

参赛队员姓名： 陈含章

中 学： 深圳高级中学

省 份： 广东省

国家 / 地 区： 中 国

指导教师姓名： 韩 乾

指导教师单位： 中山大学岭南学院

论 文 题 目： 期权交易者聪明吗？来自中国
资本市场的证据

期权交易者聪明吗？

——来自中国资本市场的证据

陈含章

深圳高级中学

摘要：本文实证了期权隐含信息与未来价格和风险的关系，证实期权交易者是“聪明”的，对市场价格变化和风险更加敏锐，具有一定预测感知能力。一是支持价格发现说，偏度指数蕴含了未来价格变化的信息，对收益率有一定预测能力。二是支持信息发现说，期权交易者对宏观经济信息有显著预测感知能力，且对“积极”消息预测感知能力强于“负面”消息。三是支持风险感知说，期权隐含信息对已实现波动和尾部风险具有显著预测能力，且对大涨和大跌两类尾部风险的预测能力存在差异。在实证发现基础上，本文建立了基于信息的两类交易者理性预期模型，解释了期权隐含信息对股票市场有预测感知能力的机制，并用计算实验仿真证明引入“聪明”期权交易者可以提高股票市场价格发现和风险定价能力。本文对于金融市场建设、金融定价理论、尾部风险管理具有重要价值。

关键词：波动率指数；偏度指数；聪明交易者；非对称信息；机器学习；尾部风险；ETF 期权

Are Option Traders Smart?

——The Evidence from China's Capital Market

Abstract: This paper empirically examines the relationship between option-implied information and future prices and risks, providing evidence that option traders are "smart" and more sensitive to market price changes and risks, and have predictive capabilities to some extent. The empirical findings are as follows: Firstly, it supports the price discovery theory, indicating that the skewness index has a certain level of predictability for returns. Secondly, it supports the information discovery theory, indicating that option traders have significant predictive perceptions of macroeconomic information, and they are better at predicting "positive" news than "negative" news. Thirdly, it supports the discovery and pricing theory, indicating that option-implied information has significant predictive power for realized volatility and tail risks, and the power differs in predicting two types of tail risks. Further, the paper constructs asymmetric information pricing model to explain empirical findings, and designs agent-based model to prove the introduction of options and spot cross-market investors can improve the efficiency of price discovery. The evidence presented in this paper suggests that option traders are "smart", having a certain level of predictive perception of market price changes and risks, especially in predicting risk perception. This paper holds important value for financial market construction, financial pricing theory, and tail risk management.

Keywords: Volatility Index; Skewness index; Smart traders; Asymmetric Information; Machine Learning; Tail risk; ETF options

目 录

一、 引言.....	4
二、 文献回顾.....	5
(一) 价格发现说.....	5
(二) 风险感知说.....	6
三、 研究设计.....	7
(一) 指标构建.....	7
(二) 实证设计.....	8
四、 实证检验.....	8
(一) 价格发现说.....	9
(二) 风险感知说.....	13
(三) 机器学习的进一步检验.....	16
五、 稳健性检验.....	18
(一) 分时段的稳健性检验.....	18
(二) 对照样本的稳健性检验.....	20
六、 理性预期均衡模型与计算实验仿真.....	20
(一) 理论模型.....	21
(二) 计算实验仿真.....	23
七、 结论与启示.....	26
参考文献.....	28
致 谢.....	38

一、引言

2022年，股票类期权成交513.3亿张，占全球衍生品成交量的6成，成交面值相当于全球股票市场成交金额的2.9倍，已成为全球最重要的风险管理工具。中国境内2015年开始试点ETF期权，发展迅速，8年多来累计上市品种9只，成交量年均增长率达89%，已成为专业机构开展资产管理、风险对冲的主流工具。2022年，境内ETF期权成交12.1亿张，跃居全球第2位，成交面值60.6万亿元，达到境内股票成交金额的3成，有4只ETF期权成交量居全球同类产品前10位。以机构为代表的交易者是期权市场主要参与者，拥有信息和专业优势，能够更敏锐地感知市场变化，其行为会导致期权市场蕴含着股市未来变化信息。在境内期权市场快速发展背景之下，自然而然的问题是：期权交易者“聪明”吗？他们对市场价格变化和风险更加敏锐吗？这些既有理论意义又具备实务价值的关键问题，还需要我们全面严谨的研究论证。

准确预测未来的市场价格变化和风险（波动），是交易者追求的最终目标，也是金融研究者孜孜不倦的努力方向。如果期权市场领先于股票市场发现价格，那么通过分析期权隐含信息，就可以提前预测感知股票市场走势获取良好收益。同理，如果期权市场对股票市场风险有领先指示作用，并且能够精确地对风险进行定价，就可以提升投资组合管理绩效。进一步地，如果能运用期权隐含信息预期到未来的暴跌风险，就可以在“黑天鹅事件”的极端风险管理方向上迈进一步。

回顾文献发现，现有研究往往通过检验期权隐含信息的预测能力，作为判断期权交易者是否能够预测感知市场价格变化和风险的方法论支撑，方法论上已较为成熟，但研究问题的广度深度上还存在三方面不足：第一，已有文献陆续证实了一些期权隐含信息与现货价格未来收益率和波动显著相关，但分析维度较为单一，系统、全面研究期权交易者对价格和风险的预测感知和定价能力问题的文献仍较少，也缺乏理论模型和统一定论。第二，很少有文献直接针对性回答期权市场是否“聪明”，并进一步揭示期权交易者“聪明”的具体表现和内在机理。第三，现有的大部分研究是针对美国等成熟市场进行的，针对境内ETF期权这一新兴市场的专门研究较为欠缺。

本文研究结论支持信息发现说和风险感知说，期权交易者对股票市场收益率和重要宏观信息有一定感知能力，对未来已实现波动（传统意义上的投资风险）和尾部风险（“黑天鹅”风险）均有较好的预测感知能力。本文的主要贡献在于：第一，首篇系统分析期权市场对现货（股票市场）的预测感知能力的文章，实证证实了期权交易者是一群“聪明”（Smart）的交易者，对重要宏观信息有显著预测感知能力，该结论对于金融理论和实践有重要意义。第二，建立了基于信息感知的两类交易者理性预期模型，分析了期权隐含信息对股票市场的预测感知的

机制，并用计算实验仿真证明了期现交易者的存在，可以改进股票市场的价格效率和波动性。第三，证明了与美国等成熟市场不同，我国期权交易者的“聪明”突出体现在风险（包括尾部风险）的感知能力上，对于投资管理尤其是极端风险管理具有重要实践价值。第四，应用计算仿真实验和机器学习方法提升结论可靠性。

