The Model Confidence Set package for **R**

Mauro Bernardi
Sapienza University of Rome

Leopoldo Catania
University of Rome Tor Vergata

Abstract

This paper presents the R package `MCS` which implements the Model Confidence Set (MCS) procedure recently developed by Hansen, Lunde, and Nason (2011). The Hansen's procedure consists on a sequence of tests which permits to construct a set of "superior" models, where the null hypothesis of Equal Predictive Ability (EPA) is not rejected at a certain confidence level. The EPA statistic tests is calculated for an arbitrary loss function, meaning that we could test models on various aspects, for example punctual forecasts. The relevance of the package is shown using an example which aims at illustrating in details the use of the functions provided by the package. The example compares the ability of different models belonging to the ARCH family to predict large financial losses. We also discuss the implementation of the ARCH–type models and their maximum likelihood estimation using the popular R package `rugarch` developed by Ghalanos (2014).

Keywords: Hypothesis testing, Model Confidence Set, Value–at–Risk, VaR combination, ARCH–Models, R–CRAN.

1. Introduction

During last decades hundred of models have been developed, estimated and validated both from an empirical and theoretical point of view with the consequence that usually several competing models are available to the econometricians to address the same empirical problem. Just to confine our considerations within a given family, and without claiming to be complete, the Autoregressive Conditional Heteroskedastic (ARCH) models of Engle (1982) and Bollerslev (1986) have seen an exponentially increasing number of different specifications in the last few decades. Despite their popularity, they do not exhaust the set of models introduced for dynamic conditional volatility modelling which includes also the stochastic volatility models initially proposed by Taylor (1994) and extensively studied by Harvey and Shephard (1996) and Gallant, Hsieh, and Tauchen (1997) within the context of non–linear state space models. The family of dynamic conditional volatility models has been recently enlarged by the Dynamic Conditional Score models of Harvey (2013) and Creal, Koopman, and Lucas (2013) also known as Generalised Autoregressive Score models. The availability of such an enormous number of models raises the question of providing a statistical method or procedure that delivers the "best" models with respect to a given criterium. Furthermore, a model selection issue appears when the usual comparing procedures does not deliver an unique result. This could happen for example, when models are compared in terms of their predictive ability so that models that produce better forecasts are preferred. Unfortunately, sometimes it is not trivial to asses which model clearly outperforms each other. This problem

is relevant from an empirical point of view especially when the set of competing alternatives is large. As observed by Hansen and Lunde (2005) and Hansen *et al.* (2011), since different competing model are usually built to answer a specific econometric question, we not expect that a single model dominates all competitors either because they are statistically equivalent or because there is not enough information coming from the data to univocally discriminate the models. Recently, a lot of effort has been devoted to develop new testing procedures being able to deliver the "best fitting" models, see e.g. the Reality Check (RC) of White (2000), the Stepwise Multiple Testing procedure of Romano and Wolf (2005) and the Superior Predictive Ability (SPA) test of Hansen and Lunde (2005) and the Conditional Predictive Ability (CPA) test of Giacomini and White (2006).

Among those multiple–testing procedures, the Model Confidence Set procedure (MCS) of Hansen, Lunde, and Nason (2003), Hansen (2005) and Hansen *et al.* (2011) consists of a sequence of statistic tests which permits to construct a set of "superior" models, the "Superior Set Models" (SSM), where the null hypothesis of equal predictive ability (EPA) is not rejected at certain confidence level α. The EPA statistic tests is evaluated for an arbitrary loss function, which essentially means that it is possible to test models on various aspects depending on the chosen loss function. The possibility of user supplied loss functions provides enough flexibility to the procedure that can be used to test competing models with respect to several different aspects. The Model Confidence Set procedure starts from an initial set of m competing models M_0 and results in a (hopefully) smaller set of superior models (i.e. SSM) denoted by $\widehat{M}^*_{1-\alpha}$; of course the best scenario is when the final set consists of a single model. At each step of the iterative procedure, the EPA hypothesis is tested, and if the null hypothesis is accepted, then the procedure stops and the SSM is created, otherwise, the EPA should be tested again after the elimination of worst model.

The models belonging to the superior set delivered by the Hansen's procedure can then be used for different purposes. For example, they can be used to forecast future volatility level, to predict the future level of observations, conditional to past information, or to deliver future Value–at–Risk levels as argued by Bernardi, Catania, and Petrella (2014). Alternatively, those models can be combined together to obtain better forecast measures. Since the original work of Bates and Granger (1969), a lot of papers have argued that combining predictions from alternative models often improves upon forecasts based on a single "best" model. In an environment where observations are subject to structural breaks and models are subject to different levels of misspecification, a strategy that pools information coming from different models typically performs better than methods that try to select the best forecasting model. In the empirical part of the paper, where several alternative GARCH specifications are compared in terms of their ability to predict future large losses, we also consider the Dynamic Model Averaging technique proposed by Bernardi *et al.* (2014) in order to form forecast combinations of the VaR levels delivered by the SSM, conditional on model's past out–of–sample performance as in Samuels and Sekkel (2011) and Samuels and Sekkel (2013). For further information about the application of the model averaging methodology the reader is referred to Bernardi *et al.* (2014) where the methodology is extensively discussed.

The R (R Development Core Team 2013) package MCS here developed provides an integrated environment for the comparison of alternative models or model's specifications within the same family using the Model Confidence Set (MCS) procedure introduced by Hansen *et al.* (2011). We believe the main feature lies in the tools that the package provides for specifying different loss functions and the iterative model selection. Throughout the paper, we only assume the reader is familiar with the basic functions and concepts of the R programming

language.

The layout of the paper is as follows. In Section 2 we present the Hansen *et al.* (2011)'s Model Confidence Set procedure detailing the alternative specifications of the test statistics. Section 3 details about the ARCH specifications and their maximum likelihood estimation using the R package `rugarch` developed by Ghalanos (2014). In Section 4, hinging on the ARCH–type models described in Section 3, we discuss how the package is used to determine the optimal superior set of models (SSM) for the purposes of the user. Section 5 covers the empirical application which aims at illustrating how the procedure in practically implemented to compare several alternative specifications of ARCH–type models. Finally, Section 6 concludes the paper.

