

doi: 10.19920/j.cnki.jmsc.2024.01.008

谁是驱动中国原油期货价格波动的关键信息?^①

马嫣然¹, 吴菲², 张大永², 姬强^{1*}

(1. 中国科学院科技战略咨询研究院, 北京 100190; 2. 西南财经大学经济与管理研究院, 成都 610074)

摘要: 本文首次对影响中国原油期货价格波动的驱动因素进行了量化分析. 运用广义动态因子模型, 结合互联网数据, 为中国原油期货价格构造了六类预测因子: 供需预测因子、市场金融化预测因子、汇率市场信息预测因子、商品市场预测因子、全球宏观经济预测因子以及事件预测因子. 基于混频 GARCH-MIDAS 模型, 本文发现上述六类因子能显著改善对原油价格波动的预测精度. 同时, 基于 MCS 检验结果, 揭示出在不同时间尺度下, 驱动中国原油价格波动的信息存在明显差异性, 即在短期和中期预测中事件预测因子起主导作用, 而供需因子则是在长期主导中国原油价格波动的关键因素. 综上, 本研究为国内原油市场参与者、政策制定者及市场监管者把握未来市场信息提供了分析工具和参考依据.

关键词: 中国原油期货; 波动预测; 预测因子; GARCH-MIDAS

中图分类号: F724.5; F224 **文献标识码:** A **文章编号:** 1007-9807(2024)01-0113-13

0 引言

2018 年 3 月中国原油期货(INE)市场作为中国首个对外开放的交易平台正式上市, 这对促进中国经济活力和推进人民币国际化具有重要作用. 近年来, 国内外市场黑天鹅事件频发, 国际原油期货市场价格波动, 直接导致全球多国股票市场触发熔断机制, 全球宏观经济不确定性明显增加. 面对国际市场价格高波动和风险传染的情况, 考虑到中国原油期货市场监管机制、风险防范机制以及预警机制方面的建设还不完善, 如何在中国原油期货市场快速发展的同时防控风险、维持价格稳定, 成为交易所面临的重要挑战. 因此, 本文以中国原油期货市场为研究对象, 旨在识别并分析不同时间尺度下影响价格波动规律的驱动因素, 为国内监管部门和投资者把握中国原油期货市场价格波动、规避市场风险提供理论指导.

目前原油价格波动预测的研究主要聚焦于 WTI 市场和 BRENT 市场, 而关于 INE 原油价格波动预测的量化研究基本处于空白状态. 通过梳理相关文献, 本文将预测油价波动的方法分为两大类. 第一类是基于原油价格历史信息预测. 学者们主要基于历史油价数据(包括高频数据或低频数据), 利用 GARCH 族类模型^[1,2]、HAR 类模型^[3,4]、MIDAS 模型^[5]等计量工具, 描述并预测原油价格波动. 这类预测对数据要求比较低, 但研究证实原油价格受一组动态且复杂多变的因素驱动, 历史价格信息不足以全面反映原油价格变化情况^[6]. 因此, 第二类预测方法在模型中引入了预测因子来提高预测精度. 这类研究重在筛选影响原油价格波动的因素, 建立其与油价波动的函数关系. 归纳文献发现, 主要因素包括: 全球原油供需基本面信息、宏观经济环境、市场金融化因素、汇率变动、地缘政治等^[7-12]. 但目前大多数研究仅聚焦于某类信息对油价波动的预测能力, 而

① 收稿日期: 2021-01-04; 修订日期: 2021-10-17.

基金项目: 国家社会科学基金资助重大项目(23&ZD093); 国家自然科学基金资助项目(71974181; 71974159; 72303219).

通讯作者: 姬强(1982—), 男, 山东威海人, 博士, 研究员, 博士生导师. Email: jqwxnj@163.com

没有对各项预测因子做分类和对比分析. 因此, 本文重点关注两个问题: 第一, 如何构造合理的预测因子; 第二, 对比分析多维信息对油价的预测效果.

本文的贡献主要包括四个方面. 首先, 是第一篇量化分析和预测中国原油期货价格波动的研究. 具体地, 从多因素预测的视角, 基于原油市场供需属性、商品属性、金融属性、汇率影响、全球宏观经济以及相关事件六个维度构造了相应的预测因子, 用于预测中国原油期货的价格波动. 其次, 重点识别和对比分析不同时间尺度下, 各类信息对 INE 原油价格波动的预测效果, 为中国原油期货市场规避市场短期、中期和长期的异常波动提供理论支持. 再次, 从信息视角, 基于广义动态因子模型提取并合成同类因素的信息, 避免因素之间的同质化影响和维数灾难, 为分析油价波动提供了新的分析工具. 最后, 本文采用互联网数据来量化突发事件. 与虚拟变量刻画事件相比, 本方法能更有效地反映事件发生的规模效应、响应模式以及持续强度^[13].

1 构建预测因子

首先, 利用广义动态因子模型构造原油市场供需预测因子、汇率市场信息预测因子、市场金融化预测因子、商品市场预测因子和全球宏观经济预测因子. 其次, 借鉴 Da 等^[14]方法, 利用互联网数据量化了原油市场相关事件, 构造了事件预测因子.

1.1 指标体系和数据

本部分重点介绍各个维度指标的选取和指标体系的构建依据, 详见表 1^②. 1) 供需基本面信息. 原油价格波动很大程度上服从供求规律^[8]. 选取 INE 原油库存体现市场短期供需; 选取波罗的海干货指数和中国沿海散货运价指数描述全球和中国原油需求情况; 选取中国原油进口金额体现中国原油市场需求侧信息. 2) 汇率市场信息. 中国原油期货以人民币计价和结算, 人民币货币信息变动会影响国际交易者的参与活跃度. 选取同业

拆放利率、美元兑人民币汇率以及人民币指数描述国内外汇率市场信息变动情况. 3) 商品市场信息. 原油作为大宗商品市场的重要组成部分, 其他大宗商品信息变动会直接传染到原油市场中^[11]. 选取 CRB 指数和南华商品指数分别用于描述国外和国内商品市场信息. 4) 市场金融化信息. 随着原油金融类资产的成熟, 原油市场金融化属性凸显, 其他金融市场信息会通过跨市场交易传递到原油市场当中, 驱动原油价格变动^[11]. 选取国内外股票市场信息(即, 上证指数和 S&P 500 指数)、中国原油期货市场场内交易信息和两市场隐含波动率描述市场金融化信息. 5) 宏观经济环境信息. 全球宏观经济不确定性增加会通过影响全球原油供需诱导原油价格剧烈波动^[15]. 选取经济政策不确定性指数(EPU) 量化宏观经济的不确定性^[16]. 根据全球主要原油需求和供给国家, 分别选取美国、中国、日本、俄罗斯、加拿大以及欧洲六个国家和地区的指数. 6) 事件信息. 与其他金融资产不同, 原油市场具有物理属性和地缘政治属性. 地缘政治事件、极端天气等突发事件容易引起原油价格波动^[13]. 本文以 OPEC 事件、地缘政治事件、自然灾害等 65 个词条为关键词, 提取互联网数据, 并以此量化事件信息冲击的规模效应和动态变化.

