CAViaR models for Value at Risk and Expected Shortfall with long range dependency features

Gelly Mitrodima[a], Jaideep Oberoi[b,*]

[a]*Department of Statistics, London School of Economics, London WC2A 2AE, United Kingdom. Tel: +44 (0) 20 7955 6009. Email: e.mitrodima@lse.ac.uk*
[b]*School of Finance and Management, SOAS University of London, London WC1H 0XG, United Kingdom. Tel: +44 (0)20 7898 4930. Email: jaideep.oberoi@soas.ac.uk*

Abstract

We consider alternative specifications of conditional autoregressive quantile models to estimate Value-at-Risk (VaR) and Expected Shortfall (ES). The proposed specifications include a slow moving component in the quantile process, along with aggregate returns from heterogeneous horizons as regressors. Using data for ten stock indices, we evaluate the performance of the models and find that the proposed features are useful in capturing tail dynamics better.

Keywords: Value-at-Risk; Expected Shortfall; CAViaR-type models; Component models; Long range dependence

JEL codes: G11, C58

*Please send correspondence to Jaideep Oberoi (jaideep.oberoi@soas.ac.uk).

1. Introduction

VaR is now a ubiquitous indicator of various forms of financial risk, used in a range of settings where summary quantification and dynamic assessment of tail risk are required. VaR is defined as a specified quantile of the loss distribution over a specified period, or a loss that will not be exceeded with a predefined level of confidence. Market risk estimation increasingly emphasises ES, which is the expected value of the loss conditional on VaR being exceeded. A large literature has been developed to find better ways of estimating VaR, within which the class of semi-parametric models known as Conditional Autoregressive Value at Risk (CAViaR) (Engle and Manganelli, 2004), is based on directly modelling the conditional quantiles of the distribution of returns. In this paper, we introduce and evaluate versions of CAViaR models with two features designed to capture long range dependence in a parsimonious manner - a slow moving component, and aggregate returns at longer horizons as regressors.

There are several approaches to estimating VaR and the forecasting literature contains various models that work reasonably well for different subsets of financial time series. Several studies have pointed to the usefulness of CAViaR models for estimation of tail risk, particularly due to the stylised features of financial asset returns (see, *e.g.*, Kuester et al., 2006; Chen et al., 2012; Taylor, 2008; Jeon and Taylor, 2013; Taylor, 2019; Patton et al., 2019, among others). CAViaR models are particularly interesting because they have also been used to produce ES estimates based on joint elicitability arguments (Fissler and Ziegel, 2016, among others). However, the potential of CAViaR models in their various specifications has not been fully explored. This is especially true for CAViaR models with long range dependence, even though the volatility literature has for a long time adopted several analogous models based on the work of Ding and Granger (1996), Engle and Lee (1999), and Corsi (2009), among others. It is therefore worthwhile to examine the potential of similar approaches to introducing long range dependence for improving VaR estimation in CAViaR models.

Our study is also motivated by practical financial considerations in relation to transaction costs and costly recapitalisation. In typical situations where investment sizes are restricted by risk-based limits, frequent changes in the VaR could lead to unnecessary transaction costs or to induce excessive conservatism leading to misallocation of capital. Similarly, it is widely recognised that clustering and asymmetry in return volatility imply that VaR increases when the value of investments falls. This in turn implies that a trader would be forced to reduce their position at an unprofitable juncture, effectively selling low and buying high, in order to bring their position in line with risk controls (see *e.g.* Adrian and Shin, 2014). We hypothesise that gradual VaR adjustments, which can capture low-frequency movements, would help make adjustments smoother and potentially reduce transaction costs without raising the average level of capital required. If the VaR series includes small gradual adjustments that come closer to anticipating the severity of future shocks, this could reduce the need for costly recapitalisation by reducing the portion of a realised loss when that loss exceeds the VaR.

To generate long range dependence we propose a two component model for the quantile in the spirit of Engle and Lee (1999) and Ding and Granger (1996). This allows the mean of the quantile to be time varying, allowing for one slower moving and one faster moving component in the quantile process.

Our second model specification is to include information about past aggregate returns at lower frequencies. There is limited evidence of predictability in daily returns for stocks and indices.[1] However, there are theoretical results (see *e.g.* Levy et al., 1994) and empirical reports (see *e.g.* Shiller, 2015) that suggest that long periods of consistently high gains are followed by larger negative shocks. If this is the case, past aggregate returns over multiple horizons could help obtain improved predictions of the severity of large losses. [2]

We extend the model specification in each case to simultaneously estimate ES. Specifically, we incorporate an additional parameter in the style of Taylor (2019), which implies a multiplicative relationship between the VaR and the corresponding ES. This allows us to produce joint estimates of VaR and ES in a tractable manner.

In order to evaluate the performance of the models, we carry out an empirical exercise on 10 stock index series, with the out of sample period extending over six years and including the Covid-19 pandemic. Particularly since the global financial crisis, there has been an accentuated move towards using ES in bank risk models as international bank regulation standards (*Basel III*) move beyond VaR. Separately estimating ES in a robust manner without moment assumptions is challenging (Gneiting, 2011, showed that ES is not an elicitable risk measure). However, Fissler and Ziegel (2016) show that VaR and ES are jointly elicitable, and related or similar backtesting procedures are proposed in Acerbi and Szekely (2014), Fissler et al. (2016), Nolde and Ziegel (2017), and Ziegel et al. (2020), among others. Patton et al. (2019) and Taylor (2019) develop models and estimation

[1]Kuester et al. (2006) extend a CAViaR model by incorporating autocorrelation in the return process and they do find an improvement in model performance. We include their model as a benchmark in this paper.

[2]Previous authors have also extended CAViaR models by including regressors other than the lags of VaR (Jeon and Taylor, 2013; Rubia and Sanchis-Marco, 2013; Schaumburg, 2012), although they do not include multi-horizon returns. Meligkotsidou et al. (2019) use combinations of regressors to produce quantile forecasts and highlight the importance of the information in the past values of the return time series, over other data.

procedures for jointly estimating VaR and ES. The former derive a scoring function based on Fissler et al. (2016) that is homogeneous of degree zero, while the latter shows that this can be achieved by maximising the log likelihood of the Asymmetric Laplace distribution. In this paper, we estimate and assess the models using the loss function in Patton et al. (2019), which we will refer to as FZ0, adopting the label used by its authors.

While estimation strategies generally rely on search-based algorithms, the optimisation surfaces for quantile estimation are notoriously problematic, requiring many starting values. We find that estimation works well when we select a relatively small set of starting values based on a rule that makes the starting values internally consistent (in the sense that the resulting time series has a stationary mean at the empirical level of the targeted quantile in sample). We argue in this paper that this approach is of practical value.

In order to assess and compare the models we report various criteria. In keeping with the literature, we report the out of sample objective function estimates, along with other tests of unconditional and conditional coverage. In addition, for comparison between models, we report the results of Diebold-Mariano tests (Diebold and Mariano, 1995) following the recent literature. We find that our proposed VaR models offer some gains over benchmarks. Even though there is no single model that performs best across all data series, we find that in each case one of the models proposed by us tends to outperform the benchmark CAViaR models. Our findings suggest that introducing long range dependence in CAViaR models in a parsimonious manner is a valuable improvement.

The rest of the paper is organised as follows. In Section 2, we introduce the model specifications and in Section 3 we describe the estimation method. In Section 4 we present and discuss the results of an out of sample comparison of the models. We then conclude in Section 5.

2. Models

Let $\{r_t\}_{t=1}^T$ denote the univariate time series of portfolio returns. Then the VaR is given by $VaR_t = -F_t^{-1}(\theta)$, where $\theta \in (0,1)$ is the specified quantile level, and $F_t(\cdot)$ represents the distribution of r_t conditional on time $t-1$ information. ES is given by $ES_t = -E_{t-1}[r_t | r_t \leq VaR_t]$.

A generic CAViaR model can be specified as follows for the time series of returns $\{r_t\}_{t=1}^T$:

$$q_{\theta,t}(\boldsymbol{\beta_\theta}) = \beta_{\theta,0} + \sum_{i=1}^{M} \beta_{\theta,i} q_{\theta,t-i}(\boldsymbol{\beta_\theta}) + \sum_{j=1}^{N} \beta_{\theta,M+j} l^j(\mathbf{x_{t-1}}) \tag{1}$$

where $q_{\theta,t}$, the θth quantile of r_t, is a function parameterised by $\boldsymbol{\beta_\theta}$, a vector with $M+N+1$ parameters. $\mathbf{x_t}$ contains the data observable at time t (including the history of returns $\{r_s\}_{s=0}^t$), and $l^j(\cdot)$ are functions of lags of the data. The VaR is usually reported as a positive number, so $-q_{\theta,t}$, because the θ of interest for VaR is in the extreme left tail of the return distribution.[3] However, in this paper, we refer to the quantile $q_{\theta,t}$ as the VaR so that readers can view the quantile model in its usual sense. It is clear that the parameter vector $\boldsymbol{\beta_\theta}$ is specific to the quantile being modelled, and that the series of quantile estimates depend on the parameters. For the rest of this paper, we adopt a simpler notation by dropping reference to θ and the dependence referred to above.

