Forecasting Value at Risk and Expected Shortfall Using a Semiparametric Approach Based on the Asymmetric Laplace Distribution

James W. Taylor

Saïd Business School

University of Oxford

Journal of Business and Economic Statistics, forthcoming.

Address for Correspondence:

James W. Taylor
Saïd Business School
University of Oxford
Park End Street
Oxford OX1 1HP, UK
Email: james.taylor@sbs.ox.ac.uk

Forecasting Value at Risk and Expected Shortfall Using a Semiparametric Approach Based on the Asymmetric Laplace Distribution

Abstract

Value at Risk (VaR) forecasts can be produced from conditional autoregressive VaR models, estimated using quantile regression. Quantile modeling avoids a distributional assumption, and allows the dynamics of the quantiles to differ for each probability level. However, by focusing on a quantile, these models provide no information regarding Expected Shortfall (ES), which is the expectation of the exceedances beyond the quantile. We introduce a method for predicting ES corresponding to VaR forecasts produced by quantile regression models. It is well known that quantile regression is equivalent to maximum likelihood based on an asymmetric Laplace (AL) density. We allow the density's scale to be time-varying, and show that it can be used to estimate conditional ES. This enables a joint model of conditional VaR and ES to be estimated by maximizing an AL log-likelihood. Although this estimation framework uses an AL density, it does not rely on an assumption for the returns distribution. We also use the AL log-likelihood for forecast evaluation, and show that it is strictly consistent for the joint evaluation of VaR and ES. Empirical illustration is provided using stock index data.

Keywords: Quantile regression; CAViaR; Elicitability.

1. INTRODUCTION

Value at Risk (VaR) is a tail quantile of the conditional distribution of the return on a portfolio. It has become the standard measure of market risk, and hence has been used by banks over the past two decades for setting regulatory capital requirements. Although it is an intuitive risk measure, VaR gives no information regarding possible exceedances beyond the quantile. A measure addressing this, and which can be viewed as a complement to VaR, is Expected Shortfall (ES), which is the conditional expectation of exceedances beyond the VaR. ES possesses a number of attractive properties (Acerbi and Tasche 2002). For example, in contrast to VaR, ES is a subadditive risk measure (Artzner et al. 1999), which means that the measure for a portfolio cannot be greater than the sum of the measure for the constituent parts of the portfolio. Future regulatory frameworks are likely to put increased emphasis on ES (Embrechts et al. 2014). Although many banks already calculate ES for their own risk measurement purposes, estimation is inherently challenging, as ES is a tail risk measure. Furthermore, there is no suitable loss function for evaluating ES forecasts (Gneiting 2011). In this paper, we provide a new approach to ES estimation, and a new loss function for jointly evaluating VaR and ES.

Forecasts of ES can be produced as a by-product of many VaR forecasting methods. The popular nonparametric methods, namely historical simulation and kernel density estimation, produce density forecasts from which VaR and ES predictions can be obtained. This is also the case for parametric approaches, which involve a model for the conditional variance, such as a GARCH model, and a distributional assumption. Semiparametric approaches to VaR forecasting include those that use extreme value theory (EVT) (see, for example, Chavez-Demoulin, Embrechts and Sardy 2014), and those that directly model the conditional quantile for a chosen probability level using quantile regression, such as conditional autoregressive VaR (CAViaR) modeling (see Engle and Manganelli's 2004). Directly modeling a quantile avoids the need for a

distributional assumption, and allows the dynamics of the quantiles to differ for each probability level. In empirical studies of VaR forecast accuracy, CAViaR models have performed well (see, for example, Sener, Baronyan and Mengütürk 2012). However, by focusing on a particular quantile, quantile regression models provide no apparent way of producing ES forecasts. In this paper, we address this using the asymmetric Laplace (AL) density.

Our approach uses the equivalence between quantile regression and maximum likelihood based on an AL density (see Koenker and Machado 1999). In this framework, the location and skewness parameters of the AL density are the quantile and probability level, respectively. The maximum likelihood estimator for the constant scale of the AL density is equal to the minimized quantile regression objective function divided by the sample size. Bassett, Koenker and Kordas (2004) highlight the simple relationship between this minimized objective function and the unconditional ES. This leads us to propose that a time-varying scale of the AL density can be used to produce an estimate of the time-varying conditional ES. This enables a joint model of conditional VaR and ES to be estimated by maximizing an AL likelihood. The approach is semiparametric because, although a model is specified for the VaR and ES, we do not make a distributional assumption for the returns.

In decision theory, a *scoring function* is the term for a loss function when used to evaluate a prediction of some measure of a probability distribution, such as the mean. The measure is referred to as being *elicitable* if the correct forecast of the measure is the unique minimizer of the expectation of at least one scoring function (Fissler and Ziegel 2016). The existence of such a scoring function enables the comparison of forecasts from different methods, with the best method deemed to be the one with the lowest value of the scoring function. It is possible that a measure may not be elicitable on its own, but is elicitable in combination with another measure; for example, although the variance is not elicitable, the mean and variance are

jointly elicitable (Gneiting 2011). Fissler and Ziegel (2016) show that, although ES is not elicitable, VaR and ES are jointly elicitable, and they provide a set of suitable scoring functions. We show that the negative of the AL log-likelihood is a member of this set, and hence we propose the use of this function to evaluate VaR and ES forecasts.

Section 2 briefly describes quantile regression and its link to unconditional ES. Section 3 explains how conditional VaR and ES estimates can be produced using maximum likelihood based on an AL density. Section 4 presents candidate joint models of VaR and ES. Section 5 proposes the use of the AL log-likelihood for jointly evaluating VaR and ES forecasts. Section 6 uses daily stock indices to illustrate the use of the models and the new evaluation measure.

2. QUANTILE REGRESSION AND ES

Quantile regression has been used in a variety of applications for the estimation of the parameters in a quantile model (see Koenker 2005). It involves the minimization of the sum of *tick loss* functions, as shown in expression (1), where y_t is the dependent variable, Q_t is the quantile with probability level α, $I(x)$ is the indicator function, and n is the sample size.

$$\min \sum_{t=1}^{n} \left(y_t - Q_t\right)\left(\alpha - I\left(y_t \le Q_t\right)\right) \tag{1}$$

As the common probability levels are 1% and 5% for VaR and ES estimation, in this paper, for simplicity, we consider only α<50%. With the VaR being the conditional quantile Q_t, the conditional ES is written as $ES_t = E\left(y_t \mid y_t \le Q_t\right)$. Although quantile regression focuses on the quantile for a chosen probability level, and seemingly involves no estimation of the distribution either side of the quantile, Bassett, Koenker and Kordas (2004) provide an interesting link between quantile regression and ES, by showing that ES can be written as:

$$ES_t = E(y_t) - \frac{1}{\alpha} E((y_t - Q_t)(\alpha - I(y_t \leq Q_t))) \quad (2)$$

Bassett, Koenker and Kordas suggest that this expression can be evaluated empirically using the sample mean $\bar{y}$ of y_t, and the minimized quantile regression objective function, as follows:

$$\hat{ES} = \bar{y} - \frac{1}{\alpha n} \sum_{t=1}^{n} \left(y_t - \hat{Q}_t\right)\left(\alpha - I\left(y_t \leq \hat{Q}_t\right)\right) \quad (3)$$

This would seem to show that an estimate of ES is a by-product of quantile regression (Komunjer 2007). However, only an unconditional estimate of ES is produced, as expression (3) involves averaging over the *n* values of the tick loss function. Our interest is in conditional ES estimation, and, given the heteroscedasticity in daily returns data, such an estimate is likely to be time-varying. Taylor (2008) uses exponentially weighted quantile regression for VaR estimation, and essentially replaces the summation in expression (3) with the resulting exponentially weighted summation to deliver a conditional ES estimate. In this paper, we use the AL distribution to provide a more flexible framework for the conditional modeling of VaR and ES.

3. USING THE AL DISTRIBUTION TO ESTIMATE CONDITIONAL VAR AND ES

Koenker and Machado (1999) point out that the quantile regression minimization of expression (1) is equivalent to maximum likelihood based on the AL density of expression (4). For this density, σ is a scale parameter, and Q_t is the time-varying location, which is the quantile of the density corresponding to the chosen probability level α.

$$f(y_t) = \frac{\alpha(1-\alpha)}{\sigma} \exp\left(-(y_t - Q_t)(\alpha - I(y_t \leq Q_t))/\sigma\right) \quad (4)$$

The likelihood framework has led to useful developments for quantile regression, such as statistical inference via quasi-maximum likelihood (see Komunjer 2005) and Bayesian quantile

regression (see, for example, Gerlach, Chen and Chan 2011). In these contexts, the observations y_t are not assumed to follow an AL distribution. To emphasize this, Gerlach, Chen and Chan (2011) note that the parameter α is not estimated, but is a chosen fixed value, and that it is only a quantile that is estimated. The AL likelihood simply provides a computationally convenient basis with which to enable their Bayesian approach to quantile regression.

