“GFC-Robust Risk Management Under the Basel Accord Using Extreme Value Methodologies”

Michael McAleer, Paulo Araújo Santos, Juan-Ángel Jiménez-Martín and Teodosio Pérez Amaral

July, 2011
GFC-Robust Risk Management Under the Basel Accord
Using Extreme Value Methodologies*

Paulo Araújo Santos
Escola Superior de Gestão e Tecnologia de Santarém
and
Center of Statistics and Applications
University of Lisbon

Juan-Ángel Jiménez-Martín
Department of Quantitative Economics
Complutense University of Madrid

Michael McAleer
Econometric Institute
Erasmus School of Economics
Erasmus University Rotterdam
and
Tinbergen Institute
The Netherlands
and
Institute of Economic Research
Kyoto University
and
Department of Quantitative Economics
Complutense University of Madrid

Teodosio Pérez Amaral
Department of Quantitative Economics
Complutense University of Madrid

July 2011

* The authors are most grateful for the helpful comments and suggestions of participants at the International Conference on Risk Modelling and Management, Madrid, Spain, June 2011. For financial support, the third author wishes to thank the Australian Research Council, National Science Council, Taiwan, and the Japan Society for the Promotion of Science. The second and fourth authors acknowledge the financial support of the Ministerio de Ciencia y Tecnologia and Comunidad de Madrid, Spain.
Abstract

In McAleer et al. (2010b), a robust risk management strategy to the Global Financial Crisis (GFC) was proposed under the Basel II Accord by selecting a Value-at-Risk (VaR) forecast that combines the forecasts of different VaR models. The robust forecast was based on the median of the point VaR forecasts of a set of conditional volatility models. In this paper we provide further evidence on the suitability of the median as a GFC-robust strategy by using an additional set of new extreme value forecasting models and by extending the sample period for comparison. These extreme value models include DPOT and Conditional EVT. Such models might be expected to be useful in explaining financial data, especially in the presence of extreme shocks that arise during a GFC. Our empirical results confirm that the median remains GFC-robust even in the presence of these new extreme value models. This is illustrated by using the S&P500 index before, during and after the 2008-09 GFC. We investigate the performance of a variety of single and combined VaR forecasts in terms of daily capital requirements and violation penalties under the Basel II Accord, as well as other criteria, including several tests for independence of the violations. The strategy based on the median, or more generally, on combined forecasts of single models, is straightforward to incorporate into existing computer software packages that are used by banks and other financial institutions.

Keywords: Value-at-Risk (VaR), DPOT, daily capital charges, robust forecasts, violation penalties, optimizing strategy, aggressive risk management, conservative risk management, Basel, global financial crisis.

JEL Classifications: G32, G11, G17, C53, C22.
1. Introduction

The Global Financial Crisis (GFC) of 2008-09 has led to substantial empirical analyses and public policy debate, and left an indelible mark on economic and financial structures worldwide, and caused a generation of investors and researchers to wonder how things could have become so bad (see, for example, Borio (2008)). There have been many questions asked about whether appropriate regulations were in place, especially in the USA, which does not enforce the Basel Accord regulations as a non-subscriber, to ensure the appropriate monitoring and encouragement of (possibly excessive) risk taking by banks and other financial institutions.

The Basel II Accord\(^1\) was designed to monitor and encourage sensible risk taking, using appropriate models of risk to calculate Value-at-Risk (VaR) and subsequent daily capital charges. VaR is defined as an estimate of the probability and size of the potential loss to be expected over a given period, and is now a standard tool in risk management. It has become especially important following the 1995 amendment to the Basel Accord, whereby banks and other Authorized Deposit-taking Institutions (ADIs) were permitted (and encouraged) to use internal models to forecast daily VaR (see Jorion (2000) for a detailed discussion). The last decade has witnessed a growing academic and professional literature comparing alternative modelling approaches to determine how to measure VaR, for portfolios of financial assets. Although such approaches are desired for portfolios of any size, especially large portfolios, statistical and computational difficulties continue to make such an analysis for large portfolios infeasible at present.

The amendment to the initial Basel Accord was designed to encourage and reward institutions with superior risk management systems. A back-testing procedure, whereby

---

\(^1\) When the Basel I Accord was concluded in 1988, no capital requirements were defined for market risk. However, regulators soon recognized the risks to a banking system if insufficient capital were held to absorb the large sudden losses from huge exposures in capital markets. During the mid-90’s, proposals were tabled for an amendment to the 1988 Accord, requiring additional capital over and above the minimum required for credit risk. Finally, a market risk capital adequacy framework was adopted in 1995 for implementation in 1998. The 1995 Basel I Accord amendment provides a menu of approaches for determining market risk capital requirements, ranging from a simple to intermediate and advanced approaches. Under the advanced approach (that is, the internal model approach), banks are allowed to calculate the capital requirement for market risk using their internal models. The use of internal models was introduced in 1998 in the European Union. The 26 June 2004 Basel II framework, implemented in many countries in 2008 (though not yet in the USA), enhanced the requirements for market risk management by including, for example, oversight rules, disclosure, management of counterparty risk in trading portfolios.
actual returns are compared with the corresponding VaR forecasts, was introduced to assess the quality of the internal models used by ADIs. In cases where internal models led to a greater number of violations than could reasonably be expected, given the confidence level, the ADI is required to hold a higher level of capital (see Table 1 for the penalties imposed under the Basel II Accord). Penalties imposed on ADIs affect profitability directly through higher capital charges, and indirectly through the imposition of a more stringent external model to forecast VaR.\(^2\) This is one reason why, in practice, financial managers generally prefer risk management strategies that are passive and conservative rather than active and aggressive.