2. The Model Confidence Set procedure

The availability of several alternative model specifications being able to adequately describe the unobserved data generating process (DGP) opens the question of selecting the "best fitting model" according to a given optimality criterion. This paper deals with the MCS procedure developed by Hansen *et al.* (2011). The Hansen's procedure consists of a sequence of statistic tests which permits to construct a set of "superior" models, the "Superior Set Models" (SSM), where the null hypothesis of equal predictive ability (EPA) is not rejected at a certain confidence level. The EPA statistic tests is calculated for an arbitrary loss function that satisfies general weak stationarity conditions, which essentially means that we could test models on various aspects, e.g. punctual forecasts as in Hansen and Lunde (2005) or the in–sample goodness of fit as in Hansen *et al.* (2011). Formally, let Y_t the observation at time t and $\hat{Y}_{i,t}$ the output of model i at time t, the loss function $\ell_{i,t}$ associated to the i–th model is defined as

$$\ell_{i,t} = \ell\left(Y_t, \hat{Y}_{i,t}\right), \tag{1}$$

and measures the difference between the output $\hat{Y}_{i,t}$ and the *a posteriori* realisation Y_t. As an example of loss function, Bernardi *et al.* (2014) consider the asymmetric VaR loss function of González-Rivera, Lee, and Mishra (2004) to compare the ability of different GARCH specification to predict extreme loss. Let VaR_t^τ the τ–level predicted VaR at time t, given information up to time $t-1$, $\mathcal{F}_{t-1}$, the asymmetric VaR loss function is defined as

$$\ell\left(y_t, \text{VaR}_t^\tau\right) = \left(\tau - d_t^\tau\right)\left(y_t - \text{VaR}_t^\tau\right), \tag{2}$$

where $d_t^\tau = \mathbb{1}\left(y_t < \text{VaR}_t^\tau\right)$ is the τ–quantile loss function, and represents the natural candidate to backtest quantile–based risk measures since it penalises more heavily observations below the τ–th quantile level, i.e. $y_t < \text{VaR}_t^\tau$. More details about the loss functions specification can be found in Hansen and Lunde (2005) and in the following Section 4.2.

We now briefly describe how the MCS procedure is implemented. The procedure starts from an initial set of models $\hat{\text{M}}^0$ of dimension m encompassing all the model specifications described in Section 3, and delivers, for a given confidence level $1 - \alpha$, a smaller set, the superior set of models, SSM, $\hat{\text{M}}_{1-\alpha}^*$ of dimension $m^* \leq m$. The best scenario is when the final set consists of a single mode, i.e. $m^* = 1$. Formally, let $d_{ij,t}$ denotes the loss differential between models i and j:

$$d_{ij,t} = \ell_{i,t} - \ell_{j,t}, \quad i,j = 1,\ldots,m, \quad t = 1,\ldots,n, \tag{3}$$

and let

$$d_{i\cdot,t} = (m-1)^{-1} \sum_{j \in M} d_{ij,t} \quad i = 1, \ldots, m, \tag{4}$$

be the simple loss of model i relative to any other model j at time t. The EPA hypothesis for a given set of models M can be formulated in two alternative ways:

$$
\begin{aligned}
H_{0,M} &: \quad c_{ij} = 0, &&\text{for all}\quad i, j = 1, 2, \ldots, m \\
H_{A,M} &: \quad c_{ij} \neq 0, &&\text{for some}\quad i, j = 1, \ldots, m,
\end{aligned}
\tag{5}
$$

or

$$
\begin{aligned}
H_{0,M} &: \quad c_{i\cdot} = 0, &&\text{for all}\quad i = 1, 2, \ldots, m \\
H_{A,M} &: \quad c_{i\cdot} \neq 0, &&\text{for some}\quad i = 1, \ldots, m,
\end{aligned}
\tag{6}
$$

where $c_{ij} = \mathbb{E}(d_{ij})$ and $c_{i\cdot} = \mathbb{E}(d_{i\cdot})$ are assumed to be finite and not time dependent. According to Hansen *et al.* (2011), in order to test the two hypothesis above, the following two statistics are constructed:

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\widehat{\text{var}}(\bar{d}_{ij})}} \quad \text{and} \quad t_{i\cdot} = \frac{\bar{d}_{i,\cdot}}{\sqrt{\widehat{\text{var}}(\bar{d}_{i,\cdot})}} \quad \text{for} \quad i, j \in M, \tag{7}$$

where $\bar{d}_{i,\cdot} = (m-1)^{-1} \sum_{j \in M} \bar{d}_{ij}$ is the simple loss of the i–th model relative to the averages losses across models in the set M, and $\bar{d}_{ij} = m^{-1} \sum_{t=1}^{m} d_{ij,t}$ measures the relative sample loss between the i–th and j–th models, while $\widehat{\text{var}}(\bar{d}_{i,\cdot})$ and $\widehat{\text{var}}(\bar{d}_{ij})$ are bootstrapped estimates of $\text{var}(\bar{d}_{i,\cdot})$ and $\text{var}(\bar{d}_{ij})$, respectively. According to Hansen *et al.* (2011), to calculate the bootstrapped variances $\widehat{\text{var}}(\bar{d}_{i,\cdot})$, we perform a block–bootstrap procedure of 5000 resamples, where the block length p is the max number of significants parameters obtained by fitting an AR(p) process on all the d_{ij} terms. However, in the MCS package we allow the user to provide an arbitrary block length p. Details about the implemented bootstrap procedure can be found in White (2000), Kilian (1999), Clark and McCracken (2001), Hansen *et al.* (2003), Hansen and Lunde (2005), Hansen *et al.* (2011) and Bernardi *et al.* (2014). The first t–statistic t_{ij} is used in the well know test for comparing two forecasts; see e.g. Diebold and Mariano (2002) and West (1996), while the second one is used in Hansen *et al.* (2003); Hansen (2005); Hansen *et al.* (2011). As discussed in Hansen *et al.* (2011) the two EPA null hypothesis presented in equation (5) and (6) map naturally into the two test statistics:

$$T_{R,M} = \max_{i,j \in M} |t_{ij}| \quad \text{and} \quad T_{\max,M} = \max_{i \in M} t_{i\cdot}, \tag{8}$$

where t_{ij} and $t_{i\cdot}$ are defined in equation (7). The test statistics defined in equation (8) can be used in order to test the two hypothesis (5) and (6), respectively. Since the asymptotic distributions of the two tests statistic is nonstandard, the relevant distributions under the null hypothesis is estimated using a bootstrap procedure similar to that used to estimate $\text{var}(\bar{d}_{i,\cdot})$ and $\text{var}(\bar{d}_{ij})$, see for details White (2000); Hansen *et al.* (2003); Hansen (2005); Kilian (1999); Clark and McCracken (2001).

As said in the Introduction, the MCS procedure consists on a sequential testing procedure, which eliminates at each step the worst model, until the hypothesis of equal predictive ability

(EPA) is accepted for all the models belonging to the SSM. The choice of the worst model to be eliminated has been made using an elimination rule that is coherent with the statistic test defined in equation (7) which are

$$e_{\max,M} = \arg\max_{i \in M} \frac{\bar{d}_{i,\cdot}}{\widehat{\mathrm{var}}\left(\bar{d}_{i,\cdot}\right)}, \qquad e_{R,M} = \arg\max_{i} \left\{ \sup_{j \in M} \frac{\bar{d}_{ij}}{\sqrt{\widehat{\mathrm{var}}\left(\bar{d}_{ij}\right)}} \right\}, \qquad (9)$$

respectively.

Summarazing, the MCS procedure to obtain the SSM, consists of the following steps:

1. set $M = M_0$

2. test for EPA–hypothesis: if EPA is accepted terminate the algorithm and set $M^*_{1-\alpha} = M$, otherwise use the elimination rules defined in equations (9) to determine the worst model.

3. remove the worst model, and go to step 2.

Since the Hansen's procedure usually delivers a SSM $\hat{M}^*_{1-\alpha}$ which contains a large number of models, in the next sections, we also describe how to implement a procedure that combines the VaRs forecasts delivered by the MCS procedure.

