Finance Research Letters 22 (2017) 169–174

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Determining risk model confidence sets

Mark Cummins[a,*], Michael Dowling[b], Francesco Esposito[a]

[a] DCU Business School, Dublin City University, Ireland
[b] Rennes School of Business, France

ARTICLE INFO

Article history:
Received 3 December 2016
Accepted 3 February 2017
Available online 9 February 2017

Keywords:
Model confidence set
Model selection
Market risk models

ABSTRACT

Two alternative approaches to identifying a model confidence set (MCS) are contrasted. Together with a specification of the established MCS test, we present a new version of a test that identifies a model set satisfying the MCS requirements and is characterised by an alternative model ranking *p*-value. We also contrast the two MCS approaches empirically, constructing a market risk model selection exercise for the Dow Jones Industrial Average. Our adapted MCS method is shown to lead to a smaller MCS, nested within the MCS determined by the popular MCS method, and allows greater distinction between models.

1. Introduction

In model selection, the model confidence set (MCS) is defined as the subset of possible models that, with a certain degree of confidence, are equivalent in some sense and contain the “best” model. This is clearly a useful technique especially in areas of economics and finance where many models make claims as to their superiority. The MCS approach of Hansen et al. (2011) is becoming a popular choice. This is an intuitive approach that primarily concentrates on reasons to exclude models. However this exclusion criterion is potentially excessively focused on downside risk and not sufficiently focused on general model performance. A second issue is the difficulty in distinguishing among the models that are accepted to the Hansen et al. MCS as there is a bias towards showing models as fully equivalent with no distinguishing features.

We contrast the Hansen et al. (2011) MCS technique (herein *max*-MCS) with an alternative MCS technique, an augmented version of Corradi and Distaso (2011) (herein *t*-MCS). We theoretically contrast the two approaches and also empirically test a range of market risk models for the Dow Jones Industrial Average (DJIA) from 1970 to 2013. We test 47 feasible market risk models for the DJIA, showing the challenge faced by risk modelling and management practitioners in choosing between the available models. We show that our *t*-MCS technique selects a smaller MCS, nested within the *max*-MCS, and most importantly allows greater distinction between models in the accepted set.

2. Model confidence set methods

Consider an initial collection of models $\mathcal{M}_0 \equiv \{\mathbf{P}_0, \mathbf{P}_1, \ldots, \mathbf{P}_m\}$. Model selection is based on the performance of model $\mathbf{P}_i \in \mathcal{M}_0$, as measured by the loss function $L_{i,t} := L(\mathbf{X}_t, \mathbf{P}_i)$; a function of the dataset $\mathbf{X}_t$. The lower $L_{i,t}$, the better the model. A *relative performance measure* $d_{ij,t} := L_{i,t} - L_{j,t}$ can be defined with expectation $\mu_{ij} := E[d_{ij,\,t}]$, such that: model i is preferred to model j ($\mathbf{P}_i \succ \mathbf{P}_j$), if $\mu_{ij} < 0$; model i is inferior to model j ($\mathbf{P}_i \prec \mathbf{P}_j$), if $\mu_{ij} > 0$; while the two models are

* Corresponding author.
E-mail address: mark.cummins@dcu.ie (M. Cummins).

http://dx.doi.org/10.1016/j.frl.2017.02.005
1544-6123/

equivalent ($\mathbf{P}_i \sim \mathbf{P}_j$), if $\mu_{ij} = 0$. The MCS is therefore

$$\mathcal{M}^\star \equiv \left\{ \mathbf{P}_i \in \mathcal{M}_0 : \mu_{ij} \le 0,\ \forall \mathbf{P}_j \in \mathcal{M}_0 \right\}.$$

The models in $\mathcal{M}^\star$ are characterised by equally superior performance, whereas $\mathcal{M}_0 \backslash \mathcal{M}^\star$ have inferior performance.

Both the *max*-MCS and *t*-MCS appeal to a central limit theorem (CLT) for dependent sequences in order to derive the model rejection rule, but differ in the algorithm implementation. With the *max*-MCS, we exploit the $T_{R,\mathcal{M}}$ test in Hansen et al. (2011). This utilises a sequentially rejective approach that relies on a sequence of scalar *equivalence tests* $H_{0,\mathcal{M}_k} : \mu_{ij} = 0\, \forall \mathbf{P}_i, \mathbf{P}_j \in \mathcal{M}_k$, yielding the model sequence $\mathcal{M}_0 \supset \mathcal{M}_1 \supset \cdots \supset \mathcal{M}_k$, by progressively discarding the lowest *p*-value models, one at a time. The test statistic is

$$T_{R,\mathcal{M}_k} = \max_{\mathbf{P}_i, \mathbf{P}_j \in \mathcal{M}_k} t_{ij} \tag{1}$$

where the t_{ij} are standardised values of $\bar{d}_{ij} = \frac{1}{T} \sum_t d_{ij,t}$. The following pseudo-code describes the algorithm:

