

https://dspace.org/

Simulation Structure for Selecting an Optimal Error Distribution Through the GAS Model

Richard Samuel; Charles Chimedza; Caston Sigauke

https://hdl.handle.net/10539/41845
SCOPUS

Downloaded from a DSpace repository. https://dspace.org/

STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **12**, July 2024, pp 1123–1148.
Published online in International Academic Press (www.IAPress.org)

Simulation Structure for Selecting an Optimal Error Distribution Through the GAS Model

Richard T. A. Samuel [1,*], Charles Chimedza [1], Caston Sigauke [2]

[1] *School of Statistics and Actuarial Science, University of the Witwatersrand, South Africa*
[2] *Department of Mathematical and Computational Sciences, University of Venda, South Africa*

Abstract In econometrics and finance, volatility modelling has long been a specialised field for addressing a variety of issues that pertain to the risks and uncertainties of an asset. However, volatility modelling for risk management is highly dependent on the underlying error distribution. Hence, this study presents a Monte Carlo simulation (MCS) structure for selecting an optimal or the most adequate error distribution that is relevant for modelling the persistence of volatility through the Generalized Autoregressive Score (GAS) model. The structure describes an organised approach to the MCS experiment that includes "background of the study (optional), defining the aim of the study, specifying the research questions, method of implementation, and summarised conclusion". The method of implementation is a process that consists of writing the simulation code, setting the seed, setting the true parameter a priori, data generation, and performance evaluation through meta-statistics. Among the findings, the study used both fat-tails and $\sqrt{N}$ consistency experiments to show that the GAS model with a lower unconditional shape parameter value ($\hat{\nu}^* = 4.1$) can generate a dataset that adequately reflects the behaviour of financial time series data, relevant for volatility modelling. This dynamic structure is intended to help interested users on MCS experiments utilising the GAS model for reliable volatility persistence calculations in finance and other areas.

Keywords Consistency, Efficiency, Score driven model, Simulation structure, Time-varying parameter estimation

AMS 2010 subject classifications 91G70, 81T80

DOI: 10.19139/soic-2310-5070-1998

1. Introduction

In finance and econometrics, the modelling of volatility has long been used to address a variety of issues concerning the risks and uncertainties of an asset [69, 26]. This study follows the simulation design steps of our earlier study on volatility persistence estimation through the family Generalized Autoregressive Conditional Heteroscedasticity (fGARCH) model created by Hentschel [45] (see Samuel et al. [68]). That is, the family GARCH model was used in our earlier study to begin a Monte Carlo simulation (MCS) design to determine appropriate innovations for volatility persistence estimation (see Samuel et al. [68]). The Generalized Autoregressive Score (GAS) model is used in this new study to extend the robust design. As a result, this study identifies a vital simulation structure that is pertinent for selecting an optimal (or the most adequate) conditional error distribution for volatility persistence modelling using the GAS model.

The GAS is a dynamic model that is used for time-varying parameter modelling, much like the GARCH model (see [6]). It is an observation-driven model that was developed to model distributional asymmetries and large changes, such as shifts or jumps, in financial time series [38]. The model has been practically applied by Oh and Patton [61] for modelling systemic risk. Other broad applications are in market risk [42], spatial econometrics

*Correspondence to: Richard T. A. Samuel (Email: richieeniks@gmail.com). School of Statistics and Actuarial Science, University of the Witwatersrand. 1 Jan Smuts Ave, Braamfontein, Johannesburg, 2017, South Africa.

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 2024 International Academic Press

[10, 15], credit risk analysis [22], high-frequency data [62, 37], and dependence modelling [43, 48], among others (see [6, 38]). Hence, this study's main contribution is to show how the GAS model, using the GAS package [6], can be effectively used to determine a suitable conditional distribution that is relevant for volatility persistence estimation through MC simulation with empirical verification. Moreover, simulation through any model largely revolves around generating an appropriate dataset that can produce a reasonable output approximation. Thus, one further contribution of this study is to investigate how to use the GAS model to generate (or simulate) a dataset that can adequately reflect the behaviour of financial time series data that is relevant for volatility modelling.

Volatility is related to variation in the price of an asset over time. In finance, the variance is used as a proxy for volatility [18, 7]. However, emphasis is usually on conditional variance because time-varying volatility clustering exists in most financial time series [7]. Volatility persistence is a measure of the speed of the decay of volatility following a shock. The level of persistence has crucial economic implications that stem from the impact it has on the predictability of future value [4]. Risk managers, policymakers, portfolio managers, financial market participants, and other financial activities can derive benefits from a clear understanding of the effect of shocks on future volatility, especially with information on whether the shocks are short-lived or (highly) persistent.

MCS studies are computer-based experiments that involve the use of known probability distributions for creating data by pseudo-random sampling. The data may be simulated through parametric models or repeated resampling [59]. MCS is used for approximating real-life scenarios, where it uses random quantities (or a probabilistic approach) to provide deterministic outcomes [49]. Since its introduction by three scientists in the field of physics [36], MCS has become an increasingly applied technique in a wide range of fields including finance, supply chain, engineering, science [50], and in all portfolio and investment types [1]. One of the reasons for its wide applications is that it is based on the concept of the central limit theorem where the distribution of its estimates follows (or converges to) the Normal distribution [53].

This study used the daily Standard & Poor South African sovereign bond index (S&P SA bond index, henceforth). They are the bond market's data from Datastream [24] for the period January 4th, 2000, to June 17th, 2021, with 5598 data points. The period under study includes the COVID-19 pandemic crisis period that began in 2020. The rest of this paper is structured as follows: Section 2 reviews the theories underlying the GAS model, the true parameter recovery (TPR) measure, and the description of the simulation design. Section 3 shows how the simulation structure is illustrated using financial bond returns data, with empirical verification. Section 4 discusses the main findings and Section 5 presents concluding remarks on the novelty of the study.

2. Methods

2.1. The GAS Model

Volatility persistence can be effectively modelled through a Score Driven (SD) model known as the Generalized Autoregressive Score (GAS) proposed by Harvey [40] and Creal et al. [21] (see [6, 5, 13]). The model applies the score of the conditional density function to determine the time variation in the parameters. The score functions are robust to outliers because they discount extreme values by reducing the number of extreme observations [40, 63, 2], and the model is suitable for modelling fat-tailed and skewed time series data like financial returns [62, 40]. In addition, extensions to other time series dynamics, like asymmetry and long memory behaviour, are possible through the GAS process. Moreover, maximum likelihood estimation (MLE) through the model is easy to implement [6].

Creal et al. [19, 21, 22] showed that the robust GAS models incorporate many familiar observation-driven models that include the Autoregressive Conditional Heteroscedasticity (ARCH) model [28], the simple (standard) GARCH model [12], the Exponential GARCH (EGARCH) model [60], the Autoregressive Conditional Duration (ACD) model [29], the Autoregressive Conditional Intensity (ACI) model [66], the Beta-t-GARCH model [40, 41], the Autoregressive Conditional Multinomial (ACM) model [67], and several related models. In other words, new time-varying parameter models of empirical interest are easy to formulate due to the flexibility of the GAS framework (see [8]).

2.1.1. Specification of the GAS Model Let $N \times 1$ vector y_t be the dependent variable of interest, ϑ_t the vector of time-varying parameter, x_t a vector of exogenous variables (i.e., the covariates), all at time t, and ξ a vector of static parameters. Define $Y^t = \{y_1, \ldots, y_t\}$, $\Theta^t = \{\vartheta_0, \vartheta_1, \ldots, \vartheta_t\}$, and $X^t = \{x_1, \ldots, x_t\}$ [20, 21]. The information set that is available at time t consists of $\{\vartheta_t, \mathcal{F}_t\}$, where

$$\mathcal{F}_t = \{Y^{t-1}, \Theta^{t-1}, X^t\}, \quad \text{for } t = 1, \ldots, n. \tag{1}$$

It is assumed that the generation of y_t is carried out by the observation density

$$y_t \sim p\left(y_t|\vartheta_t, \mathcal{F}_t; \xi\right). \tag{2}$$

It is further assumed that the mechanism required to update the time-varying parameter ϑ_t is given by the autoregressive updating equation

$$\vartheta_{t+1} = \kappa + \sum_{i=1}^{p} A_i s_{t-i+1} + \sum_{j=1}^{q} B_j \vartheta_{t-j+1}. \tag{3}$$

Ardia et al. [6] presented Equation (3) as:

$$\vartheta_{t+1} \equiv \kappa + \mathbf{A}s_t + \mathbf{B}\vartheta_t, \tag{4}$$

where κ, $\mathbf{A}$ and $\mathbf{B}$ are matrices of coefficients with appropriate dimensions and they are functions of the time-invariant parameter ξ, while s_t is the scaled score and it is a suitable function of past data, $s_t = s_t(y_t, \vartheta_t, \mathcal{F}_t; \xi)$ (see [20, 21]). The s_t indicates the direction to update the vector of parameters from ϑ_t to ϑ_{t+1}. Vector κ is used to control the level of the process ϑ_t, while vector $\mathbf{A}$ is used to determine (or control) the impact of s_t on ϑ_{t+1}, and vector $\mathbf{B}$ is used to determine the persistence of the process [6, 51]. To implement the GAS models, matrices $\mathbf{A}$ and $\mathbf{B}$ are constrained to exist as diagonals (see [6, 61]). For instance, in a GAS model with Student's t conditional distribution, $\mathbf{A} \equiv \text{diag}(a_\mu, a_\phi, a_\nu)$ and $\mathbf{B} \equiv \text{diag}(b_\mu, b_\phi, b_\nu)$, where μ, ϕ and ν are location, scale and shape parameters, respectively. Hence, b_μ denotes the persistence of the conditional mean (location), while b_ϕ refers to the persistence of the conditional variance (or scale) (see [6]).

The approach used by the GAS is based on the observation density in Equation (2) for a given parameter ϑ_t. When a y_t observation is realised, the time-varying ϑ_t can be updated to the next period $t+1$ using Equation (3) with

$$s_t = S_t \cdot \nabla_t, \qquad \nabla_t = \frac{\partial \ln p\left(y_t|\vartheta_t, \mathcal{F}_t; \xi\right)}{\partial \vartheta_t}, \qquad S_t = S\left(t, \vartheta_t, \mathcal{F}_t; \xi\right), \tag{5}$$

where $S(\cdot)$ represents a matrix function, ∇_t is the score of Equation (2) evaluated at ϑ_t, ln is the natural logarithm, and S_t denotes the scaling matrix (see [6, 20, 21]). The score of the GAS model depends on the complete density structure [20, 21]. Given that the driving mechanism in Equation (3) depends on the scaled score vector in Equation (5), the GAS model with orders p and q can be defined by Equations (2), (3) and (5). This can be referred to as the GAS(p,q) model, where the orders p and q are typically taken as $p = q = 1$ (see [20, 21]). However, readers can refer to Creal et al. [21] for details on including more lags in the GAS process [8].

To account for the variance of ∇_t, Creal et al. [21] suggested setting S_t to a power $\gamma > 0$ of the inverse of the information matrix ($\mathcal{I}$) of ϑ_t [6]. Precisely,

$$S_t = \mathcal{I}_{t|t-1}^{-\gamma}, \qquad \mathcal{I}_{t|t-1} = \mathrm{E}_{t-1}\left[\nabla_t \nabla_t^\top\right], \tag{6}$$

where the expectation E_{t-1} is taken with respect to the conditional distribution of $y_t|y_{1:t-1}$. The parameter γ usually takes 0, $\frac{1}{2}$, or 1 value [6], but other choices of S_t are also possible (see [8]). If $\gamma = 0$, S_t becomes an identity matrix ($\mathbf{I}$), which implies there is no scaling. If $\gamma = \frac{1}{2}$, then the conditional score ∇_t is pre-multiplied by the square root of its covariance matrix $\mathcal{I}_t$, but if $\gamma = 1$, the conditional score ∇_t is pre-multiplied by the inverse of its covariance matrix $\mathcal{I}_t$. The scaled score s_t is a martingale difference with respect to the distribution of $y_t|y_{1:t-1}$, i.e., $\mathrm{E}_{t-1}[s_t] = 0$ for all t (see [6]).

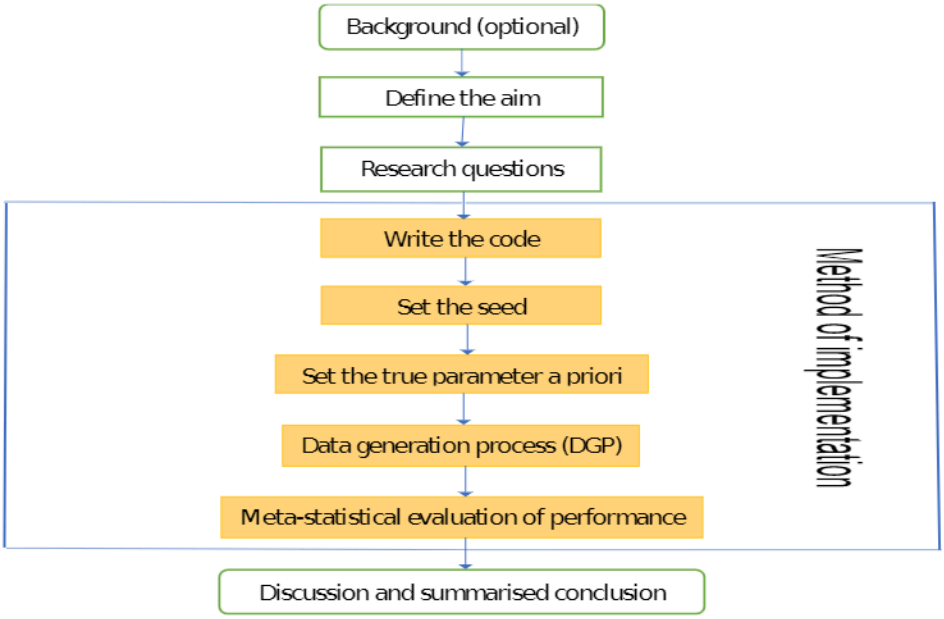


Figure 1. Flowchart on simulation design to select an optimal error distribution.

