Istanbul Business Research,
IBR 2025, 54 (1): 56–86

https://doi.org/10.26650/ibr.2025.54.1577152

Submitted: 31.10.2024
Revision Requested: 23.01.2025
Last Revision Received: 01.03.2025
Accepted: 07.03.2025

Istanbul Business Research

Research Article

Open Access

GAS or GARCH: A comparison of density and VaR forecasts in Turkish FX and stock markets

Ali Ulvi Özgül [1]

[1] Ankara Bilim University, Faculty of Humanities and Social Sciences, Department of Business Administration, Ankara, Türkiye

Abstract This paper compares the renowned GARCH model with a novel one, the Generalized Autoregressive Score (GAS) model in terms of forecasting performance. Considering the gap in the literature, this study focuses on the Turkish stock and FX markets. The analysis covers 25 years (1999-2023), of which the last 12 constitute the out-of-sample period. The selected indexes largely represent the finance (XBANK) and industry (XUSIN) sectors and the entire (XUTUM) economy, while the fourth (XU100) is the market benchmark. Likewise, FX rates are the leading factors that dominate Turkish foreign trade. Rolling density forecasts from the standard versions of the models are compared via Diebold-Mariano (DM) test with the two popular scoring rules. The GARCH model generally outperforms GAS when the conditional distribution is the Normal or its skewed version. We find some evidence for the reverse with Student-t and skewed version, but this lacks statistical support, except for the definite superiority of GAS in USD returns coupled with skewed Student-t. A deeper analysis attributed GAS's underperformance to its treatment of shocks that are more likely to occur in developing markets. We also report similar findings with DM tests using two loss functions for VaR forecasts, whereas the results of the backtesting procedures are inconsistent across risk levels.

Keywords GARCH model · GAS model · Density forecast · Density score function · Forecast evaluation · Value-at-risk

Citation: Özgül, A. U. (2025). GAS or GARCH: A comparison of density and VaR forecasts in Turkish FX and stock markets. *Istanbul Business Research*, 54(1), 56-86. https://doi.org/10.26650/ibr.2025.54.1577152

Corresponding author: Ali Ulvi Özgül ulvi.ozgul@ankarabilim.edu.tr

1. Introduction

Time-varying volatility has been at the heart of finance in both academia and practice since the 1980s. The seminal model developed by Engle (1982) opened a new scientific venue in which a large family of models has grown over decades. The workhorse model in the family is the generalized autoregressive conditional heteroscedasticity (GARCH) model (Bollerslev, 1986). Through the augmented versions in the family, GARCH models offer several tools for modeling financial time series and risk management.

Recently, as a parent to a new class of models, the generalized autoregressive score (GAS) model has been introduced to capture the time-varying parameters of a distribution that is assumed to govern a time series. Like GARCH, GAS is also an observation-driven model that portrays dynamic behavior through its autoregressive characteristic. What makes GAS very special is its ability to accommodate the complete density structure through score function (Creal, Koopman, & Lucas, 2013). In this regard, GAS model has the capacity to provide density forecasts that are statistically richer and more convenient than point forecasts. Accordingly, the GAS family has attracted the attention of many researchers who have contributed (e.g., Blasques, Hoogerkamp, Koopman, & van de Werve, 2021; Blazsek, Escribano, & Licht, 2021; Catania & Nonejad, 2020; A. Harvey, Hurn, Palumbo, & Thiele, 2024) to the development of the model.

This paper presents an extensive comparison of both models within the forecasting framework as they apply to Turkish stock and FX markets and contributes to the literature in two aspects. To the best of our knowledge, this is the first paper that compares both models in these two prominent markets that foster the Turkish economy. Second, while it confirms the superiority of the GAS model under Student's *t* law to some extent, it pinpoints the probable reason behind its shortcoming under Normal law manifested via scoring rules. We suggest that the reported shortcoming is related to an emerging market with unusual fragility in the form of sudden negative jumps compared with developed ones. Accordingly, different treatments regarding the functional constraints of the packages we employ might be the underlying cause when the distributional assumption is the Normal law or its skewed version.

The paper continues with a review of relevant literature in the following section. The third section summarizes the methodological frameworks of both models. We introduce data with a preliminary analysis and report in-sample fittings for the leading assets in selected markets in the fourth section. The forecasting and backlisting procedures are explained, and the formal density and VaR forecast comparisons are reported next. Some notable points captured during the analysis are summarized and demonstrated in the sixth section while the following one concludes the paper.

2. Literature review

Numerous studies exist on GARCH models with various asset classes and economic variables, as the concept refers to a huge family. We restrict the review to outstanding research that principally compares GARCH and GAS and to that which implements GAS in Turkish markets in line with the aim of the paper.

Bernardi and Catania (2016) compared the Value-at-Risk (VaR) forecasting performances of eight different specifications in the GARCH family and GAS under two distributional assumptions, Student-*t* and the Normal, additionally 2 specifications within dynamic quantile (QL or CAViaR as is widely known) models proposed by Engle and Manganelli (2004). The dataset comprises four major worldwide stock market indexes: Asia/Pacific 600 (SXP1E), North America 600 (SXA1E), Europe 600 (SXXP), and Global 1800 (SXW1E), spanning a 23-year

period. GAS models were eliminated in the whole out-of-sample period for SXW1E and SXXP by the model confidence set (MCS) procedure of Hansen, Lunde and Nason (2011). GAS-$\mathcal{T}$ were also eliminated for SXA1E, while the others were ranked remarkably lower than GARCH-family models. The results for the turbulent and tranquil subperiods were largely the same, which indicates the underperformance of the GAS models.

In a similar study on VaR forecasting performances of daily energy commodities by Laporta, Merlo, & Petrella (2018), the MCS procedure resulted in lower ranks for GAS models than the several alternative GARCH-type models, CAViaR, and the dynamic quantile regression model proposed by the authors.

Xu and Lien (2022) used two loss functions in conjunction with the Diebold–Mariano test to compare the volatility forecasting abilities of the GARCH, EGARCH, and GAS models on oil and gas assets. While GAS outperformed the other two models for crude oil, none of the models turned out to be the clear champion for natural gas assets. The authors emphasize that compared with the two GARCH models, GAS has a remarkable advantage in short-term forecasting of crude oil assets.

Owusu Junior, Tiwari, Tweneboah, and Asafo-Adjei (2022) used the MCS approach to evaluate 1% and 5% VaR forecasts using 32 GARCH and 6 GAS model specifications for precious metals. Rather than evaluating the model performances, they focused on the number of models included in the superior set of models across the four metals. Examining MCS rankings based on Fissler and Ziegel's (2016) loss function reveals that GAS models perform worse than GARCH; GAS (1,1) coupled with Student-*t* being the best among the six GAS models.

Concerning cryptocurrencies, the GAS model has been found to be a suitable alternative for Ethereum in estimating risk measures (Trucíos & Taylor, 2023). The GAS model was compared against several GARCH models, like NAGARCH, FIGARCH, MSGARCH, the bootstrap GARCH proposed by the authors, and some different versions of the CAViaR models. The superiority of forecast combinations was also investigated. The results show that no model outperforms others for Bitcoin, whereas GAS is found to be the single model that surpasses its competitors for Ethereum.

Research on GAS models can be reviewed on the website which is maintained to promote this class of models ("GAS Papers," n. d.). Regarding model applications to Turkish markets, we highlight three studies. The first study (Bekar, 2019) compared VaR forecasts from the Peaks-Over-Threshold (POT) model of the Extreme Value Theory (EVT) and GAS model on Turkish, European, and American banks. Daily share prices of 5 select banks from October 2010 to June 2019 comprise the dataset. GAS models conditional on Student's *t* distribution and its skewed version performed better than the model based on the Normal law for Turkish banks. The author also highlighted that the POT model is more appropriate than the GAS model due to its inherent superiority in capturing frequent shocks caused by uncertainties in emerging markets.

The second study we review concerns the main indicator of the Turkish stock market, namely the BIST 100 index. Kahyaoğlu Bozkuş (2019) implemented GAS model with the index's return series covering the period 04.01.2010 - 03.07.2018 under three distributional assumptions: Normal, Student-*t*, and skewed Student-*t*. Then, the estimated parameters were used in backlisting VaR. The findings led the author to conclude that the GAS model is not effective and consistent in risk calculations owing to the tail effects, which are more prominent in developing financial markets.

Finally, Bekar (2023) compared models from the GARCH family with Beta-t-EGARCH and its variants, which can be viewed as an unrestricted version of GAS (Sucarrat, 2013) in terms of their ability to capture the time series features of USD/TL exchange rate returns covering 06.01.2005-30.09.2021. The results indicated that

the Two-Component Beta-Skew-t-EGARCH Model with Leverage was the best owing to its robust structure against jumps, which is common in developing markets.

3. Methodology

Since this paper compares GARCH and GAS forecasts, a brief methodological background is provided for both models in the following subsections. We consider the standard GARCH model, namely the GARCH (1,1) model, as it is the simplest among the family of such models and generally sufficient to capture volatility clustering in financial time series (Brooks, 2019). Both orders of the GAS model are kept the same as the standard GARCH so that GAS (1,1) is on par with its counterpart throughout the analysis.

3.1. The generalized autoregressive conditional heteroskedasticity (GARCH) model

GARCH is one of the most popular models in econometrics and is the generalized form of the Autoregressive Conditional Heteroskedasticity (ARCH) model proposed by Engle (1982). Given a real-valued discrete-time stochastic process ε_t, the standard GARCH process can be described as follows (Bollerslev, 1986):

$$\begin{gathered} \varepsilon_t \mid \psi_{t-1} \overset{iid}{\sim} N(0, h_t) \\ h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \alpha_0 > 0,\ \alpha_1 \geq 0,\ \beta_1 \geq 0. \end{gathered} \tag{1}$$

where ψ_t is the filtration or set of information through t. In practice, the series ε_t is usually either demeaned returns or error terms from an ARMA model for returns The one-step ahead forecast of conditional variance, h_{t+1} is analytically available with ψ_t (through ε_t) and the parameters of Equation (1) are found by maximum likelihood estimation (MLE). However, predictions for two-steps and beyond suffer from uncertainties in future errors (Pascual, Romo, & Ruiz, 2006). Therefore, we proceed with forecasting one-step ahead volatility throughout the analysis.

3.2. The generalized autoregressive score (GAS) model

Compared to GARCH, GAS is a newly developed model that allows time-varying parameters through the scaled score of the likelihood function and encompasses some other well-known models like GARCH (Creal et al., 2013). The dependent variable y_t in a GAS model is assumed to be generated by the following observation density:

$$y_t \sim p(y_t \mid f_t, \mathcal{F}_t; \theta) \tag{2}$$

where f_t is the time-varying parameter vector with a history $F^t = \{f_0, f_1, \cdots, f_t\}$ and θ is the vector of static parameters. The filtration $\mathcal{F}_t$ includes three elements: $Y^{t-1} = \{y_1, y_2, \cdots, y_{t-1}\}$, namely the past history of the dependent variable up to $(t-1)$, F^{t-1} and $X^t = \{x_1, x_2, \cdots, x_t\}$ where x_i is the vector of exogenous variables at $t = i$, i.e., $\mathcal{F}_t = \{Y^{t-1}, F^{t-1}, X^t\}$. Given the information set $\{f_t, \mathcal{F}_t\}$ at time t, the time-varying parameter vector f_t is updated by the following autoregressive process:

$$f_{t+1} = \omega + A s_t + B f_t \tag{3}$$

where ω is a vector of constants, A and B are diagonal coefficient matrices of proper dimensions. The vector s_t that plays a crucial role in updating f_{t+1} is proportional to the score of the density in Equation (2):

$$s_t = S_t \nabla_t, \qquad \nabla_t = \frac{\partial \ln p(y_t \mid f_t, \mathcal{F}_t; \theta)}{\partial f_t}, \qquad S_t = S(t, f_t, \mathcal{F}_t; \theta). \tag{4}$$

S_t is a scaling matrix that adds flexibility and variations to the model because it controls how the score is integrated into the updating mechanism of f_t. The density score essentially enhances the fit of the data because it specifies the steepest ascent direction for maximizing the likelihood (Creal et al., 2013).

Finally, the scaling matrix S_t is proposed as the inverse of the information matrix of f_t raised to a power $\gamma > 0$ to account for the variance of the score as follows:

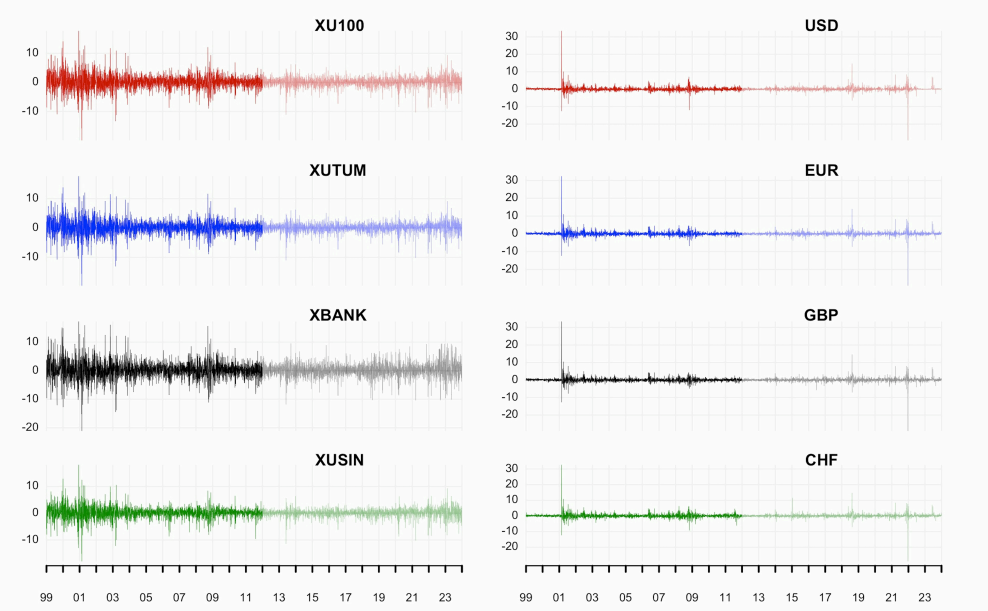
$$S_t = \mathcal{I}_{t \mid t-1}^{-\gamma}, \qquad \mathcal{I}_{t \mid t-1} = E_{t-1}[\nabla_t \nabla_t'] \tag{5}$$

E_{t-1} denotes expectation with respect to $p(y_t \mid f_t, \mathcal{F}_t; \theta)$. The parameter γ usually takes one of the three values in $\{0, 1, \frac{1}{2}\}$. The first choice corresponds to no scaling where S_t is an identity matrix of proper size. Alternatively, the conditional score is pre-multiplied by the inverse of its covariance matrix or the inverse of the square root of this for $\gamma = 1$ and $\gamma = \frac{1}{2}$, respectively (Ardia, Boudt, & Catania, 2019).

