



---

*Research article*

## Assessing green bond risk: an empirical investigation

Aikaterini (Katerina) Tsoukala<sup>1</sup> and Georgios Tsiotas<sup>2,\*</sup>

<sup>1</sup> Green Party of Canada, 116 Albert Street Suite 812 Ottawa, Ontario, K1P 5G3, Canada

<sup>2</sup> Department of Economics, University of Crete, Panepistimioupolis, Rethymnon 74100, Greece

\* **Correspondence:** Email: [tsiotas@uoc.gr](mailto:tsiotas@uoc.gr); Tel: +30 2831077412; Fax: +302831077406.

**Abstract:** Green bonds have gained a significant share in the bond market. However, dynamic risk and its spillover to other conventional bond investments plays an important role in its understanding. In this paper, we analyze the volatility and correlation dynamics between conventional bond and green bond assets under both loose and stringent eligibility green-labeled criteria. We build dynamic conditional correlation (DCC) model specifications using alternative distributional assumptions. We also assess risk dynamics expressed by Value-at-Risk (VaR) and its corresponding loss function. We illustrate risk assessment in within and out-of-sample periods using conventional and green bond returns. The results show that there is significant spillover between conventional and green bond assets, triggering significant hedging strategies. However, these spillover effects are subjected to the type of green-labeled criteria. Finally, a risk assessment using VaR forecasting and its corresponding loss function estimation also demonstrates significant differentiation between green and conventional bonds.

**Keywords:** conditional volatility; conditional correlation; value-at-risk; DCC models; forecasting; green bonds

**JEL Codes:** G15, F36

---

### 1. Introduction

Green bonds (hereafter GBs) have gained a significant and growing share in bonds markets. They are mostly used by international and local authorities to raise money for ecologically friendly and climate-friendly projects (ICMA, 2017; OECD, 2017). GBs were first launched in 2007 by the World Bank GBs program and, in 2008, by the European Investment Bank (EIB). The GBs have similarities to conventional bonds (hereafter CBs) in terms of the risk-reward equation and structure. However, by issuing GBs rather than CBs raises somehow the cost because issuers typically need to submit to third-party validation that the proceeds will be spent on environmentally friendly projects to have the bonds certified as green.

These GB securities are directly linked to fund-raising strategies in order to promote and implement investments in mostly renewable energy facilities, green buildings, clean transportation, sustainable waste management, sustainable land use, bio-diversity and clean water. However, questions are being raised regarding the risk of investing in GBs in comparison with conventional bonds. Therefore, issues such as the analysis of volatility dynamics, its effect in forecasting risk and other risk measures such as the Value-at-Risk (hereafter VaR) should be addressed. Furthermore, one needs to investigate how green bond risk is transferred to the conventional bond ones by estimating its spillover effect.

The literature data analysis on green and CBs has covered various issues. Recent research has been devoted to the estimation of the Option-Adjusted spread between these conventional bond and green bond controlled by various bond specific, macroeconomic and global factors (Zerbib, 2017; Nanayakkara and Colombage, 2019). Their findings support that there are strong incentives for investors to raise funds through issuing GBs while providing investors with an opportunity to diversify their investments. Karpf and Mandel (2018) show that, although returns on conventional bonds are on average higher than GBs, the differences can largely be explained by the fundamental properties of the bonds. Also, Flammer (2018) have shown that GBs yield: positive announcement returns, improvements in long-term value and operating performance, improvements in environmental performance, increases in green innovations, and an increase in ownership by long-term and green investors. Gianfrate and Peri (2019) find that, GBs are more financially convenient than non-green ones. The advantage is larger for corporate issuers, and it persists in the secondary market. Other research works are mostly focused on testing the hypothesis of pricing difference using yield and index pricing analyses (Zerbib, 2019; Partridge and Medda, 2020). The purpose of these papers is to test the presence of a statistically significant green premium (“greenium”) in the primary and secondary market. Both yield curve and index pricing analyses show a greenium presence in the secondary market but not in the primary issue for some GBs. Broadstock and Cheng (2019) find that the correlations between green and conventional bonds are time varying and they are sensitive to: changes in financial market volatility; economic policy uncertainty; daily economic activity; oil prices and; uniquely constructed measures of positive and negative news-based sentiment towards GBs. Kanamura (2020) by examining the performance of GBs over conventional bonds, he has demonstrated that the expected returns of green bond premiums are positive while decreasing and that the risks of green bond premiums are slightly decreasing but almost flat over time in the recent years, resulting in positive but decreasing information ratios. This implies that green bond investment performance is superior to conventional bond investment performance but the superiority is decaying over time. Huynh et al. (2020) have shown that portfolios consisting of GBs among others exhibit heavy-tail dependence which implies that in the times of economic turbulence, there will be a high probability of large joint losses. Also, they infer that volatility transmission is higher in the short term, implying that short-term shocks can cause higher volatility in the assets, but in the long run, volatility transmission decreases. In Pham and Huynh (2020), the interaction between investor attention and green bond markets’ performance is investigated.

The authors conclude that there is a time-varying feedback effect between green bond performance and investor attention. Also, Huynh (2020) has examined co-movements between GBs and triple-A government bonds during December 2008–November 2019. Using a heavy-tailed Student’s t-copulas model, the author sees that diversification of GBs with prime government bonds should be done cautiously because of return oscillations as unexpected losses in the financial markets during financial turmoil. Additionally, although GBs have been overlooked owing to their low yield, they also have

potential risks, specifically contagious risk, which should be considered when investing. Recently, Agliardi and Agliardi (2021) propose a structural model for defaultable bonds incorporating both uncertainty about corporate earnings and uncertainty due to climate-related risks. They study how bond pricing is affected by transition risks, such as those coming from an abrupt change of climate policies. They show how the issuer's credit quality changes as a result of its engagement in projects funded by GBs and study the impact of GBs on investors' portfolio allocation.

Finally, other approaches investigate hedging strategies based on dynamic hedge ratio (hereafter HR) estimation using conditional volatility models of the GARCH family (Pham, 2016; Jin et al., 2020). Pham (2016) signifies that a shock in the overall conventional bond market tends to spillover into the green bond market, where this spillover effect is variable over time. Moreover, he reveals a non-constant interaction of volatility between the green bond market and the overall broader bond market due to HR instability. On the other hand, Jin et al. (2020) point out that GBs are strong hedging instruments since the green bond index produces positive HR values in the volatile period while other market indexes such as carbon, energy and commodity index fail to do so.

However, the above estimation strategy adopted by Pham (2016), can be made more robust in the direction of the use of alternative distributional assumptions regarding to the existing competitive models. Other issues also arise regarding the use of other measures of risk such as the VaR and its corresponding loss function. In this paper, we will resolve the model selection problems using univariate and multivariate conditional volatility analysis and the estimation of conditional VaR at various nominal levels. One of the most popular multivariate volatility models is the dynamic conditional correlation (hereafter DCC) model (Engle, 2002; Tse and Tsui, 2002). The main advantage of the DCC model is the positive definiteness of the conditional covariance matrices and its ability to describe time-varying conditional correlations and covariances in a parsimonious way. Concerning parametric estimation, the DCC model can be estimated in two stages, which makes this approach relatively simple and possible to apply even for very large portfolios. Here, we adopt DCC models of alternative distributional assumptions such as: the Normal, the Student-t and the Laplace. Concerning the VaR, this is a quantile risk measure used in finance representing the minimum loss for an asset or portfolio over a pre-specified time horizon for a given probability (confidence level). In our case, time-varying VaR and its corresponding loss measure is derived from competing DCC models. Another issue that is addressed in the present paper is the investigation of the existence of differentiation between alternative GBs indexes and how this is affecting the volatility clustering and volatility spillovers. Within this context, this paper can give answers to the following questions: Is correlation between conventional and GBs subjected to the type of green-labeling? Does correlation between conventional and GBs evolve differently when it comes to adopt either loose or stringent eligibility green-labeled criteria? How does the last differentiation affects hedging strategies.

The objective of this study is three-fold. First, we assess the conditional variance and other risk measures of green bond returns in comparison with the corresponding CBs under alternative volatility specifications. Thus, by using univariate and multivariate conditional volatility specifications of the GARCH family, we compare alternative model specifications for their ability to estimate volatility, conditional covariance and conditional correlation. Special reference is given to the forecasting of the Value-at-Risk and its corresponding loss using the alternative multivariate specifications. Second, under the bivariate analysis which involves both conventional and GBs, we analyze issues such as how volatility spills over between conventional and GBs, whether there is any asymmetry in volatility

transfer and the effect of condition hedging using a Hedge Ratio (HR). Third, we investigate the above volatility and correlation dynamics when CB interacts with GBs under either loose or stringent eligibility green-labeled criteria. This can reveal the existence of potential differentiation in the way one invest when considering GBs of alternative labeling.

The paper is organized as follows. Section 2 reviews the notion of the conditional volatility and the conditional VaR. In the same section, we approach conditional volatility and VaR under a univariate and multivariate version using the DCC models. We specify these models under various distributional assumptions which involves the Normal, the Student-t and the Laplace cases. Special reference is given to the conditional VaR estimation and its corresponding loss function. In Section 3 we illustrate our risk assessment results using real data under univariate and multivariate volatility specifications. Here, a within sample period is used to assess conditional volatility, covariance and the HR between conventional and GBs. Section 4 presents forecasting evaluation results using conditional covariance, VaR and its corresponding loss function measure performed for all of the competing DCC models. Finally, we present our conclusions in Section 5.

## 2. Conditional volatility models

The time-varying risk of bond returns expressed by volatility needs to be investigated. A widely used methodology in modeling volatility is through  $GARCH(p, q)$  models. Under this framework, the bond returns at time  $t$ ,  $y_t$  is related to a conditional mean component plus an error term, such as (Bollerslev, 1986):

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim IID(0, h_t) \quad (1)$$

where the  $\mu_t \sim ARMA(p, q)$  can be an autoregressive moving average process and the  $\epsilon_t$  an independent and identically distributed error term with mean zero and variance that equals  $h_t$ . The latter component then evolves according to the autoregressive moving average structure with  $h_t$  being the conditional variance, such as:

$$h_t = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (2)$$

Estimation is usually based on maximum likelihood where the stationarity condition  $\alpha_1 + \dots + \alpha_p + \beta_1 + \dots + \beta_q < 1$  need to be considered. Here, high and statistically significant values for  $\alpha_i$  and  $\beta_j$  indicate the existence of volatility persistence, where periods of high volatility are followed by periods of high volatility and periods of low volatility are followed by periods of low volatility. Usually, in financial returns analysis and considering likelihood test criteria, the  $GARCH(1, 1)$  model is adopted. This is of the form:

$$h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (3)$$

The time-varying volatility comparison between the green bond and the conventional bond returns can be done using the above equations separately for each series. Therefore, the estimation and testing procedure of a univariate  $GARCH(1, 1)$  model has the following steps:

1. We estimate the mean values  $\hat{\mu}_t$  for each time  $t$  based on an  $ARMA(p, q)$  model (usually an  $ARMA(1, 1)$  is enough).
2. We derive the estimated errors  $\hat{\epsilon}_t = y_t - \hat{\mu}_t$  which are then used as component in the  $GARCH(1, 1)$  specification and the likelihood function derivation.

3. Given the  $\alpha_1 + \beta_1 < 1$  stationarity condition, a conditional maximum likelihood method is used to estimate the  $\omega$ ,  $\alpha_1$  and  $\beta_1$  parameters from a  $GARCH(1, 1)$  model.
4. The estimated volatility  $\hat{h}_t$  is derived together with diagnostic testing score functions such as the AIC and BIC.

However, if we want to consider the interactions between the green bond and the conventional bond market this cannot be taken into account using a univariate GARCH model. Thus, under a multivariate GARCH model the volatility of an asset return doesn't only depend on its past values but also on the volatility of an other asset expressed by the covariance function.

Thus, to understand how GBs' risk evolves dynamically in relation with other CBs we need to investigate its volatility dynamics simultaneously. Engle (2002) has extended the above univariate conditional volatility model to the so-called dynamic conditional correlation model (DCC). According to this, let's suppose a multivariate  $N \times 1$  error process  $\epsilon_t$  which can be written as:

$$\epsilon_t \sim IID(0, cov_t) \quad (4)$$

distributed according to a  $D$  distribution with a  $N \times N$  conditional covariance matrix. Then, its covariance functions evolves such as:

$$cov_t = D_t cor_t D_t \quad (5)$$

with the conditional correlation function

$$cor_t = Q_t^{*-1} Q_t Q_t^{*-1} \quad (6)$$

and

$$Q_t = \left( 1 - \sum_{i=1}^Q \zeta_i - \sum_{j=1}^P \theta_j \right) S + \sum \zeta_i (z_{t-i} z'_{t-i}) + \sum_{j=1}^P \theta_j Q_{t-j} \quad (7)$$

with  $D_t = \text{diag}(h_{1t}^{1/2}, \dots, h_{Nt}^{1/2})$ . Here, the  $z_t$  denotes the standardized IID  $N \times 1$  error such as  $z_t = D_t^{-1} \epsilon_t$ , the  $cor_t$  denotes the  $N \times N$  conditional correlation function of the  $z_t$ , the  $S$  the unconditional correlation of the errors and  $Q_t^*$  the diagonal  $N \times N$  matrix which consists of the square root of the diagonal elements of  $Q_t$ . The above specification represents the so-called DCC( $P, Q$ )-GARCH model. Concerning the parameters, these should have the following stationary condition:  $\sum_{i=1}^Q \zeta_i + \sum_{j=1}^P \theta_j < 1$ .