基于上述目标，本文的基本研究架构如下：第一和第二部分为引言和文献综述，第三部分为研究设计，第四部分为实证检验部分，主要检验中国资本市场期权隐含信息是否对未来股票价格有预测能力、是否对股票市场重要信息有感知能力，以及是否对股票市场波动率和尾部风险有预测能力，第五部分为稳健性检验，第六部分是理论建模和计算实验仿真，最后一部分是结论与启示。

二、文献回顾

综合国内外文献来看，围绕期权与现货市场关系和运行机理，主要有两大流派理论，一是**价格发现说**，主要从价格发现角度论述，主要理论包括信息发现论和市场效率论，二是**风险感知说**，主要从风险发现角度论述。

（一）价格发现说

许多研究证实了期权市场领先于股票市场发现价格（Giot, 2005; Banerjee, 2007）。信息发现论和市场效率论从各自角度解释了这种现象，前者从信息角度进行解释，后者则认为是市场之间效率差异导致两个市场之间的领先滞后关系，但是目前尚无定论。国外已有文献中，个股期权的研究普遍证实期权发现价格早于现货，信息发现论可能是更好的解释，因为期权交易者获取私有信息能力和分析公开信息能力高于其它交易者。国内文献方面，王琳玉等（2020）、朱超（2021）均证实上证 50ETF 期权隐含波动率、隐含偏度与指数收益率负相关。国内已有文献仅检验了隐含波动率、隐含偏度，其它期权隐含信息对股票市场收益率的研究较少，且对原因缺少进一步分析，笼统解释为行为金融或交易者情绪因素。

信息发现论认为期权交易者私有信息获得能力强于股票市场，且其分析解释公开信息的能力更强，其将信息先融入期权交易，表现为期权市场领先于股票市场感知价格走向。**隐含波动率方面**，An et al.（2014）发现个股期权隐含波动率与股票未来收益率显著相关，该相关性主要源于期权交易者的私有信息。**隐含偏度方面**，Xing et al.（2010）发现个股期权波动率曲线偏斜度与股票未来收益率成反比，原因在于有负面内部信息的交易者倾向于交易虚值认沽期权。**波动率价差方面**，Atilgan（2014）发现公司盈利信息公布前隐含价差显著扩大，波动率价差的变化方向可以作为股市收益率的预测因子。**认沽认购比方面**，Easley（1998）、

Pan (2006) 发现知情交易者参与个股期权市场，使得期权认沽认购比包含了未来股价信息。

市场效率论认为交易高效率是期权市场领先股票市场发现价格的本质原因。期权市场交易成本、杠杆程度和策略丰富程度均优于股票市场，新信息出现后，聪明交易者会选择流动性更好、效率更高的交易工具，因此期权市场会领先股票市场发现价格。期权交易成本更低 (Skinner, 1990; Fleming et al., 1996)、杠杆程度更高 (Chakravarty et al., 2004; Lee 和 Yi, 2001)、期权交易策略更加丰富 (Biais 和 Hillion, 1994)，都会导致交易者优先选择期权市场交易，使得期权市场领先股票市场发现价格。

(二) 风险感知说

风险感知说认为期权交易者具有风险的感知能力和定价能力，具体体现在波动率预测和尾部风险预测两个方面。

波动率预测方面，Jiang 和 Tian (2005) 发现基于标普 500 期权计算的 VIX 指数可以有效预测未来已实现波动率。Wang (2019) 发现 13 个国家股票市场的 VIX 显著影响未来已实现波动率。Pan et al. (2019) 发现 VIX 能显著提升波动率预测的效果。Kambouroudis et al. (2021) 在 HAR 模型中引入 VIX 指标后，显著改进了已实现波动的预测能力。但上述研究均专注于研究 VIX，未考虑隐含偏度等其他期权隐含信息对实际波动率的预测作用。

尾部风险预测方面，期权隐含信息可以预测未来的暴跌（左尾）风险。**隐含波动率方面**，美国 CBOE（芝加哥期权交易所）推出的 VIX 指数（波动率指数）被称为“恐慌指数”，许多成熟市场将其作为风险预警指标，Coudert 和 Gex (2008) 基于 27 个国家数据研究发现 VIX 指标能够有效预警股市危机。Li et al. (2015) 使用 VIX 等指标建立了一套预警系统来预测金融危机。**隐含偏度方面**，美国 CBOE 推出的 SKEW 指数（偏度指数），被称为“黑天鹅指数”，主要反映市场收益分布的尾部风险，Bates (1991) 发现美国 1987 年股灾前，标普 500 指数期权出现显著负偏。**国内方面**，陈蓉和林秀雀 (2016) 发现标普 500 期权市场的隐含偏度含有未来市场尾部风险的信息，但不能准确预测尾部风险的状态。

综上所述，已有研究主要存在以下几方面不足：一是已有文献陆续证实了一些期权隐含信息与现货价格未来收益率和波动显著相关，但研究采用的隐含信息指标也较为零散，综合检验各类指标、系统研究该问题的文献仍没有见到，现有文献没有直接回答期权交易者是否“聪明”，也没有进一步检验信息发现论、市场效率论的文献。二是现有研究主要针对美国等成熟市场进行，尤其是有关中国市场期权交易者对尾部风险的预测感知研究较少，也很少有文献建立理论模型分析期现市场对信息的反应机制和价格形成机理。当前中国股票类期权已经成为全

球领先的市场，理论和实证角度专门、系统研究期权交易者是否具有价格和风险的感知能力、尾部风险预测能力，及信息反应机制等问题，已经非常必要。

三、研究设计

综合前文文献回顾，如果期权交易者是“聪明”交易者，其价格发现和感知风险的能力更强，那么期权市场会领先股票市场预测感知未来价格变化和市场波动。正如信息发现说所描述，期权交易者拥有私有信息或分析解读公开信息能力更强（Jin et al., 2012），相应的购买（出售）期权会早于股票市场交易，其交易行为体现在期权隐含信息的多个指标中，导致这些指标能够预测未来指数的收益率和波动率，因此可以通过实证上述关系来分析期权交易者是否聪明的问题。

（一）指标构建

本文实证主要检验期权市场隐含信息是否对市场价格变化、未来波动、尾部风险等存在预测指示作用，选取期权市场隐含信息的代表性指标，包括波动率指数（*VIX*）、认沽认购比（*PCR*）、波动率风险溢价（*VRP*）、偏度指数（*SKEW*）、隐含的升贴水率等作为解释变量（详见附录 1），实证检验期权交易者是否为“聪明”交易者以及聪明体现在哪些方面等问题。

1. 已实现波动

低频时间序列领域有 GARCH 模型体系和 SV 模型体系两类代表性波动度量方法，但是都需要进行复杂的参数估计。随着高频金融时间序列容易获得，Andersen 等（2003）利用高频时间序列更准确、更方便地估计波动率，并提出了“已实现波动”（*RV*, Realized Volatility）的波动度量方法。