Equations (2-5) describe the GAS (1,1) model. The vector of the static parameters, along with the constant vector ω, and matrices A and B in the updating equation, are estimated using the maximum likelihood method. As with the GARCH model, one-step ahead time-varying parameters (f_{t+1}), thereby density forecasts are analytically feasible in a GAS model conditional on the information set $\{f_t, \mathcal{F}_t\}$.

Figure 1

Returns Series



The forecast period is toned down for each series.

4. Data and preliminary analysis

4.1. Data

The stock market data for the analysis include daily closing prices of the broad indexes BIST 100 (*XU100*) and BIST All Shares (*XUTUM*), as well as two sectoral indexes, namely, BIST Banks (*XBANK*) and BIST Industrials (*XUSIN*). The FX market data comprise daily US dollar (*USD*), European currency (*EUR*), British pound (*GBP*), and Swiss franc (*CHF*) rates, all denominated in Turkish lira. The entire dataset covers the period beginning

on Jan. 4, 1999 and extending to Dec. 29, 2023, for 25 years. The evolution of the price series is depicted in Figure A. 1.

The price series (P_t) are transformed into percentage-return series for further analysis using:

$$r_t = (logP_t - \log P_{t-1}) \times 100 \tag{6}$$

and can be examined in Figure 1. Descriptive statistics and results of select diagnostic tests in time series analysis, as presented in Table 1, reveal that all return series significantly depart from normality and exhibit stylized facts of financial time series with their heavy tails, being extremely apparent in FX returns, highly significant ARCH effects, and autocorrelations. Comparing the standard Ljung-Box statistics with their corrected counterparts (Q_{FZ}) for the index series suggests that autocorrelations turn out to be insignificant with a 95% confidence level for these weak white noise series. Hence, corrected portmanteau tests for GARCH processes, as proposed by Francq and Zakoian (2019), indicate that the ARCH effect should be prioritized over autocorrelation in modeling. However, autocorrelations are more nuanced in FX series even with GARCH-corrected tests.

In terms of standard deviation, indexes seem to be riskier than FX rates, with *XBANK* and *CHF* the riskiest in their categories. Despite the lower standard deviations, FX series are more vulnerable to outliers, reflecting their increased sensitivity to news with serious potential impacts on the overall economy. In this regard, the circuit breaker mechanism of Borsa İstanbul, as a risk management tool, must also be considered for lessened unusual price moves. Augmented Dickey-Fuller (ADF'79) tests strongly reject the null hypothesis of a unit root at the 5% significance level for all series. Conditioning on this level, we conclude that all series are stationary despite some contrary evidence through the stationarity test of Kwiatkowski, Phillips, Schmidt and Shin (KPSS'92) for FX series. In this respect, high persistence in financial time series should be accounted for as this feature of financial processes results in a tendency to reject null of stationarity in KPSS and LMC (Leybourne and McCabe, 1994) tests irrespective of the real situation (Caner & Kilian, 2001). Moreover, such contradictory results apply to FX series in both in-sample and forecast periods, which are specified in the appendix (see Table A. 1). The same fact is also noted for the index series during the forecast period. This leads one to conclude that size distortion in KPSS reported by Caner and Kilian (2001) might influence index results, further that persistence and long memory features and jumps are more prominent and effective in FX series. Strong rejection of the unit root via ADF tests and visual inspection (Figure 1) suggest a clear mean reversion in series. A deeper analysis of the degree of integration is likely to provide more convincing evidence for the series.

Table 1

Descriptive and Diagnostic Statistics of Return Series

Panel A: Descriptive Statistics								
Series	**N**	**Mean**	**Std. Dev.**	**Skewness**	**Kurtosis**	**Minimum**	**Median**	**Maximum**
XU100	6250	0.0896	2.1148	−0.0391	10.1850	−19.9784	0.1275	17.7649
XUTUM		0.0924	2.0405	−0.0988	10.6995	−19.6949	0.1393	17.6787
XBANK		0.0856	2.6675	0.1080	7.3968	−21.1711	0.0387	17.2617
XUSIN		0.1014	1.8708	−0.2697	12.1172	−18.0145	0.1714	18.0462
USD	6233	0.0728	1.1313	2.0575	215.4167	−29.3976	0.0270	33.4732
EUR		0.0721	1.1276	1.8467	197.7940	−29.1531	0.0260	32.4513
GBP		0.0686	1.1440	2.1102	201.6060	−29.1548	0.0385	33.4941

Panel A: Descriptive Statistics								
CHF		0.0809	1.2026	1.7674	158.1338	−29.1281	0.0338	32.6420
Panel B: Diagnostic Statistics								
Series	**JB (× 10^{-6})**	**ADF**	**KPSS**	**Arch-LM**	**Q_{LB} (5)**	**Q_{LB} (10)**	**Q_{FZ} (5)**	**Q_{FZ} (10)**
XU100	0.01***	−17.73***	0.15	938.11***	9.02	33.48***	2.72	14.73
XUTUM	0.02***	−17.50***	0.16	970.46***	10.68*	38.41***	3.07	16.49*
XBANK	0.01***	−16.74***	0.17	640.13***	8.22	21.27**	3.59	11.03
XUSIN	0.02***	−17.18***	0.17	1366.03***	13.51**	46.27***	3.63	18.29*
USD	11.73***	−12.50***	0.47**	203.72***	137.88***	180.03***	8.13	17.70*
EUR	9.86***	−13.17***	0.39*	209.29***	167.83***	200.05***	9.95*	18.80**
GBP	10.26***	−13.24***	0.44*	203.07***	188.55***	220.40***	11.15**	18.66**
CHF	6.26***	−13.26***	0.38*	211.01***	143.50***	177.84***	10.71*	21.24**

***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. N and Std. Dev. denote the number of observations and standard deviation, respectively. JB, ADF, KPSS, Arch-LM stand for Jarque-Bera, Augmented Dickey-Fuller, Kwiatkowski–Phillips–Schmidt–Shin, and Arch Lagrange Multiplier test statistics, respectively. Q_{LB} (i) and Q_{FZ} (i) denote the portmanteau (Q) test statistics of Ljung–Box and Francq-Zakoian at lag i, respectively.

For each asset studied and model estimated, the first half of the series ending on Dec. 30, 2011, constitutes the in-sample period while the forecast, or out-of-sample period starts on Jan. 2, 2012, and extends to the end of the entire period. The statistics for the in-sample and forecast periods are provided in Table A. 1. For BIST indexes, extrema at both ends occurred during the in-sample period. The maximums for all on Dec. 5, 2000, reflected positive views in anticipation of the International Monetary Fund's approval of the stand-by arrangement to Türkiye. In contrast, one of the most severe economic crises of Türkiye, which is nicknamed "Black Wednesday" (Hürriyet Newspaper, 2001) caused the minimums on Feb. 21, 2001. The same crisis resulted in maximum for all FX rate series two days later, on Feb. 23, 2001. The minimum rates, however, occurred during the forecast period, on Dec. 22, 2021, after the inception of "FX-protected deposit account" as a partial remedy for the soaring rates against the Turkish lira.

Comparing some notable statistics between periods, we report that risk in BIST is considerably (except for *XUSIN*), whereas risk in the FX market is slightly higher in the in-sample period than the forecast period in terms of standard deviation. All series have longer left tails in the forecast period, as suggested by the third standardized moment, namely skewness, which might have some implications for density and VaR forecasts. Autocorrelations turned out to be insignificant in both subperiods except for *XUSIN* with respect to Q_{FZ} statistic which assumes that the return processes are weak white noise.

4.2. In-sample estimations

The stochastic process, ε_t in GARCH (1,1) estimation for all series is the error from the mean equation, which replaces Equation (1):

$$r_t = \mu + \varepsilon_t,\ \varepsilon_t = \sqrt{h_t}\eta_t \tag{7}$$

where η_t is a series of independent and identically distributed (*iid*) random variables with zero mean and unit variance. We consider four laws for η_t, the Normal ($\mathcal{N}$), the Student-t ($\mathcal{T}$) and their skewed versions ($sk\mathcal{N}$ and $sk\mathcal{T}$), i.e., the set of conditional distributions is $\mathcal{D} = \{\mathcal{N}\ ,\ sk\mathcal{N}\ ,\ \mathcal{T}\ ,\ sk\mathcal{T}\}$. Since forecasts are estimated in *the rugarch* package of the statistical software R, the skewed distributions are Fernández and Steel's (1998) reparametrized versions so that they are standardized to have zero mean and unit variance.

The same holds for Student's *t* distribution, which was readily standardized in a 3-parameter setting with the package (Galanos, 2023).

In contrast, GAS (1,1) estimations are performed using the GAS package (Ardia et al., 2019). Densities $p(r_t \mid f_t, \mathcal{F}_t; \theta)$ generating the returns in GAS are also assumed to be those associated with the set $\mathcal{D}$. The summaries of both models for the *XU100* series for the in-sample period, i.e., $r_{1:T}$ where $T = 3238$ are presented in Table 2.

Table 2

Summary of In-Sample GARCH (1,1) and GAS (1,1) Estimations on BIST 100 Index Returns (XU100)

GARCH (1,1) Model									
Cond. Dist.	$\hat{\mu}$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\xi}$	$\hat{\nu}$	**AIC**	**SBIC**	**CAIC**
$\mathcal{N}$	0.1193	0.0812	0.1015	0.8889	-	-	4.4050	4.4125	4.4137
	(3.3778)	(2.2449)	(4.1522)	(32.4632)					
$sk\mathcal{N}$	0.1093	0.0818	0.1033	0.8870	0.9478	-	4.4037	4.4131	4.4146
	(2.8993)	(2.4152)	(4.5202)	(35.0072)	(31.9561)				
$\mathcal{S}$	0.1241	0.0904	0.0985	0.8890	-	7.3352	4.3667	4.3761	4.3776
	(3.7476)	(2.7576)	(4.8427)	(39.0890)		(8.1775)			
$sk\mathcal{S}$	0.1154	0.0897	0.0984	0.8891	0.9776	7.3886	4.3671	4.3783	4.3802
	(3.1859)	(2.7821)	(4.8962)	(39.5855)	(36.5246)	(8.2126)			
GAS (1,1) Model									
Cond. Dist.	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_4$	$\hat{a}_2$	$\hat{b}_2$	**AIC**	**SBIC**	**CAIC**
$\mathcal{N}$	0.1092	0.0327	-	-	0.1470	0.9796	4.4080	4.4155	4.4168
	(3.2526)	(4.3373)			(10.6483)	(211.1441)			
$sk\mathcal{N}$	0.1080	0.0167	−0.1189	-	0.0376	0.9791	4.4079	4.4173	4.4188
	(3.2248)	(4.3516)	(-1.5608)		(10.4973)	(207.6176)			
	0.1080		*0.9703*						
$\mathcal{S}$	0.1327	0.0272	−2.3592	-	0.2327	0.9794	4.3693	4.3787	4.3803
	(4.1471)	(3.3393)	(-8.4987)		(8.2589)	(164.8073)			
	0.1327		*7.9714*						
$sk\mathcal{S}$	0.1344	0.0166	0.0245	−2.3634	0.0583	0.9794	4.3699	4.3812	4.3830
	(4.0995)	(3.4188)	(0.2550)	(-8.4962)	(8.2298)	(163.6661)			
	0.1344		*1.0061*	*7.9563*					

$\mathcal{N}$ and S denote the Normal and Student-*t* distributions; $skN$and skS denote their skewed versions. AIC, SBIC, and CAIC denote the Akaike, Schwarz Bayesian, and Consistent Akaike information criteria, respectively. Numbers in parentheses are the t-statistics for the estimated parameters, while those in italic are the skew and shape parameter estimations with GAS, which are transformed for a direct comparison of their counterparts with GARCH.
Note: We restrict comparison of models' fits to leading assets in each class, namely *USD* and *XU100*, in line with the study's focus.