3. Model specifications

In the empirical illustration we apply the MCS procedure detailed in the previous Section 2 to compare the VaR forecasts obtained by fitting a list of popular autoregressive conditional heteroskedastic models introduced by Engle (1982) and Bollerslev (1986). Here, we chose the ARCH–type models, because of their popularity, but the procedure could be implemented for any kind of model without any further complications. To account for the stylised facts about financial returns we consider several specifications differing for the conditional mean, the volatility dynamics and the distributions of the error term. Formally, let y_t the logarithmic return at time t, we consider the following general AR(1)–GARCH–in mean specification originally proposed by Engle, Lilien, and Robins (1987)

$$y_t = \mu + \lambda\sigma_t^2 + \phi y_{t-1} + \varepsilon_t, \qquad \varepsilon_t = \sigma_t\zeta_t, \qquad \zeta_t \sim \mathcal{D}(0,1)$$
$$\sigma_t^2 = h\left(\sigma_{t-1}, \varepsilon_{t-1}, \theta_\sigma \mid \mathcal{F}_{t-1}\right),$$

where $\mathcal{F}_t$ is the information set up to time t, ζ_t is a sequence of independently and identically distributed random variables with general standardized distribution $\mathcal{D}$, σ_t is the conditional standard deviation of y_t and ϕ is the autoregressive parameter assumed to be $|\phi| < 1$ to preserve stationarity. The risk–premium parameter λ is set equal to zero if the "in mean" specification is omitted, otherwise it is jointly estimated with the other parameters. Finally, the function $h(\cdot)$ refers to one of the ARCH–type dynamics reported below, where the vector θ_σ contains all the conditional variance dynamic parameters.

Among the most popular distributions that econometricians usually choose to model the

error term ε_t we consider: the Gaussian $\mathcal{N}(0,1)$, the Student–t $\mathcal{T}_\nu\left(0, \frac{\nu-2}{\nu}\right)$, with ν degrees of freedom, the Generalised Error distribution $\mathcal{GED}(0,1,\kappa)$, where κ is the shape parameter, and their asymmetric counterparts, the Skew–Normal, the Skew Student–t and the Skew–GED, obtained by applying the skewing mechanism of Fernández and Steel (1998). Furhermore, we consider the Johnson's reparametrised SU distribution of Rigby and Stasinopoulos (2005) and Generalized Hyperbolic introduced by Barndorff-Nielsen (1977).

ARCH–type models are flexible and powerful tools for conditional volatility modelling, because they are able to consider the volatility clustering phenomena and other established stylised facts such as excess of kurtosis and asymmetry of the unconditional returns. Models belonging to this class, was principally proposed in order to describe the time–varying nature of the conditional volatility that characterises financial assets. They have also become one of the most used tool for researchers and practitioners dealing with financial market exposure. The simplest conditional volatility dynamics we consider is the GARCH(p,q) specification introduced by Bollerslev (1986)

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha \varepsilon_{t-i-1}^2 + \sum_{j=1}^{q} \beta \sigma_{t-j-1}^2, \tag{10}$$

where $\omega > 0$ and $0 \leq \alpha_i < 1, \forall i = 1, 2, \ldots, p$ and $0 \leq \beta_j < 1, \forall j = 1, 2, \ldots, q$ with $P \equiv \sum_{i=1}^{p} \alpha + \sum_{j=1}^{q} \beta < 1$ to preserve weak ergodic stationarity of the conditional variance. Sometimes it is possible to observe high persistence in financial time series volatility, i.e. it is possible to observe series for which $P \approx 1$. To account for this scenario the IGARCH(p,q) specification, where the persistence parameters is imposed to be exactly 1, i.e $P = 1$, has been proposed by Engle and Bollerslev (1986). Despite their popularity, the GARCH and the IGARCH specifications are not able to account for returns exhibiting higher volatility after negative shocks than after positive ones as theorised by the "leverage effect" of Black (1976). Consequently, in the financial econometric literature several ARCH models have been proposed. The EGARCH(p,q) model of Nelson (1991), for example, assumes that the conditional volatility dynamics follows

$$\log\left(\sigma_t^2\right) = \omega + \sum_{i=1}^{p} \left[\alpha_i \zeta_{t-i} + \gamma_i \left(|\zeta_{t-i}| - \mathbb{E}|\zeta_{t-i}|\right)\right] + \sum_{j=1}^{q} \beta_j \log\left(\sigma_{t-j}^2\right), \tag{11}$$

where the asymmetric response is introduced through the γ_i parameters: for $\gamma_i < 0$ negative shocks will obviously have a bigger impact on future volatility than positive shocks of the same magnitude. Note that $\alpha_i, \beta_j, \gamma_i$ could assume any value and no positivity constraints are required. For the EGARCH(p,q) specification the persistence parameter P is equal to $P = \sum_{j=1}^{q} \beta_j$. Another widely used asymmetric GARCH model is the GJR–GARCH(p,q) specification of Glosten, Jagannathan, and Runkle (1993), which accounts for the leverage effect simply by including a dummy variable which discriminates positive and negative lagged shocks in the following way

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \left(\alpha_i + \gamma_i \mathbb{I}_{\{\varepsilon_{t-i}<0\}}\right) \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{12}$$

where $\mathbb{I}_{\{\varepsilon_{t-i}<0\}}$ assumes value one if $\varepsilon_{t-i} < 0$ for $i = 1, 2, \ldots, p$ and zero otherwise. Because of the presence of the indicator function, the persistence of the GJR–GARCH specification

crucially depends on the asymmetry of the conditional distribution used to model the error term ε_t

$$P = \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j + \sum_{i=1}^{p} \gamma_i \mathbb{P}\left(\varepsilon_{t-i} \leq 0\right), \tag{13}$$

where $\mathbb{P}\left(\varepsilon_{t-i} \leq 0\right)$ denotes the probability of observing negative shocks and $\omega > 0$, $\alpha_i \geq 0$, for $i = 1, 2, \ldots, p$, $\beta_j \geq 0$, for $j = 1, 2, \ldots, q$, and the additional constraint $\alpha_i + \gamma_i \geq 0$ for $i = 1, 2 \ldots, p$ is imposed to preserve the positiveness of the conditional variance. Finally, the Asymmetric–Power–ARCH(p,q) (APARCH, henceforth) of Ding, Granger, and Engle (1993) imposes the following dynamic to the conditional variance

$$\sigma_t^{(\delta)} = \omega + \sum_{i=1}^{p} \alpha_i \left(|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i}\right)^{\delta} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^{(\delta)}, \tag{14}$$

where $x^{(\delta)} = \frac{x^{\delta}-1}{\delta}$ is the Box–Cox transformation of Box and Cox (1964). The parameters restrictions to ensure the positiveness of the conditional variance are $\omega > 0$, $\delta \geq 0$, $0 \leq \gamma_i \leq 0$ for $i = 1, \ldots, p$ and the usual condition $\alpha_i \geq 0$, and $\beta_j \geq 0$, for $i, j = 1, 2, \ldots, \max\{p, q\}$. As for the GJR–GARCH specification the persistence strongly depends to the probability density function chosen for the innovation term ζ_t

$$P = \sum_{i=1}^{p} \alpha_i \kappa_i + \sum_{j=1}^{q} \beta_j, \tag{15}$$

where $\kappa_i = \mathbb{E}\left[||\zeta| - \gamma_i \zeta|^{\delta}\right]$, for $i = 1, \ldots, q$. The APARCH specification results in a very flexible model that nests several of the most popular univariate ARCH parameterisations, such as

