Exploring the Impact of Sample Size on GARCH and tGARCH Models

Candidate Number: 21730

September 13, 2024

Contents

1 Introduction

Financial risk assessment is a critical component of decision-making in the dynamic landscape of investment and portfolio management. The accurate estimation of volatility, a key parameter in risk modeling, is essential for anticipating potential market movements and making informed choices in the face of uncertainty. Among the plethora of volatility models, the General Autoregressive Conditional Heteroskedasticity (GARCH) model and its variant, the tGARCH model, have emerged as prominent tools in capturing the complex dynamics of financial returns.[1]

This paper delves into the nuanced interplay between sample size and the precision of GARCH and tGARCH models in the context of financial risk assessment. As financial markets continue to exhibit intricate patterns influenced by a myriad of factors, understanding the impact of sample size on the performance of these models becomes crucial for enhancing the robustness of risk management practices.

Our exploration commences with an overview of the theoretical background, shedding light on the conceptual foundations of GARCH models and their application in capturing volatility clustering—a characteristic inherent in financial returns. The introduction of the tGARCH model, with its accommodation of fat-tailed distributions through Student-t distributed residuals, adds a layer of complexity that warrants a comprehensive investigation.

The subsequent sections delve into the integration of Monte Carlo simulation as a pivotal technique within the GARCH and tGARCH frameworks. Monte Carlo simulation, by considering a diverse array of potential outcomes, enhances risk assessments, providing a comprehensive evaluation of a portfolio's downside risk. The consideration of computation and simulation horizons becomes paramount in capturing the temporal evolution of risk factors and gaining a nuanced understanding of uncertainties associated with portfolio returns.

Our analysis focuses on the influence of sample size on parameter estimation and Value-at-Risk (VaR) accuracy. Learning curves are employed to illustrate how the Mean Squared Error (MSE) changes with the increase in the simulation horizon, revealing insights into the precision and reliability of the models. The findings prompt a compelling avenue for further analysis—determining the minimum sample size required to maintain a specified error bound, offering practical implications for optimizing risk forecasting methods.

In essence, this paper contributes valuable insights to the field of financial risk assessment, emphasizing the intricate relationship between sample size, model performance, and the dynamic nature of financial markets. The findings aim to empower practitioners with a deeper understanding of volatility modeling, guiding them towards more informed decision-making and robust risk management practices in the ever-evolving landscape of financial markets.

2 Theoretic Background

2.1 GARCH

In our exploration of volatility models within the framework of this paper, our primary focus rests on the *General Autoregressive Conditional Heteroskedasticity* (GARCH) model introduced by Bollerslev (1986). This model stands as a member of the ARCH model family, initially conceptualized by R. Engle (1982). Operating as a parametric *conditional volatility* model, the GARCH model employs a sophisticated mechanism involving optimal exponential weighting of historical returns for the purpose of forecasting volatility.[4] [1]

The conceptual underpinning of this model is rooted in the strategic adjustment of weights between squared past returns and conditional volatility. This iterative process continues until the forecasted volatility at time $t-1$ becomes progressively more reliant on recent events, thereby enhancing its responsiveness to the latest market conditions.This model's basic idea suggests that when recent volatility sharply rises, there is a higher chance of it persisting at heightened levels. This dynamic weighting scheme proves instrumental in effectively capturing the intrinsic volatility clustering characteristic of financial returns, reflecting auto-correlation in the second moment.

To streamline our analytical framework, we only consider the singular lag GARCH(1,1) model, defined as follows:

$$\begin{aligned} Y_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \omega + \alpha Y_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \tag{2.1}$$

To emulate the characteristics of typical financial data, the GARCH model also addresses the returns itself with the first equation of (2.1). Given that standard financial returns often exhibit a minimal mean, hovering around zero, the model only considers *de-meaned* returns by setting $\mu = 0$. Another pivotal consideration the model captures is the inherent unpredictability of the market, manifested through random shocks or innovations in the returns. These unforeseen elements are effectively modeled by the standardized residuals ϵ_t, constituting a sequence of independent and identically distributed (IID) random variables with a mean of 0 and a variance of 1. [3]

The distribution of these residuals is a critical aspect that requires assumption, with the standard GARCH model presuming a normal distribution for the residuals, denoted as $\epsilon_t \sim N(0,1)$. It is noteworthy that altering the assumed distribution of residuals holds the potential to fundamentally reshape the nature of the GARCH model, giving rise to a multitude of variations. An alternative GARCH model employed in this paper is the tGARCH model, where residuals follow a Student's t distribution, denoted as $\epsilon_t \sim t_{(\nu)}$. This choice aligns with the intuitive recognition of the *non-normality* often observed in financial returns, offering a more fitting representation of the underlying data dynamics.