Concerning GAS practice, the only time-varying parameter is the scale parameter, which is our deliberate choice to make it compatible with the GARCH model. All parameters, whether time-varying (f_t) or static (θ) in Equation (2) are stacked into a new vector ϕ_t following the usual order: location (μ), scale (σ), skew (ξ) and shape (ν). For a three-parameter distribution with shape but skew, the former takes the position of the latter. Hence, $\phi_t = [\mu_t \;\; \sigma_t \;\; \nu_t]'$ for $\mathcal{S}$. The vector $\kappa = [\kappa_1 \;\; \kappa_2 \;\; \kappa_3 \;\; \kappa_4]'$ (assuming a four-parameter law) replaces ω in Equation (3). Then, the coefficient matrices are $A = \text{diag}\,(0, a_2, 0, 0)$ and $B = \text{diag}\,(0, b_2, 0, 0)$

(assuming $sk\mathcal{S}$ as the underlying distribution) due to our choice explained before. The parameter γ is set to zero; therefore, the score was not scaled for all estimations. Finally, the estimators $\hat{\kappa}$, $\hat{A}$ and $\hat{B}$ reported in the model summaries are based on reparameterization through link functions to satisfy the distributional constraints on parameters during MLE. We report the skew and shape estimators mapped by the functions' inverses in italic for a better comparison with the GARCH model's estimators and refer to Ardia et al. (2019) for technical details on the estimation. The statistics in Table 2 suggest that both models' estimations of location, skew, and shape parameters are close to each other. GARCH slightly outperforms GAS in terms of information criteria for all distributional assumptions. All estimators except those for the skew parameters in the GAS models are significant. The insignificant ones under $sk\mathcal{N}$ and $sk\mathcal{T}$ assumptions are found to be significant in GARCH. The reparameterization scheme may explain the difference[1]. The sum of estimated coefficients $\hat{\alpha}_1$and $\hat{\beta}_1$ in GARCH models are in the interval $[0.9875, 0.9904]$, which clearly indicates that shocks to *XU100* returns have highly persistent effects on volatility.

Because this paper mainly focuses on forecasts, we content with the model summaries of a representative index and FX rate, *XU100* and *USD*. The model summaries for the latter, specifically information criteria, in Table 3 indicate that GAS performs worse in *the USD* series under $\mathcal{N}$ and its skewed version $sk\mathcal{N}$ since the gap is more apparent with these conditional distributions. Like the overall results for *XU100*, GARCH outperforms GAS by a small margin. The persistence in *USD* volatility is also strong, as suggested by the GARCH estimators.

Table 3

Summary of In-Sample GARCH (1,1) and GAS (1,1) Estimations of US Dollar Returns (USD)

GARCH (1,1) Model									
Cond. Dist.	$\hat{\mu}$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\xi}$	$\hat{\nu}$	**AIC**	**SBIC**	**CAIC**
$\mathcal{N}$	0.0497	0.009	0.1378	0.8568	-	-	2.2781	2.2856	2.2869
	(2.7017)	(2.8032)	(6.6970)	(41.2909)					
$sk\mathcal{N}$	0.0575	0.0103	0.1392	0.8524	1.1768	-	2.2620	2.2714	2.2729
	(3.1836)	(3.1081)	(7.3825)	(43.8390)	(32.0664)				
$\mathcal{S}$	0.0284	0.0132	0.1541	0.8399	-	6.0603	2.2222	2.2316	2.2331
	(1.4883)	(2.7007)	(6.3569)	(33.8044)		(6.0282)			
$sk\mathcal{S}$	0.0535	0.0114	0.1481	0.8495	1.1721	6.2219	2.2115	2.2227	2.2246
	(2.7914)	(2.6287)	(6.5511)	(36.1883)	(34.8772)	(5.8746)			
GAS (1,1) Model									
Cond. Dist.	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_4$	$\hat{a}_2$	$\hat{b}_2$	**AIC**	**SBIC**	**CAIC**
$\mathcal{N}$	0.0350	0.0821	-	-	0.2686	0.6636	2.9782	2.9857	2.9869
	(1.9191)	(5.8593)			(8.5144)	(33.6001)			
$sk\mathcal{N}$	0.0551	−0.0039	0.8968	-	0.0026	0.9836	2.5438	2.5532	2.5547
	(3.5471)	(-32.4207)	(5.4106)		(9.0232)	(693.0167)			
	0.0551		*1.2103*						
$\mathcal{S}$	0.0350	−0.0245	−2.9166	-	0.3636	0.9729	2.2298	2.2392	2.2407
	(3.1940)	(-3.4297)	(-10.5982)		(10.7827)	(156.4432)			

[1]The authors state that a link function $\Lambda : \mathbb{R}^j \to \mathbb{R}^j$, $\phi_t \equiv \Lambda(\tilde{\phi}_t)$ is the standard solution in the GAS framework, whereas distributional constraints can be imposed directly in GARCH models (Ardia et al., 2019). This indirect mechanism which employs $\tilde{\phi}_t$ in MLE may have statistical consequences.

	0.0350		*6.3616*						
	0.0522	−0.0060	0.4701	−2.9867	0.0840	0.9772	2.2243	2.2356	2.2374
skS	(4.4654)	(-2.3929)	(4.1185)	(-10.4418)	(10.1769)	(172.5839)			
	0.0522		*1.1154*	*6.2093*					

N and S denote the Normal and Student-t distributions; $skN$and skS denote their skewed versions. AIC, SBIC, and CAIC denote the Akaike, Schwarz Bayesian, and Consistent Akaike information criteria, respectively. Numbers in parentheses are the t-statistics for the estimated parameters, while those in italic are the skew and shape parameter estimations with GAS, which are transformed for a direct comparison of their counterparts with GARCH.
Note: We restrict comparison of models' fits to leading assets in each class, namely *USD* and *XU100*, in line with the study's focus.

5. Forecasting density and VaR

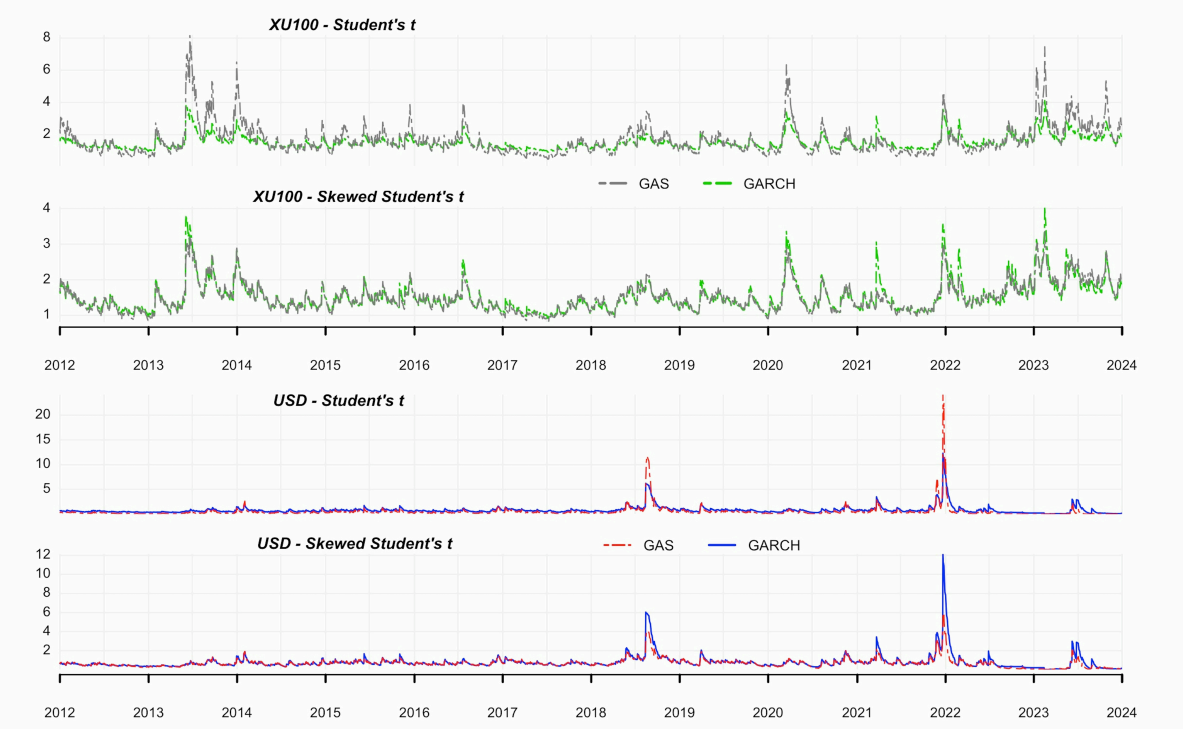
The forecasting framework is based on the autoregressive structure of both models. However, the forecast variable differs from the economic variables analyzed in Marcellino et al. (2006). In the current setting, it is the latent volatility, namely the scale parameter of the distribution of returns. The GAS estimate is based on the score vector of the preceding period, whereas a better GARCH forecast requires squared error terms or shocks to returns. Therefore, we proceed with iterated multi-step forecasts, as described in Chevillon (2007), but supplemented by the information set that contains the realized return required for the update in Equations (1) and (3).

5.1. Density forecasts

The density forecasts $\hat{\phi}^{GARCH}_{T+i \mid T+i-1}$ and $\hat{\phi}^{GAS}_{T+i \mid T+i-1}$ with $i = 1, 2, \cdots, H$ for the returns of assets in both categories, which cover the pseudo-out-of-sample period, are accomplished by taking $T = 3238\,(3241)$ and $H = 3012\,(2992)$ for indexes (rates). For each time step, a density forecast is made conditional on the available information set, that is available as at the previous one. The subscript notation expresses this conditioning. However, this set does not contain the revised density parameter estimators at all steps. Rather than updating the estimators with each marginal observation, we prefer rolling forecasts that enable refitting the model in every 63rd and alternatively 125th observation to control for parameter uncertainty. These window sizes roughly correspond to quarterly and semiannual updates, respectively. Rolling forecasts do not allow expansion of the window sizes in estimations as opposed to recursive forecasts, which use all observations since the onset. Therefore, the oldest observations are dropped, and new observations are added to maintain estimation windows of equal size throughout the forecast. This preference is primarily based on the observed strong persistence in the GARCH models, which implies that the data-generating processes are almost nonstationary. Consequently, as the parameters of the model are more likely to change, dropping the old and irrelevant data helps to avoid biased forecasts (Elliott & Timmermann, 2016).

Figure 2

Volatility Series Forecasted by GARCH and GAS models



The forecasted time-varying volatility series in Figure 2 suggest that, except for a few spikes, GAS and GARCH forecasts of *USD* volatility closely match each other with both underlying assumptions, $\mathcal{T}$ and $sk\mathcal{T}$. Two of the exceptions deserve attention in terms of both their magnitudes and the opposing behavior of GAS and GARCH on the corresponding occasions. Chronologically, the first corresponds to the Turkish lira crisis, which peaked in August 2018. The crisis was driven by the increased tension between US and Türkiye due to political and economic disputes. The era started on July 1, 2018, and it was marked by the devaluation of the Turkish lira against the US dollar by 35% in just 47 days (Hadi, Karim, Naeem, & Lucey, 2023). The second relates to the repercussions of the launching of "FX-protected deposit account" on Dec. 22, 2021. While $\mathcal{T}$ assumption produces higher volatility forecasts on these days with GAS, the reverse holds for GARCH with the skewed version, $sk\mathcal{T}$. Truncating the *USD* volatilities at the maximum value of *the XU100* series, not shown here due to space limitations, reveals that close match holds for $sk\mathcal{T}$ whereas GAS tends to produce lower forecasts than GARCH with $\mathcal{T}$ in general, the two extremes just mentioned are exceptions. The congruency of models in *XU100* volatility forecasts is retained when the underlying distribution is $sk\mathcal{T}$. However, the GAS-forecasted *XU100* volatilities with $\mathcal{T}$ tend to be higher in relatively turbulent periods and lower in tranquil periods than their counterparts with GARCH. We limit visual volatility forecasts to leading assets and two conditional distributions because comparing forecasts requires formal tests and analysis of realized returns. A further reason is that $\mathcal{N}$, and to a lesser extent $sk\mathcal{N}$ lead to enormous jumps in the forecasted volatilities with GAS, as will be discussed.

5.2. Density forecast evaluations

Rolling forecasts based on two alternative models were evaluated using two scoring rules for their predictive accuracy. All evaluations in this section are based on forecasts with a parameter update frequency

of 63 trading days or quarterly refits. Analogous to point forecasts and their associated loss functions, a scoring rule is a loss function of the actual value of a variable of interest and its density forecast. Since the forecasts performed are density forecasts of the returns with the relevant parameters, of which scale parameter changes at each period, and a density forecast can be regarded as a collection of probabilities assigned by the forecaster or forecasting model to all possible values (Amisano & Giacomini, 2007), scoring rules neatly fit the evaluation purpose here.

As a well-known and basic scoring rule, the logarithmic score (LS) is the first considered scoring rule and is defined as follows:

$$LS_{T+i}(r,p) = \log p\Big(r_{T+i}; \hat{\phi}_{T+i}\Big) \quad i = 1, 2, \cdots, H \tag{8}$$

where $p(\cdot)$ is the density function of the conditional distribution in $\mathcal{D}$.

The second scoring rule is the weighted continuous ranked probability score (wCRPS), which extends the opportunity to evaluate forecasts with an emphasis on certain regions of interest for the predicted densities. wCRPS is defined as follows (Gneiting & Ranjan, 2011) :

$$wCRPS_{T+i}(r,p) = \int_{-\infty}^{\infty} w(z)\Big[F\Big(z; \hat{\phi}_{T+i}\Big) - \mathbb{I}\{r_{T+i} \leq z\}\Big]^2 dz \qquad i = 1, 2, \cdots, H \tag{9}$$

$w(z)$ denotes the weight function evaluated at the threshold z and determines a specific region of the distribution to be emphasized. The function is nonnegative on the real line, and the most trivial specification corresponds to uniform weighting over all regions with $w(z) = 1$. Other specifications proposed assign heavier weights to the center, tails, right tail, and left tail. In the financial context, the latter gains importance because it typically directs attention to losses, which are crucial ingredients of risk modeling. We refer to Gneiting and Ranjan (2011) for the functional forms of weights associated with different regions.