The usual practice is to set $M = 1$ as models of this order are found to work well

[3] Patton et al. (2019) rely on this fact to derive their FZ0 objective function.

for financial time series like the ones considered here. With this assumption and in the simplified notation, a joint model for both VaR and ES with respect to a given quantile θ can be written, following Taylor (2019), as

$$
\begin{aligned}
q_t &= \beta_0 + \beta_1 q_{t-1} + \beta_2 l(\mathbf{x}_{t-1}) \\
e_t &= (1 + \exp(\gamma)) q_t
\end{aligned} \tag{2}
$$

where e_t is the ES and γ a parameter that helps describe, in a simple but intuitive manner, the relationship between VaR and ES in the conditional return distribution. This approach to modelling ES is appealing for several reasons. Firstly, it ensures that there are no crossings, i.e. the ES is always further out in the tail than the VaR. Secondly, the notion of a common scale parameter is generally supported by existing research (Engle and Gonzalez-Rivera, 1991; Xiao and Koenker, 2009; Gourieroux and Liu, 2012; Taylor, 2019, among others). Finally, as we are interested in the tail of the return distribution, such as the 1% VaR and ES, the number of data points that are informative about the relationship between VaR and ES is quite small - by definition, we should expect 1 hit in 100 observations. In such a situation, it would be pragmatic to limit the number of parameters that depend crucially on this subset of observations for estimation.

Engle and Manganelli (2004) propose four specifications for CAViaR with well known counterparts in the volatility literature. Of these, three have been widely adopted and adapted in subsequent research (note they also set $\mu_t = 0$):

Symmetric Absolute Value (SAV):

$$
q_t = \beta_0 + \beta_1 q_{t-1} + \beta_2 |r_{t-1}|, \tag{3}
$$

Asymmetric Slope (AS):

$$q_t = \beta_0 + \beta_1 q_{t-1} + \beta_2 r_{t-1}^{+} + \beta_3 r_{t-1}^{-}, \tag{4}$$

where the notation employed is: $r^{+} = \max(r, 0)$ and $r^{-} = -\min(r, 0)$.

Indirect GARCH (IG):

$$q_t = -\left(\beta_0 + \beta_1 q_{t-1}^2 + \beta_2 r_{t-1}^2\right)^{\frac{1}{2}}. \tag{5}$$

These models can each be extended to include ES following Equation 2. Taylor (2019) follows this approach, using the SAV and AS specifications. We will adopt the same method to include ES estimation in all models.

As we have argued, we can also extend these models by allowing for a time varying, slow moving component in the spirit of Ding and Granger (1996) and Engle and Lee (1999). The resulting models are detailed in the next subsection.

2.1. Component quantile models with multi-horizon returns

We propose models with two components that are motivated by the potential to capture long range dependence and to achieve a smoother series that could save transaction costs. We focus on the approach of Engle and Lee (1999), which has also been applied to option pricing by Christoffersen et al. (2008) and Christoffersen et al. (2010).

The component versions of CAViaR models that we propose below are written by reference to the SAV, AS and IG models, by replacing the intercept parameter β_0 with a time varying process that induces a long memory property to the VaR. Consider for example the AS recursion in Equation 4, which can first be rewritten as:

$$q_t = u + \beta_1(q_{t-1} - u) + \beta_2 r^+_{t-1} + \beta_3 r^-_{t-1} \tag{6}$$

By allowing u to vary over time, we obtain the Component-AS model (C-ASd):

$$\begin{aligned} q_t &= u_t + \beta_1 \left(q_{t-1} - u_{t-1}\right) + \beta_2 r^+_{t-1} + \beta_3 r^-_{t-1} \\ u_t &= \beta_4 + \beta_5 u_{t-1} + \beta_6 r_{t-1} \end{aligned} \tag{7}$$

Similarly, we can define the other two component counterparts.

Component-SAV (C-SAVd):

$$\begin{aligned} q_t &= u_t + \beta_1(q_{t-1} - u_{t-1}) + \beta_2 \left|r_{t-1}\right| \\ u_t &= \beta_3 + \beta_4 u_{t-1} + \beta_5 r_{t-1}, \end{aligned} \tag{8}$$

Component-IG (C-IGd):

$$\begin{aligned} q_t &= -\left\{u_t^2 + \beta_1 \left(q^2_{t-1} - u^2_{t-1}\right) + \beta_2 r^2_{t-1}\right\}^{\frac{1}{2}} \\ u_t &= \beta_3 + \beta_4 u_{t-1} + \beta_5 r_{t-1} \end{aligned} \tag{9}$$

We also propose the incorporation of multi-horizon returns as regressors in these models. Our reasoning is developed along the following lines. When we observe that the VaR from our model has not been exceeded for a long period of time, should we adjust our model? If so, should the adjustment lead to an increase or decrease in the

VaR? This is a difficult question because it depends on one's view of the return process and on the loss function. This also has significant economic implications. Historically, we have observed that large positive aggregate returns over a long period of time are followed by large negative shocks. This is not just a liberal interpretation of the adage "The higher they climb, the harder they fall." A voluminous literature on bubbles is dedicated to identifying instances where asset prices have deviated persistently from their true value. Unfortunately, it is very difficult to precisely quantify how large is 'large' for returns and how long is the 'long' period before prices revert. Hence, our approach is simply to consider the possibility that past returns at lower than daily frequencies contain important information for risk estimation.

The notion of incorporating information about different horizons has significant support in the statistics and finance literature. Glosten et al. (1993) document that monthly return volatility dynamics are different from daily dynamics. Bianco et al. (2009) find that returns have dependence at different frequencies. Venetis and Peel (2005) find changes in the serial correlation in returns based on the volatility. Among those who consider the impact of different frequencies or horizons in asset pricing are Adrian and Rosenberg (2008) and Duffie et al. (2007). The latter find that future defaults are predicted by one year lagged stock return performance. Similarly, Doshi et al. (2013) use past returns among other variables to find credit default swap prices. The model of Corsi (2009) has been widely adopted as it has been shown to improve volatility forecasting by incorporating the history of realised volatility from different horizons. In particular, Hua and Manzan (2013) use the HAR approach of Corsi (2009) in the context of quantile forecasting using a location-scale model. A simple approach to introduce long memory in our model is through the use of multi-horizon aggregate returns as regressors. Thus, the three versions of the models with regressors are given below.

Component-SAV multi-horizon (C-SAVdwm):

$$
\begin{aligned}
q_t &= u_t + \beta_1(q_{t-1} - u_{t-1}) + \beta_2 |r_{t-1}| \\
u_t &= \beta_3 + \beta_4 u_{t-1} + \beta_5 r_{t-1} + \beta_6 r^w_{t-1} + \beta_7 r^m_{t-1},
\end{aligned} \tag{10}
$$

Component-AS multi-horizon (C-ASdwm):

$$
\begin{aligned}
q_t &= u_t + \beta_1(q_{t-1} - u_{t-1}) + \beta_2 r^+_{t-1} + \beta_3 r^-_{t-1} \\
u_t &= \beta_4 + \beta_5 u_{t-1} + \beta_6 r_{t-1} + \beta_7 r^w_{t-1} + \beta_8 r^m_{t-1}
\end{aligned} \tag{11}
$$

Component IG multi-horizon (C-IGdwm)

$$
\begin{aligned}
q_t &= -\left\{u_t^2 + \beta_1 \left(q^2_{t-1} - u^2_{t-1}\right) + \beta_2 r^2_{t-1}\right\}^{\frac{1}{2}} \\
u_t &= \beta_3 + \beta_4 u_{t-1} + \beta_5 r_{t-1} + \beta_6 r^w_{t-1} + \beta_7 r^m_{t-1},
\end{aligned} \tag{12}
$$

where r^w and r^m represent aggregate returns over the past week (5 days) and month (22 days), respectively.

2.2. *Interpreting the component structure*

Unlike the traditional approach to writing component models, as discussed in detail in Christoffersen et al. (2008), here we have not attempted to write the model as a combination of two processes with zero-mean shocks.

Let us consider the AS model given in Equation (6). Assuming stationarity, if we take the unconditional expectation of the quantile, we get:

$$u = \overline{q} - \frac{\beta_2 \overline{r^+} + \beta_3 \overline{r^-}}{1 - \beta_1}, \tag{13}$$

where $\overline{q}$, $\overline{r^+}$, and $\overline{r^-}$ represent the expectations of the respective variables. Thus, we can see that, unlike the traditional approach to writing such models, here $u \neq \overline{q}$. This can also be seen by substituting u from Equation (13) into Equation (6) to retrieve a more intuitive version of the AS model.

$$q_t = \overline{q} + \beta_1 (q_{t-1} - \overline{q}) + \beta_2 \left(r^+_{t-1} - \overline{r^+} \right) + \beta_3 \left(r^-_{t-1} - \overline{r^-} \right)$$

Looking again at the Component-AS model in Equation 7, assuming stationarity and for simplicity that $\bar{r} = 0$, we can solve again for the unconditional expectation of the quantile $\bar{q}$ and the unconditional mean of the component $\bar{u}$:

$$\overline{q} = \frac{\beta_4}{1 - \beta_5} + \frac{\beta_2 \overline{r^+} + \beta_3 \overline{r^-}}{1 - \beta_1},$$
$$\text{where } \overline{u} = u = \frac{\beta_4}{1 - \beta_5}$$

Thus, in the proposed component models, the deviation $q_t - u_t$ is the component that represents an adjusted distance from the unconditional mean $\overline{q}$, while the dynamics of u_t capture the time dependence in $\overline{q}$, albeit with an adjusted mean level. Our aim is to introduce longer range dependence in the quantile process via a more persistent component. As long as the persistence of the component u_t is higher than that of $q_t - u_t$ we are able to interpret it as a ‘long-term’ component; with the caveat that q_t tends to a quantity different from u_t in the long run.