Concerning the variance equation, it adopts an ARMA model structure, comprising two distinct components. The first component is an *Autoregressive (AR)* component, much like the ARCH model, aimed at capturing the autocorrelation present in the squared returns, denoted as αY_{t-1}^2. Complementing this, the model introduces a *Moving Average (MA)* component, represented by $\beta \sigma_{t-1}^2$, offering enhanced flexibility in the lag structure. The estimation of the various unknown parameters ω, α, β is achieved through Maximum Likelihood, a characteristic made feasible by the distributional assumption underlying the standardized residuals. This distributional assumption is a crucial aspect enabling the modeling process. [1] [7]

An essential feature inherited from the ARCH model is its ability to maintain the *unconditional volatility* σ well-defined and constant. This characteristic underscores the GARCH model's ability to delineate a consistent level of volatility, unaffected by external factors, forming a foundational element in understanding the model's dynamics.

$$\sigma^2 = \frac{\omega}{1 - \alpha - \beta}$$

This formula gives rise to two important restrictions placed on the parameters:

- *positive volatility:* $\omega, \alpha, \beta > 0$
  which ensures the necessary positivity of variance
- *covariance stationarity:* $\alpha + \beta < 1$
  which ensures the unconditional volatility remains defined (unlike in the EWMA model)

In summary, our analysis will predominantly rely on the GARCH(1,1) model and its counterpart, the tGARCH model. The selection of these models is grounded in their adeptness at capturing the rapid adjustments to more recent events. This responsiveness proves crucial in our examination of market dynamics, allowing us to effectively model and interpret the impact of timely information on volatility.

2.2 VaR

The predominant risk measure widely employed in industry is the *Value-at-Risk (VaR)*. VaR is characterized as the potential loss on a portfolio, subject to a specified probability $p \in (0,1)$, where the losses are equal to or exceed the VaR. VaR is essentially a quantile on the distribution of profit and loss(P/L). Let $Q = P_t - P_{t-1}$ denote a random variable representing such distribution P/L of a given portfolio, then VaR is mathematically defined as follows:

$$\begin{aligned} p &= \Pr[Q \leq -\mathrm{VaR}(p)] \\ &= \int_{-\infty}^{-\mathrm{VaR}(p)} f_q(x)\, dx \end{aligned} \tag{2.2}$$

Its practical utility is highly valued in the industry due to its ability to reduce risk assessment into a singular, quantifiable numerical measure. This makes it a widely embraced metric, employed both by policy makers for regulatory compliance and by financial institutions seeking to enhance the efficiency of their capital allocation. Notably, the most prevalent quantiles utilized are $VaR_{1\%}$ and $VaR_{5\%}$. [3]

The implementation of VaR encompasses various methods, with the two most prevalent categories being parametric and non-parametric approaches. Additionally, semi-parametric methods have emerged, including *Extreme Value Theory (EVT)* by Embrecht et al. (1997), *CAViaR* proposed by Engle and Manganelli (1999), and *Quasi-Maximum Likelihood GARCH* developed by Bollerslev and Woolridge (1992). These semi-parametric techniques leverage the flexibility of non-parametric methods, which lack strong distributional assumptions, while also benefiting from the efficiency of parametric methods that assume a distribution for the underlying data. [5] [2]

We will now elucidate the derivation of VaR based on the parametric volatility models previously mentioned. Starting with the definition of VaR represented by equation (2.2), we proceed to derive VaR for continuously compounded returns, a choice adopted for the purposes of this paper:

Let $Y_t = \log P_t - \log P_{t-1}$ denote continuously compounded returns. In the derivation of VaR, we initiate by defining the probability p as the likelihood that the returns are less than or equal to $-\mathrm{VaR}(p)$:

$$\begin{aligned} p &= \Pr[P_t - P_{t-1} \leq -\mathrm{VaR}(p)] \\ &= \Pr[P_{t-1}(e^{Y_t} - 1) \leq -\mathrm{VaR}(p)] \\ &= \Pr\left[\frac{Y_t}{\sigma} \leq \log(-\frac{VaR_t(p)}{P_{t-1}} + 1)\frac{1}{\sigma}\right] \end{aligned} \tag{2.3}$$

where $-\frac{VaR_t(p)}{P_{t-1}} \leq 1$. Representing the distribution of standardized returns $\left(\frac{Y_t}{\sigma}\right)$ by $F_y(\cdot)$ and its inverse distribution by $F_y^{-1}(p)$, we obtain:

$$\mathrm{VaR}(p) = -\left(\exp\left(F_y^{-1}(p)\sigma\right) - 1\right) P_{t-1}$$

Let $\vartheta = \mathit{Portfolio\,value}$ and assume normality of returns. If $p = 0.05$, we get

$$\mathrm{VaR} = -\Phi^{-1}(0.05) = 1.64 \Rightarrow \mathrm{VaR}_{5\%} = \sigma 1.64\vartheta$$

If we assume a Student-t distribution of returns with ν degrees of freedom, considering the fat-tailed nature of returns, we need to rescale the volatility estimate. The degrees of freedom parameter in the Student-t distribution, denoted as ν, affects the tail behavior of the distribution. When ν is low, the tails are fatter, and as $\nu \to \infty$, the distribution approaches a normal distribution. The variance of a Student-t distribution parameterized by ν is given by:

$$\frac{\nu}{\nu - 2} \tag{2.4}$$

In the estimation of VaR, the variance implied by ν plays a crucial role. However, when ν is less than 2, the variance becomes undefined, presenting challenges in practical applications. To overcome this issue, rescaling is necessary, introducing a scaling factor denoted by (2.4). This re-scaling is essential to ensure the stability and meaningful interpretation of the variance estimate in VaR calculations, thereby enhancing the accuracy and reliability of risk assessments. [3]

2.3 Monte Carlo Simulation

In the realm of Value-at-Risk (VaR) estimation, particularly when employing GARCH and tGARCH models, the integration of Monte Carlo simulation emerges as a pivotal technique. This simulation methodology extends the capabilities of parametric volatility models, providing a computationally intensive approach to anticipate potential future scenarios for portfolio returns.