Excessive conservatism can have a negative impact on the profitability of ADIs as higher capital charges are subsequently required. Therefore, ADIs might consider a strategy that allows an endogenous decision as to how often ADIs should violate, and hence incur violation penalties, in any financial year (for further details, see McAleer and da Veiga (2008a, 2008b), McAleer (2009), Caporin and McAleer (2010a), and McAleer et al. (2009)). Additionally, ADIs need not restrict themselves to using only a single risk model. McAleer et al. (2009, 2010b) propose a risk management strategy that consists in choosing from among different combinations of risk models to forecast VaR. They discuss a combination of forecasts that was characterized as an aggressive strategy, and another that was regarded as a conservative strategy.\(^3\)

Following such an approach, this paper suggests using a combination of VaR forecasts that also includes new extreme value VaR forecast models, such as DPOT and CEVT, to obtain a GFC-robust risk management strategy. Following McAleer (2010b), defines a crisis-robust strategy as an optimal risk management strategy that remains unchanged, regardless of whether it is used before, during or after a significant financial crisis, such as the 2008-09 GFC. Parametric methods for forecasting VaR are typically fitted to historical returns, assuming specific conditional distributions of returns, such as normality, Student-t, or generalized normal distribution. The VaR forecasts depend on

\(^2\) In the 1995 amendment (p. 16), a similar capital requirement system was recommended, but the specific penalties were left to each national supervisor. The penalty structure contained in Table 1 of this paper belongs only to Basel II, and was not part of Basel I or its 1995 amendment.

\(^3\) This is a novel possibility. Technically, a combination of forecast models is also a forecast model. In principle, the adoption of a combination of forecast models by a bank or other financial institution is not forbidden by the Basel Accords, although it may be subject to regulatory approval.
the parametric model and the conditional distribution, and can be heavily affected by a few large observations.

In this paper we also consider new extreme value models such as DPOT and CEVT, in addition to previously considered parametric and semi-parametric models. Some models provide many violations, but low daily capital charges. Moreover, these results can change drastically from tranquil to turbulent periods. Regardless of economic turbulence, the purpose is to establish a model to forecast VaR that provides a reasonable number of violations and daily capital charges.

We estimate several univariate conditional volatility models to forecast VaR, assuming different returns distributions (specifically, Gaussian, Student-t and Generalized Normal). Additionally, we present 12 new strategies based on combinations of standard model VaR forecasts, namely: lowerbound, upperbound (as defined in McAleer et al. (2009)), the average, and nine additional strategies based on the 10th, … 50th, … 90th percentiles. Additionally, we consider a DPOT model with c=2/3, another DPOT model with c=3/4, and a CEVT model. These models are compared over three different time periods to investigate whether we can establish a risk management strategy that is GFC-crisis-robust. We provide evidence that using the median of the point VaR forecasts of a set of univariate conditional volatility models is a GFC-robust risk measure, such that a risk management strategy based on the median forecast is found to be superior to alternative single and combined model alternatives.

The remainder of the paper is organized as follows. In Section 2 we present the main ideas of the Basel II Accord Amendment as it relates to forecasting VaR and daily capital charges. Section 3 reviews some of the most well-known models of conditional volatility used to forecast VaR, and the three new extreme value models. In Section 4 the data used for estimation and forecasting are presented. Section 5 analyses the robust VaR forecasts before, during and after the 2008-09 GFC. Section 6 presents some conclusions.
2. Forecasting Value-at-Risk and Daily Capital Charges

As is widely known, the Basel II Accord stipulates that daily capital charges (DCC) must be set at the higher of the previous day’s VaR or the average VaR over the last 60 business days, multiplied by a factor \((3+k)\) for a violation penalty, wherein a violation involves the actual negative returns exceeding the VaR forecast negative returns for a given day:4

\[
DCC_t = \sup \left\{ -(3 + k) \overline{VaR}_{60}, \, -VaR_{t-1} \right\}
\]  

(1)

where

\(DCC_t\) = daily capital charges, which is the higher of \( (3 + k) \overline{VaR}_{60} \) and \(-VaR_{t-1}\),

\(VaR_t\) = Value-at-Risk for day \(t\),

\(VaR_t = \hat{Y}_t - z_t \cdot \hat{\sigma}_t\),

\(\overline{VaR}_{60}\) = mean VaR over the previous 60 working days,

\(\hat{Y}_t\) = estimated return at time \(t\),

\(z_t = 1\%\) critical value of the distribution of returns at time \(t\),

\(\hat{\sigma}_t\) = estimated risk (or square root of volatility) at time \(t\),

\(0 \leq k \leq 1\) is the Basel II violation penalty (see Table 1).

[Insert Table 1 here]

---

4 Our aim is to investigate the likely performance of the Basel II regulations. In this section we carry out our analysis applying the Basel II formulae to a period that includes the 2008-09 GFC, during which the Basel II Accord regulations were not fully implemented.
The multiplication factor (or penalty), $k$, depends on the central authority’s assessment of the ADI’s risk management practices and the results of a simple backtest. It is determined by the number of times actual losses exceed a particular day’s VaR forecast (Basel Committee on Banking Supervision (1996, 2006)). It should be noted that the calculation of $DCC_t$ is undertaken in terms of obtaining the lower numerical value rather than evaluating statistically significant differences between measures. In this sense, the supremum of the two values in equation (1) is similar to a two-horse race, in which there can be only one winner.

As discussed in Stahl (1997), the minimum multiplication factor of 3 is intended to compensate for various errors that can arise in model implementation, such as simplifying assumptions, analytical approximations, small sample biases and numerical errors that tend to reduce the true risk coverage of the model. Increases in the multiplication factor are designed to increase the confidence level that is implied by the observed number of violations to the 99 per cent confidence level, as required by regulators (for a detailed discussion of VaR, as well as exogenous and endogenous violations, see McAleer (2009), Jiménez-Martin et al. (2009), and McAleer et al. (2009)).

In calculating the number of violations, ADIs are required to compare the forecasts of VaR with realised profit and loss figures for the previous 250 trading days. In 1995, the 1988 Basel Accord (Basel Committee on Banking Supervision (1988)) was amended to allow ADIs to use internal models to determine their VaR thresholds (Basel Committee on Banking Supervision (1995)), However, ADIs that propose using internal models are required to demonstrate that their models are sound. Movement from the green zone to the red zone arises through an excessive number of violations. Although this will lead to a higher value of $k$, and hence a higher penalty, violations will also tend to be associated with lower daily capital charges.6

---

5 The formula in equation (1) is contained in the 1995 amendment to Basel I, while Table 1 appears for the first time in the Basel II Accord in 2004.

