econometrics

MDPI

Article

Optimal Multi-Step-Ahead Prediction of ARCH/GARCH Models and NoVaS Transformation

Jie Chen and Dimitris N. Politis *

Department of Mathematics, University of California, San Diego, CA 92093, USA
* Correspondence: dpolitis@ucsd.edu

Received: 31 October 2018; Accepted: 21 July 2019; Published: 8 August 2019

Abstract: This paper gives a computer-intensive approach to multi-step-ahead prediction of volatility in financial returns series under an ARCH/GARCH model and also under a model-free setting, namely employing the NoVaS transformation. Our model-based approach only assumes *i.i.d* innovations without requiring knowledge/assumption of the error distribution and is computationally straightforward. The model-free approach is formally quite similar, albeit a GARCH model is not assumed. We conducted a number of simulations to show that the proposed approach works well for both point prediction (under L_1 and/or L_2 measures) and prediction intervals that were constructed using bootstrapping. The performance of GARCH models and the model-free approach for multi-step ahead prediction was also compared under different data generating processes.

Keywords: bootstrap; L_1 and L_2 measures; GARCH(1,1); NoVaS transformation; multi-step prediction; Monte Carlo simulation

---

1. Introduction

Multi-step-ahead prediction in a time series amounts to predicting a sequence of future values using only the values observed over a finite time interval. Examples of time series for which multi-step-ahead prediction is useful include crop yields, stock prices, traffic volume, and electrical power consumption. In the paper at hand, we focus on multi-step-ahead prediction of squared financial returns, which is related to the so-called volatility, i.e., the conditional expectation of the squared returns. A typical approach to solve this problem, known as multi-stage or iterated prediction, is to construct a single model from the past observed time series data and then apply the model step by step to predict its future values. The iterated method uses the predicted value of the current time step to determine its value in the next time step. However, empirical evidence points to the fact that multi-stage prediction is susceptible to the error accumulation problem, i.e., errors committed in the past are propagated into future predictions.

The benchmark model for financial returns has been the GARCH(1,1); see Bollerslev et al. (1992) and the references therein. Notably, the work in Andersen et al. (2006), page 811, stated that, beyond the one-step-ahead case, there is no analytical form of the predictive density for multi-step-ahead predictions of volatility in GARCH models. However, the analytical form of the multi-step ahead predictive probability density function of a GARCH(1,1) process under Gaussian or Student t innovations has been recently derived in the working paper of Abadir et al. (2018).

The work of Abadir et al. (2018) was an important breakthrough, but it hinged on knowing the error distribution that drove the GARCH(1,1) model. By contrast, the paper at hand proposes an alternative computer-intensive approach to multi-step-ahead prediction of the squared returns in ARCH/GARCH and related models that does not require knowledge of the error distribution. Furthermore, an analogous methodology can be employed in a model-free setting using the normalizing and variance-stabilizing transformation (NoVaS) approach; for more details on NoVaS, see Politis (2003, 2007, 2015). Notably, our method is not of the multi-stage/iterated type, and therefore eliminates the errors accumulation issue.

The remainder of the paper is organized as follows. Section 2 presents the proposed method for the optimal multi-step-ahead point predictions for ARCH/GARCH processes, as well as in a model-free setting using the NoVaS transformation. Section 3 addresses the associated methods for the construction of multi-step-ahead prediction intervals using bootstrapping. Section 4 illustrates the numerical performance of the proposed methods by means of simulated examples; some concluding remarks are provided in Section 5.

2. Optimal Multi-Step-Ahead Point Prediction

Consider data $X_1, \ldots, X_n$ from a zero mean and (strictly) stationary financial returns time series $\{X_t\}$. Our goal is to predict the future squared returns X^2_{n+h} for any $h \geq 2$; the case $h = 1$ was treated in Politis (2007, 2015).

Let $\mathcal{F}_n$ be a short-hand for the observed information set, i.e., $\mathcal{F}_n = \{X_t, 1 \leq t \leq n\}$. In the L_2 sense, the optimal predictor of X^2_{n+h} based on $\mathcal{F}_n$ is the conditional mean and given by:

$$\widehat{X^2_{n+h}} = E(X^2_{n+h}|\mathcal{F}_n). \tag{1}$$

Similarly, the optimal L_1 predictor is the conditional median:

$$\widehat{X^2_{n+h}} = Median(X^2_{n+h}|\mathcal{F}_n). \tag{2}$$

In the following parts of this section, we study the multi-step-ahead prediction in the nonlinear financial models ARCH/GARCH and in the NoVaS setting, which is an application of the model-free approach to financial returns data.

2.1. *L_2 Optimal Prediction for ARCH(p) and GARCH(1,1) Models*

Suppose the data follow the ARCH(p) process of Engle (1982) defined by the recursion:

$$X_t = \sigma_t \epsilon_t, \quad and \quad \sigma_t^2 = \alpha + a_1 X^2_{t-1} + \ldots + a_p X^2_{t-p} \tag{3}$$

where $\alpha \geq 0$, $a_j \geq 0$ for all $j = 1, \ldots, p$, and $\{\epsilon_t\} \sim i.i.d.\, N(0,1)$.

First, consider the simplest case $h = 2$. Based on ARCH(P) Model in (3), we can express X_{n+1} and X_{n+2} in the following way:

$$X_{n+1} = \epsilon_{n+1}\sqrt{\sigma^2_{n+1}}, \quad and \quad \sigma^2_{n+1} = \alpha + a_1 X^2_n + \ldots + a_p X^2_{n-p+1},$$
$$X_{n+2} = \epsilon_{n+2}\sqrt{\sigma^2_{n+2}}, \quad and \quad \sigma^2_{n+2} = \alpha + a_1 X^2_{n+1} + \ldots + a_p X^2_{n-p+2}.$$

Obviously, X_{n+1} can be easily written as a function of the past observations $X_n, \ldots, X_{n+1-p}$ and the unknown future error ϵ_{n+1}. Furthermore, we can also rewrite X_{n+2} to be a function of $X_n, \ldots, X_{n+1-p}$ and the unknown future errors ϵ_{n+1} and ϵ_{n+2}. The notations are as follows:

$$\begin{aligned} X_{n+1} =& \epsilon_{n+1}\sqrt{\alpha + a_1 X_n + a_2 X^2_{n-1} + \ldots + a_p X^2_{n-p+1}} \\ =& f_1(X_n, \ldots, X_{n-p+1}; \epsilon_{n+1}) \end{aligned} \tag{4}$$

and:

$$
\begin{aligned}
X_{n+2} =& \epsilon_{n+2}\sqrt{\alpha + a_1 X_{n+1}^2 + a_2 X_n^2 + ... + a_p X_{n-p+2}^2} \\
=& \epsilon_{n+2}\sqrt{\alpha + a_1 \epsilon_{n+1}^2(\alpha + a_1 X_n + a_2 X_{n-1}^2 + ... + a_p X_{n-p+1}^2) + a_2 X_n^2 + ... + a_p X_{n-p+2}^2} \\
=& f_2(X_n, ..., X_{n-p+1}; \epsilon_{n+1}, \epsilon_{n+2}).
\end{aligned}
\tag{5}
$$

Recursively, we can express X_{n+h} for any $h \geq 1$ as a function of past observations $\{X_1, \ldots, X_n\}$ and the unknown future innovations $\{\epsilon_{n+1}, \ldots, \epsilon_{n+h}\}$ in the form:

$$X_{n+h} = f_h(X_1, ..., X_n; \epsilon_{n+1}, \ldots, \epsilon_{n+h}). \tag{6}$$

Since $\{X_1, \ldots, X_n\}$ are given and known, we can write (6) simply as:

$$X_{n+h} = f_h(\epsilon_{n+1}, \ldots, \epsilon_{n+h}), \quad \text{for any } h \geq 1. \tag{7}$$

The squared financial returns can be rewritten as f_h^2. Based on the assumption that ϵ_t is *i.i.d* $N(0,1)$, the conditional distribution function $F_{f_h^2}$ of the future squared returns $f_h^2(\cdot)$ can be derived. Hence, the optimal predictor (conditional median for L_1 or conditional mean for L_2) of x_{n+h}^2 is easy to calculate by $F_{f_h^2}$.

Take $h = 1$ and $h = 2$ as examples. By (4) and (5), the L_2 optimal predictors of X_{n+1}^2 and X_{n+2}^2 are:

$$
\begin{aligned}
\widehat{X_{n+1}^2} =& E\{\epsilon_{n+1}^2(\alpha + a_1 X_n + a_2 X_{n-1}^2 + ... + a_p X_{n-p+1}^2)|\mathcal{F}_n\} \\
=& \alpha + a_1 X_n^2 + a_2 X_{n-1}^2 + ... + a_p X_{n-p+1}^2,
\end{aligned}
\tag{8}
$$

and:

$$
\begin{aligned}
\widehat{X_{n+2}^2} =& E[\epsilon_{n+2}^2(\alpha + a_1\sigma_{n+1}^2\epsilon_{n+1}^2 + a_2 X_n^2 + \ ... \ + a_p X_{n-p+2}^2)|X_1, ... , X_n] \\
=& \alpha + a_1\sigma_{n+1}^2 + a_2 X_n^2 + \ ... \ + a_p X_{n-p+2}^2
\end{aligned}
\tag{9}
$$

since $E(\epsilon_{n+1}^2|\mathcal{F}_n) = 1$ and $E(\epsilon_{n+2}^2|\mathcal{F}_n) = 1$ by assumption. First, we can note that $\widehat{X_{n+1}^2} = f_1^2(\epsilon_{n+1}^2 = 1)$ and $\widehat{X_{n+2}^2} = f_2^2(\epsilon_{n+1}^2 = 1, \epsilon_{n+2}^2 = 1)$. Actually, we can easily verify that the L_2 optimal predictor for any $h \geq 1$ is given by:

$$\widehat{X_{n+h}^2} = f_h^2(\epsilon_{n+1}^2 = 1, \ldots, \epsilon_{n+h}^2 = 1). \tag{10}$$

Note that because all ϵ_t's are independent of each other, as well as of the past values of the X series, the h-step-ahead predictor in (10) is equivalent to the method of multi-stage/iterated prediction that uses the predicted values of the current time step to determine its value in the next time step. However, for the L_1 case, because the median function is not a linear operator, this equivalence breaks down; see Section 2.3 in what follows.

Remark 1. *As already mentioned, the benchmark for fitting financial returns is the GARCH(1,1). Nevertheless, a GARCH(1,1) model is tantamount to an ARCH(p) model with $p = \infty$ and an exponentially-decreasing coefficient; see, e.g., Francq and Zakoian (2011). Because of the exponentially decrease of the ARCH coefficients, it is customary to approximate the GARCH(1,1) models with an ARCH(p) where p is finite, albeit large. In this sense, all the above results apply verbatim to a GARCH(1,1) process as well. Of course, in fitting a GARCH(1,1) model, the GARCH equation is used to fit four parameters, which are then expanded to the p coefficients of the approximating ARCH(p) model.*

2.2. L_2 Optimal Prediction for NoVaS

Given a sequence of observations $\{X_1, \ldots, X_n\}$, we can fit the data by a special application of the model-free methodology, NoVaS, which was introduced by Politis (2003, 2007) for stationary data in prediction of squared financial returns. Let us continue considering a zero mean and (strictly) stationary financial return time series $\{X_t\}$. The NoVaS methodology is trying to map the dataset $X_1, \ldots, X_n$ to an *i.i.d* Gaussian dataset $\{W_t, t \leq n\}$.

The starting point is the ARCH model defined by:

$$X_t = Z_t \sqrt{a + \sum_{i=1}^{p} a_i X_{t-i}^2} \tag{11}$$

under which, the residual:

$$\frac{X_t}{\sqrt{a + \sum_{i=1}^{p} a_i X_{t-i}^2}} \tag{12}$$

is thought of as perfectly normalized and variance-stabilized, as it is assumed to be $i.i.d.N(0,1)$, which is actually not true here. This ratio can be interpreted as an attempt to "Studentize" the return X_t by dividing with a time-localized measure of the standard deviation of X_t. However, there seems to be no reason to exclude the value of X_t from an empirical, causal estimate of the standard deviation of X_t; recall that a causal estimate is one involving present and past data only, i.e., the data $\{X_s, s \leq t\}$.