In the initial phase of the simulation, the last volatility estimate ($\hat{\sigma_t}$) derived from the GARCH or tGARCH model serves as a foundational parameter. Subsequently, the model's parameter vector facilitates the forecasting of volatility for the upcoming period ($t + 1$). This forecasted volatility becomes a cornerstone for the subsequent steps in the simulation process.

The subsequent step involves the generation of future returns, leveraging the forecasted volatility. The simulation incorporates the assumed distribution - whether normal or Student-t — to generate standardized returns. This step encapsulates the dynamic nature of volatility and the potential autocorrelation in returns inherent in GARCH models.

Upon obtaining the simulated returns, the portfolio's value is calculated for each scenario. This entails compounding continuously compounded returns over time, aligning with the chosen time horizon for risk assessment. The generated portfolio values are then sorted, enabling the identification of quantiles that correspond to specific confidence levels, such as $VaR_{1\%}$ and $VaR_{5\%}$.

In instances where the assumption of normality may be unrealistic, the tGARCH model, featuring Student-t distributed residuals, provides a more realistic framework. The degrees of freedom parameter (ν) in the Student-t distribution accommodates the fat-tailed nature of financial returns, offering flexibility in capturing extreme events.

Monte Carlo simulation, within the context of GARCH and tGARCH models, enhances risk assessments by considering a diverse array of potential outcomes. By accounting for both volatility dynamics and non-normality, this simulation methodology provides a comprehensive evaluation of a portfolio's downside risk. These considerations contribute significantly to robust risk management practices in financial decision-making.

In the context of computation and simulation horizons, the estimation of Value-at-Risk (VaR) using Monte Carlo simulations with the models introduces a temporal dimension crucial for understanding risk dynamics. The computation horizon refers to the specific time period over which the simulations are conducted, encompassing the projection of future portfolio returns. This horizon is intricately tied to the investment strategy and risk management objectives, influencing decisions related to portfolio construction and asset allocation.

Simultaneously, the simulation horizon extends beyond the immediate timeframe of risk estimation, encompassing a broader temporal scope. It involves the generation of multiple scenarios and potential outcomes, allowing for a comprehensive evaluation of the portfolio's resilience under various market conditions. The choice of simulation horizon is pivotal, influencing the precision and reliability of risk metrics. A longer simulation horizon captures a more extensive range of potential market movements, aiding in the identification of tail risks and enhancing the robustness of risk assessments.

Therefore, the careful consideration of both computation and simulation horizons in the context of GARCH and tGARCH models is paramount. These horizons collectively contribute to the effectiveness of Monte Carlo simulations in capturing the temporal evolution of risk factors and providing a nuanced understanding of the uncertainties associated with portfolio returns. As financial markets exhibit dynamic and evolving characteristics, aligning computation and simulation horizons appropriately enhances the ability to make informed decisions and implement effective risk management strategies in the ever-changing landscape of investment environments.

Subsequent sections of this paper delve into the practical implementation of Monte Carlo simulations, explore empirical results, and discuss the implications for risk measurement and management in financial but also computational contexts. [8] [6]

3 Method used

We initiate our study by collecting daily TESLA stock returns, which were subsequently transformed into log returns. This curated dataset laid the groundwork for our comprehensive modeling and simulation procedures. In order to impart a degree of realism, we established a portfolio value of $\vartheta = \$1000$, providing a tangible context for our analyses.

The implementation of our chosen models was facilitated through the *rugarch* package in R. The specification of both GARCH and tGARCH models was accomplished using the *ugarchspec* function, allowing us to define the variance model, mean model (with the return series de-meaned), and distributional assumptions. The models were then fitted to the observed financial time series of TESLA, a crucial step that yielded estimated parameters (ω, α, β and ν for tGARCH).

Transitioning into the simulation phase, we meticulously conducted a comparative analysis across various simulation horizons. The primary objective was to dissect the properties that manifest with an increase of the simulation horizon. Deliberately selecting four distinct sizes (T), we aimed to capture a spectrum of scenarios, each contributing unique insights into the dynamic interplay of the model.

$$\begin{aligned}
T_1 &= 250\,(1 \text{ year}),\\
T_2 &= 1000\,(4 \text{ years}),\\
T_3 &= 2000\,(8 \text{ years}),\\
T_4 &= 5000\,(20 \text{ years}).
\end{aligned}$$

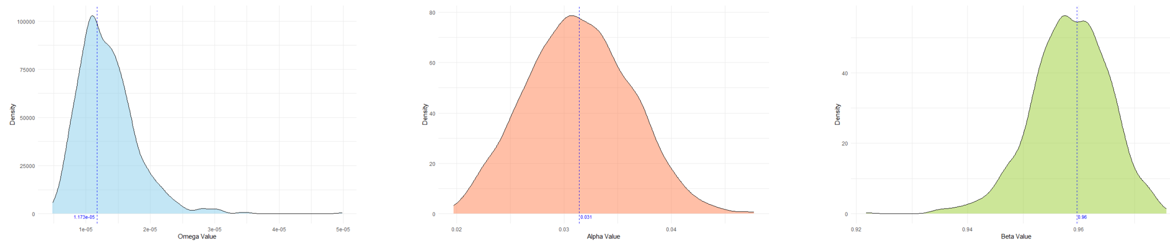
In adherence to the conventional practice of considering one year as comprising 250 observations, mirroring the standard number of trading days in a year, we embarked on an extensive simulation exercise. We executed $S = 1000$ simulations for each designated sample size, thereby generating synthetic data in accordance with the specified volatility models.