6 The number of violations in a given period is an important (though not the only) guide for regulators to approve a given VaR model.
Value-at-Risk refers to the lower bound of a confidence interval for a (conditional) mean, that is, a “worst case scenario on a typical day”. If interest lies in modelling the random variable, \( Y_t \), it could be decomposed as follows:

\[
Y_t = E(Y_t | F_{t-1}) + \varepsilon_t. \tag{2}
\]

This decomposition states that \( Y_t \) comprises a predictable component, \( E(Y_t | F_{t-1}) \), which is the conditional mean, and a random component, \( \varepsilon_t \). The variability of \( Y_t \), and hence its distribution, is determined by the variability of \( \varepsilon_t \). If it is assumed that \( \varepsilon_t \) follows a conditional distribution, such that:

\[
\varepsilon_t \sim D(\mu_t, \sigma_t^2)
\]

where \( \mu_t \) and \( \sigma_t \) are the conditional mean and standard deviation of \( \varepsilon_t \), respectively, these can be estimated using a variety of parametric, semi-parametric or non-parametric methods. The VaR threshold for \( Y_t \) can be calculated as:

\[
VaR_t = E(Y_t | F_{t-1}) - \alpha \sigma_t, \tag{3}
\]

where \( \alpha \) is the critical value from the distribution of \( \varepsilon_t \) to obtain the appropriate confidence level. It is possible for \( \sigma_t \) to be replaced by alternative estimates of the conditional standard deviation in order to obtain an appropriate VaR (for useful reviews of theoretical results for conditional volatility models, see Li et al. (2002) and McAleer (2005), who discusses a variety of univariate and multivariate, conditional, stochastic and realized volatility models).

As discussed in McAleer et al. (2010b), some recent empirical studies (see, for example, Berkowitz and O'Brien (2001), Gizycki and Hereford (1998), and Pérignon et al. (2008)) have indicated that some financial institutions overestimate their market risks in
disclosures to the appropriate regulatory authorities, which can imply a costly restriction to the banks trading activity. ADIs may prefer to report high VaR numbers to avoid the possibility of regulatory intrusion. This conservative risk reporting suggests that efficiency gains may be feasible. In particular, as ADIs have effective tools for the measurement of market risk, while satisfying the qualitative requirements, ADIs could conceivably reduce daily capital charges by implementing a context-dependent market risk disclosure policy. For a discussion of alternative approaches to optimize VaR and daily capital charges, see McAleer (2009) and McAleer et al. (2009).

The next section describes several volatility models that are widely used to forecast the 1-day ahead conditional variances and VaR thresholds.

3. Models for Forecasting VaR

It is well known that ADIs can use internal models to determine their VaR thresholds (see, for example, McAleer et al. (2010b)). There are alternative time series models for estimating conditional volatility. In what follows, we present several conditional volatility models that are widely used in the financial econometrics literature to evaluate strategic market risk disclosure, namely GARCH and GJR, with normal, Student-t and Generalized normal distribution errors, where the parameters are estimated.

These models are chosen as they are well known and widely used in the literature. For an extensive discussion of the theoretical properties of several of these models, see Ling and McAleer (2002a, 2002b, 2003a) and Caporin and McAleer (2010a). As an alternative to estimating the parameters, we also consider the exponential weighted moving average (EWMA) method by Riskmetrics (1996) and Zumbauch, (2007) that calibrates the unknown parameters. We include a section on these models to present them in a unified framework and notation, and to make explicit the specific versions we are using. Apart from EWMA, the models are presented in increasing order of complexity.

3.1 GARCH
For a wide range of financial data series, time-varying conditional variances can be explained empirically through the AutoRegressive Conditional Heteroskedasticity (ARCH) model, which was proposed by Engle (1982). When the time-varying conditional variance has both autoregressive and moving average components, this leads to the Generalized ARCH(p,q), or GARCH(p,q), model of Bollerslev (1986). It is very common to impose the widely estimated GARCH(1,1) specification in advance.

Consider the stationary AR(1)-GARCH(1,1) model for daily returns, \( y_t \):

\[
y_t = \phi_1 + \phi_2 y_{t-1} + \epsilon_t, \quad |\phi_2| < 1
\]

for \( t = 1, \ldots, n \), where the shocks to returns are given by:

\[
\epsilon_t = \eta_t \sqrt{h_t}, \quad \eta_t \sim iid(0,1)
\]

\[
h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1},
\]

and \( \omega > 0, \alpha \geq 0, \beta \geq 0 \) are sufficient conditions to ensure that the conditional variance \( h_t > 0 \). The stationary AR(1)-GARCH(1,1) model can be modified to incorporate a non-stationary ARMA(p,q) conditional mean and a stationary GARCH(r,s) conditional variance, as in Ling and McAleer (2003b).

### 3.2 GJR

In the symmetric GARCH model, the effects of positive shocks (or upward movements in daily returns) on the conditional variance, \( h_t \), are assumed to be the same as the negative shocks (or downward movements in daily returns). In order to accommodate asymmetric behaviour, Glosten, Jagannathan and Runkle (1992) proposed a model (hereafter GJR), for which GJR(1,1) is defined as follows:

\[
h_t = \omega + (\alpha + \gamma I(\eta_t)) \epsilon_{t-1}^2 + \beta h_{t-1},
\]

where \( \omega > 0, \alpha \geq 0, \alpha + \gamma \geq 0, \beta \geq 0 \) are sufficient conditions for \( h_t > 0 \), and \( I(\eta_t) \) is an indicator variable defined by:
\[ I(\eta) = \begin{cases} 1, & \varepsilon_t < 0 \\ 0, & \varepsilon_t \geq 0 \end{cases} \]  

(7)

as \( \eta \) has the same sign as \( \varepsilon_t \). The indicator variable differentiates between positive and negative shocks, so that asymmetric effects in the data are captured by the coefficient \( \gamma \). For financial data, it is expected that \( \gamma \geq 0 \) because negative shocks have a greater impact on risk than do positive shocks of similar magnitude. The asymmetric effect, \( \gamma \), measures the contribution of shocks to both short run persistence, \( \alpha + \gamma/2 \), and to long run persistence, \( \alpha + \beta + \gamma/2 \).