1.2 构建预测因子

1) 广义动态因子模型

与传统的因子模型不同, 广义动态因子模型^[17]能够描述系统统计结构的动态属性. 同时, 广义动态因子可以将各个变量之间的动态联系纳入到各个变量共同成分的计算中, 可反映原油市场系统的复杂动态. 具体模型如下

对于 $\forall n \in N, t \in T$, 各类信息 Y_{it} 可以分解为“共同信息成分” X_{it} 和“随机信息成分” Z_{it} 如下

$$Y_{it} = X_{it} + Z_{it} =: \sum_{k=1}^q b_{ik}(L) u_{kt} + Z_{it} \quad (1)$$

$i \in N, t \in T$

② 各指标的描述性统计结果如需可向作者索要.

表 1 预测因子指标体系说明
Table 1 Description of the predictive factor indicator system

类别	因素指标	频度	数据源
供需基本面信息	INE 原油库存期货	日度	万得数据库
	波罗的海干货指数	日度	
	中国原油进口金额	月度	
	中国沿海散货运价指数: 原油	周度	
汇率市场信息	上海银行间同业拆放利率	日度	万得数据库
	伦敦同业拆借利率	日度	
	离岸美元对人民币	日度	
	人民币指数	日度	
商品市场信息	CRB 指数	周度	万得数据库
	南华贵金属商品指数	日度	
	南华金属商品指数	日度	
	南华农产品商品指数	日度	
金融化信息	上证指数	日度	万得数据库
	INE 原油期货持仓量	日度	
	INE 原油期货交易量	日度	
	标准普尔 500 指数	日度	
	VIX 指数	日度	
	OVX 指数	日度	
宏观经济环境信息	美国 EPU	日度	http://www.policyuncertainty.com/index.html
	中国 EPU	月度	
	欧洲 EPU	月度	
	加拿大 EPU	月度	
	俄罗斯 EPU	月度	
	日本 EPU	月度	
事件信息	OPEC 类词条	日度	百度搜索指数
	地缘政治事件词条	日度	
	自然灾害类词条	日度	

注：事件类词条是首先通过分析泰晤士报、华尔街日报、纽约时报、经济日报等报刊关于原油市场相关报道的文本，总结出能够反映 OPEC 会议、地缘政治事件以及自然灾害事件信息的 17 个核心关键词条，如“OPEC 会议”、“飓风”、“利比亚战争”、“金融危机”等。其次，基于百度搜索相关词汇，最终将 17 个核心词汇扩展到 65 个相关词汇，这些词汇是事件预测因子的主要数据源。

其中 X_{it} 和 Z_{it} 分别表示各类信息的共同成分和随即成分； u_{it} 是共同信息 X_{it} 变动的共同因子； Q 是共同因子数目。模型假设请参考 Forni 等^[17]。

2) 量化突发事件

参考 Da 等^[14]方法，基于 65 个关键词条的互联网数据（百度搜索指数，记为 SVI ）进行分析。首先，以公式 $\frac{SVI \times 100}{SVI_{MAX}}$ 对各个词条搜索量数据做了预处理（ SVI_{MAX} 表示某词条百度搜索指数序列中最大值）。其次，针对数据中可能存在的季节性

和异方差性，本文又对数据进行对数差分预处理，即 $\Delta SVI_{j,t} = \ln(SVI_{j,t}) - \ln(SVI_{j,t-1})$ 。再次，为选出与 INE 油价波动存在有效关联的词条，分别将 65 个词条序列与 INE 油价波动进行格兰杰因果检验，删除不能显著引起油价波动的词条，筛选出 21 个有效词条（见表 2）作为构建事件预测因子的主要成分。最后，通过将这 21 个有效词条的搜索量变化值进行加和得到最终的原油市场事件预测因子 $(\sum_{j=1}^{21} \Delta SVI_{j,t})$ 。

表2 INE 油价波动与原油市场相关事件的有效关键词之间的格兰杰因果检验

Table 2 Granger causality test between INE oil price volatility and the effective keywords of oil-related events

序列	词条	滞后阶数	F 统计量
1	OPEC→INE 原油价格波动	1	45.006***
2	WTI 原油→INE 原油价格波动	1	33.900***
3	中东局势→INE 原油价格波动	1	58.013***
4	中东战争→INE 原油价格波动	1	13.650***
5	中东问题→INE 原油价格波动	1	67.178***
6	伊朗问题→INE 原油价格波动	1	8.740***
7	利比亚撤侨→INE 原油价格波动	1	3.077*
8	国际原油价格→INE 原油价格波动	1	36.435***
9	国际原油暴跌→INE 原油价格波动	1	5.843**
10	原油期货→INE 原油价格波动	1	3.197*
11	叙利亚内战→INE 原油价格波动	1	4.080**
12	次贷危机→INE 原油价格波动	1	13.487***
13	欧佩克→INE 原油价格波动	1	56.806***
14	海啸→INE 原油价格波动	1	5.178**
15	海湾战争→INE 原油价格波动	1	79.812***
16	炼油厂→INE 原油价格波动	1	71.955***
17	疫情→INE 原油价格波动	1	48.489***
18	美国制裁俄罗斯→INE 原油价格波动	1	4.608**
19	货币危机→INE 原油价格波动	1	3.563*
20	阿富汗战争→INE 原油价格波动	1	3.505*
21	飓风→INE 原油价格波动	1	4.478**

注:表中只列出格兰杰因果检验显著的21个关键词。***表示1%的显著性水平,**表示5%的显著性水平,*表示10%的显著性水平。

2 波动预测模型及检验

2.1 波动预测模型

为探究六类信息对 INE 原油价格波动的影响机制和预测效果,根据数据源频度存在不一致现象,本文选取 GARCH-MIDAS 模型^[18]用于预测。模型具体表达式如下:

假设第 t 周第 i 天 INE 油价收益服从下述过程

$$r_{i,t} = \mu + \sqrt{\tau_t g_{i,t}} \varepsilon_{i,t} \quad (2)$$

$$\varepsilon_{i,t} | \phi_{i-1,t} \sim N(0, 1)$$

其中 $r_{i,t}$ 是第 t 周第 i 天的油价收益, μ 是油价收益在基于 $\phi_{i-1,t}$ 信息集的条件期望。而 $r_{i,t}$ 的条件方差被分解为短期成分 $g_{i,t}$ 和长期成分 τ_t 。模型假设短期成分 $g_{i,t}$ 服从 GARCH(1,1) 过程,表达式如下

$$g_{i,t} = (1 - \alpha - \beta) + \alpha \frac{(r_{i-1,t} - \mu)^2}{\tau_t} + \beta g_{i-1,t} \quad (3)$$

长期成分 τ_t 采用 MIDAS 模型刻画,通过已实现波动 $RV_t = \sum_{i=1}^{N_t} r_{i,t}^2$ 解释长期成分 τ_t 表达式如下

$$RV_t = \sum_{i=1}^{N_t} r_{i,t}^2$$

$$\tau_t = m + \theta \sum_k \varphi_k RV_{t-k} \quad (4)$$

其中 RV_{t-k} 表示已实现波动相对当期(t 期)滞后 k

期的水平值. K 表示该变量水平值的最大滞后阶数.

在式(4)的基础上引入预测因子,探究这些变量对长期成分 τ_t 的影响机制,以及对油价波动的预测效果. 具体表达式如下

$$\tau_t = m + \theta_1 \sum_k^K \varphi_k RV_{t-k} + \theta_2 \sum_k^K \varphi_k X_{t-k} \quad (5)$$

其中 X_{t-k} 表示预测因子相对当期(t 期)滞后 K 期的水平值. K 表示该变量水平值的最大滞后阶数.

原油收益条件方差定义为

$$\sigma_{i,t}^2 = \tau_t g_{i,t} \quad (6)$$

在式(4)和式(5)中的权重 φ_k 为 Beta 权重函数表达式如下

$$\varphi_k(w) = \frac{\left(\frac{k}{K}\right)^{w_1-1} \left(1 - \frac{k}{K}\right)^{w_2-1}}{\sum_{j=1}^K \left(\frac{j}{K}\right)^{w_1-1} \left(1 - \frac{j}{K}\right)^{w_2-1}} \quad (7)$$

本文构建两种不同类型的模型,主要区别于对波动长期成分 τ_t 的度量,其中短期成分 $g_{i,t}$ 保持不变.

1) RV 模型. 在度量长期波动成分 τ_t 时,该模型仅引入已实现波动 RV,并未考虑预测因子,详见式(4).

2) RV + X 模型. 该模型在度量长期波动成分 τ_t 时,在上一类模型基础上引入预测因子,详见式(5).

2.2 预测效果检验模型

采用 MCS 检验方法^[19]量化六类预测因子在多个损失函数下的预测损失,识别出不同时间尺度下的最优预测模型.

首先,选用如下 4 种常用的损失函数

$$L_1: MSE = M^{-1} \sum_{t=1}^M (y_t - \hat{y}_t)^2 \quad (8)$$

$$L_2: MAE = M^{-1} \sum_{t=1}^M |y_t - \hat{y}_t| \quad (9)$$

$$L_3: QLIKE = M^{-1} \sum_{t=H+1}^{M+H} \left(\ln \hat{y}_t^2 + \frac{y_t}{\hat{y}_t} \right) \quad (10)$$

$$L_4: R^2 LOG = M^{-1} \sum_{t=H+1}^{M+H} \left(\frac{\ln y_t}{\hat{y}_t} \right)^2 \quad (11)$$

其中 y_t 表示原油价格波动的真实值,而 \hat{y}_t 表示基于不同模型得到的原油价格波动的预测值, M 表示预测值数目.

其次,将 2.1 节中的 7 个预测模型记为候选模型,即为 $M_0 = \{1, 2, \dots, 7\}$. 模型样本内估计和样本外预测长度分别为 h 和 q ,则根据式(8)~式(11)损失函数可以计算第 t 日损失值 $L_{i,j,t}$ (其中 i 表示损失函数类型 $i=1, 2, 3, 4$; j 表示相应的预测模型, t 表示预测时间点 $t=h+1, h+2, \dots, h+q$). 在 M_0 中任意两个波动率预测模型 u 和 v ,可计算出损失函数差即相对预测表现值, $d_{u,v,t} = L_{u,j,t} - L_{v,j,t}$.

根据 $d_{u,v,t}$ 可定义“高级对象的集合” M^* :

$$M^* = \{u \in M_0: E(d_{i,\mu,t}) \leq 0, \text{for } \forall v \in M_0\}$$

最后, MCS 检验本质是对集合 M_0 中的预测模型 M 进行显著性检验. 具体过程如下: 1) 根据零假设 $H_{0,M}$ 对集合 $M \subset M_0$ 进行相等性检验 δ_M ; 2) 根据剔除准则 e_M 逐步剔除集合 M_0 中聚集原假设 $H_{0,M}$ 的模型,即剔除预测能力较差的模型. 3) 通过重复如上两个检验最终得到一个“最优”模型. 其中,零假设具体表达式如下

$$H_{0,M}: E(d_{i,\mu,t}) = 0 \text{ 对于所有 } u, v \in M \subset M_0$$

为实现上述两个检验本文参考 Hansen 等^[19]选取常用的三个统计量进行统计检验.

$$T_R = \max_{u, v \in M} \frac{|\bar{d}_{i,\mu,v}|}{\sqrt{\text{var}(d_{i,\mu,v})}} \quad (12)$$

$$T_{SQ} = \max_{u, v \in M} \frac{(\bar{d}_{i,\mu,v})^2}{\text{var}(d_{i,\mu,v})} \quad (13)$$

$$T_{MAX} = \max_{i \in M} \frac{(\bar{d}_{i,\cdot})^2}{\sqrt{\text{var}(\bar{d}_{i,\cdot})}} \quad (14)$$

其中 $\bar{d}_{i,\mu,v} = \frac{1}{n} \sum_{t=1}^n d_{i,\mu,v,t}$ 表示模型 u 和 v 波动率预测值的相对损失函数值的平均值. $\bar{d}_{i,\cdot} = \frac{1}{m} \sum_{j \in M} \bar{d}_{ij}$ 表示第 i 个模型相对集合 M 中模型平均水平的损失函数值的平均值. 如果统计量 T_R, T_{SQ}, T_{MAX} 大于给定的临界值,表明拒绝零假设.