Within the iterative process of each simulation, the model was systematically fitted to the freshly generated synthetic data, and the resultant parameters were recorded. This systematic approach allowed us to scrutinize the dynamic evolution of the model's parameters as the sample sizes increased.

To offer a concrete example, we focused on the distributional properties of the three GARCH parameters— ω, α, β for sample size T_4, depicted in Figure 1.

The simulated parameters align well with the true parameter values, as seen in the distributions, which were initially calculated through the model fitting process using the stock returns. Visual inspection reveal that the disparities between the population values and the mean values of the simulated parameters are minimal.

A parallel occurrence is observed in the case of the tGARCH model. In this context, the package



(a) Distribution of ω (b) Distribution of α (c) Distribution of β

Figure 1: Simulated Parameter Distributions

provides estimations for the ν parameter, thereby indicating the most suitable degrees of freedom for optimal fitting within the Student-t distribution. This additional parameter serves to enhance the model's flexibility by accounting for the inherent non-normality in financial returns, further substantiating the effectiveness of the simulation approach in capturing and replicating key characteristics of the original data.

In the next section we will illustrate a through analysis of VaR estimates of the two models in question.

4 Analysis

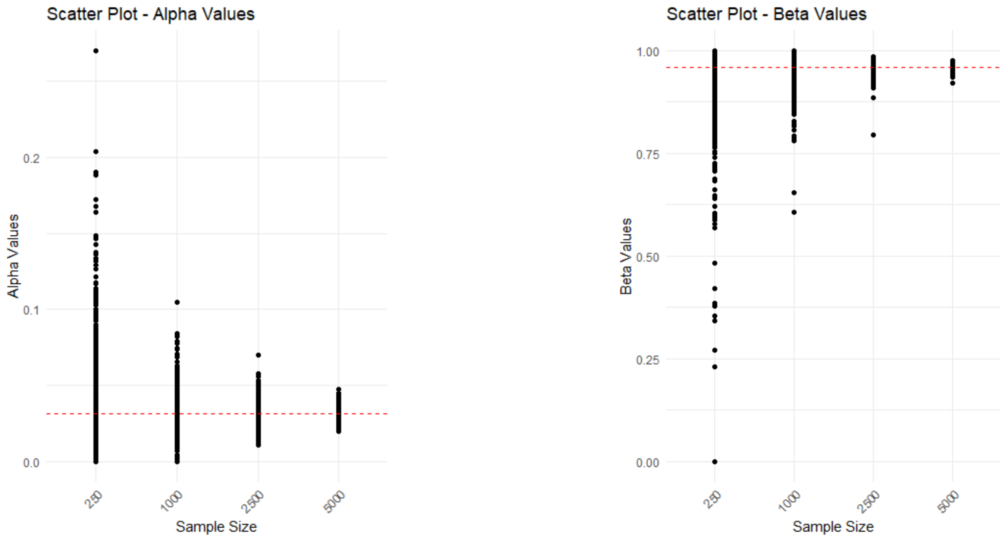
Transitioning into the analysis phase, the primary objective of this study is to scrutinize the influence of sample size on the precision of the GARCH and tGARCH models. To achieve this, we engage in the simulation of the respective parameter matrices $\hat{\theta_s}$, systematically varying the sample sizes and meticulously examining the behaviors of these models as the sample sizes progressively expand. This analytical approach allows us to gain insights into how the precision and reliability of the models respond to changes in the amount of available data.

Our initial exploration delves into the effects of the sample sizes on parameter estimation. Intuitively, one would expect that as the sample size expands, the simulation estimates of the parameters should exhibit improved accuracy, approaching the true values. This expectation is visually represented through a sequence of scatter plots presented in Figure 2, focusing initially on the parameters only of the GARCH(1,1) model.

In the plots, the x-axis denotes the four distinct sample sizes selected, while the y-axis represents the corresponding simulated parameter values. We specifically opted for α and β values due to their efficacy in illustrating the observed effects. The red dotted line serves as a reference for the true value of each parameter. A discernible pattern emerges as the spread of the simulated data points diminishes and converges towards the red line with an escalation in the sample size, notably evident in the comparison between T_1, T_3, and T_4. Interestingly, T_3 and T_4 exhibit similar spreads, distinguishing them from the pronounced disparity observed in T_1. This suggests, again, that higher sample sizes narrow down the variance of the values, therefore making the estimation more precise.

For a more detailed analysis, we narrow our focus to two specific scatter plots concerning the β parameter—one corresponding to T_2 and the other to T_4, in Figure 3. These detailed examinations further underscore the impact of sample size on the precision of parameter estimation.