The GAS framework incorporates many familiar observation-driven models as a special case for an appropriate choice of the scaling matrix S_t [21]. That is, various classes of observation-driven models can be obtained with different choices of the scaled score s_t, hence these types of models are referred to as score-driven models [51, 20, 21]. In order words, Equations (2), (3) and (5) encompass a large set of common time series models for different choices of s_t [8]. For more details on the GAS models, see [6, 40, 19, 20, 13, 21, 22, 51, 9].

2.2. *The True Parameter Recovery Measure*

The true parameter recovery (TPR) measure was developed by Samuel et al. [68] as a means of measuring the performance of the MCS estimator in recovering the true parameter. It is used as a proxy for the coverage of the MCS experiment to calculate the level of recovery of the true data-generating parameter by the MCS estimator. It is stated as:

$$\text{TPR} = \left(K - \left[\frac{(\vartheta - \hat{\vartheta})}{\vartheta} \times K \right] \right) \%, \tag{7}$$

where ϑ and $\hat{\vartheta}$ are the true parameter and estimator, respectively, while $K = 0, 1, 2, \ldots, 100$ is the nominal recovery level. A TPR outcome of 99% or 95%, for instance, denotes the recovery of 99% or 95% of the true parameter by the MCS estimator. The estimator $\hat{\vartheta}$ can recover the true parameter ϑ completely when $\hat{\vartheta} = \vartheta$, where $\vartheta > 0$. Hence, the full recovery of the true parameter occurs where the TPR estimate is equal to the specified nominal recovery level K, that is, TPR = $K\%$.

2.3. *The Design of the Simulation*

The design of the simulation structure involves "(optionally) stating the background of the study, defining the aim of the study, specifying the research questions, method of implementation, and summarised conclusion". The method of implementation is a process that consists of writing the simulation code, setting the seed, setting the true parameter a priori, data generation, and performance assessment through meta-statistics. These steps, as summarised by the flowchart [68] in Figure 1, are relevant for a successful simulation to select an optimal assumed innovation distribution using the GAS model. Each step of the structure is detailed as follows:

2.3.1. Aim of the Simulation Study After optionally stating the background of the study, the next step in the design is to define the aim of the study, and it should be concisely stated to communicate the purpose of the study to readers. The focus of many MCS studies is generally on the estimators' abilities to recover the true parameters ϑ [16], such that $\mathrm{E}(\hat{\vartheta}) = \vartheta$ for unbiasedness, standard error (SE) or root mean square error (RMSE) tends to zero as the sample size $N \to \infty$ for good precision or efficiency of the true parameter's estimator, and $\hat{\vartheta} \to \vartheta$ as $N \to \infty$ for consistency. Hence, the study's aim could be about those features, like the consistency, bias or unbiasedness, precision or efficiency of the estimator. The aim may also focus on comparisons of multiple operations, like comparing the efficiency of various conditional distributions to obtain the most efficient, or proposing a modified version of an existing model.

2.3.2. State the Research Questions Next, relevant questions that point to the objectives of the simulation should be specified and addressed in the study. The complexities of some statistical research questions may be better resolved through simulation approaches. Simulation provides a robust process for responding to many methodological and theoretical questions, and it can proffer a flexible structure for responding to certain questions that are relevant to the quest of the researchers [39].

2.3.3. Method of Implementation We carry out the simulation and empirical studies of this paper in R software, version 4.0.3, with RStudio version 2023.03.1+446, using the GAS [6, 13], tidyverse [74], patchwork [64], SimDesign [17], rugarch [33, 34], forecast [47], and aTSA [65] packages.

The method of implementing the MCS is described as follows:

- Write the simulation code: Execution of a suitable simulation experiment that reflects a real-life situation can be computationally intensive and very demanding, hence readable computer code with the right syntax must be ensured. MCS code must be well organised to avoid difficulties during debugging. Code must be efficiently and flexibly written and well arranged for easy readability.
- Set the seed: If a seed is not set, the simulation code will generate a different sequence of random numbers whenever it is run [23]. A set seed is used to initialise the random number generator [34] and to ensure reproducibility, where the same output is obtained for different runs of the simulation process [32]. It is better to use the same seed values throughout the process [59], and this (seed) needs to be set only once, for each simulation, at the start of the simulation session [59, 34]. The seed is essentially used for reproducibility. Samuel et al. [68] used the GARCH model to show that as the sample size N increases, the arrangement (or pattern) of the seed values generally does not influence the efficiency and consistency of an estimator. This however may depend on the quality of the model used.
- Next, the true parameter value of the true sampling distribution (or true model[†]) is set a priori [58, 51].
- Following this, simulated return observations are generated using the true sampling distribution or the true model given some sets of (or different sets of) parameters. Simulated data generation through the GAS model is carried out using the R GAS package. The generation of random data through this package can be implemented using either of two approaches [6, 13]. The first approach is to generate the data directly from an estimated "uGASFit" object using the UniGASSim (univariate GAS simulation) function for the simulated random data. In this study, we use "fit" as the uGASFit object. The second approach involves the full specification of a GAS model that includes the selection of the conditional error distribution of the time series process, and specifying the static parameter $\xi = (\kappa, \mathbf{A}, \mathbf{B})$ that controls the time variation in ϑ_t, as described in Section 2.1. In this latter approach, the fit object will be specified as "fit = NULL" in the UniGASSim function.

  The data-generating (or simulation) process can be replicated multiple times or run only once. However, data generation through the GAS package is generally designed to run only once. In their study through the autoregressive process, Samuel et al. [68] showed that the outcomes of MCS experiments using the (family) GARCH model are the same for datasets simulated with the same seed value, regardless of whether the

[†]The true model is described as the data-generating model fitted with the true error distribution (see [30]).

simulation or data-generating process is run once or replicated multiple times. Hence, the outcome of a single run is the same as the average of multiple runs or replications. Their findings further revealed that in any multiple generated (or simulated) datasets in the family GARCH model, each simulated series is distinct in randomness and shape, and it is different from every other series in the datasets.

- Next, the generated data are analysed and the estimates from the analysis are assessed using classic methods through meta-statistics to obtain relevant information about the estimator. Meta-statistics (see [17]) are metrics or performance measures used for evaluating the modelling outcomes by judging the closeness between the true parameter and its estimate. The root mean square error (RMSE), bias, standard error (SE), and relative efficiency (RE) as described below are part of the familiar meta-statistical summaries used in the literature. Readers can refer to [59, 17, 71] for more meta-statistics.

Bias

Bias is defined as the mean difference between the true parameter and its estimate [30]. Its optimal value is 0 [71, 44]. Bias with a negative (positive) value indicates that the true parameter value is under-estimated (over-estimated). However, the nearer the estimator is to zero (in absolute values), the better it is. Bias can be stated in Monte Carlo simulations (see [16]) as:

$$\text{bias} = \frac{1}{L}\sum_{i=1}^{L}(\hat{\vartheta}_i - \vartheta), \tag{8}$$

where $\hat{\vartheta}$ denotes the sample estimate for the ith dataset, ϑ is the true parameter, and L is the number of replications [71].

Standard Error

The standard error (SE), as presented (see [16, 75]) in Equation (9), is used to assess the sampling variability in the estimation. It is a measure of the precision or efficiency of the true parameter's estimator. The lower the sampling variability, the more efficient or precise the true parameter's estimator (see [59]). As the sample size increases, the sampling variability decreases [71].

$$\text{SE} = \sqrt{\frac{1}{L}\sum_{i=1}^{L}(\hat{\vartheta}_i - \bar{\vartheta})^2}, \quad \text{where} \quad \bar{\vartheta} = \sum_{i=1}^{L}\hat{\vartheta}/L. \tag{9}$$

RMSE

The root mean square error (RMSE) is used to assess the difference between the true value of a model and its prediction. RMSE measure denotes the sampling error of an estimator in comparison with the true parameter value [71], and it can be stated as:

$$\text{RMSE} = \sqrt{\frac{1}{L}\sum_{i=1}^{L}\left(\hat{\vartheta}_i - \vartheta\right)^2}. \tag{10}$$

An estimator with lower RMSE will recover the true parameter value more efficiently [71, 75], and minimum RMSE yields maximum precision [72]. An estimator is consistent when RMSE decreases such that $\hat{\vartheta} \to \vartheta$

as $N \to \infty$ [59, 33]. RMSE, bias and sampling variability are connected as:

$$\text{RMSE} = \sqrt{\text{bias}^2 + \text{SE}^2}. \tag{11}$$

That is, the RMSE measure includes both bias and SE, where low SE can be penalised for bias. The mean squared error (MSE) can be derived by squaring the RMSE. It is expected that as the sample size increases, the sampling error should decrease relatively, but this is not always the outcome [35].

Relative Efficiency

Sampling efficiency between one estimator and the other can be compared using the relative efficiency (RE) statistic by creating ratios between the estimators' RMSE values as shown in Equation (12). In the formula, the first RMSE or RMSE_1 is taken as the reference estimator, such that other RMSEs are assessed in relation to it. The RE $\equiv 1$ when $i = 1$, and values above (less than) 1 indicate less (greater) efficiency. Less efficiency means more variability than the statistic that is referenced, while greater efficiency indicates less variability and higher precision for recovering the true parameter ϑ [71, 16]. RE < 1 suggests that estimator i is the preferred one, i.e., if $i = 2$ for instance, estimator 1 is less efficient in relation to estimator $i = 2$ [25].

$$\text{RE} = \left(\frac{\text{RMSE}_i}{\text{RMSE}_1}\right)^2. \tag{12}$$

2.4. Discussion and Summary

After the method of implementation, the last step in the structure is the conclusion, which should reflect a summary discussion of all relevant findings from the simulations, with answers provided to the research questions. The novelty of the work is presented in the conclusion, and this may include the limitations encountered and the future directions of the work.

3. Results: Simulation and Empirical

3.1. Practical Illustrations of the Simulation Structure with Applications to Financial Returns

This section practically illustrates how the design steps can be applied using MCS with empirical verification through the real SA bond returns data. The steps are illustrated as follows:

3.2. The Background

As a preamble, simulation involving the GAS model can be conducted through the UniGASSim function with the uGASFit object (e.g., using "fit" as the uGASFit object), and through the full specification of the GAS model (with "fit = NULL" in the UniGASSim function)[‡]. In the latter approach which involves the full specification of the model, the coefficients of the matrices κ, $\mathbf{A}$ and $\mathbf{B}$ in Equation (4) have to be specified. Both of these approaches can be done with a shape parameter (or degree of freedom) that is time-varying (ν_t) or constant (ν). This illustrative study focuses on ν_t. In the simulation structure, $\mathbf{A}$ and $\mathbf{B}$ matrices are constrained to exist as diagonals as described in Section 2.1.1. Hence, a_μ, a_ϕ, and a_ν are diagonal elements of $\mathbf{A}$, while b_μ, b_ϕ, and b_ν are diagonal elements of $\mathbf{B}$. In the outputs of the estimation, a1, a2, a3, b1, b2 and b3 are estimates of a_μ, a_ϕ, a_ν, b_μ, b_ϕ and b_ν, respectively. Also $\kappa \equiv (\mathbf{I} - \mathbf{B})\Lambda^{-1}(\vartheta^*)$, where $\mathbf{I}$ is an identity matrix of a suitable size, ϑ^* is the ThetaStar object, and Λ^{-1} is a mapping function that is represented by the UniUnmapParameters() function[§]. The ThetaStar ϑ^* is a vector of

[‡]Readers can refer to Page 52 of Catania et al. [14] for the usage of the UniGASSim function.
[§]That is, the inverse of the mapping function Λ^{-1} in κ is represented by the UniUnmapParameters() function in the GAS simulation code of the R package GAS (see [6]).

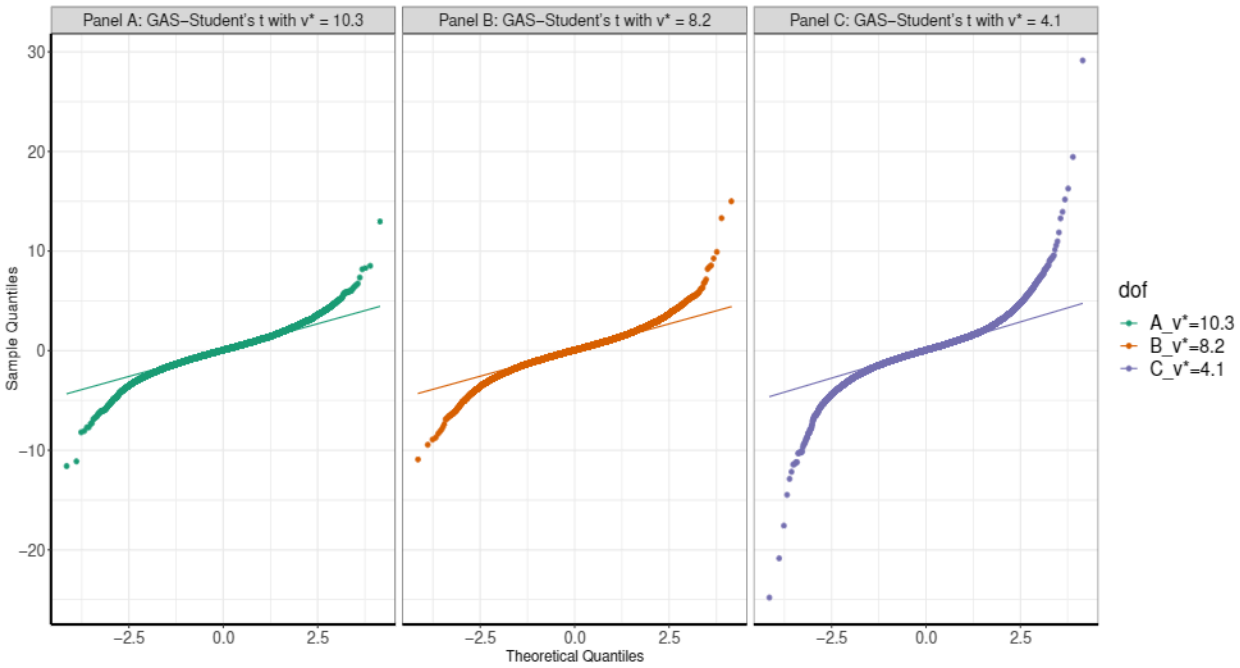


Figure 2. The QQ plots in Panels A, B, and C are obtained from GAS-Student's t model fitted with $\nu^* = 10.3$, $\nu^* = 8.2$, and $\nu^* = 4.1$, respectively. Each model is fitted to $N = 30{,}000$ simulated datasets.

the unconditional time-varying parameters, i.e., $\vartheta^* = (\mu^*, \phi^*, \nu^*)^\top$, where μ^*, ϕ^*, and ν^* are the unconditional location, scale, and shape parameters, respectively. The coefficients kappa1, kappa2 and kappa3 in the estimation's outputs are vector κ's elements, where $\kappa \equiv (\kappa_\mu, \kappa_\phi, \kappa_\nu)^\top$ (see [6] for more details).