Here, the green bond returns  $y_{t,G}$  and the conventional bond returns  $y_{t,C}$  formulate a bivariate DCC-GARCH model. The estimation of the DCC's model involves two steps. First, the estimation of the following univariate models:

$$h_{t,G} = \omega_G + \sum_{j=1}^p \alpha_{j,G} \epsilon_{t-j,G}^2 + \sum_{k=1}^q \beta_{k,G} h_{t-k,G} \quad (8)$$

$$h_{t,C} = \omega_C + \sum_{j=1}^p \alpha_{j,C} \epsilon_{t-j,C}^2 + \sum_{k=1}^q \beta_{k,C} h_{t-k,C} \quad (9)$$

for the green bond and the conventional bond index returns respectively. Here,  $\omega_i > 0$ ,  $\alpha_{j,i}, \beta_{k,i} \geq 0$  for  $i = G, C$ ,  $j = 1, \dots, p$  and  $k = 1, \dots, q$ . The covariance stationarity of  $\epsilon_{t,i}$  requires that  $\sum_{j=1}^p \alpha_{j,i} + \sum_{k=1}^q \beta_{k,i} < 1$ . After estimating the above parameters, the standardized residuals are estimated for each

case:  $\hat{z}_{t,i} = \hat{\epsilon}_{t,i} / \sqrt{\hat{h}_{t,i}}$  where  $i = G, C$ . Then, the conditional correlations between the standardized residuals are calculated. Thus, the conditional covariance function

$$\hat{c}_{t,i,j} = (1 - a - b) + a\hat{z}_{t,i}\hat{z}_{t-1,j} + b\hat{c}_{t-1,i,j} \quad (10)$$

is used to produce the conditional correlation function  $\hat{\rho}_{t,i,j} = \hat{c}_{t,i,j} / \sqrt{\hat{c}_{t,i,i} \cdot \hat{c}_{t,j,j}}$ . At the end, the first and the second steps are used to estimate the conditional variance-covariance matrix of the errors of the two univariate errors analyzed (Engle, 2002). The parameters  $a$  and  $b$  express the magnitude of volatility spillover between the analyzed returns. Thus, the estimation and testing procedure of a bivariate DCC(1,1)-GARCH model has the following steps:

1. Steps 1 and 2 of the univariate *GARCH*(1, 1) are followed here for each individual series.
2. The likelihood function is split into two parts: the pure volatility one which is like a linear combination of two univariate likelihood functions and the correlation plus volatility part. Here, volatility for each series is estimated separately using a univariate-likelihood function under the maximum likelihood method.
3. Given the univariate volatility estimates derived in step 2, the  $a$  and  $b$  parameters are estimated using the second part of the likelihood function and applying the maximum likelihood method. Here, the AIC and BIC diagnostic scores are also estimated (Engle, 2002).

In the past, various authors have used DCC-GARCH models of linear and non-linear specifications; such as the DCC-APGARCH, DCC-T-GARCH, and DCC-GJR-GARCH models (Pham, 2016; Reboredo, 2018; Jin et al., 2020). Concerning their results, the two first papers show no systematic evidences for the need for diversification to CBs when buying GBs. However, all of them used the Normality assumption for the errors. In the present analysis, we will adopt three DCC-GARCH model specifications: with Normal, Student-t and Laplace errors.\*

After deriving conditional volatility estimates, one can generate another important risk measure: the VaR. The VaR is a quantile measure that is widely used in finance as an instrument to control and manage risk. This measure represents the minimum loss for an asset or portfolio for a given probability  $\alpha$ . This is used as a mathematical tool designed to offer optimal asset allocation in financial management strategies. The VaR estimates the worst expected loss over a specific time period for a given nominal level  $\alpha$ , such as the following:  $P(y_t < -q_{t,\alpha}) = \alpha$  with  $q_{t,\alpha} \equiv \{r : F(r) \geq \alpha\} \equiv F^{-1}(\alpha)$ , where  $F$  is the conditional distribution function of the  $y_t$  observed financial return,  $F^{-1}$  the inverse conditional distribution, and  $q_{t,\alpha}$  the  $\alpha$ -level quantile of the distribution function of the  $y_t$  conditional on the information up until  $t - 1$  given a vector of returns  $y = (y_1, \dots, y_{t-1})$ . By estimating the worst risk, risk analysts can make optimal asset allocation. Conditional VaR strategies that account for the dependence structure of volatility have been modeled using mostly some fully parametric models such as the GARCH-class (for other approaches in conditional VaR estimation see: Tsiotas, 2018; Tsiotas, 2020). In such a case, the location-scale class is based on the assumption that returns belong to a location-scale family of probability distributions. Using the above approach, we can specify the conditional VaR estimate at the  $\alpha$ -level quantile as:

$$VaR_{t,\alpha} = \mu_t + D^{-1}(\alpha) \sqrt{\hat{h}_t}, \quad (11)$$

\* Here, the probability density function of the Student-t and Laplace errors is given by:  $f(x) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{v\pi}\Gamma(\frac{v}{2})} \left(1 + \frac{x^2}{v}\right)^{-\frac{v+1}{2}}$  and  $f(x | \mu, b) = \frac{1}{2b} \exp\left(-\frac{|x-\mu|}{b}\right)$  respectively.

where  $D^{-1}$  is the inverse CDF for the distribution  $D$  at  $\alpha$  nominal level. In the empirical analysis that follows, we utilize three different assumptions for the innovation distribution: the Normal; the Student-t with  $\nu$  degrees of freedom and the Laplace density.

**Table 1.** Summary statistics for the returns of the AB, GB and GBS index series.

	Statistics and Test						
	mean	sd	skewness	excess kurtosis	min.	max.	Jarque-Bera test
AB	0.0001	0.00193	-0.4482	4.578	-0.0015	0.0011	2283.416(< 2.2e-16)
GB	5.498e-05	0.0037	-0.3622	7.1844	-0.0305	0.0257	5689.947(< 2.2e-16)
GBS	7.848e-05	0.00425	-0.3714	6.9577	-0.0338	0.02713	5341.005(< 2.2e-16)

**Table 2.** VaR evaluation for the AB, GB and GBS return series using empirical, Gaussian and Cornish-Fisher (C-F) methods.

VaR	AB			GB			GBS		
	Empirical	Gaussian	C-F	Empirical	Gaussian	C-F	Empirical	Gaussian	C-F
1%	-0.0046	-0.0044	-0.0015	-0.0110	-0.0085	-0.0179	-0.0119	-0.0097	-0.0180
5%	-0.0029	-0.0030	-0.0070	-0.0056	-0.0061	-0.0066	-0.0062	-0.0068	-0.0067

**Table 3.** Box-Ljung test for squared returns.

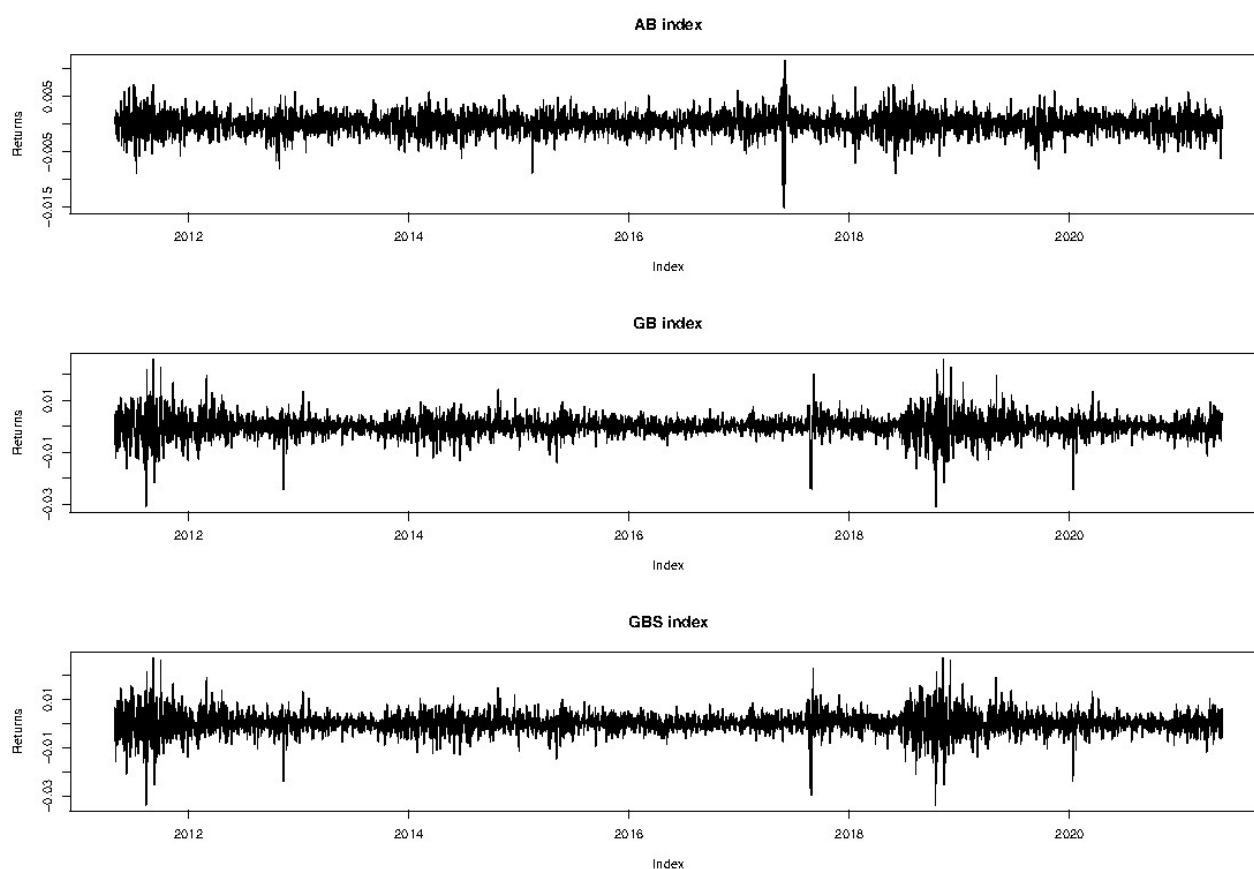
	AB returns	GB returns	GBS returns
Q-stat	606.91	47.481	64.816
p-value	< 2.2e-16	5.553e-12	7.772e-16

### 3. Real data analysis

To make a risk analysis comparison between the conventional bond and the green bond market, we consider the daily time series data from the S&P U.S. Aggregate Bond Index (denoted hereafter as AB), the S&P Green Bond Index (hereafter GB), and S&P Green Bond Select Index (hereafter GBS). The AB measures the performance of U.S. dollar-dominated investment grade U.S. fixed income market, including U.S. Treasuries, quasi-governments, supranational non-U.S. governments and agencies, corporates, covered bonds, residential mortgage pass-throughs and taxable municipal bonds. Concerning its credit rating criteria, the AB has as the minimum credit rating for inclusion in investment-grade indexes with BBB-/Baa3/BBB-. On the other hand, the GB is composed of a variety of global bonds labeled “green” by Climate Bond Initiative (CBI) and subject to eligibility criteria. This index undergoes a re-balancing process once a month, with the intend of keeping the index current. The GBS is a market value-weighted subset of the GB that seeks to measure the performance of green-labeled bonds issued globally, subject to stringent financial and extra-financial eligibility criteria. In other words, the GBS includes those issuers have provided accurate information about the use of proceeds, or whose compliance with the Green Bond Principles around the “Use of Proceeds, Process for Project Evaluation

and Selection, Management of Proceeds and Reporting” has been independently verified. Under this procedure, the GBS calls for higher standards of transparency, disclosure and accountability in the green bond market (S&P, 2021). Whereas, the the credit rating of the GBS requires as the minimum credit rating for inclusion in investment-grade indexes with BBB-/Baa3/BBB-, the GB includes sub-indexes of non-rated bonds issued by U.S. government sponsored enterprises.

For these closing price indexes, we derive the daily returns than span between the 2nd of May 2011 and the 21th of May 2021. Figure 1 presents the returns of the AB, GB and GBS indexes. What we observe is that in all returns the mean values are nearly zero and the GB and the GBS returns have spikes that are more extreme than those observed in the AB returns.