The second term in the integrand is the Brier or quadratic score, as discussed in Gneiting and Raftery (2007), along with other score functions in the forecasting literature. The cumulative density $F\Big(z; \hat{\phi}_{T+i}\Big)$ evaluated at the threshold z, conditional on the forecasted parameter vector, relates to the total density, hence probability, to the left of z by definition. Intuitively, the closer this probability to the value of the indicator function $\mathbb{I}$ which directly takes the realized return r_{T+i} as an input, the more accurate a density forecast becomes. Therefore, wCRPS can be loosely interpreted as the total weighted squared deviation in association with the density forecast of a continuous variable.

The lower wCRPS is preferred to the higher one; thus, it is a negatively oriented score and is a genuine loss function. In contrast, LS is positively oriented. Taking the negative of the score defined in Equation (8) for the sake of compatibility, the following analysis proceeds with the negative logarithmic score (NLS). The scores used to backtest GAS forecasts were estimated using the functionality provided by the GAS package. We have adapted the backlisting functions of the package in R environment (R Core Team, 2024) to implement the same for GARCH forecasts, taking the specifications of both the model and the *rugarch* package into account. The mean NLS and wCRPS statistics for the density forecasts are provided for the BIST index and FX rate series in Table A. 2 and Table A. 3, respectively. For each asset returns, the average scores with GARCH forecasts are in the first row, while those in italic relate to GAS forecasts. GARCH seems to perform better conditional on Normal law for BIST index returns, with the differentials in mean scores being remarkable in *XUSIN*. The same holds in general when the conditional distribution is the skewed version, $sk\mathcal{N}$. The mean scores, however, are mostly equal or very close to each other with $\mathcal{T}$ and $sk\mathcal{T}$. FX returns density forecasts

exhibit similar patterns under $\mathcal{N}$ and $sk\mathcal{N}$ assumptions. The wide gaps between the CHF forecast scores call for further examination of potential estimation issues. In contrast, GAS performed slightly better with $\mathcal{T}$ and $sk\mathcal{T}$. In this domain, NLS provides better predictive performance.

5.3. Formal comparison of forecasts

The formal testing framework was provided with the popular Diebold-Mariano (DM) test (Diebold & Mariano, 1995) used in forecasting literature. Denoting a score of interest and hence loss associated with a model m at t by ℓ_t^m, and the difference between scores by d_t, DM test is expressed as follows:

$$DM = \frac{\overline{d}}{\hat{\sigma}_d} \text{ where } \overline{d} = H^{-1} \sum_{i=1}^{H} \left(\ell_{T+i}^{GARCH} - \ell_{T+i}^{GAS} \right) \quad i = 1, 2, \cdots, H \tag{10}$$

Table 4

Diebold-Mariano Test Results for the Score Functions of the GARCH and GAS Density Forecasts

	NLS	wCRPS with an emphasis on				
		Uniform	Center	Tails	Right tail	Left tail
A. BIST index returns						
			Normal distribution			
XU100	−3.4770***	−2.0068**	−2.6226***	−1.6476*	−1.9732**	−2.2150**
XUTUM	−3.3943***	−2.0741**	−2.8936***	−1.6805*	−2.0564**	−2.2931**
XBANK	−3.0267***	−2.3771**	−2.1816**	−2.1444**	−1.1859	−2.5054**
XUSIN	−3.1160***	−1.1750	−3.4823***	−1.1240	−1.3658	−1.0839
			Student-t distribution			
XU100	1.4796	0.9448	1.1466	0.3860	0.8505	0.5132
XUTUM	0.8679	0.8870	1.1389	0.2381	0.9256	0.3849
XBANK	0.0343	0.2709	0.6485	−0.4046	−0.0388	0.5108
XUSIN	−0.5240	−0.9220	−0.8957	−0.7475	0.1822	−1.2873
			Skewed Normal distribution			
XU100	−4.0049***	−2.4808**	−3.1505***	−2.0012**	−2.3513**	−2.7881***
XUTUM	−4.1034***	−2.6568***	−3.5958***	−2.0537**	−2.5277**	−3.0048***
XBANK	−1.7269*	−1.7352*	−1.8792*	−1.6105	−0.9515	−1.9007*
XUSIN	−3.5093***	−1.1067	−4.9401***	−0.9992	−1.4873	−0.9441
			Skewed Student's t distribution			
XU100	0.2246	0.5351	0.7278	0.0664	0.6510	0.1594
XUTUM	−0.1407	0.2462	0.4816	−0.2462	0.6595	−0.2804
XBANK	0.3429	0.6116	1.0955	−0.3158	0.3059	0.6972
XUSIN	−2.0158**	−2.4026**	−2.8226***	−1.2919	−1.0604	−2.0435**
B. The FX returns						
			Normal distribution			
USD	−7.1531***	−2.4250**	−7.4099***	−1.8113*	−2.5860***	−2.3728**
EUR	−4.7730***	−1.4111	−7.0997***	−1.2258	−1.4677	−1.3973
GBP	−8.0647***	−2.0047**	−9.6794***	−1.4606	−2.0642**	−1.9677**
CHF	−6.0405***	−0.7064	−7.1403***	−0.6771	−0.7584	−0.6697

	NLS	wCRPS with an emphasis on				
		Uniform	Center	Tails	Right tail	Left tail
			Student-t distribution			
USD	5.6914***	3.4744***	4.8120***	2.0803**	2.5591**	2.7546***
EUR	0.3196	0.9134	1.2095	0.6347	0.3865	0.4918
GBP	2.0538**	2.0145**	2.4071**	1.6336	1.1638	1.7020*
CHF	0.7158	1.4844	1.8129*	1.1077	0.9041	0.7831
			Skewed Normal distribution			
USD	−4.4690***	−1.2626	−6.1603***	−1.0961	−1.3001	−1.2747
EUR	−4.3225***	−1.0864	−5.6540***	−0.9985	−1.0957	−1.0876
GBP	−5.2430***	−1.6618*	−7.2763***	−1.3994	−1.6923*	−1.6466*
CHF	−5.2367***	−1.1341	−6.8242***	−1.0881	−1.1727	−1.1146
			Skewed Student's t distribution			
USD	5.4934***	3.0448***	4.7348***	1.7243*	2.2036**	2.6415***
EUR	−0.1114	0.7863	1.2841	0.4406	0.3004	0.4753
GBP	1.7777*	1.8955*	2.5461**	1.4683	1.0763	1.7269*
CHF	0.2964	1.313	1.8971*	0.8786	0.7744	0.8388

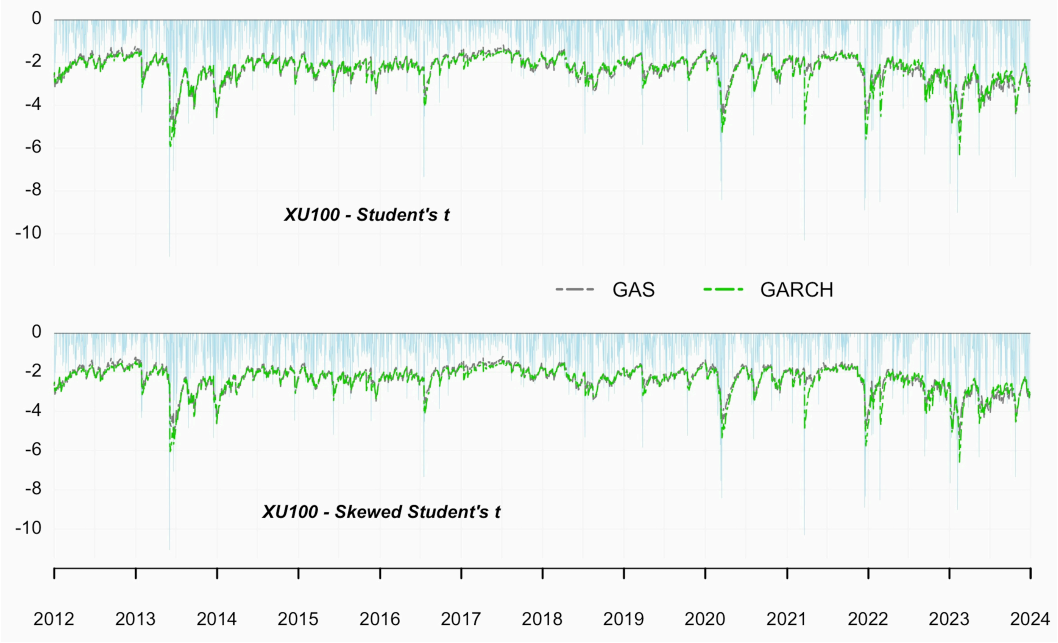
***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. NLS and wCRPS denote negative log scores and weighted continuous ranked probability scores, respectively. Diebold-Mariano test statistics are based on HAC variances.

In estimating the DM test statistic, we employed the heteroscedasticity and autocorrelation corrected (HAC) variance estimator proposed by Andrews (1991) and Andrews & Monahan (1992) and note that both the original DM test and the size-corrected version of this (D. Harvey, Leybourne, & Newbold, 1997) produce higher t-statistics, which leads to misleading results due to autocorrelation in difference series. Here, the scores, NLS, and five flavors of wCRPS across the two competing model forecasts are the corresponding losses, namely, error terms used to compute the test statistic.

DM test results for BIST index returns in Table 4 statistically confirm that GARCH mostly outperforms GAS by rejecting the null hypothesis of equal predictive accuracy (EPA) when the forecaster assumes that returns follow $\mathcal{N}$ or $sk\mathcal{N}$ distribution. A clear exception is the *XUSIN* series, in which no contrary evidence against EPA is provided with wCRPS statistics that emphasize tails collectively, each tail alone, and the whole density uniformly. The relatively increased significance of the test results with $sk\mathcal{N}$ assumption is another interesting finding, specifically for *the XU100* and *XUTUM* density forecasts. Shifting to $sk\mathcal{N}$ works the opposite way for *XBANK*, thereby rendering significant statistics with $\mathcal{N}$ insignificant.

Figure 3

VaR Series forecasted for XU100 *by GARCH and GAS models*



On the other hand, positive DM test statistics reported for *XU100* and *XUTUM* reveal that GAS outperforms GARCH with $\mathcal{T}$ albeit they present no statistical evidence. The results are more scattered with $sk\mathcal{T}$ assumption. However, reported t-statistics are generally insignificant. GARCH's forecasting capacity increases by incorporating the skew parameter for *XUSIN*.

Conditioning on Normal law, the superiority of GARCH forecasts applies to FX returns, even with higher t-statistics, as can be observed in panel B of Table 4. *EUR* and *CHF* are clear exceptions when DM statistics on wCRPS, excluding those focusing on center, are considered. GARCH's superiority is limited to NLS and wCRPS on center when the distributional assumption is $sk\mathcal{N}$ in FX returns. The only return series for which GAS forecasts are superior is *USD* when the conditional density is $\mathcal{T}$ or $sk\mathcal{T}$. *GBP* is another series with some little evidence to this end. For the rest of the models, GAS seems to be better at forecasting performance but lacks statistical support.

5.4. Value-at-Risk (VaR) forecasts

Value-at-Risk (VaR) is a practical, convenient, and concise metric that summarizes downsize risk. VaR specifies the cut-off loss typically seen in a portfolio or financial unit, above which losses have a very low probability of occurrence within a specific investment horizon. This probability coincides with a given risk level α for which 5% and 1% are conventional values in the financial industry. From a statistical perspective, VaR forecasts as a function of the estimated density parameters can be expressed as follows:

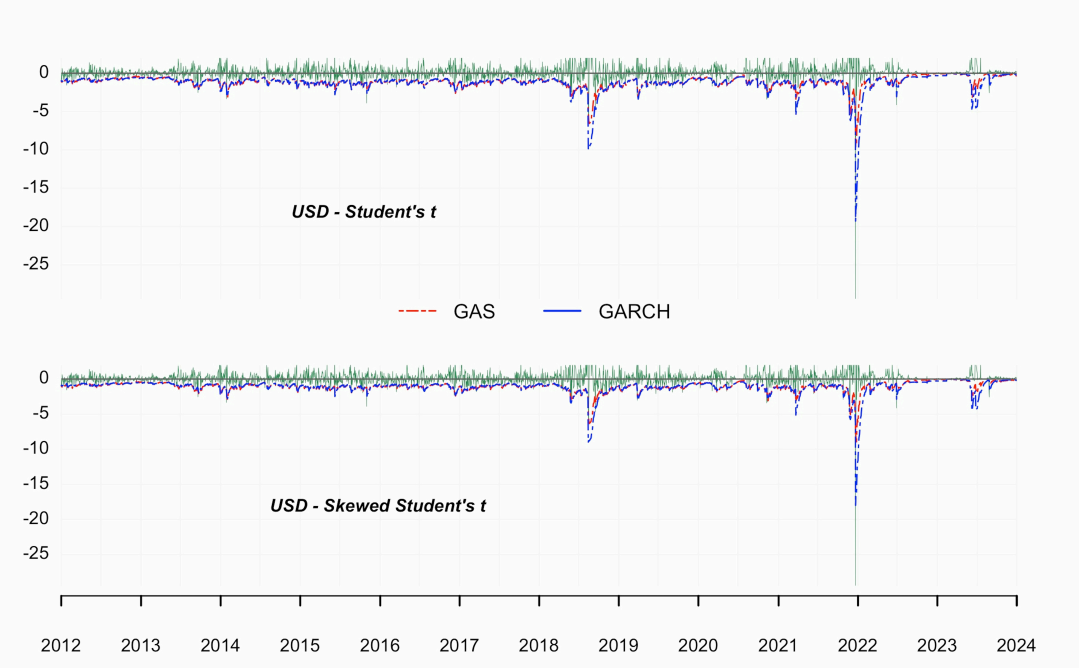
$$\Pr\left(r_{T+i} < VaR^{\alpha}_{T+i}\left(\hat{\phi}_{T+i\,|\,T+i-1}\right)\right) = \alpha \quad i = 1, 2, \cdots, H \tag{11}$$

where $Pr(\cdot)$ denotes the expected probability. This representation is based on return distribution as opposed to the preceding description, which reflects the loss-oriented approach of the risk management literature (e.g., Dowd, 2005; Jorion, 2007). Like the density scores, we estimate 1-day VaR forecasts for all series and distributional assumptions at the 5% and 1% risk levels. The forecasted 5% VaR series for the leading assets juxtaposed on returns are depicted in Figure 3 (XU100) and Figure 4 (USD). The figures suggest

that GARCH's VaR forecasts are more conservative than those of GAS in general, particularly at points with elevated downward pressure.