3. Estimation

We estimate the models by minimising the FZ0 loss function derived by Patton et al. (2019) based on the results of Fissler and Ziegel (2016):

$$L_{FZ0}(r, q, e; \theta) = -\frac{1}{\theta e} I_{r \leq q}(q - r) + \frac{q}{e} + log(-e) - 1,$$

where $I(.)$ is an indicator function. This loss function is also appropriate for out of sample evaluation of the performance of the models.

The optimisation surface for such problems is known to be multi-modal and generally problematic to minimise over. We use a relatively small set of starting values (based on predefined criteria) that take into account the relative scale of the parameters. More specifically, we select a series of ‘compatible’ parameter inputs. We first pick a grid of values for the autoregressive parameters in the model and pair each value with values from a grid of other (news impact) parameters. In each case, we modify the β_0 parameter (or the corresponding constant term in the equation for component models) to be approximately consistent with the assumption that the quantile process is stationary around the in-sample unconditional quantile of the data. Note that the parameter is not fixed - it is merely the starting value for the parameter that is chosen to be internally consistent for each combination of the other parameters.

For example, in the case of the SAV model in Equation (3), we start by choosing a range of values for β_1, viz. $\{0.65, 0.8, 0.95\}$. For each of these values, we only pair it with starting values of $\beta_2 \in -0.2, -0.1$. Using only the in-sample data, we compute $\bar{q}$ as the relevant θth empirical quantile of the data. Similarly, we compute the mean of the absolute value of the data, labelled $a\bar{b}sr$. Assuming stationarity of the quantile time series, we can then imply starting values for β_0 using the formula $(1 - \beta_2)\bar{q} - \beta_3 a\bar{b}sr$, for each pair of values of β_2 and β_3. Thus, with only 6 sets of starting values, we obtain estimates that are as good or better than the original approach of evolutionary starting parameter selection in Engle and Manganelli (2004).[4]

Additionally, when we solve for the unconditional mean of a CAViaR process, we may come across terms that require distributional assumptions, *e.g.*, the expectation of the ratio of the absolute value returns to a quantile. In such a case, it works well to simply substitute in quantities based on the normal distribution as a starting point for the estimation algorithm.

4. Empirical Results

For the empirical exercise, we adopt the three original CAViaR models and two additional benchmark models that have been shown in the literature to perform well. We first specify the additional benchmarks, then describe the data used to present comparisons across the models. Finally, we discuss the out of sample performance of the models.

[4]To ensure our approach works well, we applied it to the stocks and time period studied in the original paper by Engle and Manganelli (2004). We used the code available from the website of Simone Manganelli. http://simonemanganelli.org/Simone/Research_files/CAViaRCodes.ZIP and compared it to our modified approach, with the proviso that we cleaned the data of zero-return days before running the exercise. In this exercise, we only estimated the VaR using the Realised Quantile (tick loss) objective function. We found that we obtain a lower RQ using our approach.

4.1. Additional benchmarks

The first additional benchmark model (Kuester et al., 2006), which we label as AR-IG, is included as it allows for serial correlation in returns and performed best among CAViaR models in a horse race in the same paper.

AR-IG:

$$q_t = \alpha r_{t-1} - \left(\beta_0 + \beta_1 (q_{t-1} - \alpha r_{t-2})^2 + \beta_2 \left(\varepsilon_{t-1}\right)^2\right)^{\frac{1}{2}}, \tag{14}$$

where α is an AR(1) parameter in the return equation, such that μ_t in Equation (1) is replaced by αr_{t-1}. Note that in all other specifications, we retain the norm in the literature of setting $\mu_t = 0$.

The second is a model based on the leverage effect formulation of Glosten et al. (1993), which we label as IG-GJR. This latter model is similar to ones proposed by Gerlach et al. (2011) and Schaumburg (2012) and is used to capture asymmetry.

IG-GJR:

$$q_t = -\left(\beta_0 + \beta_1 q_{t-1}^2 + \beta_2 r_{t-1}^2 + \beta_3 r_{t-1}^2 I_{r_{t-1}<0}\right)^{\frac{1}{2}} \tag{15}$$

Finally we consider three popular parametric GARCH-type models (GARCH(1,1), EGARCH(1,1), and NGARCH(1,1)).

4.2. Data

In order to evaluate the performance of the models, we estimate them on 10 series of stock indices. These are the S&P 500 (US), Small Cap 2000 (US), DAX 30 (Germany), FTSE 100 (UK), CAC 40 (France), Euro Stoxx 50 (Europe), S&P TSX (Canada), Nikkei225 (Japan), Hang Seng (Hong Kong), and ATX (Austria).

We compute daily log returns and our only data cleaning exercise is to remove days when the return is zero as we assume they are holidays. The sample period ranges from January 01, 2010 to December 31, 2021. From the start of the sample period we retain an in-sample estimation period of 1304 days (after cleaning the data of zero-return days). As a result, we reserve the period from May 2015 as the out of sample period, which represents a time of both up and down markets with varying levels of volatility and includes the Covid-19 pandemic (see the Appendix for a plot). We roll the sample forward one day at a time, re-estimating the one day ahead VaR and ES over the out of sample period.

4.3. Performance comparisons

The focus of this analysis is on the out of sample performance of the models, whereby we present the rolling estimates and the relevant scoring statistics calculated over this period. Further, in order to keep the presentation clear, we focus on results with $\theta = 0.01$. Relative to the 5% quantile, joint estimation of VaR and ES is clearly more challenging at the 1% level.

4.3.1. Out of sample scores

We first report performance in terms of the loss function estimation in Table 1. For each data series, we rank the models from 1 to 14 in ascending order of their loss function ($FZ0$) value out of sample period. We also report the average (across assets) of the average out of sample $FZ0$ for each model in the penultimate row. In the bottom row of the table, we report the average rank for each model. While no single model completely dominates in performance, the average ranks suggest that two of the component models, C-SAVd and C-IGd tend to be among the top ranked models for most assets. Overall, the component models have the best ranks, followed by the other CAViaR models. Applying this scoring rule, it is not surprising that the GARCH models do not perform as well, because the loss function being minimised in GARCH models is not the same as the one being used to compare them to the other models. Further, looking individually at the rows, we see that in each case one of the proposed longer memory models has the best score. This suggests that there is value in incorporating a component or heterogeneous horizon aggregate returns, or both, in CAViaR models.

4.3.2. Coverage tests

As is customary, we also report the coverage tests, beginning with the unconditional coverage rates and the results of the unconditional coverage tests in Table 2. As we have a relatively large out of sample period, we use the 1% level to determine significance. Most models are not rejected by unconditional coverage tests, with the only exception being the C-ASd model. We then present the skill scores relative to the historical simulation benchmark in Table 3. The historical simulation benchmark is simply the 1% percentile of the past 1304 returns (window length). To calculate the skill score, we follow Taylor (2019) by first calculating the ratio of the score of the target model to that of the historical simulation benchmark, then subtracting it from 1 and multiplying the result by 100. Thus, a higher score is better.

Table 1: Out of sample scores

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
S&P500	1.058	0.975	1.035	0.979	1.013	0.967	0.967	1.060	0.975	**0.962**	0.973	1.219	1.333	1.240
SmallCap2000	1.252	1.218	1.242	1.221	1.228	1.212	1.216	1.254	1.217	**1.200**	1.220	1.934	1.944	2.103
DAX30	1.238	1.216	1.232	1.219	1.222	1.199	1.210	1.241	**1.185**	1.212	1.214	1.578	1.589	1.700
FTSE100	1.012	0.975	1.012	0.991	0.989	**0.959**	0.972	1.015	0.972	0.984	0.982	1.060	1.071	1.118
CAC40	1.295	1.265	1.282	1.270	1.270	1.254	1.255	1.297	1.259	**1.254**	1.260	1.504	1.556	1.557
EUROSTOXX50	1.333	1.296	1.317	1.297	1.301	1.283	1.285	1.334	1.289	**1.280**	1.289	1.505	1.553	1.563
SPTSX	0.792	0.736	0.798	0.766	0.762	**0.734**	0.736	0.797	0.737	0.757	0.757	0.991	0.969	0.951
NIKKEI225	1.442	1.406	1.457	1.430	1.437	**1.386**	1.404	1.443	1.404	1.430	1.432	1.651	1.719	1.711
HANGSENG	1.262	1.219	1.265	1.230	1.233	1.204	1.216	1.264	**1.183**	1.225	1.244	1.376	1.357	1.535
ATX	1.289	1.232	1.279	1.245	1.239	**1.220**	1.228	1.296	1.233	1.239	1.242	1.721	1.710	1.780
Average	1.197	1.154	1.192	1.165	1.169	1.142	1.149	1.200	1.145	1.154	1.161	1.454	1.480	1.526
Rank	9.7	4.7	9.6	7.1	7.5	1.6	2.7	10.7	3.3	3.5	5.6	12.4	13	13.6

Note: This table presents the average value of the FZ0 loss function from daily one day ahead estimates over the out of sample period. Lower values are the best, so for each asset, the models are ranked from 1 to 14 in ascending order of the loss function value. In each row, we mark the best score(s) in bold. The bottom row of the table shows the average across assets of the ranks of each model. The best average rank (lowest number) is 1.6 for the C-SAVd model.