3 实证结果

在此分析样本内各类预测因子在 GARCH-

MIDAS 模型中的表现,识别出最有助于预测 INE 原油价格波动的预测因子。

3.1 数据描述

中国原油期货于 2018 年 3 月 26 日正式挂牌交易,本文数据样本覆盖 2018 年 4 月 2 日至 2020 年 5 月 31 日,并采用一阶对数差分计算得到日收益率,即 $y_t = \ln(P_t/P_{t-1})$ 。

图 1 展示了 2018 年至 2020 年各类预测因子的走势情况。由图可见,2020 年前各类预测因子走势相对比较平稳,而在疫情驱动下,2020 年度各类信息变动幅度明显增加。表 3 提供了收益和各类预测因子时间序列基本统计量。首先,从偏度

和峰度来看,大部分序列均呈现左偏及尖峰厚尾的特征。其次,汇率预测因子的标准差明显比其他几个指标大,表明汇率市场信息波动明显大于其他。最后,各变量的单位根检验均拒绝原假设,说明样本中所有时间序列均是平稳的。表 4 提供了各类预测因子与 INE 原油价格波动之间的格兰杰因果检验结果,用于检验六类预测因子能否引起 INE 原油价格波动。结果表明,在 1% 的水平下,本文构造的六类信息预测因子对 INE 原油价格波动均拒绝原假设,即各类信息预测因子均是导致 INE 原油价格波动的格兰杰原因,同时反映本文所构建六类信息预测因子的有效性。

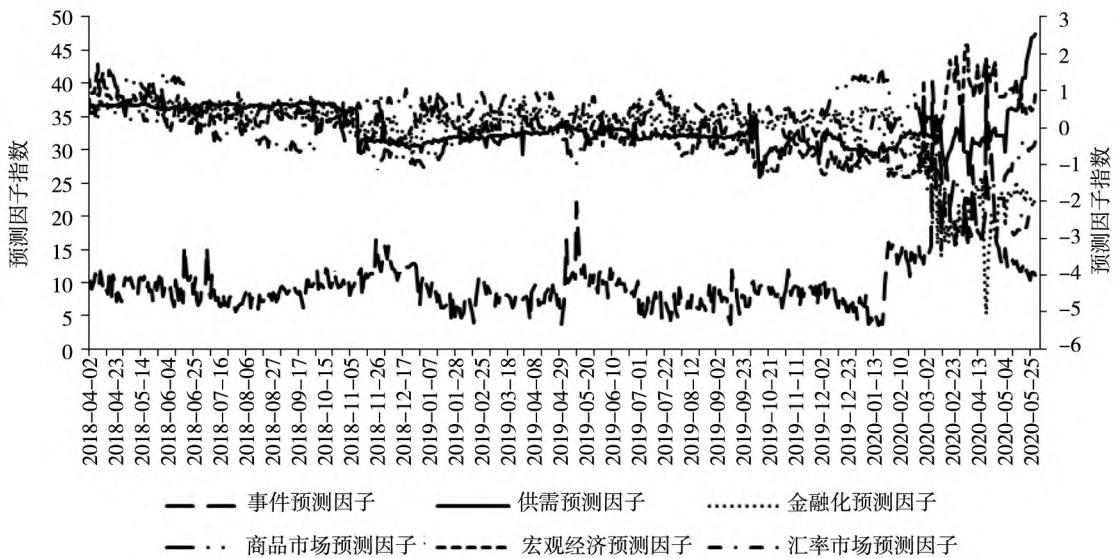


图 1 2018 年至 2020 年六类信息预测因子走势图

Fig. 1 The evolution of the six predictors, 2018 ~ 2020

表 3 INE 原油收益与六类信息预测因子的描述性统计量

Table 3 Descriptive statistics of INE crude oil returns and the six predictors

变量	INE 原油收益	供需信息	汇率信息	金融化信息	商品市场信息	宏观经济	原油相关事件
均值	-0.001	-0.001	-0.006	-0.005	-0.004	0.013	10.016
中位数	0.000	-0.185	0.307	0.272	-0.028	-0.009	9.165
最大值	0.097	2.542	1.053	0.936	1.536	2.229	43.635
最小值	-0.115	-1.344	-3.301	-5.040	-2.240	-1.384	3.519
标准差	0.027	0.522	0.971	0.888	0.774	0.678	4.479
偏度	-0.553	1.039	-1.926	-2.438	-0.191	0.285	2.833
峰度	6.024	5.532	5.802	8.999	3.012	3.022	16.241
ADF	0.001 ***	0.008 ***	0.022 ***	0.001 ***	0.008 ***	0.001 ***	0.005 ***
	-20.514	-2.647	-2.273	-4.785	-2.654	-4.747	-2.781

注: *** 表示 1% 的显著性水平, ** 表示 5% 的显著性水平, * 表示 10% 的显著性水平。

表 4 INE 原油价格波动与六类信息预测因子之间的格兰杰因果检验
Table 4 Granger causality test between oil price volatility and the six predictors

原假设	滞后阶数	F 统计量
供需预测因子 \rightarrow INE 原油价格波动	1	5.073 ***
汇率市场预测因子 \rightarrow INE 原油价格波动	1	9.884 ***
金融化预测因子 \rightarrow INE 原油价格波动	1	31.468 ***
商品市场预测因子 \rightarrow INE 原油价格波动	1	19.152 ***
宏观经济预测因子 \rightarrow INE 原油价格波动	1	6.099 ***
事件预测因子 \rightarrow INE 原油价格波动	1	66.527 ***

注: *** 表示 1% 的显著性水平, ** 表示 5% 的显著性水平, * 表示 10% 的显著性水平。

3.2 样本内估计结果

选择 2018 年 4 月 2 日到 2019 年 11 月 20 日进行样本内参数估计。表 5 展现了两类 GARCH-MIDAS 模型(共 7 个模型)样本内估计结果。表中所有模型的 β 均显著,表明 INE 原油价格波动具有记忆性。同时,表中 $\alpha + \beta$ 均超过 0.5 说明原油价格波动短期成分 $g_{i,t}$ 和所受的冲击力是持久的。