Hence, the work in Politis (2003) defined a new "Studentized" quantity as follows:

$$W_t := \frac{X_t}{\sqrt{\alpha s_{t-1}^2 + a_0 X_t^2 + \sum_{i=1}^{p} a_i X_{t-i}^2}} \quad \text{for} \quad t = p+1, p+2, \ldots, n. \tag{13}$$

In the above, s_{t-1}^2 is an estimator of $\sigma_X^2 = Var(X_1)$ based on the data up to (but not including[1]) time t; under the zero mean assumption for X_1, the natural estimator is $s_{t-1}^2 = (t-1)^{-1} \sum_{k=1}^{t-1} X_k^2$.

The definition in Equation (13) describes the proposed normalizing and variance-stabilizing transformation under which the data series $\{X_t\}$ is mapped to the new series $\{W_t\}$. The order p (≥ 0) and the vector of nonnegative parameters $(\alpha, a_0, \ldots, a_p)$ are chosen by the practitioner with the twin goals of normalization and variance stabilization.

Furthermore, the NoVaS transformation Equation (13) can be re-arranged to yield:

$$X_t = W_t \sqrt{\alpha s_{t-1}^2 + a_0 X_t^2 + \sum_{i=1}^{p} a_i X_{t-i}^2}. \tag{14}$$

Formally, the only real difference between the NoVaS of Equation (14) and the ARCH of Equation (11) is the presence of the term X_t^2 paired with the coefficient a_0. Replacing the term a in Equation (11) by the term αs_{t-1}^2 in Equation (14) is only natural since the former has, by necessity, units of variance; in other words, the term a in Equation (11) is not scale invariant, whereas the term α in Equation (14) is.

Given the assumed structure of the return series, the target of variance stabilization, which amounts to constructing a local estimator of scale for Studentization purposes, requires:

$$\alpha \geq 0, \ \ a_i \geq 0 \ \text{ for all } \ i \geq 0, \ \text{ and } \ \alpha + \sum_{i=0}^{p} a_i = 1. \tag{15}$$

---

1 The reason for not including time t in the variance estimator is for purposes of notational clarity, as well as the easy identifiability of the effect of the coefficient a_0 associated with X_t^2 in the denominator of Equation (13).

Equation (15) has the interesting implication that the $\{W_t\}$ series can be assumed to have a (unconditional) variance that is (approximately) unity. Nevertheless, note that p and $\alpha, a_0, \ldots, a_p$ must be carefully chosen to achieve a degree of conditional homoscedasticity as well; to do this, one must necessarily take p small enough, as well as α small enough or even equal to zero, so that a local (as opposed to global) estimator of scale is obtained. The work in Politis (2003) provided two structures for the a_i coefficients satisfying Equation (15). One is to let $\alpha = 0$ and $a_i = 1/(p+1)$ for all $0 \le i \le p$; this specification is called the *simple* NoVaS transformation and involves only one parameter, namely the order p, to be chosen by the practitioner. The other one is given by the *exponential* (decay) NoVaS, where $\alpha = 0$ and $a_i = c'e^{-ci}$ for all $0 \le i \le p$. The exponential scheme involves choosing two parameters: p and $c > 0$, since c' is determined by Equation (15). For more details of how to select the optimal parameters here, see Politis (2015).

The above Equation (14) can be used for one-step-ahead prediction in an analogous way to the ARCH/GARCH models already discussed. In fact, Equation (14) is formally analogous to an ARCH(p) model with *i.i.d.* errors given by $\epsilon_t = \frac{W_t}{\sqrt{1-a_0 W_t^2}}$. Hence, the construction of the previous subsection can be repeated to write X_{n+h} as some function of $\{X_1, \ldots, X_n\}$ and $\{W_t, t = 1, \ldots, h\}$ for any $h \ge 1$, i.e.,

$$X_{n+h} = f_h(X_1, \ldots, X_n; W_{n+1}, \ldots, W_{n+h}). \tag{16}$$

Since the data $\{X_1, \ldots, X_n\}$ are given, we can simplify (16) as:

$$X_{n+h} = f_h(W_{n+1}, \ldots, W_{n+h}). \tag{17}$$

In the L_2 sense, the optimal predictor of X_{n+h}^2 based on $\mathcal{F}_n$ is given by:

$$\widehat{X_{n+h}^2} = E(X_{n+h}^2 | \mathcal{F}_n) = E\{f_h^2(W_{n+1}, \ldots, W_{n+h}) | \mathcal{F}_n)\} \tag{18}$$

Since the W_t are *i.i.d*, we can get analogous results with those concerning the ARCH/GARCH models, i.e.,

$$\widehat{X_{n+h}^2} = f_h^2(W_{n+1}^2 = 1, \ldots, W_{n+h}^2 = 1)) \tag{19}$$

Therefore, for any $h \ge 1$, we can use similar ideas to that in the ARCH/GARCH cases to conduct multi-step-ahead prediction in NoVaS by approximating the conditional mean or median from their conditional distribution functions.

2.3. L_1 Optimal Prediction and Generalizations

We can generalize the above prediction method to an interesting class of prediction functions $g(\cdot)$, namely the power family where $g(x) = x^k$ for some fixed k and the power-absolute value family where $g(x) = |x|^k$. So far, we have worked on the prediction of X_{n+h}^2, that is $g(x) = x^2$. More generally, we can derive the best L_2 or L_1 predictor of $g(X_{n+h})$ given $\mathcal{F}_n$.

Regarding L_1 optimal prediction, it was already mentioned that it is *not* equivalent to the multi-stage/iterated approach. For the $h = 2$ case, we can easily get an analytic formula of the conditional distribution of $g(X_{n+h})$ assuming an ARCH(p) model with normal errors. An approximate analytic form of the conditional distribution function of X_{n+2}^2 is given as follows:

$$F_{f_2^2}(x | \{X_n, \ldots, X_1\}) = \frac{\gamma(\frac{1}{2}, \frac{x}{2A})}{B\sqrt{\pi}}, \quad x > 0$$

where:

$$A = \alpha + a_2 X_n^2 + a_3 X_{n-1}^2 + \ldots + a_p X_{n-p+2}^2,$$

$$B = a_1 \sigma_{n+1}^2 = a_1(\alpha + a_1 X_n^2 + \ldots + a_p X_{n-p+1}^2),$$

$$\gamma(s,x) = \int_0^x t^{s-1}\mathrm{e}^{-t}\mathrm{d}t, \quad s > 0.$$

Solving $F_{f_h^2}(x|\{X_n, \dots, X_1\}) = \frac{1}{2}$, we obtain $\widehat{X_{n+2}^2} \approx \frac{\pi}{8}AB^2 + \pi^4 B^4$ for the L_1 optimal predictor.

The general case $h \geq 1$ has been recently worked out by Abadir et al. (2018). However, the analytical method crucially depends on the assumption for the error distribution. A more robust way to approximate the conditional distribution of $g(X_{n+h})$ can be derived using bootstrapping[2].

Before going into bootstrapping, note that a simple Monte Carlo simulation can re-produce the analytical calculations in a straightforward manner. For example, if the ARCH/GARCH errors are assumed independent with a particular distribution function F_ϵ, we can easily generate pseudo-replicates of X_{n+h} by using Equation (7), and simulating many sets of $\{\epsilon_{n+1}, \dots, \epsilon_{n+h}\}$ where each ϵ_t is drawn *i.i.d* from F_ϵ.

If F_ϵ has a known structural form with some unknown parameters, e.g., Student-t with unknown degrees of freedom, one can use a data-based estimate of the unknown parameter in order to estimate F_ϵ and then proceed with the simulation; this is then equivalent to a parametric bootstrap procedure. In the more realistic case where F_ϵ is treated as unknown, it can be estimated by the empirical distribution of the ARCH/GARCH residuals and then used in the simulation; this is equivalent to the standard (nonparametric) bootstrap. Because of the formal analogy of NoVaS to ARCH models, a similar bootstrap method works for NoVaS as well; this is the so-called model-free bootstrap. Detailed Algorithms 1–5 are given in the following two subsections.

2.4. Bootstrap Algorithms for ARCH/GARCH Point Prediction

Assume an ARCH/GARCH model with all parameters known and errors $\{\epsilon_t\} \sim$ *i.i.d.* with known distribution F_ϵ. Under the independence of $\{\epsilon_t\}$ for all $t \geq 1$, we can generate many $\epsilon^*_{n+1}, \dots, \epsilon^*_{n+h} \sim$ *i.i.d.* from F_ϵ by Monte Carlo simulation and compute many pseudo-values of the quantity of interest $g(X^*_{n+h})$.

Algorithm 1: h-step ahead prediction with parameters known: Monte Carlo.

Step 1. Generate $\{\epsilon^*_{n+1}, \dots, \epsilon^*_{n+h}\}$ *i.i.d.* from F_ϵ, and plug them into the function (7) to obtain the pseudo-value X^*_{n+h}. Repeat the above procedure M times, and denote the M pseudo-values by $\{X^{(1)}_{n+h}, \dots, X^{(M)}_{n+h}\}$.

Step 2. Calculate the optimal predictor $\widehat{g(X_{n+h})}$ of $g(X_{n+h})$ by taking the sample median (under L_1 risk) or sample mean (under L_2 risk) of the set $\{g(X^{(1)}_{n+h}), \dots, g(X^{(M)}_{n+h})\}$.

If the parameters and F_ϵ in the ARCH model are unknown, estimates must be used, in which case the Monte Carlo simulation becomes bootstrapping. Let $\hat{F}_\epsilon$ denote the estimator of F_ϵ. The two cases, parametric and nonparametric bootstrap, depend on whether the parametric form of F_ϵ is known or not. In the former case, $\hat{F}_\epsilon$ uses the parametric form of F_ϵ with parameters estimated and plugged in. In the latter, $\hat{F}_\epsilon$ is typically taken to be the empirical distribution of the ARCH residuals normalized to unit variance.

[2] The bootstrap validity is not shown in this paper, because it is beyond the scope of the paper.

Algorithm 2: h-step ahead prediction with parameters unknown: bootstrap.

Step 1. Fit the data with an ARCH(p) or GARCH(1,1) model, and obtain the estimators $\{\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_p\}$ of $\{a_0, a_1, \ldots, a_p\}$. Furthermore, record the residuals $\{\hat{\epsilon}_1, \ldots, \hat{\epsilon}_n\}$ with the distribution function $\hat{F}_\epsilon$. We will use $\hat{F}_\epsilon$ to estimate F_ϵ in the following steps.

Step 2. Perform Step 1 and Step 2 of Algorithm 1 using $\hat{F}_\epsilon$ instead F_ϵ and $\{\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_p\}$ instead of $\{a_0, a_1, \ldots, a_p\}$.

Remark 2. *The work in Bose and Mukherjee (2009) proposed a weighted linear estimator (WLE) to estimate the ARCH parameters. This method does not involve nonlinear optimization and gives a closed-form expression, so it is computationally easier to obtain the estimator compared to maximum likelihood. In our numerical work, we used the WLE to obtain the estimators $\{\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_p\}$ of $\{a_0, a_1, \ldots, a_p\}$ in Algorithm 2.*

2.5. Bootstrap Algorithms for NoVaS-Based Point Prediction

We now go back to the model-free setting of Section 2.2. In order to estimate the conditional mean or conditional median in the NoVaS setting, we should first use one of the NoVaS methods (simple vs. exponential, generalized or not, etc.) to obtain the coefficients $\alpha, a_0, a_1, \ldots, a_p$. Based on the independence of the W_t, we can use Monte Carlo and/or bootstrap to generate different W^*_{n+k} for $k = 1, \ldots, h$, and consequently approximate the distribution of $f_h(W_{n+1}, \ldots, W_{n+h})$. Denote by $\hat{F}_W$ the empirical distribution of the transformed data $W_{p+1}, \ldots, W_n$. Similar to Algorithm 2, we can use either a (truncated)[3] standard normal distribution or $\hat{F}_W$ to generate the pseudo-values W^*_{n+k}.

