Risk Management Post Financial Crisis: A Period of Monetary Easing
Nonparametric Expectile Regression for Conditional Autoregressive Expected Shortfall Estimation
Marcelo Brutti Righi Yi Yang Paulo Sergio Ceretta

Article information:
To cite this document: Marcelo Brutti Righi Yi Yang Paulo Sergio Ceretta. "Nonparametric Expectile Regression for Conditional Autoregressive Expected Shortfall Estimation" In Risk Management Post Financial Crisis: A Period of Monetary Easing. Published online: 07 Oct 2014; 83-95.
Permanent link to this document: http://dx.doi.org/10.1108/S1569-375920140000096003

Downloaded on: 14 October 2014, At: 11:49 (PT)
References: this document contains references to 0 other documents.
To copy this document: permissions@emeraldinsight.com

Users who downloaded this article also downloaded:
Alex Yi#Hou Huang, Tsung#Wei Tseng, (2009),"Forecast of value at risk for equity indices: an analysis from developed and emerging markets", The Journal of Risk Finance, Vol. 10 Iss 4 pp. 393-409

Access to this document was granted through an Emerald subscription provided by
Token:BookSeriesAuthor:33D59305-5737-451F-8A90-4982D33D4EE8:

For Authors
If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com
Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.
Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.
*Related content and download information correct at time of download.
NONPARAMETRIC EXPECTILE REGRESSION FOR CONDITIONAL AUTOREGRESSIVE EXPECTED SHORTFALL ESTIMATION

Marcelo Brutti Righi, Yi Yang and Paulo Sergio Ceretta

ABSTRACT

In this chapter, we estimate the Expected Shortfall (ES) in conditional autoregressive expectile models by using a nonparametric multiple expectile regression via gradient tree boosting. This approach has the advantages generated by the flexibility of not having to rely on data assumptions and avoids the drawbacks and fragilities of a restrictive estimator such as Historical Simulation. We consider distinct specifications for the information sets that produce the ES estimates. The results obtained with simulated and real market data indicate that the proposed approach has good performance, with some distinctions between the specifications.

Keywords: Risk management; expected Shortfall; nonparametric expectile regression; gradient tree boosting
INTRODUCTION

In recent years, concerns with financial risk management have increased because of the frequent collapses in the financial system. This worry has spurred the development of risk management tools. A milestone of such tools is the risk measurement, which has been at the center of concepts in risk measures. Despite the fact that volatility has been synonymous with risk for a long time, when the downside risk is of primary concern, as in extreme bad events, the upside and downside movements of returns must receive different treatments. Downside risk measures are also relevant for purposes of regulation and capital requirements, issues that are crucial in modern financial systems.

The leading downside risk measure for practical use is currently Value at Risk (VaR). VaR represents the maximum loss given a confidence level during a certain period, that is, the quantile of losses distribution. However, despite its simplicity and popularity, VaR has certain drawbacks, for instance, as the quantile with a given tail probability, it does not consider losses above the quantile of interest, and it depends only on the probability of extreme losses, remaining insensitive to their magnitude. Additionally, VaR is not sub-additive, that is, despite diversification, the VaR of a portfolio could be greater than the VaR of individual assets from the same portfolio. Artzner, Delbaen, Eber, and Heath (1999) note such deficiencies of VaR by demonstrating that it is not a coherent measure of risk, and they propose the Expected Shortfall (ES) as a risk measure to overcome such shortcomings. ES is defined as the average loss given that overcomes the VaR. Thus, ES considers the magnitude of losses instead of the quantile of interest alone, and it is coherent, as noted by Acerbi and Tasche (2002), Tasche (2002), and Rockafellar and Uryasev (2002). In addition to the academic discussions of its advantages, ES is becoming increasingly widely used in the financial industry. In that sense, Yamai and Yoshida (2005) discuss the comparison between VaR and ES from a practical point of view.

