

# 基于分形理论的原油期货高频交易风险测度研究

严伟祥, 卢亚娟

(南京审计大学 金融学院, 江苏 南京 211815)

**[摘要]**为了捕捉原油期货高频波动规律,采用 WTI 原油期货五分钟数据,基于分形理论分别构建 GED 分布和 Skew-t 分布的 FIGARCH、FIAPARCH 和 HYGARCH 模型,分析其波动特征并对风险进行测度。结果显示:三种模型均较好地刻画出 WTI 原油期货波动的长记忆特征;基于 Skew-t 分布的 HYGARCH 模型在度量原油期货高频交易风险时尤为精确;多头与空头头寸的 VaR 呈现非对称性;套期保值者或高频交易者可依据模型预测波动率,防止短期波动率过大导致保证金不足而被强制平仓。高频交易在提高市场流动性和拓宽市场深度方面具有一定的作用,因此,在风险可控的条件下,政府应该鼓励高频交易,促进我国衍生品市场繁荣发展,并增强衍生品市场稳定性和国际竞争力。

**[关键词]**原油期货波动;高频交易风险;分形理论;资本市场;能源期货市场;动态风险;原油投资;金融资产

**[中图分类号]**F830.9 **[文献标识码]**A **[文章编号]**1004-4833(2016)02-0120-09

## 一、引言

随着美联储稳步退出量化宽松货币政策,与美元挂钩的大宗商品价格不断走低,波动幅度日益加剧。尤其,原油期货价格在地缘政治角力、供需失衡、美元升值和高频交易等多重因素冲击下持续下跌。原油作为国家能源战略储备和工业重要原材料,各国政府及企业对油价的走势均高度关注。原油期货市场不仅汇集了全球套期保值的投资资金,还涌入了巨额的投机资金,通过高频交易和杠杆操作,对原油期货价格波动产生不可低估的影响。因此,不论是企业的套期保值需求,还是高频交易的投资行为,人们均需了解原油投资的风险并进行科学的管理。

十八届四中全会明确提出,我国要在 2020 年之前建设完成多层次资本市场。发展金融衍生品市场作为构建多层次资本市场的重要一环,有助于优化我国资本市场结构,提高金融市场的国际竞争力。由于金融衍生品的杠杆交易,量化交易的趋势日益频繁,风险问题也日趋复杂,因此风险度量研究成为金融衍生品市场参与者的一个重要课题。为了规避风险,研究并预测金融衍生品的收益与波动规律并依此规律进行风险管理变得尤为重要。

## 二、文献回顾

自现代金融理论确立以来,众多学者对金融资产的波动是否遵循某种规律进行了大量的研究,并形成相互争辩的流派。早期,大部分金融产品的波动被认为遵循随机游走(random walk)的过程。如 Fama 提出有效市场假说<sup>[1]</sup>,认为市场能够及时、合理地传递信息,资产价格波动是对信息的完全反映。有效市场假说是建立在投资者均为理性,具有相同预期等条件基础之上,是一种均衡状态的表现,但实际资本市场的波动更加复杂,并非均衡。因此, Mandelbrot 提出分形理论(fractal theory),利

**[收稿日期]**2015-05-25

**[基金项目]**教育部人文社会科学研究规划基金(12YJA790093);江苏高校哲学社会科学研究一般项目(2015SJB195);江苏高校人文社会科学校外研究基地项目(08);江苏高校优势学科建设工程项目应用经济学(苏政办发[2014]37号)。

**[作者简介]**严伟祥(1974—),男,安徽安庆人,南京审计大学金融学院讲师,博士,从事风险管理、金融市场研究;卢亚娟(1966—),女,江苏宜兴人,南京审计大学金融学院教授,博士,从事区域金融和金融风险管理研究。

用分形几何解释自然界诸多杂乱无章的现象,通过拟合分数维来确定时间序列的波动结构<sup>[2]</sup>。分形理论揭示了资本市场的非线性特征,包括长记忆性和自相似性。一般来说,通过分形分析可提炼出金融市场波动的重要信息,进而提高资产定价、风险度量和投资决策的执行效果。Granger 和 Joyeux 基于分形理论构建自回归分数移动平均(ARFIMA)模型,证实金融资产的收益时间序列存在长记忆性(long memory)<sup>[3]</sup>。该发现为预测金融资产的走势提供了重要的计量方法,同时表明在外界条件不发生变化的情况下,具有长记忆性的金融时间序列会持续保持该特征。

此后,众多学者纷纷利用分形理论深入和细致地探索并解析时间序列波动的结构特征。在波动率建模方面,尤以 Engle 的自回归条件异方差(ARCH)模型和 Bollerslev 提出的广义自回归条件异方差(GARCH)模型最为突出<sup>[4-5]</sup>,虽然两种模型无法拟合金融时间序列波动所呈现的长记忆特征,但却成为后来学者研究波动率的基石。为了研究波动率的长记忆特征,Baillie 等首次将分形理论融入 GARCH 模型,提出分形协整广义自回归条件异方差(FIGARCH)模型,检测出德国马克兑美元汇率的波动具有非线性分形结构<sup>[6]</sup>。与前面三位学者不同,Tse 运用分形协整非对称自回归条件异方差(FI-APARCH)模型研究汇率波动时,确认了汇率波动过程中呈现长记忆性、非对称性和波动聚类特征<sup>[7]</sup>。Davidson 将 FIGARCH 模型拓展成双曲线 HYGARCH(Hyperbolic GARCH)模型,同样捕捉到金融时间序列的波动率具有长记忆性、杠杆效果和异方差特性<sup>[8]</sup>。另外,HYGARCH 模型克服了收益率的四阶矩不存在的问题,并且验证了序列的记忆长度随着参数的增大而增加。