Figure 4

VaR Series Forecasted for USD by GARCH and GAS models



5.5. Backtesting VaR forecasts

We consider four statistical tests, one metric, and two loss functions to backtest the VaR series forecasted by the two models. The primary two tests are unconditional and conditional coverage tests (UC, CC) of Kupiec (1995) and Christoffersen (1998), respectively. The third test is the dynamic quantile test (DQ) proposed by Engle and Manganelli (2004), which considers the plausibility of autoregressive dependence in financial time series as it applies to VaR. Rather than conditioning on the number of exceedances and possible dependence schemes, the final test (VaRd) by Christoffersen and Pelletier (2004) specifies a correct model as the one in which the duration between VaR violations, or equivalently exceedances, is memoryless. The ratio of realized exceedances to the expected number of such exceedances implied by the risk level (A/E) is a straightforward metric and expected to be close to one for a correctly specified model.

In contrast to tests and metrics, loss functions provide a sound comparison base using DM tests. The first is the quantile loss function (QL), also known as the asymmetric VaR loss function, which penalizes exceedances more heavily (González-Rivera, Lee, & Mishra, 2004). The other loss function (FZL) was proposed by Fissler and Ziegel (2016) and contains not only the VaR metric but also the expected shortfall (ES) in its specification. ES is superior to VaR because the latter provides the practitioner with only the loss threshold in relation to a given risk level, while ES represents the average loss to be incurred once the threshold is exceeded. In addition to its conservative and more realistic nature, ES is also a coherent risk measure (Artzner, Delbaen, Eber, & Heath, 1999).

The results of the backtest procedures summarized are presented for 5% VaR forecasts of index and FX returns in Table 5 and Table 6, respectively. Except for the DM test statistics in the rightmost two columns and

the A/E (Actual/Expected) ratios, all reported numbers are probabilities of a correctly specified VaR model with the corresponding test. The probabilities and ratios shown in italic are for GAS model forecasts. These results suggest that VaR forecasts by both the GARCH and GAS models on *XU100* and *XUTUM* return indexes are acceptable in general regardless of the conditional distribution. While GARCH forecasts are slightly better with respect to A/E ratios, DM tests with (asymmetric) quantile losses attest the same when the conditional distribution is $\mathcal{N}$ or $sk\mathcal{N}$. Other DM test statistics (DM_{FZ}) align with this finding (except *XUTUM* conditional on $\mathcal{N}$), but to a lesser degree. The GAS model provides better VaR forecasts through both DM tests with $\mathcal{T}$ and $sk\mathcal{T}$, however this eminence lacks statistical support. On the other hand, $\mathcal{T}$ results in the worst VaR models for *XUSIN* because all tests, except the VaR duration test, strongly reject the null, which proposes that the underlying model is correctly specified. This end is also apparent for the A/E ratios, which indicate that both GARCH and GAS undershot VaR exceedances by 24%. In contrast, the best VaR models for *XUSIN* were associated with $sk\mathcal{N}$. According to the results of the DM test based on FZL for the index, GARCH outperforms its counterpart when the distributional assumption is $\mathcal{N}$ or $sk\mathcal{N}$. There was no statistically sound winner in both DM tests for *XBANK* with all distributions, although all GARCH losses were lower. Another notable finding for this index is the overall inadequacy of both models due to the different forms of dependence in the VaR series, as verified by CC, DQ, and VaRd.

A thorough examination of Table 6 reveals that VaR forecasts for FX returns with both models are insufficient in capturing exceedances. Compared with loss functions of *USD* VaR forecasts, GAS clearly surpasses GARCH with $\mathcal{T}$ and $sk\mathcal{T}$ whereas the opposite holds with $\mathcal{N}$ or $sk\mathcal{N}$ except for QL under $sk\mathcal{N}$ assumption. However, the superiority of GAS disappears with other FX rates, whereas GARCH keeps resulting in significantly lower FZL values. Furthermore, GARCH's lower QLs in *GBP* VaR forecasts are statistically significant.

Table 5

Comparative Backtest Results for GARCH and GAS Value-at-Risk (VaR) Forecasts for BIST Index Returns (α=5%)

	UC	CC	A/E	DQ	VaRd	DM_{QL}	DM_{FZ}
			Normal distribution				
XU100	0.5218	0.8117	0.9495	0.2301	0.5591	−2.2330**	−1.8718*
	0.4174	**0.7106**	**0.9363**	**0.2655**	**0.7831**		
XUTUM	0.8275	0.9393	0.9827	0.2395	0.4379	−2.2809**	−1.6029
	0.5218	**0.7277**	**0.9495**	**0.2766**	**0.7708**		
XBANK	0.3261	0.0139	0.9230	0.0072	0.0156	−0.9603	−0.9919
	0.1576	**0.0112**	**0.8898**	**0.0178**	**0.0361**		
XUSIN	0.0465	0.1155	1.1620	0.0563	0.3256	−1.4299	−2.2517**
	0.3896	**0.4898**	**1.0691**	**0.0187**	**0.9298**		
			Student-t distribution				
XU100	0.8934	0.8222	0.9894	0.0935	0.2398	0.9352	1.4553
	0.3060	**0.5443**	**1.0823**	**0.0773**	**0.7928**		
XUTUM	0.3896	0.4898	1.0691	0.0241	0.3349	1.1165	0.8594
	0.1775	**0.3387**	**1.1089**	**0.0396**	**0.8431**		
XBANK	0.8275	0.0609	0.9827	0.0197	0.0634	−0.3542	−0.1140
	0.6992	**0.0985**	**0.9695**	**0.0418**	**0.0766**		
XUSIN	0.0001	0.0002	1.3280	0.0001	0.5470	0.4509	0.1580
	0.0001	**0.0003**	**1.3214**	**0.0001**	**0.6309**		

	UC	CC	A/E	DQ	VaRd	DM_{QL}	DM_{FZ}
			Skewed Normal distribution				
XU100	0.2149	0.4627	0.9031	0.3912	0.8162	−1.9648**	−1.7674*
	0.1847	**0.4147**	**0.8964**	**0.2982**	**0.9993**		
XUTUM	0.4174	0.6175	0.9363	0.1617	0.3137	−2.1943**	−1.6949*
	0.1847	**0.3838**	**0.8964**	**0.1246**	**0.6345**		
XBANK	0.1847	0.0324	0.8964	0.0319	0.0413	−0.7676	−0.7935
	0.2856	**0.0043**	**0.9163**	**0.0001**	**0.0168**		
XUSIN	0.8934	0.6189	0.9894	0.2416	0.6628	−1.5359	−2.3524**
	0.9733	**0.4654**	**1.0027**	**0.0052**	**0.5141**		
			Skewed Student's t distribution				
XU100	0.5784	0.8561	0.9562	0.1601	0.3628	1.3807	0.8081
	0.4364	**0.6421**	**1.0624**	**0.1860**	**0.6547**		
XUTUM	0.7626	0.7525	0.9761	0.1784	0.2093	0.9199	0.2697
	0.4364	**0.6421**	**1.0624**	**0.1301**	**0.6226**		
XBANK	0.8275	0.0609	0.9827	0.0197	0.0634	−0.2160	−0.0725
	0.9070	**0.1759**	**1.0093**	**0.0258**	**0.1011**		
XUSIN	0.1310	0.2257	1.1222	0.0348	0.6454	−0.3206	−1.2309
	0.0109	**0.0185**	**1.2085**	**0.0029**	**0.6461**		

***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. Numbers in bold refer to the GAS model.
UC: Unconditional coverage test of Kupiec (1995) CC: Conditional coverage test of Christoffersen (1998)
A/E: Actual over-excused ratio DQ: Dynamic quantile test of Engle and Manganelli (2004)
VaRd: VaR duration test by Christoffersen and Pelletier (2004) QL: Quantile loss function of González-Rivera et al. (2004)
FZ: Fissler & Ziegel's (2016) loss function DM_L: Diebold-Mariano test with loss model L
Diebold-Mariano test statistics are based on HAC variances. Except for A/E, the first six columns contain the acceptance probabilities of the null hypothesis with the corresponding test.

Table 6

Comparative Backtest results for GARCH and GAS Value-at-Risk (VaR) Forecasts for BIST Index Returns (α=5%)

	UC	CC	A/E	DQ	VaRd	DM_{QL}	DM_{FZ}
			Normal distribution				
USD	0.0000	0.0000	0.5281	0.0000	0.1193	−2.4538**	−9.1604***
	0.0000	**0.0000**	**0.4211**	**0.0000**	**0.1587**		
EUR	0.0000	0.0000	0.6618	0.0033	0.0159	−1.4801	−6.3473***
	0.0000	**0.0000**	**0.4947**	**0.0000**	**1.0000**		
GBP	0.0002	0.0000	0.7219	0.0025	0.3399	−1.9860**	−9.4630***
	0.0000	**0.0000**	**0.5682**	**0.0000**	**1.0000**		
CHF	0.0000	0.0000	0.5749	0.0001	0.0335	−1.1072	−7.1392***
	0.0000	**0.0000**	**0.4412**	**0.0000**	**0.6771**		
			Student-t distribution				
USD	0.0000	0.0000	0.5414	0.0000	0.0441	3.7018***	3.9988***
	0.0000	**0.0000**	**0.6083**	**0.0001**	**0.0150**		
	0.0001	0.0003	0.7086	0.0189	0.0278	0.6929	−0.2641

	UC	CC	A/E	DQ	VaRd	DM_{QL}	DM_{FZ}
EUR	**0.0014**	**0.0027**	**0.7553**	**0.0304**	**0.0597**		
GBP	0.0060	0.0018	0.7888	0.0501	0.2611	0.8494	1.4873
	0.1113	**0.0002**	**0.8757**	**0.0004**	**0.9383**		
CHF	0.0000	0.0000	0.6083	0.0005	0.0385	1.7672*	0.1277
	0.0000	**0.0001**	**0.6818**	**0.0064**	**0.0176**		
Skewed Normal distribution							
USD	0.0000	0.0000	0.6350	0.0016	0.6950	−1.4281	−6.5409***
	0.0000	**0.0000**	**0.4813**	**0.0000**	**0.4278**		
EUR	0.0019	0.0058	0.7620	0.1401	0.0187	−1.2931	−3.9723***
	0.0000	**0.0000**	**0.6751**	**0.0042**	**0.7413**		
GBP	0.0102	0.0035	0.8021	0.0551	0.4865	−1.8370*	−4.4813***
	0.0000	**0.0000**	**0.6818**	**0.0004**	**0.2344**		
CHF	0.0004	0.0015	0.7286	0.0183	0.0308	−1.4414	−5.4852***
	0.0000	**0.0000**	**0.5816**	**0.0001**	**0.4789**		
Skewed Student's t distribution							
USD	0.0002	0.0003	0.7152	0.0226	0.5577	2.2422**	2.9226***
	0.0001	**0.0002**	**0.7086**	**0.0029**	**0.2821**		
EUR	0.0339	0.0781	0.8356	0.4819	0.0068	−0.1046	−1.4076
	0.0339	**0.0524**	**0.8356**	**0.1253**	**0.0155**		
GBP	0.1113	0.0152	0.8757	0.1615	0.7562	0.3606	0.8998
	0.3244	**0.0018**	**0.9225**	**0.0049**	**0.9503**		
CHF	0.0014	0.0056	0.7553	0.0581	0.0094	0.4817	−1.6095
	0.0019	**0.0058**	**0.7620**	**0.0554**	**0.0091**		

***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. Numbers in bold refer to the GAS model.
UC: Unconditional coverage test of Kupiec (1995) CC: Conditional coverage test of Christoffersen (1998)
A/E: Actual over-excused ratio DQ: Dynamic quantile test of Engle and Manganelli (2004)
VaRd: VaR duration test by Christoffersen and Pelletier (2004) QL: Quantile loss function of González-Rivera et al. (2004)
FZ: Fissler & Ziegel's (2016) loss function DM_L: Diebold-Mariano test with loss model L
Diebold-Mariano test statistics are based on HAC variances. Except for A/E, the first six columns contain the acceptance probabilities of the null hypothesis with the corresponding test.

The backtests for VaR forecasts at the 1% risk level are also provided in Table A. 4 and Table A. 5. At first glance, the tables suggest that overall results have switched between the two asset classes. Except for the VaR duration tests (VaRd) of Christoffersen and Pelletier (2004), all statistical tests and A/E ratios on FX VaR forecasts provide statistical evidence for the correct specification of both GARCH and GAS. A notable finding with *GBP* is that DM tests using both VaR loss functions strongly favor the GAS model when the underlying distribution is $\mathcal{T}$ and $sk\mathcal{T}$. To a lesser but significant extent, GARCH replaces GAS when $\mathcal{N}$ and $sk\mathcal{N}$ are the conditional laws. GARCH's superiority in *USD* forecasts remains with $\mathcal{N}$, but not with $sk\mathcal{N}$. Similarly, $\mathcal{T}$, but not $sk\mathcal{T}$, is associated with lower GAS losses, although the evidence for FZL is weak. Concerning results for index returns in Table A. 4, the persistence of GARCH's ability to generate significantly lower QLs on *XU100* and *XUTUM* is a remarkable outcome.

