

文章编号:1003-207(2012)06-0001-08

# 中国燃油期货市场的 VaR 与 ES 风险度量

王 鹏<sup>1</sup>, 魏 宇<sup>2</sup>

(1. 西南财经大学金融学院, 四川 成都 610074;

2. 西南交通大学经济管理学院, 四川 成都 610031)

**摘 要:**近年来,我国燃油期货市场取得快速发展,但有关该市场波动特征和风险状况的研究却非常缺乏。以上海期货交易所燃油期货价格指数为例,分别在多头和空头两种头寸状况以及5种不同分位数水平下,运用条件覆盖检验、非条件覆盖检验等后验分析方法,实证对比了不同风险测度模型对 VaR 和 ES 两种不同风险指标估计的精度差异。研究结果表明:在我国燃油期货市场的风险测度估计中考虑国际燃油价格波动因素有助于获得更为精准的风险测度精度;在综合考虑了模型对价格变化动力学的刻画效果以及对极端风险的风险测度精度等因素后,FI-GARCH<sub>CST</sub>-SST 模型是一个相对合理的风险测度模型选择。

**关键词:**中国燃油期货市场; VaR; ES; 有偏学生 t 分布; 后验分析

**中图分类号:** F830.9

**文献标识码:** A

## 1 引言

近年来,国内学者在我国期货市场的价格波动特征及决定因素等方面开展了卓有成效的探索。例如,张小艳和张宗成<sup>[1]</sup>利用单位根检验与自相关检验的结合,探讨了国内铜、大豆、小麦等六大期货市场是否呈有效态势,结果显示:各种检验方法得出的结论不尽一致,除上天胶外,各大期货市场的对数期货价格序列均不能拒绝弱式有效市场的假设;杨升等<sup>[2]</sup>借助 ADF 单位根检验方法及协整理论,对大连商品交易所豆一与豆粕期货合约价格进行了实证分析,并得出了上述两交易品种期货合约价格之间存在协整关系的结论;付剑茹和张宗成<sup>[3]</sup>采用 Gibbs 抽样方法对我国铜期货市场的最优套期保值策略进行了实证分析,实证结果表明,估计风险对模型结果有重要影响,且贝叶斯统计较频率统计方法在处理估计风险方面有明显优势。

然而,需要指出的是,在对燃油期货市场的研究中,现有的绝大多数研究要么是专门探讨我国燃油

期货价格与世界上若干主要燃油期货产品价格之间的关系,要么是集中研究我国燃油期货价格本身的波动行为。例如,唐衍伟等<sup>[4]</sup>利用 Granger 因果关系检验等方法,对我国上海、美国 NYMEX 和新加坡燃油期货市场日收盘价的相互关系进行了实证研究,他们发现三个燃油期货市场间互为 Granger 成因;戴毓和周德群<sup>[5]</sup>考察了我国燃油期货市场的收益率和交易量的关系,研究结果表明,收益率与交易量之间不存在相关关系,绝对收益与交易量之间存在正相关关系;马超群等<sup>[6]</sup>运用改进的信息溢出模型,对上海燃油期货市场与国际石油市场的信息溢出(information spillover)关系进行了系统深入的研究,实证结果表明,上海燃油期货市场与新加坡燃油期货市场间有双向的均值溢出(mean spillover)。这些重要的研究成果对于深刻认识和把握我国燃油期货市场的实际波动特征和风险状况都具有积极的现实作用,但是就作者掌握的文献范围来看,目前还没有见到对我国燃油期货市场风险测度和预测模型直接相关的系统性研究。

目前,对金融市场风险的主流测度是 J. P. Morgan 投资银行在 RiskMetrics 系统中提出的风险价值 VaR(value at risk)。该指标一经提出,就逐渐成为金融行业风险度量与风险管理的一致标准。然而,其本身还是存在若干理论缺陷。比方说, VaR 只是说明了在一定的持有期(hold period)和置信水平(confidence level)下资产所面临的<sup>最大损失界限,而忽略了风险</sup>

收稿日期:2011-05-05;修订日期:2012-07-10

基金项目:国家自然科学基金资助项目(71071131,71101119);教育部新世纪优秀人才支持计划(NCET-08-0826);西南财经大学“2011工程”项目(211QJN10110)

作者简介:王鹏(1981-),男(汉族),山东宁阳人,西南财经大学金融学院讲师,管理学博士,研究方向:金融工程、风险管理、金融项目。

测度最应关心的收益分布的极端尾部状况。Bouchaud 和 Potters<sup>[7]</sup>、Yamai 和 Yoshida<sup>[8]</sup> 等都曾指出,这一重要理论缺陷使得 VaR 测度常常低估实际风险,而在低估的风险值下进行运作,可能会使金融机构面临更大的潜在破产风险。因此,Bouchaud 和 Potters<sup>[7]</sup>、Yamai 和 Yoshida<sup>[8]</sup> 以及最近的许多研究都认为<sup>[9-11]</sup>,更具价值的风险度量精度比较分析应当在另外一种更具理论优势的风险测度——预期损失 ES(excepted shortfall)测度下展开。

基于以上认识,本文的研究目的及特色在于:通过采用不同类型的波动模型来为我国燃油期货市场收益的条件波动率建模,同时选择对金融资产收益分布典型特性具有较强描述能力的有偏学生 t 分布(skewed student-t distribution,简记为 SST)来全面综合刻画我国燃油期货市场收益率的分布特征,在此基础上计算不同风险测度模型假定下燃油期货市场的风险价值 VaR 和预期损失 ES,并分别运用严谨系统的后验分析(backtesting analysis)方法,实证分析不同风险测度模型的对 VaR 和 ES 两种风险指标估计的适用范围和精确程度,探讨适用于我国燃油期货市场的风险测度方法。