These risk measures are in essence theoretical concepts. Therefore, for real applications, it is necessary to develop an estimation method to compute numeric values for these risk measures. Despite the emergence of new approaches, this is still a developing field of study because much negligence regarding the estimation of the risk measures remains. Seeking to improve the estimation for risk measures, scholars realize that current and historical information regarding financial markets, economies and even policies can
be useful. For estimating VaR, quantile regression, which directly estimates the dynamics of quantiles, is the dominant method. Despite the early approach by Chernozhukov and Umanstev (2001), work on the conditional autoregressive Value at Risk (CAViaR) models (see Engle & Manganelli, 2004), which estimate VaR conditioned to VaR lagged values, is the most prominent in the literature. Other advances for risk management from CAViaR models are presented in Taylor (2008a), Gerlach, Chen, and Chan (2011), Chen, Gerlach, Hwang, and McAleer (2012), Rubia and Sanchis-Marco (2013), and Fuertes and Olmo (2013). Regarding comparisons of the VaR estimation approaches, Bao, Lee, and Saltoglu (2006), Berkowitz, Christoffersen, and Pelletier (2011), and Mabrouk and Saadi (2012) favor the CAViaR models, emphasizing the quality of the procedures.

In the same spirit, it is possible to estimate VaR and ES by the dynamics of expectiles, which depend on both the tail realizations and their probability, whereas only the tail probability determines quantiles. Expectile regression has some advantages over quantile regression. For example, it is computationally friendlier and makes more efficient use of the available data compared to quantiles. This is because expectile estimations rely on the distance to data points, while quantiles only reflect whether an observation is above or below the predictor. Of course, the increased efficiency comes at a price of sensitivity to outliers. Taylor (2008b) and Kuan, Yeh, and Hsu (2009) extend the CAViaR concept to estimate VaR and ES using a new class of expectile models — conditional autoregressive expectiles (CARE). The basic idea is to link VaR and ES directly to conditional expectiles and then to estimate these risk measures through expectile regressions by modeling the conditional expectiles as different parametric functions of the information. However, parametric expectile regression models can be too rigid to fit complex nonlinear relationships for many real applications. Yao and Tong (1996) consider nonparametric expectile regression; however, their model is restricted by the number of explanatory variables. Aiming to improve this type of model, Yang and Zou (2013) propose a method that adopts the gradient tree-boosting algorithm (e.g., Friedman, Hastie, & Tibshirani, 2000; Schapire, 2003) to derive a fully nonparametric multiple expectile regression. This technique enjoys important advantages of tree-based methods: it can fit complex nonlinear relationships and can easily incorporate interaction effects between predictors in the final estimator, reducing the potential for modeling bias. More importantly, by combining many simple tree models adaptively, the technique can give improved predictive performance over single tree models.
With that information in mind, in this chapter, we present a new non-parametric ES estimation procedure using the model of Yang and Zou (2013). This study is a contribution to the literature because our method can estimate the dynamics of downside measures without assumptions about innovation distribution or functional forms, particularly linearity restriction, of variable relationships in a way that is not as sensitive to changes in the estimation window and the sudden or late reaction to market moves as an empirical estimator would be. Thus, nonparametric multiple expectile models can be an interesting and efficient alternative to the ES estimation, once new robust methods enter into the discussion.

The remainder of this chapter has the following structure: the section “Proposed ES Estimation Procedure” briefly presents the concept of risk measures, the principles of quantile and expectile regressions and the nonparametric multiple expectile model, and the proposed ES estimator. The section “Empirical Properties of the Proposed Procedure” presents a simulation experiment to verify certain properties of the proposed estimator as well as an empirical illustration with real data to complement the analysis. Finally, the section “Conclusion” summarizes and concludes the chapter.