至今,学者们利用上述三种分形理论模型对各类金融资产进行了广泛地研究。华仁海和陈百助运用 ARFIMA 模型分析我国铜、铝、橡胶、大豆和小麦五种期货收益的长记忆性,结果发现除了铜和铝以外,其他品种的期货价格收益均具有长记忆性<sup>[9]</sup>。该结果反映出我国期货市场处在尚未成熟的阶段。肖智等利用 FIGARCH 模型分析上证综合指数时,得出其分数协整参数为 0.48,确认了其具有长记忆性<sup>[10]</sup>。林宇在研究深沪两市、美国道琼斯工业指数和标准普尔 500 指数的 VaR 时,采用 ARFIMA-HYGARCH 模型检验出四种指数的收益波动率呈现长记忆性,同时指出该模型也是较好的风险度量方法<sup>[11]</sup>。许林等指出 HYGARCH-VaR 模型在度量基金投资风格漂移风险方面更加精确<sup>[12]</sup>。Sang 和 Seong 运用 GARCH、IGARCH 和 FIGARCH 三种模型在正态分布条件下,度量并预测了 WTI 石油期货的波动,但并未得出哪一种模型有精确测度优势的结论<sup>[13]</sup>。淳伟德等也认为 HYGARCH 模型不仅准确地反映了能源市场的波动特征,而且在 VaR 测度方面也表现出较强的可靠性<sup>[14]</sup>。

虽然国内外学者利用分形理论在研究金融资产的波动规律方面取得了一定的成就,但是对原油期货高频交易的收益和波动的长记忆性研究鲜有涉及,尤其在高频交易占原油期货市场日交易量 50% 以上的特定环境下,受众多因素冲击,原油期货的收益和波动是否为分形结构,这对套期保值者和投机者均具有重要投资决策价值。金融产品的交易每分每秒都在进行,而以往的研究大多利用低频数据,分析中可能会遗漏的大量市场信息,容易造成研究结果出现偏误。为了更精确地研究原油期货市场微观结构特征及其内在联系,减少信息损失,本文采用原油期货五分钟高频数据,通过 FIGARCH、FIAPARCH 和 HYGARCH 模型分析其收益与波动是否具有长记忆性,并在此基础上进行 VaR 度量,比较和检验各模型风险测度的有效性,为规避原油期货高频交易风险提供科学依据。

### 三、方法与模型

#### (一) 拟合长记忆性的模型

时间序列的记忆长短程度通常依据其自相关函数的衰减速度来衡量。当自相关函数呈指数速率衰减时,则认为该时间序列为短记忆;如果自相关函数呈负幂指数速度衰减,则可认为其具有长记忆性。长记忆性最早是由 Hurst 在研究尼罗河区域降水量时发现并提出用 R/S 分析方法进行检验<sup>[15]</sup>。由于 R/S 为非参数检验法,在实际检测中有一定的局限性。Granger 和 Joyeux 基于分形理论提出的 ARFIMA 模型使后人在实证过程中估算长记忆性参数变得更加灵活,而且获取了金融时间序列存在

长记忆性的强有力证据(Lobato and Savin)<sup>[16]</sup>。ARFIMA ( $m, d_1, n$ ) 模型一般形式如下:

$$\Phi L(1-L)^{d_1}(r_t - \mu) = \Theta(L)\varepsilon_t \quad (1)$$

其中,  $r_t$  为时间序列,  $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_m L^m$ ,  $\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_n L^n$  分别为自回归滞后  $m$  阶算子和移动平均滞后  $n$  阶算子,  $(1-L)^{d_1}$  为分数差分算子,  $\mu$  是时间序列  $r_t$  的均值, 假定误差项  $\varepsilon_t \sim i. i. d. N(0, \sigma_t^2)$ 。

时间序列是否具有长记忆性是由参数  $d_1$  的大小决定。如果  $0 < d_1 < 0.5$ , 则该时间序列为长记忆平稳过程; 如果  $-0.5 < d_1 < 0$ , 则该时间序列为短记忆平稳过程。短记忆平稳过程的经济意义是指金融资产受市场信息等变量冲击后, 迅速做出反应并吸收, 价格及时得以调整并达到均衡状态, 该影响不会持续很长时间, 这是成熟市场所具备的特质。长记忆平稳过程与短记忆过程正好相反, 金融资产受外界诸多因素冲击后, 影响的衰减速度非常缓慢。因此, 具有长记忆性的金融资产的风险比短记忆性的资产风险大。另外, 如果  $d_1 = 0$ , 则 ARFIMA 模型退化为 ARMA( $m, n$ ) 模型; 如果  $d_1 = 1$ , 则 ARFIMA 模型演变为 ARIMA( $m, 1, n$ ) 模型, 这两种模型均无法刻画市场的分形特征。