Volatility modelling for risk management is highly dependent on the underlying error distribution. In other words, to minimize risk and adequately maximise asset returns, volatility needs to be modelled with an appropriate error assumption when the true (underlying) error distribution is unknown. Hence, this illustrative study demonstrates the dynamics of the GAS model as the true data generating process (DGP) for the simulation, where the outcomes of the model fitted with each of seven assumed error distributions are compared to obtain the most suitable for volatility persistence estimation. The assumed error distributions are the Gaussian, skew-Gaussian, Student's t, skew-Student's t, asymmetric Student's t with one tail decay parameter (AST1), asymmetric Student's t with two tail decay parameters (AST), and asymmetric Laplace distribution (ALD). Details on the error distributions can be seen in [19, 21, 6, 33, 76, 52] or through the function DistInfo() in R package "GAS".

For further background layout, it is known that financial data are fat-tailed [55], i.e., non-Normal. Hence, to obtain a set of simulated observations that can closely reflect this fat-tailed behaviour of financial time series data under the GAS modelling, we carry out experiments to compare the fat-tails outcomes of three different datasets that are generated using three unconditional shape parameters or degrees of freedom $\nu^* = \{10.3, 8.2, 4.1\}$. We execute this by fitting GAS-Student's t model[¶] with time-varying location, scale, and shape parameters to the SA bond returns data to generate three sets of $N = 30{,}000$ simulated datasets. That is, the first dataset is obtained with $\nu^* = 10.3$, the second and third with $\nu^* = 8.2$ and $\nu^* = 4.1$, respectively. For each simulated dataset, a quantile-quantile (QQ) plot is produced as displayed in Figure 2. It can be seen from the figure that the plot in Panel C for GAS-Student's t with $\nu^* = 4.1$ displays the fattest tails than the other two panels with $\nu^* = 8.2$ and $\nu^* = 10.3$. The tails of Panel C's plot are about twice and more than twice as fat as those of Panels B and A, respectively. Hence,

[¶]GAS-Student's t model is the GAS model fitted with a Student's t error.

the GAS model fitted to a Student's t distributed error with ν^* of 4.1 appears to be the most suitable among the three to generate the required simulated financial returns data.

Based on the fat-tails outcomes, we further carry out two sets of experiments to examine the $\sqrt{N}$ consistency of datasets fitted with each of these three unconditional shape parameters $\nu^* = \{4.1, 8.2, 10.3\}$. In the first experiment, we fit GAS-Student's t model with time-varying location, scale, and shape parameters to the SA bond returns data to obtain the parameter values that are used to generate three datasets of sample sizes $N = \{30{,}000; 31{,}000; 32{,}000\}$. That is, all the outcomes from the fit (with ν^* fixed in turn to 4.1, 8.2 and 10.3) are used to generate the three datasets. This data generation process is executed through the full specification of the GAS model. Precisely, in the first set of this first experiment, we use the outcomes from the fit with ν^* fixed to 4.1 to generate the three datasets. In the second set, we also use the outcomes from the fit with ν^* fixed to 8.2 to generate the next three datasets. In the third set, we again use the outcomes from the fit with ν^* fixed to 10.3 to generate the last three datasets.

Next, we fit the GAS model with a time-varying scale parameter to each set of the three datasets under four error distributions of the Gaussian, skew-Gaussian, Student's t, and skew-Student's t. Panel A of Table 6 shows the RMSE, SE and bias outcomes of the fit under each error distribution for the GAS volatility persistence estimator $\hat{b}_\phi$. It can be seen from the table that the GAS model fitted with $\nu^* = 4.1$ gives the best consistency in terms of RMSE, SE and absolute value of bias. That is, the model fitted using $\nu^* = 4.1$ outperforms those with $\nu^* = 8.2$ and $\nu^* = 10.3$ in consistency under the four error distributions. These outcomes are visually displayed in Figure 3, where Panels A and B show the superior performance of $\nu^* = 4.1$ in terms of SE and RMSE[‖], respectively, as compared to those of $\nu^* = 8.2$ (in Panels C and D) and $\nu^* = 10.3$ in Panels E and F. The RMSE displays some inconsistency for $\nu^* = 8.2$ and $\nu^* = 10.3$ in Panels D and F, respectively.

In the second experiment, we try to avoid potential initial value effects by generating three equal datasets, each of sample size $N = 30{,}000$. From the datasets, we remove the first $N = \{13{,}000; 12{,}000; 11{,}000\}$ to avoid the effect of initial values, and then use the last $N = \{17{,}000; 18{,}000; 19{,}000\}$ under each of the four error distributions as shown in Panel B of Table 6. These trimmings (and subsequent ones in this study) are executed following the simulation design of Feng and Shi [30]. The outcomes of the experiment are presented in Panel B of Table 6 in the Appendix. It can be seen from the table that the GAS model fitted with $\nu^* = 4.1$ gives the best $\sqrt{N}$ consistency in terms of RMSE, SE and absolute value of bias when compared with those of $\nu^* = 8.2$ and $\nu^* = 10.3$ under the four error distributions. These consistency outcomes are visually displayed in Figure 4, where Panels A and B show the superior performances of $\nu^* = 4.1$ in terms of SE and RMSE, respectively, as compared to those of $\nu^* = 8.2$ in Panels C and D, and $\nu^* = 10.3$ in Panels E and F. Moreover, when compared with the outcomes of the first experiment, it is observed as shown in Panels A and B of Figure 4 that the $\sqrt{N}$ consistency of the estimator $\hat{b}_\phi$ for $\nu^* = 4.1$ becomes better and more improved in terms of SE and RMSE after removing the initial value effects. Hence, the Student's t GAS model fitted with $\nu^* = 4.1$ outperforms those fitted with $\nu^* = 8.2$ and $\nu^* = 10.3$ in consistency.

The Student's t is used as the true error distribution for this illustration because it is assumed that financial data appear to have a distribution much like it, and it is also believed that it can suitably deal with fat-tailed or leptokurtic features [56, 27] commonly found in financial data [45]. However, to buttress the illustration, we further use the AST1 as another true error distribution in addition to the Student's t, and compare their outcomes. Besides these two true sampling errors, users may choose to use any relevant non-Gaussian distributions, as stated in the DistInfo() function, for their data generation based on individual research needs. For any chosen true error distribution, the DistInfo() function reveals the parameters that must be specified as time-varying.

3.3. Aim of the Simulation Study

The aim of this study is to obtain the most suitable assumed error distribution to describe the returns for volatility persistence estimation when the underlying error distribution is not known.

[‖] We did not report the plots of the $\sqrt{N}$ consistency of the absolute value of bias for brevity reasons.

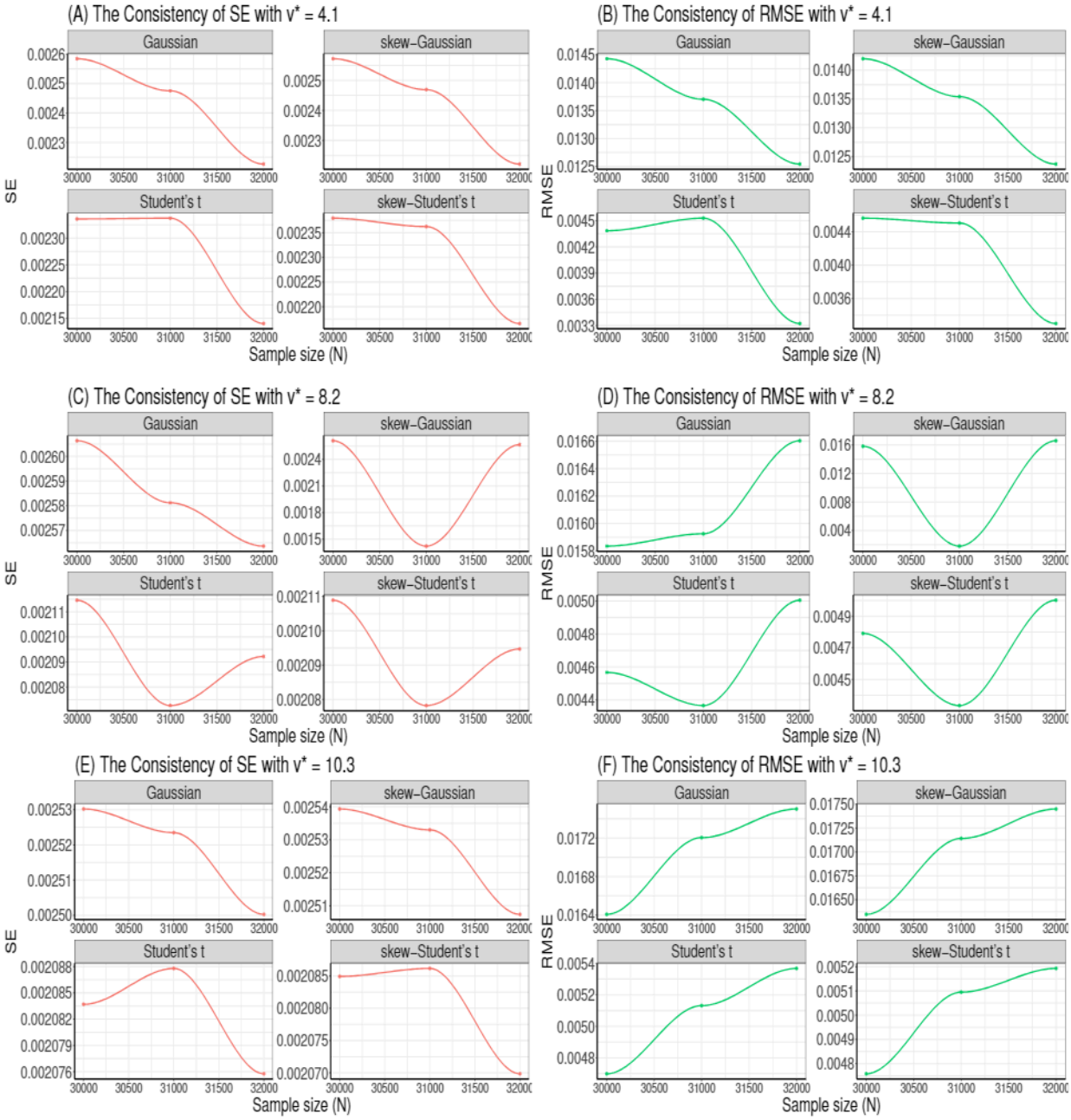


Figure 3. Panels (A) and (B) display the $\sqrt{N}$ consistency of the GAS estimator $\hat{b}_\phi$ in terms of SE and RMSE for $N = \{30{,}000; 31{,}000; 32{,}000\}$ datasets with $\nu^* = 4.1$, while Panels (C) and (D) display those for $\nu^* = 8.2$, and Panels (E) and (F) for $\nu^* = 10.3$.

3.4. Research Questions

The outcomes of this simulation study should help to respond to the following questions:

1. Which among the error assumptions is the most suitable for the GAS simulation process to estimate the volatility persistence?