For the scale σ of the AL density of expression (4), the maximum likelihood estimator is:

$$\hat{\sigma} = \frac{1}{n}\sum_{t=1}^{n}\left(y_t - \hat{Q}_t\right)\left(\alpha - I\left(y_t \le \hat{Q}_t\right)\right) \tag{5}$$

This is the average of the tick loss function, which can be interpreted as an unconditional estimator of the expectation of this loss function. The unconditional estimator of ES, presented in expression (3), can, therefore, be rewritten in terms of the scale estimator of expression (5):

$$\hat{ES} = \bar{y} - \frac{\hat{\sigma}}{\alpha}$$

Our proposal is to adapt this expression for conditional estimation. With this aim, we introduce a conditional AL scale σ_t, which can be viewed as the potentially time-varying conditional expectation of the tick loss function. We convey this in the following expression:

$$\sigma_t = E_t\left(\left(y_t - Q_t\right)\left(\alpha - I\left(y_t \le Q_t\right)\right)\right)$$

Using this, we can rewrite expression (2) so that we express the conditional ES in terms of the conditional AL scale σ_t and the conditional mean μ_t as follows:

$$ES_t = \mu_t - \frac{\sigma_t}{\alpha} \tag{6}$$

A model for the conditional scale σ_t can be estimated, along with a model for the conditional quantile Q_t, using maximum likelihood based on the following AL density:

$$f(y_t) = \frac{\alpha(1-\alpha)}{\sigma_t} \exp\left(-(y_t - Q_t)(\alpha - I(y_t \le Q_t))/\sigma_t\right)$$

Using expression (6), we can rewrite this density in terms of ES, as:

$$f(y_t) = \left(\frac{1-\alpha}{\mu_t - ES_t}\right) \exp\left(-\frac{(y_t - Q_t)(\alpha - I(y_t \le Q_t))}{\alpha(\mu_t - ES_t)}\right) \tag{7}$$

In this paper, we adopt the common assumption that the conditional mean of a series of daily returns r_t is a small constant value c, which can be estimated as the mean of the in-sample returns. We define y_t to be the residual $y_t = r_t - c$. The focus of our modeling is, therefore, a variable y_t with zero mean, and so we rewrite the AL density of expression (7) as:

$$f(y_t) = \frac{(\alpha - 1)}{ES_t} \exp\left(\frac{(y_t - Q_t)(\alpha - I(y_t \le Q_t))}{\alpha\, ES_t}\right) \tag{8}$$

Our proposal is to use maximum likelihood based on this AL density to estimate a joint model for the conditional quantile and conditional ES. We do not assume that the returns follow an AL distribution, because, instead of optimizing α, it is selected to be 1% or 5%, which are the probability levels of interest. If one also wished to model a time-varying conditional mean, the AL density of expression (7) could be used. To generate the parameter covariance matrix, one possibility is to draw on the work of Komunjer (2005) who investigates quantile model estimation using quasi-maximum likelihood based on a family of 'tick-exponential' densities, of which the AL density is a special case. An alternative is to use a bootstrapping procedure, and this is the approach that we use in our empirical work. To select between model specifications, the Bayesian Information Criterion could be calculated using the AL likelihood (see Lee, Noh and Park 2014).

In this section, we have highlighted the link between the scale of an AL density and ES; we have proposed that conditional modeling of the scale can deliver a conditional model for ES;

and we have suggested that this can be estimated simultaneously with a conditional quantile model using maximum likelihood based on an AL density. Although an AL density has previously been used within parametric approaches to ES estimation (see Chen, Gerlach and Lu 2012; Komunjer 2007), we are not aware of any studies that have linked the conditional scale of an AL density to conditional ES estimation.

4. JOINT MODELS FOR VAR AND ES

Our proposal is to model VaR and ES jointly, with parameters estimated by maximum likelihood based on the AL density of expression (8). In this section, we consider formulations for the VaR component of the model, and then present proposals for the ES component.

For the VaR component, we simply propose a CAViaR model. Expressions (9)-(10) present two of the CAViaR models introduced by Engle and Manganelli (2004). In these models, the β_i are constant parameters. The asymmetric slope CAViaR model aims to capture the leverage effect, which is the tendency for volatility to be greater following a negative return than a positive return of equal size.

Symmetric Absolute Value: $$Q_t = \beta_0 + \beta_1 |y_{t-1}| + \beta_2 Q_{t-1} \tag{9}$$

Asymmetric Slope: $$Q_t = \beta_0 + \beta_1 I(y_{t-1} > 0)|y_{t-1}| + \beta_2 I(y_{t-1} \leq 0)|y_{t-1}| + \beta_3 Q_{t-1} \tag{10}$$

For the ES component, we require model formulations that avoid ES estimates crossing the corresponding VaR estimates. For $\alpha <$ 50%, the ES estimate must be a value below the quantile estimate. It is straightforward to avoid crossing if we specify conditional ES to be a function of conditional VaR. This seems reasonable, as ES and VaR are, to some extent, likely to vary together, as both will vary with the time-varying volatility. The simple formulation for ES in expression (11) shows ES modeled as the product of the quantile and a constant multiplicative

factor (see Gourieroux and Liu 2012). To avoid crossing, we ensure this factor is greater than 1 by expressing it in terms of an exponential function of an unconstrained parameter γ_0.

$$ES_t = \left(1 + \exp\left(\gamma_0\right)\right) Q_t \tag{11}$$

The simplicity of this formulation is appealing. Furthermore, it correctly describes the relationship between ES and VaR for some data generating processes, such as a GARCH process with a Student *t* distribution. However, expression (11) is rather restrictive, as the dynamics of VaR may not be the same as the dynamics of ES. An alternative formulation for ES is presented in expressions (12)-(13), where the difference x_t between ES and the quantile is modeled using an autoregressive (AR) expression, which essentially smoothes the magnitude of exceedances beyond the quantile. To ensure that the quantile and ES estimates do not cross, we constrain the parameters γ_i to be non-negative.

$$ES_t = Q_t - x_t \tag{12}$$

$$x_t = \begin{cases} \gamma_0 + \gamma_1 \left(Q_{t-1} - y_{t-1}\right) + \gamma_2 \, x_{t-1} & \text{if } y_{t-1} \le Q_{t-1} \\ x_{t-1} & \text{otherwise} \end{cases} \tag{13}$$

In our empirical study of Section 6, we implement this AR formulation for ES, and the simpler ES formulation of expression (11). However, a variety of other models could certainly be considered for conditional ES. For example, the expression for x_t could take the same form as the CAViaR models, so that lagged values of $|y_t|$ or y_t^2 influence, in potentially differing ways, the dynamics of both the quantile and the difference between the quantile and ES. Another possibility is the use of a dynamic model within the multiplicative factor of expression (11).

5. EVALUATION OF VAR AND ES FORECASTS

5.1. Existing Approaches for Evaluating VaR and ES Forecasts

VaR forecast evaluation typically focuses on coverage tests. A quantile forecast $\hat{Q}_t$, for the α probability level, has correct unconditional coverage if the variable $Hit_t = \alpha - I(y_t \leq \hat{Q}_t)$ has zero unconditional expectation, and correct conditional coverage if Hit_t has zero conditional expectation (see Engle and Manganelli 2004). An alternative way to evaluate quantile forecasts is to use a scoring function. Given its use in quantile regression, a reasonable choice is the tick loss function (Giacomini and Komunjer 2005), and this has been termed the *quantile score*. We present this score in expression (14). A risk measure is elicitable if the correct forecast of the measure is the unique minimizer of the expectation of at least one scoring function. Such scoring functions are called *strictly consistent* for the risk measure (Fissler and Ziegel 2016). VaR is an elicitable risk measure, for which the quantile score is strictly consistent.

$$S(Q_t, y_t) = (y_t - Q_t)(\alpha - I(y_t \leq Q_t)) \tag{14}$$

ES is not elicitable (Gneiting 2011). In the absence of a suitable scoring function for ES, the test of McNeil and Frey (2000) is often used. This focuses on the discrepancy between the observed return and the ES forecast for the periods in which the return exceeds the VaR forecast. The standardized discrepancies should have zero unconditional and conditional expectation. Due to the typically small sample of discrepancies, a test of zero conditional expectation is generally not performed, which implies that the dynamic properties of the ES estimates are not evaluated. McNeil and Frey test for zero unconditional mean using a bootstrap test to avoid a distributional assumption. As this test focuses on observations exceeding the VaR forecasts, the assessment of ES forecasts is not independent of the VaR forecasts. This, along with ES not being elicitable, prompts consideration of a scoring function for jointly evaluating ES and VaR forecasts.