Although GJR permits asymmetric effects of positive and negative shocks of equal magnitude on conditional volatility, the special case of leverage, whereby negative shocks increase volatility while positive shocks decrease volatility (see Black (1976) for an argument using the debt/equity ratio), cannot be accommodated (for further details on asymmetry versus leverage in the GJR model, see Caporin and McAleer (2010b)).

### 3.3 Exponentially Weighted Moving Average (EWMA)

As an alternative to estimating the parameters of the appropriate conditional volatility models, Riskmetrics (1996) developed a model which estimates the conditional variances and covariances based on the exponentially weighted moving average (EWMA) method, which is, in effect, a restricted version of the ARCH(\( \infty \)) model. This approach forecasts the conditional variance at time \( t \) as a linear combination of the lagged conditional variance and the squared unconditional shock at time \( t-1 \). The EWMA model calibrates the conditional variance as:

\[ h_t = \lambda h_{t-1} + (1-\lambda)\epsilon^2_{t-1} \]  

(8)

where \( \lambda \) is a decay parameter. Riskmetrics (1996) suggests that \( \lambda \) should be set at 0.94 for purposes of analysing daily data. As no parameters are estimated, there are no moment or log-moment conditions.
3.4 Extreme Value Theory models

In what follows, we present two Extreme Value Theory (EVT) models, namely Conditional EVT (CEVT) and Duration based Peaks Over Threshold (DPOT). The first is well known and is widely used in the literature. The second was recently proposed by Araújo Santos and Fraga Alves (2011). Such models might be expected to be useful in explaining financial data, especially in the presence of extreme shocks that arise during a GFC.

3.4.1 CEVT

This approach is a two-stage hybrid method which combines a time-varying volatility model with the Peaks Over Threshold method from EVT (for details about the POT method, see Embrechts et al. (1997)). Diebold et al. (1998) proposed in a first step the standardization of the returns through the conditional means and variances estimated with a time-varying volatility model, and in a second step, the estimation of a p-quantile using EVT and the standardized returns. McNeil and Frey (2000) combine an AR(1)-GARCH(1,1) process, assuming normal innovations, with the POT method. The filter with normal innovations, while capable of removing the majority of clusterings, will frequently be a misspecified model for returns. In order to accommodate this misspecification, Kuester et al. (2006) suggested a filter with the skewed t distribution. We will denote this model as CEVT.

The one-day-ahead VaR forecast is calculated with the following equation:

\[
\text{VaR}_{t+1,CEVT} (p) = \hat{\mu}_{t+1|t} + \hat{\sigma}_{t+1|t} \hat{z}_p, \tag{9}
\]

where \(\hat{\mu}_{t+1|t}\) and \(\hat{\sigma}_{t+1|t}\) are the estimated conditional mean and conditional standard deviation for \(t+1\), respectively, obtained from an AR(1)-GARCH(1,1) process. Moreover, \(\hat{z}_p\) is a quantile \(p\) estimate, obtained with the POT method and the standardized residuals that are calculated as

\[
\left( z_{t-n+1}, \ldots, z_t \right) = \left( \frac{r_{t-n+1} - \hat{\mu}_{t-n+1}}{\hat{\sigma}_{t-n+1}}, \ldots, \frac{r_t - \hat{\mu}_t}{\hat{\sigma}_t} \right) \tag{10}
\]
Several studies have concluded that conditional EVT is the method with better out-of-sample performance to forecast one-day-ahead VaR (see, for example, McNeil and Frey (2000), Byström (2004), Bekiros and Georgoutsos (2005), Kuester et al. (2006), Ghorbel and Trabelsi (2008), and Ozun et al. (2010)). The POT method requires the choice of a high threshold $u$. We choose the threshold, $u$, such that 10% of the values are larger than $u$ (see McNeil and Frey (2000) for a simulation study that supports a similar choice).

### 3.4.2 DPOT Model

The POT method is based on the excesses over $u$ and on the Pickands-Balkema-de Haan Theorem (see Balkema and de Haan (1974) and Pickands(1975)). For distributions in the maximum domain of attraction of an extreme value distribution, this theorem states that when $u$ converges to the right end point ($x_F$) of the distribution, the excess distribution $P[X-u \mid X > u]$ converges to the Generalized Pareto Distribution (GPD):

$$G_{y,\sigma}(y) = \begin{cases} 1 - (1 + y / \sigma)^{-1/\gamma} , & \gamma \neq 0 \\ 1 - \exp(-y / \sigma) , & \gamma = 0, \end{cases}$$

where $\sigma > 0$, and the support is $y \geq 0$ when $\gamma \geq 0$, and $0 \leq y \leq -\sigma / \gamma$ when $\gamma < 0$.

Smith (1987) proposed a tail estimator based on a GPD approximation to the excess distribution. Inverting this estimator gives an equation to calculate the VaR forecast. With financial time series, a relation between the excesses and the durations between excesses is usually observed. Araújo Santos and Fraga Alves (2011) proposed using this dependence to improve the risk forecasts with duration-based POT models (DPOT). For estimation, these models use the durations, at time of excess $i$, as the preceding $v$ excesses ($d_{i,v}$). At time $t$, $d_{t,v}$ denotes the duration until $t$ as the preceding $v$ excesses.