Algorithm 3: h-step ahead prediction for NoVaS: bootstrap.

Step 1. Use one of the NoVaS methods (simple vs. exponential, generalized or not, etc.) to obtain the transformed data $\{W_t \text{ for } t = p+1, \ldots, n\}$ and the coefficients α, p, and $a_0, a_1, \ldots, a_p$.

Step 2. Compute the analytic form of Equation (16), i.e., express X_{n+h} as a function of $\{X_1, \ldots, X_n\}$ and $\{W_{n+1}, \ldots, W_{n+h}\}$ using the values $\{a_0, a_1, \ldots, a_p\}$ obtained in Step 1.

Step 3. Generate $\{W^*_{n+1}, \ldots, W^*_{n+h}\}$ as *i.i.d.* either from a (truncated) standard normal distribution or from $\hat{F}_W$, and plug them into the function (16) to obtain the pseudo-value X^*_{n+h}. Repeat the above procedure M times and denote the M pseudo-values by $\{X^{(1)}_{n+h}, \ldots, X^{(M)}_{n+h}\}$.

Step 4. Calculate the optimal predictor $\widehat{g(X_{n+h})}$ of $g(X_{n+h})$ by taking the sample median (under L_1 risk) or sample mean (under L_2 risk) of the set $\{g(X^{(1)}_{n+h}), \ldots, g(X^{(M)}_{n+h})\}$.

3. Optimal Multi-Step-Ahead Prediction Intervals

Going beyond point prediction, it may be desirable to construct prediction intervals for $g(X_{n+h})$ with a target coverage level $(1-\beta)100\%$. One-step ahead prediction intervals have been discussed in detail in Pan and Politis (2016); see also Chapter 10 of Politis (2015). In this section, we will propose a construction of multi-step ahead prediction intervals in the given setting of financial returns data $\{X_1, \ldots, X_n\}$.

[3] From (13), it follows that:

$$\frac{1}{W_t^2} = \frac{\alpha s^2_{t-1} + a_0 X_t^2 + \sum_{i=1}^p a_i X^2_{t-i}}{X_t^2} \geq a_0$$

since all the parameters are nonnegative; thus, $|W_t| \leq 1/\sqrt{a_0}$, i.e., the range of the W_t is finite. Typically, the $\{W_t\}$ variables have a large enough range such that the boundedness is not seen as spoiling the normality from a practical perspective, but in any theoretical works and/or simulations, it is necessary to use the standard normal distribution truncated to the range $\pm 1/\sqrt{a_0}$.

As explained in Politis (2015), the bootstrap is a *sine qua non* for the construction of prediction intervals as it allows us to incorporate the variability of estimated quantities in our estimate of the conditional distribution of $g(X_{n+h})$ given $\{X_1, \ldots, X_n\}$. The variability of estimated quantities is not so important in point prediction when only the center of the conditional distribution is of interest. However, it is crucial in order to obtain an estimate of the conditional distribution that is not too narrow, yielding prediction intervals with accurate coverage.

In what follows, we give the algorithms for bootstrap prediction intervals for $g(X_{n+h})$ in the two settings: model-based for ARCH/GARCH models and model-free based on NoVaS. The algorithms follow the "forward bootstrap" paradigm introduced in Pan and Politis (2016).

Algorithm 4: Bootstrap prediction intervals for $g(X_{n+h})$ under ARCH/GARCH models.

Step 1. Fit the ARCH(p) model to the data $\{X_1, \ldots, X_n\}$, i.e., obtain the estimators $\{\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_p\}$ and the residuals $\{\widehat{\epsilon}_{p+1}, \ldots, \widehat{\epsilon}_n\}$.

Step 2. Use Algorithm 1 or Algorithm 2 to compute $\widehat{g(X_{n+h})}$, the point predictor of $g(X_{n+h})$ of choice.

Step 3. (a) Re-sample (with replacement) the residuals $\{\widehat{\epsilon}_{p+1}, \ldots, \widehat{\epsilon}_n\}$ to create the pseudo-errors $\epsilon^*_{p+1}, \cdots, \epsilon^*_n$ and $\epsilon^*_{n+1}, \cdots, \epsilon^*_{n+h}$.
(b) Let $(X^*_1, \ldots, X^*_p)' = (X_{1+I}, \cdots, X_{p+I})'$ where I is generated as a discrete random variable uniform on the values $0, 1, \ldots, n-p$. Now, use the fitted ARCH model of Step 1 to generate bootstrap pseudo-data X^*_t for $t = p+1, \ldots, n$ in a recursive manner.
(c) Based on the bootstrap data $X^*_1, \ldots, X^*_n$, re-estimate the parameters obtaining $\{\hat{a}^*_0, \hat{a}^*_1, \ldots, \hat{a}^*_p\}$.
(d) Re-define the last p values of the bootstrap data to match the original, i.e., re-define $X^*_t = X_t$ for $t = n-p+1, \ldots, n$; this is the "forward bootstrap"[4] construction.
(e) Use the fitted ARCH model of Step 1, the bootstrap data $X^*_1, \ldots, X^*_n$, and the pseudo-errors $\epsilon^*_{n+1}, \cdots, \epsilon^*_{n+h}$ to generate recursively the future bootstrap data X^*_t for $t = n+1, \ldots, n+h$.
(f). Based on the bootstrap data $X^*_{n-p+1}, \ldots, X^*_n$ and the re-estimated parameters $\{\hat{a}^*_0, \hat{a}^*_1, \ldots, \hat{a}^*_p\}$, use Algorithm 1 or 2 (according to which one was used in Step 2 to calculate the bootstrap predictor denoted by $\widehat{g(X^*_{n+h})}$.
(g) Calculate the bootstrap root[5]: $g(X^*_{n+h}) - \widehat{g(X^*_{n+h})}$.

Step 4. Repeat Step 3 above B times; the B bootstrap root replicates are collected in an empirical distribution whose α-quantile is denoted $q(\alpha)$. The $(1-\beta)100\%$ equal-tailed prediction interval for $g(X_{n+h})$ is then given by $[\widehat{g(X_{n+h})} + q(\beta/2),\ \widehat{g(X_{n+h})} + q(1-\beta/2)]$.

---

4 Most of the bootstrap methods will not have their last p values X^*_t where $t = n, \ldots, n-p+1$, exactly equal to the original values as needed for the prediction process. Herein lies the problem, since the behavior of the predictors for future values needs to be captured conditionally on the original values. In this forward bootstrap step, we redefine the last p values of X^*_t and make them equal to the original ones; see Pan and Politis (2014).

5 Since we do not have much information of the distribution of $g(X_{n+h}) - \widehat{g(X_{n+h})}$, we could use the distribution of $g(X^*_{n+h}) - \widehat{g(X^*_{n+h})}$ to approximate it. Therefore, we can employ the quantile of the approximated distribution to calculate the prediction intervals in the following steps. See more theoretical inference in Politis (2015).

Algorithm 5: Model-free (MF) bootstrap prediction intervals for $g(X_{n+h})$ based on NoVaS.

Step 1. Use one of the NoVaS algorithms (simple vs. exponential, generalized or not, etc.) to obtain the transformed data $\{W_t \text{ for } t = p+1, \ldots, n\}$ that are assumed to be approximately *i.i.d.* Let p, α, and a_i denote the fitted NoVaS parameters.

Step 2. Use Algorithm 3 to calculate $g(\widehat{X_{n+h}})$, the chosen point predictor of $g(X_{n+h})$.

Step 3. (a) Re-sample randomly (with replacement)[6] the transformed variables $\{W_t$ for $t = p+1, \ldots, n\}$ to create the pseudo-data $W^*_{p+1}, \cdots, W^*_{n-1}, W^*_n$ and $W^*_{n+1}, \cdots, W^*_{n+h}$.
(b) Let $(X^*_1, \ldots, X^*_p)' = (X_{1+I}, \cdots, X_{p+I})'$ where I is generated as a discrete random variable uniform on the values $0, 1, \ldots, n-p$. Generate the bootstrap pseudo-data X^*_t for $t = p+1, \ldots, n$ using:

$$X^*_t = \frac{W^*_t}{\sqrt{1 - a_0 W^{*2}_t}} \sqrt{\alpha s^{*2}_{t-1} + \sum_{i=1}^{p} a_i X^{*2}_{t-i}} \;\text{ for }\; t = p+1, \ldots, n \tag{20}$$

where $s^{*2}_{t-1} = (t-1)^{-1} \sum_{k=1}^{t-1} X^{*2}_k$.
(c) Based on the bootstrap data $X^*_1, \ldots, X^*_n$, re-estimate the NoVaS parameters, obtaining p^*, α^*, and a^*_i; for simplicity, we can keep the same value for p, i.e., let p^* equal p.
(d) Re-define the last p values of the bootstrap data to match the original, i.e., re-define $X^*_t = X_t$ for $t = n-p+1, \ldots, n$; this is the "forward bootstrap" construction.
(e) Calculate the bootstrap future value X^*_{n+h} by iteration as:

$$X^*_{n+1} = \frac{W^*_{n+1}}{\sqrt{1 - a_0 W^{*2}_{n+1}}} \sqrt{\alpha s^2_n + \sum_{i=1}^{p} a_i X^2_{n-i+1}}$$

where $s^2_n = n^{-1} \sum_{i=1}^{n} X^2_i$.
If $h < p$, for $j = 2, \ldots, h$:

$$X^*_{n+j} = \frac{W^*_{n+j}}{\sqrt{1 - a_0 W^{*2}_{n+j}}} \sqrt{\alpha s^2_{n+1-j} + \sum_{k=1}^{j-1} a_k X^{*2}_{n-k+j} + \sum_{i=j}^{p} a_i X^2_{n-i+j}} \tag{21}$$

If $h \geq p$, for $j = 2, \ldots, h$:

$$X^*_{n+j} = \frac{W^*_{n+j}}{\sqrt{1 - a_0 W^{*2}_{n+j}}} \sqrt{\alpha s^2_{n+1-j} + \sum_{i=1}^{p} a_i X^{*2}_{n-i+j}} \tag{22}$$

where $s^2_{n+1-j} = (n+j-1)^{-1}(\sum_{i=1}^{n} X^2_i + \sum_{k=1}^{j-1} X^{*2}_{n+k})$.
(f) Based on the bootstrap data $X^*_{n-p+1}, \ldots, X^*_n$ and the parameters p^*, α^*, $a^*_0, a^*_1, \ldots, a^*_p$, use Algorithm 3 to calculate the bootstrap predictor $g(\widehat{X^*_{n+h}})$.
(g) Calculate the bootstrap root: $g(X^*_{n+h}) - g(\widehat{X^*_{n+h}})$.

Step 4. Repeat Step 3 above B times; the B bootstrap root replicates are collected in an empirical distribution whose α-quantile is denoted $q(\alpha)$. The $(1-\beta)100\%$ equal-tailed prediction interval for $g(X_{n+h})$ is given by $[g(\widehat{X_{n+h}}) + q(\beta/2),\ g(\widehat{X_{n+h}}) + q(1-\beta/2)]$.

---

[6] It is also possible to use other bootstrap methods, for example wild bootstrap or block bootstrap, to create pseudo-data. It depends on the information you have for W_t.

4. Simulations and Finite Sample Performance

In this section, we conduct simulations to examine the finite sample performance of our algorithms.

4.1. Settings

In the simulation, 200 datasets $X_n = (X_1, \dots, X_n)'$, each of size $n = 100$, are generated separately by the following seven different GARCH(1,1) models.