- the GARCH(p,q) of Bollerslev (1986) for $\delta = 0$ and $\gamma_i = 0$, for $i = 1, 2, \ldots, p$;

- the Absolute–Value–GARCH (AVARCH, henceforth) specification for $\delta = 1$ and $\gamma_i = 0$ for $i = 1, 2, \ldots, p$, proposed by Taylor (1986) and Schwert (1990) to mitigates the influence of large, in an absolute sense, shocks with respect to the traditional GARCH specification;

- the GJR–GARCH model of Glosten *et al.* (1993) for $\delta = 2$ and $0 \leq \gamma_i \leq 1$ for $i = 1, 2, \ldots, p$;

- the Threshold GARCH (TGARCH, henceforth) of Zakoian (1994) for $\delta = 1$, which allows different reactions of the volatility to different signs of the lagged errors;

- the Nonlinear GARCH (NGARCH, henceforth) of Higgins and Bera (1992) for $\gamma_i = 0$ for $i = 1, 2, \ldots, p$ and $\beta_j = 0$ for $j = 1, 2, \ldots, q$.

Another interesting specification is the Component–GARCH(p,q) (CGARCH, henceforth) of Engle and Lee (1993) which decomposes the conditional variance into a permanent and transitory component in a straightforward way

$$\sigma_t^2 = \xi_t + \sum_{i=1}^{p} \alpha_i \left(\epsilon_{t-i}^2 - \xi_{t-i}\right) + \sum_{j=1}^{q} \beta_j \left(\sigma_{t-j}^2 - \xi_{t-j}\right)$$

$$\xi_t = \omega + \rho \xi_{t-1} + \eta \left(\epsilon_{t-1}^2 - \sigma_{t-1}^2\right), \tag{16}$$

where to ensure the stationarity of the process at the usual $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$ condition the $\rho < 1$ constrain must be added. Further parameters restrictions for the positiveness of the conditional variance are given in Engle and Lee (1993). This solution is usually employed because it permits to investigate the long and short–run movements of volatility. The considered conditional volatility models are a minimal part of the huge number of specifications available in the financial econometric literature. We chose these models because of their heterogeneity, since each of them focuses on a different kind of stylised fact. Moreover, even if they could seem very similar, the way in which they account for the stylised fact changes. For a very extensive and up to date survey on GARCH models we will refer the reader to the works of Bollerslev (2008), Teräsvirta (2009), Bauwens, Laurent, and Rombouts (2006), Silvennoinen and Teräsvirta (2009) and the recent book of Francq and Zakoian (2011). To estimate model parameters we consider the maximum likelihood approach, see e.g. Francq and Zakoian (2011).

3.1. GARCH model estimation and forecast in R

Within the R environment a wide range of statistical packages are available in order to deal with GARCH models estimation and forecast such as the `fGarch` package of Wuertz, with contribution from Michal Miklovic, Boudt, Chausse, and others (2013). Here, in order to estimate the competing GARCH models we use some functions belonging to the library *rugarch* developed by Ghalanos (2014). Of course, the following treatment is only useful for illustrative purposes on the use of the `MCS` package, and readers are free to choose a different package to estimate models as well as their handwritten functions. Before starting the MCS procedure, it is necessary to define the set of competing GARCH models M_0. This can be done using the the `ugarchspec()` function that permits to specify a variety of GARCH models such as those described in the previous Section 3. For example, the following portion of code

```
R> library(rugarch)
R> spec <- ugarchspec(mean.model = list(armaOrder = c(0, 0)),
                  variance.model = list(model = "sGARCH",
                  garchOrder = c(1, 1))
```

creates an `uGARCHspec` object "`spec`" which defines a GARCH(1,1) model with gaussian innovations (the chosen conditional distribution can be changed by using the *distribution.model* argument). The object `spec` can be subsequently used into the `ugarchfit()` function in order to estimate the model on a given time series. For example

```
R> library(MCS)
R> data(STOXXIndexesRet)
R> ret <- STOXXIndexesRet[,"SXA1E"]
R> fit <- ugarchfit(spec = spec, data = ret)
```

creates an `uGARCHfit` object "`fit`" containing parameter estimates of the STOXX North America 600 index ("SXA1E") as well as several additional informations such as the Information Criteria, tests on standardised residuals, and several ARCH LM tests. Prediction using GARCH models can be easily performed using the `ugarchforecast()` function which takes as an argument the output of the fitting procedure `uGARCHfit` or alternatively the `uGARCHspec` object. The one step ahead forecasts can be easily obtained using the following routine

```
R> OneStepForc <- ugarchforecast(fitORspec = fit, n.ahead = 1)
```

which reports an `uGARCHforecast`, object "`OneStepForc`". The `rugarch` package also includes the `ugarchroll()` function which permits to construct a series of one step ahead rolling forecasts allowing also the user to define a "refit window" and the length of the forecast series. For example, a rolling forecast series of length 2000, using a refit window of 5 observations can be computed as follow

```
R> roll <- ugarchroll(spec = spec, data = ret, forecast.length = 2000,
                      refit.every = 5)
```

which reports an `uGARCHroll` object `roll`. Finally, a variety of methods are present in the *rugarch* package in order to deal with "`uGARCHspec`" , "`uGARCHfit`" , "`uGARCHforecast`" and "`uGARCHroll`" objects. For more informations see Ghalanos (2014) or consult the help() in `R`.

4. Using the package

As described in Section 2 the MCS procedure is used to compare different models under an user defined loss function. The loss function measures the "performance" of the competing models at a each time point t in the evaluating period. Suppose now to have m competing models and an evaluating period of length n, then using the defined loss function it is possible to construct a loss matrix of dimension $(m \times n)$. Then the `MCSprocedure()` function can be used to construct the set of superior models outlined in Section 2.

4.1. Comparing GARCH models using MCS

The MCS procedure can be used to compare models under various aspects, for example with respect to their ability to predict the future volatility, or the future returns conditional to actual and past information. Suppose instead we are interested to compare the VaR forecasts delivered by different GARCH models. As previously explained, the `ugarchroll()` function can be used to obtain a one step ahead rolling forecast series of a specified GARCH model. Furthermore, the `as.data.frame()` method permits to extract from the `uGARCHroll` object the VaR forecasts series at both the 1% and 5% confidence levels. For example, if we want to compare five different GARCH specifications, such as

- the GARCH(1,1) of Bollerslev (1986), "`sGARCH`" in the `rugarch` package;

- EGARCH(1,1) of Nelson (1991), "`eGARCH`" in the `rugarch` package;

- GJRGARCH(1,1) of Glosten *et al.* (1993), "`gjrGARCH`" in the `rugarch` package;

- APARCH(1,1) of Ding *et al.* (1993), "`apARCH`" in the `rugarch` package;