The DPOT model assumes the GPD for the excess $Y_t$ above $u$, such that

$$Y_t \sim \text{GPD}(\gamma, \sigma_t = \alpha/(d_{t,v})^c),$$

(12)
where $\gamma$ and $\alpha$ are parameters to be estimated. The proposed DPOT model implies, for $\gamma < 1$, a conditional expected value for the excess, and for $\gamma < 1/2$, a conditional variance, both of which are dependent on $d_{t,v}$:

$$
E[Y_t | \Omega_t] = \frac{\sigma}{1 - \gamma} \quad (\gamma < 1), \quad VAR[Y_t | \Omega_t] = \frac{(\sigma_t)^2}{(1 - 2\gamma)} \quad (\gamma < 1/2). \quad (13)
$$

Inverting the tail estimator based on the conditional GPD gives the equation to calculate the DPOT VaR forecast:

$$
\hat{VaR}_{t+1|t}^{DPOT(v,c)}(p) = u + \frac{\hat{\alpha}}{\hat{\gamma}(d_{t,v})} \left( \left( \frac{n_{x}}{n_{x}P} \right)^p - 1 \right), \quad (14)
$$

where $n_x$ denotes the sample size, $n$ the number of excesses, $\hat{\gamma}$ and $\hat{\alpha}$ are estimators of $\gamma$ and $\alpha$, respectively. We choose $v = 3$ and $c \in \{2/3, 3/4\}$, as values of $c$ close or equal to $3/4$ have been shown to exhibit the best results (see Araújo Santos and Fraga Alves (2011)).

### 3.5 Unconditional Coverage and Independence Tests

The primary tool for assessing the accuracy of the interval forecasts is to monitor the binary sequence generated by observing whether the return on day $t+1$ is in the tail region specified by the VaR at time $t$. This is referred to as the hit sequence, namely:

$$
I_{t+1}(p) = \begin{cases} 
1 & \text{if } R_{t+1} > \text{VaR}_{t+1|t}^{DPOT}(p) \\
0 & \text{if } R_{t+1} \leq \text{VaR}_{t+1|t}^{DPOT}(p).
\end{cases} \quad (15)
$$

Christoffersen (1998) showed that evaluating interval forecasts can be reduced to examining whether the hit sequence satisfies the unconditional coverage (UC) and independence (IND) properties. In order to test the UC hypothesis, we apply the (Kupiec, 1995) test and to to test the IND hypothesis, we apply two tests. Along the same lines as Engle and Manganelli (2004), Berkowitz et al. (2009) consider the autoregression:

$$
I_t = \alpha + \beta_1 I_{t-1} + \beta_2 \text{VaR}_{t-1|t}^{DPOT}(p) + \varepsilon_t, \quad (16)
$$
and propose the logit model. We can test the \textit{IND} hypothesis with a likelihood ratio test of the null hypothesis $\beta_1 = \beta_2 = 0$, where the asymptotic distribution is chi-square with 2 degrees of freedom. This is the CAViaR independence test of Engle and Manganelli (CAViaR).

The second independence test was recently developed by Araújo Santos and Fraga Alves (2010), and is based on the following test statistic:

$$T_{N,[N/2]} = \log 2 \left( \frac{D_{N:N}^{-1}}{D_{[N/2]:N}} \right) - \log N$$

(17)

where $D_{N:N}$ and $D_{[N/2]:N}$ are the maximum and the median, respectively, of durations between consecutive violations and until the first violation. Under the null hypothesis, the asymptotic distribution is Gumbel. This new test is suitable for detecting clusters of violations, is based on an exact distribution (see Propositions 3.1 and 3.5 in Araújo Santos and Fraga Alves (2010)), is pivotal in the sense that is based on a distribution that does not depend on an unknown parameter, and outperforms, in terms of power, existing procedures in realistic settings, with few exceptions. We refer to this test as the \textit{MM} ratio test. The R code for implementing the test is available in Araújo Santos (2010).

4. Data

Compared with the empirical analysis in McAleer et al. (2010b), the updated data used for estimation and forecasting are the closing daily prices for Standard and Poor’s Composite 500 Index (S&P500), which were obtained from the Ecowin Financial Database, initially for the period 3 January 2000 to 25 March 2011. Although it is unlikely that an ADI’s typical market risk portfolio only tracks the S&P500 index, it is used as an illustration of the broad movements of profits and losses of the equity portfolios of many large ADIs. McAleer et al. (2011) estimate similar VaR models for the following stock indexes: French CAC 40 (CAC), German DAX 30 (DAX), US Dow Jones 30 (DJ), UK FTSE100 (FTSE), Hong Kong Hang Seng (HSI), Spanish IBEX (IBEX), Japanese Nikkei 225 (Nikkei), and Swiss SMI (SMI).
If \( P_t \) denotes the market price, the returns at time \( t \) \((R_t)\) are defined as:

\[
R_t = \log \left( \frac{P_t}{P_{t-1}} \right).
\]

(8)

[Insert Figures 1-2 and Table 2 here]

Figure 1 shows the S&P500 returns, for which the descriptive statistics are given in Table 2. Extremely high positive and negative returns are evident from September 2008 onward, and these continue well into 2009. The mean is close to zero, and the range is between -9.47% and +10.96%. The Jarque-Bera Lagrange multiplier test rejects the null hypothesis of normally distributed returns. The series display high kurtosis, as can be seen in the histogram. This would seem to indicate the existence of extreme observations, which is not surprising for daily financial returns data.

Several measures of volatility are available in the literature. In order to gain some intuition, we adopt the measure proposed in Franses and van Dijk (1999), where the true volatility of returns is defined as:

\[
V_t = \left( R_t - E(R_t | F_{t-1}) \right)^2,
\]

(9)

where \( F_{t-1} \) is the information set at time \( t-1 \).