2. 波动率指数

波动率指数又称“恐慌指数”，是预测市场的重要指标之一，反映了交易者对未来已实现波动的预期。该指数是期权市场最具代表性的指标之一。

3. 偏度指数

偏度指数又称黑天鹅指数，按照无模型原理编制（Bakshi et al., 2003），综合反映交易者对未来一段时间标的市场发生极端风险的可能性。波动率指数作为平均波动率难以刻画极端风险的细致特征，偏度指数则可以反映大幅上涨或下跌的方向性预期，可以度量市场对极端风险的预期。

4. 波动率风险溢价

波动率风险溢价是隐含波动率与已实现波动之间的差值。基于期权价格的隐含波动率反映交易者预期波动率水平，基于标的资产历史价格的已实现波动率是标的市场真实波动率的测度。研究表明该指标可以刻画期权交易者的避险情绪。

5. 认沽认购比

认沽认购比是认沽持仓与认购持仓之比，综合反映交易者对后市看涨看跌的态度，认沽认购比越高，期权交易者避险意愿越强烈，看空后市的意愿越强烈，反之亦然。

6. 波动率价差

波动率价差（*CPIV*）是认购与认沽期权隐含波动率之差，反映认购期权和认沽期权的供求关系。

7. 期权隐含升贴水

期权合成期货升贴水是期权合成期货头寸与指数之间的价差，可以反映交易者看涨和看跌情绪的强烈程度。当市场看涨情绪强烈，期权合成期货升水，反之则贴水（Guo et al., 2013）。

8. 重要宏观经济信息

大量研究表明，期权交易者专业性程度高，对重要信息的反应较股票交易者更为灵敏。本研究选取 *GDP*、*CPI*、*PMI* 等重要的宏观经济指标作为宏观事件变量，研究境内期权交易者对定期发布的宏观信息的感知能力。

（二）实证设计

为了检验期权交易者是否“聪明”及其具体体现，本研究选取代表性的自回归模型、HAR 模型、Logit 模型和机器学习方法，分别检验期权隐含信息对指数收益率、已实现波动、尾部风险的预测能力，进而从理论层面呼应“价格发现说”“风险感知说”两种学说。具体实证框架如下：

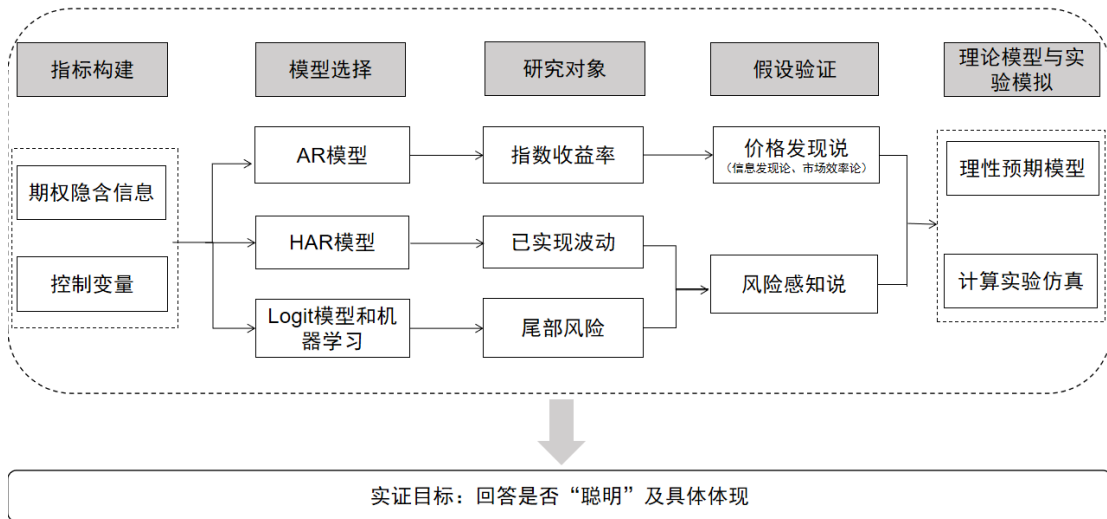


图 1 研究总体框架设计示意图

四、实证检验

沪深 300 指数由市值最大的 300 只股票构成，被视为 A 股投资的业绩基准，代表性强，相应 ETF 期权于 2019 年 12 月上市，至今已经近四年，日均成交量

超过 100 万张，样本期长、成交活跃，是最佳的期权实证样本，同时本文将另一只上市超过四年、日均成交超过 100 万张的上证 50ETF 期权选为稳健性检验样本。本研究相关期权指标选取成交更为活跃的上交所沪深 300ETF 期权进行计算¹，数据来源为 Wind，指标均通过期权公开数据计算得到。本文以沪深 300ETF 期权上市以来（2019 年 12 月 23 日-2023 年 7 月 31 日）数据作为样本，共计 868 个交易日数据，覆盖 1508 个期权合约，其中指数涨幅超过 2% 的 44 个交易日，跌幅超过 2% 的 39 个交易日，涵盖左尾和右尾各 5% 的交易日，与许多文献中尾部风险划分标准一致。分析所有解释变量和被解释变量相关性（描述性和相关性统计见附表 1、2 和 3），发现下一个交易日大涨大跌与前一个交易日的波动率风险溢价呈现一定负相关（-0.22），与前一个交易日的已实现方差呈现一定正相关（0.29）；下一个交易日的走势（指数收益率）与认沽认购比和隐含偏度均呈现一定正相关（0.31 和 0.25）；下一个交易日的已实现波动与波动率指数和基差率呈现一定正相关，相关性分别为 0.55 和 0.23，与认沽认购比、波动率风险溢价和隐含偏度呈现一定负相关，相关性分别为-0.21、-0.85 和-0.37。上述指标中解释变量之间相关性均低于 0.3，不存在多重共线问题。为了检验期权隐含信息对指数收益率、波动率以及尾部风险的影响机理，本研究在观察分析上述相关性较高的期权市场指标外，还将控制现有学术文献中的代表性期权指标以及现货指标，进一步通过实证分析来检验价格发现说和风险感知说。