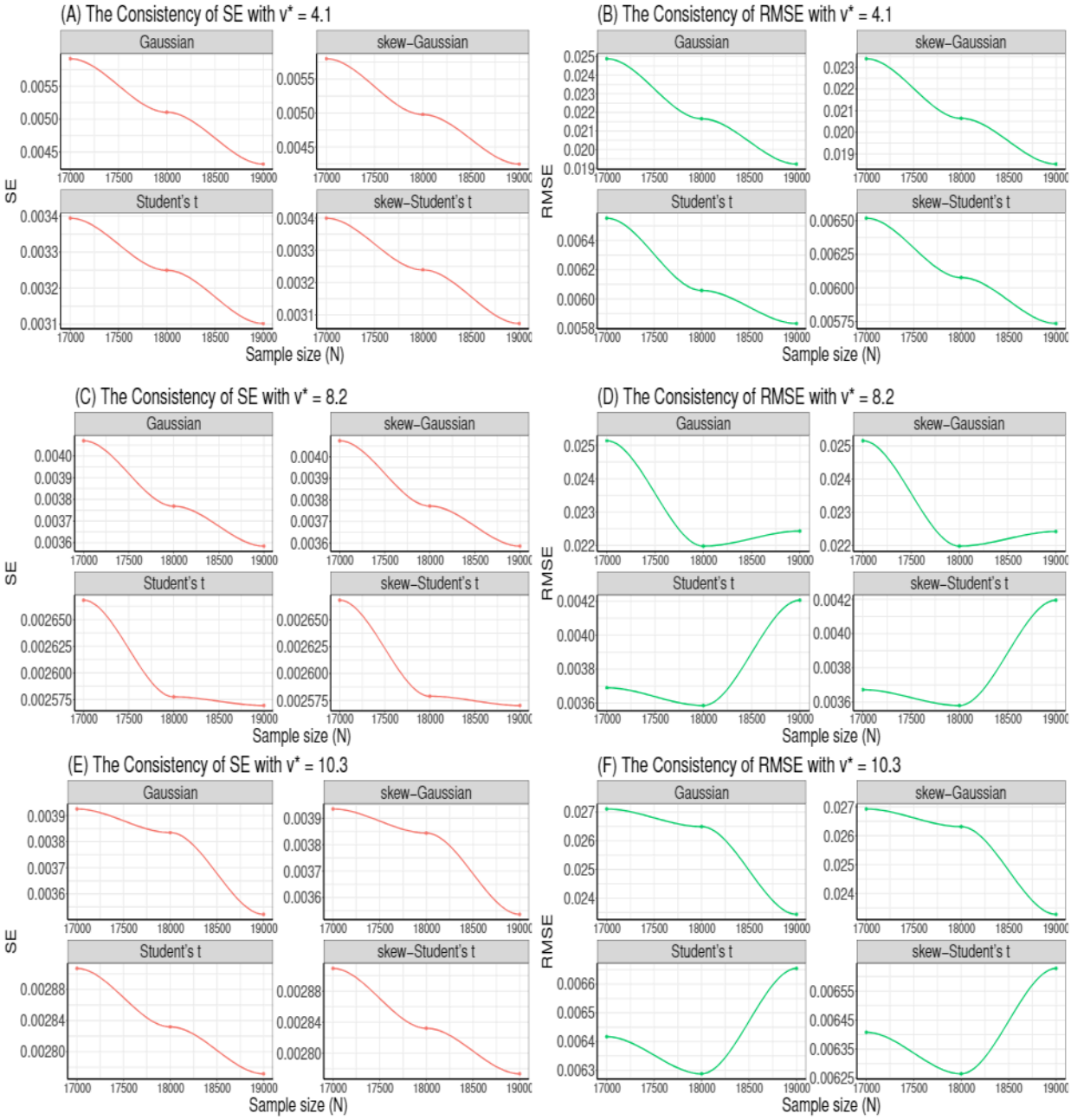


Figure 4. Panels (A) and (B) display the $\sqrt{N}$ consistency of the GAS estimator $\hat{b}_\phi$ in terms of SE and RMSE for $N = \{17{,}000; 18{,}000; 19{,}000\}$ datasets with $\nu^* = 4.1$, while Panels (C) and (D) display those for $\nu^* = 8.2$, and Panels (E) and (F) for $\nu^* = 10.3$.

2. What type (i.e., weak or strong) of $\sqrt{N}$ consistency, in terms of SE and RMSE, does the volatility persistence estimator $\hat{b}_\phi$ of the GAS process exhibit?
3. How does the estimator $\hat{b}_\phi$ of the most suitable assumed error distribution compare to the estimator $\hat{b}_\phi$ of competing assumed error distributions with regards to bias, and efficiency (or precision) in terms of RMSE, SE and RE?

4. How does the GAS estimator $\hat{b}_\phi$ perform in recovering the true parameter b_ϕ?

3.5. Method of Implementation

Since we intend to use the Student's t and AST1 as the true error distributions to generate the simulated return observations, the method of implementation is described via two Monte Carlo experiments. The implementation method is illustrated using the full specification of the GAS model in the coefficients of matrices **A**, **B** and κ. For the two true error distributions, the scale parameters' descriptions of the conditional variance are within $0.975 < b_\phi < 1$ and $0 < a_\phi < 0.15$. The b_ϕ is the GAS volatility persistence parameter in the diagonal of matrix **B** (see [21, 11]). Moreover, the b_ϕ parameter of the GAS model corresponds to the $\alpha + \beta$ parameters of the GARCH model** [11].

Next, the two MCS experiments to illustrate the method of implementation are carried out using each of the two true error distributions one after another as follows.

3.5.1. Method of Implementation with the Student's t Error With Student's t as the true innovation distribution, the written code is first used to fit the true model GAS-Student's t with time-varying†† location, scale and shape parameters to the SA bond returns data through the UniGASFit function, to obtain the true parameter values for the data generating process. The true parameter values are the MLE estimates obtained from fitting the GAS-Student's t model to the SA bond returns data. The time variation in the location, scale and shape parameters of the fit are graphically described as displayed in Figure 5, with the location in the top panel, followed by the scale and then the shape parameter graphics. The true parameter values used to generate the simulated returns data are $a_\mu = 0.0000$, $b_\mu = 0.3461$, $a_\phi = 0.1400$, $b_\phi = 0.9831$, $a_\nu = 9.8356$, $b_\nu = 0.5253$, $\mu^* = 0.0578$, $\phi^* = 1.0695$ (and ν^* fixed to 4.1). The a_μ, b_μ denote the location parameters, a_ϕ, b_ϕ indicate the scale parameters, and a_ν, b_ν are the shape parameters, while μ^*, ϕ^* and ν^* are the unconditional location, scale and shape parameters, respectively.

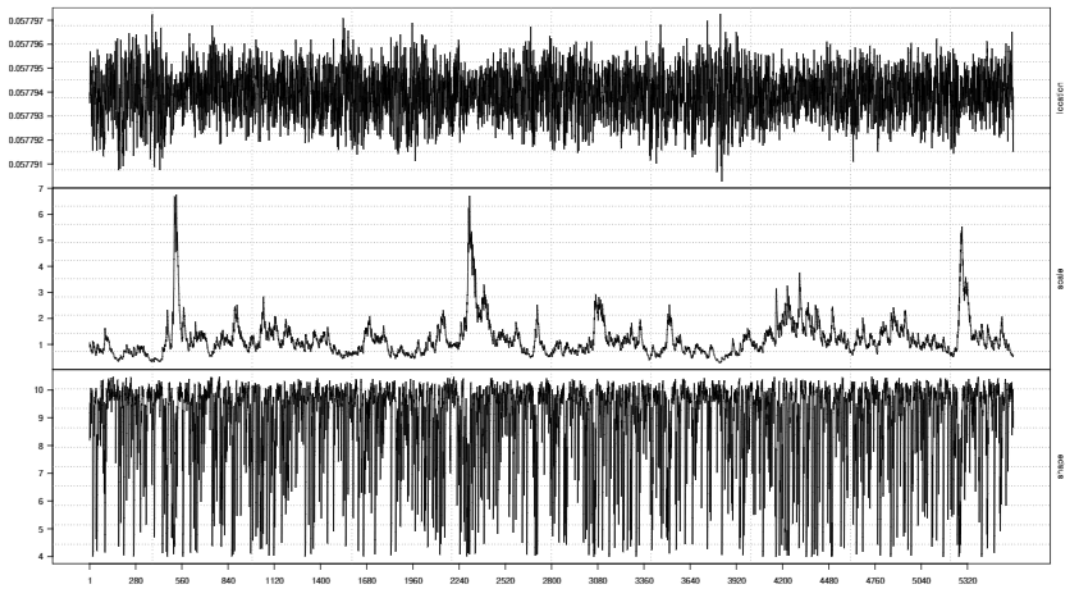


Figure 5. GAS-Student's t: Time-varying location (top panel), scale (middle panel), and shape (bottom panel) parameters.

Next, through the UniGASSim function, using seed 12345 in the code, the value of the scale parameter b2 = $b_\phi = 0.9831$ from the fit is set a priori as the true parameter for the MCS process, as shown in Table 1. This value

**We examined this by comparing the estimate of the GAS persistence parameter $\hat{b}_\phi$, fitted with a time-varying scale parameter, and the estimate from the GARCH(1,1) model's persistence parameter $\hat{\alpha}_1 + \hat{\beta}_1$. We fitted both models to the SA bond returns data where the innovation is Gaussian, and their results yielded $\hat{b}_\phi \equiv \hat{\alpha}_1 + \hat{\beta}_1 \approx 0.98$.

††Time-varying parameters are specified as "TRUE" in the UniGASSpec function.

Table 1. Simulation outcomes.

True model: GAS-Student's t True parameter: $b_\phi = 0.9831$							
Assumed error	N	$\hat{b}_\phi$	**RMSE**$_{\hat{b}_\phi}$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**RE**$_{\hat{b}_\phi}$	**TPR**$_{\hat{b}_\phi}$ **(95%)**
Gaussian	17,000	0.9589	0.0249	-0.0242	0.0059	–	92.66%
	18,000	0.9620	0.0217	-0.0210	0.0051	0.7575	92.97%
	19,000	0.9643	0.0192	-0.0187	0.0043	0.7872	93.19%
skew Gaussian	17,000	0.9604	0.0234	-0.0227	0.0058	–	92.81%
	18,000	0.9630	0.0206	-0.0200	0.0050	0.7781	93.06%
	19,000	0.9650	0.0185	-0.0180	0.0042	0.8055	93.26%
Student's t	17,000	0.9775	0.0066	-0.0056	0.0034	–	94.46%
	18,000	0.9780	0.0061	-0.0051	0.0032	0.8547	94.51%
	19,000	0.9781	0.0058	-0.0049	0.0031	0.9268	94.52%
skew Student's t	17,000	0.9775	0.0065	-0.0056	0.0034	–	94.46%
	18,000	0.9779	0.0061	-0.0051	0.0032	0.8692	94.50%
	19,000	0.9782	0.0057	-0.0048	0.0031	0.8906	94.53%
AST	17,000	0.9774	0.0066	-0.0057	0.0034	–	94.45%
	18,000	0.9777	0.0063	-0.0054	0.0033	0.9061	94.48%
	19,000	0.9782	0.0058	-0.0049	0.0031	0.8398	94.53%
AST1	17,000	0.9775	0.0065	-0.0056	0.0034	–	94.46%
	18,000	0.9780	0.0060	-0.0051	0.0032	0.8581	94.51%
	19,000	0.9788	0.0053	-0.0043	0.0031	0.7623	94.59%
ALD	17,000	0.9768	0.0073	-0.0063	0.0037	–	94.39%
	18,000	0.9775	0.0066	-0.0056	0.0035	0.8140	94.46%
	19,000	0.9778	0.0062	-0.0052	0.0033	0.8827	94.49%

Note: Simulation outcomes from the true model GAS-Student's t.

and all the other estimates (with the ν^* that is fixed to 4.1) from the uGASFit object are used in the code to generate three sets of $N = 30{,}000$ return observations.

Next, due to the effect of the initial values in the generation of the data, which may cause size distortion [70], we discard the first $N = \{13{,}000; 12{,}000; 11{,}000\}$ sets of observations at each stage of the generated 30,000 return observations to avoid the distortion. Hence, we only use the last $N = \{17{,}000; 18{,}000; 19000\}$ simulated returns under each of the seven assumed errors as presented in Table 1. After generating the simulated return datasets, the GAS model is fitted to them under each of the seven assumed error distributions. The full analytic code that shows the lines of commands for each stage of the method of implementation is available from the authors on request.

Next, comparisons through the four meta-statistical summaries comprising of the RMSE, bias, SE and RE are carried out for the volatility persistence estimator $\hat{b}_\phi$ in **B**. The most adequate assumed innovation for volatility persistence estimation will be obtained from the estimator $\hat{b}_\phi$ with the best efficiency and precision (for recovering the true parameter) in the meta-statistical comparisons executed under the seven innovation assumptions.

Now, comparing RMSE for $\hat{b}_\phi$ in Table 1, the AST1 assumed error distribution outperforms the other six assumed errors, including the true error Student's t, in efficiency with the least value as N tends to the peak at 19,000. For bias comparison, the AST1 also outperforms the rest of the error assumptions with the lowest absolute value of bias as N tends to the peak. Comparing SE, the Student's t, skew-Student's t, and AST1 are equal in performance for efficiency and precision in terms of SE, and they outperform the rest. The AST assumed error follows them in performance. Lastly, for RE comparison, we use each preceding RMSE as the reference estimator, i.e., RMSE_1 as the reference estimator for RMSE_2, and RMSE_2 for RMSE_3 under each error assumption. The outcomes show that the AST1 displays the highest precision for recovering the true parameter than the rest, with the smallest RE value as N tends to the peak. The RMSE and SE performance measures also showcase adequate $\sqrt{N}$ consistency

under the seven assumed error distributions. To summarise, when the true error distribution is Student's t, the AST1 assumed error distribution relatively outperforms the other six error assumptions in efficiency and precision, and in the overall performance.

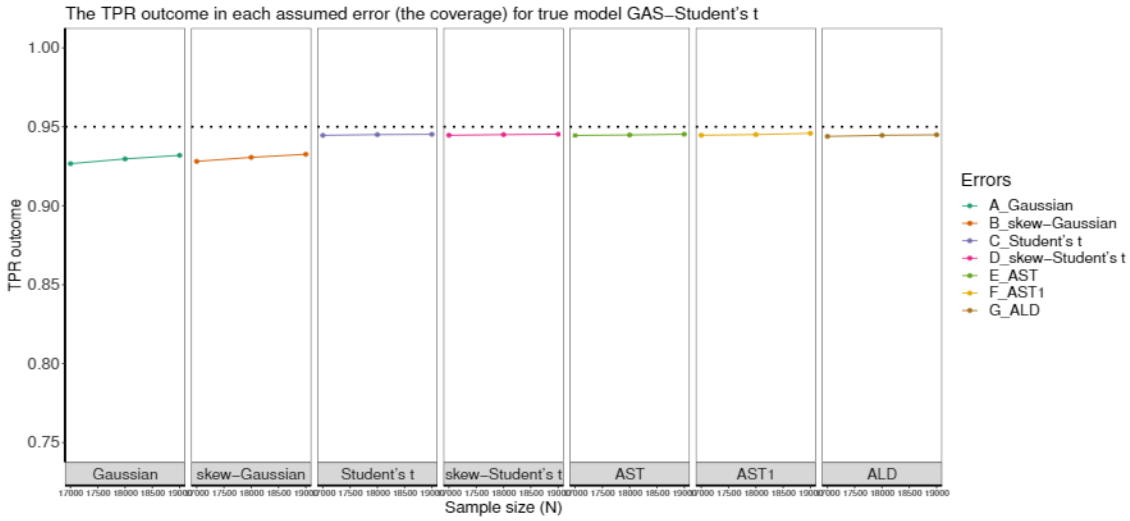


Figure 6. The true parameter recovery (TPR) outlook. The dotted line denotes the 0.95 (i.e., 95%) nominal recovery level.