Model 1. Standard GARCH with Gaussian errors and finite fourth moment:
$X_t = \sigma_t \epsilon_t,\ \sigma_t^2 = 0.00001 + 0.73\sigma_{t-1}^2 + 0.10X_{t-1}^2,\ \{\epsilon_t\} \sim i.i.d.\ N(0,1)$.
Model 2. Standard GARCH with Gaussian errors and infinite fourth moment:
$X_t = \sigma_t \epsilon_t,\ \sigma_t^2 = 0.00001 + 0.8895\sigma_{t-1}^2 + 0.10X_{t-1}^2,\ \{\epsilon_t\} \sim i.i.d.\ N(0,1)$.
Model 3. Standard GARCH with Student-*t* errors:
$X_t = \sigma_t \epsilon_t,\ \sigma_t^2 = 0.00001 + 0.73\sigma_{t-1}^2 + 0.10X_{t-1}^2,\ \{\epsilon_t\} \sim i.i.d.\ t$ distributed with five degrees of freedom.
Model 4. GARCH with time-varying parameters (TV-GARCH):
The value of β decreases as a linear function of t, starting at $\beta_1 = 0.10$ for $t = 1$, and ending at $\beta = 0.05$ for $t = n$. At the same time, the value of α increases as a linear function of t, starting at $\alpha = 0.73$ for $t = 1$, and ending at $\alpha = 0.93$ for $t = n$. $\omega = 0.00001$ and $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$.
Model 5. Two-state Markov Switching GARCH(1,1) (MS-GARCH):

$$X_t = \sigma_t \epsilon_t, \sigma_t^2 = \sum_{s=1}^{2} 1\{P(S_t = s)\}[\omega_s + \alpha_t \sigma_{t-1}^2 + \beta_s X_{t-1}^2]$$

In the first regime, we set $\alpha_1 = 0.9$, $\beta_1 = 0.07, \omega_1 = 2.4e-5$. In the second regime, we set $\alpha_2 = 0.7$, $\beta_2 = 0.22, \omega_2 = 1.2e-4$. The transition probabilities for the first regime are $p_{11} = 0.9$ and $p_{12} = 0.1$, while for the second regime, we use $p_{21} = 0.3$ and $p_{22} = 0.7$. $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$.
Model 6. Smooth transition GARCH (ST-GARCH):

$$X_t = [a - b(t/T)]\sigma_t \epsilon_t,\ \sigma_t^2 = \omega + \alpha\sigma_{t-1}^2 + \beta X_{t-1}^2$$

where $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$. $\omega = 1.2e-5, \alpha = 0.9, \beta = 0.07, a = \alpha + \beta = 0.97$, and $b = \beta/\alpha \approx 0.078$.
Model 7. Stochastic volatility model (SV-GARCH):

$$X_t | h_t \sim N(0, \exp(h_t)),$$

$$h_t | h_{t-1} \sim N(\mu + \phi(h_{t-1} - \mu), \eta^2),\ h_0 \sim N(\mu, \eta^2/(1-\phi^2)),$$

where $\mu = -10$, $\phi = 0.95$, $\eta = 0.2$.

We performed up to five-step-ahead point predictions and interval predictions for each dataset. We used $M = 5000$ simulations to compute the point predictions. For the bootstrap prediction intervals, we used B = 300 replications; we would have liked to use a higher number of bootstrap replications, but it was practically infeasible having 200 simulated datasets in each of which the point predictors were computed in a computer-intensive way. However, a practitioner having a single dataset at hand could (and should) use a higher B, e.g., $B = 1000$ or more.

Five models or transformations were used to fit the data in both point predictions and interval predictions as follows: fitting a GARCH(1,1) model, simple-NoVaS, exponential NoVaS (Exp-NoVaS), generalized simple NoVaS (GS-NoVaS), and generalized exponential NoVaS (GE-NoVaS).

In point predictions, the mean absolute deviations (MAD) and mean squared errors (MSE) for five-step-ahead point predictions in both the L_1 and L_2 sense (the absolute value or the square of the prediction error at the updated time point averaged over the 200 replications) were recorded. Furthermore, the bootstrap prediction interval (L_i, U_i) with a nominal coverage 95% was constructed for the future values X_{n+h} with $h = 1, \dots, 5$.

The corresponding empirical average coverage level (CVR) and the average length (LEN) of the constructed intervals and the standard error (St.err) associated with each length of the constructed intervals are calculated as:

$$CVR = \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{[L_i,U_i]} X_{(n+h,i)}$$

$$LEN = \frac{1}{N}\sum_{i=1}^{N} LEN_i \text{ and } St.err = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(LEN_i - LEN)^2}$$

where $LEN_i = U_i - L_i$.

4.2. Results and Discussions

The simulation results for point predictions are shown in Tables 1–28. The following conclusions can be obtained from the results:

- When comparing the MADs between the L_1 and L_2 predictions by fitting the same models, we can find that the MADs of L_1 predictions were always smaller than those of L_2 predictions. Furthermore, we can find that the MSEs of L_1 predictions were always bigger than those of L_2 predictions, when comparing the MSEs between the L_1 and L_2 predictions with the same model settings. This was expected since in the L_1 sense, we tried to minimize the mean absolute deviations, while for L_2, the loss function to be minimized was the mean squared error.
- Furthermore, for each model's fitting results, there were no obvious error accumulation problems in the multi-step-ahead prediction for both L_1 and L_2 measures.
- NoVaS methods consistently performed better than GARCH(1,1) for all data generating processes. When the prediction step h was higher, the difference of their respective performances became smaller.

Table 1. MADs of L_1 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.8895, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$. Exp, exponential; GS, generalized simple; GE, generalized exponential.

Prediction Step	1	2	3	4	5
Fitting a GARCH	8.23×10^{-5}	7.35×10^{-5}	6.93×10^{-5}	8.86×10^{-5}	1.49×10^{-4}
Simple-NoVaS	6.99×10^{-5}	7.88×10^{-5}	8.29×10^{-5}	1.06×10^{-4}	1.63×10^{-4}
Exp-NoVaS	7.21×10^{-5}	8.28×10^{-5}	8.72×10^{-5}	1.14×10^{-4}	1.70×10^{-4}
GS-NoVaS	6.30×10^{-5}	7.31×10^{-5}	8.07×10^{-5}	8.71×10^{-5}	9.94×10^{-5}
GE-NoVaS	7.02×10^{-5}	8.44×10^{-5}	8.76×10^{-5}	1.16×10^{-4}	1.71×10^{-4}

Table 2. MADs of L_2 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.8895, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$.

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.48×10^{-4}	2.07×10^{-4}	2.67×10^{-4}	3.20×10^{-4}	4.60×10^{-4}
Simple-NoVaS	5.22×10^{-5}	6.44×10^{-5}	6.61×10^{-5}	8.82×10^{-5}	1.49×10^{-4}
Exp-NoVaS	5.13×10^{-5}	6.41×10^{-5}	6.58×10^{-5}	8.84×10^{-5}	1.50×10^{-4}
GS-NoVaS	4.71×10^{-5}	5.97×10^{-5}	6.06×10^{-5}	8.48×10^{-5}	1.46×10^{-4}
GE-NoVaS	4.84×10^{-5}	6.26×10^{-5}	6.38×10^{-5}	8.73×10^{-5}	1.48×10^{-4}

Table 3. MSEs of L_1 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.8895, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$.

Prediction Step	**1**	**2**	**3**	**4**	**5**
Fitting a GARCH	2.48×10^{-8}	4.42×10^{-8}	4.36×10^{-8}	1.94×10^{-7}	1.12×10^{-6}
Simple-NoVaS	1.23×10^{-8}	4.24×10^{-8}	4.17×10^{-8}	1.94×10^{-7}	1.12×10^{-6}
Exp-NoVaS	1.18×10^{-8}	4.21×10^{-8}	4.14×10^{-8}	1.94×10^{-7}	1.12×10^{-6}
GS-NoVaS	1.08×10^{-8}	4.09×10^{-8}	3.94×10^{-8}	1.92×10^{-7}	1.13×10^{-6}
GE-NoVaS	1.10×10^{-8}	4.14×10^{-8}	4.06×10^{-8}	1.92×10^{-7}	1.12×10^{-6}

Table 4. MSEs of L_2 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.8895, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$.

Prediction Step	**1**	**2**	**3**	**4**	**5**
Fitting a GARCH	1.03×10^{-7}	2.11×10^{-7}	3.92×10^{-7}	6.26×10^{-7}	1.86×10^{-6}
Simple-NoVaS	1.36×10^{-8}	4.15×10^{-8}	4.10×10^{-8}	1.93×10^{-7}	1.11×10^{-6}
Exp-NoVaS	1.31×10^{-8}	4.13×10^{-8}	4.07×10^{-8}	1.93×10^{-7}	1.12×10^{-6}
GS-NoVaS	1.03×10^{-8}	3.88×10^{-8}	3.73×10^{-8}	1.91×10^{-7}	1.13×10^{-6}
GE-NoVaS	1.13×10^{-8}	3.98×10^{-8}	3.93×10^{-8}	1.88×10^{-7}	1.11×10^{-6}

Table 5. MADs of L_1 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$.

Prediction Step	**1**	**2**	**3**	**4**	**5**
Fitting a GARCH	4.84×10^{-5}	4.69×10^{-5}	5.18×10^{-5}	5.45×10^{-5}	5.99×10^{-5}
Simple-NoVaS	4.94×10^{-5}	4.73×10^{-5}	5.26×10^{-5}	5.59×10^{-5}	6.00×10^{-5}
Exp-NoVaS	4.87×10^{-5}	4.69×10^{-5}	5.24×10^{-5}	5.55×10^{-5}	6.00×10^{-5}
GS-NoVaS	4.83×10^{-5}	4.67×10^{-5}	5.17×10^{-5}	5.44×10^{-5}	5.94×10^{-5}
GE-NoVaS	4.84×10^{-5}	4.69×10^{-5}	5.25×10^{-5}	5.43×10^{-5}	6.00×10^{-5}

Table 6. MADs of L_2 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$.

Prediction Step	**1**	**2**	**3**	**4**	**5**
Fitting a GARCH	5.68×10^{-5}	5.45×10^{-5}	5.54×10^{-5}	6.25×10^{-5}	0.26×10^{-5}
Simple-NoVaS	6.20×10^{-5}	6.05×10^{-5}	6.06×10^{-5}	6.96×10^{-5}	6.67×10^{-5}
Exp-NoVaS	6.20×10^{-5}	6.09×10^{-5}	6.15×10^{-5}	7.05×10^{-5}	6.87×10^{-5}
GS-NoVaS	5.98×10^{-5}	5.81×10^{-5}	5.84×10^{-5}	6.55×10^{-5}	6.37×10^{-5}
GE-NoVaS	5.50×10^{-5}	5.30×10^{-5}	5.75×10^{-5}	6.22×10^{-5}	6.23×10^{-5}

Table 7. MSEs of L_1 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$.

Prediction step	1	2	3	4	5
Fitting a GARCH	7.94×10^{-9}	8.48×10^{-9}	7.80×10^{-9}	8.67×10^{-9}	1.03×10^{-8}
Simple-NoVaS	8.00×10^{-9}	8.35×10^{-9}	7.85×10^{-9}	8.88×10^{-9}	1.03×10^{-8}
Exp-NoVaS	7.95×10^{-9}	8.33×10^{-9}	7.83×10^{-9}	8.84×10^{-9}	1.03×10^{-8}
GS-NoVaS	7.78×10^{-9}	8.35×10^{-9}	7.63×10^{-9}	8.59×10^{-9}	1.02×10^{-8}
GE-NoVaS	8.02×10^{-9}	8.64×10^{-9}	7.99×10^{-9}	8.92×10^{-9}	1.02×10^{-8}

Table 8. MSEs of L_2 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.\ N(0,1)$.

Prediction Step	**1**	**2**	**3**	**4**	**5**
Fitting a GARCH	7.28×10^{-9}	7.47×10^{-9}	6.50×10^{-9}	7.64×10^{-9}	8.65×10^{-9}
Simple-NoVaS	7.79×10^{-9}	7.50×10^{-9}	6.84×10^{-9}	8.67×10^{-9}	8.85×10^{-9}
Exp-NoVaS	7.48×10^{-9}	7.34×10^{-9}	6.64×10^{-9}	8.39×10^{-9}	8.70×10^{-9}
GS-NoVaS	7.12×10^{-9}	7.40×10^{-9}	6.41×10^{-9}	7.90×10^{-9}	8.43×10^{-9}
GE-NoVaS	6.99×10^{-9}	7.44×10^{-9}	6.43×10^{-9}	7.68×10^{-9}	8.53×10^{-9}

Table 9. MADs of L_1 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.$ t distributed with degrees of freedom of 5.