Table 2: Unconditional Coverage (in %)

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
S&P500	1.4	1.5	1.4	1.5	1.2	1.7	1.8	1.4	1.6	1.7	1.2	1.0	1.0	1.0
SmallCap2000	1.2	1.4	1.1	1.4	1.2	1.2	1.6	1.2	1.3	1.6	1.2	0.6	0.6	0.6
DAX30	1.6	1.6	1.7	1.7	1.6	1.5	2.1	1.6	1.6	1.3	1.5	0.5	0.6	0.4
FTSE100	1.4	1.5	1.4	1.5	1.5	1.6	1.3	1.5	1.5	1.3	1.4	0.7	0.9	0.6
CAC40	1.3	1.5	1.4	1.4	1.4	1.4	1.8	1.3	1.4	1.3	1.3	0.8	1.0	0.8
EUROSTOXX50	1.4	1.5	1.5	1.5	1.8	1.3	1.8	1.4	1.5	1.5	1.4	0.8	1.1	0.7
SPTSX	1.1	1.1	1.2	1.0	1.2	1.1	1.7	1.1	1.3	1.1	1.1	0.8	0.7	0.6
NIKKEI225	1.2	1.2	1.1	1.2	0.9	1.1	1.9	1.1	1.4	1.2	1.1	0.7	0.6	0.5
HANGSENG	1.2	1.3	1.5	1.6	1.3	1.4	1.7	1.2	1.8	1.5	1.5	0.6	0.7	0.5
ATX	1.0	1.4	1.1	1.3	1.1	1.6	2.0	1.1	1.4	1.3	1.2	0.5	0.7	0.5
NS count	10	10	9	9	9	9	4	10	9	10	10	10	10	9

Note: This table presents the unconditional coverage rate of the models in percentage points, for each of the models. A number closest to 1 represents the best unconditional coverage rate. The final row counts the number of times the unconditional coverage test was not significant at the 1% level for each model.

Table 3: Skill scores

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
S&P500	39.04	43.84	40.36	43.61	41.66	44.28	44.28	38.94	43.87	**44.59**	43.97	29.80	23.23	28.58
SmallCap2000	35.73	37.49	36.27	37.34	36.94	37.77	37.59	35.61	37.55	**38.40**	37.37	0.70	0.229	-7.95
DAX30	24.84	26.14	25.21	26.01	25.81	27.23	26.52	24.67	**28.06**	26.38	26.30	4.16	3.50	-3.23
FTSE100	34.65	37.05	34.65	35.99	36.14	**38.08**	37.22	34.46	37.23	36.45	36.59	31.54	30.85	27.76
CAC40	24.01	25.75	24.75	25.46	25.45	26.39	26.34	23.89	26.11	**26.44**	26.04	11.73	8.69	8.64
EUROSTOXX50	21.70	23.89	22.63	23.83	23.56	**24.63**	24.50	21.61	24.26	24.80	24.28	11.57	8.79	8.19
SPTSX	54.70	57.90	54.37	56.18	56.42	**57.98**	57.92	54.42	57.86	56.68	56.70	43.30	44.56	45.59
NIKKEI225	12.07	14.29	11.18	12.81	12.36	**15.51**	14.42	12.02	14.40	12.84	12.67	-0.66	-4.83	-4.33
HANGSENG	16.50	19.35	16.30	18.60	18.40	20.31	19.49	16.33	**21.71**	18.91	17.68	8.91	10.16	-1.59
ATX	30.03	33.12	30.59	32.44	32.75	**33.80**	33.34	29.68	33.06	32.78	32.62	6.62	7.20	3.40

Note: The skill score is reported relative to historical simulation. We take the score of the relevant model, divide it by the score of historical simulation using a window length of 1304 returns, then subtract this ratio from 1 and multiply the result by 100. Thus, a bigger value is better. In each row, we mark the best score(s) in bold.

We next report the results of a conditional coverage test for the out of sample period for all assets and models in Table 4. Following Christoffersen (1998), the idea behind such tests is to assess to what extent the hits or exceedances are predictable.[5] Berkowitz et al. (2011) suggest that the DQ test of Engle and Manganelli (2004) performs relatively well overall, and we follow them and subsequent literature in performing the DQ test. In the table, we report p-values of the tests, where a low p-value suggests rejection of the model. All models perform in line with or better than other implementations in the literature. At the 1% significance level, each model is rejected for between 1 and 5 out of 10 indices, with the exception of the C-ASd model that is rejected for most assets. In this case, the GARCH models perform best.

4.3.3. Forecast comparison

Following Nolde and Ziegel (2017) and Patton et al. (2019), we next consider results of Diebold–Mariano tests. Table 5 presents a summary of the tests by counting the signs of the t-statistics for pairwise comparisons across the 10 assets. The number reported in the table represents the number of assets for which the column model outperformed the row model. With the exception of the C-SAVdwm model, the other component models tend to outperform, and are challenged only by the AS model, suggesting that asymmetry is a key feature to include. In the interest of completeness, we provide the actual test statistics for all pairs across all assets in Appendix B. Here, we can see that while most results in favour of the component models among themselves and against the benchmark CAViaR

[5] Regulators also consider failure rates (hit ratios) of own-model VaR forecasts of financial institutions to assess their reliability. All models appear to be within a reasonable range when compared to the failure rates that attract the attention of regulators in terms of requiring higher capital.

models are statistically significant, this is not necessarily the case for the GARCH models. The GARCH models appear to have a large number of inferior statistics, but these tests are not statistically significant in the majority of cases. Thus, GARCH models tend to hold their performance in forecast comparison tests, although the summary table does not show this.

4.3.4. Excess losses

A further analysis that we conduct is to assess whether the models indeed assist in avoiding costly recapitalisations, as per the motivation. To this end, in Table 6, we examine the losses on exceedance days (when VaR is exceeded by the loss). We calculate the absolute value of the difference between ES and the return on all exceedance days and then report its mean, times 1000. A larger number signifies that greater capital mismatches arise on these dates. A similar exercise, looking only at the losses when they exceed the ES has similar results. With the exception of the DAX index, the best model in each row is one of the component models, suggesting that there may be some benefit in reducing capital shortfalls if we take long range dependence features into account.

Thus, to summarise the empirical analyses, the overall comparisons provide evidence of the potential for improving the financial performance of CAViaR models by incorporating longer memory features. The out of sample scores show clear gains from considering longer term dependence.

Table 4: Out of sample DQ test p-values for the 1% VaR

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
S&P500	0.000	0.000	0.002	0.147	0.055	0.000	0.000	0.000	0.000	0.000	0.063	0.000	0.379	0.000
SmallCap2000	0.595	0.067	0.917	0.283	0.337	0.404	0.075	0.008	0.172	0.154	0.723	0.109	0.000	0.004
DAX30	0.000	0.002	0.001	0.001	0.002	0.008	0.000	0.000	0.063	0.010	0.000	0.028	0.056	0.003
FTSE100	0.000	0.152	0.003	0.207	0.160	0.004	0.251	0.000	0.187	0.004	0.005	0.187	0.943	0.062
CAC40	0.003	0.012	0.002	0.005	0.003	0.012	0.000	0.003	0.009	0.108	0.004	0.347	0.927	0.338
EUROSTOXX50	0.002	0.004	0.002	0.007	0.000	0.011	0.000	0.000	0.000	0.003	0.007	0.352	0.821	0.257
SPTSX	0.122	0.203	0.524	0.475	0.543	0.219	0.000	0.001	0.191	0.419	0.553	0.355	0.011	0.001
NIKKEI225	0.896	0.719	0.953	0.711	0.915	0.924	0.000	0.951	0.031	0.424	0.832	0.912	0.858	0.653
HANGSENG	0.321	0.420	0.000	0.000	0.054	0.401	0.000	0.187	0.004	0.004	0.013	0.032	0.190	0.012
ATX	0.539	0.439	0.631	0.515	0.619	0.127	0.000	0.087	0.358	0.557	0.610	0.030	0.003	0.011
%	50	70	60	60	70	70	20	30	60	60	60	90	80	60

Note: We report p-values for the DQ test, performed for the series of out of sample estimates. A low p-value (≤ 0.01) suggests model rejection. The final row reports the percentage of times a model was not rejected.