（一）价格发现说

1. 价格发现说的实证

选取沪深 300 指数的每日收益率作为被解释变量，选取滞后一期的指数收益率 $Return$ 、波动率指数 VIX 、波动率风险溢价 VRP 、偏度指数 $SKEW$ 、波动率价差 $CPIV$ 、认沽认购比 PCR 、期权隐含升贴水 BR 等期权隐含信息指标作为解释变量，进行多元回归，结果见表 1 模型 1。同时，进一步选取滞后一期的市盈率 PE 、换手率 TR 等股票市场指标作为控制变量，进行多元回归，结果见表 1 模型 2。

$$Return_{t+1} = \beta_0 + \beta_1 Return_t + \beta_2 VIX_t + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 CPIV_t + \beta_6 PCR_t + \beta_7 BR_t + \beta_8 PE_t + \beta_9 TR_t + \varepsilon_{t+1} \quad (1)$$

模型 1 中除偏度指数显著外（系数为-0.0328，在 5% 的显著性水平下显著），波动率指数、波动率风险溢价、波动率价差、认沽认购比、期权隐含升贴水等期权隐含信息指标均不显著，表明除了隐含偏度外，其他期权隐含信息指标对下一

¹ 考虑到上交所和深交所分别上市了 1 个沪深 300ETF 期权品种，两个品种价格高度关联，成交持仓存在一定差异，所以与“价”相关的期权指标选取成交活跃的沪市期权计算，与成交量相关的认沽认购比指标合并两个沪深 300ETF 期权计算。

期的指数收益率不具备显著的预测作用。加入市盈率和换手率两个股票市场控制指标后，模型 2 中的偏度指数系数为-0.0462，依然在 5%的显著性水平下显著。期权隐含信息中，除了偏度指数外，其他期权隐含信息指标均不显著，其主要原因是偏度指数能够反映出交易者忽视的尾部风险，信息反映更加灵敏。进一步将沪深 300 指数走势划分为上行阶段（2019 年 12 月 23 日至 2021 年 2 月 9 日）和下行阶段（2021 年 2 月 18 日至 2023 年 8 月 21 日），分别对两个阶段的期权隐含信息预测作用进行检验，结果见表 1。模型 2（上行）中偏度指数系数为-0.0565，在 5%显著性水平下显著，模型 2（下行）中偏度指数系数为-0.0422，在 5%显著性水平下显著。偏度指数在全样本、上行周期、下行周期的系数均为负数，且数值差异不大，表明偏度系数对收益率的预测能力较为稳健。同时，表 1 中模型 1 和模型 2 的拟合度均较低（调整的 R^2 分别为 0.0034 和 0.0119），说明期权隐含信息对指数收益率仅有一定预测能力。

表 1 期权隐含信息对指数收益率预测作用估计结果

变量	模型 1	模型 2	模型 3	模型 4
<i>Return</i>	0.0170 (0.4600)	0.0234 (0.64)	0.0295 (0.8036)	0.0368 (1.0080)
<i>VIX</i>	-0.0040 (-0.3584)	-0.0092 (-0.71)	-0.0097 (-0.7547)	-0.0091 (-0.7121)
<i>VRP</i>	-0.0035 (-0.5178)	-0.0014 (-0.1912)	-0.0041 (-0.5567)	-0.0030 (-0.4168)
<i>SKEW</i>	-0.0328** (-2.2324)	-0.0462*** (-3.0244)	-0.0313* (-1.9319)	-0.0346 (-2.1577)
<i>CPIV</i>	-0.8465 (-0.5242)	-1.8637 (-1.1351)	-1.3119 (-0.7997)	-0.9448 (-0.5759)
<i>PCR</i>	0.3709 (1.3188)	0.2706 (0.8478)	0.3050 (0.9601)	0.3119 (0.9870)
<i>BR</i>	-0.0577 (-0.2224)	0.0870 (0.3295)	-0.0175 (-0.0662)	-0.0547 (-0.2077)
<i>PE</i>		-0.0932*** (-2.7528)	-0.0882*** (-2.6180)	-0.0884*** (-2.6301)
<i>TR</i>		0.1032 (0.2851)	0.0691 (0.1919)	0.0923 (0.2578)
<i>Event</i>			8.7301** (2.2959)	
<i>Event_SKEW</i>			-0.0820** (-2.2103)	
<i>Event_positive</i>				18.4925*** (3.2411)
<i>Event_negative</i>				-2.1824 (-0.4297)
<i>Event_positive_SKEW</i>				-0.1752*** (-3.1309)
<i>Event_negative_SKEW</i>				0.0226 (0.4558)
<i>Adj.R²</i>	0.0034	0.0119	0.0237	0.0334

注：“***”、“**”、“*”代表显著性水平为 1%、5%、10%，括号内标注 t 值（下同）。

期权隐含信息指标可以揭示期权交易者对未来市场的判断。信息的扩散和市场共识的形成存在一定时滞，偏度指数会领先波动率指数做出反应，因此偏度指数更能捕捉尾部风险因素。波动率指数是交易者对波动率预期的“温度计”，偏度指数则是“放大镜”。总的来看，实证结果支持价格发现说，期权市场的偏度指数蕴含了未来市场价格变化的信息，其对下一交易日的指数收益具有稳健的负向预测作用，与陈蓉、林秀雀（2016）研究结论一致。

2. 信息发现说的实证

为了检验信息发现说，考察期权交易者对定期的重要宏观信息发布的感知能力，我们在模型（1）中进一步引入事件变量 $Event_t$ 。若某一交易日存在重要宏观经济数据发布²（GDP、CPI、PPI、PMI、工业增加值、社会零售销售总额等，具体结果见附表 4），则将事件变量设定为 1，否则设定为 0。同时，引入事件变量与偏度指数的交乘项 $Event_SKEW_t$ ，检验重要宏观经济信息发布是否会影响偏度指数的预测能力。表 1 模型 3 表明事件发布对指数收益率影响显著为正。表 1 模型 3 交乘项系数为-0.082，在 5%水平下显著，表明重要宏观经济信息发布日偏度指数对收益率的预测能力较强。结果证实了信息发现说，表明期权交易者对重要宏观经济信息具有显著感知力，将相关信息在期权市场提前定价。

$$Return_{t+1} = \beta_0 + \beta_1 Return_t + \beta_2 VIX_t + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 CPIV_t + \beta_6 PCR_t + \beta_7 BR_t + \beta_8 PE_t + \beta_9 TR_t + \beta_{10} Event_t + \beta_{11} Event_SKEW_t + \varepsilon_{t+1} \quad (2)$$

进一步，将事件变量分类，发布的宏观数据达到或者超过市场预期（取自 Wind），将该事件定义为积极事件，若不及预期，则定义为负面事件。同样的，在模型中引入事件变量与偏度指数的交乘项，检验“积极”和“负面”事件变量对偏度指数预测能力的影响。

$$Return_{t+1} = \beta_0 + \beta_1 Return_t + \beta_2 VIX_t + \beta_3 VRP_t + \beta_4 SKEW_t + \beta_5 CPIV_t + \beta_6 PCR_t + \beta_7 BR_t + \beta_8 PE_t + \beta_9 TR_t + \beta_{10} Event_positive_t + \beta_{11} Event_negative_t + \beta_{12} Event_positive_SKEW_t + \beta_{13} Event_negative_SKEW_t + \varepsilon_{t+1} \quad (3)$$