众所周知, GARCH 模型已被广泛地应用于拟合时间序列的波动率, 但它无法解决波动率的长记忆性问题。为此, Baillie 等提出 FIGARCH( $p, d_2, q$ ) 模型来捕捉波动率的长记忆性<sup>[6]</sup>, 模型如下:

$$\sigma_t^2 = \omega[1 - \beta(L)]^{-1} + \{1 - [1 - \beta(L)]^{-1}\phi(L)(1-L)^{d_2}\}\varepsilon_t^2 \quad (2)$$

其中,  $\sigma_t^2$  为时间序列的条件方差,  $\omega > 0$ ,  $\phi(L)$  和  $\beta(L)$  为滞后算子多项式。 $\phi(L)$  和  $\beta(L)$  系数可以捕捉短期内的动态波动, 如果  $d_2 > 0$ , 则表明时间序列的波动率具有长记忆性, 即该资产的波动率在信息冲击下, 无法及时消化(大幅波动不会衰减), 预示存在巨大风险。Davidson 认为 FIGARCH 模型存在假设条件的限制, 且无法对参数大小做出解释<sup>[8]</sup>。因此, 将 FIGARCH 模型扩展为双曲线 HYGARCH 模型, 该模型如下:

$$\sigma_t^2 = \omega[1 - \beta(L)]^{-1} + \{1 - [1 - \beta(L)]^{-1}\phi(L)[1 + (\alpha(1-L)^{d_2} - 1)]\}\varepsilon_t^2 \quad (3)$$

当公式(3)中  $\alpha = 0$  的时候, 则 HYGARCH 模型变成了 GARCH 模型, 如果  $\alpha = 1$ , 则该模型退化成了 FIGARCH 模型。该模型能够刻画时间序列记忆长度, 且该长度随着参数  $d_2$  的增大而增加。 $d_2$  值越大, 显示金融资产的波动越难以衰减。通过建模识别资产的长记忆程度, 才能选择期限相匹配的金融工具进行风险对冲。FIAPARCH 模型刻画分形市场的波动聚类 and 波动非对称性问题, 其函数表达式为:

$$\sigma_t^\delta = \omega + \{1 - [1 - \beta(L)]^{-1}\phi(L)(1-L)^{d_2}\}(|\varepsilon_t| - \gamma\varepsilon_t)^\delta \quad (4)$$

其中,  $\gamma$  为杠杆系数, 如果  $\gamma > 0$ , 就表明前一期的负收益率导致当期有更高的收益波动, 即风险越大。参数  $d_2$  和前面模型相同, 如果大于零, 表明时间序列的波动率呈现长记忆性。

公式(2)、公式(3)和公式(4)中  $\varepsilon_t = z_t \sigma_t$ , 在估计 GARCH 族模型时, 通常假定误差  $\varepsilon_t$  服从正态分布进行拟合, 但正态分布不能刻画金融资产收益分布的尖峰、厚尾(fat tail)特征。在处理尖峰、厚尾的金融时间序列时, 广义误差分布(GED)被认为具有明显优势, GED 分布的密度函数为:

$$f_{(GED)}(z_t; v) = \frac{v \cdot \exp(-\frac{1}{2} |z_t/\lambda_v|^v)}{\lambda_v 2^{(1+\frac{1}{v})} \Gamma(1/v)}, v > 0, \lambda_v = \sqrt{\frac{\Gamma(1/v) 2^{(-2/v)}}{\Gamma(3/v)}} \quad (5)$$

当  $v = 2$  时,  $z_t$  为标准正态分布。当  $v < 2$  时,  $z_t$  为厚尾分布。通过模型(5)拟合出  $v$  的大小, 我们可以准确地判定分布, 避免因误判导致人为地增加风险。金融资产的收益有时也会呈现偏态学生分布(Skew-t), 偏态学生分布的密度函数为:

$$f_{(Skew-t)}(z_t; v, \xi) = \frac{\Gamma((v+1)/2)}{\Gamma(v/2) \sqrt{\pi(v-2)}} \left(\frac{2s}{\xi + 1/\xi}\right) \left(1 + \frac{sz_t + m}{v-2} \xi^{-2t}\right)^{-(v+1)/2}, v > 2 \quad (6)$$

其中,

$$I_t = \begin{cases} 1 & \text{if } z_t \geq -m/s \\ -1 & \text{if } z_t < -m/s \end{cases}, m = \frac{\Gamma(v+1)/2 \sqrt{v-2}}{\sqrt{\pi}\Gamma(v/2)} \left( \xi - \frac{1}{\xi} \right), s = \sqrt{\left( \xi + \frac{1}{\xi} - 1 \right) - m^2}$$

$v$  的大小表示分布的厚尾程度。 $\xi$  为非对称参数, 当  $\xi = 1$  或者  $\ln(\xi) = 0$  时表示左右对称的  $t$  分布, 说明多空头寸的风险相同; 当  $\xi > 1$  或者  $\ln(\xi) > 0$  时表示分布的右尾比左尾厚, 即空头头寸的风险比多头头寸的风险大; 相反, 当  $\xi < 1$  或者  $\ln(\xi) < 0$  时表示分布的左尾比右尾厚, 此时持有多头头寸的风险显著高于空头头寸。衍生品市场均为双向交易, 不同头寸持有者的风险分布不同, 需要精确地刻画, 才能有的放矢地进行风险管理。为了精确地拟合误差项, 本文利用 GED 分布和 Skew-t 分布, 基于极大似然估计 (MLE) 来估计各模型参数。