**Figure 1.** Plots of the AB, GB and GBS return series (from top to the bottom).

Before we proceed with our analysis, we summarize some descriptive statistics measures of the AB, GB and GBS return series described in Table 1. These include the mean, standard deviation, minimum, maximum, skewness, excess kurtosis coefficients, and the Jarque-Bera normality test. The mean values show that in all series the return is nearly zero a phenomenon that is widely observed in all stock return series. As regards the standard deviation this indicates an overall variation of the series. Thus, we observe that the GBS return series exhibits the highest variation followed by the GB one. Also, the maximum and the minimum values indicate the upper and the lower range of values returns can take. Again, the GB return indexes show higher risk than the conventional bond return index with



the GB having the lowest minimum value and the GBS the highest maximum value among the bond return indexes. The skewness and excess kurtosis measures assess how much the returns deviate from the Normality assumption. Therefore, we observe that both the AB and GBs have negative skewness. Concerning the excess kurtosis, they all have leptokurtosis with the GB returns having more than the others. Finally, the Jarque-Bera test is a test that examines the existence or not of Normality. It is based on a chi-square distribution with two degrees of freedom. In practice the higher its value the more probable is to reject the Normality assumption. Under this test and using the p-value results for all return series we have Normality violation with the stronger one being observed for the GB return index.

An important statistical measure that can quantify the bond market risk between the green and the conventional market is the VaR. VaR is a number that indicates how much a financial institution can lose with some probability over a given time horizon.

Here, together with the empirical VaR we demonstrated the values under the Normal and Cornish-Fisher (C-F) VaR estimates. The Normal VaR estimates ( $VaR_N$ ) are generated by:

$$VaR_N = \mu + s \cdot q_{\alpha,N} \quad (12)$$

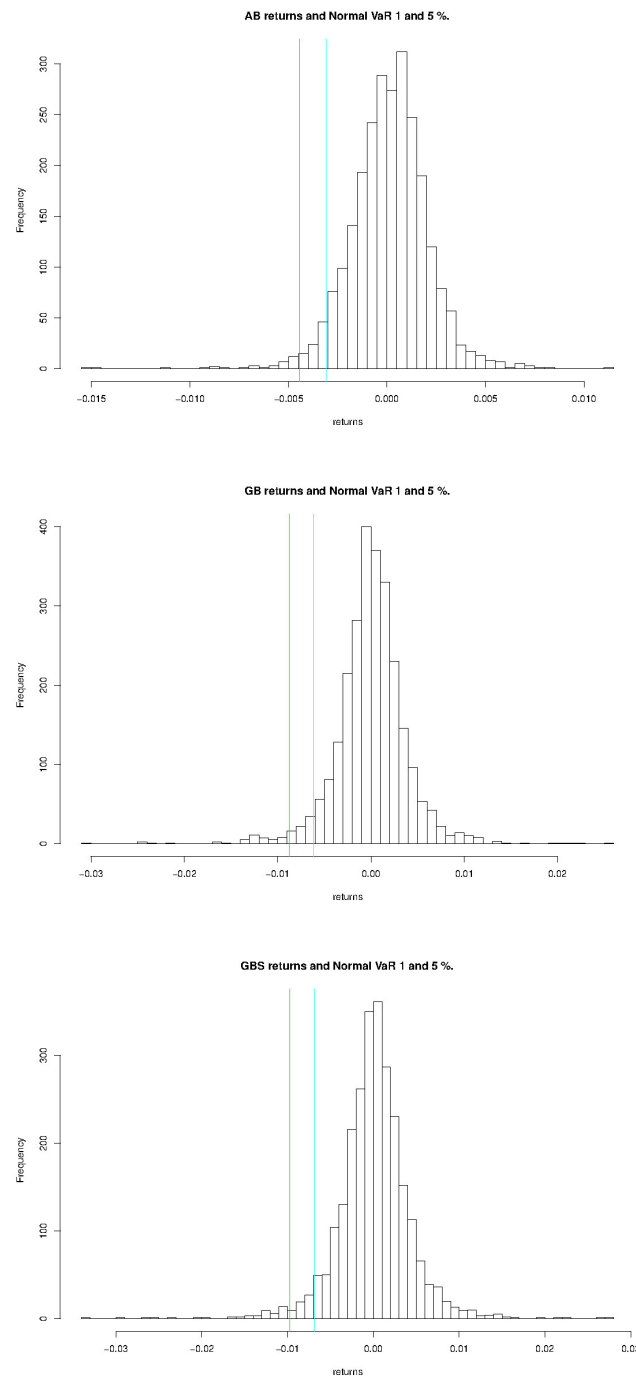
where  $\mu$  and  $s$  are the empirical mean and standard deviation values of the return series respectively and  $q_{\alpha,N}$  is the  $\alpha$ -level percentile under standard Normality (here  $(P(Z \leq q_{N,\alpha}) = \alpha)$ ). A robust type of VaR that can count for Normality violations such as excess kurtosis and skewness can be estimated using C-F VaR estimates. These are generated by first deriving the C-F  $\alpha$ -level percentile as illustrated below:

$$\begin{aligned} q_{CF,\alpha} &= q_{N,\alpha} + \frac{a_1 \cdot (q_{N,\alpha}^2 - 1)}{6} \\ &+ \frac{K \cdot (q_{N,\alpha}^3 - 3q_{N,\alpha})}{24} \\ &+ \frac{a_1^2 \cdot (2q_{N,\alpha}^3 - 5q_{N,\alpha})}{36} \end{aligned} \quad (13)$$

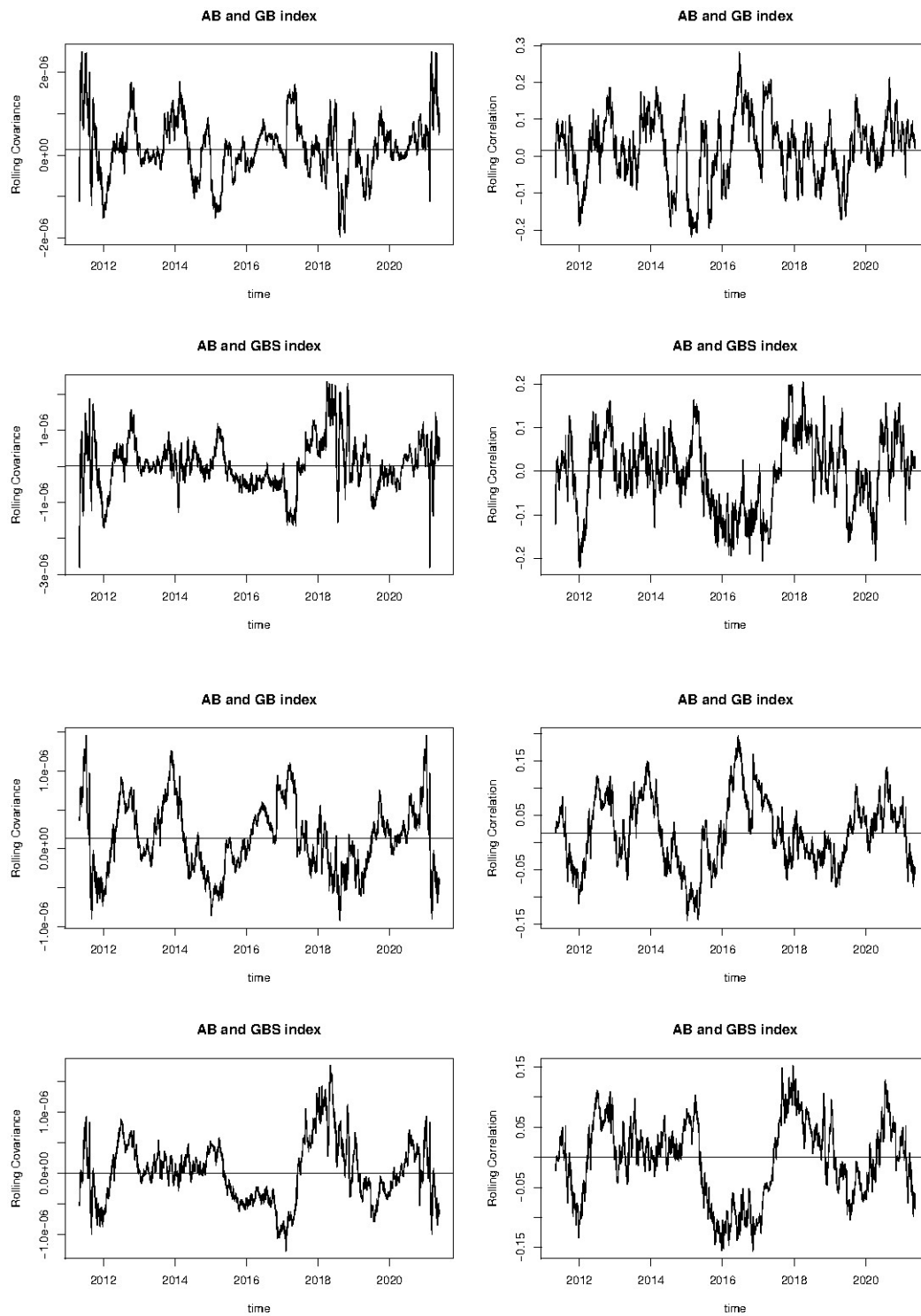
where  $a_1$  is the skewness coefficient and  $K = a_2 - 3$  is the excess kurtosis coefficient. Therefore, C-F VaR estimates ( $VaR_{CF}$ ) are given by:

$$VaR_{CF} = \mu + s \cdot q_{CF,\alpha} \quad (14)$$

In Table 2, we present the  $\alpha = 1\%$  and  $5\%$  VaR of the AB, GB and GBS return series using: the empirical VaR, the Normal VaR and Cornish-Fisher VaR. The empirical one is in accordance to the formal VaR definition. Under all methods the GBs seem riskier to the conventional bond with the GBS showing the highest risk under all methods of estimation for both  $\alpha = 1\%$  and  $5\%$  nominal levels. In Figure 2 we present the  $\alpha = 1\%$  and  $5\%$  VaR of the AB, GB and GBS return using the Normal VaR estimates.



**Figure 2.** Normal VaR at  $\alpha = 1\%$  (in green vertical line) and  $\alpha = 5\%$  (in blue vertical line) for the AB, GB and GBS returns (from the top to the bottom).



**Figure 3.** Rolling covariance and correlation functions between AB and GB and AB and GBS using a 100-day and 200-day rolling window. The horizontal line represents the zero value.

An additional statistical measure that reveals the overall return series characteristics is the rolling conditional covariance and correlation functions. By selecting a 100 and a 200 day window, we estimate the empirical covariance and the correlation between AB and GB and between AB and GBS index return series. Figure 3 demonstrates the results for the 100 and 200 day window respectively. The rolling covariance and correlation are generally very stable with the stability to be increased with the increase of the rolling window. What we can observe is that the rolling covariance and correlation between the AB and the GBS are more stable than those between the AB and the GB a phenomenon that shows that there is more causal relationship between the AB and the GBS than between the AB and the GB indexes.

### 3.1. Univariate GARCH modeling

In this subsection we will do conditional volatility estimation for each return series separately. Before we proceed with a univariate GARCH estimation, we test whether a GARCH model is valid for the available return series. To do so, we need to test for the existence of correlation in the squared returns using the Ljung-Box test. The Ljung-Box statistic uses the squared residuals of an  $ARMA(p, q)$  model to check for GARCH model adequacy (for details see Tsay, 2005). Table 3 reports the statistic of this test using the squared residuals of the  $ARMA(1, 1)$  model. Under the p-value results, the null hypothesis which signifies that there are not GARCH effects is rejected even at the 1% significance level for all return series. Another, less formal test for the validity or not of GARCH modeling is to assess the absolute and the squared return values for serial correlation using autocorrelation functions (ACF). In Figure 4, we report the ACF of the observed, the absolute and the squared AB, GB and GBS return series. These indicate that there is nearly zero correlation in the returns and significant correlation in both the absolute and squared returns for all series. This signifies that the series can exhibit significant correlation in their volatility (volatility clustering). In such a case, periods of high volatility may be followed by periods of high volatility and periods of low volatility may be followed by periods of low volatility.