## 2 数据

本文的研究样本为上海期货交易所燃油期货指数的每日收盘价,时间区间为 2004 年 8 月 25 日(即燃油期货合约上市交易首日,亦即燃油期货指数编制的起始日)至 2011 年 1 月 14 日,共有 T=1556 个样本点。数据来源于文华财经指数系统。

若令  $p_t$  表示燃油期货指数的每日收盘价,则每日收益率  $r_t(t = 1, 2, \dots, T)$  可以表示为:

$$r_t = 100 \times [\ln(p_t) - \ln(p_{t-1})] \quad (1)$$

表 1 报告了燃油期货指数日收益率  $r_t$  的描述性统计结果。

从表 1 的描述性统计结果可以看到:燃油期货指数的非条件收益率并不服从正态分布(J-B 检验在 1%水平上显著),且呈现出较为明显的“尖峰厚尾”(leptokurtic and fat tailed)和“有偏”(skewed)

形态;非条件收益序列不存在单位根(见 ADF 单位根检验结果);在较高的显著性水平上(1%),可以拒绝燃油期货指数的非条件收益率在较长时间范围内(10 期)都不具有自相关性的原假设(见  $Q_1(10)$  检验结果);较大收益率和较小收益率都在某段特定时期内交替集中出现,即我国燃油期货市场的非条件收益率表现出较为明显的波动聚集(volatility clustering)特征(见  $Q_2(10)$  检验结果)。

## 3 燃油期货指数收益率的计量模型描述

在金融计量研究中,一般假定金融资产的日收益率序列  $r_t$  满足以下离散形式:

$$r_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t z_t \quad (2)$$

其中,  $\mu_t$  是收益率的条件均值(conditional mean),  $\sigma_t^2$  是条件方差(conditional variance),而新生量  $z_t$  (innovation)则是一个满足均值为 0、方差为 1 的独立同分布(i. i. d.)随机变量。

### 3.1 条件均值的建模方法

有研究表明,AR(1)模型是一种简单但非常实用的刻画条件均值  $\mu_t$  的模型<sup>[12]</sup>,因此结合表 1 中燃油期货指数收益所展现出的统计特征,我们假定其条件均值  $\mu_t$  满足一个 AR(1)过程,即指数收益率  $r_t$  通过下式刻画:

$$r_t = \rho r_{t-1} + \sigma_t z_t \quad (3)$$

### 3.2 条件方差的建模方法

表 1 中的  $Q_2(10)$  统计检验结果显示,我国燃油期货指数收益率具有非常明显的波动聚集效应。为了描述收益率的这种典型特征,我们采用在金融计量研究中最常见的经典 GARCH 模型及其变型形式为其建模。Engle 和 Patton<sup>[13]</sup> 的研究表明,在实证研究当中,GARCH(1,1)模型是权衡计算精度和模型复杂程度的一种比较合适的折中。因此,我们首先假定条件方差满足具有如下形式的 GARCH(1,1)模型:

$$\sigma_t^2 = \beta_0 + \beta_1 \varepsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 \quad (4)$$

另外,为了将金融市场的其它很多典型事实

表 1 燃油期货指数收益率的描述性统计

均值	标准差	偏度	峰度	J-B	ADF	$Q_1(10)$	$Q_2(10)$
0.051	1.577	-0.476 <sup>†</sup>	2.229 <sup>†</sup>	380.743 <sup>†</sup>	-27.064 <sup>†</sup>	20.826 <sup>*</sup>	782.520 <sup>†</sup>

注:表中,“†”和“\*”分别代表在 1%和 10%水平上显著。其中,峰度为超额峰度系数(Excess kurtosis coefficient),正态分布的超额峰度系数为 0;J-B 为检验收益率是否服从正态分布的 Jarque-Bera 统计量;ADF 是以最小 AIC 准则确定最优检验滞后阶数后得到的 Augmented Dickey-Fuller 单位根检验统计量; $Q_1(10)$  为燃油期货指数收益率滞后 10 阶的 Ljung-Box Q 统计量; $Q_2(10)$  为燃油期货指数平方收益率滞后 10 阶的 Ljung-Box Q 统计量。

(stylized facts) 纳入 GARCH 模型的分析框架, 一些学者还发展出了许多其它类型的非线性 GARCH 族模型<sup>[14]</sup>。结合表 1 中指数收益率所展现的若干描述性统计特征, 本文对条件方差的建模继续考虑 FIAPARCH 模型。常用的 FIAPARCH(1,  $d$ , 1) 模型可以具体表示为以下形式:

$$\sigma_t^\delta = \beta_0 + [1 - (1 - \beta_2 L)^{-1} \beta_1 L (1 - L)^d] (|\epsilon_{t-1}| - \gamma \epsilon_{t-1})^\delta \quad (5)$$

FIAPARCH 模型中, 对价格波动非对称性的捕捉通过参数  $\gamma$  来实现。另外, FIAPARCH 模型中的参数“ $d$ ”是“长记忆参数”。

进一步, 为了增强本文研究方法的可比较性, 进而提升研究结论的稳健性, 除了 GARCH 模型和 FIAPARCH 模型外, 我们还将考虑以下两种不同的波动率模型。

FIGARCH(1,  $d$ , 1):

$$\sigma_t^2 = \beta_0 + [1 - \beta_2 L - (1 - \beta_1 L)(1 - L)^d] \epsilon_t^2 + \beta_2 \sigma_{t-1}^2 \quad (6)$$