**PROPOSED ES ESTIMATION PROCEDURE**

We begin by defining ES. To that purpose, let $X$ be the series of returns of a financial position with probability distribution function $F$. Thus, VaR at a significance level $\alpha \in (0,1)$ is the quantile $q_\alpha$ of $F$. Mathematically,

$$\text{VaR}_\alpha = q_\alpha(X) = \inf \{ q : F(q) \geq \alpha \}$$

(1)

Based on this definition, we note that VaR does not consider information after the quantile of interest. ES can address this drawback. The ES at significance level $\alpha$ is then the expectation of $X$, once $X$ is below VaR,

$$\text{ES}_\alpha = E[X | X < \text{VaR}_\alpha = q_\alpha(X)] = \alpha^{-1} \int_0^\alpha q_\alpha(X) ds$$

(2a)

One can directly estimate the dynamics of the risk measures VaR and ES by expectile regression. We focus here on the ES, although VaR estimation is also possible. We noted the connection between the ES and
expectiles. First, we should consider that the population \( \omega \)-expectile of \( X \) is the solution of the following minimization problem over \( m \),

\[
\tau_\omega = \arg\min_m E[(\omega - 1_{X < m})(X - m)^2] 
\]  

(2b)

By straightforward algebra, one can show that the solution \( \tau \) satisfies the expression

\[
\left( \frac{1 - 2\omega}{\omega} \right) E[(X - \tau_\omega)1_{X < \tau_\omega}] = \tau_\omega - E(X) 
\]  

(2c)

This suggests a link between ES and expectiles \( \tau_\omega \). We can rearrange Eq. (2c) as

\[
E(X|X < \tau_\omega) = \left( 1 + \frac{\omega}{(1 - 2\omega)F(\tau_\omega)} \right) T_\omega - \frac{\omega}{(1 - 2\omega)F(\tau_\omega)} E(X) 
\]  

(2d)

where \( F \) is the cdf of \( X \). Because Jones (1994) demonstrates that there is a one-to-one correspondence between expectiles and the quantiles of a related distribution, (2d) actually provides an expression for the ES through expectiles when the \( \alpha \)-quantile coincides with the \( \omega \)-expectiles, that is, \( \text{VaR}_\alpha = q_\alpha = \tau_\omega \) and \( F(\tau_\omega) = \alpha \). One can rewrite (2d) as

\[
\text{ES}_\alpha^\omega = \left( 1 + \frac{\omega}{(1 - 2\omega)\alpha} \right) \tau_\omega - \frac{\omega}{(1 - 2\omega)\alpha} E(X) 
\]  

(2e)

If \( X \) is defined as a zero mean term \( E(X) = 0 \), this further simplifies to the expression,

\[
\text{ES}_\alpha^\omega = \left( 1 + \frac{\omega}{(1 - 2\omega)\alpha} \right) \tau_\omega 
\]  

(2f)

Because the estimation of \( \text{ES}_\alpha^\omega \) is equivalent to the estimation of the corresponding expectile, the idea is to estimate the conditional expectile as a function \( h \) of past information \( \Psi_T \). As there is no need for distributional assumptions for \( X \), this type of model is semi-parametric if \( h_\omega \) is a complex
function of the linear combination of $\Psi_T$; or, it is nonparametric if $h_\omega$ is simply a complex function of $\Psi_T$. To estimate the function $h_\omega$, we use the Asymmetric Least Squares proposed by Newey and Powell (1987):

$$\hat{\tau}_\omega(X_T|\Psi_T) = \hat{h}_\omega(\Psi_T) = \arg\min_{h_\omega} \sum_{X_T} |\omega - 1_{X_T < h_\omega(\Psi_T)}(X_T - h_\omega(\Psi_T))^2$$

(3)

Based on this framework, Taylor (2008a, 2008b) and Kuan et al. (2009) extend the CAViaR model of Engle and Manganelli (2004) using a CARE approach for ES estimation, that is, lagged values of ES are present in $\Psi_T$. In the literature, distinct functional forms for $h_\omega$ in CARE modeling are used. The argument is that these parametric forms can restrict the dynamics of the conditional expectile. In that sense, one can use a nonparametric specification. Considering the canonical bivariate nonparametric model $X_T = \mu(\Psi_T) + \sigma(\Psi_T)\epsilon$, where $\mu$ and $\sigma$ are, respectively, the mean and standard deviation operators, and $\epsilon$ is a random residual term independent of $\Psi_T$. From this approach, it is easy to obtain the relation $\tau_\omega(X_T; \Psi_T = \psi) = h_\omega(\psi) = \mu(\psi) + \sigma(\psi)\tau_\omega(\epsilon)$. Yao and Tong (1996) explore such a relation, developing a local linear estimator of the expectile $\tau_\omega(X_T; \Psi_T)$ when $\Psi_T$ has one dimension. Although theoretically feasible, in practice, it is very difficult to extend their method to the multiple regression case because local regression suffers from the so-called “curse-of-dimensionality.”