## (二) 风险测度方法

当前, VaR (Value at Risk, 在险价值) 是巴塞尔协议中指定的市场风险管理指标, 是全球金融机构和市场参与者均采用的重要风险测度指标。VaR 是指在给定的置信水平下, 金融资产在未来特定时期内的最大可能损失。VaR 的由来及计算原理在诸多论著中都有阐述, 本文就不再赘述。由于原油期货合约存在多空两种头寸, 如果金融资产其波动具有非对称性, 持有原油期货多头头寸和持有原油期货空头头寸所面临的市场风险理应不同。为此, 本文定义持有多头头寸的 VaR 测度方法为:

$$VaR_{long} = \mu_t - z_q \sigma_t \quad (7)$$

同理, 持有空头头寸的 VaR 的度量方法为:

$$VaR_{short} = \mu_t + z_{1-q} \sigma_t \quad (8)$$

其中,  $z_q$  和  $z_{1-q}$  分别为左尾和右尾分位数。不难看出, 多头头寸的 VaR 实际为收益分布的左尾风险, 空头头寸的 VaR 为收益分布的右尾风险。本文基于分形理论的 GARCH 模型估算  $\mu_t$  和  $\sigma_t$ , 然后将其代入上式即可求出多头和空头的 VaR。对 VaR 进行测度, 可以帮助套期保值的企业或者其它投资者有效管理风险资本金, 实现最大收益。

## (三) 风险测度的准确性检验

为了检验上述各类模型度量 VaR 的精度, 我们采用 Kupiec 的似然比 (LR) 检验法来回测其准确性<sup>[17]</sup>。假定计算的置信水平为  $\alpha$ , 失败的期望概率为  $p^* = 1 - \alpha$ , 回测检验的观测样本为  $T$ ,  $N$  为回测检验的实际失败次数, 则失败频率为  $p = N/T$ 。如果风险测度模型有效,  $N$  是符合  $B(T, p)$  二项分布的。这样, 评估 VaR 模型准确性问题就转化为检验失败频率  $p$  是否显著不同  $p^*$ 。Kupiec 的似然比检验统计量为:

$$LR = -2 \ln[(1 - p^*)^{T-N} (p^*)^N] + 2 \ln[(1 - N/T)^{T-N} (N/T)^N] \quad (9)$$

如果零假设成立, LR 统计量服从  $\chi^2(1)$  分布。当  $LR > 3.84$ , 拒绝零假设。似然比统计量  $P$  值越小, 零假设成立的概率越低, 即拒绝零假设的成立。相反如果  $P$  值越大, 说明构建的风险测度模型有效。

## 四、数据选取及基本统计

原油期货以纽约商品交易所 (NYMEX) 的西德克萨斯原油期货 (WTI) 交易最为活跃, 是全球原油期货价格的风向标之一。该品种包括公开喊价和电子交易, 除了周六, 基本上每天 24 小时均在交易, 价格变化非常频繁。本文从 Wind 数据库中提取 2015 年 1 月 1 日至 2015 年 2 月 28 日 WTI 原油期货交易的五分钟连续数据, 样本容量为 10536 个。本文的统计分析在 OxMetrics6 软件上完成。

WTI 原油期货的五分钟收益率可以通过公式 (10) 计算得到:

$$r_t = \ln P_t - \ln P_{t-1} \quad (10)$$

其中,  $P_t$  为  $t$  期原油期货合约的价格,  $P_{t-1}$  为  $t-1$  期的原油期货合约的价格。WTI 原油期货的收益基本描述统计如表 1 所示。

表1显示,WTI原油期货合约五分钟最大收益率为2.8556%,最小收益率为-1.6124%,可知其在五分钟内的波动幅度非常大,但平均收益接近零。从偏度和峰度的数值来看,原油期货价格明显呈现出尖峰的特征。

表1 WTI原油期货收益率的描述性统计量及单位根检验

平均值	最大值	最小值	标准差	偏度	峰度	ADF
-0.0001	2.8556	-1.6124	0.0663	0.0968	33.0928	-135.59***