其中,  $L$  为滞后算子,  $d$  为长记忆参数, 它刻画了波动的长短记忆性, 当  $0 < d < 1$  时, 市场波动服从一个长记忆过程。

APARCH(1, 1):

$$\sigma_t = [\beta_0 + \beta_1 (|\epsilon_{t-1}| - \gamma \epsilon_{t-1})^\delta + \beta_2 \sigma_{t-1}^2]^\frac{1}{\delta} \quad (7)$$

该模型中各参数的含义与 FIAPARCH 模型相同, 不再赘述。

我们进一步的考虑是, 现有文献的研究结果已经表明, 中国燃油市场的价格波动与国际上一些主要燃油市场价格波动之间存在密切联系, 因此为了考察国际燃油市场价格波动是否有助于我国燃油期货市场的风险度量, 我们将国际燃油市场价格波动作为一个解释变量加入到上述 4 种波动模型中。在具体品种的选择上, 由于目前新加坡是亚洲的成品油定价中心, 国内燃油价格主要受到新加坡燃油市场价格的影响, 而新加坡燃油市场价格有两个代表性价格数据——新加坡高硫 180CST 现货价格和纸货价格, 其中现货交易的成交量和持仓量要比纸货交易大得多, 另外国内进口的燃油价格也基本参照新加坡燃油现货报价, 因此与马超群等(2009)一致, 这里首先将新加坡高硫 180CST 现货价格确定为国际燃油市场价格的代表(数据来源于 Bloomberg 资讯)。然后, 按照金融资产价格波动研究中的惯例<sup>[15-16]</sup>, 我们将滞后一期 180CST 现货日收益率的平方  $r_{t-1}^2$  作为解释变量纳入上文中的 4 类波动模型, 即 GARCH<sub>CST</sub>

(1, 1) 模型的具体形式为:

$$\sigma_t^2 = \beta_0 + \beta_1 \epsilon_{t-1}^2 + \beta_2 \sigma_{t-1}^2 + \beta_3 r_{t-1}^2 \quad (8)$$

同理, FIAPARCH<sub>CST</sub>(1,  $d$ , 1):

$$\sigma_t^\delta = \beta_0 + [1 - (1 - \beta_2 L)^{-1} \beta_1 L (1 - L)^d] (|\epsilon_{t-1}| - \gamma \epsilon_{t-1})^\delta + \beta_3 r_{t-1}^2 \quad (9)$$

FIGARCH<sub>CST</sub>(1,  $d$ , 1):

$$\sigma_t^2 = \beta_0 + [1 - \beta_2 L - (1 - \beta_1 L)(1 - L)^d] \epsilon_t^2 + \beta_2 \sigma_{t-1}^2 \quad (10)$$

APARCH<sub>CST</sub>(1, 1):

$$\sigma_t = [\beta_0 + \beta_1 (|\epsilon_{t-1}| - \gamma \epsilon_{t-1})^\delta + \beta_2 \sigma_{t-1}^2]^\frac{1}{\delta} \quad (11)$$

最后需要指出的是, 由于上海期货交易所原油期货和 180CST 现货的交易日期并不完全一致, 因此在计算  $r_{t-1}^2$  时是以上海期货交易所原油期货的交易日期为准, 对 180CST 现货的收益率进行累积计算。

### 3.3 新生量 $z_t$ 的建模方法

本文采用能同时刻画“有偏”和“尖峰胖尾”分布特征的有偏学生  $t$  分布来为新生量(条件收益)  $z_t$  建模, 即假定  $z_t \sim SST(0, 1, \xi, \nu)$ 。

SST 的概率密度函数表示为<sup>[14-15]</sup>:

$$f(z_t | \xi, \nu) = \begin{cases} \frac{2s}{\xi + \frac{1}{\xi}} g[\xi(sz_t + m) | \nu] & \text{if } z_t < -\frac{m}{s} \\ \frac{2s}{\xi + \frac{1}{\xi}} g[\frac{sz_t + m}{\xi} | \nu] & \text{if } z_t \geq -\frac{m}{s} \end{cases} \quad (12)$$

其中,  $g(\cdot | \nu)$  是标准对称学生  $t$  分布(均值为 0, 方差为 1)的概率密度函数;  $\nu$  是其自由度(degree of freedom);  $\xi$  是“非对称参数”(parameter of asymmetry), 如果  $\ln(\xi) > 0$  则表明该新生量的分布为右偏, 反之如果  $\ln(\xi) < 0$ , 则表明其为左偏;  $m$  和  $s$  分别是该 SST 分布的均值和标准差。另外, 除了 SST 分布以外, 我们还将运用常用的正态分布为新生量  $z_t$  建模, 即假定  $z_t \sim N(0, 1)$ 。

### 3.4 风险测度计算模型的参数估计及诊断检验结果

我们对上文提出的 16 种风险测度计算模型进行了参数估计和诊断检验。限于篇幅, 表 2 只报告了 GARCH-SST 等 4 种模型的参数估计和诊断检验结果。

由表 2 的估计结果可见: 我国原油期货指数收益波动中存在非常明显的长记忆性, 但并不存在明显的杠杆效应( $d$  参数都显著小于 1, 但  $\gamma$  估计值不显著); 与非条件收益率一样, 条件收益率同样具有较为明显的有偏和厚尾分布特征(参数  $\ln(\xi)$  和  $\nu$  都