Prediction Step	**1**	**2**	**3**	**4**	**5**
Fitting a GARCH	1.46×10^{-4}	1.26×10^{-4}	1.29×10^{-4}	1.84×10^{-4}	1.77×10^{-4}
Simple-NoVaS	1.44×10^{-4}	1.24×10^{-4}	1.27×10^{-4}	1.82×10^{-4}	1.76×10^{-4}
Exp-NoVaS	1.43×10^{-4}	1.22×10^{-4}	1.26×10^{-4}	1.80×10^{-4}	1.75×10^{-4}
GS-NoVaS	1.43×10^{-4}	1.23×10^{-4}	1.27×10^{-4}	1.80×10^{-4}	1.76×10^{-4}
GE-NoVaS	1.43×10^{-4}	1.23×10^{-4}	1.29×10^{-4}	1.81×10^{-4}	1.77×10^{-4}

Table 10. MADs of L_2 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.$ t distributed with degrees of freedom of 5.

Prediction Step	**1**	**2**	**3**	**4**	**5**
Fitting a GARCH	1.72×10^{-4}	1.53×10^{-4}	1.63×10^{-4}	2.17×10^{-4}	2.16×10^{-4}
Simple-NoVaS	1.58×10^{-4}	1.35×10^{-4}	1.44×10^{-4}	1.97×10^{-4}	1.92×10^{-4}
Exp-NoVaS	1.57×10^{-4}	1.35×10^{-4}	1.46×10^{-4}	1.98×10^{-4}	1.93×10^{-4}
GS-NoVaS	1.56×10^{-4}	1.35×10^{-4}	1.45×10^{-4}	1.99×10^{-4}	1.92×10^{-4}
GE-NoVaS	1.59×10^{-4}	1.34×10^{-4}	1.48×10^{-4}	2.01×10^{-4}	1.94×10^{-4}

Table 11. MSEs of L_1 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.$ t distributed with degrees of freedom of 5.

Prediction Step	**1**	**2**	**3**	**4**	**5**
Fitting a GARCH	1.06×10^{-7}	7.11×10^{-8}	8.97×10^{-8}	2.95×10^{-7}	3.81×10^{-7}
Simple-NoVaS	1.03×10^{-7}	6.86×10^{-8}	9.00×10^{-8}	2.93×10^{-7}	3.83×10^{-7}
Exp-NoVaS	1.03×10^{-7}	6.83×10^{-8}	9.00×10^{-8}	2.93×10^{-7}	3.82×10^{-7}
GS-NoVaS	1.03×10^{-7}	6.90×10^{-8}	9.05×10^{-8}	2.94×10^{-7}	3.84×10^{-7}
GE-NoVaS	1.05×10^{-7}	6.88×10^{-8}	9.23×10^{-8}	2.96×10^{-7}	3.85×10^{-7}

Table 12. MSEs of L_2 predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\{\epsilon_t\} \sim i.i.d.$ t distributed with degrees of freedom of 5.

Prediction Step	**1**	**2**	**3**	**4**	**5**
Fitting a GARCH	1.05×10^{-7}	6.94×10^{-8}	8.91×10^{-8}	2.84×10^{-7}	3.79×10^{-7}
Simple-NoVaS	9.24×10^{-8}	6.03×10^{-8}	8.32×10^{-8}	2.77×10^{-7}	3.66×10^{-7}
Exp-NoVaS	9.16×10^{-8}	5.96×10^{-8}	8.28×10^{-8}	2.76×10^{-7}	3.64×10^{-7}
GS-NoVaS	9.17×10^{-8}	6.07×10^{-8}	8.35×10^{-8}	2.77×10^{-7}	3.67×10^{-7}
GE-NoVaS	9.48×10^{-8}	6.12×10^{-8}	8.53×10^{-8}	2.80×10^{-7}	3.67×10^{-7}

Table 13. MADs of L_1 predictions for data generated from GARCH(1,1) with slowing-varying parameters (TV-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.94×10^{-4}	2.17×10^{-4}	2.01×10^{-4}	1.76×10^{-4}	2.09×10^{-4}
Simple-NoVaS	1.91×10^{-4}	2.12×10^{-4}	2.03×10^{-4}	1.72×10^{-4}	2.07×10^{-4}
Exp-NoVaS	1.91×10^{-4}	2.12×10^{-4}	2.02×10^{-4}	1.72×10^{-4}	2.06×10^{-4}
GS-NoVaS	1.91×10^{-4}	2.13×10^{-4}	2.02×10^{-4}	1.73×10^{-4}	2.06×10^{-4}
GE-NoVaS	1.97×10^{-4}	2.17×10^{-4}	2.02×10^{-4}	1.79×10^{-4}	2.13×10^{-4}

Table 14. MADs of L_2 predictions for data generated from GARCH(1,1) with slowing-varying parameters (TV-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.90×10^{-4}	2.13×10^{-4}	2.03×10^{-4}	1.73×10^{-4}	2.02×10^{-4}
Simple-NoVaS	1.94×10^{-4}	2.11×10^{-4}	2.07×10^{-4}	1.74×10^{-4}	2.10×10^{-4}
Exp-NoVaS	1.94×10^{-4}	2.11×10^{-4}	2.05×10^{-4}	1.74×10^{-4}	2.09×10^{-4}
GS-NoVaS	1.91×10^{-4}	2.10×10^{-4}	1.99×10^{-4}	1.71×10^{-4}	2.03×10^{-4}
GE-NoVaS	1.89×10^{-4}	2.09×10^{-4}	1.99×10^{-4}	1.72×10^{-4}	2.07×10^{-4}

Table 15. MSEs of L_1 predictions for data generated from GARCH(1,1) with slowing-varying parameters (TV-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.21×10^{-7}	1.70×10^{-7}	1.28×10^{-7}	1.06×10^{-7}	1.58×10^{-7}
Simple-NoVaS	1.15×10^{-7}	1.61×10^{-7}	1.23×10^{-7}	9.86×10^{-8}	1.53×10^{-7}
Exp-NoVaS	1.16×10^{-7}	1.61×10^{-7}	1.23×10^{-7}	9.95×10^{-8}	1.52×10^{-7}
GS-NoVaS	1.16×10^{-7}	1.62×10^{-7}	1.24×10^{-7}	9.99×10^{-8}	1.53×10^{-7}
GE-NoVaS	1.25×10^{-7}	1.70×10^{-7}	1.31×10^{-7}	1.08×10^{-7}	1.61×10^{-7}

Table 16. MSEs of L_2 predictions for data generated from GARCH(1,1) with slowing-varying parameters (TV-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.07×10^{-7}	1.55×10^{-7}	1.12×10^{-7}	9.44×10^{-8}	1.41×10^{-7}
Simple-NoVaS	9.60×10^{-8}	1.41×10^{-7}	1.06×10^{-7}	8.36×10^{-8}	1.32×10^{-7}
Exp-NoVaS	9.60×10^{-8}	1.40×10^{-7}	1.05×10^{-7}	8.39×10^{-8}	1.30×10^{-7}
GS-NoVaS	9.70×10^{-8}	1.41×10^{-7}	1.07×10^{-7}	8.51×10^{-8}	1.32×10^{-7}
GE-NoVaS	1.04×10^{-7}	1.47×10^{-7}	1.12×10^{-7}	9.02×10^{-8}	1.38×10^{-7}

Table 17. MADs of L_1 predictions for data generated from two-state Markov switching GARCH(1,1) (MS-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	6.76×10^{-4}	7.67×10^{-4}	8.44×10^{-4}	7.80×10^{-4}	7.14×10^{-4}
Simple-NoVaS	7.00×10^{-4}	7.74×10^{-4}	8.85×10^{-4}	7.90×10^{-4}	7.26×10^{-4}
Exp-NoVaS	7.02×10^{-4}	7.75×10^{-4}	8.87×10^{-4}	7.92×10^{-4}	7.26×10^{-4}
GS-NoVaS	6.97×10^{-4}	7.70×10^{-4}	8.80×10^{-4}	7.91×10^{-4}	7.23×10^{-4}
GE-NoVaS	7.06×10^{-4}	7.75×10^{-4}	8.85×10^{-4}	7.98×10^{-4}	7.27×10^{-4}

Table 18. MADs of L_2 predictions for data generated from two-state Markov switching GARCH(1,1) (MS-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	7.80×10^{-4}	9.01×10^{-4}	9.21×10^{-4}	9.17×10^{-4}	8.80×10^{-4}
Simple-NoVaS	6.77×10^{-4}	7.50×10^{-4}	8.59×10^{-4}	7.73×10^{-4}	7.10×10^{-4}
Exp-NoVaS	6.78×10^{-4}	7.50×10^{-4}	8.61×10^{-4}	7.75×10^{-4}	7.09×10^{-4}
GS-NoVaS	6.76×10^{-4}	7.50×10^{-4}	8.59×10^{-4}	7.76×10^{-4}	7.08×10^{-4}
GE-NoVaS	6.81×10^{-4}	7.50×10^{-4}	8.62×10^{-4}	7.79×10^{-4}	7.09×10^{-4}

Table 19. MSEs of L_1 predictions for data generated from two-state Markov switching GARCH(1,1) (MS-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.27×10^{-6}	2.30×10^{-6}	2.35×10^{-6}	2.78×10^{-6}	1.49×10^{-6}
Simple-NoVaS	1.45×10^{-6}	2.46×10^{-6}	2.73×10^{-6}	3.09×10^{-6}	1.76×10^{-6}
Exp-NoVaS	1.45×10^{-6}	2.46×10^{-6}	2.73×10^{-6}	3.10×10^{-6}	1.76×10^{-6}
GS-NoVaS	1.43×10^{-6}	2.45×10^{-6}	2.71×10^{-6}	3.10×10^{-6}	1.74×10^{-6}
GE-NoVaS	1.46×10^{-6}	2.47×10^{-6}	2.74×10^{-6}	3.12×10^{-6}	1.76×10^{-6}

Table 20. MSEs of L_2 predictions for data generated from two-state Markov switching GARCH(1,1) (MS-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.29×10^{-6}	2.28×10^{-6}	2.05×10^{-6}	2.55×10^{-6}	1.51×10^{-6}
Simple-NoVaS	1.34×10^{-6}	2.33×10^{-6}	2.58×10^{-6}	2.97×10^{-6}	1.64×10^{-6}
Exp-NoVaS	1.34×10^{-6}	2.33×10^{-6}	2.59×10^{-6}	2.98×10^{-6}	1.64×10^{-6}
GS-NoVaS	1.33×10^{-6}	2.32×10^{-6}	2.57×10^{-6}	2.97×10^{-6}	1.62×10^{-6}
GE-NoVaS	1.35×10^{-6}	2.33×10^{-6}	2.59×10^{-6}	2.99×10^{-6}	1.64×10^{-6}

Table 21. MADs of L_1 predictions for data generated from smooth transition GARCH(1,1) (ST-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.83×10^{-4}	1.78×10^{-4}	2.01×10^{-4}	2.02×10^{-4}	2.22×10^{-4}
Simple-NoVaS	1.83×10^{-4}	1.79×10^{-4}	2.03×10^{-4}	2.02×10^{-4}	2.24×10^{-4}
Exp-NoVaS	1.82×10^{-4}	1.78×10^{-4}	2.03×10^{-4}	2.02×10^{-4}	2.24×10^{-4}
GS-NoVaS	1.81×10^{-4}	1.78×10^{-4}	2.02×10^{-4}	1.98×10^{-4}	2.21×10^{-4}
GE-NoVaS	1.82×10^{-4}	1.79×10^{-4}	2.05×10^{-4}	1.99×10^{-4}	2.24×10^{-4}

Table 22. MADs of L_2 predictions for data generated from smooth transition GARCH(1,1) (ST-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	2.23×10^{-4}	2.14×10^{-4}	2.18×10^{-4}	2.39×10^{-4}	2.42×10^{-4}
Simple-NoVaS	1.91×10^{-4}	1.88×10^{-4}	2.02×10^{-4}	2.09×10^{-4}	2.19×10^{-4}
Exp-NoVaS	1.89×10^{-4}	1.86×10^{-4}	2.03×10^{-4}	2.09×10^{-4}	2.19×10^{-4}
GS-NoVaS	1.90×10^{-4}	1.89×10^{-4}	2.03×10^{-4}	2.09×10^{-4}	2.17×10^{-4}
GE-NoVaS	1.86×10^{-4}	1.83×10^{-4}	2.06×10^{-4}	2.03×10^{-4}	2.20×10^{-4}