5.2. A New Scoring Function for Jointly Evaluating VaR and ES Forecasts

Fissler and Ziegel (2016) explain that VaR and ES are jointly elicitable, even though ES is not elicitable individually. They show that strictly consistent scoring functions, for jointly evaluating VaR and ES forecasts, are of the following form:

$$\begin{aligned} S(Q_t, ES_t, y_t) = &\left(I(y_t \le Q_t) - \alpha\right)G_1(Q_t) - I(y_t \le Q_t)G_1(y_t) \\ &+ G_2(ES_t)\left(ES_t - Q_t + I(y_t \le Q_t)(Q_t - y_t)/\alpha\right) - \zeta_2(ES_t) + a(y_t) \end{aligned} \tag{15}$$

where G_1, G_2, ζ_2 and a are functions satisfying a number of conditions, including the properties that $G_2=\zeta_2'$; G_1 is increasing; and ζ_2 is increasing and convex. (The domain of ζ_2 contains only negative values, because we are considering α<50%, which implies that ES_t is negative.) These conditions clearly allow a variety of alternative functions to be chosen. We consider here three examples from the set of scoring functions of expression (15).

Our first example is the score used in the empirical analysis of Fissler, Ziegel and Gneiting (2016). They consider the scoring function produced by using $G_1(x)=x$ and $G_2(x)=\exp(x)/(1+\exp(x))$ in expression (15). We set $a=\ln(2)$ to give positive values for the scoring function. We refer to this as the *FZG score*, and present it in expression (16).

$$\begin{aligned} S(Q_t, ES_t, y_t) = &\left(I(y_t \le Q_t) - \alpha\right)Q_t - I(y_t \le Q_t)y_t \\ &+ \frac{ES_t}{\left(1+\exp(ES_t)\right)}\left(ES_t - Q_t + I(y_t \le Q_t)(Q_t - y_t)/\alpha\right) + \ln\left(\frac{2}{1+\exp(ES_t)}\right) \end{aligned} \tag{16}$$

A second example from the set of scoring functions of expression (15) is the function proposed by Acerbi and Székeley (2014). We present this in expression (17), where W is a constant parameter that is large enough to ensure $WQ_t<ES_t$ for α<50%. (Note that $ES_t<0$ and $Q_t<0$.) We refer to this as the *AS score*. Fissler and Ziegel (2016) explain that, if $WQ_t<ES_t$, the AS score is a strictly consistent scoring function that can be produced by setting $G_1(x)=-(W/2)x^2$, $G_2(x)=\alpha x$ and $a=0$ in expression (15). In our empirical study of Section 6, we implemented the

AS score with W=4, as this was the smallest integer that ensured $WQ_t<ES_t$ for all pairs of forecasts of ES_t and Q_t from all methods considered in our study.

$$S(Q_t, ES_t, y_t) = \alpha\left(ES_t^2/2 + W\,Q_t^2/2 - Q_t ES_t\right) + I(y_t \le Q_t)\left(-ES_t(y_t - Q_t) + W\left(y_t^2 - Q_t^2\right)/2\right) \quad (17)$$

As a third example of a scoring function of the form of expression (15), let us consider G_1=0, $G_2(x)$=-1/x, $\zeta_2(x)$=-ln(-x), and a=1-ln(1-α). Expression (15) then becomes:

$$S(Q_t, ES_t, y_t) = -\ln\left(\frac{\alpha - 1}{ES_t}\right) - \frac{(y_t - Q_t)(\alpha - I(y_t \le Q_t))}{\alpha\, ES_t} + \frac{y_t}{ES_t} \quad (18)$$

As we have defined y_t to have zero mean, the expectation of the final summand of expression (18) is zero. Therefore, forecasts of VaR and ES that minimize the expectation of expression (18) also minimize the expectation of this scoring function if the final summand is removed, as in expression (19). This implies that expression (19) is also a strictly consistent scoring function. This function is the negative of the AL log-likelihood. We refer to it as the *AL log score*. Averaging the score across a sample gives a joint measure of VaR and ES forecast accuracy.

$$S(Q_t, ES_t, y_t) = -\ln\left(\frac{\alpha - 1}{ES_t}\right) - \frac{(y_t - Q_t)(\alpha - I(y_t \le Q_t))}{\alpha\, ES_t} \quad (19)$$

Note that if a scoring function is strictly consistent, it can also be used as the loss function in model estimation (Gneiting and Raftery 2007). This section, therefore, provides support for our proposal of estimating joint VaR and ES models by maximizing the AL log-likelihood.

Our proposal of using the AL log score to compare the forecast accuracy of methods could be viewed as advantageous for methods estimated using the AL log-likelihood. However, a similar criticism could be made for other popular scoring functions, such as the quantile score, as it is not the only strictly consistent scoring function for quantile forecasts (Gneiting 2011). Using the AL log score to evaluate VaR and ES forecasts has the theoretical appeal of being a member

of the set of scoring functions proposed by Fissler and Ziegel (2016), and it has the intuitive appeal that the AL likelihood is well established in the literature on quantile estimation.

6. EMPIRICAL STUDY OF VAR AND ES FORECASTS USING STOCK INDICES

We evaluated day-ahead VaR and ES forecasts for daily log returns of the FTSE 100, NIKKEI 225 and S&P 500 stock indices. Following common convention, we considered the 1% and 5% probability levels. Each series consisted of the 3500 daily log returns ending on 16 April 2013. We used a rolling window of 2500 observations for repeated re-estimation of each method, and evaluated day-ahead VaR and ES forecasts for the final 1000 observations. As we stated in Section 3, our modeling focuses on a residual term, defined as $y_t = r_t - c$, where r_t is the daily return and c is a constant term, which we estimated using the mean of the in-sample returns.

6.1. VaR and ES Forecasting Methods

Historical Simulation and GARCH Methods

As a simple benchmark, we produced VaR and ES forecasts using historical simulation with a moving window consisting of the 2500 observations in each estimation sample. We also considered historical simulation with moving windows of 100 and 25 observations, as in the work of Chen et al. (2012a). A short moving window has the potential advantage of enabling fast adaptation in VaR and ES estimation during periods when the market experiences major change.

We estimated GARCH(1,1) and GJRGARCH(1,1) models using maximum likelihood based on a Student *t* distribution. We produced VaR and ES forecasts using three approaches:

(i) A Student *t* distribution with degrees of freedom optimized with the model parameters.

(ii) Filtered historical simulation, which applied historical simulation to all 2500 in-sample residuals standardized by the estimated volatility.

(iii) The method of McNeil and Frey (2000), which applies the peaks-over-threshold (POT) EVT method to the standardized residuals.

Standard CAViaR with Simple ES Methods

We fitted the two CAViaR models of expressions (9)-(10), with the parameters estimated using quantile regression, as in the work of Engle and Manganelli (2004). We first sampled 10^4 candidate parameter vectors from uniform distributions with lower and upper bounds based on initial experimentation. Of these vectors, the three giving the lowest quantile regression objective function were used, in turn, as the initial vector in a quasi-Newton algorithm. The resulting vector, with lowest objective function, was chosen as the final parameter vector. When estimating the parameters for the second moving window of 2500 observations, and for all subsequent moving windows, we included, as an additional candidate, the parameter vector that had been optimized for the previous window of observations. After producing CAViaR model quantile forecasts, we used the following two approaches to forecast ES:

(i) As suggested by Manganelli and Engle (2004), we performed least squares regression, with dependent variable set as the vector of observations that exceeded the quantile estimates, and regressor set as the vector of quantile estimates. Forecasts from this model were used as ES predictions. In our results tables, we refer to this as "QR for VaR: ES = multiple of VaR", where QR emphasizes the use of quantile regression to estimate the CAViaR model.

(ii) We produced ES forecasts by summing the quantile forecast and the average in-sample quantile exceedance. In our tables, we refer to this as "QR for VaR: ES = mean exceedance".