在油价波动长期成分 τ_t 方程中,本文重点关注已实现波动 $RV(\theta_{RV})$ 和各类信息预测因子 ($\theta_{predictor}$) 对 INE 原油价格波动的影响。表 5 结果显示大部分模型中已实现波动参数 θ_{RV} 和各类信息预测因子 $\theta_{predictor}$ 均显著为正,表明该因素是驱动 INE 油价波动长期成分 τ_t 的关键成分,且这些因素会增加 INE 原油价格波动。1) 供需预测因子 $\theta_{predictor}$ 显著为正,表明国内外原油市场供需基本面信息发生变动会显著增加 INE 原油价格波动^[8]。2) 汇率市场预测因子 $\theta_{predictor}$ 显著为负,表明人民币无风险利率和外汇信息变动是显著缓解 INE 原油价格波动的关键因素。说明除对标的资产价值变动会产生影响外,汇率变化也是关键驱动因素。3) 金融化预测因子 $\theta_{predictor}$ 显著为正,表明其他金融资产与中国原油期货市场之间的跨市场交易行为和跨市场风险传染会引起 INE 原油价格剧烈波动。这说明中国原油期货市场与国际原油市场一样存在金融化特征^[20],与 Ma 等研究结论相一致^[11]。Ma 等研究强调市场金融化不仅增加市场投机,同时丰富了原油类金融资产,使得投

机资本在全球范围内流动,为市场外部信息冲击原油价格变动和风险传染到原油市场中提供新渠道。4) 商品市场预测因子 $\theta_{predictor}$ 显著为负,表明大宗商品市场信息会显著影响 INE 原油价格波动。5) 宏观经济预测因子 $\theta_{predictor}$ 显著为正,表明全球宏观经济政策不确定性增加原油价格波动。Wei 等^[12]强调宏观经济不确定性会通过改变市场供应商、消费者以及投机者的预期,影响 INE 油价波动。6) 事件预测因子 $\theta_{predictor}$ 显著为正,表明突发性事件和地缘政治事件所包含的信息会增加 INE 原油价格波动。Ji 和 Guo^[13]强调突发性事件以及资源国地缘政治事件的信息释放到原油市场会扰乱投资者的市场预期,从而增加市场的波动。

3.3 样本外预测

为识别各类信息的预测能力、确保模型结果的稳健性,本文对比分析已实现波动和六类预测因子在样本外的预测能力。对 2018 年 4 月 2 日到 2019 年 11 月 20 日样本外数据进行 2 d、5 d、22 d 和 125 d 的预测,并通过 MCS 方法检验模型预测效果。参照魏宇等^[21]和张一锋等^[22],选取 $d = 2$ 和模拟次数 $B = 10\ 000$ 次作为 MCS 检验中 Bootstrap 过程的控制参数,并将 MCS 检验的显著性水平 α 取值为 0.1。若 p 值小于 0.1,则认为波动率预测模型样本外预测能力较差,将在 MCS 检验过程中被剔除,反之,若波动率预测模型 p 值大于 0.1,则在 MCS 检验中保留。

预测样本外 2 d、5 d、22 d 和 125 d INE 原油价格波动情况,并通过预测效果识别出最有助于

预测 INE 原油价格波动的预测因子,详见表 6 ~ 表 9. 表 6 ~ 表 9 是检验已实现波动和六类预测因子对 INE 原油价格波动的预测效果,表中的第 1 列是本文所涉及的全部波动率预测模型(共 7 个),第 1 行是各损失函数 L_i ,表中数字是在各损失函数下统计量 T_R 、 T_{SQ} 、 T_{MAX} 对应的 p 值.

由表 6 ~ 表 8 得到如下结论. 第一,中短期,与其他类信息预测因子相比,事件预测因子的预测效果最佳. 表 6 ~ 表 8 结果表示基于已实现波动和事件预测因子的模型(模型 7)的预测损失最低,说明该模型具有最高的预测精度. 并且在所有的损失函数 MSE 、 MAE 、 $QLIKE$ 、 $R^2 \ln$ 标准下,只有模型 7 被保留(p 值大于 0.1),其他模型均被剔除(p 值小于 0.1). 第二,随着预测时间尺度的增长,已实现波动信息、市场供需预测因子以及商品市场信息预测因子对 INE 原油价格波动的预测效果逐渐提高. 由表 6 可知,在所有损失函数下,除模型 7 外,其他模型均被剔除. 表 7 至表 9 结果显示,在 $QLIKE$ 和 $R^2 \ln$ 两个标准的 MCS 检验中已实现波动信息(模型 1)、市场供需预测因子(模型 2)以及商品市场信息预测因子(模型 5)均通过检验成为幸存模型. 因此,与短期预测相比,在中长期预测中,这些预测因子的预测能力明显增强.

表 9 展示了预测样本外 125 d INE 原油价格波动的评估结果,可见市场供需预测因子(模型 2)对 INE 原油价格波动预测能力最强. 在四个损失函数下,仅有模型 2 通过检验,其他模型均被剔除. 反映在长期预测中,市场供需信息对 INE 原油价格波动的预测效果最好.

总体上,首先,在预测样本外 2 d、5 d 和 22 d 的短期和中期 INE 原油价格波动时,基于已实现波动和原油市场相关事件的预测模型(模型 7) (p 值大于 0.1 或 p 值等于 1)的预测效果最好. 短、中期内,国内外原油市场相关事件对 INE 原油价格波动起到主要驱动作用. 这符合当前中国原油期货市场特点,理由如下: 1) 与国际原油市场相比,中国原油期货市场刚处于初级发展阶段,市场成熟度和风险防控机制建设水平有待提高,市场

更易受到地缘政治事件和黑天鹅事件的影响,而造成市场价格短期波动. 2) 与其他国际市场不同,中国原油期货市场对胜利油和阿曼、上扎库姆、巴士拉轻油、迪拜、卡塔尔海洋、马西拉等 6 种中东油种作为交易标的. 一旦中东局势震荡,将直接影响我国原油期货市场交易. 与国际市场相比,中国原油期货价格对地缘政治等突发事件的反应更加敏感. 因此,面对原油市场特有“地缘政治属性”,市场参与者在短期内应更多地关注原油市场相关事件的影响,以提高 INE 原油价格波动的预测精度.