Table 1 further reveals that the MCS estimator $\hat{b}_\phi$ considerably recovers the true parameter b_ϕ at the 95% (i.e., 0.95) nominal recovery level. The recovery outcomes for the seven assumed error distributions are visually displayed in Figure 6, where it is revealed that the MCS estimates perform considerably well in recovering the true parameter as shown by the closeness of the TPR outcomes to the 95% nominal recovery level. This indicates that the MCS experiments are well executed with suitably valid outcomes. However, the non-Gaussian error assumptions perform better in the recovery than the Gaussian and skew-Gaussian error assumptions, as shown in the plots. It is also observed from the results in the table that the MCS estimate $\hat{b}_\phi$ is directly proportional to the TPR outcome. Hence, the smaller (larger) the MCS estimate, the smaller (larger) the TPR outcome.

3.5.2. Method of Implementation with AST1 Error The second illustration of the method of implementation is carried out with the AST1 as the true error distribution. The written lines of code are first used to fit the true model GAS-AST1 with time-varying location, scale, skewness, and shape parameters to the SA bond returns data through the UniGASFit function. The time variations in these parameters are graphically displayed in Figure 7, with the location in the top panel, followed by the scale, skewness, and shape parameter graphics downwards in subsequent panels. The MLE estimates from the fit are used as the true parameter values to generate the simulated returns for the MCS. The true parameter values used to generate the returns data are $a_\mu = 0.0000$, $b_\mu = 0.8802$, $a_\phi = 0.0420$, $b_\phi = 0.9788$, $a_\eta = 0.1186$, $b_\eta = 0.0000$, $a_\nu = 0.0001$, $b_\nu = 0.9702$, $\mu^* = 0.1548$, $\phi^* = 2.6392$, $\eta^* = 0.5329$ (and ν^* fixed to 4.1). The a_μ, b_μ are the location parameters, a_ϕ, b_ϕ denote the scale parameters, a_η, b_η are the skewness parameters, and a_ν, b_ν are the shape parameters. The μ^*, ϕ^*, η^* and ν^* are the unconditional location, scale, skewness, and shape parameters. Next, through the UniGASSim function, using seed 12345 in the code, the value of the scale parameter b2 = $b_\phi = 0.9788$ from the fit is set a priori as the true parameter for the MCS process as presented in Table 2. This value and all the other estimates (with the ν^* that is fixed to 4.1) from the uGASFit object are used in the code to generate three sets of N = 30,000 return observations. Next, following the previous trimming pattern, $N = \{17{,}000; 18{,}000; 19000\}$ datasets are used under each of the seven assumed errors, as shown in Table 2. The full code is available from the authors on request.

Now, comparing RMSE and SE for $\hat{b}_\phi$ in Table 2, the AST1 error assumption outperforms the other six assumed errors in efficiency and precision with the least values as N tends to the peak. For bias comparison, the AST1 also outperforms the rest, with the lowest absolute value of bias as N tends to the peak. The relative efficiency (RE)

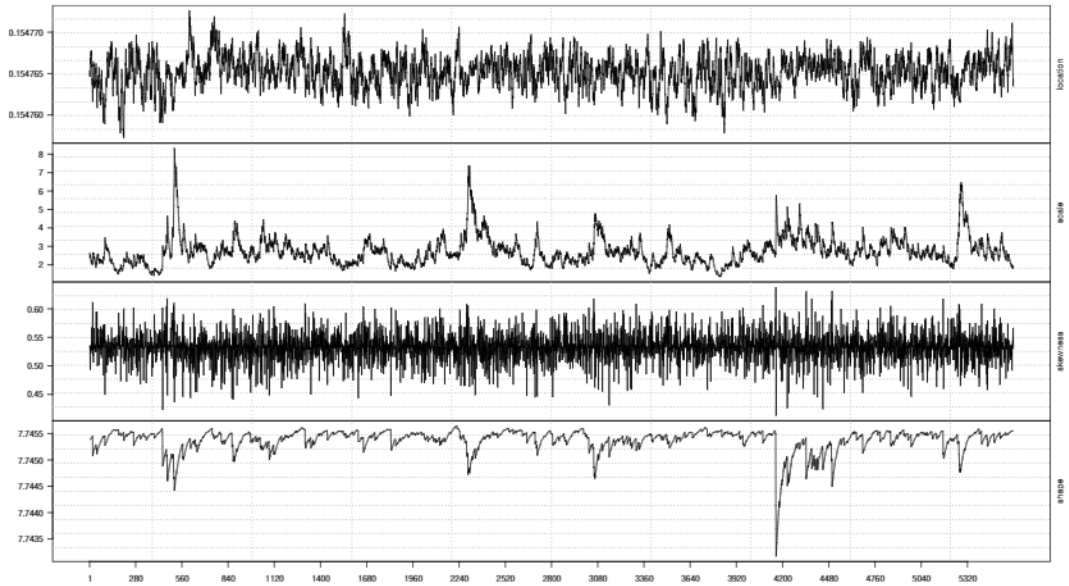


Figure 7. GAS-AST1: Time-varying location (top panel), scale (upper-middle panel), skewness (lower-middle panel) and shape (bottom panel) parameters.

Table 2. Simulation outcomes.

True model: GAS-AST1 True parameter: $b_\phi = 0.9788$						
Assumed error	N	$\hat{b}_\phi$	**RMSE**$_{\hat{b}_\phi}$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**TPR**$_{\hat{b}_\phi}$ **(95%)**
Gaussian	17,000	0.9444	0.0346	-0.0344	0.0036	91.66%
	18,000	0.9488	0.0302	-0.0300	0.0031	92.09%
	19,000	0.9493	0.0300	-0.0295	0.0051	92.13%
skew Gaussian	17,000	0.9522	0.0268	-0.0266	0.0034	92.42%
	18,000	0.9527	0.0264	-0.0262	0.0034	92.46%
	19,000	0.9531	0.0259	-0.0257	0.0033	92.50%
Student's t	17,000	0.9747	0.0051	-0.0042	0.0030	94.60%
	18,000	0.9746	0.0051	-0.0042	0.0030	94.59%
	19,000	0.9752	0.0046	-0.0036	0.0028	94.65%
skew Student's t	17,000	0.9744	0.0053	-0.0044	0.0030	94.57%
	18,000	0.9744	0.0053	-0.0044	0.0030	94.57%
	19,000	0.9750	0.0047	-0.0038	0.0028	94.63%
AST	17,000	0.9749	0.0049	-0.0039	0.0030	94.62%
	18,000	0.9744	0.0053	-0.0044	0.0029	94.57%
	19,000	0.9753	0.0045	-0.0035	0.0028	94.66%
AST1	17,000	0.9750	0.0049	-0.0039	0.0030	94.63%
	18,000	0.9749	0.0049	-0.0040	0.0029	94.62%
	19,000	0.9755	0.0043	-0.0033	0.0027	94.68%
ALD	17,000	0.9745	0.0054	-0.0044	0.0031	94.58%
	18,000	0.9740	0.0057	-0.0048	0.0031	94.54%
	19,000	0.9748	0.0049	-0.0040	0.0030	94.61%

Note: Simulation outcomes from the true model GAS-AST1.

comparison is not done here due to the weak $\sqrt{N}$ consistency of the RMSE under the AST and ALD assumed errors. In this study, consistency is termed "strong" ("weak") when the estimator's RMSE/SE value decreases as the sample size N increases without (with) distortion. To summarise, when the true error distribution is AST1, the AST1 assumed error distribution outperforms the other six error assumptions in efficiency and precision, and in overall performance. It is also observed from the outcomes of the $\text{SE}_{\hat{b}_\phi}$ for both true models in Tables 1 and 2 that the GAS model with the Gaussian error distribution is considerably consistent but not efficient when the true error is not Gaussian.

Furthermore, as revealed in Table 2 and as visualised in Figure 8, the MCS estimator $\hat{b}_\phi$ considerably recovers the true parameter b_ϕ at the 95% nominal recovery level, for each of the seven assumed errors. This suggests that the MCS experiments are well carried out with suitably valid results. However, as observed under the true model GAS-Student's t, it is also observed here that the non-Gaussian error assumptions perform better in the recovery than the Gaussian and skew-Gaussian error assumptions, as shown in the plots.

It can be seen from Figures 6 and 8, and in Tables 1 and 2 that the MCS estimator $\hat{b}_\phi$ performs slightly better at recovering the true parameter b_ϕ under the Gaussian and skew-Gaussian assumed errors for the true model GAS-Student's t than it is for true model GAS-AST1. Moreover, the true model GAS-Student's t displays better $\sqrt{N}$ consistency for the RMSE and SE metrics than the GAS-AST1 as revealed in the tables. Hence, these show the preference of the Student's t error distribution for the data generation process over the AST1 error distribution.

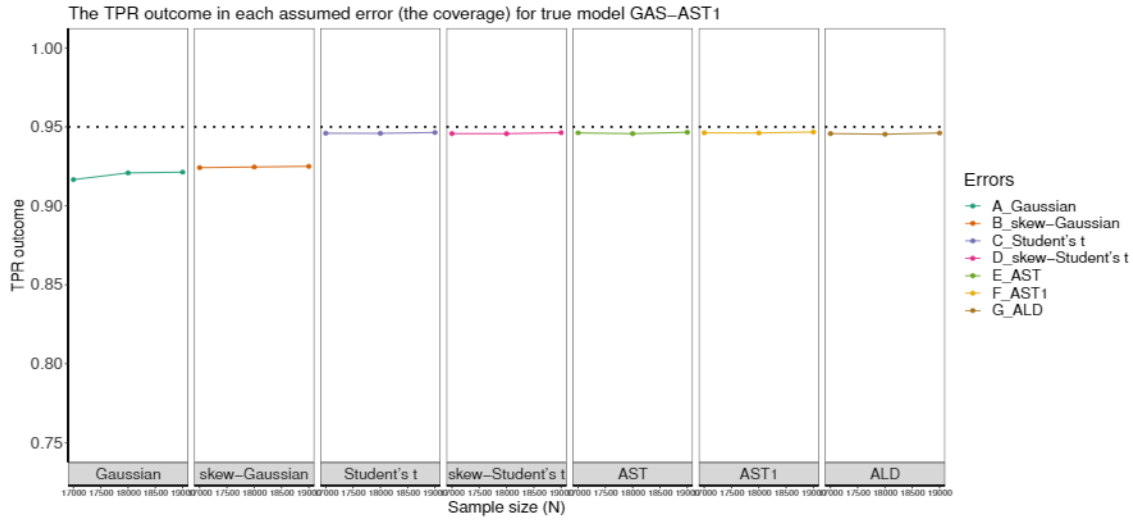


Figure 8. The true parameter recovery (TPR) outlook. The dotted line denotes the 0.95 (i.e., 95%) nominal recovery level.

3.6. Empirical Study

Here, the results of the simulation modelling are verified empirically by using the actual returns from the SA bond market index. Among the selected seven innovations, the most suitable for the GAS model to describe the market's returns for volatility persistence estimation is investigated. We transform the daily price data to the log returns by taking the log-difference of the value of the index as:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \times 100, \tag{13}$$

where r_t is the current returns, ln denotes the natural logarithm, and P_t is the daily closing bond price at time t.

3.6.1. Exploratory Data Analysis To begin with, the price and return series are first examined via exploratory data analysis (EDA) as shown in Figures 9 to 11. The EDA visually inspects the content of the dataset to show possible



Figure 9. Histograms of price (in Panel A) and returns (Panel B). The two colours merely demarcate the distribution where the price is < 200 (with blue colour) and above 200 (with an indianred colour). This is also used where the returns < 0.

outliers and other vital information. The histogram plot in Panel A of Figure 9 shows a skewed non-symmetric distribution that suggests non-stationarity in the price index, while Panel B reveals approximately symmetric or stationary distribution in the return series. Panel A of Figure 10 also reveals the non-stationarity in the price index as observed in the price series plot, while the return series plot in Panel B shows stationarity in the returns. Figure 10 further reveals four obvious extreme spikes in the distributions of price and returns during the period under study. These spikes are a result of severe global events and economic downturns that led to extreme market volatilities, notably the 2002/2003 SARS epidemic outbreak [73], the 2008 global financial crisis caused by the U.S. loans crisis, i.e., the defaults on consolidated mortgage-backed securities [3], the 2016 Brexit issues, and the 2020 COVID-19 pandemic outbreak. Figure 11 shows the price and return densities in Panels A and B respectively, while Panels C and D display the movements of the price and return distributions with their densities.

3.6.2. Tests for Autocorrelation and Heteroscedasticity Next, we fit the Autoregressive Moving Average (ARMA(u,v)) models (see [33]), with a Gaussian innovation distribution and a non-Gaussian (Student's t) innovation, to the stationary return series to check for the presence of autocorrelation (or serial correlation). The u and v are the orders of the AR and MA processes, respectively. Student's t innovation distribution is used because it adequately reflects the distribution of financial data [46]. Among the candidate ARMA(u,v) models examined, both ARMA(1,1) and ARMA(2,2) models as stated in Equations (14) and (15) are found adequate to remove possible autocorrelation in the return residuals. The autocorrelation function (ACF) and partial ACF (PACF) plots of the fit of the ARMA(2,2) model are displayed in Figure 12[‡‡]. Table 3 shows the results of the Weighted Ljung-Box (WLB) tests (see [31]) on the "standardised" and "standardised squared" residuals for fitting the two models. The p-values of the tests at lag order 9 for ARMA(1,1) and lag order 19 for ARMA(2,2) are all greater than 5% under each error distribution. Hence, we cannot reject the null hypothesis of "no serial correlation" in the market's returns. This implies that serial correlation is captured in the return residuals.

$$r_t = \phi_0 + \phi_1 r_{t-1} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t \tag{14}$$

$$r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \varphi_2 \varepsilon_{t-2} + \varphi_1 \varepsilon_{t-1} + \varepsilon_t \tag{15}$$

Through the outcomes from the autocorrelation test, we carry out Engle's ARCH test (see [28]) using the Portmanteau-Q (PQ) and Lagrange Multiplier (LM) tests to examine the existence of heteroscedasticity in the

[‡‡]The ACF and PACF diagrams were plotted using the auto.arima() function in the R forecast package developed by Hyndman and Khandakar [47].