CAViaR with EVT

Manganelli and Engle (2004) adapt the EVT-based method of McNeil and Frey (2000). A CAViaR model is first estimated for a tail quantile that is not as extreme as the VaR of interest. We followed Manganelli and Engle by estimating the 7.5% quantile. POT EVT is then applied to exceedances beyond this quantile, after standardizing the exceedances by the corresponding quantile estimates. The fitted EVT distribution is then used to obtain the 1% and 5% quantile and ES estimates of the returns. The CAViaR model, therefore, only provides the EVT threshold for the method. In our results tables, we refer to the method as "QR for 7.5% with EVT".

Joint Models for VaR and ES Estimated using AL Density

We implemented our proposed approach of Sections 3 and 4, which involves maximum likelihood based on the AL density of expression (8). We considered four joint models for VaR and ES, which each involved one of the two CAViaR formulations of expressions (9)-(10), and one of the two ES formulations of expressions (11)-(13). In our results tables, we refer to the ES formulation of expression (11) as "AL: ES = multiple of VaR", and the ES formulation of expressions (12)-(13) as "AL: ES = AR model", where AL is used to emphasize that the models have been estimated with maximum likelihood based on an AL density.

The likelihood maximization followed a similar optimization procedure to the one that we described for CAViaR models, with two notable differences. First, the quantile regression objective function was replaced by the negative of the AL log-likelihood. Second, for the candidate parameter vectors, we set the CAViaR parameters to be the values optimized separately using quantile regression, while the ES model parameters were randomly sampled. We did this to assist the optimization when the AR model of expressions (12)-(13) was used for the ES, due to the relatively large number of parameters involved. For this model, we used 10^4

candidate parameter vectors, while, for the simpler ES formulation of expression (11), we found that 10^3 was sufficient. In using, as starting values, CAViaR model parameters, estimated separately using quantile regression, we have followed the approach employed by White, Kim and Manganelli (2015) for their multi-equation models.

For the 5% probability level, expressions (20)-(22) present a joint model with asymmetric slope CAViaR formulation for VaR, and the AR formulation for ES. The parameters were estimated using the first moving window of 2500 S&P 500 returns. The expressions also present parameter standard errors in parentheses below each parameter. The parameter covariance matrix was estimated using bootstrapping, as described in the supplementary material to this paper. In the quantile model of expression (20), there is asymmetry in the response to the size of the previous period's return, and the AR parameter is relatively close to 1, which is typical of CAViaR, as well as GARCH, models. The AR parameter is also quite high in the model for x_t in expression (22).

$$\underset{}{Q_t} = \underset{(0.000159)}{-0.000298} + \underset{(0.052)}{0.023}\, I(y_{t-1} > 0)|y_{t-1}| \underset{(0.046)}{-0.174}\, I(y_{t-1} \le 0)|y_{t-1}| + \underset{(0.022)}{0.949}\, Q_{t-1} \tag{20}$$

$$ES_t = Q_t - x_t \tag{21}$$

$$x_t = \begin{cases} \underset{(0.00165)}{0.000176} + \underset{(0.076)}{0.152}(Q_{t-1} - y_{t-1}) + \underset{(0.224)}{0.840}\, x_{t-1} & \text{if } y_{t-1} \le Q_{t-1} \\ x_{t-1} & \text{otherwise} \end{cases} \tag{22}$$

For the post-sample period, Figure 1 shows the S&P 500 returns and the 5% VaR and ES forecasts from the model of expressions (20)-(22), implemented with parameter re-estimation. The plot also shows the difference between VaR and ES forecasts, which is represented by x_t in the model. As shown in expression (22), x_t responds to exceedances beyond the VaR. We

highlight these exceedances in Figure 1. The figure shows x_t varying across the post-sample period, with a clear response to the increased volatility around period 3100.

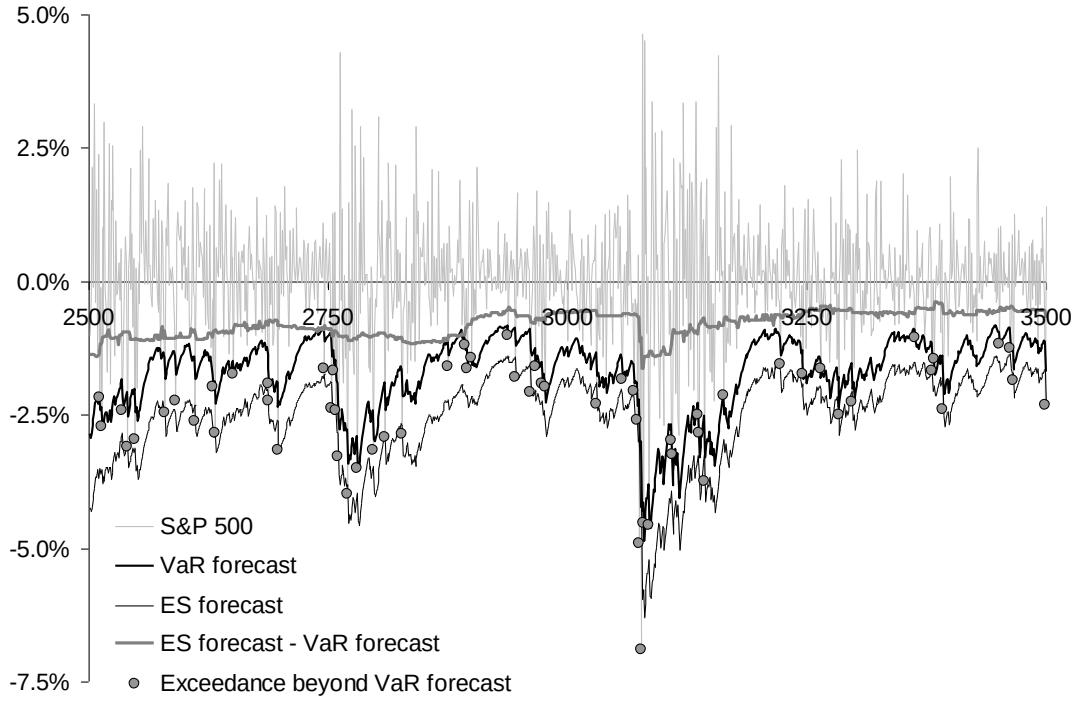

Figure 1. For the S&P 500 and 5% probability level, forecasts from asymmetric slope CAViaR with AR model for ES, jointly estimated by maximizing the AL likelihood.

The simpler formulation for ES is used in the joint model of expressions (23)-(24), which was also estimated for the 5% probability level using the first moving window of 2500 S&P 500 returns. The parameters of expression (23) are quite similar to those of expression (20).

$$Q_t = \underset{(0.000084)}{-0.000321} + \underset{(0.014)}{0.019}\, I(y_{t-1} > 0)|y_{t-1}| - \underset{(0.029)}{0.174}\, I(y_{t-1} \le 0)|y_{t-1}| + \underset{(0.034)}{0.947}\, Q_{t-1} \tag{23}$$

$$ES_t = \underset{(0.054)}{(1+\exp(-1.11))} Q_t \tag{24}$$

Figure 2 relates to the model of expressions (23)-(24), implemented with repeated re-estimation of parameters. In addition to the variables plotted in Figure 1, Figure 2 presents the re-estimated multiplicative factor (1+exp(%)) of expression (11) and (24). Although this factor is

constant for a given estimation sample, Figure 2 shows that it varies a little over time, motivating the possible use of more complex formulations for the ES. In Figure 2, the difference between the VaR and ES forecasts is generally smaller and more variable than in Figure 1. Informally, one might take the view that, in Figure 1, the ES forecasts look too extreme in comparison with the VaR exceedances. In the next section, we formally evaluate the VaR and ES forecasts.