表 1 模型 4 中的“积极”事件交乘项系数为-0.1752，在 1%显著性水平下显著，表明“积极”事件的发布增强了偏度指数对收益率的预测能力；“负面”事件交乘项系数为 0.0226，不显著，表明“负面”事件的发布对偏度指数预测能力的影响中性。“积极”与“负面”事件的影响具有明显的不对称性，说明期权交易者对“积极”事件反应更为灵敏。本研究认为这种不对称影响可能与中国期权市场发展阶段有关，当前境内期权市场还不够成熟，其市场参与者中散户还占有

² GDP、CPI、PPI 等宏观经济指标定期公布，受到市场广泛关注。其中 GDP 按季度发布，CPI、PPI、PMI 等其他指标按月度发布。数据发布前，市场会对相应宏观指标形成不同，Wind 汇总了不同机构预测值，形成市场整体预期。本研究比较真实值和预期值，若真实值大于等于预期值，则表示宏观经济数据超预期，反之则表示低于预期。

重要部分（3成），其中的投机交易者往往更加倾向于在利好消息发生时利用期权杠杆性博取方向性收益，因此预期事件“积极”时反应会比“负面”时更加敏感。

3. 市场效率说的实证

文献综述提到，市场效率论还是信息发现论更有解释力尚无定论。根据市场效率论，当新信息出现后，聪明交易者会选择流动性更好、效率更高的市场或者交易工具进行交易。据此我们推论，如果市场效率论成立，那么流动性好、交易成本低的期权合约上应该包含更多的信息，对指数收益的预测作用应该更加明显。通常而言，偏度指数是选用流动性好的近月和次月合约进行编制（结果见表1），而远月合约由于流动性较差，不参与偏度指数编制。为了比较不同流动性期权合约隐含的信息，我们进一步采用远月合约编制偏度指数，结果如表2所示。模型4结果显示，采用远月合约编制的偏度指数仍然具有一定的预测作用，模型调整 R^2 （0.0313）与近月合约编制的偏度指数模型的调整 R^2 （0.0334）之间差异并不明显。因此，本研究结论不支持市场效率论，结合前文信息发现论的支持证据，本文认为期权隐含信息对指数收益率的预测能力，主要源于新信息的获取和分析方面的差异。

表2 期权隐含信息对指数收益率预测作用估计结果（远月合约）

变量	模型 1	模型 2	模型 3	模型 4
<i>Return</i>	0.0027 (0.0746)	0.0026 (0.0717)	0.0167 (0.4582)	0.0230 (0.6328)
<i>VIX</i>	0.0035 (0.3175)	0.0020 (0.1570)	-0.0026 (-0.2028)	-0.0015 (-0.1177)
<i>VRP</i>	-0.0053 (-0.7849)	-0.0037 (-0.5196)	-0.0055 (-0.7696)	-0.0049 (-0.6811)
<i>SKEW</i>	-0.0145 (-1.2958)	-0.0237** (-2.0380)	-0.0202* (-1.7485)	-0.0192* (-1.6684)
<i>CPIV</i>	-0.9037 (-0.5582)	-1.8669 (-1.1312)	-1.3070 (-0.7960)	-0.9096 (-0.5532)
<i>PCR</i>	0.4160 (1.3716)	0.3854 (1.1357)	0.4442 (1.3170)	0.4303 (1.2821)
<i>BR</i>	-0.0628 (-0.2414)	0.0692 (0.2613)	-0.0314 (-0.1189)	-0.0743 (-0.2820)
<i>PE</i>		-0.0846** (-2.4948)	-0.0863*** (-2.5644)	-0.0841** (-2.5031)
<i>TR</i>		0.0907 (0.2498)	0.0506 (0.1406)	0.0773 (0.2158)
<i>Event</i>			10.8945*** (3.0467)	
<i>Event_SKEW</i>			-0.1031*** (-2.9530)	
<i>Event_positive</i>				20.6476*** (3.7044)
<i>Event_negative</i>				0.4552 (-0.0925)

<i>Event_positive</i>				-0.1962***
<i>_SKEW</i>				(-3.5895)
<i>Event_negative</i>				-0.0030
<i>_SKEW</i>				(-0.0621)
<i>Adj.R²</i>	0.0000	0.0062	0.0229	0.0313

(二) 风险感知说

1. 波动预测作用的实证

Corsi (2009) 首次提出了 HAR-RV 模型刻画日波动率的动态特征。HAR-RV 模型能够反映波动率的特征，包括长记忆、厚尾和自相似等，因而被广泛应用。该模型中日、周、月的三种已实现波动是由市场中三种不同的交易者所驱动的，HAR 模型结合了交易者异质性和叠加了不同期限的波动率。不同期限交易对应的已实现波动率之间有相互影响：

$$RV_{t+1} = \beta_0 + \beta_1 RV_t + \beta_2 RVW_t + \beta_3 RVM_t + \varepsilon_{t+1} \quad (4)$$

其中， RV_{t+1} 为未来 1 天的已实现波动率， RV_t 、 RVW_t 和 RVM_t 分别为日、周、月已实现波动率。周已实现波动率和月已实现波动率构建如下：

$$RVW_t = \frac{1}{5}(RV_t + RV_{t-1} + RV_{t-2} + RV_{t-3} + RV_{t-4}) \quad (5)$$

$$RVM_t = \frac{1}{22}(RV_t + RV_{t-1} + RV_{t-2} + \dots + RV_{t-20} + RV_{t-21}) \quad (6)$$

为了探究期权市场隐含信息对已实现波动的影响机理，本研究在经典的 HAR-RV 框架下，进一步引入波动率指数、偏度指数、波动率风险溢价、认沽认购比等期权市场代表性指标，并对指数相关特征进行控制，检验期权市场信息是否有助于增强已实现波动的预测能力。公式如下所示：

$$RV_{t+1} = \beta_0 + \beta_1 RV_D_t + \beta_2 RV_W_t + \beta_3 RV_M_t + \beta_4 VIX_t + \beta_5 \Delta VIX_t + \beta_6 PCR_t + \beta_7 SKEW_t + \beta_8 CPIV_t + \beta_9 BR_t + \beta_{10} PE_t + \beta_{11} TR_t + \varepsilon_{t+1} \quad (7)$$