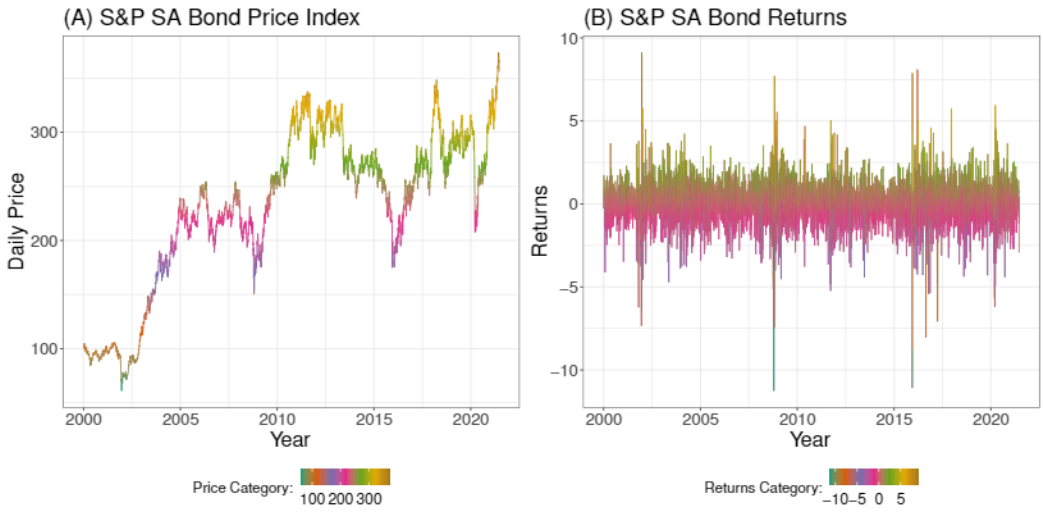


Figure 10. Distributions of price (in Panel A) and returns (Panel B).

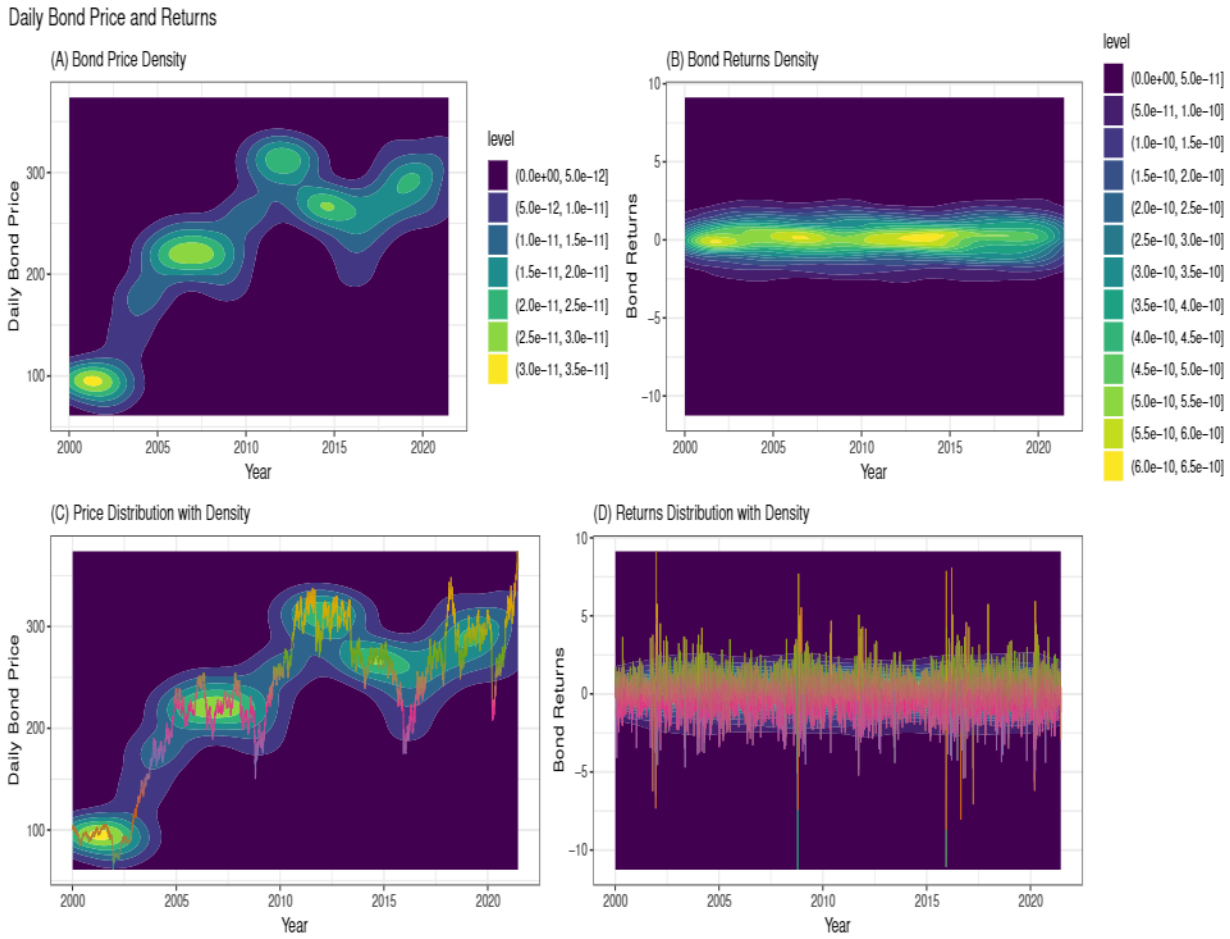


Figure 11. Price and return densities in Panels A and B respectively, while Panels C and D display the price and return distributions with their densities.

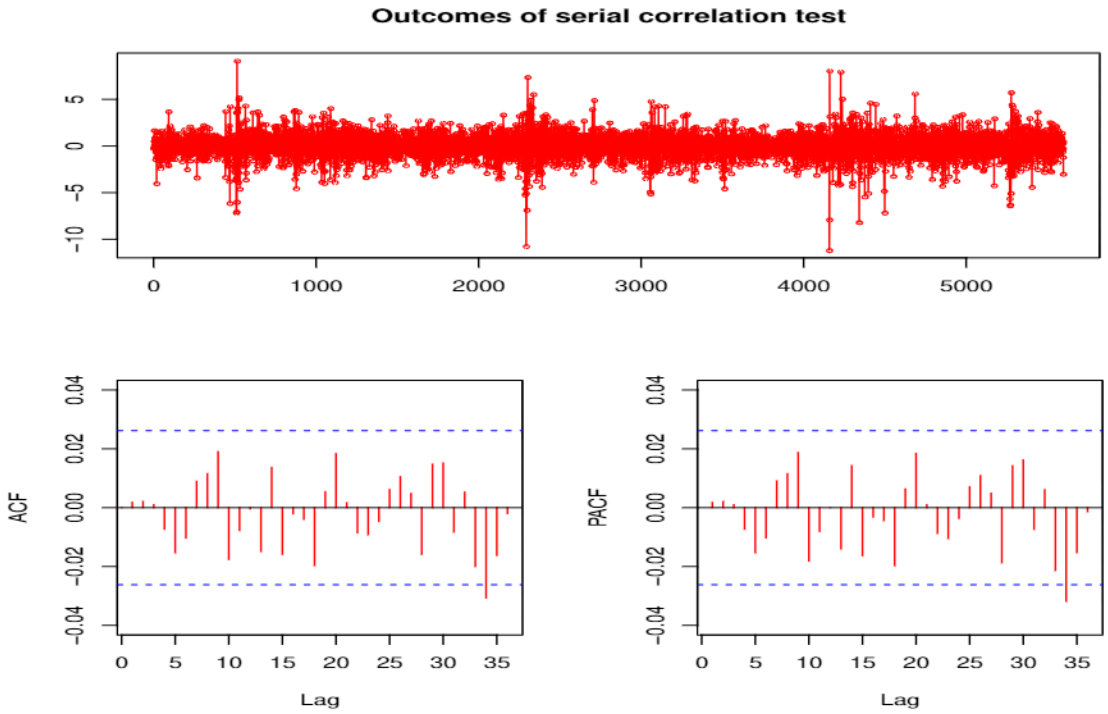


Figure 12. Outcomes of serial correlation test.

Table 3. ARMA(1,1) and ARMA(2,2) models' outcomes on Bond returns.

		Panel A ARMA(1,1)			Panel B ARMA(2,2)	
		Gaussian	Student's t		Gaussian	Student's t
Test on standardised residuals	WLB(9)	2.2258	3.2100	WLB(19)	6.2949	10.3830
	p-value(9)	(0.9728)	(0.8563)	p-value(19)	(0.9573)	(0.4060)
Test on standardised squared residuals	WLB(9)	6.5740	7.4880	WLB(19)	6.4740	7.5450
	p-value(9)	(0.2376)	(0.1617)	p-value(19)	(0.2473)	(0.1577)

Note: WLB denotes the Weighted Ljung-Box test, where "(9)" and "(19)" are lags 9 and 19, respectively.

residuals. These tests are carried out based on the null hypothesis of homoscedasticity in the residuals of an Autoregressive Integrated Moving Average (ARIMA) model. The tests are executed using the arima() function, in R package aTSA, for ARMA(1,1) and ARMA(2,2) models. The results of both tests reveal highly significant p-values of 0 from lag order 4 to 24 as presented in Panels A and B of Table 4. Thus, the null hypothesis of "no heteroscedasticity" in the residuals is rejected, which indicates the presence of volatility clustering. Volatility clustering implies that big (small) changes in returns tend to be followed by big (small) changes of either sign [57]. Based on the outcomes, we fit the GAS model with each of the seven error distributions to filter out the heteroscedasticity in the returns.

3.6.3. Selection of the Most Appropriate Error Assumption Next, we use the Akaike information criterion (AIC) and Bayesian information criterion (BIC) (see [54, 33] for details) to select the optimal (or the most adequate) assumed error distribution to describe the market's returns for volatility persistence estimation through the fit of

Table 4. Engle's ARCH test results.

	Panel A ARMA(1,1) model				Panel B ARMA(2,2) model			
	PQ test		LM test		PQ test		LM test	
Lag order	**PQ**	**P-value**	**LM**	**P-value**	**PQ**	**P-value**	**LM**	**P-value**
4	904	0	4729	0	888	0	4658	0
8	1127	0	2196	0	1104	0	2184	0
12	1400	0	1361	0	1377	0	1354	0
16	1517	0	1015	0	1497	0	1010	0
20	1575	0	807	0	1556	0	803	0
24	1624	0	666	0	1602	0	663	0

Note: PQ is the Portmanteau-Q statistic, and LM is the Lagrange-Multiplier statistic.

the GAS model. The error distribution with the lowest values of the information criteria is selected as the most suitable to describe the market for volatility persistence modelling.

Table 5 shows that the GAS parameter estimates $\hat{\kappa}_\phi$, $\hat{a}_\phi$ and $\hat{b}_\phi$ are statistically significant at 1% level under the seven assumed innovations. The only exception to these is the $\hat{\kappa}_\phi$ that is insignificant under the Student's t distribution. The comparisons of the seven assumed innovations are carried out with a time-varying scale parameter in the GAS modelling, and the empirical outcomes of the information criteria as presented in Table 5 show that the values of the AIC and BIC are lowest under the AST1 innovation. Hence, the AST1 innovation assumption outperforms the other six assumed innovations. This outcome is consistent with the results of the MCS studies. The result implies that the persistence of volatility in the returns of the SA bond market can be most suitably described through the GAS model when the assumed innovation is AST1. As shown in the table, the volatility persistence estimate $\hat{b}_\phi$ in **B** under this optimal AST1 error is 0.9790, indicating a considerably high persistence in the returns.

Table 5. Empirical outcomes of GAS modelling on SA Bond returns data.

Panel A	**Normal**	**skew-Normal**	**Student's t**	**skew-Student's t**	**AST**	**AST1**	**ALD**
$\hat{\kappa}_\phi$	0.0071*	0.0033*	0.0011	0.0034*	0.0206*	0.0204*	0.0053*
$\hat{a}_\phi$	0.0872*	0.0221*	0.1653*	0.0395*	0.0412*	0.0411*	0.0587*
$\hat{b}_\phi$	0.9804*	0.9819*	0.9790*	0.9804*	0.9788*	0.9790*	0.9784*
AIC	17936.36	17900.84	17682.49	17672.13	17668.28	17666.39	17910.63
BIC	17962.88	17933.99	17715.64	17711.91	17714.69	17706.17	17943.78

Note: The "$*$" indicates 1% level of significance. The table shows the outcomes of the fit of the GAS model for time-varying scale parameters.

From the empirical outcomes of the GAS model fitted with the AST1 innovation in Table 5, the equation of the GAS model for the time-varying scale (or volatility) parameter can be stated as:

$$\vartheta_{t+1} = 0.0204 + 0.0411 s_t + 0.9790 \vartheta_t. \tag{16}$$

4. Discussion and Summarised Conclusion

In conclusion, it is observed from the MCS studies that using the UniGASSim function (with full specification of the GAS model) approach, the AST1 consistently emerges as the optimal (or the most adequate) assumed error distribution to use with the GAS model for the returns' volatility persistence estimation when the underlying distribution of errors is unknown. The result was verified empirically. The persistence of volatility under the optimal assumed innovation revealed considerably high persistence in the volatility of the SA bond market's returns.