表 2 燃油期货指数 4 种风险测度计算模型的参数估计和诊断检验结果

	GARCH-SST	FIAPARCH-SST	FIGARCH-SST	APARCH-SST
$\rho$	-0.083 (0.001)	-0.086 (0.001)	-0.084 (0.001)	-0.089 (0.001)
$\beta_0$	0.031 (0.033)	0.043 (0.082)	0.043 (0.118)	0.025 (0.031)
$d$	—	0.932 (0.000)	0.875 (0.000)	—
$\beta_1$	0.097 (0.000)	-0.005 (0.956)	0.018 (0.871)	0.103 (0.000)
$\beta_2$	0.899 (0.000)	0.864 (0.000)	0.836 (0.000)	0.905 (0.000)
$\gamma$	—	-0.069 (0.256)	—	-0.081 (0.189)
$\delta$	—	1.667 (0.000)	—	1.476 (0.000)
$\ln(\xi)$	-0.097 (0.002)	-0.010 (0.002)	-0.097 (0.002)	-0.102 (0.001)
$v$	5.785 (0.000)	6.050 (0.000)	5.683 (0.000)	5.846 (0.000)
$\ln L$	-2711.249	-2708.764	-2710.156	-2709.263
AIC	3.495	3.496	3.494	3.495
$Q_1(20)$	21.980 (0.285)	24.272 (0.186)	22.7697 (0.246)	24.033 (0.195)
$Q_2(20)$	12.298 (0.831)	15.287 (0.642)	14.831 (0.674)	12.653 (0.812)
ARCH-F	0.454 (0.635)	0.510 (0.988)	0.538 (0.981)	0.609 (0.953)
APC(20)	14.736 (0.324)	11.495 (0.321)	11.727 (0.468)	13.219 (0.060)

注：表中，上半部分是对我国燃油期货指数 4 种风险测度计算模型的估计结果，下半部分是对模型估计的诊断检验结果。圆括号中的数字是各检验的显著性 p 值； $\ln L$  是对数极大似然函数值；AIC 是平均赤池信息； $Q_1(20)$  是残差序列的 Ljung-Box  $Q(20)$  检验值； $Q_2(20)$  是平方残差序列的 Ljung-Box  $Q(20)$  检验值；ARCH-F 是残差序列 ARCH 效应检验的 F 统计量值；APC(20) 是 20 等分区间的 Adjusted Pearson Chi-square Goodness-of-fit test 检验值。

在较高水平上显著)；由 APC(20) 统计量的检验结果来看，FIGARCH-SST 模型在表 2 所列出的 4 种模型中取得了最优的 Adjusted Pearson Chi-square 拟合效果。

#### 4 燃油期货指数 VaR 测度值的计算及后验分析

VaR 定义为某项资产或资产组合在一定持有期和置信水平  $a$  下的最大可能损失，即：

$$Pr\{r_t > VaR_t^a\} = 1 - a \quad (13)$$

由上述理论表达式可以看出，VaR 相当于  $r_t$  概率分布在  $a$  水平上的分位数。若假定：

$$f\{r_t | \Psi_{t-1}\} \sim N(\mu_t, \sigma_t^2) \quad (14)$$

式中， $f\{\cdot | \Psi_{t-1}\}$  表示基于信息集  $\Psi_{t-1}$  的条件分布， $\Psi_t = \sigma(r_t, r_{t-1}, \dots)$ ，其中  $\sigma(r_t, r_{t-1}, \dots)$  表示由  $(r_t, r_{t-1}, \dots)$  生成的  $\sigma$  代数， $N(\cdot)$  为正态分布，则可得  $t$  时刻  $q$  分位数下的 VaR 估计方法为：

$$VaR_t^q = \mu_t + z_q \sigma_t \quad (15)$$

其中， $z_q$  为所要考察的金融收益分布的  $q$  损失分位数。对于多头头寸， $q$  应取左尾分位数；对于空头头寸， $q$  应取右尾分位数。为保证研究结论的可靠性，本文将多头头寸分位数  $z_q$  分别取为 10%、5%、2.5%、1%、0.1%，对应的空头头寸分位数分别为 90%、95%、97.5%、99%、99.9%。同时，条件均值  $\mu_t$  和波动率(条件标准差)  $\sigma_t$  通过前述各类风险测度模型估计得出，由此我们可以计算不同模型假

定下的 VaR 风险测度值并开展 Backtesting 分析。

为了确保研究结论的稳健性，我们考虑两种不同的 VaR Backtesting 方法，即 Kupiec<sup>[11]</sup> 提出的非条件覆盖检验(unconditional coverage testing)，以及 Engle 和 Manganelli<sup>[18]</sup> 提出的条件覆盖检验(conditional coverage testing)。为了定量比较不同风险测度模型的 VaR 测度精度，我们在考虑评判波动模型的 VaR 测度精度时，所采用的定量判断标准是对比相应上述两种检验的显著性 p 值。也就是说，如果对由某一风险测度模型所计算的 VaR 值的检验 p 值越大，则说明我们越不能拒绝原假设  $H_0$ ，即表明该风险测度模型的 VaR 测度精度越高。非条件覆盖检验和条件覆盖检验的具体步骤可分别参见 Kupiec<sup>[17]</sup> 和 Engle 和 Manganelli<sup>[18]</sup>。

#### 4.1 VaR 值计算结果

图 1 报告了不同模型对 1% 分位数下多头头寸 VaR 的估计结果，图 2 是对应空头头寸在 99% 分位数下的情况。为了图形清晰起见，我们只选择了 GARCH-N、APARCH-SST、FIGARCH-N、FIAPARCH-SST 等 4 种模型和全样本中一段期间内 ( $t = 800, 801, \dots, 1000$ ) 的结果进行展示。同时，为了对风险测度模型 VaR 估计值的准确性进行初步判断，图 1 和图 2 中用针头图标识收益率  $r_t$ 。