A noticeable disparity in the spread of data points between the two sample sizes becomes apparent upon examining the two graphs. In the scatter plot on the right, the data points are confined within the range of 0.93 to 0.99, while in the plot on the left, they span between 0.8 and 1.0, as seen in Table 1. The red dot, denoting the true value, provides a reference point. This discrepancy underscores a reduction in the confidence interval as the sample size increases, affirming the influence of sample size



(a) Scatterplot of α (b) Scatterplot of β

Figure 2: GARCH parameter Scatter-Plots

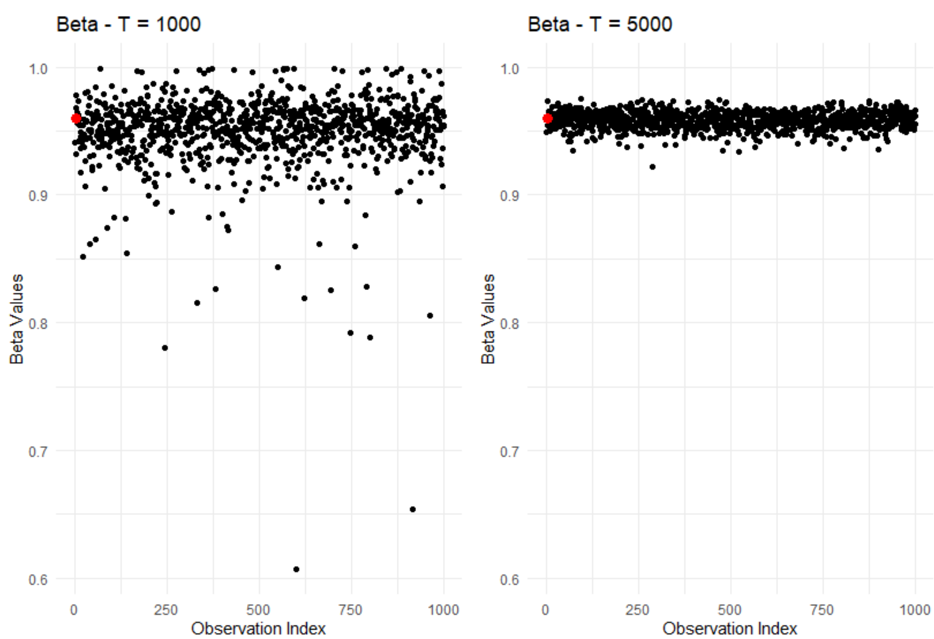
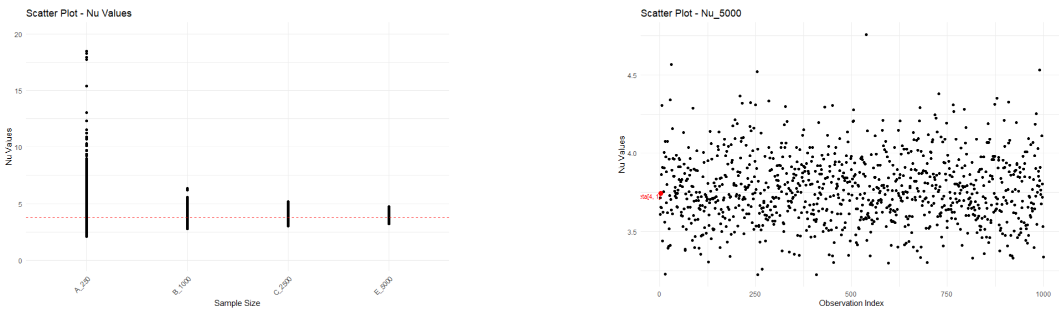


Figure 3: Scatterplot of β zoomed in

on the precision of parameter estimation.

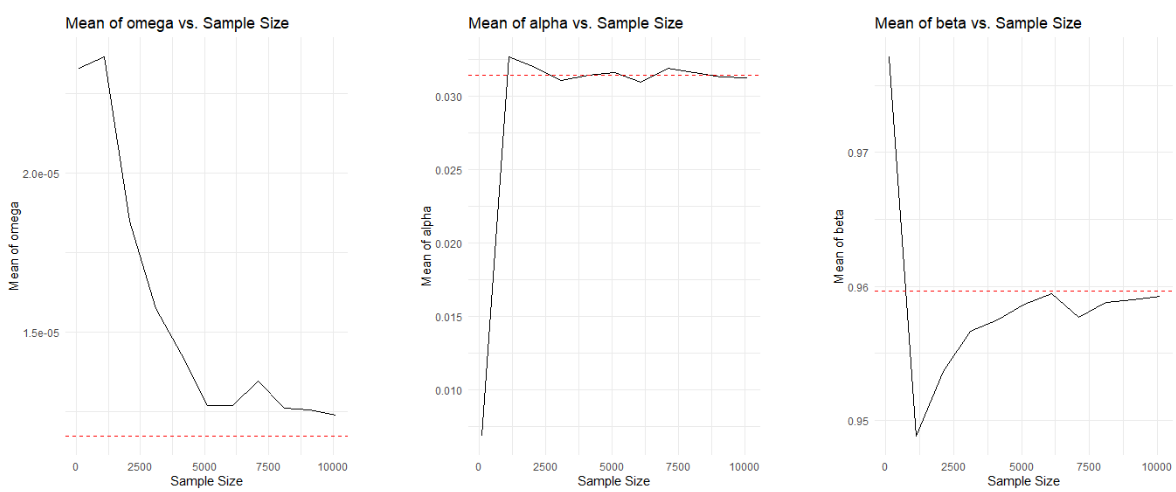
We find similar results when analysing the tGARCH parameters as we can see from the scatter-plots of the ν parameter in Figure 4.