Table 23. MSEs of L_1 predictions for data generated from smooth transition GARCH(1,1) (ST-GARCH)

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.12×10^{-7}	1.23×10^{-7}	1.14×10^{-7}	1.13×10^{-7}	1.43×10^{-7}
Simple-NoVaS	1.16×10^{-7}	1.24×10^{-7}	1.19×10^{-7}	1.21×10^{-7}	1.48×10^{-7}
Exp-NoVaS	1.16×10^{-7}	1.24×10^{-7}	1.20×10^{-7}	1.21×10^{-7}	1.49×10^{-7}
GS-NoVaS	1.12×10^{-7}	1.22×10^{-7}	1.16×10^{-7}	1.16×10^{-7}	1.45×10^{-7}
GE-NoVaS	1.18×10^{-7}	1.27×10^{-7}	1.22×10^{-7}	1.21×10^{-7}	1.50×10^{-7}

Table 24. MSEs of L_2 predictions for data generated from smooth transition GARCH(1,1) (ST-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.05×10^{-7}	1.13×10^{-7}	9.61×10^{-8}	1.04×10^{-7}	1.25×10^{-7}
Simple-NoVaS	1.01×10^{-7}	1.09×10^{-7}	9.96×10^{-8}	1.04×10^{-7}	1.23×10^{-7}
Exp-NoVaS	9.93×10^{-8}	1.08×10^{-7}	1.00×10^{-7}	1.03×10^{-7}	1.23×10^{-7}
GS-NoVaS	9.84×10^{-8}	1.08×10^{-7}	9.85×10^{-8}	1.00×10^{-7}	1.21×10^{-7}
GE-NoVaS	1.00×10^{-7}	1.11×10^{-7}	1.02×10^{-7}	1.02×10^{-7}	1.24×10^{-7}

Table 25. MADs of L_1 predictions for data generated from the stochastic volatility model (SV-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	5.35×10^{-5}	5.77×10^{-5}	4.72×10^{-5}	4.28×10^{-5}	3.86×10^{-5}
Simple-NoVaS	5.34×10^{-5}	5.69×10^{-5}	4.79×10^{-5}	4.45×10^{-5}	3.78×10^{-5}
Exp-NoVaS	5.35×10^{-5}	5.66×10^{-5}	4.74×10^{-5}	4.35×10^{-5}	3.75×10^{-5}
GS-NoVaS	5.23×10^{-5}	5.68×10^{-5}	4.74×10^{-5}	4.23×10^{-5}	3.79×10^{-5}
GE-NoVaS	5.23×10^{-5}	5.71×10^{-5}	4.79×10^{-5}	4.31×10^{-5}	3.85×10^{-5}

Table 26. MADs of L_2 predictions for data generated from the stochastic volatility model (SV-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	5.77×10^{-5}	6.23×10^{-5}	5.56×10^{-5}	5.29×10^{-5}	5.12×10^{-5}
Simple-NoVaS	6.02×10^{-5}	6.17×10^{-5}	5.85×10^{-5}	5.99×10^{-5}	5.41×10^{-5}
Exp-NoVaS	6.10×10^{-5}	6.19×10^{-5}	5.99×10^{-5}	6.04×10^{-5}	5.68×10^{-5}
GS-NoVaS	5.74×10^{-5}	6.29×10^{-5}	6.10×10^{-5}	5.85×10^{-5}	5.69×10^{-5}
GE-NoVaS	5.66×10^{-5}	5.96×10^{-5}	5.45×10^{-5}	5.28×10^{-5}	4.77×10^{-5}

Table 27. MSEs of L_1 predictions for data generated from the stochastic volatility model (SV-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	1.04×10^{-8}	1.69×10^{-8}	1.08×10^{-8}	7.03×10^{-9}	5.06×10^{-9}
Simple-NoVaS	1.02×10^{-8}	1.65×10^{-8}	1.05×10^{-8}	7.06×10^{-9}	4.82×10^{-9}
Exp-NoVaS	1.04×10^{-8}	1.65×10^{-8}	1.06×10^{-8}	6.99×10^{-9}	4.78×10^{-9}
GS-NoVaS	1.00×10^{-8}	1.66×10^{-8}	1.04×10^{-8}	6.92×10^{-9}	4.80×10^{-9}
GE-NoVaS	1.03×10^{-8}	1.70×10^{-8}	1.08×10^{-8}	6.94×10^{-9}	5.04×10^{-9}

Table 28. MSEs of L_2 predictions for data generated from the stochastic volatility model (SV-GARCH).

Prediction Step	1	2	3	4	5
Fitting a GARCH	8.99×10^{-9}	1.51×10^{-8}	1.02×10^{-8}	6.76×10^{-9}	5.29×10^{-9}
Simple-NoVaS	8.70×10^{-9}	1.42×10^{-8}	9.99×10^{-9}	7.62×10^{-9}	5.49×10^{-9}
Exp-NoVaS	8.82×10^{-9}	1.39×10^{-8}	1.00×10^{-8}	7.41×10^{-9}	5.51×10^{-9}
GS-NoVaS	8.10×10^{-9}	1.42×10^{-8}	1.02×10^{-8}	6.92×10^{-9}	5.32×10^{-9}
GE-NoVaS	8.37×10^{-9}	1.46×10^{-8}	1.00×10^{-8}	6.62×10^{-9}	5.27×10^{-9}

As regards prediction intervals, the simulation results are summarized in Tables 29–35. Generally, the conclusions were similar to those of the point predictions. Furthermore, NoVaS methods gave more accurate coverage than GARCH(1,1) in the L_1 sense prediction for all seven data generating processes. In the L_2 sense, when the data were generated from a standard GARCH(1,1) with normal errors, GARCH(1,1) gave as good coverage as NoVaS. When we used other models to generate data, for example, GARCH(1,1) with t distributed errors or MS-GARCH(1,1) or ST-GARCH(1,1) etc., GARCH(1,1) performed very poorly, while NoVaS methods were still performing well.

Table 29. Interval predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.8895, \theta = 0.10$, and $\epsilon \sim i.i.d\ N(0,1)$. CVR, average coverage level; LEN, average length.

L2				**L1**			
GARCH(1,1)				GARCH(1,1)			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.714	1.28×10^{-2}	1.77×10^{-2}	1	0.744	1.01×10^{-2}	1.55×10^{-2}
2	0.746	1.28×10^{-2}	1.79×10^{-2}	2	0.734	1.13×10^{-2}	1.83×10^{-2}
3	0.746	1.37×10^{-2}	1.87×10^{-2}	3	0.768	1.19×10^{-2}	1.95×10^{-2}
4	0.766	1.18×10^{-2}	1.59×10^{-2}	4	0.734	1.20×10^{-2}	1.99×10^{-2}
5	0.786	1.28×10^{-2}	1.75×10^{-2}	5	0.744	1.28×10^{-2}	2.04×10^{-2}
EXP-NoVaS				EXP-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.958	2.00×10^{-2}	1.35×10^{-2}	1	0.936	1.99×10^{-2}	1.39×10^{-2}
2	0.952	2.07×10^{-2}	1.44×10^{-2}	2	0.936	2.19×10^{-2}	1.52×10^{-2}
3	0.952	1.98×10^{-2}	1.27×10^{-2}	3	0.944	1.94×10^{-2}	1.27×10^{-2}
4	0.946	2.17×10^{-2}	1.43×10^{-2}	4	0.944	2.08×10^{-2}	1.28×10^{-2}
5	0.950	2.19×10^{-2}	1.32×10^{-2}	5	0.936	2.19×10^{-2}	1.51×10^{-2}
Simple-NoVaS				Simple-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.946	1.98×10^{-2}	1.35×10^{-2}	1	0.934	1.84×10^{-2}	1.32×10^{-2}
2	0.942	1.78×10^{-2}	1.41×10^{-2}	2	0.946	1.91×10^{-2}	1.49×10^{-2}
3	0.946	1.92×10^{-2}	1.35×10^{-2}	3	0.936	2.04×10^{-2}	1.47×10^{-2}
4	0.946	2.07×10^{-2}	1.34×10^{-2}	4	0.964	1.93×10^{-2}	1.40×10^{-2}
5	0.956	2.21×10^{-2}	1.43×10^{-2}	5	0.954	2.07×10^{-2}	1.45×10^{-2}
GS-NoVaS				GS-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.948	2.05×10^{-2}	1.39×10^{-2}	1	0.95	1.61×10^{-2}	1.21×10^{-2}
2	0.942	1.78×10^{-2}	1.41×10^{-2}	2	0.94	2.07×10^{-2}	1.47×10^{-2}
3	0.952	1.93×10^{-2}	1.43×10^{-2}	3	0.936	1.62×10^{-2}	1.25×10^{-2}
4	0.948	2.14×10^{-2}	1.30×10^{-2}	4	0.936	1.70×10^{-2}	1.42×10^{-2}
5	0.954	2.26×10^{-2}	1.40×10^{-2}	5	0.94	1.78×10^{-2}	1.32×10^{-2}
GE-NoVaS				GE-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.948	1.73×10^{-2}	1.14×10^{-2}	1	0.958	1.98×10^{-2}	1.13×10^{-2}
2	0.948	1.66×10^{-2}	1.05×10^{-2}	2	0.942	1.68×10^{-2}	1.16×10^{-2}
3	0.952	1.98×10^{-2}	1.27×10^{-2}	3	0.944	1.94×10^{-2}	1.27×10^{-2}
4	0.946	2.17×10^{-2}	1.43×10^{-2}	4	0.944	2.08×10^{-2}	1.28×10^{-2}
5	0.950	2.19×10^{-2}	1.32×10^{-2}	5	0.942	2.32×10^{-2}	1.13×10^{-2}

Table 30. Interval predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\epsilon \sim i.i.d\ N(0,1)$.

L2				**L1**			
GARCH(1,1)				GARCH(1,1)			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.948	4.39×10^{-3}	2.11×10^{-3}	1	0.92	3.76×10^{-3}	3.18×10^{-3}
2	0.940	4.53×10^{-3}	2.23×10^{-3}	2	0.936	5.49×10^{-3}	5.21×10^{-3}
3	0.950	4.47×10^{-3}	2.74×10^{-3}	3	0.938	5.99×10^{-3}	5.52×10^{-3}
4	0.952	4.02×10^{-3}	1.89×10^{-3}	4	0.922	7.16×10^{-3}	6.31×10^{-3}
5	0.934	3.77×10^{-3}	2.74×10^{-3}	5	0.92	4.57×10^{-3}	4.21×10^{-3}