Figure 2 shows the S&P500 volatility, as the square root of \( V_t \) in equation (9). The series exhibit clustering that should be captured by an appropriate time series model. The volatility of the series appears to be high during the early 2000s, followed by a quiet period from 2003 to the beginning of 2007. Volatility increases dramatically after August 2008, due in large part to the worsening global credit environment. This increase in volatility is even higher in October 2008. In less than four weeks in October 2008, the S&P500 index plummeted by 27.1%. In less than three weeks in November 2008, starting the morning after the US elections, the S&P500 index plunged a further 25.2%. Overall, from late August 2008, US stocks fell by an unbelievable 42.2% to reach a low on 20 November 2008.
An examination of daily movements in the S&P500 index from 2000 suggests that large changes by historical standards are 4% in either direction. From January 2000 to August 2008, there was a 0.31% chance of observing an increase of 4% or more in one day, and a 0.18% chance of seeing a reduction of 4% or more in one day. Therefore, 99.5% of movements in the S&P500 index during this period had daily swings of less than 4%. Prior to September 2008, the S&P500 index had only 7 days with massive 4% gains, but since September 2008, there have been a further 12 such days. On the downside, before the current stock market meltdown, the S&P500 index had only 4 days with huge 4% or more losses whereas, during the recent panic, there were a further 17 such days.

This comparison is between more than 99 months and less than 6 months. During the GFC the chances of 4% or more gain days increased 80 times, while the chances of 4% or more loss days increased 32 times. Such movements in the S&P500 index are truly exceptional.

5. Robust Forecasting of VaR and Evaluation Framework

As observed in McAleer et al. (2010a,b), the GFC has affected the best risk management strategies by changing the optimal model for minimizing daily capital charges. The objective here is to provide a robust risk management strategy, namely one that does not change over time, even in the context of a GFC. This robust risk management strategy also has to lead to daily capital charges that are not excessive, and violation frequencies that are compatible with the Basel II Accord. As stated previously, the calculation of daily capital charges, and the evaluation of a supremum to satisfy the Basel requirements, is based on numerical rather than statistical considerations.

The Basel II Accord does not stipulate that ADIs should restrict themselves to using only a single risk model. We propose a risk management strategy that consists of choosing a forecast from among different combinations of alternative risk models to forecast VaR. McAleer et al. (2010a) developed a risk management strategy that used combinations of several models for forecasting VaR. It was found that an aggressive risk management strategy (namely, choosing the supremum of VaR forecasts, or upperbound) yielded the lowest mean capital charges and largest number of violations. On the other hand, a conservative risk management strategy (namely, by choosing the
infinum, or lowerbound) had far fewer violations, and correspondingly higher mean daily capital charges.

In this paper, we forecast VaR using combinations of the forecasts of individual VaR models, namely the $r$th percentile of the VaR forecasts of a set of univariate conditional volatility models. Alternative single models with different error distributions, several combinations, and alternative methodologies are compared over three different time periods to investigate which, if any, of the risk management strategies may be robust.

We conduct an exercise to analyze the performance of existing VaR forecasting models, as permitted under the Basel II framework, when applied to the S&P500 index. Additionally, we analyze twelve new strategies based on combinations of the previous standard single-model forecasts of VaR, namely: lowerbound (0th percentile), upperbound (100th percentile), average, and nine additional strategies based on the 10th through to the 90th percentiles. We also consider the output of extreme value models such as DPOT(C=2/3), DPOT(c=3/4) and CEVT. It is intended to determine whether we can select a robust VaR forecast irrespective of the time period, and to provide reasonable daily capital charges and number of violation penalties.

[Insert Figure 3 here]

In Figure 3 we show the S&P500 returns together with VaR forecasts of a selection of forecasting models. The upper line is the S&P500 returns. The lower thick line is the median of the forecasts of the individual models. The thinner red line in between is the VaR forecast by the DPOT(C=2/3) model and the black line is the VaR forecast of Riskmetrics. It can be seen that, while the median and Riskmetrics are not far from each other most of the time, DPOT is more aggressive during the GFC while it is often times more conservative after the crisis.

[Insert Figure 4 here]

Figure 4 shows the daily capital charges corresponding to each of the above models, together with the S&P500 returns. The upper line is the S&P500 returns, while the thick line corresponds to the median, the thin red line corresponds to DPOT(c=2/3), and
the thin blue line corresponds to Riskmetrics. It can be seen first that the three strategies lead to different capital charges. Notice that DPOT(c=2/3) is more aggressive than the median before and during the crisis, while it is generally more conservative after the crisis. The lines are sufficiently distinct as to lead to significant differences in the performance of the different forecasting models.

5.1 Evaluating Crisis-Robust Risk Management Strategies

In Table 3, we compare the performance of the different VaR forecasting models using several economic and statistical criteria. The individual VaR models are Riskmetrics and GARCH, GJR, with, respectively normal, t and generalized normal errors. Additionally, we use three extreme value models: conditional EVT, DPOT(c=2/3), and DPOT(c=3/4). We also use three combination models, namely: infinum, supremum and the median.

[Insert Table 3 here]

We also evaluate the forecasting behaviour before, during and after the GFC. Before the GFC is from 2 January 2008 to 11 August 2008, during the GFC is from 12 August 2008 to 9 March 2009, and after the GFC is from 10 March 2009 to 16 March 2011. In each of the calculated measures, the numerical value is dominant relative to any statistical accuracy measures.

We evaluate the models according to the following criteria:

1. The percentage of time for which the model would keep the ADI in the red zone of the Basel II Accord (see Table 1).
2. The average daily capital charges incurred by the ADI using a given forecasting model (entries are percentages).
3. The Failure rate, which measures the percentage of violations incurred during the period.
4. The Kupiec independence test (entries are p-values).
5. The MM independence test (entries are p-values).
6. The CAViaR independence test (entries are p-values).
We exclude from consideration models that lead the ADIs to the red zone in at least one period, namely: Riskmetrics, GARCH-n, GJR-n, DPOT(c=2/3), and Supremum.

- The best model before the crisis is GARCH-gnd, with no days in red, and the lowest average daily capital charges of 9.32%, while the temporal independence of the violations is not rejected by any of the 2 tests.
- During the crisis, the best model is DPOT(c=3/4), with no days in red and the lowest average daily capital charges of 19.73%, while independence is not rejected by either of the 2 tests, although it has a high failure rate of 4.8%.
- After the crisis, the best model is GJR-gnd, with no days in red, minimum average daily capital charges of 10.47%, and independence not rejected by the 2 independence tests.