注:\*\*\*表示在1%的水平上显著。

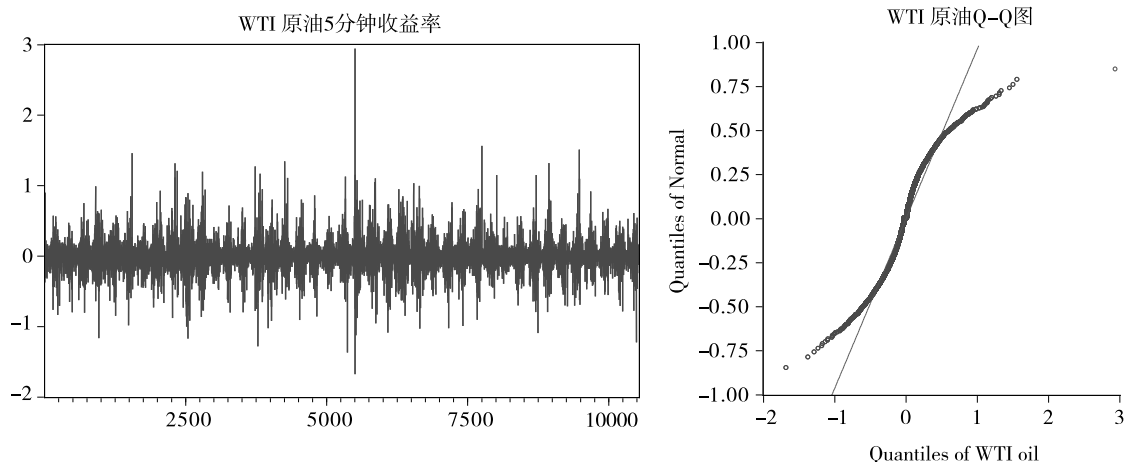


图1 原油期货5分钟收益走势及Q-Q图

从图1(右图)中可见,原油期货的收益率序列并没有集中散布在正态直线上,即表明拒绝正态分布的原假设,两侧均有厚尾现象,且左尾比右尾厚,右尾比较长,因此基于正态分布假设下的标准差估算风险存在较大偏差。由图1(左图)中可以看出,原油期货收益率序列波动具有明显的时变性,且不同时期波动性的大小也不相同,同时波动出现聚集现象,从而印证了原油期货的收益率序列存在条件异方差。鉴于此,利用GARCH族三种模型估算标准差十分必要。同时为了避免伪回归,需要对原油期货收益序列进行单位根检验。表1中的ADF单位根检验结果显示在1%显著水平上,WTI原油的五分钟收益率均拒绝存在单位根的零假设。

## 五、实证分析与结果

### (一) 模型参数估计

为了精确度量VaR,同时考虑WTI原油期货的五分钟收益序列的尖峰厚尾特征,我们采用GED分布和Skew-t分布刻画FIGARCH、FIAPARCH和HYGARCH模型,在此基础上估算波动率并进行比较。在建模之前,依据AIC信息准则,经过反复测试,判断方差方程分别为FIGARCH(1,  $d_2$ , 1)、FIAPARCH(1,  $d_2$ , 1)和HYGARCH(1,  $d_2$ , 1)比较合适,均值方程采用ARFIMA(1,  $d_1$ , 1)。基于极大似然估计(MLE)的各方程参数结果如表2所示。

表2显示WTI原油期货GED分布三个模型的 $d_1$ 分别为-0.0815、-0.0830和-0.0813,Skew-t分布的 $d_1$ 分别为-0.0588、-0.0556和-0.0589,且都在1%的水平上显著。因为 $-0.5 < d_1 < 0$ ,所以可以判定WTI原油期货的收益率为短记忆平稳过程。原油收益率序列的短记忆性体现了期货交易具有双向性的特征,只要投资者对走势预判准确,就可通过建立长短头寸获利。从另一个角度来说,纽约原油期货市场是高度成熟的市场,其合约价格能快速吸收各类信息冲击,继而趋向平衡。

方差方程中,波动率长记忆性参数 $d_2$ 均大于0.5,并且都在1%的水平上显著,表明WTI原油期货合约的波动率为长记忆特征,即原油期货市场呈非线性,为分形结构。进一步可知,原油期货波动序列的绝对值或幂的自相关呈现十分缓慢地衰减,相距较远的时间间隔仍然具有显著的自相关性。

由此可见,只要原油市场供给关系、经济环境不发生根本转变,当前原油期货的波动就将持续影响未来,套利行为成为可能。FIAPARCH 模型中的  $\gamma$  系数分别为 0.1122 和 0.03,均大于零,表明原油期货价格前一期的大幅下跌导致当期有更高的收益波动,更进一步凸显了原油期货的风险。Skew - t 分布模型的  $\ln\xi$  均大于零,表示原油期货合约的多头面临的风险比空头大,这与原油期货合约的价格不断下跌,多头容易发生损失是高度吻合的。

表 2 模型参数估计结果

	FIGARCH		FIAPARCH		HYGARCH	
	GED	Skew - t	GED	Skew - t	GED	Skew - t
$Cst(M) \times 10^4$	-1.45 (-1.26)	0.80 (0.75)	-3.19* (-1.68)	-1.17 (-0.98)	-1.46 (-1.23)	0.75 (0.71)
$d - Arfima(d_1)$	-0.0815*** (-8.96)	-0.0588*** (-11.25)	-0.0830*** (-8.72)	-0.0556*** (-5.49)	-0.0813*** (-12.53)	-0.0589*** (-13.22)
AR(1)	0.5247*** (5.28)	0.5162** (2.24)	0.5261* (1.89)	0.5479*** (3.47)	0.5235* (1.92)	0.5144** (2.38)
MA(1)	-0.5495** (-2.40)	-0.5352** (-2.54)	-0.5499* (-1.91)	-0.5689*** (-5.21)	-0.5485*** (-6.45)	-0.5334** (-2.55)
$Cst(V) \times 10^4(\omega)$	0.5571*** (3.77)	1.3787*** (4.62)	1.569*** (4.82)	4.9490** (5.11)	0.5811*** (5.23)	1.2497*** (9.76)
$d-Figarch(d_2)$	0.7672*** (10.25)	0.5641*** (11.45)	0.7332*** (9.85)	1.1341*** (9.93)	0.7747*** (12.37)	0.5445*** (11.89)
ARCH( $\phi$ )	0.2799*** (4.65)	0.2893*** (5.10)	0.2836*** (4.76)	0.0126* (1.84)	0.2737*** (6.01)	0.2966*** (6.17)
GARCH( $\beta$ )	0.7801*** (5.11)	0.6359*** (4.92)	0.7758*** (6.21)	0.9277*** (6.35)	0.7829*** (5.71)	0.6281*** (5.13)
APARCH( $\gamma$ )			0.1122*** (4.88)	0.0300* (1.87)		
APARCH( $\delta$ )			1.7435*** (8.46)	1.2626*** (8.24)		
G. E. D.	1.1867*** (25.62)		1.1954*** (26.71)		1.188*** (21.22)	
Asymmetry( $\ln\xi$ )		0.0119 (1.04)		0.0121 (0.95)		0.0120 (0.99)
Tail		3.7482*** (5.76)		3.9256*** (6.29)		3.6581*** (6.14)
$\ln\alpha$					-0.0058 (-1.23)	0.0275 (1.37)
AIC	-3.16	-3.28	-2.97	-3.17	-3.02	-3.31