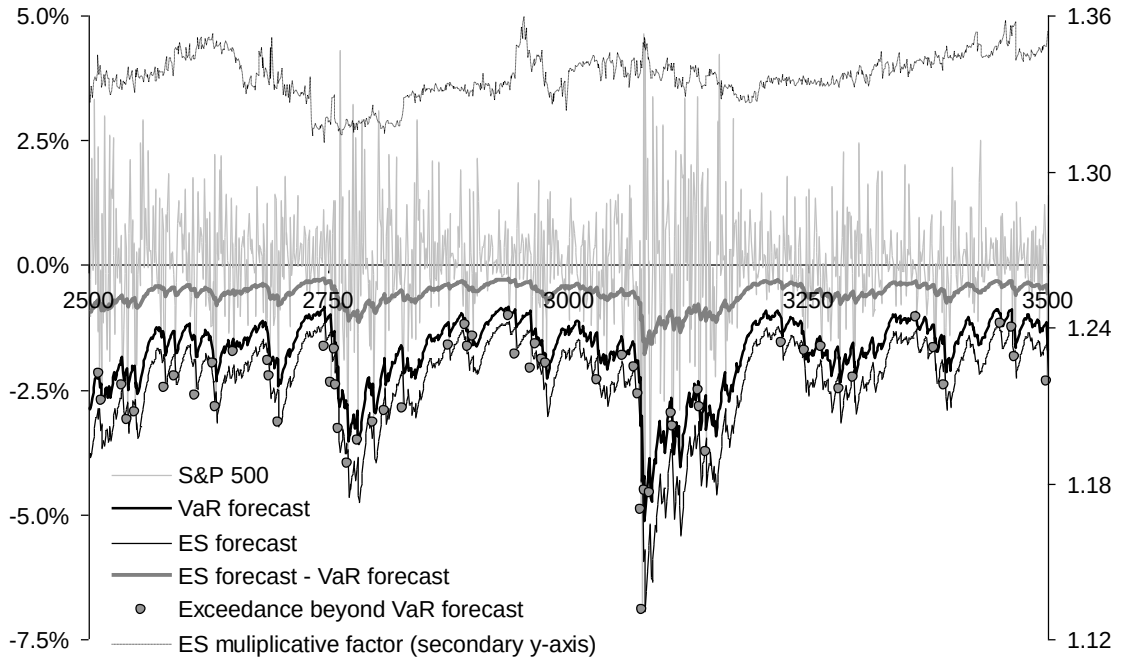

Figure 2. For the S&P 500 and 5% probability level, forecasts from asymmetric slope CAViaR with ES modeled as multiple of VaR, jointly estimated by maximizing the AL likelihood.

6.2. Evaluation of Post-Sample VaR and ES Forecasts

For the 1000 post-sample periods, we evaluated the unconditional coverage of the VaR predictions using a test based on the binomial distribution to examine whether the percentage of observations falling below the corresponding quantile estimates is significantly different from the VaR probability level. We refer to the proportion as the *hit percentage*, and present the results in Table 1 for both the 1% and 5% probability levels. To save space, we do not report the results for historical simulation based on a moving window of 25 observations, as this was

comfortably outperformed by the use of 100 observations in the moving window. Table 1 shows reasonable results for all methods, except perhaps the historical simulation methods.

Table 1. VaR unconditional coverage hit percentages.

	1%				5%			
	FTSE	NIKKEI	S&P	No. sig. at 5%	FTSE	NIKKEI	S&P	No. sig. at 5%
Historical simulation								
2500 observations	0.2*	0.3*	0.5	2	3.8	3.2*	3.9	1
100 observations	2.1*	2.1*	2.1*	3	**5.0**	**5.2**	5.5	**0**
GARCH								
Student t	1.1	**1.0**	1.8*	1	6.6*	**5.2**	6.0	1
Filtered historical simulation	0.7	0.9	1.7*	1	5.3	4.1	5.2	**0**
EVT	0.6	0.7	1.3	**0**	5.8	4.2	5.3	**0**
GJRGARCH								
Student t	1.8*	1.2	1.8*	2	6.3	4.7	6.2	**0**
Filtered historical simulation	**1.0**	0.9	1.8*	1	5.5	4.3	5.3	**0**
EVT	0.6	0.8	1.5	**0**	5.8	4.4	5.4	**0**
Symmetric absolute value CAViaR								
QR for VaR: ES = multiple of VaR	0.7	0.9	1.7*	1	5.6	4.0	5.6	**0**
QR for VaR: ES = mean exceedance	0.7	0.9	1.7*	1	5.6	4.0	5.6	**0**
QR for 7.5% with EVT	0.6	0.6	1.3	**0**	5.8	4.1	5.7	**0**
AL: ES = AR model	0.7	0.9	1.7*	1	5.8	4.2	5.4	**0**
AL: ES = multiple of VaR	0.7	0.9	1.7*	1	5.8	4.1	5.6	**0**
Asymmetric slope CAViaR								
QR for VaR: ES = multiple of VaR	0.9	0.7	1.7*	1	5.5	4.0	6.1	**0**
QR for VaR: ES = mean exceedance	0.9	0.7	1.7*	1	5.5	4.0	6.1	**0**
QR for 7.5% with EVT	0.9	0.9	**1.2**	**0**	6.0	4.0	**5.0**	**0**
AL: ES = AR model	**1.0**	0.7	1.6	**0**	5.6	3.7	6.1	**0**
AL: ES = multiple of VaR	0.9	0.7	1.7*	1	5.7	3.8	5.9	**0**

Notes. Bold indicates best method in each column. Significance at 5% level indicated by *.

We tested for conditional coverage using Engle and Manganelli's (2004) dynamic quantile test. We included four lags in the test's regression to give a test statistic that, under the null hypothesis of correct coverage, is distributed $\chi^2(6)$. Table 2 provides the p-values for the test. The results for the simplistic historical simulation methods are poor. For the GARCH

models, there was benefit in using the asymmetric model for the 1% quantile. The results are reasonable for the models that involve CAViaR quantile formulations, with no clear superiority of one of these models over another.

Table 2. VaR conditional coverage dynamic quantile test p-values.

	1%				5%			
	FTSE	NIKKEI	S&P	No. sig. at 5%	FTSE	NIKKEI	S&P	No. sig. at 5%
Historical simulation								
2500 observations	0.000	0.000	0.000	3	0.000	0.013	0.008	3
100 observations	0.000	0.000	0.000	3	0.000	0.000	0.000	3
GARCH								
Student t	0.273	0.148	0.005	1	0.223	**0.715**	0.099	**0**
Filtered historical simulation	0.001	0.059	0.004	2	0.868	0.289	0.090	**0**
EVT	0.704	0.002	0.000	2	0.873	0.358	0.095	**0**
GJRGARCH								
Student t	0.390	**0.415**	0.133	**0**	0.103	0.544	0.020	1
Filtered historical simulation	0.849	0.081	0.165	**0**	0.371	0.356	**0.248**	**0**
EVT	0.400	0.019	0.228	1	0.312	0.408	0.136	**0**
Symmetric absolute value CAViaR								
QR for VaR: ES = multiple of VaR	0.923	0.081	0.009	1	**0.936**	0.201	0.184	**0**
QR for VaR: ES = mean exceedance	0.923	0.081	0.009	1	**0.936**	0.201	0.184	**0**
QR for 7.5% with EVT	0.783	0.000	0.226	1	0.615	0.215	0.042	1
AL: ES = AR model	0.926	0.081	0.010	1	0.754	0.161	0.069	**0**
AL: ES = multiple of VaR	0.926	0.081	0.010	1	0.752	0.297	0.186	**0**
Asymmetric slope CAViaR								
QR for VaR: ES = multiple of VaR	0.972	0.002	0.372	1	0.642	0.169	0.060	**0**
QR for VaR: ES = mean exceedance	0.972	0.002	0.372	1	0.642	0.169	0.060	**0**
QR for 7.5% with EVT	0.630	0.064	0.195	**0**	0.500	0.180	0.243	**0**
AL: ES = AR model	**0.991**	0.002	0.346	1	0.646	0.194	0.060	**0**
AL: ES = multiple of VaR	0.988	0.002	**0.377**	1	0.621	0.333	0.146	**0**

Notes. Bold indicates best method in each column.

In addition to the coverage tests, we evaluated the VaR forecasts using the quantile score of expression (14). For each method, we calculated the ratio of the score to that of the historical simulation method involving 2500 observations, then subtracted this ratio from one, and multiplied the result by 100. We term this the *quantile skill score*, and present the results in

Table 3. It is essentially the quantile model pseudo R^2 presented by Koenker and Machedo (1999). Higher values indicate superior accuracy. To summarize performance across the three series for each probability level, we calculated the geometric mean of the ratios of the score for each method to the score for the historical simulation reference method, then subtracted this from one, and multiplied the result by 100. The resulting values are presented in Table 3 in the columns entitled "Geo. Mean". For both the GARCH and CAViaR-based methods, we see that the asymmetric versions were more accurate. The CAViaR-based methods compare well with the GARCH models. Although our main motivation for jointly modeling VaR and ES is to improve ES estimation, it is interesting to see that the best quantile score results, overall, are in the final row of Table 3, which corresponds to one of the new joint models.