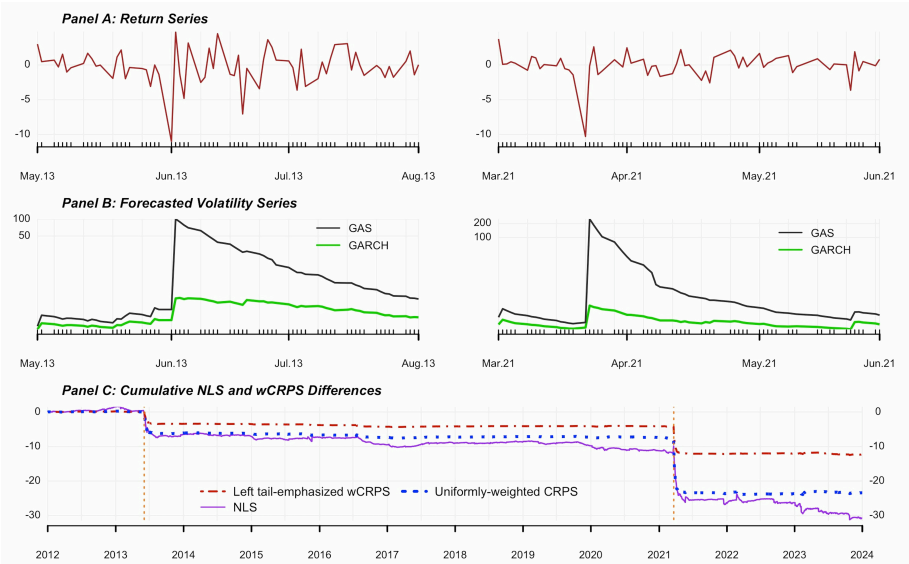
6. Further remarks and forecast analysis

As mentioned earlier, all forecasts are also performed using model refits in every 125th observation, i.e., semiannually. Although the results are not presented for the sake of tractability, they largely confirm the findings with quarterly refits. All significant DM tests with density and VaR forecasts were found to be robust in the sense that they preserved their significance regardless of the parameter update frequency. We noticed some changes in the sign with insignificant DM test statistics and minor overall differences in magnitudes, as expected. In addition, alternative values for the scaling parameter of GAS model are also considered in our experiments, i.e., $\gamma = 1$ and $\gamma = \frac{1}{2}$. The attempts with the former resulted in failures in convergence with $\mathcal{N}$ in both the BIST index and FX rate returns. While the other choice led to success in estimation with BIST index returns, it also caused failure in FX rate returns along with $\mathcal{N}$.

Regarding GARCH's overall superiority with $\mathcal{N}$and $sk\mathcal{N}$, a deeper investigation reveals that this end is closely related to the GAS model's handling of market shocks, which are more common in emerging economies than in developed economies. In particular, the density score difference series of the leading index *XU100*, subject to $\mathcal{N}$, exhibits two spectacular jumps that can be observed in the third panel of Figure 5. The first shock relates to turbulence in markets on June 3, 2013, in relation to Gezi Park events (Bloomberg HT, 2013), which is manifested with a 11.06% loss on *XU100* in the left chart of the first panel. The chart on the right depicts a recent shock that can be considered the market's reaction to worries resulting from the possible global effects of US inflation and bond rate hikes coupled with a recurring change in the Central Bank of Türkiye's governor (Gedik Yatırım Menkul Değerler A.Ş., 2021). The plots in the middle panel show the forecasted volatilities by both models in logarithmic scale, which allows clear visualization of both forecasts although the GAS volatilities are tremendously higher. These inflated volatilities propagate even after the following refit dates in both cases because of autoregression. Therefore, GAS forecasts result in higher loss values, which makes the model less attractive in the case of shocks. In general, this finding applies to other indexes, to a greater degree to FX rates, and similarly with $sk\mathcal{N}$.

Figure 5

A. XU100 *Return series (May - Aug. 2013 & Mar. - Jun. 2021)*
B. XU100 *volatility forecasts (May - Aug. 2013 & Mar. - Jun. 2021) with the Normal distribution*
C. Cumulative score differences for XU100 *with the Normal distribution*



However, GAS does not suffer such a drawback when the underlying law is $\mathcal{S}$ or $sk\mathcal{S}$. Figure A. 2 in the appendix justifies that both models produce comparable volatility forecasts with $\mathcal{S}$ subject to the same shocks. We observed temporary differences in score functions due to this comparability. Moreover, the way models treat both occurrences differently is remarkable because they offer alternative risk management tools to researchers and practitioners.

7. Conclusion

Aiming at comparing the forecast performances of the GARCH and GAS models for Turkish stock and FX markets, this research employs scoring rules to evaluate density forecasts and popular backlisting procedures for VaR forecasts. The results of the analyses demonstrate that GARCH outperforms GAS through the density forecasts for *XU100* and *XUTUM* with respect to NLS and wCRPS, while the same holds for *XUSIN* with NLS and wCRPS with an emphasis on the center of distribution conditional on $\mathcal{N}$and $sk\mathcal{N}$. *XBANK* provides evidence to this end only with $\mathcal{N}$. The same findings apply to FX rate series with respect to NLS and center-emphasized wCRPS under $\mathcal{N}$and $sk\mathcal{N}$ assumptions. More evidence is obtained for *USD* and *GBP* in this regard. Concerning the remaining two distributional assumptions $\mathcal{T}$ and $sk\mathcal{T}$, GAS model's better performance with indexes, except *XUSIN* with $\mathcal{T}$, is not substantiated by statistical tests. However, GAS's superiority is statistically justified for *USD* and partially justified for *GBP*.

Similar results were obtained with fewer assets when the DM tests with the two VaR loss functions were considered. While backlisting of VaR forecasts for BIST index returns with popular tests supports correctly specified model hypothesis at %5 risk level except *for XUSIN*, none is found to be adequate for FX returns. Coupled with the overall result at 1%, which turns out to be just the reverse across asset groups, the findings support those of Kahyaoğlu Bozkuş (2019) in that VaR forecasts are inconsistent with respect to risk level. However, this inconsistency concerns not only GAS but also GARCH.

The shocks of July 2013 and March 2021 are the predominant factors underlying the underperformance of the GAS model. Although these do not result in much difference when the conditional distribution is $\mathcal{T}$ or its skewed variant, they have substantial effects through sky-rocketed variance estimates with $\mathcal{N}$and $sk\mathcal{N}$. This finding agrees with that of Bekar (2019) in that Student's t- and variates are better for modeling tail behavior. Such shocks are more frequent in developing markets, like Türkiye, and this fact calls for a remedy for GAS models. On the other hand, observing different patterns with $\mathcal{T}$ in models' volatility estimates for *XU100* during the shocks suggests that models can complement each other. Therefore, the GAS model should be viewed as a complementary tool for risk managers, policymakers, and regulators. Moreover, POT and some other practices in EVT can model shocks, more precisely jumps in return processes in emerging markets, as suggested by Bekar (2019).

Finally, the forecasts are limited to the basic specifications of each model. It is possible to extend the analysis further by letting other parameters vary in GAS specification and by incorporating numerous advanced models within the GARCH family. Like the GARCH model, the GAS family is becoming richer with newly developed variates, specifically Beta-t-EGARCH, which is more convenient for financial series (Bekar, 2023). In this respect, the optimal combination of models, as in Bernardi and Catania (2016), can be considered a sensible strategy for improving VaR forecasts.

Peer Review Externally peer-reviewed.
Conflict of Interest The author has no conflict of interest to declare.
Grant Support The author declared that this study has received no financial support.

Author Details **Ali Ulvi Özgül (Asst. Prof. Dr.)**
[1] Ankara Bilim University, Faculty of Humanities and Social Sciences, Department of Business Administration, Ankara, Türkiye
0000-0002-1082-2652 ulvi.ozgul@ankarabilim.edu.tr

References

Amisano, G., & Giacomini, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, *25*(2), 177–190. https://doi.org/10.1198/073500106000000332

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, *59*(3), 817–858.

Andrews, D. W. K., & Monahan, J. C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, *60*(4), 953–966.

Ardia, D., Boudt, K., & Catania, L. (2019). Generalized autoregressive score models in R: The GAS package. *Journal of Statistical Software*, *88*(6). https://doi.org/10.18637/jss.v088.i06

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

Bekar, E. (2019). Les Previsions Statiques et Dynamiques des Valeurs a Risque (VaR) des Actions de Banque: Le Modele de Depassements de Seuil (POT) et Les Modeles de Score Autoregressifs Generalises (GAS). *The Journal of Academic Social Science*, (97), 142–169. https://doi.org/10.29228/ASOS.36752

Bekar, E. (2023). Döviz kuru volatilite modellemesinde Beta-t-EGARCH modelleri: Amerikan doları / Türk lirası döviz kuru üzerinden bir değerlendirme. *Sosyoekonomi*, *31*(55), 371–395. https://doi.org/10.17233/sosyoekonomi.2023.01.19

Bernardi, M., & Catania, L. (2016). Comparison of Value-at-Risk models using the MCS approach. *Computational Statistics*, *31*(2), 579–608. https://doi.org/10.1007/s00180-016-0646-6

Blasques, F., Hoogerkamp, M. H., Koopman, S. J., & van de Werve, I. (2021). Dynamic factor models with clustered loadings: Forecasting education flows using unemployment data. *International Journal of Forecasting*, *37*(4), 1426–1441. https://doi.org/10.1016/j.ijforecast.2021.01.026

Blazsek, S., Escribano, A., & Licht, Adrian. (2021). Identification of seasonal effects in impulse responses using score-driven multivariate location models. *Journal of Econometric Methods*, *10*(1), 53–66.

Bloomberg HT. (2013, June 3). BIST 100'den büyük düşüş. Retrieved October 27, 2024, from https://www.bloomberght.com/haberler/haber/1368565-bist-100den-buyuk-dusus

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Brooks, C. (2019). *Introductory econometrics for finance* (4th ed.). Cambridge: Cambridge University Press.

Caner, M., & Kilian, L. (2001). Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate. *Journal of International Money and Finance*, *20*(5), 639–657. https://doi.org/10.1016/S0261-5606(01)00011-0

Catania, L., & Nonejad, N. (2020). Density forecasts and the leverage effect: Evidence from observation and parameter-driven volatility models. *European Journal of Finance*, *26*(2–3), 100–118. https://doi.org/10.1080/1351847X.2019.1586744

Chevillon, G. (2007). Direct multi-step estimation and forecasting. *Journal of Economic Surveys*, *21*(4), 746–785. https://doi.org/10.1111/j.1467-6419.2007.00518.x

Christoffersen, P. F. (1998). Evaluating Interval Forecasts. *International Economic Review*, *39*(4), 841–862.

Christoffersen, P. F., & Pelletier, D. (2004). Backtesting Value-at-Risk: A duration-based approach. *Journal of Financial Econometrics*, *2*(1), 84–108. https://doi.org/10.1093/jjfinec/nbh004

Creal, D., Koopman, S. J., & Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, *28*(5), 777–795. https://doi.org/10.1002/jae.1279

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*(3), 134–144. https://doi.org/10.1198/073500102753410444

Dowd, K. (2005). *Measuring market risk* (2nd ed.). Hoboken, NJ: John Wiley & Sons Ltd.

Elliott, G., & Timmermann, A. (2016). *Economic Forecasting*. New Jersey: Princeton University Press.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, *50*(4), 987–1007. https://doi.org/10.2307/1912773

Engle, R. F., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, *22*(4), 367–381. https://doi.org/10.1198/073500104000000370

Fernández, C., & Steel, M. F. J. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, *93*(441), 359–371. https://doi.org/10.1080/01621459.1998.10474117

Fissler, T., & Ziegel, J. F. (2016). Higher order elicitability and Osband's principle. *The Annals of Statistics*, *44*(4), 1680–1707. https://doi.org/10.1214/16-AOS1439

Francq, C., & Zakoian, J.-M. (2019). *GARCH models: Structure, statistical inference, and financial applications* (2nd ed.). Hoboken, NJ: John Wiley & Sons Ltd.

Galanos, A. (2023). *rugarch: Univariate GARCH models*. R package version 1.5-1.

GAS papers. (n.d.). Retrieved January 28, 2025, from https://www.gasmodel.com/gaspapers.htm

Gedik Yatırım Menkul Değerler A.Ş. (2021). *Günlük Bülten*. Retrieved from https://gedik-cdn.foreks.com/yatirim/reports/files/000/005/912/original/Gunluk_Bulten_22.03.2021.pdf?1616392483

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, *102*(477), 359–378. https://doi.org/10.1198/016214506000001437

Gneiting, T., & Ranjan, R. (2011). Comparing density forecasts using threshold and quantile-weighted scoring rules. *Journal of Business and Economic Statistics*, *29*(3), 411–422. https://doi.org/10.1198/jbes.2010.08110

González-Rivera, G., Lee, T. H., & Mishra, S. (2004). Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of Forecasting*, *20*(4), 629–645. https://doi.org/10.1016/j.ijforecast.2003.10.003

Hadi, D. M., Karim, S., Naeem, M. A., & Lucey, B. M. (2023). Turkish Lira crisis and its impact on sector returns. *Finance Research Letters*, *52*. https://doi.org/10.1016/j.frl.2022.103479

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The Model Confidence Set. *Econometrica*, *79*(2), 453–497. https://doi.org/10.3982/ECTA5771

Harvey, A., Hurn, S., Palumbo, D., & Thiele, S. (2024). Modelling circular time series. *Journal of Econometrics*, *239*(1). https://doi.org/10.1016/j.jeconom.2023.02.016

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, *13*(2), 281–291. https://doi.org/10.1016/S0169-2070(96)00719-4

Hürriyet Newspaper. (2001, February 22). Kara Çarşamba. Retrieved July 13, 2024, from https://bigpara.hurriyet.com.tr/haberler/ekonomi-haberleri/kara-carsamba_ID358956/

Jorion, P. (2007). *Value at risk: The new benchmark for managing financial risk* (3rd ed.). New York: McGraw-Hill.