其次,随着预测时间尺度的延展,供需预测因子对 INE 原油价格波动的预测能力逐渐增强,并成为主导. 首先,这一结果反映不同影响因素对不同时间尺度 INE 原油价格波动的预测效果存在差异性. 为市场参与者在不同时间尺度下如何重点关注市场预测指标和防范市场异常波动提供理论依据. 其次,这一结论证实从长期预测来看,原油供需信息仍是决定 INE 原油价格波动的关键. Degiannakis 和 Filis^[23] 同样证实原油市场供需基本面信息对长期原油价格波动的预测效果较好. Degiannakis 和 Filis 认为这主要是由于市场供需基本面信息更全面,包含全球商业周期、全球石油产量和全球原油库存之间的长期均衡关系等.

因此,在六类信息预测因子中,事件预测因子和原油市场供需预测因子对 INE 原油价格波动的预测能力较强. 其中,前者在短期、中期内起主导作用,后者在长期波动预测中占主导. 本文认为,参考原油市场相关事件信息和市场供需信息,可有效提高原油价格波动的预测精度.

3.4 稳健性检验

本文通过改变样本内时间窗口长度检验了预测效果的稳健性. 表 10 描述了预测未来 2 d、5 d、22 d 和 125 d INE 原油价格波动的最优预测模型. 可见,对未来 2 d、5 d 和 22 d 的 INE 油价波动预测时,原油市场事件预测因子的预测效果最佳. 随着预测时间尺度的增加,市场供需基本面信息的预测能力逐渐增强. 这与文中表 6 ~ 表 9 的结果相一致,表明本文研究结果和结论是稳健的.

表 5 GARCH-MIDAS 模型样本内参数估计结果

Table 5 In-sample estimation results of the GARCH-MIDAS models

序号	模型	μ	α	β	θ_{RV}	$\theta_{predictor}$	w_{RV}	$w_{predictor}$	m	LLF/BIC
(I)	基于已实现波动模型	0.102 (1.127)	0.277*** (3.563)	0.274** (1.987)	0.312*** (3.518)		1.693*** (3.737)		1.715*** (4.984)	-837.394 4.501
(II)	基于已实现波动和供需预测因子模型	0.107 (1.094)	0.280*** (3.573)	0.268** (2.108)	-0.268*** (-2.950)	2.444*** (5.953)	2.235*** (3.414)	500.505*** (3.003)	0.007*** (3.924)	-836.054 4.525
(III)	基于已实现波动和汇率预测因子模型	0.102 (0.453)	0.277*** (3.320)	0.274* (1.830)	-0.312** (-2.042)	-3.476*** (-3.168)	1.693** (1.702)	499.979*** (4.418)	1.715** (2.982)	-837.394 4.532
(IV)	基于已实现波动和金融化预测因子模型	0.105 (1.039)	0.283*** (3.514)	0.278** (2.239)	0.307*** (3.729)	2.065* (1.454)	1.634** (2.104)	1.079** (2.774)	0.772*** (3.581)	-837.139 4.531
(V)	基于已实现波动和商品市场预测因子模型	0.195* (1.877)	0.900*** (16.666)	0.098* (1.829)	4.548** (2.683)	-2.257** (-1.970)	1.444* (1.876)	43.562*** (16.481)	26.903*** (3.642)	-852.389 4.611
(VI)	基于已实现波动和宏观经济预测因子模型	0.106 (1.096)	0.279*** (3.553)	0.273** (2.317)	-0.314*** (-4.003)	0.622** (2.035)	1.692** (1.691)	3.380** (1.968)	1.637*** (5.283)	-837.372 4.532
(VII)	基于已实现波动和事件预测因子模型	0.205** (1.975)	0.921*** (20.797)	0.078* (1.763)	6.578** (2.642)	5.950** (2.633)	1.511** (2.081)	134.473*** (5.406)	35.066*** (3.397)	-853.268 4.616

注：***表示 1% 的显著性水平，**表示 5% 的显著性水平，*表示 10% 的显著性水平；括号内为 t 统计量。

表 6 已实现波动和六类预测因子对未来 2 d 原油价格波动的预测效果

Table 6 Based on RV and the six predictors, MCS results for different volatility forecasting results at 2-day-ahead horizon

损失函数	MSE			MAE			QLIKE			R^2 ln	
	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}	T_{SQ}	T_{MAX}
模型 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
模型 2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
模型 3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.000	0.028
模型 4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.000	0.028
模型 5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.000	0.028
模型 6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.000	0.028
模型 7	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

注：模型 1 为基于已实现波动模型，模型 2 为基于已实现波动和供需预测因子模型，模型 3 为基于已实现波动和汇率预测因子模型，模型 4 为基于已实现波动和金融化预测因子模型，模型 5 为基于已实现波动和商品市场预测因子模型，模型 6 为基于已实现波动和宏观经济预测因子模型，模型 7 为基于已实现波动和事件预测因子模型。表中加粗数字对应的模型为在特定损失函数下通过 MCS 检验保留的预测模型。

表7 已实现波动和六类预测因子对未来5 d原油价格波动的预测效果

Table 7 Based on RV and the six predictors, MCS results for different volatility forecasting results at 5-day-ahead horizon

损失函数	MSE			MAE			QLIKE			R ² ln		
	T _R	T _{SQ}	T _{MAX}	T _R	T _{SQ}	T _{MAX}	T _R	T _{SQ}	T _{MAX}	T _R	T _{SQ}	T _{MAX}
模型1	0.000	0.000	0.000	0.008	0.002	0.000	0.185	0.196	0.145	0.172	0.172	0.142
模型2	0.000	0.000	0.000	0.098	0.084	1.000	0.203	0.219	1.000	0.191	0.188	1.000
模型3	0.000	0.000	0.000	0.039	0.029	0.122	0.000	0.001	0.111	0.000	0.002	0.097
模型4	0.000	0.000	1.000	0.098	0.084	1.000	0.204	0.220	1.000	0.192	0.191	1.000
模型5	0.000	0.000	1.000	0.098	0.084	1.000	1.000	1.000	1.000	1.000	1.000	1.000
模型6	0.000	0.000	0.000	0.020	0.009	0.105	0.011	0.014	0.113	0.008	0.013	0.101
模型7	1.000	1.000	1.000	1.000	1.000	1.000	0.195	0.209	0.157	0.182	0.183	0.164

注：同表6。

表8 已实现波动和五类预测因子对未来22 d原油价格波动的预测效果

Table 8 Based on RV and the six predictors, MCS results for different volatility forecasting results at 22-day-ahead horizon