To solve this limitation, Yang and Zou (2013) introduce a tree-based boosting estimator for multiple nonparametric expectile regression. The technical details are available in their paper, and we summarize the idea here. They model the conditional expectiles by combining multiple regression trees. The final model is obtained by solving the minimization problem (3) using functional gradient descent. The fundament of a boosting procedure is to combine many prediction models in a way such that the combined model has a superior prediction performance. Gradient boosting uses an iterative procedure that sequentially updates the estimator and then stops after a sufficient number of iterations. These authors provide an implementation of the algorithm in the R package erboost.

Thus, with the estimates of conditional expectiles $\tau_\omega(X_T; \Psi_T)$ obtained through the nonparametric multiple expectile regression, it is possible to compute ES at the significance level $\alpha$. As we discussed, we consider the direct relation between estimated conditional expectiles and ES through a correction factor, conform $ES_T^\alpha = \left[1 + \frac{\omega}{(1-2\alpha)\alpha}\right] \tau_\omega(X_T; \Psi_T)$. We use expectiles
as estimators of quantiles. For a fixed value of $\alpha$, we select the value of $\tau$ such that the proportion of in-sample observations lying below the estimated conditional expectiles $\tilde{\tau}_\omega(X_T|\Psi_T)$ is $\alpha$.

Once we establish the estimation procedure, the choice remains of which variables to include in the information set $\Psi_T$. Under the parametric functional forms, the most used in the literature are the Symmetric Absolute Value (SAV), the Asymmetric Slope (AS), and the Indirect GARCH - Generalized Autoregressive Conditional Heteroscedastic – (IG). The SAV considers that the past information of $X$ equally affects the conditional expectile, while the AS assumes that there is a difference regarding the impact of positive and negative values for previous returns. For its part, the IG forces data to follow a pattern of GARCH specification. Formulations (4–6) define such specifications, respectively.

$$\text{ES}^\alpha_T = \text{CF}^\alpha [\gamma_0 + \gamma_1 \tau_\omega(X_{T-1}|\Psi_{T-1}) + \gamma_2 |X_{T-1}|]$$

(4)

$$\text{ES}^\alpha_T = \text{CF}^\alpha [\gamma_0 + \gamma_1 \tau_\omega(X_{T-1}|\Psi_{T-1}) + \gamma_2 (X_{T-1})^+ + \gamma_3 (X_{T-1})^-]$$

(5)

$$\text{ES}^\alpha_T = \{ (1 - 21_{\alpha < 0.5})[\gamma_0 + \gamma_1 (\tau_\omega(X_{T-1}|\Psi_{T-1}))^2 + \gamma_2 X_{T-1}^2] \}$$

(6)

Where $\text{CF}^\alpha = \left[1 + \frac{\omega}{(1 - 2\alpha)\alpha}\right]$, $(x)^+ = \max(x, 0)$, $(x)^- = -\min(x, 0)$, and $1_p$ is the indicator function used to ensure the correct sign for the ES over long and short positions once the functional term of IG gives only positive values by definition. Although more lagged information can be included in the sets $\Psi_T$, as is usually done in the literature, we restrict it to one lag. We adapt these specifications here for a nonparametric approach by bearing in mind the relation $\tau_\omega(X_T|\Psi_T) = \mu(\Psi_T) + \sigma(\Psi_T) \cdot \tau_\omega(\epsilon)$. Despite the nonparametric functional form, $h_\omega$ is the same for all models, and the information set varies. Thus, keeping the nomenclature, we get that for the SAV, AS, and IG approaches, the information set $\Psi_T$ is, respectively, $\{\tau_\omega(X_{T-1}|\Psi_{T-1}), X_{T-1}\}$, $\{\tau_\omega(X_{T-1}|\Psi_{T-1}), (X_{T-1})^+, (X_{T-1})^-\}$, and $\{[(\tau_\omega(X_{T-1}|\Psi_{T-1}))^2], (X_{T-1})^2\}$. Hence, it is possible to estimate the ES in a flexible way, without being exposed to the dangers of the HS estimation procedures.
EMPIRICAL PROPERTIES OF THE PROPOSED PROCEDURE