由图 1 和图 2 的直观表象来看，总体来讲，在 1%(99%) 的分位数下，对于多头头寸(空头头寸)来说，基于正态分布的风险测度模型似乎都有低估

VaR 值的倾向。当然,要得到更为精确的结论,必须对基于各个风险测度模型的 VaR 序列进行 Back-testing 检验。

#### 4.2 不同 VaR 计算模型的 Backtesting 检验结果

表 3 和表 4 中的数字为 VaR 后验分析检验的显著性  $p$  值。 $p$  值越大,表明 VaR 准确度越高。

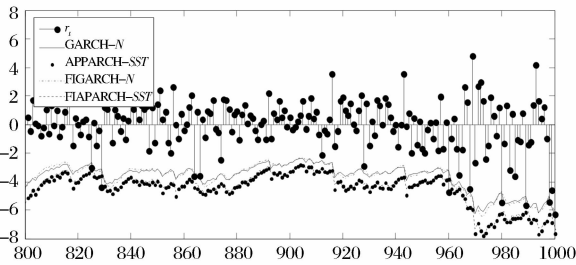


图 1 1%分位数下多头头寸的部分 VaR 估计结果

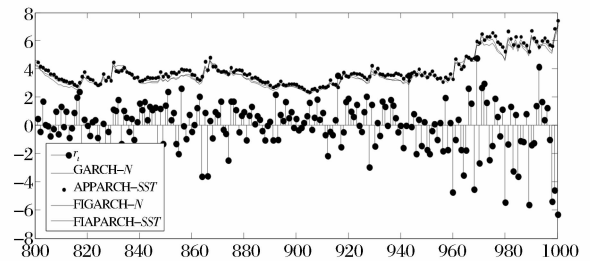


图 2 99%分位数下空头头寸的部分 VaR 估计结果

表 3 不同 VaR 计算模型的非条件覆盖检验结果

	多头 VaR					空头 VaR				
	10%	5%	2.5%	1%	0.1%	90%	95%	97.5%	99%	99.9%
GARCH-N	0.063	0.300	0.083	0.015	0.007	0.326	0.930	0.984	0.122	0.001
GARCH-SST	0.111	0.076	0.887	0.221	1.000	0.062	0.133	0.732	0.395	0.733
APARCH-N	0.027	0.300	0.153	0.002	0.007	0.418	0.660	0.887	0.122	0.007
APARCH-SST	0.093	0.099	0.856	0.221	1.000	0.090	0.345	0.732	0.542	0.733
FIGARCH-N	0.013	0.201	0.201	0.027	0.007	0.248	0.359	0.984	0.122	0.002
FIGARCH-SST	0.132	0.076	0.522	0.221	1.000	0.145	0.201	0.616	0.542	0.634
FIAPARCH-N	0.034	0.201	0.083	0.151	0.001	0.248	0.577	0.732	0.122	0.007
FIAPARCH-SST	0.156	0.201	0.616	0.221	0.634	0.075	0.345	0.616	0.395	0.733
GARCH <sub>CST</sub> -N	0.037	0.296	0.172	0.294	0.272	0.182	0.136	0.464	0.297	0.177
GARCH <sub>CST</sub> -SST	0.136	0.301	0.707	0.349	0.967	0.493	0.468	0.286	0.962	0.321
APARCH <sub>CST</sub> -N	0.459	0.047	0.322	0.030	0.317	0.328	0.026	0.080	0.024	0.377
APARCH <sub>CST</sub> -SST	0.994	0.996	0.193	0.809	0.073	0.119	0.345	0.077	0.321	0.378
FIGARCH <sub>CST</sub> -N	0.017	0.059	0.319	0.105	0.160	0.301	0.303	0.084	0.038	0.319
FIGARCH <sub>CST</sub> -SST	0.100	0.203	0.624	0.266	0.192	0.417	0.310	0.325	0.377	0.601
FIAPARCH <sub>CST</sub> -N	0.179	0.015	0.262	0.330	0.534	0.290	0.311	0.035	0.235	0.045
FIAPARCH <sub>CST</sub> -SST	0.294	0.319	0.404	0.248	0.993	0.234	0.927	0.242	0.258	0.952

注:表中数字为非条件覆盖检验的  $p$  值。 $p$  值越大,说明由该模型所计算的 VaR 精确度越高。

表 4 不同 VaR 计算模型的条件覆盖检验结果

	多头 VaR					空头 VaR				
	10%	5%	2.5%	1%	0.1%	90%	95%	97.5%	99%	99.9%
GARCH-N	0.012	0.092	0.118	0.047	0.000	0.377	0.438	0.453	0.024	0.000
GARCH-SST	0.039	0.081	0.731	0.357	0.669	0.024	0.034	0.035	0.015	0.664
APARCH-N	0.007	0.054	0.076	0.004	0.000	0.624	0.416	0.481	0.014	0.000
APARCH-SST	0.072	0.062	0.717	0.305	0.669	0.070	0.127	0.082	0.011	0.599
FIGARCH-N	0.019	0.083	0.163	0.091	0.000	0.285	0.175	0.113	0.009	0.000
FIGARCH-SST	0.217	0.043	0.457	0.358	0.669	0.087	0.050	0.014	0.022	0.806
FIAPARCH-N	0.009	0.067	0.010	0.039	0.000	0.476	0.395	0.293	0.008	0.000
FIAPARCH-SST	0.052	0.054	0.573	0.312	0.905	0.149	0.282	0.032	0.015	0.676
GARCH <sub>CST</sub> -N	0.694	0.496	0.090	0.280	0.062	0.158	0.182	0.699	0.143	0.084
GARCH <sub>CST</sub> -SST	0.009	0.208	0.377	0.744	0.174	0.739	0.999	0.172	0.021	0.228
APARCH <sub>CST</sub> -N	0.092	0.070	0.075	0.267	0.724	0.383	0.312	0.497	0.034	0.019
APARCH <sub>CST</sub> -SST	0.516	0.457	0.182	0.393	0.262	0.809	0.058	0.364	0.017	0.798
FIGARCH <sub>CST</sub> -N	0.086	0.058	0.104	0.079	0.035	0.270	0.325	0.050	0.193	0.022
FIGARCH <sub>CST</sub> -SST	0.301	0.083	0.489	0.284	0.010	0.326	0.880	0.015	0.329	0.078
FIAPARCH <sub>CST</sub> -N	0.042	0.327	0.235	0.036	0.102	0.093	0.130	0.042	0.014	0.280
FIAPARCH <sub>CST</sub> -SST	0.587	0.503	0.870	0.367	0.327	0.158	0.313	0.163	0.327	0.081