Algorithm *max* -MCS

- Let $k = -1$; $\mathcal{M}_0 \equiv \{\mathbf{P}_i\}_{i=0,\ldots,m}$;
- **do** $k = k + 1$
  - **compute** $c(1-\alpha)$ quantile of the $T_{R,\mathcal{M}_k}$ distribution under $H_{0,\mathcal{M}_k}$
  - **if** any $t_{ij} > c$
    Let $\mathbf{P}_r$ be model producing the highest t_{ij}
    $\mathcal{M}_{k+1} \equiv \mathcal{M}_k \backslash \{\mathbf{P}_r\}$
  - **endif**
- **while** $H_{0,\mathcal{M}_k}$ not accepted
- **set** $\mathcal{M}^\star \equiv \mathcal{M}_k$ and **stop**;

At each iteration, the max statistic distribution is generated with Monte Carlo simulation as the max of a zero mean standard multivariate normal distribution, whose correlation matrix is determined via block bootstrap (Politis and Romano, 1994) applied to the t_{ij} sample of the surviving models. The worst expected performance is compared to the critical value of the max statistic. Each time $H_{0,\mathcal{M}_k}$ is rejected, the model with the worst target statistic and lowest *p*-value is expelled. This is the *elimination rule*. The test sequence terminates the first time that the null hypothesis of model equivalence is accepted. The *max*-MCS produces a model ranking as a result of the elimination sequence.

Some aspects of the *max*-MCS approach leave room for further consideration. First, the test exhibits some conservatism, isolating larger preferable model sets, due to the nature of the max $\bar{Z}$ statistic. In practice, this approach corresponds to the usage of a "worst case scenario" as a term of reference, in order to check for the existence of an inferior model. Second, the *p*-value associated with each model represents the confidence in the normality of that model's worst performance, in the context of the current model set $\mathcal{M}_k$. The latter value represents the confidence in the MCS. Nevertheless, this *p*-value is silent about the overall probability of each model belonging to the MCS and therefore generating superior performances, which is clearly useful information.

In the second method, the *t*-MCS, we modify the procedure presented in Corradi and Distaso (2011). The main feature of this alternative test is that of direct model performance comparison. Instead of constructing $m' \times (m' - 1)$ scalar model comparisons at each iteration, the MCS performs a random sequence of model benchmarking, whereby at each iteration inferior models are rejected until all the surviving models are equivalent. The rejected model set may include the current benchmark, if inferior to any competitors. The following pseudo-code describes the algorithm:

Algorithm *t*-MCS

- Let $k = 0$; $\mathcal{M}_0 \equiv \{\mathbf{P}_i\}_{i=0,\ldots,m}$; $\mathcal{B}_0 \equiv \emptyset$,
- **do**
  - **pick** any $\mathbf{P}_j \in \mathcal{M}_k \backslash \mathcal{B}_k$
  - **compute** t_{ij} $\forall i \neq j$
  - **call** *relative performance test* and let
    $\mathcal{E} \equiv \{\mathbf{P}_u \in \mathcal{M}_k : \mathbf{P}_u \prec \mathbf{P}_j\}$
  - **if** there is a $\mathbf{P}_s \in \mathcal{M}_k : \mathbf{P}_s \succ \mathbf{P}_j$,
    then $\mathcal{E} \equiv \mathcal{E} \cup \mathbf{P}_j$
    endif
  - $k = k + 1$, $\mathcal{M}_k \equiv \mathcal{M}_{k-1} \backslash \mathcal{E}$, $\mathcal{B}_k \equiv \mathcal{B}_{k-1} \cup \mathbf{P}_j$
- **while** $\mathcal{M}_k \backslash \mathcal{B}_k \not\equiv \emptyset$
- **set** $\mathcal{M}^\star \equiv \mathcal{M}_k$ and **stop**;

The random sequence of model benchmarking will generate a unique outcome if the decision rule is independent from the sequencing. This can be achieved by considering a test statistic and hence critical values that remain unchanged irrespective of the benchmark picking process. Nonetheless, the randomised sequence is not strictly necessary for the construction of the test. In fact, if we consider all possible benchmark sequencing we see that a model belongs to the MCS only if the