Table 5: Summary of Diebold–Mariano tests

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	10	7	10	10	10	10	0	10	10	10	0	0	0
AS	0	-	0	0	0	10	10	0	8	5	4	0	0	0
IG	2	10	-	10	10	10	10	3	10	10	10	0	0	0
AR-IG	0	10	0	-	3	10	10	0	10	10	9	0	0	0
IG-GJR	0	10	0	7	-	10	10	0	10	10	8	0	0	0
C-SAVd	0	0	0	0	0	-	1	0	2	4	0	0	0	0
C-ASd	0	0	0	0	0	9	-	1	3	4	0	0	0	0
C-SAVdwm	9	10	7	2	10	10	9	-	10	10	10	0	0	0
C-ASdwm	0	2	0	0	0	8	7	0	-	4	2	0	0	0
C-IGd	0	5	0	0	0	6	6	0	6	-	2	0	0	0
C-IGdwm	0	5	0	2	2	10	10	0	8	8	-	0	0	0
GARCH	10	10	10	10	10	10	10	10	10	10	10	-	1	9
NGARCH	10	10	10	10	10	10	10	10	10	10	10	9	-	9
EGARCH	10	10	10	10	10	10	10	10	10	10	10	1	1	-

Note: We computed pairwise t-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. This table counts the sign of the t-statistic across the 10 assets studied in this paper. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior. However, many of the test values are not statistically significant. The actual t-statistics, asset by asset, are reported in Appendix Appendix B

Table 6: Mean Absolute Error around ES on exceedance dates

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
S&P500	0.048	0.096	0.095	0.133	0.046	0.157	**0.043**	0.087	0.075	**0.043**	0.089	2.521	3.965	1.911
SmallCap2000	0.194	0.245	0.249	**0.147**	0.189	0.169	**0.147**	0.212	0.204	0.244	0.158	2.791	2.648	1.781
DAX30	**0.096**	0.158	0.148	0.113	0.1	0.11	0.098	0.145	0.127	0.118	0.123	3.662	3.845	2.956
FTSE100	0.089	0.099	0.132	0.089	0.091	**0.087**	0.108	0.112	0.113	0.103	0.136	1.136	1.316	1.094
CAC40	0.073	0.091	0.077	0.087	0.08	**0.046**	0.064	0.082	0.07	0.065	0.084	0.704	0.722	0.786
EUROSTOXX50	0.106	0.048	0.053	0.034	0.144	**.031**	0.051	0.05	0.032	0.065	0.093	0.873	0.973	0.703
SPTSX	0.152	0.159	0.229	0.117	0.137	0.146	**0.11**	0.152	0.139	0.153	0.176	1.35	1.747	1.148
NIKKEI225	0.086	0.064	0.148	**0.048**	0.083	0.186	0.052	0.054	0.07	0.117	0.079	1.23	1.427	1.149
HANGSENG	0.139	0.162	0.238	**0.12**	0.158	0.146	0.266	0.14	0.314	0.132	0.147	1.439	1.987	1.198
ATX	0.109	0.066	0.1	**0.051**	0.098	0.062	0.075	0.093	0.108	0.095	0.151	2.291	2.871	1.552

Note: In this table, we report the mean size of the difference between the realised return and ES on exceedance days (when $r_t < q_t$). The number is multiplied by 1000 for ease of reading. The smaller this number, the lower the need to raise extra capital immediately at the worst times. With bold we indicate the best performing model.

5. Conclusion

We propose and evaluate CAViaR models that have long range dependence features through the introduction of a component structure. The inclusion of past aggregate returns at heterogeneous horizons as regressors in the component process serves to capture the possibility of lower frequency data being informative in predicting extreme risks. In the sense implied by Hanson et al. (2011), our approach is guided by making micro-prudential methods more consistent with macro-prudential goals. Finally, we introduce a simple practically useful approach to selecting starting values for parameters in the estimation algorithm. Our approach is to ensure the starting values are approximately compatible with each other in the sense that they are consistent with stationarity of the quantile process.

The models are motivated by the need to allow VaR to change steadily while taking into account variation in the underlying mean level of VaR. The intuition behind this motivation is that it may assist in controlling transaction costs and in reducing the severity of the extreme losses (relative to VaR) when VaR is exceeded.

By analysing a particularly challenging period that covered the Covid-19 pandemic, we have provided more evidence that CAViaR models in general are good at capturing the underlying quantile process. Although our models introduce more parameters, they perform very well out of sample, pointing to the usefulness of considering long range dependence among modelling choices for CAViaR models.

Acknowledgements

We are grateful to the Editors and an anonymous reviewer for helpful feedback. We are also grateful, for helpful discussions on the current and/or previous versions of the paper, to Peter Christoffersen, Christian Dorion, Hitesh Doshi, John Galbraith, Jim Griffin, Ekaterini Panopoulou, Gareth Peters, James W. Taylor, David Veredas, and seminar participants at the Computational and Financial Econometrics conference.

Funding and conflicts of interest This research did not receive any specific funding. The authors have no conflicts of interest related to this work.

Data availability The data used in this paper consists of publicly available stock index values. The authors downloaded the data provided by the Center for Research in Security Prices, LLC (CRSP) via Wharton Research Data Services (WRDS). Code is available upon request from the authors.

References

Acerbi, C., Szekely, B., 2014. Back-testing expected shortfall. Risk Magazine 27, 76–81.

Adrian, T., Rosenberg, J., 2008. Stock returns and volatility: Pricing the short-run and long-run components of market risk. The Journal of Finance 63, 2997–3030.

Adrian, T., Shin, H.S., 2014. Procyclical leverage and value-at-risk. Review of Financial Studies 27, 373–403.

Berkowitz, J., Christoffersen, P., Pelletier, D., 2011. Evaluating value-at-risk models with desk-level data. Management Science 57, 2213–2227.

Bianco, S., Corsi, F., Reno, R., 2009. Intraday LeBaron effects. Proceedings of the National Academy of Sciences 106, 11439–11443.

Chen, Q., Gerlach, R., Lu, Z., 2012. Bayesian value-at-risk and expected shortfall forecasting via the asymmetric Laplace distribution. Computational Statistics & Data Analysis 56, 3498–3516.

Christoffersen, P., 1998. Evaluating interval forecasts. International Economic Review 39, 841–862.

Christoffersen, P., Dorion, C., Jacobs, K., Wang, Y., 2010. Volatility components, affine restrictions, and nonnormal innovations. Journal of Business & Economic Statistics 28, 483–502.

Christoffersen, P., Jacobs, K., Ornthanalai, C., Wang, Y., 2008. Option valuation with long-run and short-run volatility components. Journal of Financial Economics 90, 272–297.

Corsi, F., 2009. A simple approximate long- memory model of realized volatility. Journal of Financial Econometrics 7, 174–196.

Diebold, F.X., Mariano, R.S., 1995. Comparing predictive accuracy. Journal of Business & Economic Statistics 13.

Ding, Z., Granger, C.W., 1996. Modeling volatility persistence of speculative returns: A new approach. Journal of Econometrics 73, 185 – 215.

Doshi, H., Ericsson, J., Jacobs, K., Turnbull, S.M., 2013. Pricing credit default swaps with observable covariates. Review of Financial Studies , hht015.

Duffie, D., Saita, L., Wang, K., 2007. Multi-period corporate default prediction with stochastic covariates. Journal of Financial Economics 83, 635–665.

Engle, R.F., Gonzalez-Rivera, G., 1991. Semiparametric ARCH models. Journal of Business & Economic Statistics 9, 345–359.

Engle, R.F., Lee, G., 1999. A permanent and transitory component model of stock return volatility, in: Cointegration, Causality and Forecasting: A festschrift in honor of Clive W.J. Granger. Oxford University Press, pp. 475–497.

Engle, R.F., Manganelli, S., 2004. CAViaR: Conditional autoregressive value at risk by regression quantiles. Journal of Business & Economic Statistics 22, 367–381.

Fissler, T., Ziegel, J.F., 2016. Higher order elicitability and Osband's principle. The Annals of Statistics 44, 1680–1707.

Fissler, T., Ziegel, J.F., Gneiting, T., 2016. Expected shortfall is jointly elicitable with value at risk-implications for backtesting. Risk Magazine .

Gerlach, R.H., Chen, C.W., Chan, N.Y., 2011. Bayesian time-varying quantile forecasting for value-at-risk in financial markets. Journal of Business & Economic Statistics 29.

Glosten, L.R., Jagannathan, R., Runkle, D.E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. The Journal of Finance 48, 1779–1801.

Gneiting, T., 2011. Making and evaluating point forecasts. Journal of the American Statistical Association 106, 746–762.

Gourieroux, C., Liu, W., 2012. Converting Tail-VaR to VaR: An Econometric Study. Journal of Financial Econometrics 10, 233–264.

Hanson, S.G., Kashyap, A.K., Stein, J.C., 2011. A macroprudential approach to financial regulation. The Journal of Economic Perspectives 25, 3–28.

Hua, J., Manzan, S., 2013. Forecasting the return distribution using high-frequency volatility measures. Journal of Banking & Finance 37, 4381–4403.

Jeon, J., Taylor, J.W., 2013. Using CAViaR models with implied volatility for value-at-risk estimation. Journal of Forecasting 32, 62–74.

Kuester, K., Mittnik, S., Paolella, M., 2006. Value-at-risk prediction: A comparison of alternative strategies. Journal of Financial Econometrics 4(1), 53–89.

Levy, M., Levy, H., Solomon, S., 1994. A microscopic model of the stock market. Economics Letters 45, 103 – 111.

Meligkotsidou, L., Panopoulou, E., Vrontos, I.D., Vrontos, S.D., 2019. Quantile forecast combinations in realised volatility prediction. Journal of the Operational Research Society 70, 1720–1733.

Nolde, N., Ziegel, J.F., 2017. Elicitability and backtesting: Perspectives for banking regulation. Annals of Applied Statistics 11, 1833–1874.

Patton, A.J., Ziegel, J.F., Chen, R., 2019. Dynamic semiparametric models for expected shortfall (and value-at-risk). Journal of Econometrics 211, 388–413.