(a) Scatterplot of ν (b) Scatterplot of ν for T_4

Figure 4: Scatterplots for ν

The preamble suggests that as we allow T_i to approach infinity, the mean of the estimated parameter will converge to the population value. This trend is evident in the subsequent graphs in Figure 4, where we examine the behavior of the compounded mean for each parameter as it approaches a high sample size, set to $T = 10000$ to emulate infinity. In consideration of computational constraints, we streamlined our empirical analysis by reducing the number of simulations to $S = 100$ and incrementally increasing the simulation sizes by 1000 observations on each iteration. This adjustment yielded meaningful results while staying within the confines of required computing resources.



(a) Convergence of ω (b) Convergence of α (c) Convergence of β

Figure 5: Convergence of Parameters

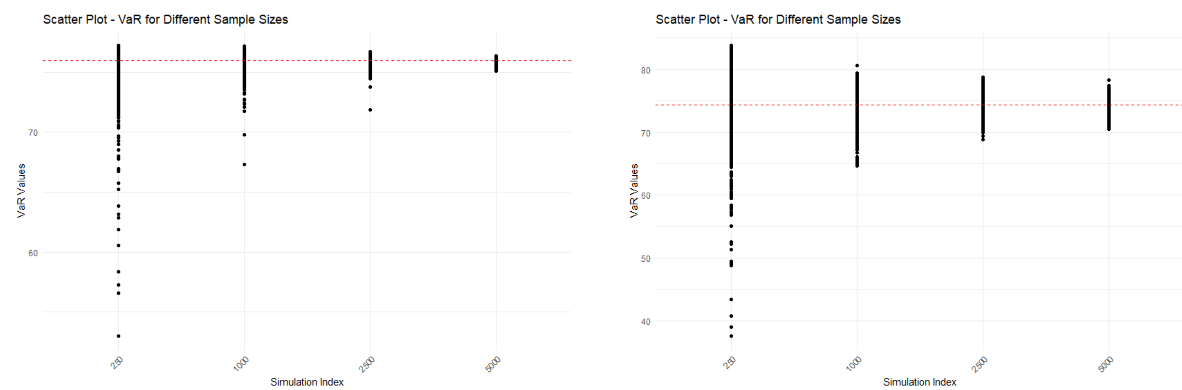
4.1 VaR

We proceed with the subsequent phase of our analysis, focusing on the estimation of Value-at-Risk (VaR) using the simulated parameters. Once again, our objective is to discern how sample size influences the precision of VaR estimates. Gaining a profound understanding of risk forecasting methods

Parameter		Sample Size = 250			Sample Size = 5000			True Values
		Mean	SD	Max/Min	Mean	SD	Max/Min	
GARCH	ω	0.00003	0.00009	0.00164/0.00	0.00005	0.00	0.00005/0.00	0.00001
	α	0.02243	0.03402	0.27015/0.00	0.03138	0.00478	0.04770/0.01972	0.03143
	β	0.95335	0.09207	0.999/0.00	0.95840	0.00693	0.97579/0.92177	0.95965
tGARCH	ω	0.00006	0.00013	0.00206/0.00	0.00001	0.00	0.00003/0.00	0.00002
	α	0.04448	0.04413	0.25905/0.00	0.04627	0.00628	0.06858/0.02612	0.04645
	β	0.92668	0.09597	0.999/0.00	0.94533	0.00652	0.9654/0.92000	0.94589
	ν	4.75	3.78	99.89/2.1	3.77	0.21	4.75/3.22	3.74349

Table 1: Mean and Standard Deviation Values for GARCH and tGARCH Parameters with True Values

and optimizing them in all aspects is crucial. For both the GARCH(1,1) and tGARCH models, we calculate the σ for each simulation and subsequently derive the corresponding VaR5%. It is noteworthy that the squared returns factor $Yt-1$ employed represents the last observed return of the empirical time series, while the $t-1$ volatility factor is the last observed conditional variance obtained from the respective fitted model. In the tGARCH model, VaR values are further adjusted using the re-scaling factor defined in equation (2.4). Figure 6 illustrates scatter-plots of the VaR estimates for both models.



(a) Scatterplot GARCH VaR

(b) Scatterplot tGARCH VaR

Figure 6: Caption for the entire figure

Upon initial inspection, distinctions between the two models are apparent. The estimates from the GARCH model display a more skewed distribution, whereas those from the tGARCH model are more evenly dispersed. This translates to the GARCH model having a tendency to underestimate VaR, as it has trouble replicating the non-normality of returns. Interestingly however, as the sample size increases, the confidence intervals of the GARCH estimates appear to become more precise compared to the tGARCH estimates. This observation becomes particularly evident when zooming into the respective scatter plots for T_4.

The discernibly narrower confidence bounds for GARCH, as opposed to tGARCH, lead us to consider the possibility that tGARCH may require larger sample sizes to achieve comparable levels of precision to GARCH.