In this section, we explore the efficiency of the proposed estimation procedure with regard to risk management matters. A good risk model produces estimates that match the real evolution of financial data. We explore results from simulated and real market data. For best comprehension, we split them into two sub-sections.

Simulated Data

We assess the efficiency of the proposed ES estimation procedure in the context of a Monte Carlo study. Consider a portfolio where the returns $X$ are drawn from an AR (1) – GARCH (1,1), conform formulations (7–9).

\[ X_T = 0.50X_{T-1} + \varepsilon_T \]  
\[ \varepsilon_T = \sigma_T z_T, z_T \sim t_8 \]  
\[ \sigma_T^2 = 4e - 6 + 0.10\varepsilon_{T-1}^2 + 0.85\sigma_T^2 \]

Where for period $T$, $X_T$ is the return, $\sigma_T^2$ is the conditional variance, $\varepsilon_T$ is the innovation in the conditional mean, and $Z_T$ represents a student distributed white noise series. We choose this particular data-generating process to match a realistic representation of equity portfolio returns to contemplate stylized facts such as volatility clusters and heavy tails. Under this specification, we simulate 10,000 processes of length 1,000. Once we know for sure the data-generating process, we also have the real VaR and ES values, which are presented in formulations (10) and (11).

\[ \text{VaR}_T^\alpha = 0.50X_{T-1} + \sigma_T t_8^{-1}(\alpha) \]  
\[ \text{ES}_T^\alpha = 0.50X_{T-1} + \sigma_T \left( \alpha^{-1} \int_0^\alpha t_8^{-1}(s) ds \right) \]
For each simulated process, we estimate nonparametric expectile regression, conform explained in the previous section, under the SAV, AS, and IG information sets. We choose this exercise, which is similar to an in-sample estimation, because the computational cost of using an estimation window and updating parameters at each step would be too costly. We consider the significance level of 1% because it is recommended for regulation by the Basel committee. For all of these sets, we compute the bias and standard error between the estimated and real ES. Moreover, we perform the backtesting for ES estimates of McNeil and Frey (2000), which verifies if returns over a VaR standardized by ES have a mean of zero, with 1,000 bootstraps. The alternative hypothesis of the test is that such returns have a mean larger than zero, that is, it is a one-sided test for risk underestimation. Figure exhibits plots of the densities of the bias, standard error and backtesting p-values for the three specifications used. Table 1 numerically summarizes these results.

Fig. 1 and Table 1 indicate, initially, that the bias from the SAV and AS models is slightly negative, with values in the third decimal. The IG specification is more conservative, with a bias around four times larger. This emphasizes a conservative pattern of in-sample estimations, which is corroborated by the p-values obtained for the backtesting approach. The estimates do not come close to the rejection of the null hypothesis for any simulated sample, which would indicate risk underestimation. Regarding standard errors, the dispersion of the bias is very similar for all three specifications. In short, simulated results point to a good performance of the proposed CARE estimation procedure, especially under the SAV and AS specifications. The IG information set can lead to risk overestimation, which, despite being preferable to risk underestimation, implies that the IG set could make, for instance, a financial institution retain an excess of capital that would otherwise be applied to earnings.

Table 1. Mean Bias, Standard Error, and p-Value of the Backtesting for the ES Estimates at the 1% Significance Level for the Simulated Samples.