Rubia, A., Sanchis-Marco, L., 2013. On downside risk predictability through liquidity and trading activity: A dynamic quantile approach. International Journal of Forecasting 29, 202–219.

Schaumburg, J., 2012. Predicting extreme value at risk: Nonparametric quantile regression with refinements from extreme value theory. Computational Statistics & Data Analysis 56, 4081 – 4096.

Shiller, R.J., 2015. Irrational exuberance. Princeton university press.

Taylor, J., 2008. Estimating value at risk and expected shortfall using expectiles. Journal of Financial Econometrics 6, 231–252.

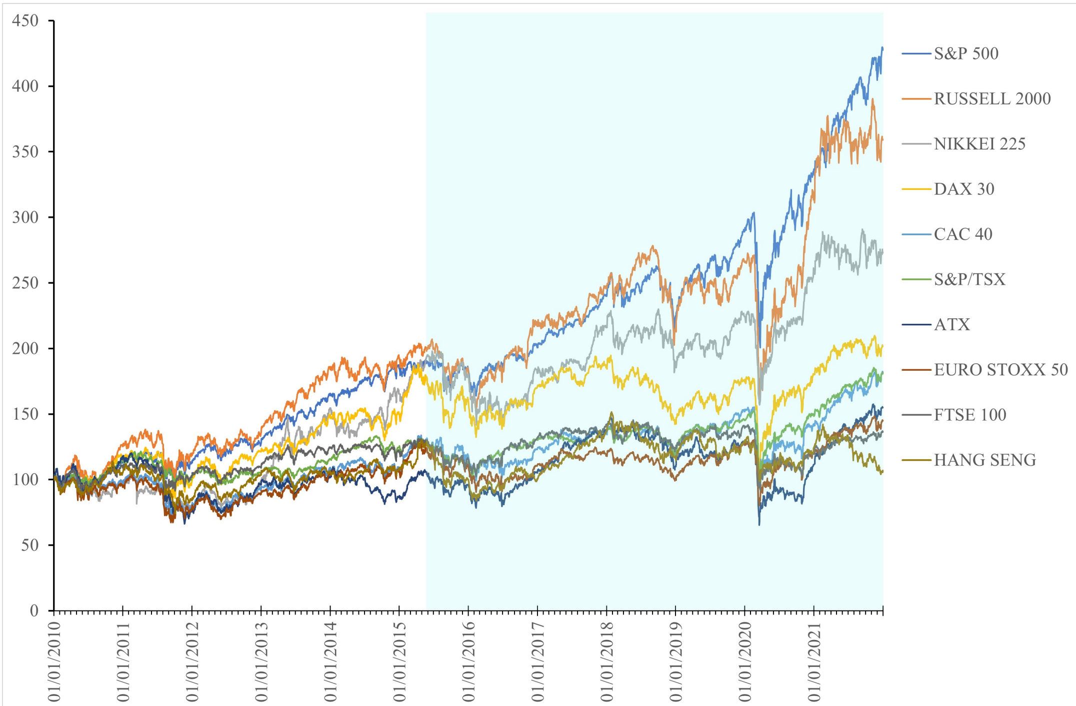
Taylor, J.W., 2019. Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric Laplace distribution. Journal of Business & Economic Statistics 37, 121–133.

Venetis, I.A., Peel, D., 2005. Non-linearity in stock index returns: the volatility and serial correlation relationship. Economic Modelling 22, 1–19.

Xiao, Z., Koenker, R., 2009. Conditional quantile estimation for generalized autoregressive conditional heteroscedasticity models. Journal of the American Statistical Association 104, 1696–1712.

Ziegel, J.F., Krüger, F., Jordan, A., Fasciati, F., 2020. Robust Forecast Evaluation of Expected Shortfall. Journal of Financial Econometrics 18, 95–120.

Appendix A. Index levels over the data sample



Note: Plot of index levels with Jan 01, 2010 set at 100 for all indices. We can see that all indices went through extreme price movements, but with a certain degree of heterogeneity across indices. The shaded zone represents the out of sample period.

Appendix B. Diebold–Mariano t-statistics on out of sample loss differences

Table B.7: S&P 500

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	8.606	-5.644	7.753	7.228	8.540	8.690	-6.161	8.667	7.595	8.367	-1.075	-1.070	-1.089
AS	-8.606	-	-8.948	-8.025	-8.659	1.372	0.178	-8.555	-2.024	-7.368	-6.781	-1.080	-1.075	-1.099
IG	5.644	8.948	-	8.537	8.008	8.834	9.029	0.421	8.998	8.455	9.086	-1.075	-1.069	-1.088
AR-IG	-7.753	8.025	-8.537	-	1.951	7.695	8.099	-8.103	8.011	4.252	4.500	-1.078	-1.072	-1.094
IG-GJR	-7.228	8.659	-8.008	-1.951	-	8.642	8.702	-7.339	8.663	2.173	1.814	-1.078	-1.073	-1.094
C-SAVd	-8.540	-1.372	-8.834	-7.695	-8.642	-	-1.091	-8.466	-2.367	-6.746	-6.354	-1.080	-1.075	-1.099
C-ASd	-8.690	-0.178	-9.029	-8.099	-8.702	1.091	-	-8.630	-2.265	-7.413	-6.994	-1.080	-1.075	-1.099
C-SAVdwm	6.161	8.555	-0.421	8.103	7.339	8.466	8.630	-	8.600	7.841	8.497	-1.075	-1.069	-1.088
C-ASdwm	-8.667	2.024	-8.998	-8.011	-8.663	2.367	2.265	-8.600	-	-7.116	-6.750	-1.080	-1.075	-1.099
C-IGd	-7.595	7.368	-8.455	-4.252	-2.173	6.746	7.413	-7.841	7.116	-	-0.257	-1.078	-1.073	-1.095
C-IGdwm	-8.367	6.781	-9.086	-4.500	-1.814	6.354	6.994	-8.497	6.750	0.257	-	-1.078	-1.073	-1.095
GARCH	1.075	1.080	1.075	1.078	1.078	1.080	1.080	1.075	1.080	1.078	1.078	-	1.251	1.058
NGARCH	1.070	1.075	1.069	1.072	1.073	1.075	1.075	1.069	1.075	1.073	1.073	-1.251	-	1.043
EGARCH	1.089	1.099	1.088	1.094	1.094	1.099	1.099	1.088	1.099	1.095	1.095	-1.058	-1.043	-

Note: We computed pairwise t-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.

Table B.8: S&P/TSX

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	8.606	-5.644	7.753	7.228	8.540	8.690	-6.161	8.667	7.595	8.367	-1.075	-1.070	-1.089
AS	-8.606	-	-8.948	-8.025	-8.659	1.372	0.178	-8.555	-2.024	-7.368	-6.781	-1.080	-1.075	-1.099
IG	5.644	8.948	-	8.537	8.008	8.834	9.029	0.421	8.998	8.455	9.086	-1.075	-1.069	-1.088
AR-IG	-7.753	8.025	-8.537	-	1.951	7.695	8.099	-8.103	8.011	4.252	4.500	-1.078	-1.072	-1.094
IG-GJR	-7.228	8.659	-8.008	-1.951	-	8.642	8.702	-7.339	8.663	2.173	1.814	-1.078	-1.073	-1.094
C-SAVd	-8.540	-1.372	-8.834	-7.695	-8.642	-	-1.091	-8.466	-2.367	-6.746	-6.354	-1.080	-1.075	-1.099
C-ASd	-8.690	-0.178	-9.029	-8.099	-8.702	1.091	-	-8.630	-2.265	-7.413	-6.994	-1.080	-1.075	-1.099
C-SAVdwm	6.161	8.555	-0.421	8.103	7.339	8.466	8.630	-	8.600	7.841	8.497	-1.075	-1.069	-1.088
C-ASdwm	-8.667	2.024	-8.998	-8.011	-8.663	2.367	2.265	-8.600	-	-7.116	-6.750	-1.080	-1.075	-1.099
C-IGd	-7.595	7.368	-8.455	-4.252	-2.173	6.746	7.413	-7.841	7.116	-	-0.257	-1.078	-1.073	-1.095
C-IGdwm	-8.367	6.781	-9.086	-4.500	-1.814	6.354	6.994	-8.497	6.750	0.257	-	-1.078	-1.073	-1.095
GARCH	1.075	1.080	1.075	1.078	1.078	1.080	1.080	1.075	1.080	1.078	1.078	-	1.251	1.058
NGARCH	1.070	1.075	1.069	1.072	1.073	1.075	1.075	1.069	1.075	1.073	1.073	-1.251	-	1.043
EGARCH	1.089	1.099	1.088	1.094	1.094	1.099	1.099	1.088	1.099	1.095	1.095	-1.058	-1.043	-

Note: We computed pairwise t-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.