We now quantify the precision of the models by examining the Mean Squared Errors (MSE) of the respective VaR estimates, defined as follows:

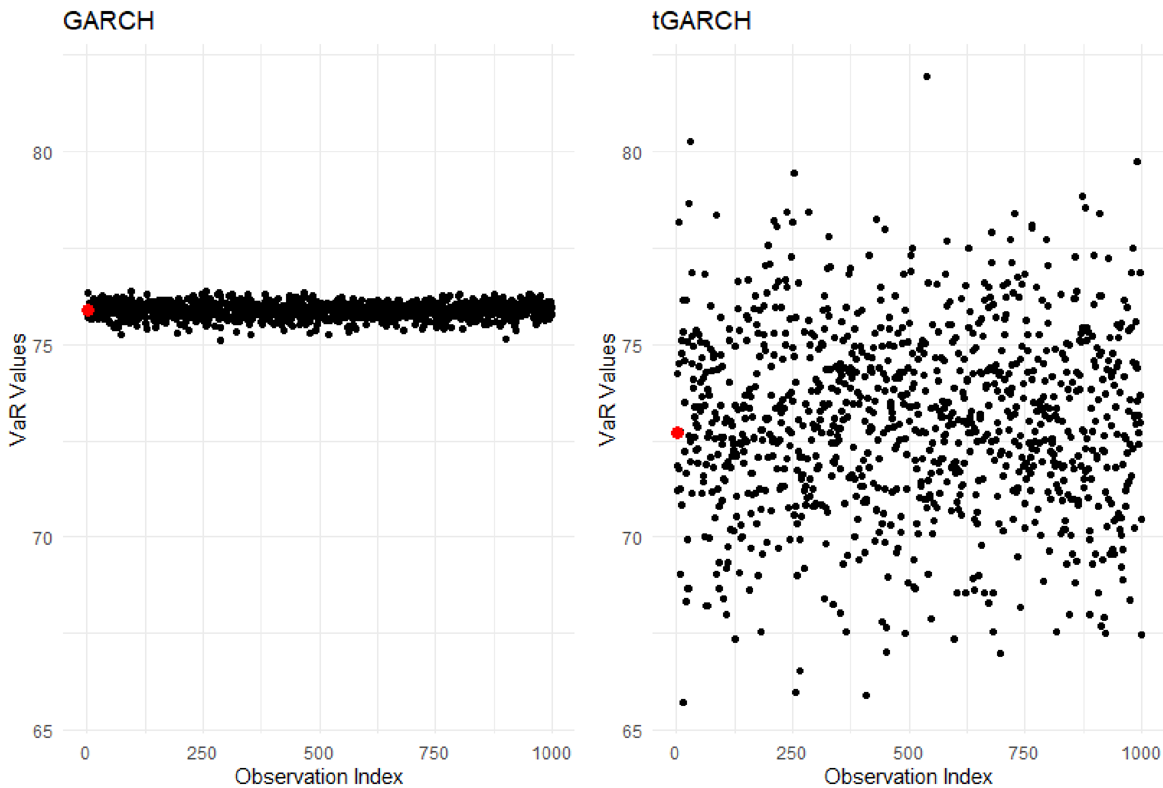


Figure 7: Scatterplot T = 5000, GARCH vs tGARCH

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(VaR_{\text{simulated},i} - VaR_{\text{true}})^2$$

This well-known statistical metric provides insights into how far the simulated VaR deviates from the population VaR.

Armed with the MSE metric, we aim to comprehend how the error, i.e., the model's precision, evolves across different sample sizes. In Figure 8, we present learning curves for both GARCH and tGARCH, illustrating how the Mean Squared Error (MSE) changes as the amount of training data increases, i.e., the simulation horizon. This plot reaffirms the earlier analysis. In both models, there is a noticeable decrease in the estimation error, indicating an improvement in performance as the simulation horizon increases. Notably, the tGARCH curve consistently remains higher than the GARCH curve throughout the entire graph. This observation suggests that tGARCH tends to be a less precise model in this context. This is most likely caused by the ν value of the tGARCH. As we can see in Table 1, the ν parameter has the highest standard deviation out of all parameters, which, when compounded through the VaR calculations, causes increased imprecision in the model.

In the realm of financial modeling, the consideration of computational efficiency holds particular significance. Beyond assessing the precision of models through metrics like Mean Squared ErThe observation of a plateau in precision on the learning curves prompts an intriguing avenue for further analysis—determining the minimum sample size required to maintain a specified error bound. This analysis would address the practical question of resource optimization and efficiency in financial modeling.

Estimating the minimum sample size needed for a given level of precision is a critical aspect of model development. It provides insights into the trade-off between computational resources and the de-