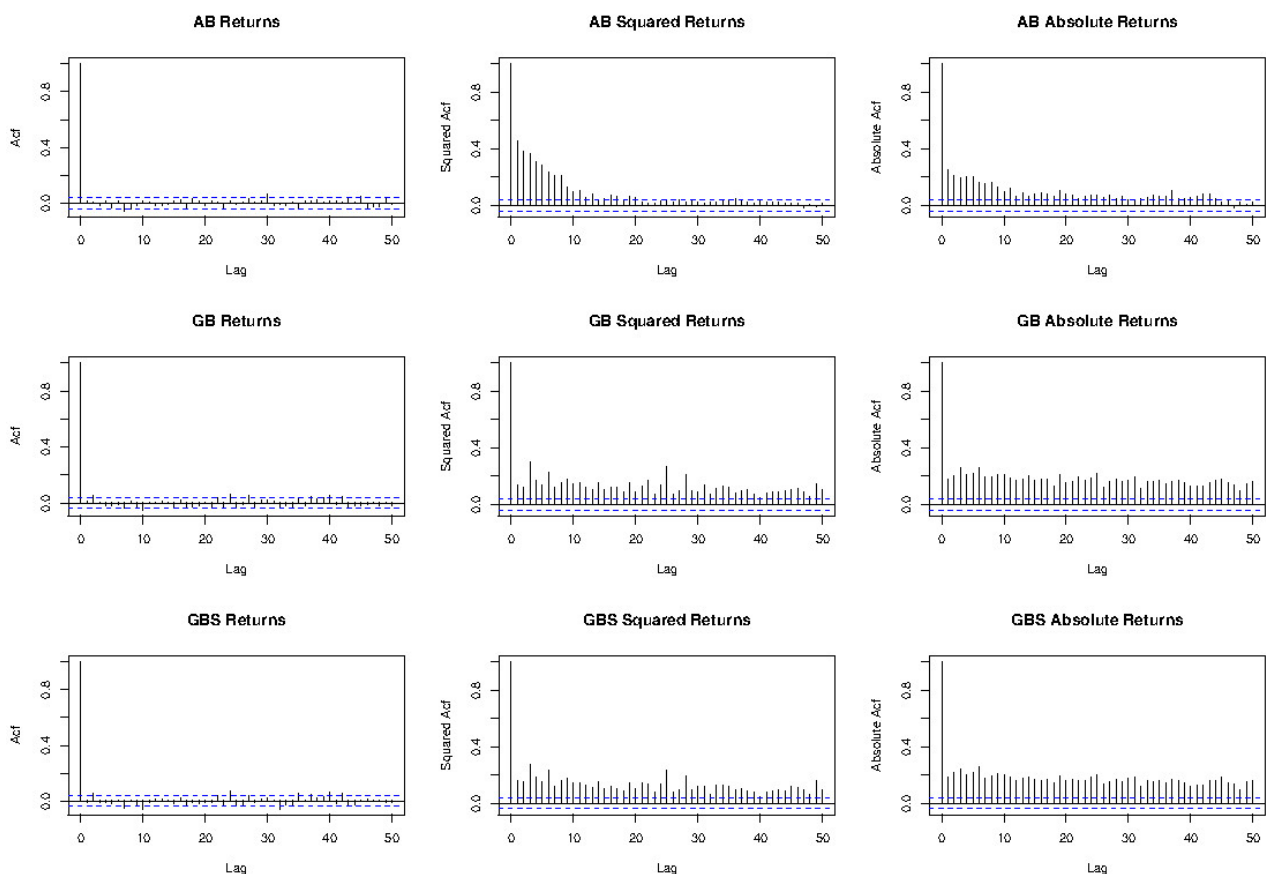
After this formal and informal test for the existence of the conditional volatility effect, we run univariate volatility estimation for the three return series using the whole sample. Initially, we consider alternative Autoregressive Conditional heteroscedasticity (ARCH) and GARCH models of various  $p$  and  $q$  specifications. However, as the extended financial literature suggests, the  $GARCH(1, 1)$  process is the best specification when it comes to estimate volatility dynamics (Tsay, 2005). The parametric model that is considered here is the  $GARCH(1, 1)$  model with Gaussian and Student-t errors. Much of the literature on conditional volatility estimation and forecasting uses these models as benchmarks. The models are specified as:

$$y_t = \mu_t + \epsilon_t, \quad \epsilon_t \sim IID(0, h_t) \quad (15)$$

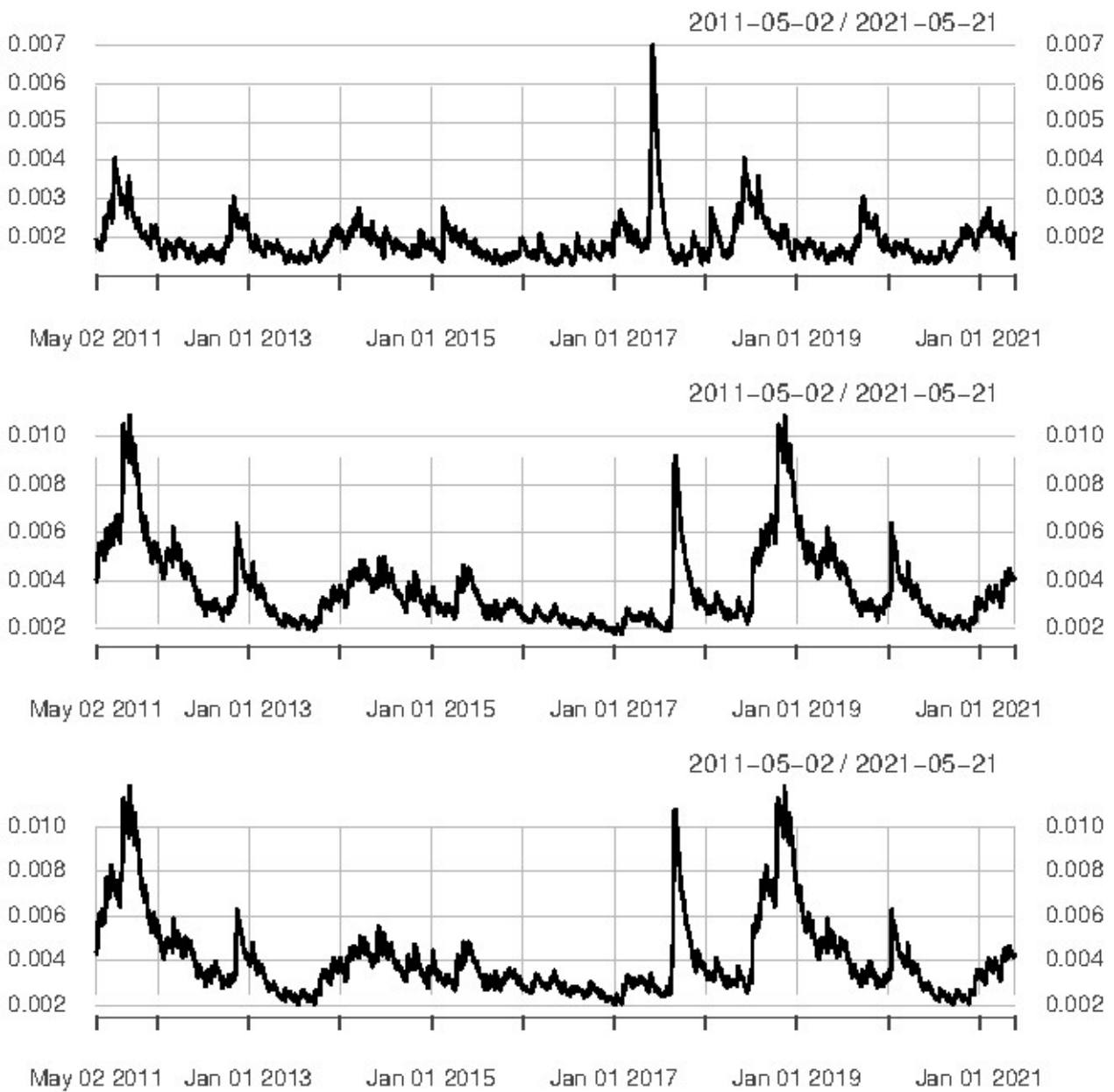
$$h_t = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (16)$$

for distribution D with Normal and Student-t errors. Table 4 presents the estimated parameters of the univariate  $GARCH(1, 1)$  model for the AB, the GB and the GBS return series under the Normal and the Student-t distribution assumptions. To perform the models fitness we use the AIC and the BIC score functions. Under these measures the Student-t  $GARCH(1, 1)$  model is the best performer since it minimizes both scores under all data cases. The reported parametric estimates that correspond to the  $\alpha_1$ ,  $\beta_1$  and  $\nu$  have a high level of significance in all data cases. In terms of the measure of persistence in conditional volatility ( $\alpha_1 + \beta_1$ ), the GB return series have the highest persistence with the GBS following

and the AB in the third place. This is translated into more volatility clustering for the green bond indexes compared to the conventional one. Also, using the  $\ln 0.5 / (\ln(\alpha_1 + \beta_1))$  measure one can count the days needed to reduce a volatility shock by 50% (see Pham, 2016 for details). This is estimated in 17.7207 days for the AB index, 137.2096 days for the GB index and 132.2114 days for the GBS one. Again, the volatility impact of the GB index is considered the highest compared to the other two. If we see the conditional variance estimates using the Student-t model, presented in Figure 5, we can diagnose that although the GB index has a higher volatility clustering compared to the GBS the latter has higher minimum and maximum volatility values with a very high volatility spike observed in mid of 2014. This characteristic comes in accordance to the VaR results where the GBS index has been considered as the riskiest asset compared to the AB and GB ones. Concerning the kurtosis parameter  $\nu$ , again the GB index returns show higher deviation from Normality than the CB. Here, the GBS series show the highest deviation since the estimated degrees of freedom parameter  $\nu$  is the lowest one.



**Figure 4.** Autocorrelation functions of the returns, absolute returns and squared returns of the AB, GB and GBS series.



**Figure 5.** Conditional standard deviation using the univariate GARCH(1,1) model under Student-t errors and using the AB index returns, the GB index returns and the GBS index returns (from the top to the bottom).

**Table 4.** Univariate *GARCH*(1, 1) estimation for AB, GB and GBS return series under the Normal and the Student-t (Std) model.

Parameter	AB-Normal				AB-Std			
	Estimate	Std. Error	t value	$Pr(>  t )$	Estimate	Sd	t value	$Pr(>  t )$
$\mu$	1.197e-04	2.870e-05	4.170	3.05e-05	1.114e-04	2.869e-05	3.883	0.000103
$\omega$	1.479e-07	3.384e-08	4.373	1.23e-05	1.384e-07	3.220e-08	4.298	1.72e-05
$\alpha_1$	0.0715	9.845e-03	7.266	3.69e-13	0.06984	9.537e-03	7.323	2.42e-13
$\beta_1$	0.8875	1.677e-02	52.923	< 2e-16	0.8918	1.610e-02	55.392	< 2e-16
$\nu$					9.0207	0.9812	9.1935	< 2e-16
AIC	-9.7446				-9.754			
BIC	-9.7479				-9.729			
Parameter	GB-Normal				GB-Std			
	Estimate	Sd	t value	$Pr(>  t )$	Estimate	Sd	t value	$Pr(>  t )$
$\mu$	6.124e-05	4.942e-05	1.239	0.215276	0.000109	0.000054	2.01696	0.043699
$\omega$	8.604e-08	2.285e-08	3.765	0.000167	0.000000	0.000000	0.39728	0.691158
$\alpha_1$	0.0576	5.830e-03	9.890	< 2e-16	0.053898	0.006589	8.17991	0.000000
$\beta_1$	0.9378	5.930e-03	158.144	< 2e-16	0.941063	0.005748	163.7170	0.000000
$\nu$					6.736408	0.781254	8.62256	0.000000
AIC	-8.5107				-8.864			
BIC	-8.5040				-8.625			
Parameter	GBS-Normal				GBS-Std			
	Estimate	Sd	t value	$Pr(>  t )$	Estimate	Sd	t value	$Pr(>  t )$
$\mu$	6.752e-05	5.541e-05	1.219	0.223	0.000117	0.000060	1.94982	0.051197
$\omega$	1.239e-07	3.150e-08	3.935	8.33e-05	0.000000	0.000000	0.49532	0.620374
$\alpha_1$	0.0543	5.812e-03	9.349	< 2e-16	0.050485	0.006015	8.39386	0.000000
$\beta_1$	0.9393	6.243e-03	150.463	< 2e-16	0.944286	0.005254	179.7283	0.000000
$\nu$					6.396175	0.697954	9.16418	0.000000
AIC	-8.3067				-8.405			
BIC	-8.2999				-8.389			

Note: Entries in each table represent the estimated parameter (Estimate), the standard deviation (sd), the t-stat under the zero parameter hypothesis and the p-value  $Pr(> |t|)$  of the same test.

**Table 5.** AIC and BIC scores for alternative Bivariate DCC(1, 1)-GARCH models on AB-GB and AB-GBS return series .

		scores	
		AIC	BIC
AB-GB	MV-Normal	-18.323	-18.297
	MV-Student-t	-18.366	-18.337
	MV-Laplace	-18.164	-18.144
AB-GBS	MV-Normal	-18.057	-18.036
	MV-Student-t	-18.103	-18.083
	MV-Laplace	-17.954	-17.936

Note: The boxed numbers indicate the favored estimation case.

### 3.2. Bivariate GARCH modeling

In the univariate GARCH analysis we have captured volatility dynamics for the individual time series. However, this fails to capture a potential volatility interaction (or spillover) between the time series especially those between GBs and CBs. As seen on Figure 5, there are periods where the bond returns' volatility moves towards the same and in some other periods the opposite direction. In this subsection, we intend to analyze the volatility movements of the GB and the GBS returns in relation with those expressed by conventional bond returns such as the AB ones. In practice we will demonstrate the results of the bivariate DCC(1, 1)-GARCH model applied in two cases: the AB with the GB (hereafter AB-GB) return series and the AB with the GBS return series (hereafter AB-GBS).