The lowest average daily capital charges across the whole sample used for comparison, of all the models corresponds to GJR-gnd, with a value of 12.87%, while the second lowest corresponds to the Median, with a value of 13.03%, as seen in the last column.

The median is, respectively, third, third and second across the three periods in terms of daily capital charges. No risk model is always found to be superior to its competitors, as there is no strategy that optimizes every evaluation statistic for the three sub-periods. Nonetheless, the 50th percentile strategy (namely, the median) is found to be robust, as it produces adequate VaR forecasts that exhibit stable results across different periods relative to the other risk models. In general, the median strategy provides a robust VaR forecast, regardless of whether there is a GFC.

6. Conclusion

In this paper we proposed robust risk forecasts that use combinations of several conditional volatility models for forecasting VaR. These include parametric as well as extreme value models. Different strategies for combining models were compared over three different time periods, using S&P500 to investigate whether we can determine a GFC-robust risk management strategy.
Backtesting provided evidence that a risk management strategy based on VaR forecast corresponding to the 50th percentile (median) of the VaR forecasts of a set of univariate conditional volatility models is robust in that it yields reasonable daily capital charges, numbers of violations that do not jeopardize institutions that might use it and, more importantly, is invariant before, during and after the 2008-09 GFC.

It is worth noting that, as in McAleer et al. (2010a), the VaR model that minimizes DCC before, during and after the GFC can, and does, change frequently. In our case, they were, respectively: GARCH-gnd, DPOT(c=3/4) and GJR-gnd. Although the median is not derived as necessarily the best model for minimizing DCC and the number of violation penalties, it is nevertheless a model that usefully balances daily capital charges and violation penalties in minimizing DCC.

The idea of combining different VaR forecasting models is entirely within the spirit of the Basel II Accord, although its use may require approval by the regulatory authorities, as for any forecasting model. This approach is not computationally demanding, even though several models have to be specified and estimated over time. Further research is needed to compute the standard errors of the forecasts of the combination models, including the median forecast, using numerical methods.
References


Ling, S. and McAleer, M. (2002b), Necessary and sufficient moment conditions for the GARCH(r,s) and asymmetric power GARCH(r, s) models, Econometric Theory, 18, pp. 722-729.


Figure 1. Daily Returns on S&P500 Index
13 January 2003 – 25 March 2011
Figure 2. Daily Volatility in S&P500 Returns
13 January 2003 – 25 March 2011
Figure 3. VaR for S&P500 Returns
13 January 2003 – 25 March 2011
Figure 4. Daily Capital Charges and S&P500 Returns
13 January 2003 – 25 March 2011
Table 1: Basel Accord Penalty Zones

<table>
<thead>
<tr>
<th>Zone</th>
<th>Number of Violations</th>
<th>$k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>0 to 4</td>
<td>0.00</td>
</tr>
<tr>
<td>Yellow</td>
<td>5</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.85</td>
</tr>
<tr>
<td>Red</td>
<td>10+</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Note: The number of violations is given for 250 business days. The penalty structure under the Basel II Accord is specified for the number of violations and not their magnitude, either individually or cumulatively.
Table 2. Descriptive Statistics for S&P500 Returns (%)
13 January 2000 – 25 March 2011

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
<td>RETURNS</td>
</tr>
<tr>
<td>Sample</td>
<td>13/01/2003 25/03/2011</td>
</tr>
<tr>
<td>Observations</td>
<td>2065</td>
</tr>
<tr>
<td>Mean</td>
<td>0.016926</td>
</tr>
<tr>
<td>Median</td>
<td>0.084217</td>
</tr>
<tr>
<td>Maximum</td>
<td>10.95720</td>
</tr>
<tr>
<td>Minimum</td>
<td>-9.469514</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>1.330519</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.263573</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>14.03100</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>10493.72</td>
</tr>
<tr>
<td>Probability</td>
<td>0.000000</td>
</tr>
</tbody>
</table>
Table 3. Comparing Alternative Models of Volatility