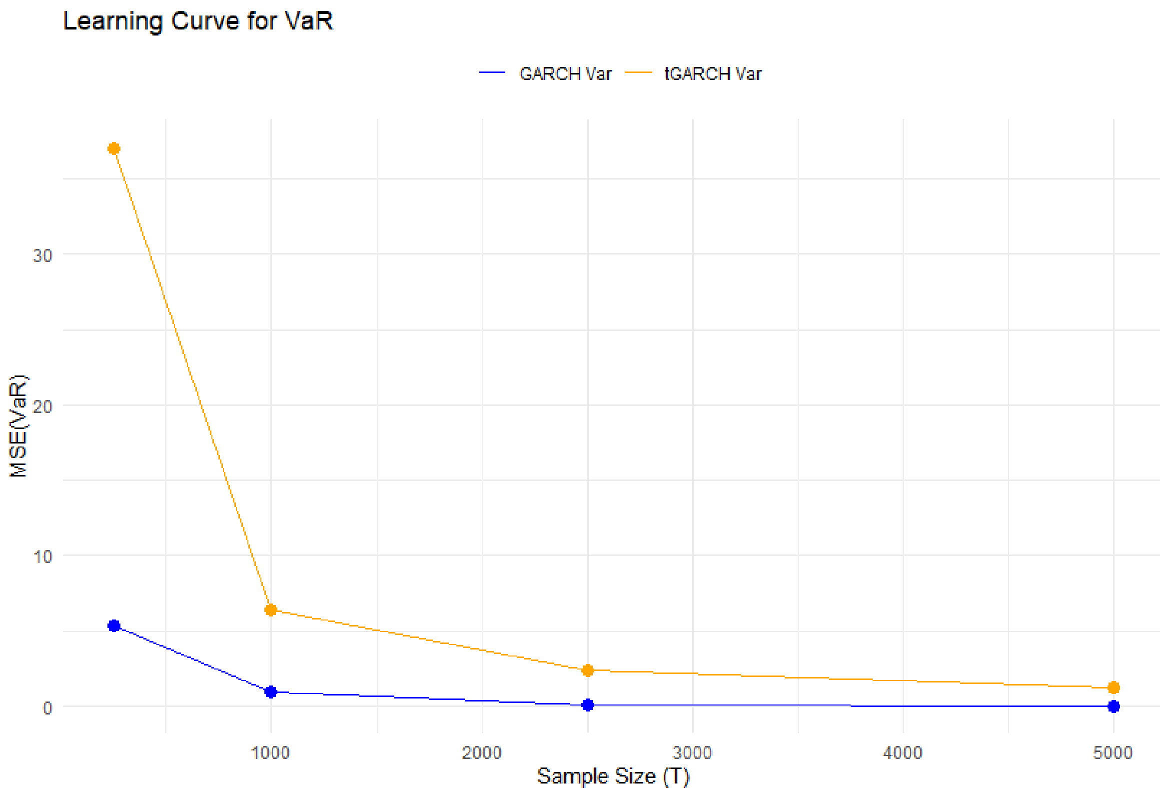


Figure 8: Learning Curves VaR

sired level of accuracy, allowing practitioners to streamline data collection efforts and computational expenses. By establishing a threshold for sample size that ensures the model's performance stays within an acceptable error margin, financial professionals can make informed decisions about resource allocation and model deployment.

This analysis becomes particularly valuable in scenarios where data acquisition is costly or time-consuming. Identifying the minimum sample size required for reliable predictions enables a more targeted and resource-efficient approach to model training and validation. It aligns with the overarching goal of achieving optimal predictive performance while minimizing computational overhead.

In conclusion, investigating the minimum sample size necessary to meet a predefined precision threshold represents a logical next step in refining the applicability of GARCH and tGARCH models in financial contexts. This analysis would contribute valuable insights to practitioners seeking to strike an optimal balance between model accuracy and computational efficiency in their decision-making processes.

5 Conclusion

In conclusion, this paper has provided a detailed exploration of volatility modeling within the realm of financial risk assessment. Focusing primarily on the GARCH and tGARCH model, we have delved into their theoretical underpinnings and practical applications in Value-at-Risk (VaR) estimation. Our study has underscored the significance of Monte Carlo simulation as a pivotal technique within the context of GARCH and tGARCH models. We scrutinized the influence of sample size on the precision of GARCH and tGARCH models, examining parameter estimation and VaR accuracy. Learning curves demonstrated a decrease in estimation error with an increase in the simulation horizon, indicating improved model performance. Notably, tGARCH exhibited higher Mean Squared Error (MSE) values, suggesting a potential trade-off between precision and model complexity.
The key discovery in this study centers around Figure 6. Our conclusion emphasizes that when dealing

with a data generating process characterized by fat-tailed distributions (tGARCH), achieving more accurate asymptotic approximations necessitates a larger dataset.

Furthermore, our study has prompted a compelling avenue for further analysis—determining the minimum sample size required to maintain a specified error bound. This consideration holds practical significance in optimizing risk forecasting methods and aligns with the broader goal of achieving optimal predictive performance while minimizing computational overhead.

References

[1] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

[2] Tim Bollerslev and Jeffrey M Wooldridge. Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric reviews*, 11(2):143–172, 1992.

[3] Jon Danielsson. *Financial risk forecasting: The theory and practice of forecasting market risk with implementation in R and Matlab*. John Wiley & Sons, 2011.

[4] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987–1007, 1982.

[5] Robert F Engle and Simone Manganelli. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of business & economic statistics*, 22(4):367–381, 2004.

[6] L Jeff Hong, Zhaolin Hu, and Guangwu Liu. Monte carlo methods for value-at-risk and conditional value-at-risk: a review. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 24(4):1–37, 2014.

[7] Simone Manganelli and Robert F Engle. Value at risk models in finance. 2001.

[8] Jones Osei, Peter Sarpong, and Samuel Amoako. Comparing historical simulation and monte carlo simulation in calculating var. *Quant. Financ*, 3:22–35, 2018.