<table>
<thead>
<tr>
<th>Model</th>
<th>Bias</th>
<th>Standard Error</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symmetric absolute value</td>
<td>−0.0062</td>
<td>0.0130</td>
<td>0.9996</td>
</tr>
<tr>
<td>Asymmetric slope</td>
<td>−0.0061</td>
<td>0.0126</td>
<td>0.9996</td>
</tr>
<tr>
<td>Indirect GARCH</td>
<td>−0.0225</td>
<td>0.0129</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
Any risk model must be able to forecast risk measures in practice. For this empirical illustration, we consider daily data for the S&P500 since its inception in March 4, 1957 to October 25, 2013, totaling 14,262 observations. This is the main index for the equity market, representing the most relevant stocks in the United States. This is an extensive sample, which considers both turbulent and calm periods. Regarding ES computation issues, we apply the same SAV, AS, and IG specifications for the 1% significance level for the data log-returns. We consider an estimation window of 2,000 days, that is, each ES prediction is based on the last four years of data as a result of the updates to the models. To simplify, we use VaR as the empirical 1% quantile of the estimation window. Fig. 2 exhibits plots of the data and estimated ES with the three specifications. Complementing the models,
Table 2 numerically summarizes the results, just as Table 1 does for the simulated results.

The plots in Fig. 2 visually indicate that the conditional nature of the proposed ES estimation procedure produces estimates that follow the dynamics of the data. It is perceptible that especially in turbulent periods such as Black Monday in 1987 (the worst result in the sample), actual returns can exceed the estimates, and because by definition ES is an expectation, it is reasonable for deviations to occur. The SAV and AS lead to very similar estimates, whereas the IG, conform pointed out in the simulated results, is more conservative. Corroborating this distinction concerning the information sets are the results in Table 2. Regarding dispersion, as emphasized by Danielsson (2002), risk forecasts fluctuate considerably from one period to the next so that if two estimation models give good risk predictions, the less volatile tends to be preferred. Again, the IG presents the more volatile estimates, while the SAV and AS are quite similar. Finally, none of

Table 2. Mean, Standard Deviation, and p-Value of the Backtesting for the ES Estimates at the 1% Significance Level for the S&P500.

<table>
<thead>
<tr>
<th>Model</th>
<th>Mean</th>
<th>Deviation</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symmetric absolute value</td>
<td>−0.0339</td>
<td>0.0106</td>
<td>0.2264</td>
</tr>
<tr>
<td>Asymmetric slope</td>
<td>−0.0339</td>
<td>0.0106</td>
<td>0.2003</td>
</tr>
<tr>
<td>Indirect GARCH</td>
<td>−0.0589</td>
<td>0.0245</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Fig. 2. Daily Log-Returns and ES Estimates for S&P500 from SAV, AS, and IG Specifications for the CARE Proposed Method.
the models rejects the null hypothesis of the backtest at the usual statistical levels. However, as this is a one-sided test, the IG that has very conservative predictions does not necessarily have the best performance. Thus, as for the simulated samples, the SAV and AS specifications are more recommended, indicating that the proposed procedure leads to good results.

CONCLUSION

In this chapter, we present a procedure for ES estimation based on nonparametric multiple expectile regression with tree-based gradient boosting. Such an approach borrows from the flexibility of an empirical estimator, without incurring the risks and fragilities of an HS method. We consider the SAV, AS, and IG specifications for the CARE model, which are often used in the financial literature. The results based on both simulations and real market data indicate that the proposed procedure has had good performance, especially under the SAV and AS information sets. Thus, this approach emerges as an interesting option for ES-based risk management in practical application. For future research, we suggest that the procedure be applied to distinct asset classes and estimation scenarios, such as different significance levels and estimation windows. Moreover, one can also conduct a comparison between the proposed model and other concurring techniques.

NOTE

1. A risk measure is coherent if it simultaneously meets the axioms of translation invariance, positive homogeneity, subadditivity, and monotonicity.

REFERENCES