The conclusion proceeds with answers provided to the research questions. In answer to the first question, AST1 is the most appropriate conditional distribution, among the seven assumed error distributions from the GAS process simulation, to estimate the volatility persistence. Second, the volatility persistence estimator $\hat{b}_\phi$ displays considerably strong $\sqrt{N}$ consistency. To break this down, there is strong $\sqrt{N}$ consistency for the estimator $\hat{b}_\phi$ under all the seven assumed error distributions in the Student's t true error distribution (or true model GAS-Student's t) for both RMSE and SE in Table 1. The AST1 true error distribution or true model GAS-AST1, in Table 2, also shows a considerably strong $\sqrt{N}$ consistency under the error assumptions, except the weak consistency in the AST and ALD assumed errors for RMSE metric, and under the Gaussian assumed error for SE metric. Third, when compared to competing assumed errors, the estimator $\hat{b}_\phi$ of the optimal AST1 error is the best overall in efficiency and precision in terms of RMSE, SE and RE, and in the absolute value of bias under the two true error distributions, as revealed in Tables 1 and 2.

Fourth, with the TPR measure used as a proxy for the coverage of the simulation experiments, the study's outcomes revealed a satisfactory performance of the GAS volatility persistence estimator $\hat{b}_\phi$ in recovering the true parameter b_ϕ at the 95% nominal recovery level. This indicates that the MCS experiments are well carried out with suitably valid outcomes under the two true models. However, the non-Gaussian error assumptions performed better in the recovery than the Gaussian and skew-Gaussian assumed errors. The study also observed from the outcomes of the TPR measure and the $\sqrt{N}$ consistency of the RMSE and SE, that the Student's t is a preferred error distribution for data generation than the AST1.

This illustrative study through the GAS simulation modelling focused on time-varying shape parameter or degree of freedom ν_t, where the "shape = TRUE". Interested users and readers can try the other approach with "shape = FALSE" (i.e., constant shape parameter or degree of freedom ν). Any argument specified as "FALSE" in the UniGASSpec function for the GAS specification is constrained to 0 (see [6] or use help(UniGASSpec) in the R GAS package for details).

5. Concluding Remarks

This novel study presents a step-by-step simulation structure for selecting an optimal assumed distribution for the error term to model volatility persistence through the observation-driven GAS model using the GAS package, with relevant findings. The structure discloses an organised approach to the Monte Carlo simulation (MCS) experiment that includes "background of the study (optional), defining the aim of the study, specifying the research questions, method of implementation, and summarised conclusion". The method of implementation is a process that consists of writing the simulation code, setting the seed, setting the true parameter a priori, data generation, and performance evaluation through meta-statistics. This handy structure is exemplified using financial returns data, which makes it easily applicable to users for effective MCS studies.

Some of the works that have been used in the literature to generally describe the simulation approach to solving problems include [59, 17, 30], and [71], among others, but none of these has directly designed a simulation structure for volatility modelling involving autoregressive models. Specifically, whilst the empirical approach of volatility modelling involving autoregressive models is pervasive in the literature, an organised simulation framework to select an optimal conditional distribution for the error term to model volatility has been rarely studied. Hence, this tutorial study is motivated by the inclination to derive a flexible but robust simulation framework for improved volatility persistence estimation using the autoregressive GAS model. The findings in this paper may help current practitioners and upcoming researchers in finance and other sectors involved with volatility modelling to better understand the relevant simulation steps necessary for an improved volatility persistence estimation through this model. Hence, this paper has contributed to addressing inadequate literature on the relevant simulation steps needed to select a suitable conditional error distribution for modelling the persistence of volatility.

Among the findings, this study used both the fat-tails and $\sqrt{N}$ consistency experiments to show that the GAS model with a lower unconditional shape parameter or degree of freedom value ($\hat{\nu}^* = 4.1$) can generate a dataset that adequately reflects the behaviour of financial time series data, relevant for volatility modelling. This dynamic approach, to the best of the authors' knowledge, has not been used in prior works, and the finding could be a

relevant guide to users on how to obtain appropriate simulated financial time series data through the GAS model. It is also observed that to determine the sample size that is adequate for a particular simulation, focus should be put on the level of accuracy or precision (in this case, on efficiency and/or consistency) of the target statistic of interest. If a set of sample sizes yields the same efficiency outcomes, the outcome with the best consistency in the set can be used for a final decision on sample size determination.

Furthermore, in line with the findings of Gilli et al. [35] and Feng and Shi [30], the outcomes from this study also showed that increasing the samples or sample size does not always reduce the sampling error (RMSE). The RMSE results may be strongly or weakly $\sqrt{N}$ consistent. Moreover, it is observed that the volatility persistence estimator of the GAS model displays appreciable $\sqrt{N}$ consistency. In addition, consistent with the outcomes of other autoregressive models like the GARCH models (see [30, 68]), this study also revealed that the GAS model with the Gaussian error distribution is consistent but not efficient when the true error distribution is not Gaussian. Also, the outcomes of the illustrative studies showed that the AST1 error assumption is the most suitable among the seven assumed errors to describe the financial returns for volatility persistence modelling. The application of this (AST1) error distribution with the GAS model revealed a considerably high persistence of volatility in the SA bond returns. In a broader context, the outcome of this robust estimation approach using the AST1 fitted with the GAS specification could be useful to investors and financial practitioners in enhancing the accuracy of their risk measurement for investment decisions and portfolio management.

Lastly, the true parameter recovery (TPR) measure was used as a proxy for the coverage of the MCS experiments. This measure is a robust but flexible presentation of the concept of the estimator's coverage, and it is easy to use for measuring the level of recovery of the true parameter value by the MCS estimates. Its outcomes revealed considerably good performance of the MCS estimator at recovering the true parameter. Hence, in the absence of a reliable and accurate means of calculating the coverage of a simulation experiment involving the GAS or any other model, this study has further largely contributed to providing for such calculation through the proxy TPR measure. It is expected that this robust structure will adequately guide interested researchers on MCS experiments using the GAS model and package for selecting an optimal assumed innovation for volatility modelling in finance and other areas.

5.1. Limitations in the Study

The following limitations are noticed in this study. First, MCS is somewhat complex and time-demanding, hence it requires adequate planning with critical and analytical reasoning to be properly executed. Second, the quality of the simulation modelling may depend on the quality of the model used. Third, implementing the experiments may sometimes involve large computational time. Furthermore, this study shows that a considerably large sample size may sometimes be required to obtain a reasonable output approximation. It is also noticed that the GAS package does not have a provision to test for serial correlation in the returns, hence we used the Weighted Ljung-Box (WLB) statistic through the R rugarch package and the auto.arima() function in the R forecast package to carry out the test. Lastly, since the package also does not make provision for estimating the coverage of the MCS experiments, this study used the TPR measure as a proxy for the coverage, and it is observed that the MCS estimates considerably recovered the true parameters.

5.2. Future Research Interest

For future work, the authors intend to extend the simulation structure ideas to volatility forecasting through the model.

Acknowledgement

The authors are grateful to Thomson Reuters Datastream for providing the data, and to the University of the Witwatersrand and the University of Venda for their resources.

Table 6. The $\sqrt{N}$ consistency outcomes of datasets with different unconditional shape parameters $\nu^* = \{4.1, 8.2, 10.3\}$.

True parameter: $b_\phi = 0.9831$

Panel (A)		$\nu^* = 4.1$				$\nu^* = 8.2$				$\nu^* = 10.3$			
	N	$\hat{b}_\phi$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**RMSE**$_{\hat{b}_\phi}$	$\hat{b}_\phi$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**RMSE**$_{\hat{b}_\phi}$	$\hat{b}_\phi$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**RMSE**$_{\hat{b}_\phi}$
	30,000	0.9689	-0.0142	0.0026	0.0144	0.9674	-0.0156	0.0026	0.0158	0.9669	-0.0162	0.0025	0.0164
Gaussian	31,000	0.9696	-0.0135	0.0025	0.0137	0.9674	-0.0157	0.0026	0.0159	0.9660	-0.0170	0.0025	0.0172
	32,000	0.9707	-0.0123	0.0022	0.0125	0.9667	-0.0164	0.0026	0.0166	0.9657	-0.0173	0.0025	0.0175
	30,000	0.9691	-0.0140	0.0026	0.0142	0.9675	-0.0156	0.0026	0.0158	0.9669	-0.0161	0.0025	0.0163
skew	31,000	0.9698	-0.0133	0.0025	0.0135	0.9842	0.0011	0.0014	0.0018	0.9661	-0.0170	0.0025	0.0171
Gaussian	32,000	0.9709	-0.0122	0.0022	0.0124	0.9667	-0.0164	0.0026	0.0166	0.9658	-0.0173	0.0025	0.0174
	30,000	0.9794	-0.0037	0.0023	0.0044	0.9790	-0.0040	0.0021	0.0046	0.9789	-0.0042	0.0021	0.0047
Student's t	31,000	0.9792	-0.0039	0.0023	0.0045	0.9792	-0.0038	0.0021	0.0044	0.9784	-0.0047	0.0021	0.0051
	32,000	0.9805	-0.0025	0.0021	0.0033	0.9785	-0.0045	0.0021	0.0050	0.9781	-0.0050	0.0021	0.0054
	30,000	0.9792	-0.0039	0.0024	0.0046	0.9788	-0.0043	0.0021	0.0048	0.9788	-0.0043	0.0021	0.0048
skew	31,000	0.9792	-0.0038	0.0024	0.0045	0.9793	-0.0038	0.0021	0.0043	0.9784	-0.0046	0.0021	0.0051
Student's t	32,000	0.9806	-0.0025	0.0022	0.0033	0.9785	-0.0045	0.0021	0.0050	0.9783	-0.0048	0.0021	0.0052

Panel (B)		$\nu^* = 4.1$				$\nu^* = 8.2$				$\nu^* = 10.3$			
	N	$\hat{b}_\phi$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**RMSE**$_{\hat{b}_\phi}$	$\hat{b}_\phi$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**RMSE**$_{\hat{b}_\phi}$	$\hat{b}_\phi$	**Bias**$_{\hat{b}_\phi}$	**SE**$_{\hat{b}_\phi}$	**RMSE**$_{\hat{b}_\phi}$
	17,000	0.9589	-0.0242	0.0059	0.0249	0.9583	-0.0248	0.0041	0.0251	0.9562	-0.0268	0.0039	0.0271
Gaussian	18,000	0.9620	-0.0210	0.0051	0.0217	0.9614	-0.0217	0.0038	0.0220	0.9569	-0.0262	0.0038	0.0265
	19,000	0.9643	-0.0187	0.0043	0.0192	0.9609	-0.0221	0.0036	0.0224	0.9599	-0.0232	0.0035	0.0234
	17,000	0.9604	-0.0227	0.0058	0.0234	0.9582	-0.0248	0.0041	0.0252	0.9564	-0.0266	0.0039	0.0269
skew	18,000	0.9630	-0.0200	0.0050	0.0206	0.9614	-0.0217	0.0038	0.0220	0.9570	-0.0260	0.0038	0.0263
Gaussian	19,000	0.9650	-0.0180	0.0042	0.0185	0.9609	-0.0221	0.0036	0.0224	0.9601	-0.0230	0.0035	0.0233
	17,000	0.9775	-0.0056	0.0034	0.0066	0.9805	-0.0025	0.0027	0.0037	0.9773	-0.0057	0.0029	0.0064
Student's t	18,000	0.9780	-0.0051	0.0032	0.0061	0.9806	-0.0025	0.0026	0.0036	0.9775	-0.0056	0.0028	0.0063
	19,000	0.9781	-0.0049	0.0031	0.0058	0.9797	-0.0033	0.0026	0.0042	0.9770	-0.0060	0.0028	0.0067
	17,000	0.9775	-0.0056	0.0034	0.0065	0.9805	-0.0025	0.0027	0.0037	0.9774	-0.0057	0.0029	0.0064
skew	18,000	0.9779	-0.0051	0.0032	0.0061	0.9806	-0.0025	0.0026	0.0036	0.9775	-0.0056	0.0028	0.0063
Student's t	19,000	0.9782	-0.0048	0.0031	0.0057	0.9797	-0.0033	0.0026	0.0042	0.9770	-0.0060	0.0028	0.0066

Note: The $\sqrt{N}$ consistency outcomes of datasets with different unconditional shape parameters or degrees of freedom $\nu^* = \{4.1, 8.2, 10.3\}$ for sample sizes $N = \{30{,}000; 31{,}000; 32{,}000\}$ in Panel A, and $N = \{17{,}000; 18{,}000; 19{,}000\}$ in Panel B.