Table 30. *Cont.*

L2				**L1**			
EXP-NoVaS				EXP-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.954	4.46×10^{-3}	2.69×10^{-3}	1	0.954	4.74×10^{-3}	2.29×10^{-3}
2	0.972	4.40×10^{-3}	2.55×10^{-3}	2	0.95	4.62×10^{-3}	2.19×10^{-3}
3	0.938	4.17×10^{-3}	2.59×10^{-3}	3	0.95	4.64×10^{-3}	2.12×10^{-3}
4	0.958	4.62×10^{-3}	2.59×10^{-3}	4	0.948	4.58×10^{-3}	2.12×10^{-3}
5	0.950	4.58×10^{-3}	2.49×10^{-3}	5	0.942	4.45×10^{-3}	1.88×10^{-3}
Simple-NoVaS				Simple-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.960	4.54×10^{-3}	2.50×10^{-3}	1	0.946	4.28×10^{-3}	2.45×10^{-3}
2	0.958	4.27×10^{-3}	2.97×10^{-3}	2	0.95	4.26×10^{-3}	2.33×10^{-3}
3	0.968	4.63×10^{-3}	2.87×10^{-3}	3	0.952	4.21×10^{-3}	2.62×10^{-3}
4	0.960	4.73×10^{-3}	2.85×10^{-3}	4	0.954	4.25×10^{-3}	2.55×10^{-3}
5	0.948	4.15×10^{-3}	2.93×10^{-3}	5	0.948	4.19×10^{-3}	2.32×10^{-3}
GS-NoVaS				GS-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.949	4.37×10^{-3}	2.53×10^{-3}	1	0.946	4.26×10^{-3}	2.37×10^{-3}
2	0.948	4.63×10^{-2}	2.78×10^{-3}	2	0.95	4.26×10^{-3}	2.33×10^{-3}
3	0.938	4.17×10^{-3}	2.59×10^{-3}	3	0.95	4.22×10^{-3}	2.42×10^{-3}
4	0.945	3.76×10^{-3}	2.71×10^{-3}	4	0.948	4.20×10^{-3}	1.91×10^{-3}
5	0.950	4.58×10^{-3}	2.49×10^{-3}	5	0.948	4.19×10^{-3}	2.32×10^{-3}
GE-NoVaS				GE-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.946	4.92×10^{-3}	2.57×10^{-3}	1	0.96	5.37×10^{-3}	2.00×10^{-3}
2	0.946	4.68×10^{-3}	2.38×10^{-3}	2	0.948	5.13×10^{-3}	3.34×10^{-3}
3	0.958	4.39×10^{-3}	2.35×10^{-3}	3	0.952	4.03×10^{-3}	2.05×10^{-3}
4	0.954	4.30×10^{-3}	2.03×10^{-3}	4	0.95	4.80×10^{-3}	2.03×10^{-3}
5	0.948	4.15×10^{-3}	2.93×10^{-3}	5	0.944	4.42×10^{-3}	2.78×10^{-3}

Table 31. Results of interval predictions for data generated from GARCH(1,1) with $\omega = 0.00001, \alpha = 0.73, \theta = 0.10$, and $\epsilon \sim i.i.d\ t_5$.

L2				**L1**			
GARCH(1,1)				GARCH(1,1)			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.924	4.17×10^{-3}	3.18×10^{-3}	1	0.936	2.35×10^{-3}	8.60×10^{-3}
2	0.931	3.75×10^{-3}	2.57×10^{-3}	2	0.928	2.15×10^{-3}	7.50×10^{-3}
3	0.922	4.47×10^{-3}	2.24×10^{-3}	3	0.92	2.37×10^{-3}	8.52×10^{-3}
4	0.925	4.02×10^{-3}	2.63×10^{-3}	4	0.938	2.92×10^{-3}	6.95×10^{-3}
5	0.922	4.56×10^{-3}	2.79×10^{-3}	5	0.92	2.79×10^{-3}	8.20×10^{-3}
EXP-NoVaS				EXP-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.945	4.46×10^{-3}	2.66×10^{-3}	1	0.95	3.38×10^{-3}	2.95×10^{-3}
2	0.942	4.25×10^{-3}	2.72×10^{-3}	2	0.958	3.80×10^{-3}	2.72×10^{-3}
3	0.943	4.54×10^{-3}	2.77×10^{-3}	3	0.946	3.75×10^{-3}	2.40×10^{-3}
4	0.949	4.94×10^{-3}	2.72×10^{-3}	4	0.952	3.76×10^{-3}	2.43×10^{-3}
5	0.954	4.72×10^{-3}	3.08×10^{-3}	5	0.946	3.40×10^{-3}	2.84×10^{-3}

Table 31. *Cont.*

L2				**L1**			
Simple-NoVaS				Simple-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.950	4.82×10^{-3}	2.35×10^{-3}	1	0.953	4.01×10^{-3}	3.07×10^{-3}
2	0.963	4.86×10^{-3}	3.26×10^{-3}	2	0.942	3.58×10^{-3}	2.73×10^{-3}
3	0.966	4.85×10^{-3}	2.82×10^{-3}	3	0.952	3.28×10^{-3}	2.48×10^{-3}
4	0.954	5.04×10^{-3}	3.05×10^{-3}	4	0.952	3.51×10^{-3}	2.54×10^{-3}
5	0.944	4.36×10^{-3}	2.51×10^{-3}	5	0.944	4.20×10^{-3}	3.03×10^{-3}
GS-NoVaS				GS-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.950	4.82×10^{-3}	2.35×10^{-3}	1	0.948	3.64×10^{-3}	4.49×10^{-3}
2	0.954	4.47×10^{-3}	2.93×10^{-3}	2	0.956	3.22×10^{-3}	5.92×10^{-3}
3	0.952	4.69×10^{-3}	3.03×10^{-3}	3	0.946	3.31×10^{-3}	4.08×10^{-3}
4	0.950	4.57×10^{-3}	2.99×10^{-3}	4	0.948	3.23×10^{-3}	4.52×10^{-3}
5	0.954	4.50×10^{-3}	3.06×10^{-3}	5	0.95	3.62×10^{-3}	4.54×10^{-3}
GE-NoVaS				GE-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.950	4.90×10^{-3}	2.57×10^{-3}	1	0.954	3.65×10^{-3}	2.78×10^{-3}
2	0.946	4.36×10^{-3}	2.93×10^{-3}	2	0.954	3.86×10^{-3}	2.84×10^{-3}
3	0.948	4.48×10^{-3}	2.82×10^{-3}	3	0.946	3.64×10^{-3}	2.81×10^{-3}
4	0.952	4.58×10^{-3}	2.78×10^{-3}	4	0.95	3.53×10^{-3}	2.89×10^{-3}
5	0.952	4.53×10^{-3}	2.94×10^{-3}	5	0.966	6.38×10^{-3}	3.92×10^{-3}

Table 32. Results of interval predictions for data generated from TV-GARCH(1,1).

L2				**L1**			
GARCH(1,1)				GARCH(1,1)			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.850	1.57×10^{-3}	1.07×10^{-2}	1	0.726	1.12×10^{-3}	4.97×10^{-3}
2	0.848	1.20×10^{-3}	1.60×10^{-3}	2	0.716	1.32×10^{-3}	5.88×10^{-3}
3	0.858	1.84×10^{-3}	2.77×10^{-3}	3	0.718	8.87×10^{-3}	3.90×10^{-3}
4	0.844	2.32×10^{-3}	2.02×10^{-3}	4	0.716	9.78×10^{-3}	4.39×10^{-3}
5	0.856	2.29×10^{-3}	1.68×10^{-3}	5	0.72	1.24×10^{-3}	5.77×10^{-3}
EXP-NoVaS				EXP-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.948	2.40×10^{-3}	2.20×10^{-3}	1	0.952	2.76×10^{-3}	2.45×10^{-3}
2	0.950	2.23×10^{-3}	2.17×10^{-3}	2	0.952	2.74×10^{-3}	2.50×10^{-3}
3	0.952	2.93×10^{-3}	2.15×10^{-3}	3	0.95	2.69×10^{-3}	2.62×10^{-3}
4	0.954	3.02×10^{-3}	2.18×10^{-3}	4	0.942	2.78×10^{-3}	2.65×10^{-3}
5	0.950	2.86×10^{-3}	2.12×10^{-3}	5	0.948	2.82×10^{-3}	2.61×10^{-3}
Simple-NoVaS				Simple-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.950	2.32×10^{-3}	2.37×10^{-3}	1	0.942	2.20×10^{-3}	2.30×10^{-3}
2	0.956	2.20×10^{-3}	2.36×10^{-3}	2	0.956	2.82×10^{-3}	2.45×10^{-3}
3	0.960	2.88×10^{-3}	2.15×10^{-3}	3	0.948	2.60×10^{-3}	2.50×10^{-3}
4	0.952	2.50×10^{-3}	2.27×10^{-3}	4	0.946	2.79×10^{-3}	2.36×10^{-3}
5	0.954	2.80×10^{-3}	2.01×10^{-3}	5	0.946	2.47×10^{-3}	2.62×10^{-3}

Table 32. *Cont.*

L2				**L1**			
GS-NoVaS				GS-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.950	2.32×10^{-3}	2.37×10^{-3}	1	0.942	2.20×10^{-3}	2.30×10^{-3}
2	0.950	2.52×10^{-3}	2.57×10^{-3}	2	0.948	2.57×10^{-3}	2.33×10^{-3}
3	0.952	2.59×10^{-3}	2.53×10^{-3}	3	0.952	2.52×10^{-3}	2.18×10^{-3}
4	0.950	2.37×10^{-3}	2.05×10^{-3}	4	0.946	2.97×10^{-3}	2.22×10^{-3}
5	0.950	2.62×10^{-3}	2.12×10^{-3}	5	0.95	2.64×10^{-3}	2.21×10^{-3}
GE-NoVaS				GE-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.948	2.40×10^{-3}	2.20×10^{-3}	1	0.948	2.77×10^{-3}	2.66×10^{-3}
2	0.950	2.23×10^{-3}	2.17×10^{-3}	2	0.952	2.71×10^{-3}	2.83×10^{-3}
3	0.942	2.47×10^{-3}	2.25×10^{-3}	3	0.95	2.53×10^{-3}	2.50×10^{-3}
4	0.948	2.44×10^{-3}	2.17×10^{-3}	4	0.95	2.75×10^{-3}	2.49×10^{-3}
5	0.949	2.29×10^{-3}	2.12×10^{-3}	5	0.954	2.64×10^{-3}	2.47×10^{-3}

Table 33. Results of interval predictions for data generated from MS-GARCH(1,1).

L2				**L1**			
GARCH(1,1)				GARCH(1,1)			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.868	3.32×10^{-2}	1.78×10^{-2}	1	0.856	3.27×10^{-2}	1.25×10^{-2}
2	0.872	3.42×10^{-2}	1.50×10^{-2}	2	0.89	3.04×10^{-2}	1.10×10^{-2}
3	0.868	3.58×10^{-2}	1.62×10^{-2}	3	0.882	3.10×10^{-2}	1.07×10^{-2}
4	0.858	3.67×10^{-2}	1.86×10^{-2}	4	0.886	3.09×10^{-2}	1.12×10^{-2}
5	0.87	3.60×10^{-2}	2.10×10^{-2}	5	0.908	7.66×10^{-3}	1.28×10^{-2}
EXP-NoVaS				EXP-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.946	4.14×10^{-2}	2.18×10^{-2}	1	0.952	3.97×10^{-2}	2.67×10^{-2}
2	0.948	4.02×10^{-2}	2.26×10^{-2}	2	0.944	4.22×10^{-2}	2.80×10^{-2}
3	0.96	4.78×10^{-2}	2.19×10^{-2}	3	0.958	3.99×10^{-2}	2.74×10^{-2}
4	0.958	4.16×10^{-2}	2.06×10^{-2}	4	0.938	3.86×10^{-2}	2.80×10^{-2}
5	0.956	4.27×10^{-2}	2.07×10^{-2}	5	0.944	4.20×10^{-2}	2.97×10^{-2}
Simple-NoVaS				Simple-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.954	4.21×10^{-2}	2.87×10^{-2}	1	0.958	3.05×10^{-2}	2.03×10^{-2}
2	0.948	3.98×10^{-2}	2.81×10^{-2}	2	0.936	3.08×10^{-2}	2.26×10^{-2}
3	0.94	4.47×10^{-2}	2.91×10^{-2}	3	0.936	3.45×10^{-2}	2.13×10^{-2}
4	0.948	4.26×10^{-2}	2.80×10^{-2}	4	0.94	3.42×10^{-2}	2.28×10^{-2}
5	0.946	4.32×10^{-2}	2.93×10^{-2}	5	0.938	3.52×10^{-2}	2.16×10^{-2}
GS-NoVaS				GS-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.954	4.21×10^{-2}	2.87×10^{-2}	1	0.948	3.21×10^{-2}	2.16×10^{-2}
2	0.948	3.98×10^{-2}	2.81×10^{-2}	2	0.946	3.26×10^{-2}	2.18×10^{-2}
3	0.942	4.75×10^{-2}	2.84×10^{-2}	3	0.948	3.46×10^{-2}	2.28×10^{-2}
4	0.948	4.26×10^{-2}	2.80×10^{-2}	4	0.952	3.17×10^{-2}	2.21×10^{-2}
5	0.946	4.32×10^{-2}	2.93×10^{-2}	5	0.946	3.22×10^{-2}	2.02×10^{-2}

Table 33. *Cont.*

L2				L1			
GE-NoVaS				GE-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.954	4.30×10^{-2}	2.08×10^{-2}	1	0.948	3.40×10^{-2}	2.19×10^{-2}
2	0.952	3.93×10^{-2}	2.09×10^{-2}	2	0.942	3.24×10^{-2}	2.09×10^{-2}
3	0.948	4.36×10^{-2}	2.03×10^{-2}	3	0.946	3.63×10^{-2}	2.15×10^{-2}
4	0.948	4.20×10^{-2}	2.05×10^{-2}	4	0.95	3.08×10^{-2}	2.39×10^{-2}
5	0.95	4.29×10^{-2}	2.07×10^{-2}	5	0.944	3.55×10^{-2}	2.90×10^{-2}

Table 34. Results of interval predictions for data generated from ST-GARCH(1,1).