注:括号内为 t 值;Cst(M) 为均值方程的常数项,Cst(V) 为方差方程的常数项;括号内为 t 值;\*、\*\*和\*\*\* 分别表示在 10%、5% 和 1% 水平上显著。

### (二) 模型残差检验

为了检验 ARFIMA(1,  $d_1$ , 1) - FIGARCH(1,  $d_2$ , 1)、ARFIMA(1,  $d_1$ , 1) - FIAPARCH(1,  $d_2$ , 1)、ARFIMA(1,  $d_1$ , 1) - HYGARCH(1,  $d_2$ , 1) 三个模型的残差项是否还存在自相关性和异方差效应,我们采用 Ljung - Box 统计量和 ARCH - LM 检验。由于实证分析是五分钟高频数据,所以采用滞后 10 阶和 50 阶的残差来进行分析,检验结果如表 3 所示。

从表 3 的检验结果来看,WTI 原油期货合约的三个模型在残差滞后 10、50 阶是在显著水平 1% 下,不能拒绝存在不相关的零假设,但各模型的残差平方都不存在相关性。另外,拟合后的模型残差也都消除了 ARCH 效应,表明三个模型的 GED 分布和 Skew - t 分布能够较好地拟合了 WTI 原油期货的波动率的长记忆特征。

### (三) VaR 回测检验

本文利用公式(2)、公式(3)和公式(4)分别估算出在 GED 和 Skew - t 分布下的均值和波动率,然后代进公式(7)计算多头头寸 VaR 和公式(8) 计算空头头寸 VaR,由于样本容量巨大,无法一一列

出 VaR 值。为了检验所构建的风险模型的有效性,下面对不同模型计算出的 VaR 分别进行 Kupiec 回测检验。我们比较各模型在 95%、97.5%、99%、99.5% 和 99.75% 五种置信水平下,未来五分钟 VaR 值的准确性。

表 3 残差检验结果

	FIGARCH		FIAPARCH		HYGARCH	
	GED	Skew - t	GED	Skew - t	GED	Skew - t
Q(10)	24.2458 [0.0000]	98.4753 [0.0000]	106.016 [0.0000]	109.949 [0.0000]	200.876 [0.0000]	98.6648 [0.0000]
Q(50)	81.9771 [0.0025]	111.610 [0.0000]	117.979 [0.0000]	122.558 [0.0000]	219.801 [0.0000]	111.803 [0.0000]
Q2(10)	0.8957 [0.9986]	0.8882 [0.8282]	1.6411 [0.6501]	1.6295 [0.6527]	0.8348 [0.8411]	0.9009 [0.8252]
Q2(50)	37.8945 [0.7581]	2.5201 [0.9607]	3.1343 [0.9256]	3.1369 [0.9255]	2.4418 [0.9644]	2.5231 [0.9606]
ARCH(10)	0.0862 [0.9979]	0.1779 [0.9710]	0.3280 [0.8964]	0.3256 [0.8978]	0.1670 [0.9747]	0.1804 [0.9701]
ARCH(50)	0.5217 [0.8769]	0.2502 [0.9908]	0.3085 [0.9794]	0.3087 [0.9793]	0.2429 [0.9919]	0.2505 [0.9908]

注:中括号内数值为 P 值。

从表 4 回测检验结果可知,不论是空头还是多头,FIGARCH、FIAPARCH 和 HYGARCH 模型 GED 分布的 P 值均比 Skew - t 分布的 P 值小,表明基于 GED 分布的模型拟合风险测度较差,Skew - t 分布能较好地刻画原油期货的风险。又知 HYGARCH 的 Skew - t 分布的 P 值普遍高于 FIGARCH 和 FIAPARCH 模型,这进一步说明双曲线 HYGARCH - Skew - t 模型不仅验证了原油期货市场风险的非线性,而且在度量高频交易风险中存在较大的优势。