- CGARCH(1,1) of Engle and Lee (1993), "`csGARCH`" in the `rugarch` package;

we can simply define the five GARCH specifications combined with the distributions presented in the previous section for the innovation terms using the `ugarchspec()` function as follows

```
R> models <- c("sGARCH", "eGARCH", "gjrGARCH", "apARCH", "csGARCH")
R> distributions <- c("norm", "std", "ged", "snorm", "sstd", "sged", "jsu", "ghyp")
R> spec.comp <- list()
R> for( m in models ) {
     for( d in distributions ) {
   spec.comp[[paste( m, d, sep = "-" )]] <-
       ugarchspec(mean.model = list(armaOrder = c(0, 0)),
       variance.model = list(model = m, garchOrder = c(1, 1)),
       distributtion.model=d)
                             }
                   }
R> specifications <- names( spec.comp )
```

In this way we have defined a list with the 40 combinations of GARCH specifications and innovation term distributions, and we can perform the rolling forecast of length 2000 using a "refit window" of 200 observations

```
R> roll.comp <- list()
R> for( s in specifications ){
   roll.comp[[s]] <-  ugarchroll(spec = spec.comp[[s]], data = ret,
                     forecast.length = 2000, refit.every = 200)
                   }
```

and, using the `as.data.frame()` method, extract the VaR forecast series at the confidence level $\tau = 1\%$

```
R> VaR.comp=list()
R> for( s in specifications ) {
   VaR.comp[[s]] <-  as.data.frame(roll.comp[[s]], which = "VaR")[, 1]
   }
```

Now it is possible to calculate the loss associated to each model at each time. Here, we consider the asymmetric VaR loss function of González-Rivera *et al.* (2004) and considered also by Bernardi *et al.* (2014), which is implemented in the MCS package through the `LossVaR()` function, more details of the available loss functions are reported in Section 4.2

```
R> Loss <- do.call(cbind,lapply(specifications,
                 function(s) LossVaR(tau=0.01, realized=tail(ret, 2000)/100,
                 evaluated=VaR.comp[[s]]/100)))
R> colnames(Loss) <- specifications
```

The object "Loss" is a matrix of dimension (2000×40) that contains the VaR losses associated to each of the different GARCH specifications we considered in this example. The object "Loss" is also included in the MCS packages and can be easily loaded using `data(Loss)`, since it represents the main input of the `MCSprocedure()` function described in Section 4.3. Nevertheless, in the next section we describe some alternative loss functions which are particularly suitable for volatility forecast assessment as well as to forecast future observations conditional to past information.

4.2. Loss functions

As previously discussed the MCS procedure is able to discriminate models under a user defined loss function. The choice of the loss function is somewhat arbitrary, and crucially depends on the nature of the competing models and the scope of their usage. For more considerations about the choice of the loss function for model comparison purposes we refer to Hansen and Lunde (2005), Bollerslev, Engle, and Nelson (1994), Diebold and Lopez (1996) and Lopez (2001). In what follows, we report the loss functions available within the MCS package. However, since the MCSprocedure() function described in the previous section accepts a pre–defined loss matrix, named "Loss", the user is free to define and use its own loss function. Three different loss functions are freely available within the MCS package:

1. the LossVaR() that can be used to check the performances associated to VaR (or more generally quantile) forecasts;

2. the LossVol() for volatility forecasts assessment;

3. the LossLevel() that can be used instead for level forecasts, as the punctual mean forecast of a regression model.

These loss functions accept three common arguments. The first two arguments are realized that consists of a vector of realised observations (i.e. the ones that a model hopes to accurately forecast or describe), and evaluated which is a vector or a matrix of models output. Note that we decided to call the second argument of those functions "evaluated" instead of "forecasted" since the MCS procedure is more general than simply a procedure suited for forecasts evaluation. In fact, as reported by Hansen *et al.* (2011) the MCS procedure also adapts to in sample studies. The third argument which instead is function dependent. The available choices and other function specific arguments are reported below.

- For LossVaR() only which = "asymmetricLoss" is available. This coincides with the asymmetric VaR loss function of González-Rivera *et al.* (2004) which is particularly suited to assess quantile risk measures, such as the VaR, since it penalises more heavily observations below the τ–th quantile level, i.e. $y_t < \text{VaR}_t^\tau$. The asymmetric loss function is defined as

$$\ell\left(y_t, \text{VaR}_t^\tau\right) = \left(\tau - d_t^\tau\right)\left(y_t - \text{VaR}_t^\tau\right), \tag{17}$$

where $d_t^\tau = \mathbb{1}\left(y_t < \text{VaR}_t^\tau\right)$ is the τ–quantile loss function. Further arguments are tau, which represents the VaR confidence level and type with possible choices "normal" and "differentiable". The type argument permits to discriminate between the normal and the differentiable versions of the loss function: "normal" permits the specification of the loss function of González-Rivera *et al.* (2004) defined in equation (17) while "differentiable" considers the following loss function

$$\ell\left(r_t, \text{VaR}_t^\tau\right) = \left(\tau - m_\delta\left(r_t, \text{VaR}_t^\tau\right)\right)\left(r_t - \text{VaR}_t^\tau\right), \tag{18}$$

where $m_\delta\left(a, b\right) = \left[1 + \exp\left\{\delta\left(a - b\right)\right\}\right]^{-1}$. Note that the δ parameter, controlling the function smoothness, can be chosen by the delta argument in the LossVaR() function and it is set equal to 25 by default.

- For `LossVol()`, the six loss functions reported in Hansen and Lunde (2005) are implemented. Note that for this kind of loss functions the `realized` and the `evaluated` arguments should be some realised volatility measures $\tilde{\sigma}_{t+1}$ and the punctual volatility forecasts $\hat{\sigma}_{t+1}$. In this context, we use the term volatility as for the standard deviation σ. The implemented loss functions are:

  1. $\text{SE}_{1,t+1} = (\tilde{\sigma}_{t+1} - \hat{\sigma}_{t+1})^2$, by setting `which = "SE1"`,
  2. $\text{SE}_{2,t+1} = (\tilde{\sigma}_{t+1}^2 - \hat{\sigma}_{t+1}^2)^2$, by setting `which = "SE2"`,
  3. $\text{QLIKE}_{t+1} = \log(\hat{\sigma}_{t+1}^2) + \tilde{\sigma}_{t+1}^2 \hat{\sigma}_{t+1}^{-2}$, by setting `which = "QLIKE"`,
  4. $\text{R}^2\text{LOG}_{t+1} = \left[\log(\tilde{\sigma}_{t+1}^2 \hat{\sigma}_{t+1}^{-2})\right]^2$, by setting `which = "R2LOG"`,
  5. $\text{AE}_{1,t+1} = |\tilde{\sigma}_{t+1} - \hat{\sigma}_{t+1}|$, by setting `which = "AE1"`,
  6. $\text{AE}_{2,t+1} = |\tilde{\sigma}_{t+1}^2 - \hat{\sigma}_{t+1}^2|$, by setting `which = "AE2"`.

- For `LossLevel()`, the `which` argument accepts values: `"SE"` and `"AE"` that coincide with the squared error and the absolute error.

4.3. Constructing the Superior Set of Models

The function `MCSprocedure()` returns a S4 object of the class "SSM", which has several arguments we now briefly describe here. The main inputs of the function `MCSprocedure()` are

- `Loss`, which must be matrix or something coercible to that (using the as.matrix() function) which contains the loss series for each model to be compared.