We implemented Diebold-Mariano tests to compare the quantile score for pairs of methods. We draw on the asymptotic results of Giacomini and White (2006) to justify our use of the Diebold-Mariano test without the need for a correction for parameter estimation error. This seems reasonable, as we are using moving windows of 2500 observations for estimation and a post-sample period of 1000 observations. (For insight into the conditions under which the asymptotic results of Giacomini and White apply, see Clark and McCracken 2012.) Due to the inherent variability in the quantile score, the test's standard errors were high, and this resulted in few cases of statistical significance. In Table 3, the symbol * indicates that the quantile score for the method in that row was significantly worse (at the 5% significance level) than that of the method in the final row of the table, which corresponds to the model estimated using the AL likelihood, and ES modeled as a multiple of VaR. The symbol † indicates that the method in that row was significantly worse than that of the GJRGARCH model with EVT, which is one of the more competitive methods in Table 3. The symbol * occurs many more times than the symbol †, providing support for the new method in the final row of the table.

Table 3. VaR evaluated using quantile skill score.

	1%				5%			
	FTSE	NIKKEI	S&P	Geo. Mean	FTSE	NIKKEI	S&P	Geo. Mean
Historical simulation								
100 observations	10.5†*	-0.9†*	3.4†*	4.4	-2.6†*	-0.5*	-1.1†*	-1.4
GARCH								
Student t	25.4	15.2	16.6*	19.2	7.3	1.8	7.1	5.4
Filtered historical simulation	24.0*	15.2	16.6*	18.7	7.9	1.1†*	7.3	5.5
EVT	23.5*	14.1†	17.6*	18.5	7.7	1.3*	7.3	5.5
GJRGARCH								
Student t	25.6*	20.0	20.4	22.1	9.0	3.0	8.3	6.8
Filtered historical simulation	26.0*	**20.1**	20.1	22.1	9.0	2.7	8.1	6.6
EVT	24.9*	19.8	21.5	22.1	9.1	2.7	8.2	6.7
Symmetric absolute value CAViaR								
QR for VaR: ES = multiple of VaR	26.0	17.8	18.5	20.8	8.7	1.1*	6.8	5.6
QR for VaR: ES = mean exceedance	26.0	17.8	18.5	20.8	8.7	1.1*	6.8	5.6
QR for 7.5% with EVT	26.1	12.2	17.2	18.7	8.6	1.8	6.8	5.8
AL: ES = AR model	26.0	17.8	18.5	20.9	8.5	0.8*	6.5	5.3
AL: ES = multiple of VaR	26.0	18.1	18.7	21.0	8.4	1.3*	6.9	5.6
Asymmetric slope CAViaR								
QR for VaR: ES = multiple of VaR	26.8	18.8	20.5	22.1	9.7	3.1	8.1	7.0
QR for VaR: ES = mean exceedance	26.8	18.8	20.5	22.1	9.7	3.1	8.1	7.0
QR for 7.5% with EVT	26.6*	16.7	**21.9**	21.8	**10.0**	2.9	**8.9**	**7.3**
AL: ES = AR model	27.7	18.5	20.6	22.4	9.8	3.1*	8.0	7.0
AL: ES = multiple of VaR	**28.3**	19.2	20.7	**22.8**	**10.0**	**3.5**	8.3	**7.3**

Notes. Higher values are better. Bold indicates best method in each column.
† indicates score is significantly worse than 'GJRGARCH-EVT' at 5% level using Diebold-Mariano test.
* indicates score is significantly worse than 'AL: ES = multiple of VaR' at 5% level using Diebold-Mariano test.

To evaluate the ES forecasts, we first used McNeil and Frey's (2000) bootstrap test, which we discussed in Section 5.1. We standardized by dividing each discrepancy by the corresponding VaR estimate. The results, which are shown in Table 4, provide little insight into the relative performance of the methods. This motivates the use of an additional approach to evaluating ES forecast accuracy.

Table 4. ES evaluated using p-values for bootstrap test for zero mean VaR exceedances.

	1%				5%			
	FTSE	NIKKEI	S&P	No. sig. at 5%	FTSE	NIKKEI	S&P	No. sig. at 5%
Historical simulation								
2500 observations	0.511	0.513	0.647	**0**	0.003	0.414	0.111	1
100 observations	0.456	0.339	0.157	**0**	0.038	0.404	0.705	1
GARCH								
Student t	0.328	0.283	**0.992**	**0**	0.597	0.883	0.054	**0**
Filtered historical simulation	0.068	0.739	0.516	**0**	0.137	0.888	0.155	**0**
EVT	0.020	0.445	0.710	1	0.013	**0.960**	0.216	1
GJRGARCH								
Student t	0.264	0.884	0.657	**0**	0.459	0.571	0.012	1
Filtered historical simulation	0.449	0.931	0.195	**0**	0.666	0.953	0.017	1
EVT	**0.695**	**0.950**	0.288	**0**	0.237	0.820	0.027	1
Symmetric absolute value CAViaR								
QR for VaR: ES = multiple of VaR	0.624	0.901	0.987	**0**	0.040	0.824	0.117	1
QR for VaR: ES = mean exceedance	0.020	0.815	0.468	1	0.001	0.540	0.653	1
QR for 7.5% with EVT	0.208	0.347	0.880	**0**	0.017	0.951	0.422	1
AL: ES = AR model	0.371	0.608	0.510	**0**	0.103	0.782	0.408	**0**
AL: ES = multiple of VaR	0.354	0.830	0.571	**0**	0.027	0.924	0.374	1
Asymmetric slope CAViaR								
QR for VaR: ES = multiple of VaR	0.485	0.291	0.854	**0**	0.041	0.177	0.025	2
QR for VaR: ES = mean exceedance	0.374	0.655	0.324	**0**	0.197	0.578	**0.827**	**0**
QR for 7.5% with EVT	0.360	0.426	0.967	**0**	0.565	0.476	0.008	1
AL: ES = AR model	0.124	0.303	0.753	**0**	0.202	0.581	0.080	**0**
AL: ES = multiple of VaR	0.437	0.459	0.643	**0**	**0.804**	0.606	0.036	1

Notes. Bold indicates best method in each column.

In Section 5.2, we introduced the AL log score of expression (19) for jointly evaluating VaR and ES forecasts. In Table 5, we show the *AL log skill score*, which we calculated as the ratio of a method's AL log score to that of historical simulation using 2500 observations, then subtracted one from this ratio, and multiplied the result by 100. Higher values are preferable for this skill score, which can be viewed as a pseudo R^2 for jointly evaluating VaR and ES predictions. As with the quantile skill score, we summarize across the three series using the geometric mean. The table shows that, overall, the best AL log skill score results are in the final

row, which corresponds to the joint model estimated by maximizing the AL likelihood, and with ES modeled as a multiple of VaR. The asymmetric slope CAViaR model with EVT and the GJRGARCH model with EVT also performed well.

As with the quantile score, we implemented Diebold-Mariano tests to compare the AL log score for pairs of methods. In Table 5, the symbols * and † have the same interpretations as in Table 3. In Table 5, the symbol * occurs approximately twice as many times as the symbol †, providing support for the new method in the final row of the table.

Table 5. VaR and ES evaluated using AL log skill score.

	1%				5%			
	FTSE	NIKKEI	S&P	Geo. Mean	FTSE	NIKKEI	S&P	Geo. Mean
Historical simulation								
100 observations	6.4†*	-3.5†*	-0.4†*	0.7	0.0†*	-0.8†*	0.4†*	-0.2
GARCH								
Student t	16.6	12.8	11.4*	13.6	4.9*	3.1	4.4	4.1
Filtered historical simulation	15.8	13.0	11.8*	13.5	5.3	2.9†*	4.8	4.3
EVT	15.4	12.4†	13.1	13.6	5.2	3.0†*	4.8	4.3
GJRGARCH								
Student t	16.2	**17.7**	13.0	15.6	5.6*	4.2	5.1	5.0
Filtered historical simulation	16.6*	17.6	13.0	15.8	5.9	4.1	5.5	5.2
EVT	16.3*	17.4	14.9	16.2	5.9	4.1	5.5	5.2
Symmetric absolute value CAViaR								
QR for VaR: ES = multiple of VaR	17.6	16.2	13.1	15.6	5.9	2.9*	4.9	4.6
QR for VaR: ES = mean exceedance	17.4	15.8	13.3	15.5	5.8	2.8†*	5.1	4.5
QR for 7.5% with EVT	17.4	10.3†	12.9	13.5	5.9	2.8*	4.8	4.5
AL: ES = AR model	17.6	15.7	13.0	15.4	5.8	2.8*	4.9	4.5
AL: ES = multiple of VaR	17.6	16.3	13.7	15.8	5.8	3.0*	5.0	4.6
Asymmetric slope CAViaR								
QR for VaR: ES = multiple of VaR	17.1	17.0	14.3	16.1	6.0	4.1	5.5	5.2
QR for VaR: ES = mean exceedance	17.4	16.6	14.5	16.2	6.3	4.0*	5.8	5.4
QR for 7.5% with EVT	17.2*	14.9	**15.7**	15.9	**6.5**	3.8	**6.2**	**5.5**
AL: ES = AR model	18.0	16.6	14.3	16.3	6.4	4.2	5.9	**5.5**
AL: ES = multiple of VaR	**18.1**	17.2	14.4	**16.6**	6.4	**4.4**	5.6	**5.5**

Notes. Higher values are better. Bold indicates best method in each column.
† indicates score is significantly worse than 'GJRGARCH-EVT' at 5% level using Diebold-Mariano test.
* indicates score is significantly worse than 'AL: ES = multiple of VaR' at 5% level using Diebold-Mariano test.