其中 RV_{t+1} 为未来 1 天的已实现波动率， RV_t 、 RVW_t 和 RVM_t 分别为日、周、月已实现波动率，本文分别选用过去 5 天、22 天已实现波动率的平均值作为周、月已实现波动率。 VIX_t 、 ΔVIX_t 、 PCR_t 、 $SKEW_t$ 、 $CPIV_t$ 和 BR_t 分别为波动率指数、波动率指数变动、认沽认购持仓比、偏度指数、波动率价差、期权隐含升贴水等期权市场隐含信息指标。 PE_t 和 TR_t 分别为市盈率和换手率等指数特征指标。

表 3 期权隐含信息对指数波动预测作用估计结果

变量	模型 1	模型 2	模型 3
RV_D	0.2023*** (5.0184)	0.1491*** (3.4918)	0.1224*** (2.8208)

<i>RV_W</i>	0.4071 ^{***} (5.6909)	0.2366 ^{***} (3.2083)	0.1783 ^{**} (2.3635)
<i>RV_M</i>	0.1438 [*] (1.8746)	-0.0541 (-0.5908)	-0.0340 (-0.3707)
<i>VIX</i>		0.6099 ^{***} (6.6718)	0.5667 ^{***} (6.1706)
ΔVIX		0.6316 ^{***} (2.9488)	0.5822 ^{***} (2.7251)
<i>PCR</i>		-4.6643 ^{***} (-3.0154)	-7.6228 ^{***} (-4.2541)
<i>SKEW</i>		0.1113 (1.3628)	0.1255 (1.4966)
<i>CPIV</i>		1.1948 (0.1381)	0.8403 (0.0952)
<i>BR</i>		0.7278 (0.5233)	0.2431 (0.1722)
<i>PE</i>			-0.2568 (-1.3963)
<i>TR</i>			6.6127 ^{***} (3.3024)
<i>Adj.R²</i>	0.2553	0.3051	0.3124

表 3 给出了 HAR-RV 族模型对未来波动率预测作用的估计结果。首先，在标准 HAR-RV 模型中，日、周已实现波动率能够有效预测未来日波动率，且模型解释力较强， R^2 达到 0.2553，这表明市场存在明显的波动聚集效应。其次，在 HAR-RV 模型中引入期权隐含信息，能够明显改善模型对未来日波动率的预测效果。其中，波动率指数和波动率指数变动对未来日波动率具有正向的预测作用，这表明期权价格信息包含了的交易者对未来波动的预期，交易者通过期权对波动率进行交易，能够更加准确地对波动率进行定价；认沽认购持仓比对未来日波动率具有反向的预测作用，这表明期权持仓结构信息包含了交易者对未来波动的预期。最后，在 HAR-RV 模型中对指数的相关特征进行控制后，期权隐含信息对未来日波动率仍然具有预测作用。

2. 尾部风险预测作用的实证

本节构建 Logit 模型实证对股票市场尾部风险（价格大幅波动风险）的预测指示作用。其模型假设为：

$$Y_i = X_i' \beta + \varepsilon_i, i = 1, 2, \dots, M$$

其中， X_i 是包含常数项的多元解释变量， Y_i 是二元变量：

$$Y_i = \begin{cases} 1, & \text{涨跌幅超过设定阈值} \\ 0, & \text{涨跌幅不超过设定阈值} \end{cases}$$

假设 $\text{Prob}(Y_i = 1 | X_i) = F(X_i' \beta)$ 定义为 Logistic 分布函数，则为 Logit 模型。

$$\begin{aligned}
Big_Movement_{t+1} = & \beta_0 + \beta_1 VIX_t + \beta_2 \Delta VIX_t + \beta_3 PCR_t + \beta_4 SKEW_t \\
& + \beta_5 CPIV_t + \beta_6 BR_t + \beta_7 PE_t + \beta_8 TR_t + \varepsilon_{t+1}
\end{aligned} \tag{8}$$

选取沪深 300 指数涨幅处于前 5%（涨幅约为 2%）和跌幅处于前 5%（跌幅约为 -2%）的尾部事件样本，将涨跌超过 2% 的交易日赋值为 1，其他交易日赋值为 0。首先，选取波动率指数等期权隐含信息指标检验尾部风险的预测能力，具体结果见表 4 模型 1。结果表明，波动率指数的系数为 0.0669，波动率风险溢价的系数为 -0.0435，均在 1% 的显著性水平下显著，表明期权隐含信息对市场大涨和大跌的尾部风险具有一定预测能力。其次，在模型 1 中引入市盈率、换手率等现货市场指标作为控制变量，发现波动率指数、波动率风险溢价的系数变化不大，依然在 5% 的显著性水平下显著（见模型 2），表明期权隐含信息对市场大涨大跌的预测能力较为稳健。

在表 4 模型 2 的基础上，对大涨和大跌进行分类，分别进行估计，结果见模型 3（大涨）和模型 4（大跌）。模型 3 中，波动率指数和偏度指数的系数均在 5% 水平下显著，表明期权市场隐含波动的大幅上升以及偏度指数的大幅下降，均意味着下一个交易日沪深 300 指数更可能出现超过 2% 的上涨。模型 4 中，波动率指数、认沽认购比、波动率风险溢价以及偏度指数均在 5% 显著性水平下显著，波动率指数、偏度指数以及认沽持仓的快速上升可能意味着下一个交易日沪深 300 指数可能出现超过 2% 的下跌。

上述研究结果表明，期权隐含信息对市场尾部风险，特别是大幅下跌风险具有良好的预测能力。本研究结论与已有隐含信息预测能力的实证结果是一致的，也符合金融原理，即波动率指数、偏度指数、认沽认购比均蕴含了交易者对未来市场波动的预期信息，当交易者感知到市场可能出现大幅下跌时，持有现货交易者和投机交易者愿意以更贵的价格买入更多认沽期权来为现货资产进行保险，进而推高波动率指数、偏度指数和认沽认购比。期权隐含信息对于市场大涨和大跌的预测能力具有不对称性，上涨时期波动率推高效应要强于下跌时期波动率压低效应，因为大涨时期仅有纯投机交易的交易者会买入更多认购期权，而持有现货的交易者并不需要为资产进行保险。

表 4 期权隐含信息对指数大涨大跌预测作用估计结果

变量	模型 1	模型 2	模型 3（上涨）	模型 4（下跌）
<i>VIX</i>	0.0669*** (3.3689)	0.0602*** (2.6612)	0.1187*** (2.7039)	0.1116*** (2.6037)
ΔVIX	0.1118* (1.6709)	0.1066 (1.5925)	0.0549 (0.4729)	0.0289 (0.2385)
<i>PCR</i>	0.2196 (0.4525)	-0.0301 (-0.0541)	0.9341 (0.8043)	-3.4503*** (-2.6143)
<i>VRP</i>	-0.0435*** (-2.9777)	-0.0381** (-2.4390)	-0.0446* (-1.8396)	-0.0448** (-2.1114)