注:表中数字为条件覆盖检验的  $p$  值。 $p$  值越大,说明由该模型所计算的 VaR 精确度越高。

由表3和表4的两种 Backtesting 检验结果可以发现:

(1)回顾表2中的参数估计结果,我国原油期货指数收益率并未展现出明显的杠杆效应。因此,这里通过比较不包含杠杆效应项的波动模型与包含杠杆效应项的对应模型的 Backtesting 检验结果同样可以看出,采用包含杠杆效应项的波动模型并未取得明显优于不包含杠杆效应项对应模型的检验  $p$  值,即在波动模型中纳入杠杆效应项对提高我国原油期货市场的 VaR 测度精度并无太多帮助;

(2)观察同一波动模型分别在正态分布和 SST 分布下的不同检验  $p$  值可以发现,无论是非条件覆盖检验,还是条件覆盖检验,基于正态分布的风险价值计算模型在较低的分位数水平上往往能够取得较基于 SST 分布模型更高的 Backtesting 检验  $p$  值,而在较高的分位数水平上,基于 SST 分布的 VaR 计算模型的表现却相对更优。因此可以认为,实务界在估计 VaR 值时普遍采用的正态分布假设适用于对我国原油期货非极端风险的测度,而使用 SST 分布则有助于提高对原油期货极端风险的估计精度。

(3)对比不同记忆性波动模型的 Backtesting 检验结果可以看到,尽管在原油期货指数收益波动中存在较为明显的长记忆性(见表2中参数  $d$  的估计结果),但长记忆波动模型与短记忆对应模型的 Backtesting 检验  $p$  值却不存在系统性差异,即采用长记忆模型为我国原油期货市场的收益率建模,并不能提高对其 VaR 测度的估计精度。这一结果似乎令人惊讶,但我们认为,可能的原因在于,本文的研究专注于对未来1天 VaR 值的估计,而对1天这样相对较短时间的市场风险测度来讲,短记忆模型的精度已经足够<sup>[9]</sup>。

(4)对比普通波动模型和将新加坡 180CST 燃油现货价格波动作为解释变量的对应模型的风险测度精度可以发现,总体上讲,包含解释变量的波动模型取得了较普通模型更好的 VaR 计算效果,这表现为无论是非条件覆盖检验还是条件覆盖检验,在大多数情况下,包含解释变量的波动模型都取得了较对应的普通模型更大的检验  $p$  值;

(5)再次对表3和表4的后验分析结果进行全面审视,可以发现,许多风险测度模型 Backtesting 检验的  $p$  值都远大于常用的统计检验水平(如1%、5%或10%等)。也就是说,如果采用传统的检验判断标准,大多数风险测度模型都是可以通过显著性

检验的。但重要的是,即使可以通过检验,不同模型的 Backtesting 检验  $p$  值间确实还是存在较大差别,即不同模型对我国原油期货市场的风险测度效果确实是存在较大差异的。以这里的几种风险价值计算模型为例,就我国原油期货市场的风险测度和预测而言,在综合考虑了模型对价格变化动力学的刻画效果(见表2)以及对极端风险的测度精度等两方面因素后,我们认为 FIGARCH<sub>CST</sub>-SST 是一个相对合理的模型选择。

## 5 ES 测度值的计算及后验分析

鉴于 VaR 测度在描述金融市场极端波动风险时暴露出的种种缺陷(如忽略极端尾部风险状况、不满足次可加性等),Artzner<sup>[20-21]</sup>提出了一种更具理论优势的风险测度——预期损失 ES。

ES 测度的理论定义为<sup>[20]</sup>:

$$ES_t^q = -E_{t-1}(r_t | r_t < -VaR_t^q) \quad (16)$$

在具体的估计过程中,可以采用首先将区间(0,  $q$ )进行  $M$  等分并由此得到  $M$  个 VaR 值,然后求取  $M$  个 VaR 均值的方法获得对 ES 的估计,具体步骤可见 McNeil 和 Frey<sup>[12]</sup>。

McNeil 和 Frey<sup>[12]</sup>提出了基于自举法(bootstrap)的后验分析方法。考虑如下形式的超出残差(exceedance residuals):

$$y_t = \frac{x_t - ES_t^q}{h^{\frac{1}{p}}} \quad (17)$$

其中,  $x_t$  指的是损失超出 VaR 值的日收益率  $r_t$ 。ES 测度的后验分析过程基于超出残差  $y_t$ ,后续具体步骤可参见 McNeil 和 Frey<sup>[12]</sup>。与对 VaR 的后验分析一样,在判断不同波动模型的 ES 测度精度时,所采用的定量标准仍是其后验分析的显著性  $p$  值。按照 McNeil 和 Frey<sup>[12]</sup>中的步骤,我们对 16 种波动模型在不同分位数水平下的 ES 估计精度进行了后验分析,并将结果报告于表5中。