Table 6. VaR and ES evaluated using FZG skill score.

	1%				5%			
	FTSE	NIKKEI	S&P	Geo. Mean	FTSE	NIKKEI	S&P	Geo. Mean
Historical simulation								
100 observations	10.6†*	-0.9†*	3.5†*	4.3	-2.6†*	-0.5*	-1.1†*	-1.4
GARCH								
Student t	25.5	15.4	16.8*	19.2	7.4	1.9	7.2	5.5
Filtered historical simulation	24.1*	15.4	16.8*	18.7	8.0	1.3†*	7.5	5.5
EVT	23.6*	14.3†	17.8*	18.5	7.8	1.5*	7.4	5.5
GJRGARCH								
Student t	25.8*	20.3	20.6	22.2	9.1	3.2	8.5	6.9
Filtered historical simulation	26.1*	**20.4**	20.3	22.2	9.2	2.8	8.3	6.7
EVT	25.0*	20.1	21.7	22.3	9.2	2.8	8.4	6.8
Symmetric absolute value CAViaR								
QR for VaR: ES = multiple of VaR	26.1	18.1	18.7	20.9	8.8	1.3*	7.0	5.6
QR for VaR: ES = mean exceedance	26.1	18.1	18.7	20.9	8.8	1.3*	7.0	5.6
QR for 7.5% with EVT	26.3	12.3	17.4	18.5	8.7	1.9	6.9	5.8
AL: ES = AR model	26.2	18.1	18.7	20.9	8.6	1.0*	6.7	5.4
AL: ES = multiple of VaR	26.2	18.4	18.9	21.1	8.6	1.4*	7.0	5.6
Asymmetric slope CAViaR								
QR for VaR: ES = multiple of VaR	27.0	19.1	20.7	22.2	9.9	3.2	8.3	7.1
QR for VaR: ES = mean exceedance	27.0	19.1	20.7	22.2	9.9	3.2	8.3	7.1
QR for 7.5% with EVT	26.8*	16.9	**22.2**	21.9	**10.2**	3.1	**9.1**	**7.4**
AL: ES = AR model	27.9	18.7	20.8	22.4	9.9	3.2*	8.1	7.1
AL: ES = multiple of VaR	**28.4**	19.5	20.9	**22.9**	10.1	**3.6**	8.5	**7.4**

Notes. Higher values are better. Bold indicates best method in each column.
† indicates score is significantly worse than 'GJRGARCH-EVT' at 5% level using Diebold-Mariano test.
* indicates score is significantly worse than 'AL: ES = multiple of VaR' at 5% level using Diebold-Mariano test.

Our use of the AL log score to evaluate VaR and ES forecasts could perhaps be viewed as giving an unfair advantage to methods estimated using the AL log-likelihood. We, therefore, also evaluate VaR and ES using the FZG and AS scores, which we described in Section 5.2, and presented in expressions (16) and (17), respectively. Tables 6 and 7 present the skill score values corresponding to these two scores, with benchmark method again chosen as historical simulation using 2500 observations. Higher values are again preferable for the skill scores. The results of Tables 6 and 7 are broadly consistent with those for the AL log score in Table 5, with the

GJRGARCH methods performing relatively well, and the best results overall produced by the joint model estimated using the AL likelihood, and with ES modeled as a multiple of VaR.

Table 7. VaR and ES evaluated using AS skill score.

	1%				5%			
	FTSE	NIKKEI	S&P	Geo. Mean	FTSE	NIKKEI	S&P	Geo. Mean
Historical simulation								
100 observations	20.0†*	1.8†	10.9†*	10.6	-1.6†*	-2.2*	-2.2†*	-2.0
GARCH								
Student t	42.7	20.3	26.8*	29.6	13.7	3.2	12.3	9.6
Filtered historical simulation	40.3*	20.4	26.4*	28.8	14.0	2.0*	12.4	9.4
EVT	38.9*	19.1	26.6*	27.9	13.9	2.3	12.4	9.4
GJRGARCH								
Student t	43.5*	28.1	34.3	35.1	15.9	5.2	**14.4**	11.7
Filtered historical simulation	43.3*	**28.5**	34.0	35.2	15.4*	4.5	13.4	11.0
EVT	40.5*	28.4	**34.4**	34.3	15.7	4.5	13.8	11.2
Symmetric absolute value CAViaR								
QR for VaR: ES = multiple of VaR	42.2*	24.9	27.3	31.3	15.3	2.0*	10.2	9.0
QR for VaR: ES = mean exceedance	42.2*	24.4	27.5	31.1	15.3	1.9*	10.3	9.0
QR for 7.5% with EVT	42.3*	15.0	24.4*	26.8	15.1	2.7	10.0	9.1
AL: ES = AR model	42.1*	24.6	27.1	31.0	14.9	1.4*	9.8	8.6
AL: ES = multiple of VaR	41.9*	25.4	27.7	31.5	14.9	2.3*	10.3	9.0
Asymmetric slope CAViaR								
QR for VaR: ES = multiple of VaR	44.9	26.6	32.1	34.3	16.6	5.1	13.6	11.7
QR for VaR: ES = mean exceedance	45.0	26.2	32.0	34.2	16.8	4.8*	13.6	11.6
QR for 7.5% with EVT	43.9*	22.7	33.6	33.1	**17.1**	5.1	14.2	**12.0**
AL: ES = AR model	46.4	25.6	31.8	34.3	16.8	4.7*	13.6	11.6
AL: ES = multiple of VaR	**47.1**	27.9	32.7	**35.7**	16.9	**5.5**	13.9	**12.0**

Notes. Higher values are better. Bold indicates best method in each column.
† indicates score is significantly worse than 'GJRGARCH-EVT' at 5% level using Diebold-Mariano test.
* indicates score is significantly worse than 'AL: ES = multiple of VaR' at 5% level using Diebold-Mariano test.

Table 8. VaR and ES evaluated using AL log skill score. Method parameters estimated without the period of the Global Financial Crisis.

	1%				5%			
	FTSE	NIKKEI	S&P	Geo. Mean	FTSE	NIKKEI	S&P	Geo. Mean
Historical simulation								
100 observations	6.4†*	-3.5†*	-0.4†*	0.7	0.0†*	-0.8†*	0.4†*	-0.2
GARCH								
Student t	16.4	12.0	10.8	13.0	4.8	3.0	4.2	4.0
Filtered historical simulation	15.9	12.3	11.0	13.0	5.3	2.9†*	4.6	4.2
EVT	15.6	11.8†	12.5	13.3	5.1	2.9†*	4.6	4.2
GJRGARCH								
Student t	15.6	**17.1**	12.5	15.1	5.5†*	4.1	5.0†	4.9
Filtered historical simulation	16.6	**17.1**	13.0†	15.5	5.9	4.0	**5.5**	5.1
EVT	16.8	16.9	**14.8**	**16.2**	5.9	4.0	**5.5**	5.1
Symmetric absolute value CAViaR								
QR for VaR: ES = multiple of VaR	18.2	15.4	11.1	14.9	5.8	3.1*	4.7	4.6
QR for VaR: ES = mean exceedance	17.9	15.1	11.4	14.8	5.8	2.9*	4.8	4.5
QR for 7.5% with EVT	17.4	11.3	12.4	13.7	5.6	3.2*	4.6	4.5
AL: ES = AR model	18.3	15.0	10.5	14.5	5.6	3.1*	4.7	4.5
AL: ES = multiple of VaR	**18.5**	16.0	11.2	15.2	5.7	3.2*	4.8	4.6
Asymmetric slope CAViaR								
QR for VaR: ES = multiple of VaR	16.9	16.4	11.2	14.8	5.9	4.3	3.9*	4.7
QR for VaR: ES = mean exceedance	17.1	16.1	11.6	14.9	6.2	4.1*	4.6	4.9
QR for 7.5% with EVT	16.5	**17.1**	11.3	14.9	6.2	4.4	4.5	5.0
AL: ES = AR model	17.1	16.1	11.2	14.7	**6.4**	4.3*	5.3	**5.3**
AL: ES = multiple of VaR	17.4	16.8	11.4	15.2	6.3	**4.6**	5.0	**5.3**

Notes. Higher values are better. Bold indicates best method in each column.
† indicates score is significantly worse than 'GJRGARCH-EVT' at 5% level using Diebold-Mariano test.
* indicates score is significantly worse than 'AL: ES = multiple of VaR' at 5% level using Diebold-Mariano test.