Table B.9: Small Cap 2000

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	8.785	4.379	7.866	5.805	9.084	8.821	-6.953	8.786	9.202	8.119	-1.227	-1.227	-1.422
AS	-8.785	-	-7.925	-2.480	-4.281	5.772	3.117	-8.872	2.459	5.755	-1.501	-1.232	-1.232	-1.430
IG	-4.379	7.925	-	8.849	7.065	8.360	8.168	-5.066	8.031	8.354	7.310	-1.228	-1.229	-1.425
AR-IG	-7.866	2.480	-8.849	-	-4.892	5.096	3.735	-8.037	3.218	6.005	0.092	-1.231	-1.231	-1.429
IG-GJR	-5.805	4.281	-7.065	4.892	-	5.411	4.799	-6.135	4.561	6.237	3.227	-1.230	-1.230	-1.428
C-SAVd	-9.084	-5.772	-8.360	-5.096	-5.411	-	-4.304	-9.147	-5.288	5.057	-4.019	-1.232	-1.233	-1.431
C-ASd	-8.821	-3.117	-8.168	-3.735	-4.799	4.304	-	-8.905	-2.229	5.392	-2.335	-1.232	-1.232	-1.430
C-SAVdwm	6.953	8.872	5.066	8.037	6.135	9.147	8.905	-	8.881	9.230	8.256	-1.226	-1.227	-1.422
C-ASdwm	-8.786	-2.459	-8.031	-3.218	-4.561	5.288	2.229	-8.881	-	5.553	-2.032	-1.232	-1.232	-1.430
C-IGd	-9.202	-5.755	-8.354	-6.005	-6.237	-5.057	-5.392	-9.230	-5.553	-	-6.310	-1.234	-1.234	-1.434
C-IGdwm	-8.119	1.501	-7.310	-0.092	-3.227	4.019	2.335	-8.256	2.032	6.310	-	-1.231	-1.232	-1.429
GARCH	1.227	1.232	1.228	1.231	1.230	1.232	1.232	1.226	1.232	1.234	1.231	-	-0.991	0.832
NGARCH	1.227	1.232	1.229	1.231	1.230	1.233	1.232	1.227	1.232	1.234	1.232	0.991	-	0.899
EGARCH	1.422	1.430	1.425	1.429	1.428	1.431	1.430	1.422	1.430	1.434	1.429	-0.832	-0.899	-

Note: We computed pairwise t-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.

Table B.10: NIKKEI 225

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	8.114	-5.162	5.388	2.119	7.793	8.244	-2.928	8.293	4.271	2.804	-2.148	-1.912	-2.483
AS	-8.114	-	-7.566	-6.920	-7.037	5.899	2.815	-8.124	3.796	-4.671	-4.769	-2.208	-1.947	-2.562
IG	5.162	7.566	-	7.762	7.502	7.561	7.684	5.028	7.757	8.462	8.302	-2.120	-1.897	-2.446
AR-IG	-5.388	6.920	-7.762	-	-5.205	7.080	7.159	-5.648	7.332	0.368	-0.812	-2.166	-1.923	-2.506
IG-GJR	-2.119	7.037	-7.502	5.205	-	7.351	7.208	-2.504	7.347	3.383	2.220	-2.153	-1.916	-2.489
C-SAVd	-7.793	-5.899	-7.561	-7.080	-7.351	-	-5.108	-7.808	-5.244	-5.634	-5.787	-2.241	-1.966	-2.605
C-ASd	-8.244	-2.815	-7.684	-7.159	-7.208	5.108	-	-8.260	-0.402	-5.049	-5.120	-2.212	-1.949	-2.567
C-SAVdwm	2.928	8.124	-5.028	5.648	2.504	7.808	8.260	-	8.303	4.550	3.081	-2.146	-1.911	-2.481
C-ASdwm	-8.293	-3.796	-7.757	-7.332	-7.347	5.244	0.402	-8.303	-	-5.098	-5.160	-2.212	-1.949	-2.567
C-IGd	-4.271	4.671	-8.462	-0.368	-3.383	5.634	5.049	-4.550	5.098	-	-1.731	-2.168	-1.925	-2.509
C-IGdwm	-2.804	4.769	-8.302	0.812	-2.220	5.787	5.120	-3.081	5.160	1.731	-	-2.163	-1.921	-2.501
GARCH	2.148	2.208	2.120	2.166	2.153	2.241	2.212	2.146	2.212	2.168	2.163	-	-1.258	1.055
NGARCH	1.912	1.947	1.897	1.923	1.916	1.966	1.949	1.911	1.949	1.925	1.921	1.258	-	1.309
EGARCH	2.483	2.562	2.446	2.506	2.489	2.605	2.567	2.481	2.567	2.509	2.501	-1.055	-1.309	-

Note: We computed pairwise t-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.

Table B.11: HANG SENG

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	8.044	-2.878	7.106	6.798	8.719	8.180	-5.088	9.204	8.740	5.261	-2.268	-2.233	-3.162
AS	-8.044	-	-8.525	-8.034	-7.042	6.358	2.721	-7.976	8.394	-2.539	-6.786	-2.503	-2.450	-3.358
IG	2.878	8.525	-	7.999	7.919	9.044	8.600	0.356	9.328	9.127	7.034	-2.251	-2.217	-3.147
AR-IG	-7.106	8.034	-7.999	-	-2.601	7.886	7.195	-7.114	8.710	1.825	-4.923	-2.443	-2.393	-3.308
IG-GJR	-6.798	7.042	-7.919	2.601	-	7.950	7.082	-6.855	8.814	3.014	-4.429	-2.424	-2.377	-3.292
C-SAVd	-8.719	-6.358	-9.044	-7.886	-7.950	-	-5.922	-8.692	6.901	-7.026	-8.733	-2.582	-2.523	-3.423
C-ASd	-8.180	-2.721	-8.600	-7.195	-7.082	5.922	-	-8.103	8.348	-3.325	-7.056	-2.514	-2.461	-3.367
C-SAVdwm	5.088	7.976	-0.356	7.114	6.855	8.692	8.103	-	9.162	8.737	5.676	-2.253	-2.219	-3.149
C-ASdwm	-9.204	-8.394	-9.328	-8.710	-8.814	-6.901	-8.348	-9.162	-	-8.816	-8.892	-2.695	-2.628	-3.522
C-IGd	-8.740	2.539	-9.127	-1.825	-3.014	7.026	3.325	-8.737	8.816	-	-6.159	-2.474	-2.424	-3.336
C-IGdwm	-5.261	6.786	-7.034	4.923	4.429	8.733	7.056	-5.676	8.892	6.159	-	-2.363	-2.322	-3.239
GARCH	2.268	2.503	2.251	2.443	2.424	2.582	2.514	2.253	2.695	2.474	2.363	-	-0.319	-5.545
NGARCH	2.233	2.450	2.217	2.393	2.377	2.523	2.461	2.219	2.628	2.424	2.322	0.319	-	-2.149
EGARCH	3.162	3.358	3.147	3.308	3.292	3.423	3.367	3.149	3.522	3.336	3.239	5.545	2.149	-

Note: We computed pairwise t-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.

Table B.12: FTSE 100

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	9.380	0.310	8.426	8.392	9.338	9.237	-5.926	9.233	8.385	8.879	-1.238	-1.168	-1.536
AS	-9.380	-	-9.501	-8.895	-8.024	8.043	4.620	-9.351	4.767	-5.772	-5.796	-1.264	-1.187	-1.586
IG	-0.310	9.501	-	8.870	8.773	9.494	9.554	-2.484	9.498	9.361	9.517	-1.238	-1.168	-1.536
AR-IG	-8.426	8.895	-8.870	-	2.678	8.771	8.842	-8.496	8.599	4.300	6.471	-1.252	-1.179	-1.563
IG-GJR	-8.392	8.024	-8.773	-2.678	-	8.504	8.208	-8.473	7.988	2.839	4.718	-1.254	-1.180	-1.566
C-SAVd	-9.338	-8.043	-9.494	-8.771	-8.504	-	-7.551	-9.354	-7.825	-8.572	-8.472	-1.275	-1.195	-1.607
C-ASd	-9.237	-4.620	-9.554	-8.842	-8.208	7.551	-	-9.224	0.304	-6.950	-6.785	-1.266	-1.188	-1.589
C-SAVdwm	5.926	9.351	2.484	8.496	8.473	9.354	9.224	-	9.225	8.544	8.949	-1.236	-1.167	-1.532
C-ASdwm	-9.233	-4.767	-9.498	-8.599	-7.988	7.825	-0.304	-9.225	-	-6.916	-6.650	-1.266	-1.188	-1.589
C-IGd	-8.385	5.772	-9.361	-4.300	-2.839	8.572	6.950	-8.544	6.916	-	2.061	-1.257	-1.182	-1.573
C-IGdwm	-8.879	5.796	-9.517	-6.471	-4.718	8.472	6.785	-8.949	6.650	-2.061	-	-1.259	-1.183	-1.576
GARCH	1.238	1.264	1.238	1.252	1.254	1.275	1.266	1.236	1.266	1.257	1.259	-	-1.019	0.887
NGARCH	1.168	1.187	1.168	1.179	1.180	1.195	1.188	1.167	1.188	1.182	1.183	1.019	-	0.929
EGARCH	1.536	1.586	1.536	1.563	1.566	1.607	1.589	1.532	1.589	1.573	1.576	-0.887	-0.929	-

Note: We computed pairwise t-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.