损失函数	MSE			MAE			QLIKE			R ² ln		
	T _R	T _{SQ}	T _{MAX}	T _R	T _{SQ}	T _{MAX}	T _R	T _{SQ}	T _{MAX}	T _R	T _{SQ}	T _{MAX}
模型1	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000
模型2	0.000	0.000	0.993	0.000	0.000	0.441	0.202	0.257	0.207	0.243	0.230	0.207
模型3	0.000	0.000	0.000	0.000	0.000	0.011	0.208	0.257	0.883	0.243	0.230	0.898
模型4	0.000	0.000	0.000	0.000	0.000	0.068	0.255	0.269	0.902	0.270	0.278	0.919
模型5	0.000	0.000	1.000	0.000	0.000	0.999	0.208	0.257	0.229	0.243	0.230	0.229
模型6	0.000	0.000	0.993	0.000	0.000	0.999	0.255	0.269	0.892	0.270	0.278	0.909
模型7	1.000	1.000	1.000	1.000	1.000	1.000	0.208	0.257	0.226	0.243	0.230	0.225

注：同表6。

表9 已实现波动和五类预测因子对未来125 d原油价格波动的预测效果

Table 9 Based on RV and the six predictors, MCS results for different volatility forecasting results at 125-day-ahead horizon

损失函数	MSE			MAE			QLIKE			R ² ln		
	T _R	T _{SQ}	T _{MAX}	T _R	T _{SQ}	T _{MAX}	T _R	T _{SQ}	T _{MAX}	T _R	T _{SQ}	T _{MAX}
模型1	0.000	0.000	0.000	1.000	1.000	1.000	0.521	0.579	0.914	0.558	0.564	0.908
模型2	1.000	1.000	1.000	0.848	0.849	0.760	1.000	1.000	1.000	1.000	1.000	1.000
模型3	0.055	0.040	1.000	0.848	0.849	0.607	0.007	0.014	0.169	0.025	0.019	0.170
模型4	0.123	0.122	1.000	0.848	0.839	0.573	0.007	0.014	0.892	0.025	0.019	0.894
模型5	0.055	0.040	1.000	0.848	0.839	0.569	0.007	0.014	0.889	0.025	0.019	0.879
模型6	0.123	0.122	1.000	0.848	0.849	0.607	0.030	0.026	0.900	0.043	0.049	0.904
模型7	1.000	1.000	1.000	0.848	0.849	0.607	0.096	0.096	0.902	0.109	0.107	0.904

注：同表6。

表 10 预测未来 2 d、5 d、22 d 和 125 d INE 原油价格波动的最优预测模型^③
Table 10 The best forecasting model for oil price volatility at different day ahead horizon

未来 2 d INE 原油价格波动最优预测模型						
损失函数	MSE			MAE		
	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}
模型 7	1.000	1.000	1.000	1.000	1.000	1.000
损失函数	QLIKE			$R^2 \ln$		
	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}
模型 7	1.000	1.000	1.000	1.000	1.000	1.000
未来 5 d INE 原油价格波动最优预测模型						
损失函数	MSE			MAE		
	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}
模型 7	1.000	1.000	1.000	1.000	1.000	1.000
损失函数	QLIKE			$R^2 \ln$		
	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}
模型 7	1.000	1.000	1.000	1.000	1.000	1.000
未来 22 d INE 原油价格波动最优预测模型						
损失函数	MSE			MAE		
	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}
模型 7	1.000	1.000	1.000	1.000	1.000	1.000
损失函数	QLIKE			$R^2 \ln$		
	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}
模型 7	1.000	1.000	1.000	1.000	1.000	1.000
未来 125 d INE 原油价格波动最优预测模型						
损失函数	MSE			MAE		
	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}
模型 2	1.000	1.000	1.000	1.000	1.000	1.000
损失函数	QLIKE			$R^2 \ln$		
	T_R	T_{SQ}	T_{MAX}	T_R	T_{SQ}	T_{MAX}
模型 2	1.000	1.000	1.000	1.000	1.000	1.000

注: 同表 6.

4 结束语

本文从原油市场供需、外汇市场、商品市场、原油市场金融属性、全球宏观经济以及原油市场相关事件六个维度提取相关信息构造预测因子, 并对 INE 原油价格波动进行分析与预测. 具体结论如下.

首先, 与其他指标相比, 本文发现原油市场相关事件信息对 INE 油价波动的预测精度更高. 市场短期波动已经偏离供需框架的范畴, 因此, 面对

原油价格持续低迷的情况, 突发性事件和地缘政治事件等成为刺激原油价格波动的重要信息. 对 INE 原油市场投资者来说, 未来更加复杂的市场交易、地缘政治环境, 以及突发的自然灾害等黑天鹅事件, 将成为短期内需要应对的主要风险点. 其次, 随着预测尺度的增加, 原油市场供需因子的预测能力增强. 这一结果反映出 INE 原油价格波动在长期仍由供需决定, 同时本文也证实了, 在不同时间尺度下决定 INE 原油价格波动的信息是变化的.

本文的研究结论为原油市场参与者、政策制

^③ 详细稳健性检验结果可向作者索要.

定者、市场监管者提供了把握未来市场的分析工具和参考依据。本文强调了原油市场相关信息和市场供需信息在分析和预测 INE 原油价格波动中的重要性。并且强调不同影响因素在不同时间尺度下对 INE 原油价格波动的预测效果存在差异。这一发现提示 INE 原油市场参与者在把握未来 INE 油价波动时应更多地考虑原油市场

突发事件,同时对于价格监管者来说,应加强对国际资本投机的监管从而有效地抑制原油价格过度波动。同时,在预测 INE 原油价格波动时,市场参与者可根据预测时间尺度选择合适的预测因子。此外,本文构建的原油价格波动预测模型和研究框架具有一般性特点,可推广应用到其他原油市场的研究中。

参考文献:

- [1]Marchese M, Kyriakou I, Tamvakis M, et al. Forecasting crude oil and refined products volatilities and correlations: New evidence from fractionally integrated multivariate GARCH models[J]. *Energy Economics*, 2020, (88): 104757.
- [2]张跃军,张 晗,王金丽. 考虑结构变化和长记忆性的国际原油价格波动率预测研究[J]. *中国管理科学*, 2021, 29(9): 54-64.
Zhang Yuejun, Zhang Han, Wang Jinli. Volatility forecasting of crude oil market based on structural changes and long memory[J]. *Journal of Management Science in China*, 2021, 29(9): 54-64. (in Chinese)
- [3]Chen W, Ma F, Wei Y, et al. Forecasting oil price volatility using high-frequency data: New evidence[J]. *International Review of Economics and Finance*, 2020, (66): 1-12.
- [4]龚 旭,林伯强. 跳跃风险、结构突变与原油期货价格波动预测[J]. *中国管理科学*, 2018, 26(11): 11-21.
Gong Xu, Lin Boqiang. Jump risk, structural breaks and forecasting crude oil futures volatility[J]. *Chinese Journal of Management Science*, 2018, 26(11): 11-21. (in Chinese)
- [5]Pan Z, Wang Y, Wu C, et al. Oil price volatility and macroeconomic fundamentals: A regime switching GARCH-MIDAS model[J]. *Journal of Empirical Finance*, 2017, (43): 130-142.
- [6]Hamilton J D. Understanding crude oil prices[J]. *Energy Journal*, 2009, (30): 179-206.
- [7]王盼盼,夏 婷,石建勋,等. “石油-美元”动态关联的时变特征及影响因素研究[J]. *国际金融研究*, 2020, (11): 35-44.
Wang Panpan, Xia Ting, Shi Jianxun, et al. Time-varying characteristics and influencing factors of “oil-dollar” dynamic correlation[J]. *Studies of International Finance*, 2020, (11): 35-44. (in Chinese)
- [8]Nonejad N. Forecasting crude oil price volatility out-of-sample using news-based geopolitical risk index: What forms of non-linearity help improve forecast accuracy the most? [J]. *Finance Research Letters*, 2021, (46): 102310.
- [9]尹力博,杨清元,韩立岩. 技术指标能够预测商品期货价格吗? 来自中国的证据[J]. *管理科学学报*, 2018, 21(6): 99-109.
Yin Libo, Yang Qingyuan, Han Liyan. Can technical indicators forecast commodity prices? Evidence from China[J]. *Journal of Management Sciences in China*, 2018, 21(6): 99-109. (in Chinese)
- [10]Miao H, Ramchander S, Wang T, et al. Influential factors in crude oil price forecasting[J]. *Energy Economics*, 2017, (68): 77-88.
- [11]Ma Y R, Ji Q, Pan J F. Oil financialization and volatility forecast: Evidence from multidimensional predictors[J]. *Journal of Forecasting*, 2019, 38(6): 564-581.
- [12]Wei Y, Liu J, Lai X, et al. Which determinant is the most informative in forecasting crude oil market volatility: Fundamental speculation or uncertainty? [J]. *Energy Economics*, 2017, (68): 141-150.
- [13]Ji Q, Guo J F. Oil price volatility and oil-related events: An Internet concern study perspective[J]. *Applied Energy*, 2015, (137): 256-264.
- [14]Da Z, Engelberg J, Gao P. The sum of all FEARS investor sentiment and asset prices[J]. *The Review of Financial Studies*, 2014, 28(1): 1-32.
- [15]Aloui R, Gupta R, Miller S M. Uncertainty and crude oil returns[J]. *Energy Economics*, 2016, (55): 92-100.

- [16]彭涛,黄福广,孙凌霞. 经济政策不确定性与风险承担: 基于风险投资的证据[J]. 管理科学学报, 2021, 24(3): 98–114.
Peng Tao, Huang Fuguang, Sun Lingxia. Economic policy uncertainty and risk-taking: Evidence from venture capital[J]. Journal of Management Sciences in China, 2021, 24(3): 98–114. (in Chinese)
- [17]Forni M, Hallin M, Lippi M, et al. The generalized dynamic-factor model: Identification and estimation[J]. Review of Economics and Statistics, 2000, 82(4): 540–554.
- [18]Engle R F, Ghysels E, Sohn B. Stock market volatility and macroeconomic fundamentals[J]. Review of Economics and Statistics, 2013, 95(3): 776–797.
- [19]Hansen P R, Lunde A, Nason J M. The model confidence set[J]. Econometrica, 2011, 79(2): 453–497.
- [20]张大永,姬强. 中国原油期货动态风险溢出研究[J]. 中国管理科学, 2018, 26(11): 42–49.
Zhang Dayong, Ji Qiang. Studies on the dynamic risk spillovers for China's crude oil futures[J]. Journal of Management Science in China, 2018, 26(11): 42–49. (in Chinese)
- [21]魏宇,马锋,黄登仕. 多分形波动率预测模型及其 MCS 检验[J]. 管理科学学报, 2015, 18(8): 61–72.
Wei Yu, Ma Feng, Huang Dengshi. Multi-fractal volatility forecasting model and its MCS test[J]. Journal of Management Sciences in China, 2015, 18(8): 61–72. (in Chinese)
- [22]张一锋,雷立坤,魏宇. 羊群效应的新测度指数及其对我国股市波动的预测作用研究[J]. 系统工程理论与实践, 2020, 40(11): 2810–2824.
Zhang Yifeng, Lei Likun, Wei Yu. A new herd index and volatility forecasting of China's stock market[J]. System Engineering: Theory and Practice, 2020, 40(11): 2810–2824. (in Chinese)
- [23]Degiannakis S, Filis G. Forecasting oil prices: High-frequency financial data are indeed useful[J]. Energy Economics, 2018, (76): 388–402.

What is the key factor driving price volatility of China's crude oil futures?

MA Yan-ran¹, WU Fei², ZHANG Da-yong², JI Qiang^{1*}

1. Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China;
2. Research Institute of Economics and Management, Southwestern University of Finance and Economics, Chengdu 610074, China

Abstract: This paper quantifies for the first time the driving factors predicting price volatility of China's crude oil futures. Using the generalized dynamic factor model, and the internet data, this paper constructs six factors to predict the price movements of China's crude oil futures: the demand-supply factor, market financialization factor, exchange rate factor, commodity market factor, global macroeconomic factor, and event-driven factor. Results from the GARCH-MIDAS model show that these factors can significantly improve the forecasting accuracy of crude oil futures price volatility. Further, results based on the MCS tests demonstrate that the underlying information driving price volatility varies across different time horizons. Specifically, the event-driven factor plays the leading role in the short-term and medium-term forecasting, whereas the demand-supply factor is the key for explaining long-term price volatility. Overall, this study provides a useful framework and key reference for market participants, policymakers and regulators to utilize market information in China's crude oil market.

Key words: China's crude oil futures; volatility forecasting; the predictors; GARCH-MIDAS