Kahyaoğlu Bozkuş, S. (2019). Generalized autoregressive score (Gas) model: An applied analysis on the BIST 100 index. In Ç. Başarır, B. Darıcı, & H. M. Ertuğrul (Eds.), *New trends in banking and finance*. Berlin: Peter Lang GmbH.

Kupiec, P. H. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *The Journal of Derivatives*, *3*(2), 73–84. https://doi.org/10.3905/jod.1995.407942

Laporta, A. G., Merlo, L., & Petrella, L. (2018). Selection of Value at Risk Models for Energy Commodities. *Energy Economics*, *74*, 628–643. https://doi.org/10.1016/j.eneco.2018.07.009

Marcellino, M., Stock, J. H., & Watson, M. W. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, *135*(1–2), 499–526. https://doi.org/10.1016/j.jeconom.2005.07.020

Owusu Junior, P., Tiwari, A. K., Tweneboah, G., & Asafo-Adjei, E. (2022). GAS and GARCH based value-at-risk modeling of precious metals. *Resources Policy*, *75*. https://doi.org/10.1016/j.resourpol.2021.102456

Pascual, L., Romo, J., & Ruiz, E. (2006). Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics and Data Analysis*, *50*(9), 2293–2312. https://doi.org/10.1016/j.csda.2004.12.008

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Retrieved from https://www.R-project.org/

Sucarrat, G. (2013). Betategarch: Simulation, estimation and forecasting of Beta-Skew-t-EGARCH models. *R Journal*, *5*(2), 137–147. https://doi.org/10.32614/rj-2013-034

Trucíos, C., & Taylor, J. W. (2023). A comparison of methods for forecasting value at risk and expected shortfall of cryptocurrencies. *Journal of Forecasting*, *42*(4), 989–1007. https://doi.org/10.1002/for.2929

Xu, Y., & Lien, D. (2022). Forecasting volatilities of oil and gas assets: A comparison of GAS, GARCH, and EGARCH models. *Journal of Forecasting*, *41*(2), 259–278. https://doi.org/10.1002/for.2812

Appendix

Figure A. 1

Evolution of price series

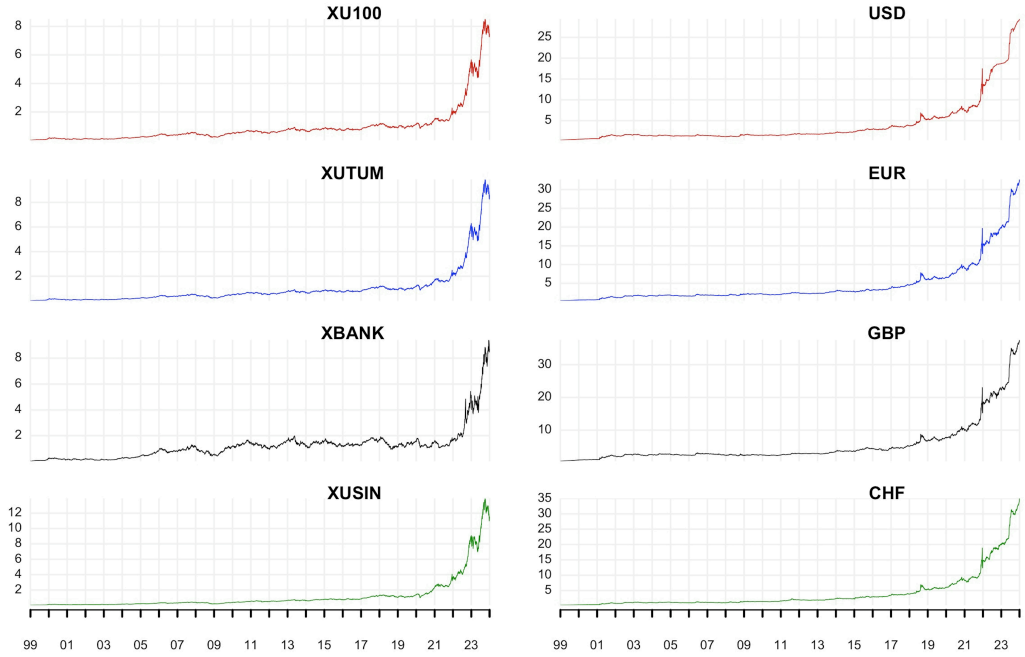


Figure A. 2

A. XU100 volatility forecasts (May - Aug. 2013 & Mar. - Jun. 2021) with Student-t

B. Cumulative score differences for XU100 with Student-t

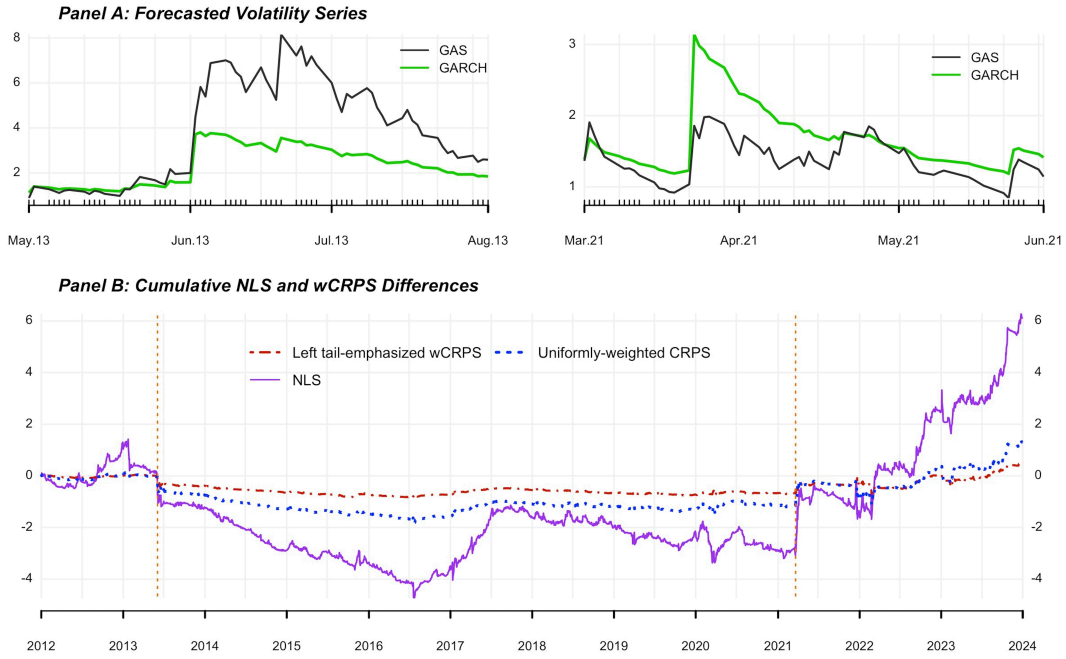


Table A. 1

Descriptive and diagnostic statistics of return series for in-sample and forecast periods

I. In-sample Period (04.01.1999-30.12.2011)								
Panel A: Descriptive Statistics								
Series	**N**	**Mean**	**Std. Dev.**	**Skewness**	**Kurtosis**	**Minimum**	**Median**	**Maximum**
XU100	3238	0.0903	2.5235	0.1006	8.5794	−19.9784	0.1014	17.7649
XUTUM		0.0915	2.4376	0.0591	8.9880	−19.6949	0.1142	17.6787
XBANK		0.0979	3.0026	0.1798	7.1009	−21.1711	0.0696	17.2617
XUSIN		0.0978	2.1909	−0.0853	10.7118	−18.0145	0.1480	18.0462
USD	3241	0.0557	1.1824	7.1849	214.0849	−12.5637	0.0030	33.4732
EUR		0.0588	1.1701	7.0610	195.3471	−12.2662	−0.0030	32.4513
GBP		0.0533	1.1719	7.6055	219.1162	−12.6949	0.0196	33.4941
CHF		0.0674	1.2685	5.5605	146.2326	−12.3507	0.0118	32.6420
Panel B: Diagnostic Statistics								
Series	**JB (× 10^{-5})**	**ADF**	**KPSS**	**Arch-LM**	**Q_{LB} (5)**	**Q_{LB} (10)**	**Q_{FZ} (5)**	**Q_{FZ} (10)**
XU100	0.04***	−12.46***	0.26	482.91***	6.11	30.60***	2.22	16.91*
XUTUM	0.05***	−12.91***	0.27	505.80***	6.62	33.46***	2.30	18.13*
XBANK	0.02***	−11.83***	0.23	323.63***	7.95	25.20***	3.97	14.93
XUSIN	0.08***	−12.38***	0.29	776.80***	7.58	39.89***	2.29	18.80**
USD	60.52***	−8.24***	0.80***	110.77***	98.37***	121.44***	4.75	9.06
EUR	50.29***	−9.28***	0.63**	120.66***	125.58***	143.18***	6.84	11.97
GBP	63.46***	−9.45***	0.88***	120.01***	141.74***	157.03***	9.12	11.78
CHF	27.91***	−9.27***	0.45*	116.13***	87.85***	105.00***	6.52	11.90
II. Forecast Period (02.01.2012-29.12.2023)								
Panel A: Descriptive Statistics								
Series	**N**	**Mean**	**Std. Dev.**	**Skewness**	**Kurtosis**	**Minimum**	**Median**	**Maximum**
XU100	3012	0.0889	1.5607	−0.6613	8.2066	−11.0633	0.1395	9.4219
XUTUM		0.0934	1.5010	−0.7840	8.6174	−11.0503	0.1552	9.0932
XBANK		0.0723	2.2529	−0.0995	6.0697	−11.8611	0.0089	9.4921
XUSIN		0.1054	1.4504	−0.8749	9.6532	−11.3998	0.1892	9.2309
USD	2992	0.0914	1.0732	−5.3754	213.5568	−29.3976	0.0414	14.7066
EUR		0.0865	1.0798	−5.3438	199.0808	−29.1531	0.0483	14.0194
GBP		0.0852	1.1130	−4.8385	177.8978	−29.1548	0.0544	14.5365
CHF		0.0954	1.1269	−4.1067	173.5646	−29.1281	0.0575	14.6878
Panel B: Diagnostic Statistics								
Series	**JB (× 10^{-5})**	**ADF**	**KPSS**	**Arch-LM**	**Q_{LB} (5)**	**Q_{LB} (10)**	**Q_{FZ} (5)**	**Q_{FZ} (10)**
XU100	0.04***	−29.72***	0.56**	179.33***	12.36**	17.45*	6.19	9.96
XUTUM	0.04***	−29.32***	0.63**	174.98***	15.52***	19.65**	7.52	10.56
XBANK	0.01***	−15.71***	0.38*	241.27***	6.59	18.86**	3.65	11.93
XUSIN	0.06***	−15.41***	0.60**	256.08***	14.78**	19.92**	6.19	9.77
USD	55.49***	−11.45***	0.62**	97.49***	87.48***	107.66***	6.10	15.02

EUR	48.14***	−11.93***	0.70**	94.56***	96.13***	111.76***	6.99	13.61
GBP	38.30***	−11.90***	0.62**	86.87***	96.15***	114.71***	7.56	17.01*
CHF	36.40***	−11.80***	0.71**	99.52***	108.54***	128.74***	8.36	16.31*

***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. N and Std. Dev. denote the number of observations and standard deviation, respectively. JB, ADF, KPSS, Arch-LM stand for Jarque-Bera, Augmented Dickey-Fuller, Kwiatkowski–Phillips–Schmidt–Shin, and Arch Lagrange Multiplier test statistics, respectively. Q_{LB} (i) and Q_{FZ} (i) denote the portmanteau (Q) test statistics of Ljung–Box and Francq-Zakoian at lag i, respectively.

Table A. 2

Mean negative loss scores (NLS) and weighted continuous ranked probability scores (wCRPS) of density forecasts on BIST index returns with GARCH and GAS models.