表 4 Kupiec 检验结果

	FIGARCH		FIAPARCH		HYGARCH		
	GED	Skew - t	GED	Skew - t	GED	Skew - t	
short position	95%	0.0000	0.0006	0.0000	0.0749	0.0000	0.2009
	97.5%	0.0000	0.0814	0.0986	0.2147	0.0002	0.3296
	99%	0.8294	0.6795	0.0143	0.0472	0.6722	0.7317
	99.5%	0.0359	0.6237	0.0002	0.2085	0.0225	0.6959
	99.75%	0.0000	0.6173	0.0000	0.0616	0.0000	0.6773
long position	5%	0.0000	0.5129	0.0000	0.9114	0.0000	0.5337
	2.5%	0.0078	0.6820	0.0068	0.4953	0.0133	0.6185
	1%	0.2982	0.1677	0.5748	0.0403	0.2677	0.3479
	0.5%	0.0081	0.1481	0.0062	0.0449	0.0062	0.1485
	0.25%	0.0000	0.0005	0.0000	0.0000	0.0000	0.0005

注:表中数据为 Kupiec 检验的 P 值,P 值越大表明风险度量越准确。

从 HYGARCH 模型 Skew - t 分布的 P 值大小可知,空头头寸的 VaR 在置信水平 99% 以上时,比较有效,但在 95% 的置信水平上,VaR 未较好地刻画风险,此处存在厚尾的可能性非常大。因此,空头头寸持有者需要准备较多的风险资金应对不可预知的波动所造成的损失。多头头寸的 VaR 在 95% (1 - 5%) 置信的水平上较好地捕捉到风险,但在 99.5% 和 99.75% 的置信水平上,无法准确度量风险。这和原油期货的收益序列的 Q - Q 图中得知左尾较厚,右尾较长一致,表明其风险为非对称性。鉴于此,当多头或空头在调整持仓头寸后,出现多翻空或空翻多时,头寸持有者不能按照先前的模型进行风险度量,必须及时变更置信水平重新度量风险,才能满足有效控制风险的目的。

## 六、高频交易的思考

原油价格的形成机制十分复杂,一般认为原油供求关系是影响其价格的最直接和最关键因素,同时其他各种不确定因素在短期内也会冲击原油价格。近年来,全球经济增速放缓导致各国对原油需

求逐渐减少,另外,OPEC 成员国保持原油供应不变,美国页岩气开发技术的提升等因素的叠加导致原油供需失衡,促使原油现货预期价格不断走低。衍生品市场具有价格发现功能,原油现货预期价格是通过期货价格表现出来,期货价格的波动又会传递至现货市场,影响标的资产的价格走势。

衍生品的交易,尤其是交易所内期货合约的交易方式比较接近,不论是纽约商品交易所还是我国上海等地的期货交易所,从原油、铜等工业原材料到大豆、玉米等农产品的期货合约都是标准化交易,其价格走势和波动均会受到宏观经济、供需变化等因素的冲击。正是期货交易的相似性,不同标的资产的期货合约同样可以通过分形理论构筑 FIGARCH、FIAPARCH 和 HYGARCH 模型对其高频交易风险进行检测,根据风险测度的结果调整交易策略和管理头寸。高频交易是期货市场中不可或缺的交易行为,是实现衍生品投机功能的一种重要手段。它不仅可以为市场和套期保值者提供流动性,增加市场活力,而且具有纠正衍生品价格扭曲的作用。高频交易的出现是技术进步所引起的金融市场演变的结果,在众多影响市场变革的因素中,技术进步无疑最为深远。技术的进步改变了信息传播的途径,加快了市场参与者之间的沟通速度,使市场由无效变为有效。高频交易者利用计算机程序拟合模型,对市场海量数据、信息进行整理、统计和分析,找到市场无效率的瞬间实现获利。市场无效率的短暂出现是由于投资者的非理性行为和市场对信息的反应滞后或者反应过度,从而造成市场失效,高频交易的积极交易行为会促使市场无效率向效率市场转变。由于高频交易投机者逐利的共性,容易诱发金融市场大幅波动,在向效率市场转变过程中出现矫枉过正,再次造成期货价格严重偏离标的资产的真实估值水平。虽然高频交易者利用分形理论构建模型控制自身风险,但是很有可能将风险转嫁至其他投资者或市场。

随着科技的普及和推广,高频交易让市场微观结构变得更加复杂,同时在资本市场中的作用也越来越明显。美国证券交易委员会(SEC)在应对资本市场的科技创新时,不断加强高频交易的监控,防止市场出现闪电崩盘(flash crash)。2015年6月中旬以来,我国沪深股市出现了暴跌,恶意做空股指期货的行为被认为是导致本次股灾的原因之一。恶意做空往往会借助高频交易或程序化交易等手段达到获利的目的。因此,有关部门需要对高频交易风险进行全面认识,提高对高频交易的监管,防止高频交易中的程序化出错或恶意高频报价行为,引发恐慌造成市场交易失序。监管和限制高频交易者的持仓数量,可以削弱冲击市场的力度,阻断衍生市场的恐慌传染到现货市场,从而达到避免引发局部系统金融危机。

在控制好风险的前提下,我们认为应该鼓励高频交易,通过高频交易可以提高市场流动性和拓宽市场深度,繁荣我国衍生品市场。只有做好风险控制,才能稳步发展我国衍生品市场,继而提高我国衍生品市场的国际竞争力和掌握大宗商品定价话语权。