<i>SKEW</i>	-0.0088 (-0.3266)	0.0059 (0.2101)	-0.1404** (-2.3704)	0.1177** (2.2192)
<i>CPIV</i>	0.7903 (0.2837)	1.5757 (0.5537)	-3.8327 (-0.6016)	-5.9277 (-0.9715)
<i>BR</i>	-0.1961 (-0.4251)	-0.3403 (-0.7269)	0.5014 (0.4885)	0.9205 (0.9504)
<i>PE</i>		0.0500 (0.8469)	-0.2312 (-1.6237)	0.2625* (1.9580)
<i>TR</i>		0.6147 (0.9687)	-1.0211 (-0.7497)	1.0893 (0.7745)
<i>Adj.R²</i>	0.0312	0.0328	0.0654	0.0524

(三) 机器学习的进一步检验

为了进一步验证风险发现说，本节以机器学习方法构建预测模型，实证检验期权隐含信息对极端下跌风险的预警功能。

1. 机器学习模型设计

相比 Logit、KLR 模型等传统计量模型，机器学习模型预测更准确，更适应数据丰富的金融市场模型，因此本文进一步应用有监督学习的机器学习模型研究期权隐含信息对极端风险的预测作用。本文选择主流的支持向量机（SVM）、判别分析（DA）、朴素贝叶斯（NBM）、随机森林（RF）、逻辑回归（LR）³ 5 个机器学习模型。限于篇幅不再赘述模型原理。

基于前文对尾部风险的计量模型分析，本模型选择显著的 *VIX*、 ΔVIX 、*PCR*、*VRP*、*SKEW*、*PE* 这六个变量，作为机器学习的特征变量。尾部风险指标使用了常用的股市风险系数 *CMD*（Coefficient of Maximum Drawdown）。先计算 *MD*（Maximum Drawdown），该指标是当前价格与过去一段时间内（*k* 期）最高价格之比，用来表示股市下跌程度或回撤程度（Coudert and Gex, 2008）：

$$MD_t = \frac{P_t}{\text{Max}(p_{t-k}, p_{t-k+1}, \dots, p)}$$

其中， P_t 表示第 *t* 日收盘股价，本文将滞后期数 *k* 设置为 252 天。在 *MD* 的基础上，计算得到股市风险系数：

$$CMD = \begin{cases} 1 & \text{如果 } MD_t < \overline{MD}_t - \sigma_t \\ 0 & \text{其他} \end{cases}$$

其中 \overline{MD}_t 和 σ_t 分别为 *k* 期 *MD* 序列的均值、标准差。风险系数 *CMD* 将 *MD* 序列分为两类：1 为发生大幅下跌事件，0 则为不发生。

样本时间选取方面，为保证预测检验次数充足，本文使用了区间后 500 个样本点进行预测检验，使用区间前期数据（368 个样本点）进行模型训练。按照经

³ 支持向量机（Support Vector Machine）、判别分析（Discriminant Analysis）、随机森林（Random Forest）、朴素贝叶斯（Naive Bayesian Model）、逻辑回归（Logistic Regression）。

典做法，混淆矩阵（表 5）根据模型预测值与真实值的对比得到。模型预测准确度高，对应到混淆矩阵中，即是 TP 与 FN 的数量大，而 FP 与 TN 的数量小。预测性分类模型的评价效果主要看预测准确度，评价指标见表 6。

表 5 混淆矩阵示意图

混淆矩阵		真实值	
		真 (True)	假 (False)
预测值	发出信号 (Positive)	True Positive (TP)	False Positive (FP)
	未发出信号 (Negative)	True Negative (TN)	False Negative (FN)

表 6 机器学习模型评价指标

指标	公式	含义
<i>Precision</i> (精确率)	$TP/(TP+FP)$	模型预测是 Positive 的结果中，结果正确的比重
<i>Recall</i> (召回率)	$TP/(TP+TN)$	在真实值是 True 的结果中，预测正确的比重
<i>Accuracy</i> (准确率)	$(TP+FN)/(TP+FP+TN+FN)$	分类模型中所有判断正确的结果占总样本比重
<i>F1 Score</i>	$2 * \frac{Precision * Recall}{Precision + Recall}$	Precision 和 Recall 的调和平均数
<i>AUC</i>	$\frac{\sum_{i=1}^M (Rank(p_i) - i)}{M * N}$	<i>AUC</i> ⁴ 取值高于 0.5 即认为具备预测能力，越接近 1 预测能力越强

2. 模型实证结果

对 5 个模型机器学习训练并进行预测检验，结果见表 7。从结果来看，SVM 模型综合评分最高，其 *Precision* 达到 0.75，意味着模型发出预警信号的预测成功率为 75%，即发出 4 次预警有 3 次确实出现了尾部风险，其 *Recall* 达到 0.35，意味着模型能够识别出 35% 的风险事件，即成功预测超过 1/3 的下跌极端事件。同时，整体模型预测准确率达到 65%，这都证明了期权隐含信息对市场尾部风险具有较好的预警能力，已经具备了一定的预警应用价值。

总体而言，本节基于机器学习的风险预警模型的实证，证实了期权市场相关指标对股市风险具有较好的预测能力。

表 7 机器学习预测结果

模型名称	AUC	Accuracy	Precision	Recall	F1 Score
SVM	0.63	0.65	0.75	0.35	0.48
DA	0.62	0.65	0.81	0.30	0.44
RF	0.61	0.64	0.85	0.26	0.40
NBM	0.55	0.58	0.61	0.22	0.32
LR	0.57	0.60	0.97	0.13	0.24

注：准确率最高的为支持向量机模型，除了精度指标，其它指标均最高。

⁴ AUC(Area Under the ROC Curve) 是 ROC (Receiver Operating Characteristic) 曲线下的面积。

五、稳健性检验

采用两种方式检验前文研究结果的稳健性。首先，将沪深 300ETF 期权数据划分为上行阶段和下行阶段，分段估计不同时段结论稳健性。其次，选取境内上市时间最早的上证 50ETF 期权作为对照样本，选取与沪深 300ETF 期权同时段数据，估计收益率、已实现波动、大涨大跌三个模型结果稳健性。

（一）分时段的稳健性检验

沪深 300ETF 期权上市以来，沪深 300 指数呈现先涨后跌走势，其中 2019 年 12 月 23 日至 2021 年 2 月 9 日为上行阶段，2021 年 2 月 18 日至 2023 年 7 月 31 日为下行阶段，选取模型（1）分别检验两个阶段期权隐含信息对收益率的预测指示作用，结果见表 8。上行阶段中偏度指数系数为-0.0565，下行阶段中偏度指数系数为-0.0422，两个阶段均在 5%显著性水平下显著。偏度指数在全样本、上行周期、下行周期估计系数均为负数，且数值接近，表明偏度系数对收益率的预测能力较为稳健。