	NLS	**wCRPS with an emphasis on**				
		Uniform	**Center**	**Tails**	**Right tail**	**Left tail**
			Normal distribution			
XU100	1.8038	0.8146	0.1084	0.1285	0.4204	0.3942
	1.8140	**0.8223**	**0.1088**	**0.1338**	**0.4240**	**0.3983**
XUTUM	1.7646	0.7798	0.1075	0.1226	0.4047	0.3751
	1.7753	**0.7891**	**0.1079**	**0.1290**	**0.4092**	**0.3799**
XBANK	2.1691	1.1884	0.1260	0.2393	0.5960	0.5925
	2.1731	**1.1899**	**0.1261**	**0.2401**	**0.5966**	**0.5933**
XUSIN	1.7021	0.7366	0.1100	0.1319	0.3856	0.3510
	1.7249	**0.7919**	**0.1106**	**0.1837**	**0.4131**	**0.3788**
			Student-t distribution			
XU100	1.7581	0.8107	0.1079	0.1280	0.4182	0.3925
	1.7560	**0.8103**	**0.1078**	**0.1279**	**0.4179**	**0.3924**
XUTUM	1.7139	0.7757	0.1069	0.1220	0.4024	0.3733
	1.7126	**0.7753**	**0.1068**	**0.1220**	**0.4021**	**0.3732**
XBANK	2.1376	1.1855	0.1257	0.2387	0.5945	0.5909
	2.1376	**1.1853**	**0.1257**	**0.2388**	**0.5946**	**0.5908**
XUSIN	1.6355	0.7326	0.1094	0.1313	0.3834	0.3492
	1.6364	**0.7331**	**0.1094**	**0.1315**	**0.3833**	**0.3498**
			Skewed Normal distribution			
XU100	1.7972	0.8140	0.1083	0.1285	0.4201	0.3939
	1.8073	**0.8197**	**0.1087**	**0.1319**	**0.4227**	**0.3970**
XUTUM	1.7546	0.7791	0.1074	0.1225	0.4044	0.3747
	1.7651	**0.7852**	**0.1078**	**0.1263**	**0.4072**	**0.3780**
XBANK	2.1701	1.1886	0.1260	0.2394	0.5960	0.5925
	2.1897	**1.1927**	**0.1263**	**0.2416**	**0.5981**	**0.5947**
XUSIN	1.6811	0.7348	0.1097	0.1317	0.3849	0.3500
	1.7027	**0.7674**	**0.1104**	**0.1606**	**0.4002**	**0.3673**
			Skewed Student's t distribution			
	1.7560	0.8104	0.1078	0.1279	0.4181	0.3924

	NLS	wCRPS with an emphasis on				
		Uniform	Center	Tails	Right tail	Left tail
XU100	**1.7556**	**0.8102**	**0.1078**	**0.1279**	**0.4179**	**0.3923**
XUTUM	1.7106	0.7753	0.1069	0.1219	0.4022	0.3731
	1.7108	**0.7752**	**0.1068**	**0.1220**	**0.4020**	**0.3732**
XBANK	2.1380	1.1855	0.1257	0.2387	0.5946	0.5909
	2.1376	**1.1853**	**0.1257**	**0.2388**	**0.5945**	**0.5908**
XUSIN	1.6293	0.7316	0.1092	0.1312	0.3829	0.3487
	1.6334	**0.7329**	**0.1094**	**0.1315**	**0.3832**	**0.3497**

Figures in bold refer to GAS model forecasts.

Table A. 3

Mean negative loss scores (NLS) and weighted continuous ranked probability scores (wCRPS) of density forecasts on FX rate returns with GARCH and GAS models.

	NLS	wCRPS with an emphasis on				
		Uniform	Center	Tails	Right tail	Left tail
		Normal distribution				
USD	1.0013	0.4133	0.1039	0.1049	0.1968	0.2165
	1.1087	**0.4672**	**0.1089**	**0.1440**	**0.2246**	**0.2426**
EUR	1.1095	0.4335	0.1116	0.1059	0.2088	0.2247
	1.1611	**0.4737**	**0.1134**	**0.1407**	**0.2288**	**0.2450**
GBP	1.1773	0.4598	0.1170	0.1157	0.2230	0.2368
	1.2724	**0.4975**	**0.1205**	**0.1431**	**0.2420**	**0.2555**
CHF	1.1978	0.4597	0.1096	0.1107	0.2195	0.2402
	1.3220	**0.7203**	**0.1136**	**0.3585**	**0.3502**	**0.3701**
		Student-t distribution				
USD	0.8887	0.4099	0.1032	0.1036	0.1948	0.2151
	0.8346	**0.4040**	**0.1021**	**0.1009**	**0.1917**	**0.2123**
EUR	1.0362	0.4305	0.1110	0.1046	0.2067	0.2237
	1.0354	**0.4297**	**0.1109**	**0.1041**	**0.2064**	**0.2233**
GBP	1.1098	0.4582	0.1166	0.1153	0.2216	0.2366
	1.1028	**0.4554**	**0.1163**	**0.1135**	**0.2203**	**0.2351**
CHF	1.0878	0.4535	0.1087	0.1074	0.2158	0.2377
	1.0855	**0.4518**	**0.1085**	**0.1064**	**0.2148**	**0.2369**
		Skewed Normal distribution				
USD	1.0017	0.4136	0.1041	0.1048	0.1969	0.2167
	1.0781	**0.5021**	**0.1080**	**0.1818**	**0.2440**	**0.2581**
EUR	1.1079	0.4339	0.1117	0.1059	0.2089	0.2250
	1.1668	**0.4872**	**0.1132**	**0.1549**	**0.2364**	**0.2509**
GBP	1.1792	0.4600	0.1171	0.1157	0.2231	0.2369
	1.2554	**0.4980**	**0.1191**	**0.1476**	**0.2422**	**0.2558**
	1.1869	0.4597	0.1097	0.1104	0.2194	0.2402

	NLS	wCRPS with an emphasis on				
		Uniform	Center	Tails	Right tail	Left tail
CHF	**1.3077**	**0.7232**	**0.1135**	**0.3616**	**0.3589**	**0.3642**
		Skewed Student's t distribution				
USD	0.8888	0.4104	0.1034	0.1036	0.1951	0.2153
	0.8364	**0.4046**	**0.1023**	**0.1011**	**0.1919**	**0.2127**
EUR	1.0358	0.4306	0.1110	0.1045	0.2068	0.2239
	1.0361	**0.4299**	**0.1109**	**0.1042**	**0.2064**	**0.2234**
GBP	1.1096	0.4583	0.1167	0.1152	0.2217	0.2367
	1.1032	**0.4556**	**0.1163**	**0.1136**	**0.2203**	**0.2352**
CHF	1.0867	0.4537	0.1088	0.1074	0.2159	0.2378
	1.0856	**0.4520**	**0.1085**	**0.1065**	**0.2149**	**0.2371**

Figures in bold refer to GAS model forecasts.

Table A. 4

Comparative backtest results for GARCH and GAS Value-at-Risk (VaR) forecasts for BIST index returns (α=1%)

	UC	CC	A/E	DQ	VaRd	DM_{QL}	DM_{FZ}
			Normal distribution				
XU100	0.0000	0.0000	2.0916	0.0000	0.0631	−2.8123***	−0.9945
	0.0000	**0.0000**	**1.9920**	**0.0000**	**0.0647**		
XUTUM	0.0000	0.0000	2.1248	0.0000	0.0625	−2.5900***	−0.4909
	0.0000	**0.0000**	**2.0252**	**0.0000**	**0.0641**		
XBANK	0.0015	0.0000	1.6268	0.0000	0.0769	−1.6288	−1.4810
	0.0043	**0.0001**	**1.5604**	**0.0000**	**0.0778**		
XUSIN	0.0000	0.0000	2.7224	0.0000	0.0641	−2.0630**	−0.8617
	0.0000	**0.0000**	**2.6228**	**0.0000**	**0.0713**		
			Student-t distribution				
XU100	0.0401	0.0671	1.3944	0.0233	0.0611	−0.2314	0.4511
	0.0070	**0.0248**	**1.5272**	**0.0000**	**0.0613**		
XUTUM	0.0267	0.0772	1.4276	0.1125	0.0604	−0.8651	−0.8695
	0.0070	**0.0114**	**1.5272**	**0.0000**	**0.0610**		
XBANK	0.3836	0.0018	1.1620	0.0000	0.0753	−1.2548	−1.0050
	0.1655	**0.0002**	**1.2616**	**0.0000**	**0.0764**		
XUSIN	0.0000	0.0000	2.0584	0.0000	0.0595	−0.4524	−0.7720
	0.0000	**0.0000**	**2.1580**	**0.0000**	**0.0600**		
			Skewed Normal distribution				
XU100	0.0003	0.0014	1.7264	0.0000	0.0616	−3.0790***	−1.9955**
	0.0015	**0.0034**	**1.6268**	**0.0000**	**0.0633**		
XUTUM	0.0009	0.0039	1.6600	0.0001	0.0609	−3.0975***	−1.6834*
	0.0005	**0.0023**	**1.6932**	**0.0000**	**0.0626**		
XBANK	0.0026	0.0000	1.5936	0.0000	0.0767	−1.0822	−1.0474
	0.0009	**0.0000**	**1.6600**	**0.0000**	**0.0827**		

	UC	CC	A/E	DQ	VaRd	DM_{QL}	DM_{FZ}
XUSIN	0.0000	0.0000	2.1580	0.0000	0.0606	−2.4232**	−1.3033
	0.0000	**0.0000**	**2.2576**	**0.0000**	**0.0660**		
	Skewed Student's t distribution						
XU100	0.1655	0.2352	1.2616	0.0224	0.0602	−1.3356	−1.4439
	0.0174	**0.0542**	**1.4608**	**0.0009**	**0.0613**		
XUTUM	0.1655	0.2352	1.2616	0.1372	0.0595	−1.0369	−1.2973
	0.0174	**0.0542**	**1.4608**	**0.0000**	**0.0603**		
XBANK	0.6035	0.0172	1.0956	0.0010	0.0754	−1.2169	−1.0909
	0.0849	**0.0002**	**1.3280**	**0.0000**	**0.0765**		
XUSIN	0.0070	0.0248	1.5272	0.1617	0.0573	−1.2118	−1.6848*
	0.0001	**0.0004**	**1.7928**	**0.0000**	**0.0584**		

***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. Numbers in bold refer to the GAS model.
UC: Unconditional coverage test of Kupiec (1995) CC: Conditional coverage test of Christoffersen (1998)
A/E: Actual over-excused ratio DQ: Dynamic quantile test of Engle and Manganelli (2004)
VaRd: VaR duration test by Christoffersen and Pelletier (2004) QL: Quantile loss function of González-Rivera et al. (2004)
FZ: Fissler & Ziegel's (2016) loss function DM_L: Diebold-Mariano test with loss model L
Diebold-Mariano test statistics are based on HAC variances. Except for A/E, the first six columns contain the acceptance probabilities of the null hypothesis with the corresponding test.

Table A. 5

Comparative backtest results for GARCH and GAS Value-at-Risk (VaR) forecasts for FX returns (α=1%)

	UC	CC	A/E	DQ	VaRd	DM_{QL}	DM_{FZ}
	Normal distribution						
USD	0.8436	0.7088	1.0361	0.9796	0.0307	−2.0945**	−2.5763**
	0.5854	**0.4378**	**0.9024**	**0.7980**	**0.0377**		
EUR	0.3522	0.5254	0.8356	0.7805	0.0313	−1.3083	−2.8937***
	0.4613	**0.3649**	**0.8690**	**0.5006**	**0.0376**		
GBP	0.5777	0.5927	1.1029	0.8960	0.0343	−2.0645**	−5.1428***
	0.8651	**0.5584**	**0.9693**	**0.6224**	**0.0404**		
CHF	0.4613	0.6069	0.8690	0.8042	0.0329	−1.1215	−3.1820***
	0.3522	**0.2914**	**0.8356**	**0.2902**	**0.1204**		
	Student-t distribution						
USD	0.0833	0.1925	0.7019	0.8583	0.0313	2.5526**	1.7530*
	0.1850	**0.3477**	**0.7687**	**0.9128**	**0.0299**		
EUR	0.0833	0.1925	0.7019	0.4703	0.0319	0.1639	−0.3590
	0.5854	**0.4378**	**0.9024**	**0.6160**	**0.0318**		
GBP	0.1850	0.3477	0.7687	0.6470	0.0354	4.1451***	2.7054***
	0.2599	**0.2226**	**0.8021**	**0.3630**	**0.0338**		
CHF	0.0833	0.1925	0.7019	0.4722	0.0331	1.0690	−0.4296
	0.5854	**0.4378**	**0.9024**	**0.6033**	**0.0325**		
	Skewed Normal distribution						
	0.0364	0.0992	1.4037	0.3311	0.0312	−1.1946	−0.8155

	UC	CC	A/E	DQ	VaRd	DM_{QL}	DM_{FZ}
USD	**0.4631**	**0.5392**	**1.1364**	**0.7145**	**0.0428**		
EUR	0.3634	0.4841	1.1698	0.7836	0.0315	−1.2719	−2.3364**
	0.4631	**0.1323**	**1.1364**	**0.0002**	**0.0418**		
GBP	0.0241	0.0420	1.4372	0.0589	0.0345	−2.2409**	−2.5752**
	0.0539	**0.0493**	**1.3703**	**0.0107**	**0.0414**		
CHF	0.2790	0.4212	1.2032	0.7187	0.0329	−1.3934	−2.0441**
	0.7055	**0.6068**	**1.0695**	**0.6563**	**0.0706**		
	Skewed Student's t distribution						
USD	0.2096	0.3555	1.2366	0.7024	0.0310	1.5431	1.4650
	0.4631	**0.5392**	**1.1364**	**0.8449**	**0.0301**		
EUR	0.3522	0.2914	0.8356	0.4939	0.0316	−0.3808	−1.1306
	0.8651	**0.5584**	**0.9693**	**0.6952**	**0.0319**		
GBP	0.7055	0.6587	1.0695	0.8979	0.0351	3.5339***	2.4303**
	0.7214	**0.5043**	**0.9358**	**0.2690**	**0.0339**		
CHF	0.5854	0.4378	0.9024	0.6333	0.0327	0.3158	−1.1851
	0.7214	**0.5043**	**0.9358**	**0.6811**	**0.0325**		

***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. Numbers in bold refer to the GAS model.
UC: Unconditional coverage test of Kupiec (1995) CC: Conditional coverage test of Christoffersen (1998)
A/E: Actual over-excused ratio DQ: Dynamic quantile test of Engle and Manganelli (2004)
VaRd: VaR duration test by Christoffersen and Pelletier (2004) QL: Quantile loss function of González-Rivera et al. (2004)
FZ: Fissler & Ziegel's (2016) loss function DM_L: Diebold-Mariano test with loss model L
Diebold-Mariano test statistics are based on HAC variances. Except for A/E, the first six columns contain the acceptance probabilities of the null hypothesis with the corresponding test.