We have specified three DCC-GARCH specifications depending on the distribution assumption assumed: The Multivariate Normal (hereafter MV-Normal), the Multivariate Student-t (hereafter MV-Std) and the Multivariate Laplace (hereafter MV-Laplace). Under these specifications, we estimate the Bivariate models for the AB-GB and AB-GBS index return series. In Table 5 we demonstrate two diagnostic testing measures such as the AIC and the BIC scores. Results show that the MV-Std is the best performer under the two multivariate data cases. Here, we need to point out that other DCC-GARCH specifications such as the asymmetric (DCC-GJR-GARCH), with threshold (DCC-T-GARCH) and the flexible DCC-GARCH cases do not perform better than the "standard" DCC-GARCH models. More specifically, the asymmetric DCC-GARCH model aims to count for the leverage effect observed in many financial data. As results have shown (NB: not demonstrated in the present manuscript), the asymmetric coefficient is very small and statistically insignificant signifying that the volatility transfer between green and CBs is not affected by the leverage effect.

In Table 6 we summarize our results under the MV-Std DCC-GARCH model. First, in the AB-GB case we see that in the p-value results of all important variables that influence volatility estimation are statistically significant at the 1% significance level (they have p-value < 1%). Again, the GB index exhibit higher volatility clustering compared to the AB index. Moreover, the positive and statistical significant value of  $b_{AB,GB}$  signifies a volatility spillover between the GB and the AB index. This is translated into a positive conditional correlation of a small value. In Figure 6, we see how the conditional standard deviation of the GB return series is larger compared to that of the AB ones. This comes in accordance with the univariate GARCH results. Also, in Figure 8, we see that the conditional correlation between these two

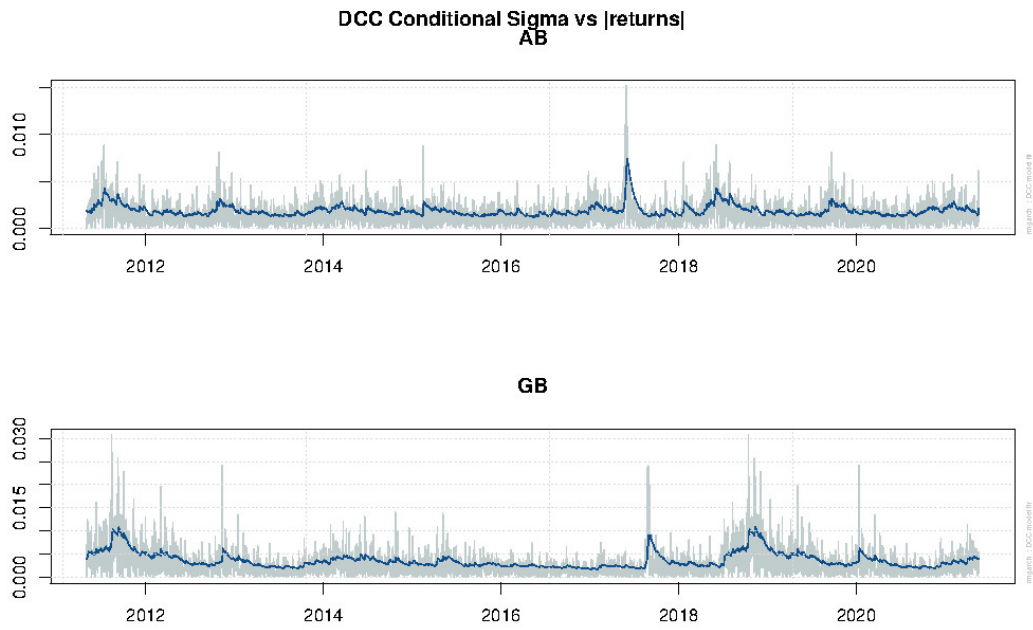


return series is unstable. However, on average, the GB is positively correlated with the market benchmark AB index. This signifies that the use of environmentally friendly projects has sometimes created different patterns from conventional bonds. Secondly, in the AB-GBS case we see that the p-value results of all important variables that influence volatility estimation are statistically significant at the 1% significance level (they have p-value < 1%) with the exception of the  $\alpha_{1,GBS}$  which is significant at 10% significance level (they have p-value < 10%). Here, the AB and the GBS indexes exhibit similar volatility clustering. The positive and statistical significant value of  $b_{AB,GBS}$  also signifies a volatility spillover between the GBS and the AB indexes. In Figure 7 we see the conditional standard deviation of the GBS return series to be systematically larger compared to that of the AB ones. This comes in accordance with the univariate GARCH results.

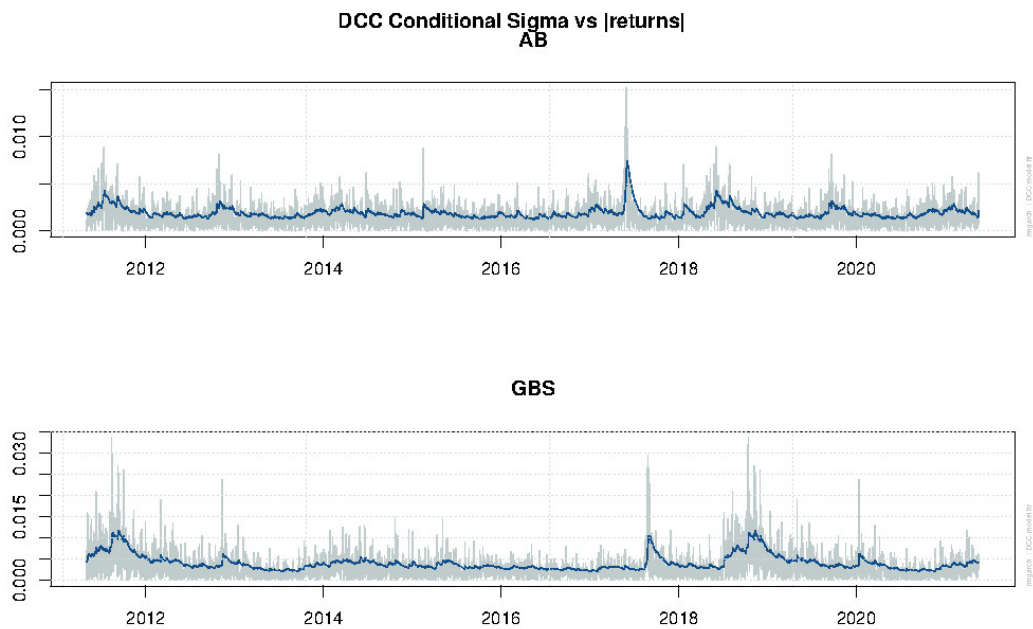
**Table 6.** Bivariate Student-t DCC(1, 1)-GARCH model results for AB-GB return series and AB-GBS return series.

<b>AB-GB</b>				
Parameter	Estimate	Std. Error	t value	$Pr(>  t )$
$\mu_{AB}$	0.000137	0.000033	4.119006	0.000038
$\omega_{AB}$	0.000000	0.000000	0.106180	0.915440
$\alpha_{1,AB}$	0.084308	0.024715	3.411237	0.000647
$\beta_{1,AB}$	0.913232	0.020377	44.816288	0.000000
$\mu_{GB}$	0.000089	0.000059	1.502241	0.133035
$\omega_{GB}$	0.000000	0.000002	0.036071	0.971226
$\alpha_{1,GB}$	0.061456	0.059537	1.032232	0.301963
$\beta_{1,GB}$	0.935438	0.054711	17.097768	0.000000
$a_{AB,GB}$	0.008254	0.011714	0.704645	0.481031
$b_{AB,GB}$	0.841273	0.144074	5.839178	0.000000
$\nu$	9.316885	0.494763	18.830997	0.000000
<b>AB-GBS</b>				
Parameter	Estimate	Std. Error	t value	$Pr(>  t )$
$\mu_{AB}$	0.000114	0.000029	3.996301	0.000064
$\omega_{AB}$	0.000000	0.000002	0.028356	0.977378
$\alpha_{1,AB}$	0.058638	0.115101	0.509447	0.610439
$\beta_{1,AB}$	0.924675	0.108233	8.543392	0.000000
$\mu_{GBS}$	0.000060	0.000056	1.055131	0.291366
$\omega_{GBS}$	0.000000	0.000002	0.071799	0.942762
$\alpha_{1,GBS}$	0.053057	0.037221	1.425444	0.154029
$\beta_{1,GBS}$	0.941117	0.035894	26.219468	0.000000
$a_{AB,GBS}$	0.000000	0.000272	0.001721	0.998627
$b_{AB,GB}$	0.906743	0.106205	8.537646	0.000000
$\nu$	10.351443	1.198938	8.633844	0.000000

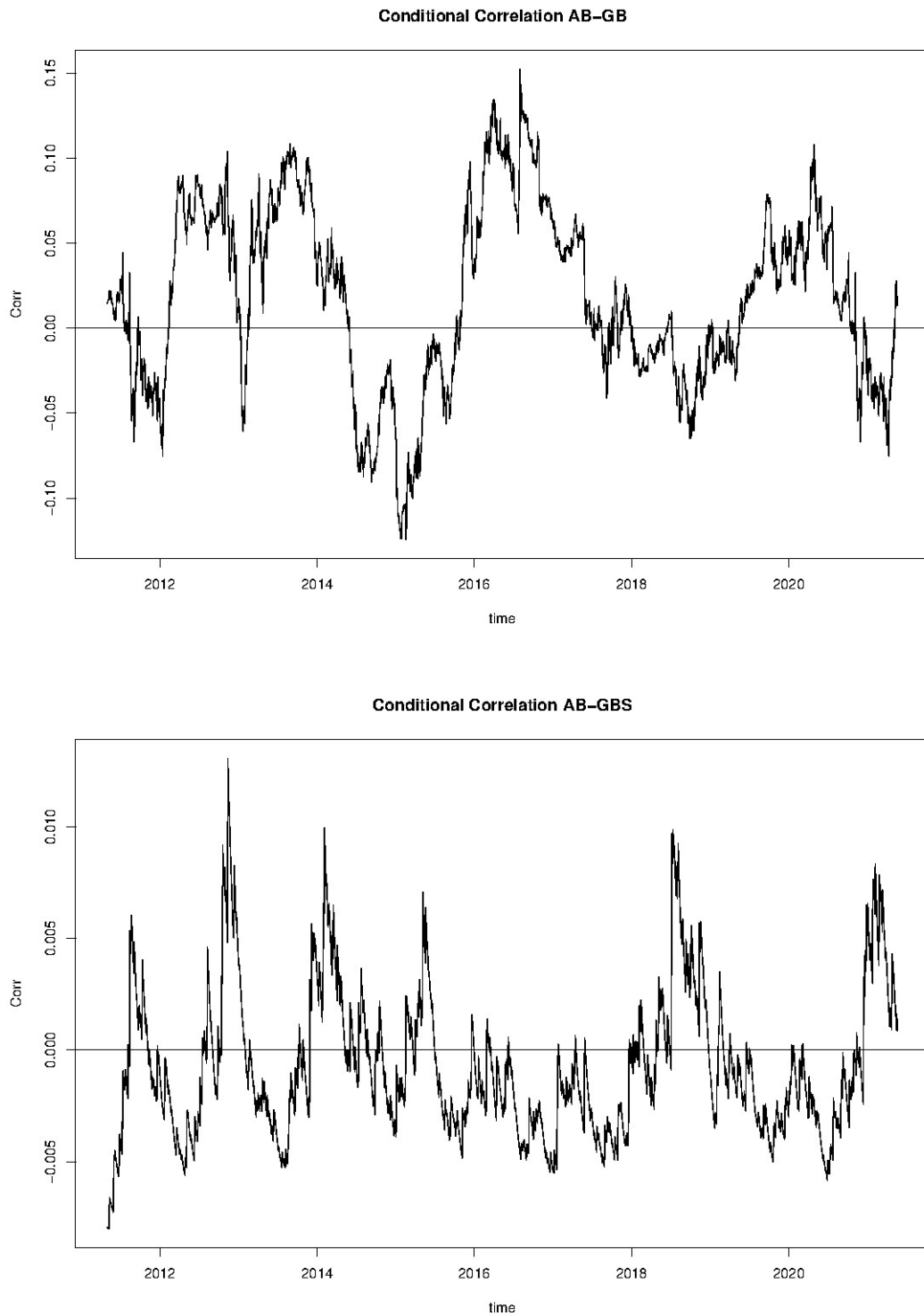
Note: Entries in each table represent the estimated parameter (Estimate), the standard deviation (sd), the t-stat under the zero parameter hypothesis and the p-value  $Pr(> |t|)$  of the same test.