- `alpha`, that must be a positive scalar in $(0, 1)$ indicating the confidence level of the MCS tests.

- `B`, which is an integer indicating the number of bootstrapped samples used to construct the statistic test.

- `cluster`, that coincides with a cluster object created by calling makeCluster from the parallel package. By default this is set to NULL but if an appropriate cluster object is submitted, then this will be used for parallel processing.

- `statistic`, which is the statistic that should be used to test the EPA at each step of the iteration. Possible choices are "Tmax" and "TR", which coincide with $\text{T}_{\text{max},M}$ and $\text{T}_{\text{R,M}}$ of Section 2 respectively.

For sample purposes, in the MCS package the "Loss" dataset is included. This coincide with the `Loss` matrix obtained in the previous Section and can be loaded using `data(Loss)`. The construction of the Superior Set of Models can be easily done using the following portion of code:

```
R> library(MCS)
R> data(Loss)
R> SSM <- MCSprocedure(Loss = Loss, alpha = 0.2, B = 5000, statistic = "Tmax")
```

```
R> SSM
```

-			Superior Set of Models			-

	Rank_M	v_M	MCS_M	Rank_R	v_R	MCS_R	Loss
sGARCH-ged	27	0.996797601	0.4486	27	1.46532329	0.2352	0.0003986329
sGARCH-snorm	25	0.896954417	0.5920	26	1.38772567	0.2998	0.0003982803
sGARCH-sstd	26	0.938712179	0.5248	25	1.37978372	0.3166	0.0003977886
sGARCH-sged	20	0.521486029	0.9732	20	1.20336841	0.5346	0.0003956815
sGARCH-jsu	23	0.820277282	0.7052	24	1.35541006	0.3166	0.0003971334
sGARCH-ghyp	22	0.685040892	0.8628	22	1.27425551	0.4478	0.0003964821
eGARCH-ged	30	1.136085890	0.2480	31	1.55570183	0.1574	0.0003994099
eGARCH-sstd	17	-0.004284375	1.0000	17	0.99146387	0.7916	0.0003933537
eGARCH-sged	12	-0.538418981	1.0000	12	0.68337534	0.9802	0.0003910679
eGARCH-jsu	15	-0.132370421	1.0000	15	0.92181102	0.8454	0.0003928127
eGARCH-ghyp	16	-0.107676711	1.0000	16	0.94757871	0.8260	0.0003929190
gjrGARCH-norm	18	0.213742841	1.0000	18	1.10231033	0.6710	0.0003943638
gjrGARCH-std	13	-0.427196927	1.0000	13	0.75538350	0.9588	0.0003916026
gjrGARCH-ged	6	-0.954946367	1.0000	8	0.37895234	1.0000	0.0003891435
gjrGARCH-snorm	1	-1.295629237	1.0000	2	0.04458467	1.0000	0.0003870702
gjrGARCH-sstd	5	-0.962068788	1.0000	5	0.33202562	1.0000	0.0003887004
gjrGARCH-sged	2	-1.252823785	1.0000	1	-0.04527337	1.0000	0.0003867926
gjrGARCH-jsu	7	-0.942033667	1.0000	6	0.33760463	1.0000	0.0003887359
gjrGARCH-ghyp	8	-0.929362792	1.0000	4	0.32763998	1.0000	0.0003886976
apARCH-norm	21	0.535927772	0.9668	21	1.25059782	0.4788	0.0003959991
apARCH-std	19	0.367561226	0.9980	19	1.19805785	0.5490	0.0003949103
apARCH-ged	14	-0.317682324	1.0000	14	0.80503488	0.9304	0.0003920526
apARCH-snorm	4	-0.977717453	1.0000	7	0.33884203	1.0000	0.0003889586
apARCH-sstd	9	-0.628176498	1.0000	9	0.63183157	0.9904	0.0003905724
apARCH-sged	3	-1.022889502	1.0000	3	0.29488276	1.0000	0.0003884919
apARCH-jsu	10	-0.588780764	1.0000	10	0.63897617	0.9892	0.0003906327
apARCH-ghyp	11	-0.582730804	1.0000	11	0.64705812	0.9878	0.0003906581
csGARCH-sstd	31	1.165077099	0.2170	30	1.53265345	0.1782	0.0003992507
csGARCH-sged	24	0.839136228	0.6792	23	1.34145394	0.3166	0.0003972841
csGARCH-jsu	29	1.063487090	0.3454	28	1.47330668	0.2280	0.0003985447
csGARCH-ghyp	28	1.046751336	0.3706	29	1.47854362	0.2244	0.0003984870

```
Details
Number of eliminated models : 9
Statistic : Tmax
Elapsed Time : Time difference of 6.771243 mins
```

5. Application

For the empirical study, a panel of four major worldwide stock markets indices is considered. The four daily stock price indices includes the Asia/Pacific 600 (SXP1E), the North America 600 (SXA1E) and the Europe 600 (SXXP) as well as the Global 1800 (SXW1E). The data are freely available and can be download from the STOXX website http://www.stoxx.com/indices/types/benchmark.html. The data were obtained over a 23–years time period, from 31 December 1991 to 24 July 2014, for a total of 5874 observations. For each market, the returns are calculated as the logarithmic difference of the daily price index and multiplied by 100

$$y_t = (\log(p_t) - \log(p_{t-1})) \times 100,$$

where p_t is the closing index value on day t. To examine the performance of the models to predict extreme VaR levels, the complete dataset of daily returns is divided into two samples: an in–sample period from 1 January 1992 to 06 October 2006, for a total of 3814 observations, and a forecast or validation period, containing the remaining 2000 observations: from 09 October 2006 to 24 July 2014. A rolling window approach is used to produce 1–day ahead forecasts of the 1% VaR thresholds $\text{VaR}_{t+1}^{0.01}$, for $t = 1, 2, \ldots, 1999$ in the forecast samples. To apply the MCS procedure we consider 160 ARCH–type specifications obtained by combining the models reported in the previous Section 3 estimated on each of the four international indexes. More precisely, the 160 model specifications have been obtained by considering all the possible combinations of the 10 ARCH dynamics, the 8 conditional distributions and the *in–mean/non in–mean* options as outlined in Section 3. Estimated coefficients for each fitted model are not reported but they are available upon request to the second author. VaR estimates are performed by inverting the conditional cumulative density function of the corresponding estimated model. Then, the MCS procedure of Hansen *et al.* (2011) described in the previous Section 2 is applied to obtain the set of models with superior predictive ability in term of the VaR forecast at the 1% confidence level.