## 七、结论

原油作为工业重要原料,在套期保值、套利和投机的多方博弈下,原油期货波动愈来愈大,风险日益突出。实证分析发现,原油期货的收益序列为短记忆平稳过程,这证实了期货市场的双向性,只要顺势建立多空净头寸即可获利;其波动率序列呈长记忆特征,表明原油期货市场为分形结构,即原油期货高频波动具有非线性;进而发现基于双曲线 HYGARCH - Skew - t 模型在度量风险中具有较大的潜力。

鉴于以上发现,企业在进行套期保值时,需要时刻提防原油期货短期波动过大,导致原油期货头寸保证金低于维持保证金,一旦未及时追加保证金,头寸将被强制平仓,从而导致套期保值失败。因此,企业可以依据 FIGARCH、FIAPARCH 和 HYGARCH 模型预测波动率,对保证金账户的资金额度进行科学和有效的管理。另外,高频交易投资者应该积极搜集国际局势、供求关系和宏观数据等信息,并做出科学预判,及时调整原油期货的多空头寸。当头寸出现空翻多或多翻空时,投资者可以运用双曲线 HYGARCH 模型预测原油期货波动率并进行风险度量,达到有效控制投资风险的目标。

我们认为在鼓励高频交易的同时,为了切实防止高频交易行为冲击金融市场,监管部门应该对高



频交易进行相应的限仓,对高频交易行为进行跟踪审计,抑制高频交易的过度投机行为,保证我国衍生品乃至金融市场的安全稳定。

参考文献:

- [1] Fama E F. Efficient capital markets: a review of theory and empirical work[J]. the Journal of Finance,1970,25(2):383-417.
- [2] Mandelbrot B. How long Is the coast of Britain? statistical self-similarity and fractional dimension [J]. Science,1967,156(5):636-638.
- [3] Granger J, Joyeux R. An introduction to long-memory time series models and fractional differencing[J]. Journal of Time Series Analysis, 1980,1(1):15-29.
- [4] Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation[J]. Econometrica,1982, 50(4):987-1007.
- [5] Bollerslev T. Generalized autoregressive conditional heteroskedasticity[J]. Journal of Econometrics,1986,31(3):307-327.
- [6] Baillie R T, Bollerslev T, Mikkelsen H. Fractionally integrated generalized autoregressive conditional heteroskedasticity[J]. Journal of Econometrics,1996,74(1):3-30.
- [7] Tse Y K. The conditional heteroscedasticity of the yen-dollar exchange rate[J]. Journal of Applied Econometrics,1998,13(1):49-55.
- [8] Davidson J. Moment and memory properties of linear conditional heteroscedasticity models, and a new model[J]. Journal of Business and Economics Statistics,2004,22(1):16-29.
- [9] 华仁海,陈百助. 我国期货市场期货价格收益及波动方差的长记忆性研究[J]. 金融研究,2004(2):52-61.
- [10] 肖智,傅肖肖,钟波. 基于 EVT-BM-FIGARCH 的动态 VaR 风险测度[J]. 中国管理科学,2008(4):18-23.
- [11] 林宇. 基于双曲线记忆 HYGARCH 模型的动态风险 VaR 测度能力研究[J]. 中国管理科学,2011(6):15-24.
- [12] 许林,宋光辉,郭文伟. 基于 SKT-ARFIMA-HYGARCH-VaR 模型的股票型基金投资风格漂移风险测度研究[J]. 中国管理科学, 2011(5):10-20.
- [13] Sang H K, Seong M Y. Modeling and forecasting the volatility of petroleum futures prices[J]. Energy Economics,2013(36):354-362.
- [14] 淳伟德,张德园,林宇. 能源期货市场动态极端风险测度研究[J]. 投资研究,2014(7):110-125.
- [15] Hurst H E. Long term storage capacity of reservoirs[J]. Transactions of the American Society of Civil Engineers,1951,116(3):770-799.
- [16] Lobato I N, Savin N E. Real and spurious long-memory properties of stock-market data[J]. Journal of Business & Economic Statistics, 1998,16(3):261-268.
- [17] Kupiec P H. Techniques for verifying the accuracy of risk measurement models [J]. Journal of Derivatives,1995,3(2):73-84.

[责任编辑:杨志辉]

## A Study on Risk Measurement of High-frequency Traded Crude Oil Futures Based on Fractal Theory

YAN Weixiang, LU Yajuan

(School of Finance, Nanjing Audit University, Nanjing 211815, China)

**Abstract:** In order to capture the characteristics of crude oil futures, and based on the WTI crude oil futures five minutes high frequency data to construct the GED distribution and Skew-t distribution of FIGARCH, FIAPARCH and HYGARCH model, this paper tried to analyze the volatility characteristics and to measure risk by using the fractal theory. The results show that three kinds of models that better depict the WTI crude oil futures volatility has a long memory feature; HYGARCH model with Skew-t distribution in the risk measurement of crude oil futures is demonstrated precisely. The VaR of long and short positions are asymmetric. Hedgers or high-frequency traders can forecast the volatility based on three models, prevent the short-term volatility broaden abruptly, thus leading to insufficient margin to be liquidated. High frequency trading has a certain role in improving market liquidity and expanding market depth. Therefore, in the condition of controllable risk, it should encourage high-frequency trading, which can improve the boom of derivatives market, enhance the stability and the international competitiveness of derivatives market.

**Key Words:** volatility of crude oil future; high-frequency trading risks; fractal theory; capital market; energy future market; VaR measurement; dynamic risks; crude oil investment; financial assets