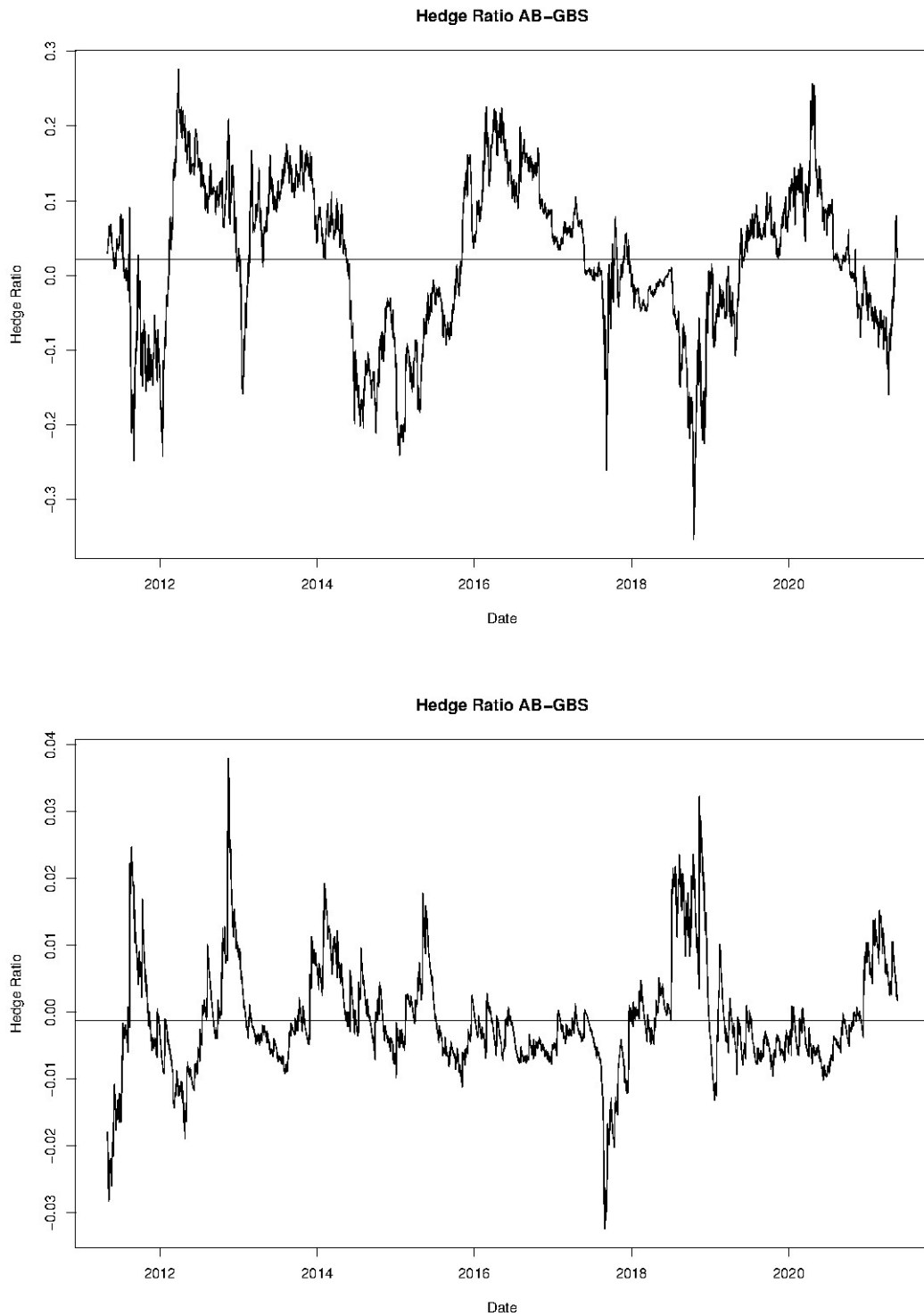
**Figure 6.** Conditional standard deviation for AB and GB cases computed using the Bivariate DCC(1, 1)-GARCH model with Student-t errors.



**Figure 7.** Conditional standard deviation for AB and GBS cases computed using the Bivariate DCC(1, 1)-GARCH model with Student-t errors.



**Figure 8.** Conditional correlation for the AB-GB and the AB-GBS cases computed using the Bivariate DCC(1, 1)-GARCH model with Student-t errors. The horizontal line represents the zero value.



**Figure 9.** Hedge ratio for the AB-GB and the AB-GBS cases computed using the Bivariate DCC(1, 1)-GARCH model with Student-t errors. The horizontal line represents the mean hedge ratio value.

Moreover, in Figure 8, we see the conditional correlation between these two return series. In contrary to the AB-GB case, the AB-GBS ones show on average negative conditional correlation which is less stable than the corresponding in the AB-GB case. Another important finding is that, for the period after the 2015, positive correlation values between the AB-GB indexes is associated with negative correlation values between the AB-GBS indexes and vice versa. This signifies an important sign of differentiation between the GB and GBS.

Using the above conditional estimates of standard deviations and covariance functions, we use these results to derive HRs. The HR between the AB and a green bond returns is given by (see Kroner and Sultan, 1993):

$$\gamma_{AB,j,t} = \frac{\sigma_{AB,j,t}}{\sigma_{AB,t}^2} \quad (17)$$

with  $j = GB, GBS$  where  $\sigma_{AB,j,t}$  to denote the covariance between the AB and the green bond return indexes at time  $t$  and  $\sigma_{AB,t}^2$  to denote the variance of the AB index at time  $t$ <sup>†</sup>. Generally, a positive HR shows the extent to which a long position in a green bond can be hedged by a short position in the conventional bond. On the other hand, a negative HR shows the extent to which a short position in the green bond can be hedged by a long position in the conventional bond.

In Figure 9 we show the time-varying HR for the AB-GB and the AB-GBS cases. First, we observe that on average the HR value of the AB-GB case is positive and the corresponding one for the AB-GBS case is negative. This indicates that, for the AB-GB case, a long position in the GB should be hedged by a short position in the AB asset which price is expected to drop. By having an average HR value of 0.0219, one can say that the optimal quantity of hedge for 1 million dollar of investment is 0.0219 million dollars. On the other hand, a negative HR value for the AB-GBS case indicates that, a short position in the GBS can be hedged by a long position in the AB which price is expected to rise. By having an average HR value of  $-0.0012$ , one can say that the optimal quantity of hedge for 1 million dollar of investment is 0.0012 million dollars. Looking at the dynamics of the HRs, we observe significant differences in hedging strategies where periods with a positive HR values between the AB-GB indexes is associated with negative HR values between the AB-GBS indexes and vice versa. This again signifies the existence of differentiation strategies one should adopt when he/she considers hedging. Also, the magnitude of the absolute HR value is larger in the AB-GB than in the AB-GBS case signifying that the level of hedging is stronger in the AB-GB than in the AB-GBS case.

To robustify the above estimation results, we present the AIC and the BIC scores when our total sample is divided into three sub-samples. Here, we denote as period I the one that than span between the 2nd of May 2011 and the 22th of August 2014, as period II the one that than span between the 25th of August 2014 and the 15th of January 2018, and as period III the one that than span between the 16th of January 2018 and the 21th of May 2021. For these periods and for the AB-GB and the AB-GBS cases, we have estimated the MV-Normal, the MV-Std and MV-Laplace models. In Table 7 we demonstrate two diagnostic testing measures such as the AIC and the BIC scores. Again, results show that the MV-Std is the best performer under the two multivariate data cases for all three periods.

<sup>†</sup> As stated by Chen and Sutcliffe (2012), the average position in the hedging instrument, or the average HR, can be considered as a proxy for the transaction cost.

3.2.1. Portfolio assessment using a portfolio rebalancing method

In this paragraph, we assess the portfolio weights adopted for both the AB-GB and the AB-GBS cases using portfolio rebalancing. Portfolio rebalancing is an important part of the portfolio management process. Usually, once a portfolio manager determines a target portfolio, maintaining this balance of assets is non-trivial. Therefore, due to changes in portfolio return and risk, a manager must rebalance actively the portfolio weights in an optimal way. Conventional approaches to portfolio rebalancing include periodic and tolerance band rebalancing (Masters, 2003). With periodic rebalancing, the portfolio manager adjusts to the target weights at a consistent time interval (e.g., monthly or quarterly). The time-varying weights are estimated using an objective function which is usually based on mean and variance estimates of the portfolio assets.

Here, we use the quadratic utility function, with it expected value being of the form:

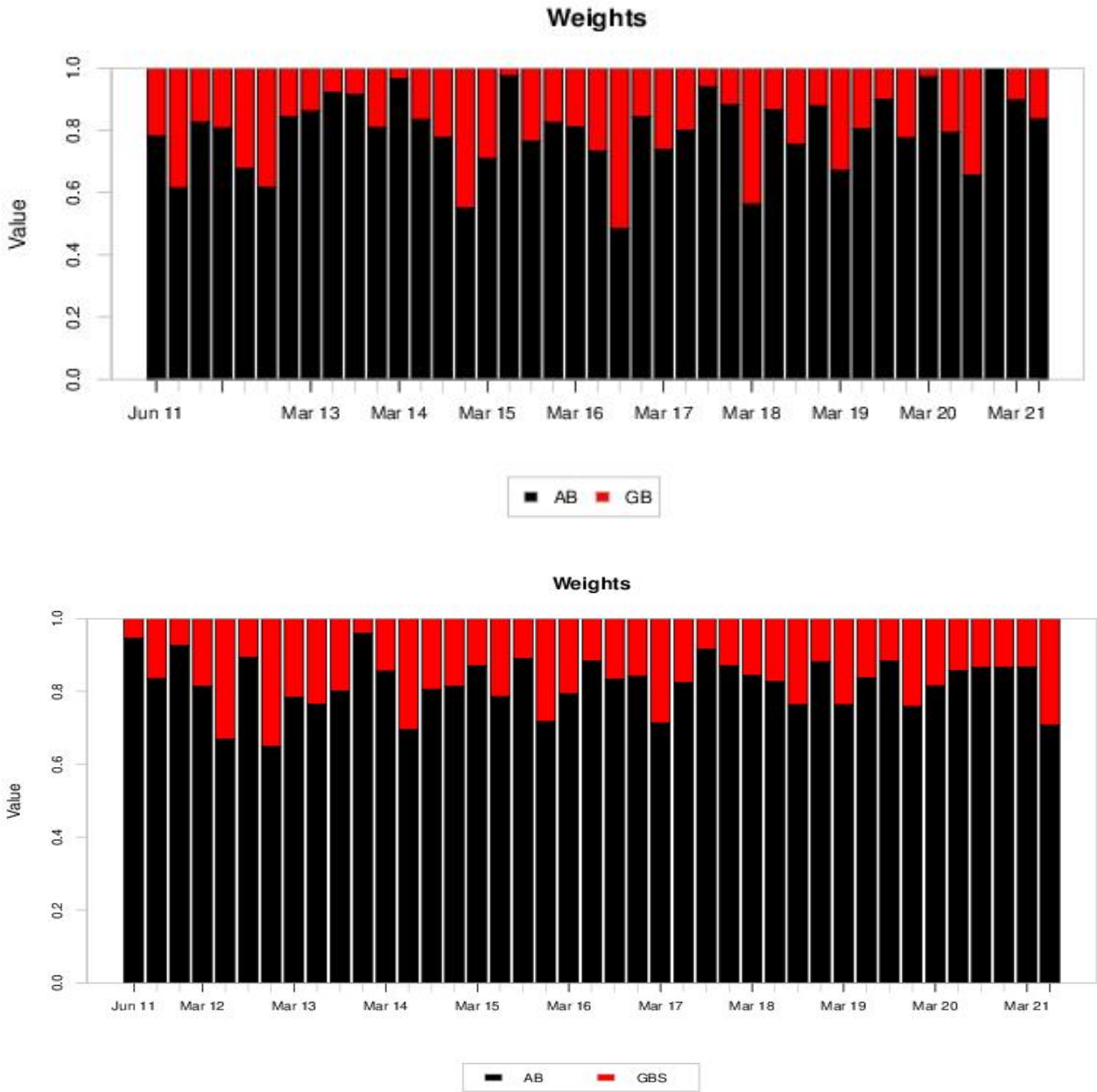
$$U_q(\mu_p, \sigma_p) = \mu_p - \frac{\phi}{2} \sigma_p^2 \tag{18}$$

with  $\mu_p$  the portfolio mean,  $\sigma_p$  the portfolio standard deviation and  $\phi$  the risk aversion parameter<sup>‡</sup>. In practice, from the selected MV-Std DCC-GARCH model, we draw random samples for the mean return values from each series  $i$  based on  $\hat{\mu}_i + \hat{\sigma}_{t,i} \sqrt{\frac{\hat{\nu}-2}{\hat{\nu}}} \cdot T_{\hat{\nu}}(0, 1)$ . Here,  $T_{\hat{\nu}}(0, 1)$  denotes a Student-t random variable with mean 0, variance 1 and  $\hat{\nu}$  degrees of freedom which are estimated. The  $\hat{\mu}_i$  and  $\hat{\sigma}_{t,i}$  components represent the corresponding mean and variance estimates of each series. After, we derive the portfolio mean value  $\mu_p = \sum_{i=1}^2 w_i \mu_i$  and the portfolio variance value  $\hat{\sigma}_p^2 = \sum_{i=1}^2 w_i \hat{\sigma}_i^2 + 2w_1 w_2 \hat{c}_{t,1,2}$ . Thus, the  $w_1$  and  $w_2$  weight components are derived optimally under the  $w_1 + w_2 = 1$  restriction. In Figure 10, we show the time-varying portfolio weights for both the AB-GB and AB-GBS cases. What is diagnosed is that the AB dominates both the GB and GBS indexes in most of the observed quarters. Also, one can observe that the GB weights are more volatile and of marginally higher value than the corresponding GBS weights. A recent work by Nguyen and Huynh (2019) has used an alternative approach to derive portfolio weights. Under this paper’s approach, a risk measure based on conditional value-at-risk (CVaR) is used in order to achieved the maximal expected return for investments.