Table 1 reports the compositions of the Superior Set of Models discriminating by model, distribution and *in–mean* options. The different entries in each column represent the number of models that belong to the SSM at the end of the MCS procedure discriminated by model, distribution and *in–mean* options. From Table 1 we can observe that for the SXA1E and SXP1E indexes the SSM is quite homogeneous with respect to the dynamics and innovation assumptions. In these case the one step ahead 1% VaR forecasting performance of the competing models are quite similar, suggesting that for those series the use of complicated nonlinear models are not entirely justified. Moreover, it is interestingly to note that the EGARCH specification is the model most frequently eliminated, probably because the logarithmic specification of the conditional volatility is too much sensitive to the previous volatility changes. Concerning the distribution specifications, we observe that the MCS procedure confirms the common finding that the Gaussian distribution poorly describes the financial time series behaviour. Conversely, looking at the third and the fourth columns, it is possible to asses that for the SXW1E and SXXP indexes the MCS procedure reports more discriminant results. For the SXW1E index, all the specifications that include the GARCH, IGARCH, C–GARCH and N–GARCH dynamics were eliminated while also the GJR–GARCH dynamic is not present in the SSM for the SXXP index. The exclusion of those models suggests that for the SXW1E and SXXP series, more complicated dynamics are necessary to describe the shape of the conditional returns density function.

In order to test the benefit of using the MCS procedure technique we report a simple com-

	Asset			
	SXA1E	SXP1E	SXW1E	SXXP
	Models			
eGARCH	7	8	7	5
sGARCH	16	16	0	0
gjrGARCH	16	16	16	0
iGARCH	12	9	0	0
apARCH	16	16	16	10
csGARCH	12	16	0	0
TGARCH	13	13	14	14
AVGARCH	15	15	7	11
NGARCH	15	16	0	0
NAGARCH	16	16	13	13
	Distributions			
norm	12	19	5	0
snorm	16	18	9	8
std	16	18	10	5
sstd	19	17	11	9
ged	18	18	8	5
sged	19	18	11	9
ghyp	19	16	9	9
jsu	19	17	10	8
	In–Mean Specification			
Not in mean	72	76	43	29
in mean	66	65	30	24
Total Number	138	141	73	53

Table 1: Composition of remaining models in the Superior Set for each index.

parison study using two different VaR forecasts. The first VaR forecasts series (VaR_{avg}) is performed using the simple average VaR across all the 160 available models, while the second (VaR_{Dyn}) is performed using the dynamic VaR combination technique proposed by Bernardi *et al.* (2014) which average across the models belonging to the SSM. Table 2 reports three VaR backtesting measures. The first is the Actual over Expected ratio AE, defined as the ratio between the realised VaR exceedances over a given time horizon and their "a priori" expected values; VaR forecasts series for which the AE ratio is closer to the unity are preferred. The second and the third backtesting measures are the mean and maximum Absolute Deviation (ADmean and ADmax) of VaR violating returns described in McAleer and da Veiga (2008). The AD in general provides a measure of the expected loss given a VaR violation; of course models with lower mean and/or maximum ADs are preferred. As showed in Table 2 the $VaR_{dynamic}$ series always report lower ADmean and ADmax compared with the VaR_{avg}. Moreover, in two cases (for the SXW1E and the SXXP indices) also the AE ratio is improved.

6. Conclusion

	VaR$_{\text{Dyn}}$			VaR$_{\text{Avg}}$		
Asset	AE	ADmean	ADmax	AE	ADmean	ADmax
SXA1E	1.35	0.633	2.658	1.35	0.653	2.693
SXP1E	1.00	0.973	3.630	1.00	0.975	3.638
SXW1E	1.20	0.406	1.936	1.20	0.466	1.954
SXXP	1.25	0.533	2.201	1.35	0.568	2.616

Table 2: VaR backtesting measures of the dynamic VaR combination VaR$_{\text{Dyn}}$ and the static average VaR$_{\text{Avg}}$.

In the previous Sections we have illustrated the main features of the R package MCS which implements the Model Confidence Set procedure introduced by Hansen *et al.* (2011). The technique proposed by Hansen *et al.* (2011) is especially useful when more models are available and it is not obvious which one is the best. The MCS sequence of tests delivers the Superior Set of Models (SSM) having Equal Predictive Ability (EPA) in terms of an user supplied loss function discriminating models with respect to desired model characteristics such as, for example, forecasts performances. The MCS package is very flexible in the types of model and loss function that can be specified by the researcher. This freedom allows the user to concentrate on substantive issues, such as the construction of the initial set of model's specifications M^0, without being limited by the constraints imposed by the software. An empirical example shows the relevance of the package by illustrating in details the use of the provided functions. In particular, the example compares the ability of different models belonging to the ARCH family to predict large financial losses and discuss the ARCH–type models and their maximum likelihood estimation using the popular R package rugarch developed by Ghalanos (2014). The Model Confidence Set procedure is firstly performed in order to reduce the number of models, and then we show that accounting for the VaR dynamic model averaging technique of Bernardi *et al.* (2014) improves the VaR forecast performance.

Acknowledgments

References

Barndorff-Nielsen O (1977). "Exponentially decreasing distributions for the logarithm of particle size." *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, **353**(1674), 401–419.

Bates JM, Granger CW (1969). "The combination of forecasts." *OR*, pp. 451–468.

Bauwens L, Laurent S, Rombouts JVK (2006). "Multivariate GARCH models: a survey." *Journal of Applied Econometrics*, **21**(1), 79–109. ISSN 1099-1255. doi:10.1002/jae.842. URL http://dx.doi.org/10.1002/jae.842.

Bernardi M, Catania L, Petrella L (2014). "Are News Important to Predict Large Losses ?" *Working Paper, Arxiv Preprint*. 1306.2834.

Black F (1976). "Studies of Stock Price Volatility Changes." *In Proceedings of the 1976 American Statistical Association, Business and Economical Statistics Section, Alexandria, VA: American Statistical Association*, pp. 177–181.

Bollerslev T (1986). "Generalized autoregressive conditional heteroskedasticity." *Journal of Econometrics*, **31**(3), 307 – 327. ISSN 0304-4076. doi:http://dx.doi.org/10.1016/0304-4076(86)90063-1. URL http://www.sciencedirect.com/science/article/pii/0304407686900631.

Bollerslev T (2008). "Glossary to ARCH (GARCH)." *CREATES Research Papers 2008-49*, School of Economics and Management, University of Aarhus. URL http://ideas.repec.org/p/aah/create/2008-49.html.

Bollerslev T, Engle RF, Nelson DB (1994). "ARCH models." *Handbook of econometrics*, **4**, 2959–3038.

Box GE, Cox DR (1964). "An analysis of transformations." *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 211–252.

Clark TE, McCracken MW (2001). "Tests of equal forecast accuracy and encompassing for nested models." *Journal of econometrics*, **105**(1), 85–110.

Creal D, Koopman SJ, Lucas A (2013). "Generalized autoregressive score models with applications." *Journal of Applied Econometrics*, **28**(5), 777–795.

Diebold FX, Lopez JA (1996). "Forecast evaluation and combination."

Diebold FX, Mariano RS (2002). "Comparing predictive accuracy." *Journal of Business & economic statistics*, **20**(1).

Ding Z, Granger CW, Engle RF (1993). "A long memory property of stock market returns and a new model." *Journal of Empirical Finance*, **1**(1), 83 – 106. ISSN 0927-5398. doi:http://dx.doi.org/10.1016/0927-5398(93)90006-D. URL http://www.sciencedirect.com/science/article/pii/092753989390006D.

Engle RF (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." *Econometrica*, **50**(4), 987–1007. URL http://www.jstor.org/stable/1912773.