6.3. Influence of the Global Financial Crisis on VaR and ES Forecast Accuracy

A consequence of our dataset of 3500 periods ending on 16 April 2013 was that all estimation samples contained the period covering the height of the global financial crisis. To try to assess the influence of the crisis period on our results, we repeated our empirical study with observations 2251 to 2500 omitted from the objective functions used for parameter estimation for the various methods. These omitted observations covered the approximately 1-year periods

starting on 6 May 2008, 11 March 2008 and 29 April 2008, for the FTSE 100, NIKKEI 225 and S&P 500, respectively. Table 8 reports the resulting AL log skill scores. Comparing this table with our original results for this measure in Table 5, we see that, in general, the accuracy of the methods worsened when the crisis period was omitted from the estimation sample. This was most noticeable for the CAViaR-based methods when estimating the 1% quantile of the S&P 500. It is interesting to note that the GJRGARCH model with EVT was least affected by the removal of the crisis period from the parameter estimation.

7. SUMMARY

Using quantile regression to estimate VaR models has the appeal that it allows the quantile dynamics to differ for different probability levels. However, it leaves open the question of how to estimate ES. To address this, we have proposed that estimation is performed by maximum likelihood based on an AL density. The location of the density is the quantile, and the scale is a simple function of ES. This estimation framework avoids a distributional assumption, and enables joint modeling of the time-varying conditional VaR and ES. Estimating VaR and ES in one step has theoretical appeal in terms of efficiency, as well as being convenient from a practical perspective. In addition to its use for estimation, we have proposed that the AL likelihood be used to evaluate post-sample VaR and ES forecasts. The work of Fissler and Ziegel (2016) has enabled us to provide theoretical support for this, and hence for the use of an AL likelihood to estimate joint models for VaR and ES. Using stock index data, we evaluated the forecasts from joint models of VaR and ES estimated in this way. The results were promising, with benchmark methods not able to outperform the model consisting of an asymmetric slope CAViaR formulation for VaR, and ES expressed simply as a constant multiple of VaR. This model, estimated using the AL likelihood, provides a simple extension of the CAViaR approach

to enable simultaneous VaR and ES estimation in a semiparametric framework. Although we have focused on CAViaR models, our approach can be used for other types of VaR models estimated using quantile regression, such as models with independent variables, or models for multiple dependent variables and probability levels (see White, Kim and Manganelli 2015).

ACKNOWLEDGEMENTS

We are very grateful to the Editor, an associate editor and three referees for providing comments that helped greatly to improve the paper.

REFERENCES

Acerbi, C., and Székeley, B. (2014), "Backtesting Expected Shortfall," *Risk*, December, 76-81.

Acerbi, C., and Tasche, D. (2002), "On the Coherence of Expected Shortfall," *Journal of Banking and Finance*, 26, 1487-1503.

Artzner, P., Delbaen, F., Eber, J.M., and Heath, D. (1999), "Coherent Measures of Risk," *Mathematical Finance*, 9, 203-228.

Bassett, G., Koenker, R., and Kordas, G. (2004), "Pessimistic Portfolio Allocation and Choquet Expected Utility," *Journal of Financial Econometrics*, 2, 477–492.

Chavez-Demoulin, V., Embrechts, P., and Sardy, S. (2014), "Extreme-Quantile Tracking for Financial Time Series," *Journal of Econometrics*, 181, 44-52.

Chen, C.W.S., Gerlach, R., Hwang, B.B.K., and McAleer, M. (2012), "Forecasting Value-at-Risk using Nonlinear Regression Quantiles and the Intra-day Range," *Computational Statistics and Data Analysis*, 56, 3498-3516.

Chen, Q., Gerlach, R., and Lu, Z. (2012), "Bayesian Value-at-Risk and Expected Shortfall Forecasting via the Asymmetric Laplace Distribution," *Computational Statistics and Data Analysis*, 56, 3498-3516.

Clark, T.E., and McCracken, M.W. (2012), "Tetsing for Unconditional Predictive Ability," in *Oxford Handbook of Economic Forecasting*. Oxford, UK: Oxford University Press, 415-440.

Embrechts, P., Puccetti, G., Ruschendorf, L., Wang, R., and Beleraj, A. (2014), "An Academic Response to Basel 3.5," *Risks*, 2, 25-48.

Engle, R.F., and Manganelli, S. (2004), "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," *Journal of Business and Economic Statistics*, 22, 367-381.

Fissler, T., and Ziegel, J.A. (2016), "Higher Order Elicitability and Osband's Principle," *Annals of Statistics* forthcoming.

Fissler, T., Ziegel, J.A., and Gneiting, T. (2016), "Expected Shortfall is Jointly Elicitable with Value at Risk - Implications for Backtesting," *Risk* forthcoming.

Gerlach, R., Chen, C.W.S., and Chan, N.Y.C. (2011), "Bayesian Time-Varying Quantile Forecasting for Value-at-Risk in Financial Markets," *Journal of Business and Economic Statistics*, 29, 481-492.

Giacomini, R., and Komunjer, I. (2005), "Evaluation and Combination of Condiitonal Quantile Forecasts," *Journal of Business and Economic Statistics*, 23, 416-431.

Giacomini, R., and White, H. (2006), "Tests of Conditional Predictive Ability," *Econometrica*, 74, 1545-1578.

Gneiting, T. (2011), "Making and Evaluating Point Forecasts," *Journal of the American Statistical Association*, 106, 746-762.

Gneiting, T., and Raftery, A.E. (2007), "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, 102, 359-378.

Gourieroux, C., and Liu, W. (2012), "Converting Tail-VaR to VaR: An Econometric Study," *Journal of Financial Econometrics*, 10, 233-264.

Koenker, R.W. (2005) *Quantile Regression*. Cambridge, UK: Cambridge University Press.

Koenker, R., and Machado, J.A.F. (1999), "Goodness of Fit and Related Inference Processes for Quantile Regression," *Journal of the American Statistical Association*, 94, 1296-1310.

Komunjer, I. (2005), "Quasi-Maximum Likelihood Estimation for Conditional Quantiles," *Journal of Econometrics*, 128, 137-164.

Komunjer, I. (2007), "Asymmetric Power Distribution: Theory and Applications to Risk Measurement," *Journal of Applied Econometrics*, 22, 891-921.

Lee, E.R., Noh, H., and Park, B.U. (2014), "Model Selection via Bayesian Information Criterion for Quantile Regression Models," *Journal of the American Statistical Association*, 109, 216-229.

Manganelli, S., and Engle, R.F. (2004), "A Comparison of Value-at-Risk Models in Finance," In *Risk Measures for the 21st Century*, ed. G. Szegö, Chichester: Wiley, pp. 123-144.

McNeil, A.J., and Frey, R. (2000), "Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach," *Journal of Empirical Finance*, 7, 271-300.

Sener, E., Baronyan, S., and Mengütürk, L.A. (2012), "Ranking the Predictive Performances of Value-at-Risk Estimation Methods," *International Journal of Forecasting*, 28,849-873.

Taylor, J.W. (2008), "Using Exponentially Weighted Quantile Regression to Estimate Value at Risk and Expected Shortfall," *Journal of Financial Econometrics*, 6, 382-406.

White, H., Kim, T.-H., and Manganelli, S. (2015), "VAR for VaR: Measuring Tail Dependence Using Multivariate Regression Quantiles," *Journal of Econometrics*, 187, 169-188.