REFERENCES

1. K. Agarwal, *The Monte Carlo simulation: Understanding the basics*, 2022. https://tinyurl.com/3vepmtv3.
2. W. Alanya-Beltran, *Modelling stock returns volatility with dynamic conditional score models and random shifts*. Finance Research Letters, vol. 45, 102121, 2022.
3. K. Amadeo, *The stock market crash of 2008*, 2022. https://www.thebalance.com/stock-market-crash-of-2008-3305535.
4. V. Arago, and F. Izquierdo, *GARCH models with changes in variance: An approximation to risk measurements*. Journal of Asset Management, vol. 4, pp. 277–287, 2003.
5. D. Ardia, K. Boudt, and L. Catania, *Downside risk evaluation with the R package GAS*. The R Journal, vol. 10, pp. 410–421, 2018. https://doi.org/10.32614/RJ-2018-064.
6. D. Ardia, K. Boudt, and L. Catania, *Generalized autoregressive score models in R: The GAS package*. Journal of Statistical Software, vol. 88, pp. 1–28, 2019. https://doi.org/10.18637/jss.v088.i06.
7. R. Beg, and S. Anwar, *Detecting volatility persistence in GARCH models in the presence of the leverage effect*. Quantitative Finance, vol. 14, pp. 2205–2213, 2014. https://doi.org/https://doi.org/10.1080/14697688.2012.716162.
8. F. Blasques, D.D. Creal, P. Janus, S.J. Koopman, A. Lucas, M. Scharth, and B. Schwaab, *Generalized autoregressive score models for time-varying parameters: new models and applications*. Workshop ISF 2014 Rotterdam, pp. 1–79, 2014.
9. F. Blasques, P. Gorgi, and S.J. Koopman, *Accelerating GARCH and score-driven models: Optimality, estimation and forecasting*. Estimation and Forecasting, pp. 1–37, 2017. https://doi.org/10.2139/ssrn.3004598.
10. F. Blasques, S. Koopman, and A. Lucas, *Maximum likelihood estimation for generalized autoregressive score models*. TI 2014-029/III, Tinbergen Institute Discussion Paper, pp. 1–67, 2014.
11. F. Blasques, S. Koopman, and A. Lucas, *Stationarity and ergodicity of univariate generalized autoregressive score*. Electronic Journal of Statistics, vol. 8, pp. 1088–1112, 2014.
12. T. Bollerslev, *Generalized autoregressive conditional heteroskedastic*. Journal of Econometrics, vol. 31, pp. 307–327, 1986.
13. L. Catania, D. Ardia, and K. Boudt, *Generalized autoregressive score models. Package 'GAS' Version 0.3.3*, pp. 1–61, 2020. https://doi.org/doi:10.18637/jss.v088.i06.
14. L. Catania, D. Ardia, and K. Boudt, *Generalized autoregressive score models. Package 'GAS' Version 0.3.4*, 2022. https://cran.r-project.org/web/packages/GAS/GAS.pdf.
15. L. Catania, and A. Billé, *Dynamic spatial autoregressive models with autoregressive and heteroskedastic disturbances*. J. Appl. Econom., vol. 32, pp. 1178–1196, 2017. https://doi.org/https://doi.org/10.1002/jae.2565.
16. P. Chalmers, *Introduction to Monte Carlo simulations with applications in R using the SimDesign package*, pp. 1–46, 2019. https://philchalmers.github.io/SimDesign/pres.pdf.
17. R.P. Chalmers, and M.C. Adkins, *Writing effective and reliable Monte Carlo simulations with the SimDesign package*. The Quantitative Methods for Psychology, vol. 16, pp. 248–280, 2020. https://doi.org/10.20982/tqmp.16.4.p248.
18. G. Covarrubias, *Persistence, sudden changes, and modeling volatility of financial time series*. Doctor of philosophy, Texas Tech University, 2004.
19. D. Creal, S.J. Koopman, and A. Lucas, *A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations*. Journal of Business and Economic Statistics, vol. 29, pp. 552–563, 2011. https://doi.org/10.1198/jbes.2011.10070.
20. D. Creal, S.J. Koopman, and A. Lucas, *Generalized autoregressive score models*. Aenorm, vol. 20, pp. 37–41, 2012.
21. D. Creal, S.J. Koopman, and A. Lucas, *Generalized autoregressive score models with applications*. Journal of Applied Econometrics, vol. 28, pp. 777–795, 2013. https://doi.org/10.1002/jae.1279.
22. D. Creal, B. Schwaab, S. Koopman, and A. Lucas, *Observation-driven mixed-measurement dynamic factor models with an application to credit risk*. Rev. Econ. Stat., vol. 96, pp. 898–915, 2014.
23. Danielsson, J. *Financial risk forecasting: the theory and practice of forecasting market risk with implementation in R and Matlab*, John Wiley \& Sons: Chichester, West Sussex, pp. 1–274, 2011.
24. Datastream *Thomson reuters datastream. [Online]. Available at: Subscription Service (Accessed: June 2021)*, 2021. https://solutions.refinitiv.com/datastream-macroeconomic-analysis?.
25. R. Davidian, *Simulation studies in statistics*. North Carolina State University, North Carolina, pp 1–6, 2016.
26. F.X. Diebold, and J.A. Lopez, *Modeling volatility dynamics*. in Macroeconometrics: Developments, Tensions, and Prospects, Hoover, K.D., Ed., Springer Netherlands: Dordrecht, pp. 427–472, 1995. https://doi.org/10.1007/978-94-011-0669-6_11.
27. M. Duda, and H. Schmidt, *Evaluation of various approaches to value at risk*. PhD thesis, Lund University, 2009.
28. R.F. Engle, *Autoregressive conditional heteroscedacity with estimates of variance of United Kingdom inflation*. Econometrica, vol. 50, pp. 987–1008, 1982.
29. R. Engle, and J. Russell, *Autoregressive conditional duration: A new model for irregularly spaced transaction data*. Econometrica, vol. 66, pp. 1127–1162, 1998.
30. L. Feng, and Y. Shi, *A simulation study on the distributions of disturbances in the GARCH model*. Cogent Economics and Finance, vol. 5, 1355503, 2017. https://doi.org/10.1080/23322039.2017.1355503.
31. T.J. Fisher, and C.M. Gallagher, *New weighted portmanteau statistics for time series goodness of fit testing*. Journal of the American Statistical Association, vol. 107, pp. 777–787, 2012. https://doi.org/10.1080/01621459.2012.688465.
32. W.G. Foote, *Financial engineering analytics: A practice manual using R*, 2018. https://bookdown.org/wfoote01/faur/.
33. A. Ghalanos, *Introduction to the rugarch package. (Version 1.3-8)*, 2018. https://cran.r-project.org/web/packages/rugarch/vignettes.
34. A. Ghalanos, *Univariate GARCH models. Package 'rugarch' version 1.4-7*, 2022. https://cran.r-project.org/web/packages/rugarch/rugarch.pdf.
35. M. Gilli, D. Maringer, and E. Schumann, *Generating random numbers*. in Numerical Methods and Optimization in Finance, Janco, C., Ed., Academic Press, pp. 103–132, 2019. https://doi.org/doi:10.1016/B978-0-12-815065-8.00017-0.

36. R. Giovanni, *Essential Monte Carlo analysis*. in Elements of Numerical Mathematical Economics with Excel, Romer, B., Ed., Elsevier Academic Press, pp. 763–795, 2020. https://doi.org/10.1016/B978-0-12-817648-1.00014-1.
37. P. Gorgi, P. Hansen, P. Janus, and S. Koopman, *Realized Wishart-GARCH: a score-driven multi-asset volatility model*. J. Financ. Economet., vol. 17, pp. 1–32, 2019. https://doi.org/https://doi.org/10.1093/jjfinec/nby007.
38. M.F.C. Haddad, S. Blazsek, P. Arestis, F. Fuerst, and H.H. Sheng, *The two-component Beta-t-QVAR-M-lev: a new forecasting model*. Financial Markets and Portfolio Management, vol. 37, pp. 379–401, 2023. https://doi.org/https://doi.org/10.1007/s11408-023-00431-4.
39. K.A. Hallgren, *Conducting simulation studies in the R programming environment*. Tutorials in Quantitative Methods for Psychology, vol. 9, pp. 43–60, 2013. https://doi.org/10.20982/tqmp.09.2.p043.
40. A.C. Harvey, *Dynamic models for volatility and heavy tails: With applications to financial and economic time series*, Cambridge University Press: New York, USA, pp. 1–261, 2013. https://doi.org/https://doi.org/10.1017/CBO9781139540933.
41. A. Harvey, and T. Chakravarty, *Beta-t-EGARCH. Cambridge working papers in economics 0840*. Faculty of Economics, University of Cambridge, pp. 137, 2008.
42. A. Harvey, and G. Sucarrat, *EGARCH models with fat tails, skewness and leverage*. Computational Statistics and Data Analysis, vol. 26, pp. 320–338, 2014.
43. A. Harvey, and S. Thiele, *Testing against changing correlation*. J. Empir. Financ., vol. 38, pp. 575–589, 2016. https://doi.org/https://doi.org/10.1016/j.jempfin.2015.09.003.
44. M. Harwell, *A strategy for using bias and RMSE as outcomes in Monte Carlo studies in statistics*. Journal of Modern Applied Statistical Methods, vol. 17, pp. 1–16, 2018. https://doi.org/10.22237/jmasm/1551907966.
45. L. Hentschel, *All in the family nesting symmetric and asymmetric GARCH models*. Journal of Financial Economics, vol. 39, pp. 71–104, 1995. https://doi.org/10.1016/0304-405X(94)00821-H.
46. M.S. Heracleous, *Sample kurtosis, GARCH-t and the degrees of freedom issue*. EUR Working Papers, pp. 1–22, 2007.
47. R.J. Hyndman, and Y. Khandakar, *Automatic time series forecasting: The forecast package for R*. Journal of Statistical Software, vol. 27, pp. 1–22, 2008.
48. P. Janus, S.J. Koopman, and A. Lucas, *Long memory dynamics for multivariate dependence under heavy tails*. Journal of Empirical Finance, vol. 29, pp. 187–206, 2014. https://doi.org/10.1016/j.jempfin.2014.09.007.
49. A. Johansen, *Monte Carlo methods*. in International Encyclopedia of Education, Peterson, P., Baker, E., MCGaw, B., Eds., Elsevier Academic Press, pp. 296–303, 2010.
50. W. Kenton, *Monte Carlo simulation*, 2021. https://www.investopedia.com/terms/m/montecarlosimulation.asp.
51. S.J. Koopman, A. Lucas, and M. Zamojski, *Dynamic term structure models with score driven time varying parameters: estimation and forecasting*. NBP Working Paper, pp. 3–80, 2017.
52. S. Kotz, T. Kozubowski, and K. Podgorski, *The laplace distribution and generalizations*, Springer-Verlag: New York, pp. 1–441, 2001. https://doi.org/doi:10.1007/978-1-4612-0173-1.
53. R. Kwiatkowski, *Monte Carlo simulation — a practical guide*, 2022. https://towardsdatascience.com/monte-carlo-simulation-a-practical-guide-85da45597f0e.
54. F. Li, A.S. Cohen, S. Kim, and S. Cho, *Model selection methods for mixture dichotomous IRT models*. Applied Psychological Measurement, vol. 33, pp. 353–373, 2009.
55. Q. Li, *Three essays on stock market volatility*. Doctor of philosophy thesis, Utah State University, 2008.
56. C.H. Lin, and S.S. Shen, *Can the student-t distribution provide accurate value at risk?* Journal of Risk Finance, vol. 7, pp. 292–300, 2006. https://doi.org/10.1108/15265940610664960.
57. B. Mandelbrot, *The variation of certain speculative prices*. The Journal of Business, vol. 36, pp. 394–419, 1963. https://doi.org/10.1086/294632.
58. C.Z. Mooney, *Monte Carlo simulation*, first ed., Vol. 116, SAGE Publications: Thousand Oaks, CA, 1997.
59. T.P. Morris, I.R. White, and M.J. Crowther, *Using simulation studies to evaluate statistical methods*. Statistics in Medicine, vol. 38, pp. 2074–2102, 2019. https://doi.org/https://doi.org/10.1002/sim.8086.
60. B.D. Nelson, *Conditional heteroskedasticity in asset returns : A new approach*. Econometrica, vol. 59, pp. 347–370, 1991.
61. D. Oh, and A. Patton, *Time-varying systemic risk: Evidence from a dynamic copula model of CDS spreads*. Journal of Business and Economic Statistics, vol. 36, pp. 181–195, 2016. https://doi.org/doi:10.1080/07350015.2016.1177535.
62. A. Opschoor, P. Janus, A. Lucas, and D. Van Dijk, *New HEAVY models for fat-tailed realized covariances and returns*. Journal of Business and Economic Statistics, vol. 36, pp. 643–657, 2018. https://doi.org/10.1080/07350015.2016.1245622.
63. P. Owusu Junior, and I. Alagidede, *Risks in emerging markets equities: Time-varying versus spatial risk analysis*. Physica A: Statistical Mechanics and its Applications, vol. 542, 123474, 2020. https://doi.org/10.1016/j.physa.2019.123474.
64. T.L. Pedersen, *patchwork: The composer of plots*, 2020. https://cran.r-project.org/package=patchwork.
65. D. Qiu, *Alternative time series analysis. Package 'aTSA'*, 2015. https://cran.r-project.org/web/packages/aTSA/aTSA.pdf.
66. J. Russell, *Econometric modeling of multivariate irregularly-spaced high-frequency data*, 2001. University of Chicago, Graduate School of Business.
67. T. Rydberg, and N. Shephard, *Dynamics of trade-by-trade price movements: decomposition and models*. Journal of Financial Econometrics, vol. 1, no. 1, pp. 2-25, 2003.
68. T. Samuel, C. Chimedza, and C. Sigauke, *Simulation framework to determine suitable innovations for volatility persistence estimation: The GARCH approach*. J. Risk Financial Manag., vol. 16, no. 9, 392, 2023. https://doi.org/https://doi.org/10.3390/jrfm16090392.
69. S. Satchell, and J. Knight, *Forecasting volatility in the financial markets*, Elsevier, 2011.
70. J.J. Su, *On the oversized problem of Dickey-Fuller-type tests with GARCH errors*. Communications in Statistics: Simulation and Computation, vol. 40, pp. 1364–1372, 2011. https://doi.org/10.1080/03610918.2011.575502.

71. M.J. Sigal, and P.R. Chalmers, *Play it again: Teaching statistics with Monte Carlo simulation*. Journal of Statistics Education, vol. 24, pp. 136–156, 2016. https://doi.org/10.1080/10691898.2016.1246953.
72. C. Wang, R. Gerlach, R., and Q. Chen, *A semi-parametric realized joint value-at-risk and expected shortfall regression framework*. arXiv preprint arXiv:1807.02422, pp. 1–39, 2018. [1807.02422].
73. J. Weismer, *History as guide : Lessons from the 2002-2003 SARS outbreak*, 2020. https://www.mizuhogroup.com/americas/insights/2020/03/history-as-guide-lessons-from-the-2002-2003-sars-outbreak.html.
74. H. Wickham, M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester et al. *Welcome to the Tidyverse*. Journal of Open Source Software, vol. 4, pp. 1–6, 2019. https://doi.org/10.21105/joss.01686.
75. K.H. Yuan, X. Tong, and Z. Zhang, *Bias and efficiency for SEM with missing data and auxiliary variables: Two-stage robust method versus two-stage ML*. Structural Equation Modeling: A Multidisciplinary Journal, vol. 22, pp. 178–192, 2015. https://doi.org/10.1080/10705511.2014.935750.
76. D. Zhu, and J. Galbraith, *A generalized asymmetric Student-t distribution with application to financial econometrics*. Journal of Econometrics, vol. 157, pp. 297–305, 2010.