观察表5的实证结果,可以发现:

(1)各种波动模型对多头头寸 ES 的估计效果落后于其在空头头寸中的表现。这体现在:在表5进行的针对多头头寸的 80 次 Backtesting 检验中(5 种分位数,16 种波动模型),如果采用常用的 5% 的显著性水平,那么仅有 34 次估计可以通过 Backtesting 检验。而在表5右半部分所进行的针对空头头寸的 80 次 Backtesting 检验中,这个数字为 65 次。也就是说,在空头头寸中,我们在很多时候都无法拒绝“波动模型准确”的原假设;

表5 不同 ES 计算模型的 Backtesting 检验结果

	多头 ES					空头 ES				
	10%	5%	2.5%	1%	0.1%	90%	95%	97.5%	99%	99.9%
GARCH-N	0.001	0.000	0.000	0.023	0.081	0.057	0.094	0.075	0.129	0.322
GARCH-SST	0.003	0.000	0.023	0.072	0.143	0.092	0.108	0.076	0.144	0.297
APARCH-N	0.003	0.000	0.001	0.029	0.044	0.075	0.081	0.079	0.170	0.318
APARCH-SST	0.001	0.000	0.005	0.034	0.101	0.185	0.076	0.097	0.160	0.288
FIGARCH-N	0.004	0.001	0.001	0.012	0.154	0.045	0.051	0.030	0.085	0.310
FIGARCH-SST	0.005	0.000	0.024	0.072	0.156	0.025	0.148	0.031	0.194	0.344
FIAPARCH-N	0.003	0.000	0.000	0.015	0.088	0.043	0.043	0.023	0.062	0.337
FIAPARCH-SST	0.002	0.000	0.012	0.009	0.111	0.082	0.094	0.105	0.197	0.349
GARCH <sub>CST</sub> -N	0.092	0.017	0.087	0.015	0.032	0.321	0.099	0.228	0.166	0.224
GARCH <sub>CST</sub> -SST	0.010	0.095	0.257	0.130	0.613	0.319	0.018	0.035	0.328	0.087
APARCH <sub>CST</sub> -N	0.294	0.103	0.034	0.051	0.042	0.081	0.894	0.014	0.082	0.302
APARCH <sub>CST</sub> -SST	0.156	0.263	0.016	0.070	0.016	0.315	0.180	0.433	0.275	0.850
FIGARCH <sub>CST</sub> -N	0.082	0.036	0.056	0.163	0.149	0.612	0.013	0.635	0.547	0.799
FIGARCH <sub>CST</sub> -SST	0.038	0.061	0.150	0.192	0.217	0.328	0.048	0.437	0.548	0.545
FIAPARCH <sub>CST</sub> -N	0.025	0.091	0.168	0.003	0.067	0.338	0.043	0.168	0.341	0.261
FIAPARCH <sub>CST</sub> -SST	0.105	0.008	0.081	0.018	0.175	0.076	0.035	0.040	0.870	0.247

(2)与观察不同模型 VaR 估计精度差异的方法一样,这里通过比较不包含杠杆效应项的波动模型与包含杠杆效应项的对应模型的 Backtesting 检验结果同样可以看出,采用含杠杆效应项的波动模型并未取得明显大于不含杠杆效应项对应模型的检验 p 值,也就是说,在波动模型中纳入杠杆效应项对提高我国原油期货市场的 ES 测度精度也没有太大的效果;

(3)再次对比不同记忆性波动模型的 Backtesting 检验结果可以看到,长记忆波动模型与短记忆对应模型的 Backtesting 检验 p 值并无太大差别,即与基于表 3 和表 4 的研究结论一样,这里的实证结果再次表明,采用长记忆模型为我国原油期货市场的收益率建模,并不能提高对其风险指标的估计精度;

(4)再次对比普通波动模型和将新加坡 180CST 燃油现货价格波动作为解释变量的对应模型的 ES 测度精度可以发现,带有解释变量模型的 ES 估计效果总体上优于普通模型,这一结论与基于 VaR 精度分析的第(4)条结论一致,共同说明了在我国原油期货市场的风险测度估计中考虑国际燃油价格波动因素的重要性和有效性;

(5)最后,全面观察表 5 中的 Backtesting 分析结果可以发现,总体上讲,在表 3 和表 4 中取得了不错的 VaR 估计精度的 FIGARCH<sub>CST</sub>-SST 模型在对 ES 测度的估计中再次取得了相对较好的风险测度精度。

## 6 结语

本文以上海期货交易所原油期货指数为例,深

入研究了我国原油期货市场的价格变化统计特征及波动模式,并运用严谨系统的 Backtesting 方法,实证对比了 16 种风险测度模型对原油期货市场的 VaR 和 ES 两种风险指标的测度精度。为了确保研究结论的稳健性和实用性,本文有关不同模型的风险指标估计及后验分析的研究都是在多头和空头两种不同的头寸状况以及 5 种不同的分位数水平下展开的。