L2				L1			
GARCH(1,1)				GARCH(1,1).			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.89	2.80×10^{-3}	2.14×10^{-3}	1	0.894	2.53×10^{-3}	3.80×10^{-3}
2	0.888	2.19×10^{-3}	2.79×10^{-3}	2	0.902	2.88×10^{-3}	4.28×10^{-3}
3	0.904	2.15×10^{-3}	2.87×10^{-3}	3	0.884	2.48×10^{-3}	3.20×10^{-3}
4	0.908	2.07×10^{-3}	2.04×10^{-3}	4	0.9012	2.73×10^{-3}	4.21×10^{-3}
5	0.896	2.00×10^{-3}	2.09×10^{-3}	5	0.89	2.02×10^{-3}	4.92×10^{-3}
EXP-NoVaS				EXP-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.956	3.35×10^{-3}	2.57×10^{-3}	1	0.958	3.70×10^{-3}	2.69×10^{-3}
2	0.96	3.40×10^{-3}	2.58×10^{-3}	2	0.944	3.60×10^{-3}	2.54×10^{-3}
3	0.946	3.60×10^{-3}	2.73×10^{-3}	3	0.966	3.64×10^{-3}	2.65×10^{-3}
4	0.942	3.42×10^{-3}	2.51×10^{-3}	4	0.956	3.52×10^{-3}	2.46×10^{-3}
5	0.944	3.56×10^{-3}	2.70×10^{-3}	5	0.962	3.66×10^{-3}	2.53×10^{-3}
Simple-NoVaS				Simple-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.952	3.44×10^{-3}	2.32×10^{-3}	1	0.944	3.44×10^{-3}	2.75×10^{-3}
2	0.956	3.49×10^{-3}	2.24×10^{-3}	2	0.938	3.65×10^{-3}	2.75×10^{-3}
3	0.958	3.46×10^{-3}	2.04×10^{-3}	3	0.934	3.67×10^{-3}	2.81×10^{-3}
4	0.954	3.40×10^{-3}	2.10×10^{-3}	4	0.946	3.60×10^{-3}	2.72×10^{-3}
5	0.946	3.73×10^{-3}	2.19×10^{-3}	5	0.938	3.51×10^{-3}	2.51×10^{-3}
GS-NoVaS				GS-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.952	3.49×10^{-3}	2.18×10^{-3}	1	0.946	3.73×10^{-3}	2.16×10^{-3}
2	0.956	3.30×10^{-3}	2.29×10^{-3}	2	0.956	3.86×10^{-3}	2.31×10^{-3}
3	0.948	3.51×10^{-3}	2.29×10^{-3}	3	0.954	3.84×10^{-3}	2.27×10^{-3}
4	0.948	3.54×10^{-3}	2.36×10^{-3}	4	0.95	3.82×10^{-3}	2.15×10^{-3}
5	0.944	3.56×10^{-3}	2.70×10^{-3}	5	0.946	3.76×10^{-3}	2.22×10^{-3}
GE-NoVaS				GE-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.946	3.37×10^{-3}	2.10×10^{-3}	1	0.954	3.64×10^{-3}	2.60×10^{-3}
2	0.944	3.51×10^{-3}	2.76×10^{-3}	2	0.952	3.60×10^{-3}	2.55×10^{-3}
3	0.946	3.60×10^{-3}	2.73×10^{-3}	3	0.956	3.61×10^{-3}	2.55×10^{-3}
4	0.948	3.44×10^{-3}	2.49×10^{-3}	4	0.95	3.76×10^{-3}	2.68×10^{-3}
5	0.948	3.69×10^{-3}	2.61×10^{-3}	5	0.944	3.46×10^{-3}	2.31×10^{-3}

Table 35. Results of interval predictions for data generated from SV-GARCH(1,1).

L2				**L1**			
GARCH(1,1)				GARCH(1,1)			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.878	2.24×10^{-2}	2.99×10^{-2}	1	0.884	1.22×10^{-2}	1.93×10^{-2}
2	0.862	2.11×10^{-2}	2.51×10^{-2}	2	0.882	1.28×10^{-2}	2.10×10^{-2}
3	0.896	2.49×10^{-2}	2.78×10^{-2}	3	0.872	1.17×10^{-2}	1.60×10^{-2}
4	0.87	2.14×10^{-2}	2.18×10^{-2}	4	0.878	1.26×10^{-2}	1.46×10^{-2}
5	0.892	2.32×10^{-2}	2.33×10^{-2}	5	0.876	1.33×10^{-2}	1.63×10^{-2}
EXP-NoVaS				EXP-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.954	2.71×10^{-2}	2.59×10^{-2}	1	0.944	1.84×10^{-2}	2.33×10^{-2}
2	0.952	2.80×10^{-2}	2.67×10^{-2}	2	0.93	2.01×10^{-2}	2.46×10^{-2}
3	0.956	2.79×10^{-2}	2.66×10^{-2}	3	0.954	1.96×10^{-2}	2.45×10^{-2}
4	0.95	2.84×10^{-2}	2.65×10^{-2}	4	0.942	2.19×10^{-2}	2.62×10^{-2}
5	0.968	3.07×10^{-2}	2.84×10^{-2}	5	0.93	1.88×10^{-2}	2.43×10^{-2}
Simple-NoVaS				Simple-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.952	2.68×10^{-2}	2.79×10^{-2}	1	0.942	2.08×10^{-2}	2.40×10^{-2}
2	0.95	2.77×10^{-2}	2.92×10^{-2}	2	0.932	2.46×10^{-2}	2.53×10^{-2}
3	0.946	2.78×10^{-2}	2.74×10^{-2}	3	0.95	2.63×10^{-2}	2.84×10^{-2}
4	0.958	3.05×10^{-2}	2.90×10^{-2}	4	0.926	2.25×10^{-2}	2.65×10^{-2}
5	0.954	2.66×10^{-2}	2.73×10^{-2}	5	0.934	2.07×10^{-2}	2.41×10^{-2}
GS-NoVaS				GS-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.952	2.68×10^{-2}	2.79×10^{-2}	1	0.95	2.17×10^{-2}	2.42×10^{-2}
2	0.95	2.77×10^{-2}	2.92×10^{-2}	2	0.96	2.05×10^{-2}	2.40×10^{-2}
3	0.952	2.84×10^{-2}	2.96×10^{-2}	3	0.95	2.18×10^{-2}	2.44×10^{-2}
4	0.944	2.28×10^{-2}	2.34×10^{-2}	4	0.95	2.22×10^{-2}	2.48×10^{-2}
5	0.946	2.34×10^{-2}	2.18×10^{-2}	5	0.942	2.28×10^{-2}	2.55×10^{-2}
GE-NoVaS				GE-NoVaS			
STEPS	CVR	LEN	ST.ERR	STEPS	CVR	LEN	ST.ERR
1	0.954	2.71×10^{-2}	2.59×10^{-2}	1	0.952	2.02×10^{-2}	2.35×10^{-2}
2	0.948	2.52×10^{-2}	2.29×10^{-2}	2	0.948	2.18×10^{-2}	2.47×10^{-2}
3	0.956	2.79×10^{-2}	2.66×10^{-2}	3	0.948	1.95×10^{-2}	2.29×10^{-2}
4	0.95	2.84×10^{-2}	2.65×10^{-2}	4	0.942	2.29×10^{-2}	2.58×10^{-2}
5	0.946	2.74×10^{-2}	2.46×10^{-2}	5	0.946	2.10×10^{-2}	2.34×10^{-2}

Our results accentuate the drawbacks of GARCH(1,1) associated with one-step ahead predictions from previous work; see Politis (2015) and the references therein. In particular, NoVaS methods were invariably more robust than GARCH(1,1) fitting when the data have a stochastic structure that deviates from a stationary GARCH(1,1) model, e.g., a time-varying GARCH, a GARCH with structure breaks, etc.

5. Conclusions

In this paper, we derived a new way of multi-step-ahead predictions for ARCH/GARCH and NoVaS methods only based on the basic assumptions of models or transformation. This method has good properties based on our theoretical methodology and simulated results. To sum up:

- The ARCH/GARCH version of our algorithms worked well for data that are generated by a stationary GARCH(1,1) model.
- The NoVaS version of our algorithms worked well for time series data that are either GARCH or have a stochastic structure that deviates from a stationary GARCH(1,1) model.

- There was no apparent error accumulation issue in the multi-step-ahead prediction.
- The methods were theoretically and computationally straightforward.
- Combined with the one-step ahead prediction results in Politis (2015), NoVaS was shown to outperform GARCH model fitting most of the time, for h-step-ahead prediction for all $h \geq 1$.

Author Contributions: Conceptualization, D.N.P.; methodology, J.C. and D.N.P.; software, J.C.; validation, J.C.; formal analysis, J.C.; investigation, J.C.; resources, J.C.; data curation, J.C.; writing original draft preparation, J.C.; writing review and editing, D.N.P.; visualization, J.C.; supervision, D.N.P.; project administration, D.N.P.; funding acquisition, D.N.P.

Funding: This research was partially supported by the National Science Foundation Grant DMS 16-13026.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

References

Abadir, Karim, Alessandra Luati, and Paolo Paruolo. 2018. *The Forecast Density of a Garch(1,1)*. Working Paper. Bologna: University of Bologna.

Andersen, Torben G., Tim Bollerslev, Peter F. Christoffersen, and Francis X. Diebold. 2006. Volatility and Correlation Forecasting. In *Handbook of Economic Forecasting*. Amsterdam: Elsevier, vol. 1, chp. 15, pp. 777–878.

Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner. 1992. ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics* 52: 5–59. [CrossRef]

Bose, Arup, and Kanchan Mukherjee. 2009. Bootstrapping a weighted linear estimator of the arch parameters. *Journal of Time Series Analysis* 30: 315–31. doi:10.1111/j.1467-9892.2009.00613.x. [CrossRef]

Engle, Robert F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50: 987–1007. [CrossRef]

Francq, Christian, and Jean-Michel Zakoian. 2011. *GARCH Models: Structure, Statistical Inference and Financial Applications*. Chichester: John Wiley & Sons.

Pan, Li, and Dimitris N. Politis. 2014. Bootstrap prediction intervals for markov processes. *Computational Statistics & Data Analysis* 100: 467–94.

Pan, Li, and Dimitris N. Politis. 2016. Bootstrap prediction intervals for linear, nonlinear and nonparametric autoregressions. *Journal of Statistical Planning and Inference* 177: 1–27. [CrossRef]

Politis, Dimitris N. 2003. A normalizing and variance–stabilizing transformation for financial time series. In *Recent Advances and Trends in Nonparametric Statistics*. Edited by Michael G. Akritas and Dimitris N. Politis. Amsterdam: JAI, pp. 335–47. doi:10.1016/B978-044451378-6/50022-3. [CrossRef]

Politis, Dimitris N. 2007. Model-free versus Model-based Volatility Prediction. *Journal of Financial Econometrics* 5: 358–59. doi:10.1093/jjfinec/nbm004. [CrossRef]

Politis, Dimitris N. 2015. *Model-Free Prediction and Regression: A Transformation-Based Approach to Inference*. New York: Springer.