<table>
<thead>
<tr>
<th></th>
<th>% in red</th>
<th>AvDCC</th>
<th>% FailRa</th>
<th>Kupiec</th>
<th>MM</th>
<th>CAViaR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Before</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSKM</td>
<td>0.0</td>
<td>9.23</td>
<td>2.0</td>
<td>0.29</td>
<td>0.52</td>
<td>0.28</td>
</tr>
<tr>
<td>GARCH-n</td>
<td>0.0</td>
<td>8.52</td>
<td>3.9</td>
<td>0.01</td>
<td>0.30</td>
<td>0.16</td>
</tr>
<tr>
<td>GARCH-t</td>
<td>0.0</td>
<td>10.94</td>
<td>0.7</td>
<td>0.65</td>
<td>-</td>
<td>0.36</td>
</tr>
<tr>
<td>GARCH-gnd</td>
<td>0.0</td>
<td>9.32</td>
<td>2.6</td>
<td>0.09</td>
<td>0.44</td>
<td>0.25</td>
</tr>
<tr>
<td>GJR-n</td>
<td>0.0</td>
<td>8.57</td>
<td>3.3</td>
<td>0.03</td>
<td>0.55</td>
<td>0.45</td>
</tr>
<tr>
<td>GJR-t</td>
<td>0.0</td>
<td>10.75</td>
<td>0.0</td>
<td>0.08</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GJR-gnd</td>
<td>0.0</td>
<td>9.33</td>
<td>2.0</td>
<td>0.29</td>
<td>0.67</td>
<td>0.64</td>
</tr>
<tr>
<td>Cond_EVT</td>
<td>0.0</td>
<td>10.05</td>
<td>1.3</td>
<td>0.72</td>
<td>0.23</td>
<td>0.21</td>
</tr>
<tr>
<td>DPOT(c=2/3)</td>
<td>0.0</td>
<td>9.16</td>
<td>2.0</td>
<td>0.29</td>
<td>0.60</td>
<td>0.71</td>
</tr>
<tr>
<td>DPOT(c=3/4)</td>
<td>0.0</td>
<td>9.71</td>
<td>0.7</td>
<td>0.65</td>
<td>-</td>
<td>0.63</td>
</tr>
<tr>
<td>Inf.</td>
<td>0.0</td>
<td>11.84</td>
<td>0.0</td>
<td>0.08</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sup.</td>
<td>0.0</td>
<td>7.95</td>
<td>3.9</td>
<td>0.01</td>
<td>0.31</td>
<td>0.58</td>
</tr>
<tr>
<td>50th Per</td>
<td>0.0</td>
<td>9.38</td>
<td>3.3</td>
<td>0.09</td>
<td>0.45</td>
<td>0.44</td>
</tr>
<tr>
<td><strong>During</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSKM</td>
<td>0.0</td>
<td>24.68</td>
<td>4.1</td>
<td>0.00</td>
<td>0.45</td>
<td>0.02</td>
</tr>
<tr>
<td>GARCH-n</td>
<td>71.0</td>
<td>24.44</td>
<td>4.8</td>
<td>0.00</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>GARCH-t</td>
<td>0.0</td>
<td>26.11</td>
<td>1.4</td>
<td>0.66</td>
<td>0.59</td>
<td>0.01</td>
</tr>
<tr>
<td>GARCH-gnd</td>
<td>0.0</td>
<td>25.72</td>
<td>2.1</td>
<td>0.26</td>
<td>0.84</td>
<td>0.00</td>
</tr>
<tr>
<td>GJR-n</td>
<td>0.0</td>
<td>24.16</td>
<td>3.5</td>
<td>0.02</td>
<td>0.33</td>
<td>0.00</td>
</tr>
<tr>
<td>GJR-t</td>
<td>0.0</td>
<td>26.06</td>
<td>1.4</td>
<td>0.66</td>
<td>0.59</td>
<td>0.01</td>
</tr>
<tr>
<td>GJR-gnd</td>
<td>0.0</td>
<td>25.16</td>
<td>2.1</td>
<td>0.26</td>
<td>0.84</td>
<td>0.00</td>
</tr>
<tr>
<td>Cond_EVT</td>
<td>0.0</td>
<td>23.64</td>
<td>1.4</td>
<td>0.66</td>
<td>0.59</td>
<td>0.01</td>
</tr>
<tr>
<td>DPOT(c=2/3)</td>
<td>0.0</td>
<td>37.2</td>
<td>5.5</td>
<td>0.00</td>
<td>0.43</td>
<td>0.13</td>
</tr>
<tr>
<td>DPOT(c=3/4)</td>
<td>0.0</td>
<td>19.73</td>
<td>4.8</td>
<td>0.00</td>
<td>0.81</td>
<td>0.19</td>
</tr>
<tr>
<td>Inf.</td>
<td>0.0</td>
<td>27.32</td>
<td>1.4</td>
<td>0.66</td>
<td>0.59</td>
<td>0.30</td>
</tr>
<tr>
<td>Sup.</td>
<td>0.0</td>
<td>24.15</td>
<td>1.4</td>
<td>0.66</td>
<td>0.59</td>
<td>0.35</td>
</tr>
<tr>
<td>50th Per</td>
<td>0.0</td>
<td>82.1</td>
<td>18.50</td>
<td>8.3</td>
<td>0.00</td>
<td>0.61</td>
</tr>
<tr>
<td><strong>After</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSKM</td>
<td>21.7</td>
<td>11.34</td>
<td>2.5</td>
<td>0.00</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>GARCH-n</td>
<td>19.5</td>
<td>11.55</td>
<td>2.9</td>
<td>0.00</td>
<td>0.27</td>
<td>0.04</td>
</tr>
<tr>
<td>GARCH-t</td>
<td>0.0</td>
<td>13.53</td>
<td>0.2</td>
<td>0.02</td>
<td>-</td>
<td>0.18</td>
</tr>
<tr>
<td>GARCH-gnd</td>
<td>0.0</td>
<td>11.31</td>
<td>1.5</td>
<td>0.25</td>
<td>0.75</td>
<td>0.02</td>
</tr>
<tr>
<td>GJR-n</td>
<td>11.8</td>
<td>11.43</td>
<td>3.1</td>
<td>0.00</td>
<td>0.00</td>
<td>0.11</td>
</tr>
<tr>
<td>Cond_EVT</td>
<td>0.0</td>
<td>12.61</td>
<td>0.6</td>
<td>-</td>
<td>0.10</td>
<td>0.00</td>
</tr>
<tr>
<td>DPOT(c=2/3)</td>
<td>0.0</td>
<td>12.99</td>
<td>0.4</td>
<td>0.11</td>
<td>0.43</td>
<td>0.53</td>
</tr>
<tr>
<td>DPOT(c=3/4)</td>
<td>0.0</td>
<td>12.71</td>
<td>0.4</td>
<td>0.11</td>
<td>0.53</td>
<td>0.51</td>
</tr>
<tr>
<td>Inf.</td>
<td>0.0</td>
<td>14.37</td>
<td>0.0</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sup.</td>
<td>88.8</td>
<td>11.11</td>
<td>3.9</td>
<td>0.00</td>
<td>0.23</td>
<td>0.29</td>
</tr>
<tr>
<td>50th Per</td>
<td>0.0</td>
<td>8.00</td>
<td>1.4</td>
<td>0.44</td>
<td>0.14</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Notes: Percentage of Days in red zone (% red), Average of Daily Capital Charges (AvDCC), Failure Rate (FailRa), P-values of the Unconditional coverage Kupiec test (Kupiec), independence test of Araújo Santos and Fraga Alves (2010) (MM), and the CAViaR independence test of Engle and Manganelli (CAViaR). Average of Daily Capital Charges for the whole period I denoted as (AV_DCC_all).