Engle RF, Bollerslev T (1986). "Modelling the persistence of conditional variances." *Econometric Reviews*, **5**(1), 1–50. doi:10.1080/07474938608800095. http://dx.doi.org/10.1080/07474938608800095, URL http://dx.doi.org/10.1080/07474938608800095.

Engle RF, Lee GG (1993). "A permanent and transitory component model of stock return volatility." *University of California at San Diego, Economics Working Paper Series.*

Engle RF, Lilien DM, Robins RP (1987). "Estimating Time Varying Risk Premia in the Term Structure: The Arch-M Model." *Econometrica*, **55**(2), 391 – 407. URL http://www.jstor.org/stable/1913242.

Fernández C, Steel MF (1998). "On Bayesian modeling of fat tails and skewness." *Journal of the American Statistical Association*, **93**(441), 359–371.

Francq C, Zakoian JM (2011). *GARCH models: structure, statistical inference and financial applications.* John Wiley & Sons.

Gallant A, Hsieh D, Tauchen G (1997). "Estimation of stochastic volatility models with diagnostics." *Journal of Econometrics*, **81**(1), 159 – 192. ISSN 0304-4076. doi:http://dx.doi.org/10.1016/S0304-4076(97)00039-0. URL http://www.sciencedirect.com/science/article/pii/S0304407697000390.

Ghalanos A (2014). *rugarch: Univariate GARCH models.* R package version 1.3-1.

Giacomini R, White H (2006). "Tests of conditional predictive ability." *Econometrica*, **74**(6), 1545–1578.

Glosten LR, Jagannathan R, Runkle DE (1993). "On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks." *The Journal of Finance*, **48**(5), 1779–1801. ISSN 1540-6261. doi:10.1111/j.1540-6261.1993.tb05128.x. URL http://dx.doi.org/10.1111/j.1540-6261.1993.tb05128.x.

González-Rivera G, Lee TH, Mishra S (2004). "Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood." *International Journal of Forecasting*, **20**(4), 629 – 645. ISSN 0169-2070. doi:http://dx.doi.org/10.1016/j.ijforecast.2003.10.003. URL http://www.sciencedirect.com/science/article/pii/S0169207003001420.

Hansen PR (2005). "A test for superior predictive ability." *Journal of Business & Economic Statistics*, **23**(4).

Hansen PR, Lunde A (2005). "A forecast comparison of volatility models: does anything beat a GARCH(1,1)?" *Journal of Applied Econometrics*, **20**(7), 873–889. ISSN 1099-1255. doi:10.1002/jae.800. URL http://dx.doi.org/10.1002/jae.800.

Hansen PR, Lunde A, Nason JM (2003). "Choosing the best volatility models: The model confidence set approach." *Oxford Bulletin of Economics and Statistics*, **65**(s1), 839–861.

Hansen PR, Lunde A, Nason JM (2011). "The model confidence set." *Econometrica*, **79**(2), 453–497.

Harvey AC (2013). *Dynamic Models for Volatility and Heavy Tails: With Applications to Financial and Economic Time Series.* 52. Cambridge University Press.

Harvey AC, Shephard N (1996). "Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns." *Journal of Business & Economic Statistics*, **14**(4), 429–434. doi:10.1080/07350015.1996.10524672. http://amstat.tandfonline.com/doi/pdf/10.1080/07350015.1996.10524672, URL http://amstat.tandfonline.com/doi/abs/10.1080/07350015.1996.10524672.

Higgins ML, Bera AK (1992). "A class of nonlinear ARCH models." *International Economic Review*, pp. 137–158.

Kilian L (1999). "Exchange rates and monetary fundamentals: What do we learn from long-horizon regressions?" *Journal of Applied Econometrics*, **14**(5), 491–510. URL http://hdl.handle.net/2027.42/34956.

Lopez JA (2001). "Evaluating the predictive accuracy of volatility models." *Journal of Forecasting*, **20**(2), 87–109.

McAleer M, da Veiga B (2008). "Single-index and portfolio models for forecasting value-at-risk thresholds." *Journal of Forecasting*, **27**(3), 217–235. ISSN 1099-131X. doi:10.1002/for.1054. URL http://dx.doi.org/10.1002/for.1054.

Nelson DB (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica*, **59**(2), 347–70. URL http://ideas.repec.org/a/ecm/emetrp/v59y1991i2p347-70.html.

R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rigby R, Stasinopoulos D (2005). "Generalized additive models for location, scale and shape." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **54**(3), 507–554.

Romano JP, Wolf M (2005). "Stepwise multiple testing as formalized data snooping." *Econometrica*, **73**(4), 1237–1282.

Samuels JD, Sekkel RM (2011). "Forecasting with Large Datasets: Trimming Predictors and Forecast Combination." *Working paper*.

Samuels JD, Sekkel RM (2013). "Forecasting with many models: Model confidence sets and forecast combination." *Technical report*, Bank of Canada Working Paper.

Schwert G (1990). "Stock volatility and the crash of '87." *Review of Financial Studies*, **3**(1), 77–102. doi:10.1093/rfs/3.1.77. http://rfs.oxfordjournals.org/content/3/1/77.full.pdf+html, URL http://rfs.oxfordjournals.org/content/3/1/77.abstract.

Silvennoinen A, Teräsvirta T (2009). "Multivariate GARCH models." In *Handbook of Financial Time Series*, pp. 201–229. Springer.

Taylor SJ (1986). *Modelling Financial Times Series*. Wiley.

Taylor SJ (1994). "MODELING STOCHASTIC VOLATILITY: A REVIEW AND COMPARATIVE STUDY." *Mathematical Finance*, **4**(2), 183–204. ISSN 1467-9965. doi:10.1111/j.1467-9965.1994.tb00057.x. URL http://dx.doi.org/10.1111/j.1467-9965.1994.tb00057.x.

Teräsvirta T (2009). "An introduction to univariate GARCH models." In *Handbook of Financial Time Series*, pp. 17–42. Springer.

West KD (1996). "Asymptotic inference about predictive ability." *Econometrica: Journal of the Econometric Society*, pp. 1067–1084.

White H (2000). "A reality check for data snooping." *Econometrica*, **68**(5), 1097–1126.

Wuertz D, with contribution from Michal Miklovic YC, Boudt C, Chausse P, others (2013). *fGarch: Rmetrics - Autoregressive Conditional Heteroskedastic Modelling*. R package version 3010.82, URL http://CRAN.R-project.org/package=fGarch.

Zakoian JM (1994). "Threshold heteroskedastic models." *Journal of Economic Dynamics and Control*, **18**(5), 931 – 955. ISSN 0165-1889. doi:http://dx.doi.org/10.1016/0165-1889(94)90039-6. URL http://www.sciencedirect.com/science/article/pii/016518994900396.

Affiliation:

Mauro Bernardi
Department of Methods and Models for Economics Territory and Finance
Faculty of Economics
Sapienza University of Rome
Via del Castro Laurenziano, 9
00161 Rome, Italy
E-mail: mauro.bernardi@uniroma1.it

Leopoldo Catania
Department of Economics and Finance
Faculty of Economics
University of Rome Tor Vergata
Via Columbia, 2
00133 Rome, Italy
E-mail: leopoldo.catania@uniroma2.it