**Table 7.** AIC and BIC scores for alternative Bivariate DCC(1, 1)-GARCH models on AB-GB and AB-GBS return series using three sub-periods (boxed numbers indicate the favored estimation case).

		period I		period II		period III	
		AIC	BIC	AIC	BIC	AIC	BIC
AB-GB	MV-Normal	-18.025	-17.994	-18.742	-18.690	-18.639	-18.581
	MV-Student-t	<span style="border: 1px solid black; padding: 2px;">-18.070</span>	<span style="border: 1px solid black; padding: 2px;">-18.031</span>	<span style="border: 1px solid black; padding: 2px;">-18.789</span>	<span style="border: 1px solid black; padding: 2px;">-18.724</span>	<span style="border: 1px solid black; padding: 2px;">-18.680</span>	<span style="border: 1px solid black; padding: 2px;">-18.607</span>
	MV-Laplace	-17.914	-17.882	-18.646	-18.594	-18.549	-18.490
AB-GBS	MV-Normal	-17.889	-17.858	-18.451	-18.399	-18.439	-18.302
	MV-Student-t	<span style="border: 1px solid black; padding: 2px;">-17.943</span>	<span style="border: 1px solid black; padding: 2px;">-17.904</span>	<span style="border: 1px solid black; padding: 2px;">-18.483</span>	<span style="border: 1px solid black; padding: 2px;">-18.418</span>	<span style="border: 1px solid black; padding: 2px;">-18.837</span>	<span style="border: 1px solid black; padding: 2px;">-18.453</span>
	MV-Laplace	-17.789	-17.757	-18.324	-18.272	-18.038	-17.938

<sup>‡</sup> Here, we suppose a risk aversion parameter that equals 0.25.



**Figure 10.** Time-varying portfolio weights for the AB-GB and the AB-GBS cases computed using the Bivariate DCC(1, 1)-GARCH model with Student-t errors.

**Table 8.** Bivariate Student-t DCC(1, 1)-GARCH model mean results for AB-GB return series and AB-GBS return series using the pro-COVID-19 and COVID-19 periods.

	<b>AB-GB</b>	
	pro-COVID-19	COVID-19
$\alpha_{1,AB}$	0.035578	0.256625
$\beta_{1,AB}$	0.959976	0.663642
$\alpha_{1,GB}$	0.034632	0.156793
$\beta_{1,GB}$	0.961384	0.760333
$a_{AB,GB}$	0.013729	0.017160
$b_{AB,GB}$	0.856911	0.931651
Correlation	0.018410	0.01443
HR	0.019897	0.034698
	<b>AB-GBS</b>	
	pro-COVID-19	COVID-19
$\alpha_{1,AB}$	0.034412	0.249544
$\beta_{1,AB}$	0.951625	0.676667
$\alpha_{1,GBS}$	0.036109	0.144773
$\beta_{1,GBS}$	0.959335	0.773899
$a_{AB,GBS}$	0.0000	0.038845
$b_{AB,GBS}$	0.856166	0.931651
Correlation	-0.000834	-0.00021
HR	-0.0013739	-0.000330

### 3.2.2. Assessing the COVID-19 Pandemic effect

Finally in this section, we investigate the effect of the COVID-19 Pandemic on the Bivariate GARCH Modeling results. In practice we divide the sample period to two sub-periods: from the 2nd of May 2011 till the end of December 2019, and from the 2nd of January 2020 till the 21th of May 2021. The first period is called as the pre-COVID-19 period and the second as the COVID-19 period. Our intention is to assess the COVID-19 Pandemic effect on the volatility clustering, the conditional covariance-correlation and the HR. Thus, we investigate whether one can observe differences in risk analysis when comparing the pre-COVID-19 with the COVID-19 period.

In Table 8 we present the mean estimate results under the MV-Std DCC(1,1)-GARCH model for these two time-periods. Here, we need to point out that for all estimation results we have statistical significance with p-value < 1%. First, we diagnose that the volatility persistence of the COVID-19 period is lower than the corresponding in the pre-COVID-19 period. More specifically, under the AB-GB case, the volatility persistence for the AB index equals 0.995554 for the pre-COVID-19 period and 0.920267 for the COVID-19 period which is translated to the corresponding 155.5567 and 8.341981 days needed to reduce a volatility shock by 50%. For the GB index, we have volatility persistence that equals 0.996037 for the pre-COVID-19 period and 0.988603 for the COVID-19 period. Again, this is translated to 174.5579 and 60.47115 days needed to reduce a volatility shock by 50% for the pre-COVID-19 and COVID-19 periods respectively. Now, under the AB-GBS case, the volatility persistence for the AB



index equals 0.996037 for the pre-COVID-19 period and 0.988603 for the COVID-19 period. This is translated to the 155.5567 and 8.341981 days needed to reduce a volatility shock by 50% for the pre-COVID-19 and COVID-19 periods respectively. For the GBS index, we have volatility persistence that equals 0.995444 for the pre-COVID-19 period and 0.988278 for the COVID-19 period. Again, this is translated to 151.7926 and 58.78491 days needed to reduce a volatility shock by 50% for the pre-COVID-19 and COVID-19 periods respectively.

Concerning the volatility spillover, expressed by the  $a$  and  $b$  parameters, we diagnose that this is higher for the COVID-19 period when compared with the pre-COVID-19 period. This is observed for both AB-GB and AB-GBS cases.. Finally, as regards the level of conditional correlation and that of HR, we see that for both periods we have mean values of the same sign with no significant differences. As it is more obvious in Figures 8 and 9, where we present the time-varying correlation and HR functions for the AB-GB and AB-GBS cases, the COVID-19 period is characterized by a period of negative and a period of positive correlation. The latter signifies non-constant interaction of volatility between the GBs and the AB within the COVID-19 period. A recent work in Huynh et al. (2021) has analyzed the effect of the COVID-19 period in the equity market using the predictive power of a newly constructed “feverish” index on stock returns and volatility.

**Table 9.** MAFE and MSFE results for alternative Bivariate DCC(1, 1)-GARCH models on AB-GB and AB-GBS return series.

		MAFE	MSFE
AB-GB	MV-Normal	0.1196	0.1210
	MV-Student-t	0.1190	0.1198
	MV-Laplace	0.1233	0.1237
AB-GBS	MV-Normal	0.1790	0.2264
	MV-Student-t	0.1765	0.2219
	MV-Laplace	0.1944	0.2364

Note: The boxed numbers indicate the favored estimation case.

#### 4. Forecasting

In this section, we compare the forecasting performance of the three DCC-GARCH models. Based on a within-sample estimation where parameters are estimated separately for each competing model, we formulate out-of-sample forecasts for the conditional variance and covariance functions. The out-of-sample period is derived by extracting the last  $S = 1000$  observations from a period of 2619 observations. Thus, the forecasting period is from the 23th of May 2017 to the 21th of May 2021. The purpose of this exercise is to investigate the effect of the use of alternative distributional assumptions in forecasting the covariance and VaR estimates. Also, by estimating the VaR its corresponding loss function, one can have a view of the tail of the distribution as regards both conventional and GBs.

**Table 10.** Models'  $\hat{\alpha}/\alpha$  evaluation applied to financial data series using alternative specifications.

		Normal				$\alpha = 1\%$ Student-t				Laplace			
		$\hat{\alpha}/\alpha$	Loss	$LR_{icc}$	$LR_{cc}$	$\hat{\alpha}/\alpha$	Loss	$LR_{icc}$	$LR_{cc}$	$\hat{\alpha}/\alpha$	Loss	$LR_{icc}$	$LR_{cc}$
AB-GB	AB	3.900	-0.08628	4.901e+01 (2.541e-12)	3.170 (0.0750)	2.8000	-0.09811	2.1892 (0.1390)	0.4573 (0.4989)	3.100	-0.08587	2.860e+01 (8.919e-08)	1.986 (0.1588)
	GB	1.80000	-0.1772	5.22514 (0.02226)	0.66059 (0.41635)	1.80000	-0.1772	5.22514 (0.02226)	0.66059 (0.41635)	1.80000	-0.1772	5.22514 (0.02226)	0.66059 (0.41635)
AB-GBS	AB	2.600	-0.08591	1.795e+01 (2.272e-05)	1.390 (0.2385)	2.100	-0.08975	1.795e+01 (2.272e-05)	1.390 (0.2385)	2.600	-0.1953	9.284046 (0.002312)	0.901910 (0.342270)
	GBS	2.000	-0.1983	7.827239 (0.005146)	2.93348 (0.365995)	1.70000	-0.2002	0.9019 (0.3422)	0.58862 (0.44295)	1.7000	-0.17997	4.09097 (0.04311)	0.58862 (0.44295)
		Normal				$\alpha = 5\%$ Student-t				Laplace			
		$\hat{\alpha}/\alpha$	Loss	$LR_{icc}$	$LR_{cc}$	$\hat{\alpha}/\alpha$	Loss	$LR_{icc}$	$LR_{cc}$	$\hat{\alpha}/\alpha$	Loss	$LR_{icc}$	$LR_{cc}$
AB-GB	AB	1.760	-0.2732	3.09097 (0.06311)	0.58862 (0.44295)	1.1400	-0.2733	3.392 (0.6551)	0.6194 (0.5050)	1.740	-0.2785	1.2843 (0.3926)	0.4396 (0.5073)
	GB	1.04000	-0.5814	0.817217 (0.77305)	0.817217 (0.21431)	1.0400	-0.5814	0.8317 (0.77305)	1.54204 (0.21431)	1.04000	-0.5814	0.08317 (0.77305)	1.54204 (0.21431)
AB-GBS	AB	1.600	-0.2917	3.076 (0.07943)	0.4396 (0.5073)	1.1200	-0.2868	3.07655 (0.07943)	2.20183 (0.47047)	1.600	-0.2792	0.7308 (0.3926)	0.5253 (0.4686)
	GBS	1.1600	-0.6473	0.328658 (0.566450)	0.002497 (0.960150)	1.0800	-0.6348	1.2843 (0.2571)	0.0467 (0.8289)	1.1200	-0.6374	0.73078 (0.3926)	0.00701 (0.9332)

Note: Entries in each table represent the mean  $\hat{\alpha}/\alpha$  ratio, the loss function value and the p-values of the unconditional ( $LR_{icc}$ ) and conditional ratio tests ( $LR_{cc}$ ) for the Bivariate DCC(1, 1)-GARCH Normal, Student-t and Laplace models. The p-values in red indicate violation of the null hypothesis of the test. Boxed numbers indicate the favored model under the  $\hat{\alpha}/\alpha$  ratio criterion and the underlined number the corresponding criterion under the minimum loss function.

#### 4.1. Conditional covariance

For the conditional covariance evaluation, we use the index return products as a proxy. Here, we perform a one-step ahead conditional covariance function estimation. As evaluation measures, we use the Mean Absolute Forecast Error (MAFE) and the Mean Squared Forecast Error (MSFE). Compared to the MSFE, the MAFE measure puts less weight on possible outliers of our forecasts. The forecasting performance results are presented in Table 9. As the results indicate the DCC-GARCH under the multivariate Student-t errors outperforms the other two in the AB-GBS conditional covariance estimation, whereas, in the AB-GB conditional covariance estimation, the DCC-GARCH under the multivariate Normal errors is the best performer under MAFE measure and the corresponding under the Student-t errors under the MSFE measure.

#### 4.2. Conditional value-at-risk

For the same out-of-sample period, we perform one-step-ahead volatility forecasts for the three competing models using the two bivariate settings. These forecasts are then used to formulate the one-step-ahead  $\alpha$ -level conditional VaR estimate for each univariate series, such as:

$$VaR_{t+1,\alpha} = \mu_{t+1} + D^{-1}(\alpha)\hat{\sigma}_{t+1} \quad (19)$$

where  $D^{-1}$  is the inverse CDF for the distribution  $D$  at  $\alpha$  nominal level. In the empirical analysis below, we utilize three different assumptions for the innovation distribution: the normal; the Student-t with  $\nu$  degrees of freedom and the Laplace density. In the case of the Student-t model, the quantile measure is formulated as  $\mu_{t+1} + D^{-1}(\alpha)\hat{\sigma}_{t+1} \sqrt{\frac{\nu-2}{\nu}}$  for the  $\nu$  degrees of freedom.