Table B.13: EURO STOXX

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	7.709	7.114	7.766	6.861	7.715	8.397	-4.643	8.098	8.458	7.890	-1.330	-1.252	-1.636
AS	-7.709	-	-6.540	-1.057	-3.417	5.697	5.090	-7.767	5.332	5.105	2.625	-1.352	-1.266	-1.673
IG	-7.114	6.540	-	7.013	5.900	7.036	8.258	-7.415	7.721	8.686	7.863	-1.339	-1.258	-1.651
AR-IG	-7.766	1.057	-7.013	-	-3.668	5.434	6.079	-7.828	5.907	6.210	3.514	-1.351	-1.265	-1.671
IG-GJR	-6.861	3.417	-5.900	3.668	-	6.853	6.953	-6.993	6.949	7.679	6.060	-1.348	-1.263	-1.666
C-SAVd	-7.715	-5.697	-7.036	-5.434	-6.853	-	-0.465	-7.751	-2.797	1.092	-1.670	-1.358	-1.270	-1.684
C-ASd	-8.397	-5.090	-8.258	-6.079	-6.953	0.465	-	-8.447	-3.621	2.150	-2.101	-1.357	-1.269	-1.681
C-SAVdwm	4.643	7.767	7.415	7.828	6.993	7.751	8.447	-	8.148	8.493	7.969	-1.329	-1.252	-1.634
C-ASdwm	-8.098	-5.332	-7.721	-5.907	-6.949	2.797	3.621	-8.148	-	3.849	0.102	-1.355	-1.268	-1.678
C-IGd	-8.458	-5.105	-8.686	-6.210	-7.679	-1.092	-2.150	-8.493	-3.849	-	-5.413	-1.359	-1.271	-1.686
C-IGdwm	-7.890	-2.625	-7.863	-3.514	-6.060	1.670	2.101	-7.969	-0.102	5.413	-	-1.354	-1.268	-1.677
GARCH	1.330	1.352	1.339	1.351	1.348	1.358	1.357	1.329	1.355	1.359	1.354	-	-1.059	0.896
NGARCH	1.252	1.266	1.258	1.265	1.263	1.270	1.269	1.252	1.268	1.271	1.268	1.059	-	1.006
EGARCH	1.636	1.673	1.651	1.671	1.666	1.684	1.681	1.634	1.678	1.686	1.677	-0.896	-1.006	-

Note: We computed pairwise *t*-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.

Table B.14: DAX

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	6.600	6.048	6.767	6.797	7.311	7.140	-5.478	8.006	6.993	7.532	-1.594	-1.433	-2.191
AS	-6.600	-	-5.369	-2.414	-4.430	7.502	4.858	-6.866	8.133	2.015	1.929	-1.609	-1.445	-2.216
IG	-6.048	5.369	-	5.404	5.378	6.816	6.398	-6.659	7.622	6.614	7.024	-1.598	-1.436	-2.198
AR-IG	-6.767	2.414	-5.404	-	-4.016	6.774	5.078	-7.023	7.768	2.991	3.911	-1.608	-1.443	-2.213
IG-GJR	-6.797	4.430	-5.378	4.016	-	7.096	5.841	-7.060	7.722	4.947	6.316	-1.605	-1.442	-2.209
C-SAVd	-7.311	-7.502	-6.816	-6.774	-7.096	-	-4.862	-7.413	5.850	-4.936	-5.575	-1.622	-1.454	-2.236
C-ASd	-7.140	-4.858	-6.398	-5.078	-5.841	4.862	-	-7.341	6.916	-1.500	-2.628	-1.614	-1.448	-2.222
C-SAVdwm	5.478	6.866	6.659	7.023	7.060	7.413	7.341	-	8.081	7.152	7.631	-1.592	-1.432	-2.188
C-ASdwm	-8.006	-8.133	-7.622	-7.768	-7.722	-5.850	-6.916	-8.081	-	-6.360	-6.929	-1.632	-1.461	-2.252
C-IGd	-6.993	-2.015	-6.614	-2.991	-4.947	4.936	1.500	-7.152	6.360	-	-0.957	-1.612	-1.446	-2.219
C-IGdwm	-7.532	-1.929	-7.024	-3.911	-6.316	5.575	2.628	-7.631	6.929	0.957	-	-1.611	-1.446	-2.218
GARCH	1.594	1.609	1.598	1.608	1.605	1.622	1.614	1.592	1.632	1.612	1.611	-	-0.778	0.555
NGARCH	1.433	1.445	1.436	1.443	1.442	1.454	1.448	1.432	1.461	1.446	1.446	0.778	-	0.706
EGARCH	2.191	2.216	2.198	2.213	2.209	2.236	2.222	2.188	2.252	2.219	2.218	-0.555	-0.706	-

Note: We computed pairwise *t*-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.

Table B.15: CAC 40

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	7.920	7.678	7.303	7.173	7.753	8.418	-3.458	8.140	8.526	8.136	-1.318	-1.210	-1.574
AS	-7.920	-	-6.688	-4.760	-4.339	5.620	5.954	-7.796	5.411	5.125	2.400	-1.333	-1.218	-1.598
IG	-7.678	6.688	-	5.749	5.712	6.881	7.724	-7.435	7.389	8.395	7.663	-1.324	-1.213	-1.584
AR-IG	-7.303	4.760	-5.749	-	-0.309	6.376	6.766	-7.251	6.556	6.508	4.618	-1.330	-1.216	-1.594
IG-GJR	-7.173	4.339	-5.712	0.309	-	6.878	6.678	-7.120	6.698	6.645	4.611	-1.330	-1.216	-1.594
C-SAVd	-7.753	-5.620	-6.881	-6.376	-6.878	-	-0.154	-7.712	-2.858	0.529	-1.975	-1.337	-1.220	-1.606
C-ASd	-8.418	-5.954	-7.724	-6.766	-6.678	0.154	-	-8.329	-4.229	0.831	-2.434	-1.337	-1.220	-1.605
C-SAVdwm	3.458	7.796	7.435	7.251	7.120	7.712	8.329	-	8.037	8.428	8.031	-1.317	-1.209	-1.572
C-ASdwm	-8.140	-5.411	-7.389	-6.556	-6.698	2.858	4.229	-8.037	-	3.035	-0.535	-1.335	-1.219	-1.602
C-IGd	-8.526	-5.125	-8.395	-6.508	-6.645	-0.529	-0.831	-8.428	-3.035	-	-4.669	-1.337	-1.221	-1.606
C-IGdwm	-8.136	-2.400	-7.663	-4.618	-4.611	1.975	2.434	-8.031	0.535	4.669	-	-1.334	-1.219	-1.601
GARCH	1.318	1.333	1.324	1.330	1.330	1.337	1.337	1.317	1.335	1.337	1.334	-	-1.041	0.947
NGARCH	1.210	1.218	1.213	1.216	1.216	1.220	1.220	1.209	1.219	1.221	1.219	1.041	-	1.016
EGARCH	1.574	1.598	1.584	1.594	1.594	1.606	1.605	1.572	1.602	1.606	1.601	-0.947	-1.016	-

Note: We computed pairwise t-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.

Table B.16: ATX

	SAV	AS	IG	AR-IG	IG-GJR	C-SAVd	C-ASd	C-SAVdwm	C-ASdwm	C-IGd	C-IGdwm	GARCH	NGARCH	EGARCH
SAV	-	9.159	7.475	8.754	8.518	8.744	9.090	-7.220	9.096	8.727	8.518	-1.220	-1.120	-1.647
AS	-9.159	-	-9.158	-7.473	-4.335	5.609	4.411	-9.144	-3.115	-4.370	-4.433	-1.235	-1.131	-1.686
IG	-7.475	9.158	-	8.859	8.327	8.768	9.051	-7.945	9.100	8.668	8.614	-1.223	-1.122	-1.654
AR-IG	-8.754	7.473	-8.859	-	3.583	8.048	7.395	-8.774	7.165	3.799	2.558	-1.232	-1.129	-1.678
IG-GJR	-8.518	4.335	-8.327	-3.583	-	6.879	5.263	-8.603	3.767	0.148	-1.072	-1.234	-1.130	-1.682
C-SAVd	-8.744	-5.609	-8.768	-8.048	-6.879	-	-4.194	-8.750	-6.057	-7.003	-7.336	-1.239	-1.133	-1.695
C-ASd	-9.090	-4.411	-9.051	-7.395	-5.263	4.194	-	-9.078	-5.538	-5.305	-5.175	-1.237	-1.132	-1.689
C-SAVdwm	7.220	9.144	7.945	8.774	8.603	8.750	9.078	-	9.085	8.767	8.572	-1.218	-1.119	-1.642
C-ASdwm	-9.096	3.115	-9.100	-7.165	-3.767	6.057	5.538	-9.085	-	-3.756	-4.173	-1.235	-1.131	-1.686
C-IGd	-8.727	4.370	-8.668	-3.799	-0.148	7.003	5.305	-8.767	3.756	-	-1.296	-1.234	-1.130	-1.682
C-IGdwm	-8.518	4.433	-8.614	-2.558	1.072	7.336	5.175	-8.572	4.173	1.296	-	-1.233	-1.129	-1.680
GARCH	1.220	1.235	1.223	1.232	1.234	1.239	1.237	1.218	1.235	1.234	1.233	-	-0.899	0.928
NGARCH	1.120	1.131	1.122	1.129	1.130	1.133	1.132	1.119	1.131	1.130	1.129	0.899	-	0.923
EGARCH	1.647	1.686	1.654	1.678	1.682	1.695	1.689	1.642	1.686	1.682	1.680	-0.928	-0.923	-

Note: We computed pairwise t-statistics for the Diebold–Mariano test comparing the average losses over the out of sample period for the 14 different forecasting models. A positive value indicates that the row model has higher average loss than the column model, so the column model is superior.