本文所取得的主要研究结论包括:(1)我国原油期货市场价格收益呈现出较为明显的有偏和尖峰厚尾分布,收益波动具有明显的聚集特征,但没有证据表明原油期货市场的波动具有与中国股票市场类似的“杠杆效应”;(2)采用有偏学生 t 分布有助于提高对我国原油期货市场价格风险测度指标的估计精度,而在波动模型中包含杠杆效应项和运用长记忆波动模型对提高 VaR 和 ES 的估计精度并无太多帮助;(3)在我国原油期货市场的风险测度估计中考虑国际燃油价格波动因素有助于获得更为精准的风险测度精度;(4)在综合考虑了模型对价格变化动力学的刻画效果以及对极端风险的风险测度精度等两方面因素后,我们认为就我国原油期货市场的风险测度而言,FI-GARCH<sub>CST</sub>-SST 是一个相对合理的模型选择。

论文的研究方法和实证结果为我国原油期货市场的风险测度和管理以及市场监管工作提供了一些有益的理论借鉴和极具可操作性的方法选择。比方说,原油期货价格波动的有偏特性说明,同一标的资产的多头和空头在相同市场条件下的风险也是有显著差异的,这时对原油期货交易的多头头寸和空头头寸应该分别做出有差异的限制规定,例如对多头和空头头寸的持仓上限以及保证金收取比例应该有

所不同等。另外,在选择适用于不同期货品种的风险测度模型时,也必须预先对该品种的收益分布及波动状况进行深入检验,并通过定量比较不同模型 Backtesting 检验的  $p$  值大小来确定最优风险测度模型。

### 参考文献:

- [1] 张小艳, 张宗成. 期货市场有效性理论与实证检验[J]. 中国管理科学, 2005, 13(6): 1-5.
- [2] 杨升, 何凌云, 周曙东, 等. 大商所豆一与豆粕期货合约价格的协整性分析[J]. 中国管理科学, 2008, 16(S1): 302-305.
- [3] 付剑茹, 张宗成. 套期保值, 估计风险与贝叶斯统计——基于中国铜期货市场的经验研究[J]. 中国管理科学, 2009, 17(4): 21-29.
- [4] 唐衍伟, 陈刚, 李海英. 我国与国际燃料油期货市场长期均衡的实证研究[J]. 系统工程, 2007, 25(10): 51-57.
- [5] 戴毓, 周德群. 我国燃料油期货收益、交易量关系的实证研究[J]. 系统工程, 2008, 26(7): 35-39.
- [6] 马超群, 余升翔, 陈彦玲, 等. 中国上海燃料油期货市场信息溢出效应研究[J]. 管理科学学报, 2009, 12(3): 92-101.
- [7] Bouchaud J P, Potters M. Theory of financial risk: From statistical physics to risk management [M]. Cambridge: Cambridge University Press, 1999.
- [8] Yamai Y, Yoshida T. Value-at-Risk versus expected shortfall: a practical perspective [J]. Journal of Banking and Finance, 2005, 29(4): 997-1015.
- [9] Robert J E, Hong M. VaR and expected shortfall: a non-normal regime switching framework [J]. Quantitative Finance, 2009, 9(6): 747-755.
- [10] Abderrahim T. Analytical value-at-risk and expected shortfall under regime switching [J]. Financial Research Letters, 2009, 6 (1): 138-151.
- [11] Lennart H, Herman K. Bayesian forecasting of Value at Risk and expected shortfall using adaptive importance sampling [J]. International Journal of Forecasting, 2010, 26(2): 231-247.
- [12] McNeil A J, Frey R. Estimation of tail related risk measures for heteroscedastic financial time series: an extreme value approach [J]. Journal of Empirical Finance, 2000, 7(3-4): 271-300.
- [13] Engle R F, Patton A. What good is volatility model? [J]. Quantitative Finance, 2001, 1(2): 237-245.
- [14] Down K. Measuring market risk [M]. Chi Chester: John Wiley & Sons Press, 2005.
- [15] Chung I Y, Lu J R, Lee P H. Forecasting volatility in the financial markets: a comparison of alternative distributional assumptions [J]. Applied Financial Economics, 2007, 17(13): 1051-1060.
- [16] Agnolucci P. Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models [J]. Energy Economics, 2009, 31(2): 316-321.
- [17] Kupiec P. Techniques for verifying the accuracy of risk measurement models [J]. Journal of Derivatives, 1995, 3(2): 173-184.
- [18] Engle R F, Manganelli S. CAViaR: Conditional autoregressive Value at Risk by regression quantiles [J]. Journal of Business and Economic Statistics, 2004, 22(3): 367-381.
- [19] Christofferson P F. Elements of financial risk management [M]. San Diego: Academic Press, 2003.
- [20] Artzner P, Delbaen F, Eber J M, et al. Thinking coherently [J]. Risk, 1997, 10(11): 68-71.
- [21] Artzner P, Delbaen F, Eber J M, Heath D. Coherent measures of risk [J]. Mathematical Finance, 1999, 9(3): 203-228.

## Backtesting Risk Models for Chinese Fuel-oil Futures Market

WANG Peng<sup>1</sup>, WEI Yu<sup>2</sup>

(1. School of Finance, Southwest University of Finance and Economics, Chengdu 610074, China;

2. School of Economics and Management, Southwestern Jiaotong University, Chengdu 610031, China)

**Abstract:** By taking four representative indices of Chinese fuel-oil futures market as sample, VaR predicting is used for eight risk models. Furthermore, two robust backtesting methodologies, unconditional coverage test and conditional coverage test, are introduced to estimate the accuracy for VaR predictions produced by different models. The main results show that adding international fuel-oil price volatility as explanatory variable in typical models is helpful to improve risk estimation accuracy of Chinese fuel-oil futures market. In addition, FIGARCH<sub>CST</sub>-SST is moderately good in overall consideration of description efficiency and estimation accuracy to extreme risk.

**Key words:** Chinese fuel-oil futures market; value at risk; expected shortfall; skewed student-t distribution; backtesting analysis