We will evaluate the conditional VaR estimates using three measures: the Violation Rate (hereafter VRate), the quantile loss function and the two-hypothesis testing methods based on the unconditional and conditional tests (see Kupiec, 1995; Christoffersen, 1998). First, the VRate is a common non-test criterion to compare alternative VaR estimates defined as the proportion of days for which the actual return is more extreme than the forecasted VaR level, over the forecast period. Thus, one should investigate how these specifications can estimate VaR at a pre-specified nominal level given the alternative models. Here, we calculate the VRate which takes the following form for an  $S$  forecasting period  $S^{-1} \sum_{t=1}^S I(y_t < VaR_{t,\alpha})$ . This non-statistical validation test estimates the empirical nominal level  $\hat{\alpha}$ . This becomes the proportion which represents the fraction of days for which the actual return yields bigger losses than the estimated VaR over a given forecast evaluation period of  $S$ . When the estimated  $\hat{\alpha}$  equals to the given nominal level  $\alpha$  this is highly desirable. However, this  $\hat{\alpha} = \alpha$  condition is only rarely met. In practice, when  $\hat{\alpha} < \alpha$ , the losses are underestimated (conservative), while on the other hand, when  $\hat{\alpha} > \alpha$  the losses are overestimated and this results in capital allocation with potential losses. In practice, underestimation is more desirable than overestimation unless  $\hat{\alpha}/\alpha$  equals one.

Then, we calculate a quantile criterion function based on the log-transformed loss function

$$\log \mathbf{E}[L_{\alpha}(e_t)] = \log \sum_t (\alpha - I_{\{e_t < 0\}}) \cdot e_t \quad (20)$$

where  $e_t = y_t - VaR_{t,\alpha}$ . The above criterion function should be minimized for the forecasting period adopted.

Finally, we use the unconditional coverage test by Kupiec (1995) and the conditional coverage test by Christoffersen (1998). The first, is a likelihood ratio test of the hypothesis that the nominal level equals the estimated VRate when conditional dependence is assumed. According to this, we test whether

the VRate  $\hat{\alpha}$  obtained by the candidate model is significantly different from the proposed nominal level  $\alpha$ . According to this, the likelihood function of the i.i.d. Bernoulli sequences is defined as:

$$L(\hat{\alpha}) = \prod_{t=1}^S (1 - \hat{\alpha})^{1-I_{t+1}} \hat{\alpha}^{I_{t+1}} = (1 - \hat{\alpha})^{S_0} \hat{\alpha}^{S_1} \quad (21)$$

where  $S_0$  and  $S_1$  are the number of 0s and 1s in the sample. Therefore, the VRate is expressed as  $\hat{\alpha} = S_1/S$ . Under the  $H_0$  hypothesis  $\hat{\alpha} = \alpha$  and its likelihood function is defined as:  $L(\alpha) = (1 - \alpha)^{S_0} \alpha^{S_1}$ . Then, we derive the likelihood ratio unconditional coverage test, such as:

$$LR_{uc} = -2 \ln[L(\alpha)/L(\hat{\alpha})] \sim \chi_1^2 \quad (22)$$

distributed as a  $\chi^2$  distribution with one degree of freedom. Under the second test: the conditional coverage test derived by Christoffersen (1998) this is based on a first-order Markov sequence with a transition probability matrix such as:

$$\hat{A} = \begin{bmatrix} 1 - \hat{\alpha}_{01} & \hat{\alpha}_{01} \\ 1 - \hat{\alpha}_{11} & \hat{\alpha}_{11} \end{bmatrix} \quad (23)$$

Under this matrix, the  $\hat{\alpha}_{01}$  corresponds to the probability given 0 today to have a 1 result the next day; i.e.  $Pr(I_t + 1 = 1, I_t = 0)$ , whereas the  $\hat{\alpha}_{11}$  corresponds to the probability given 1 today to have a 1 result the next day; i.e.  $Pr(I_t + 1 = 1, I_t = 1)$ . Thus, the probability of a 0 event following a 0 event is  $1 - \hat{\alpha}_{01}$ , and the probability of a 0 event following a 1 event is  $1 - \hat{\alpha}_{11}$ . Then the likelihood function for the  $S$  forecast period is defined as:

$$L(\hat{A}) = (1 - \hat{\alpha}_{01})^{S_{00}} \hat{\alpha}_{01}^{S_{01}} (1 - \hat{\alpha}_{11})^{S_{10}} \hat{\alpha}_{11}^{S_{11}}$$

Now, under the null hypothesis of the test, the probability of a violation tomorrow does not depend on today being a violation or not having  $\alpha_{01} = \alpha_{11} = \alpha$ . Now, the likelihood ratio conditional coverage test becomes as follows:

$$LR_{cc} = -2 \ln[L(A)/L(\hat{A})] \sim \chi_1^2 \quad (24)$$

In practice, in order for the VRate to be validated, we intend for the forecasting data to show acceptance of the null hypothesis with p-values at least equal to 5%.

Table 10 reports the one-step-ahead  $\hat{\alpha}/\alpha$  ratio for all data series under the bivariate DCC-GARCH models under MV-Normal, MV-Std and MV-Laplace specifications at the 1% and 5% significance level. Also, for the same models, we report a quantile criterion function based on the log-loss function  $\log \mathbf{E}[L_\alpha(e_t)]$  appearing in (4). Finally, in the same table we report two formal statistical tests: the unconditional coverage test (denoted as  $LR_{uc}$ ) and the conditional coverage test appearing in (5) and (6) respectively. Results at  $\alpha = 1\%$  show that under the  $\hat{\alpha}/\alpha$  ratio criterion, the MV-Std model specification ranked first in three out of four data cases, whereas the MV-Normal and MV-Laplace specification show similar VRate results in one case. Concerning the unconditional and the conditional coverage tests the first seems very sensitive with nine rejections out of twenty four test cases. The conditional coverage test is never rejected. At  $\alpha = 5\%$  and under the  $\hat{\alpha}/\alpha$  ratio criterion, we have the MV-Normal specification coming first in two out of four cases. Concerning the unconditional and the conditional coverage tests they both have no rejections. Under the loss function criterion, we observe that GBs systematically have larger losses than the convectional ones in all model specifications

and under both nominal levels. Now, concerning the distribution specification, at  $\alpha = 1\%$  nominal level the MV-Std specification outperforms the others in all cases whereas for  $\alpha = 5\%$  the MV-Normal one dominates the rest as in the case of the  $\hat{\alpha}/\alpha$  ratio criterion.

The conditional VaR forecasting results demonstrate a number of general results:

1. Among the DCC(1, 1)-GARCH models with different distribution assumptions, the one selected under the VRate criterion depends on the nominal level assumed.
2. At  $\alpha = 1\%$  nominal level the MV-Std model seems to outperform the MV-Normal and MV-Laplace whereas at  $\alpha = 5\%$  the MV-Normal model outperforms the MV-Std and the MV-Laplace models.
3. Model selection results under the loss function measure seem to coincide with those of the VRate one.
4. Under the loss function measure, GBs convey higher risk than CBs.

## 5. Conclusions

The need for financing the green economy has urged the financial market to issue green bond products. These have been in the interest of investors who mostly want to promote investments that are environmentally friendly. However, growing interest in these products has raised questions regarding whether GBs differ to CBs in terms of their risk and whether there are any systematic patterns between these two different products. To answer these questions we have collected daily closing prices of green and conventional bond indexes. For the purposes of analyzing their risk, we took the daily returns of these series. The preliminary and conditional variance and covariance statistical analysis implemented using univariate and bivariate GARCH models has revealed the following results:

1. Risk expressed by conditional volatility as well as VaR measures is higher for green bond than for conventional bond returns.
2. Volatility clustering exists within conventional bond and green bond returns with the GBs demonstrating a marginally higher level of clustering than the CBs.
3. There are cases where we observe volatility spillover between the green bond and the conventional bond returns.
4. There are significant differences in conditional correlations when comparing the conventional bonds with green bonds of either loose or stringent green-labeled criteria.
5. The adoption of loose or stringent eligibility green-labeled criteria can significantly affect hedging strategies.
6. The COVID-19 Pandemic period is associated with lower volatility persistence for both green and conventional bonds when compared to the pre-COVID-19 period. Also, we observe a higher spillover effect between conventional and GBs.

These results have several consequences for investors and decision makers. These are as follows: First, the estimation of time-varying volatility as well as the VaR measure can help investors to construct an optimal portfolio where assets of high risk like the GBs can be mixed with assets of lower risk such as the CBs. Second, the different signs of correlations and HRs observed in the AB-GB and AB-GBS cases indicate that, the way these indexes are generated has a significant effect in their real diversification.

Finally, the present approach can be easily applied in cases of other sub-indexes of green bonds, indexes of specific issuers, sectors or geographical areas, stocks, gold, crude oil, cryptocurrencies etc.

## Conflict of interest

All authors declare no conflicts of interest in this paper.

## References

- Agliardi E, Agliardi R (2021) Pricing Climate-related Risks in the Bond Market. *Jour of Fin Stab* 54: 100868.
- Bollerslev T (1986) Generalized autoregressive conditional heteroscedasticity. *J Economet* 31: 307–327.
- Broadstock DC, Cheng LTW (2019) Time-varying relation between black and green bond price benchmarks: macroeconomic determinants for the first decade. *Financ Res Let* 29: 17–22.
- Chen F, Sutcliffe C (2012) Better cross hedges with composite hedging? Hedging equity portfolios using financial and commodity futures. *Eur J Financ* 18: 575–595.
- Christoffersen P (1998) Evaluating interval forecasts. *Int Econ Rev* 39: 841–862.
- Engle RF (2002) Dynamic conditional correlation: a simple class of multivariate GARCH models. *J Bus Econ Stat* 20: 339–350.
- Flammer C (2018) Corporate Green Bonds. *Working Paper. SSRN*.
- Huynh TLD, Hillec E, Nasir MA (2020) Diversification in the age of the 4th industrial revolution: The role of artificial intelligence, green bonds and cryptocurrencies. *Technol Forecast Soc Change* 159: 120188.
- Huynh TLD (2020) When ‘green’ challenges ‘prime’: empirical evidence from government bond markets. *J Sust Financ Investments*, 1–14.
- Huynh TLD, Foglia M, Nasir MA, et al. (2021) Feverish sentiment and global equity market during COVID-19 Pandemic. *J Econ Behav Organ* [In Press].
- Gianfrate G, Peri M (2019) The green advantage: exploring the convenience of issuing green bonds. *J Clean Prod* 219: 127–135.
- ICMA (2017) The Green Bond Principles 2017: Voluntary Process Guidelines for Issuing Green Bonds. Annual Report. Switzerland. Available from: <https://www.icmagroup.org/assets/documents/Regulatory/\Green-Bonds/GreenBondsBrochure-JUNE2017.jpg>.
- Jin J, Hanb L, Wu L (2020) The hedging effect of green bonds on carbon market risk. *Int Rev Financ Anal* 71: 1057–1072.
- Kanamura T (2020) Are green bonds environmentally friendly and good performing assets? *Energy Econ* 88: 47–67.
- Karpf A, Mandel A (2018) The changing value of the ‘green’ label on the US municipal bond market. *Nat Climate Change* 8: 61–165.

- Kroner KF, Sultan J (1993) Time-Varying Distributions and Dynamic Hedging with Foreign Currency Futures. *J Financ Quant Anal* 28: 535–551.
- Kupiec PH (1995) Techniques for verifying the accuracy of risk management model. *J Deriv* 3: 73–84.
- Lee HT (2010) Regime switching correlation hedging. *J Bank Financ* 34: 2728–2741.
- Lee HT, Yorder JK (2007) A bivariate markov regime switching GARCH approach to estimate time varying minimum variance hedge ratios. *Appl Econ* 39: 1253–1265.
- Nanayakkara M, Colombage S (2019) Do investors in Green Bond market pay a premium? *Appl Econ* 40: 4425–4437.
- Nguyen SP, Huynh TLD (2019) Portfolio optimization from a Copulas–GJR–GARCH–EVT–CVAR model: Empirical evidence from ASEAN stock indexes. *Quant Financ Econ* 3: 562–585.
- OECD (2017) Mobilising Bond Markets for a Low-Carbon Transition, Green Finance and Investment. Paris: EOCED Publishing. Available from: <http://doi.org/10.1787/9789264272323-en>.
- Partridge C, Medda FR (2020) The evolution of pricing performance of green municipal bonds. *J Sust Financ Investment* 10: 44–64.
- Pham L (2016) Is it risky to go green? A volatility analysis of the green bond market. *J Sust Financ Investment* 6: 263–291.
- Pham L, Huynh TLD (2020) How does investor attention influence the green bond market? *Financ Res Lett* 35: 101533J.
- Reboredo JC (2018) Green bond and financial markets: Co-movement, diversification and price spillover effects. *Energy Econ* 74: 38–50.
- S&P Green Bond Indices Methodology. S&P Dow Jones Indices: Indices Methodology, February 2021.
- Tsay RS (2005) *Analysis of financial time series*, NJ: Wiley 2nd edition.
- Tse YK, Tsui AKC (2002) A multivariate GARCH model with time-varying correlations. *J Bus Econ Stat* 20: 351–362.
- Tsiotas G (2018) A Bayesian encompassing test using combined Value-at-Risk estimates. *Quant Financ* 18: 395–417.
- Tsiotas G (2020) An ABC approach for CAViaR models with an asymmetric decision rule. *J Stat Comput Simul* 90: 1373–1398.
- Zerbib OD (2017) *The Green Bond Premium*, University of Chicago Press.
- Zerbib OD (2019) The Effect of Pro-Environmental Preferences on Bond Prices: Evidence from Green Bonds. *J Bank Financ* 98: 39–60.