表 8 上行和下行阶段期权隐含信息对指数收益率预测作用

变量	全样本	上行阶段	下行阶段
<i>Return</i>	0.0602*** (2.6612)	0.0715 (1.0406)	-0.0040 (-0.0905)
<i>VIX</i>	0.1066 (1.5925)	-0.0547** (-2.2246)	0.0150 (0.7528)
<i>VRP</i>	-0.0301 (-0.0541)	0.0124 (1.0422)	-0.0161 (-1.3791)
<i>SKEW</i>	-0.0381** (-2.4390)	-0.0565** (-2.1182)	-0.0422** (-2.0563)
<i>CPIV</i>	0.0059 (0.2101)	-3.8250 (-1.1701)	-0.5923 (-0.2999)
<i>PCR</i>	1.5757 (0.5537)	0.0592 (0.0885)	0.1173 (0.3045)
<i>BR</i>	-0.3403 (-0.7269)	-0.0728 (-0.4233)	0.1767 (0.5099)
<i>PE</i>	0.0500 (0.8469)	-0.0728 (-0.8271)	-0.0702* (-1.9604)
<i>TR</i>	0.6147 (0.9687)	1.0874 (1.3897)	-0.3033 (-0.6954)
<i>Adj.R²</i>	0.0328	0.0231	0.0079

选取模型（7）分别检验两个阶段期权隐含信息对已实现波动的预测指示作用，结果见表 9。上行阶段中，波动率指数系数为 0.9444，认沽认购比系数为 -10.7418，均在 5%显著性水平下显著；下行阶段中，波动率指数系数为 0.7646，认沽认购比系数为-6.5001，均在 5%显著性水平下显著。波动率指数、认沽认购比两类期权市场指标在全样本、上行周期、下行周期的系数数值和显著性差异很

小，表明结果较为稳健。

表 9 上行和下行阶段期权隐含信息对指数波动预测作用

变量	全样本	上行阶段	下行阶段
<i>RV_D</i>	0.12*** (2.82)	0.1508* (1.7147)	0.0877* (1.7784)
<i>RV_W</i>	0.18** (2.36)	-0.0090* (-0.0598)	0.2751*** (3.4408)
<i>RV_M</i>	-0.03 (-0.37)	-0.4377* (-1.9510)	-0.0799 (-0.8865)
<i>VIX</i>	0.57*** (6.17)	0.9444*** (3.7873)	0.7646*** (6.7932)
ΔVIX	0.58*** (2.73)	0.2167 (0.5032)	0.8277*** (3.8823)
<i>PCR</i>	-7.62*** (-4.25)	-10.7418** (-2.0452)	-6.5001*** (-4.5606)
<i>SKEW</i>	0.13 (1.50)	0.1880 (0.9617)	0.0710 (0.9327)
<i>CPIV</i>	0.84 (0.10)	14.8551 (0.6374)	-0.7870 (-0.1074)
<i>BR</i>	0.24 (0.17)	0.6405 (0.1896)	-0.8695 (-0.6582)
<i>PE</i>	-0.26 (-1.40)	-1.4002** (-2.0854)	-0.2430* (-1.7924)
<i>TR</i>	6.61*** (3.30)	9.2146 (1.5856)	5.0636 (3.0507)
<i>Adj.R²</i>	0.31	0.2017	0.4976

选取模型（8）分别检验上行阶段和下行阶段期权隐含信息对大涨大跌的预测指示作用，结果见表 10。上行阶段中，波动率指数系数为 0.0695，在 10%显著性水平下显著；下行阶段中，波动率指数系数为 0.0840，在 5%显著性水平下显著。波动率指数在全样本、上行周期、下行周期的系数数值和显著性差异不大，表明结果较为稳健。

表 10 上行和下行阶段期权隐含信息对指数大涨大跌预测作用

变量	全样本	上行阶段	下行阶段
<i>VIX</i>	0.1187*** (2.7039)	0.0695* (1.7417)	0.0840** (2.1880)
ΔVIX	0.0549 (0.4729)	0.2342** (2.3817)	-0.0391 (-0.3892)
<i>PCR</i>	0.9341 (0.8043)	1.6428 (1.5710)	-0.8376 (-1.1856)
<i>VRP</i>	-0.0446* (-1.8396)	-0.0779*** (-2.6011)	-0.0157 (-0.7371)
<i>SKEW</i>	-0.1404** (-2.3704)	-0.0683 (-1.5981)	0.0707* (1.8585)
<i>CPIV</i>	-3.8327 (-0.6016)	5.4731 (1.0608)	1.1331 (0.3099)
<i>BR</i>	0.5014	-0.4556	-0.4406

	(0.4885)	(-0.5486)	(-0.6948)
<i>PE</i>	-0.2312 (-1.6237)	0.1570 (1.1114)	0.0193 (0.2828)
<i>TR</i>	-1.0211 (-0.7497)	-0.9507 (-0.7628)	1.0790 (1.3199)
<i>Adj.R²</i>	0.0654	0.0760	0.0197

综合来看，指数上行阶段、下行阶段以及全样本中，期权隐含信息对指数收益率、已实现波动和大涨大跌具有不同程度预测能力，且结果稳健。一是隐含偏度对指数收益率具有较为一致的预测能力，表明期权交易者对于指数走势的预判，更多通过对虚值期权的偏好体现，当然模型解释力总体较弱。二是波动率指数和认沽认购比对已实现波动具有较为一致的预测能力，且解释力较强，与期权是波动率交易与定价工具的普遍认识一致。三是波动率指数对大涨大跌具有较为一致的预测能力，表明波动率指数作为期权市场的情绪指标，能够提前探知现货市场“温度”，并提前在期权价格中反映。

（二）对照样本的稳健性检验

本部分利用境内首只期权品种——上证 50ETF 期权数据，进一步检验收益率、波动率、尾部风险三个模型的稳健性，结果见附表 5。

从收益率来看，模型所有变量均不显著。上证 50ETF 期权的实证结果之所以跟沪深 300 估计结果有一定差异，可能在于上证 50 指数本身波动较小，其指数集中于金融股和石油股等权重股，波动性和偏度均较低，降低了偏度等指标对收益率的预测作用。

从波动率来看，模型中的日已实现波动率、周已实现波动、波动率指数、波动率指数变动、认沽认购比、换手率等变量均在 1% 的显著性水平下显著，表明期权隐含信息对上证 50 指数的波动率存在稳健的预测作用。

从尾部风险来看，期权隐含信息对上证 50 指数的尾部风险存在一定的预测作用，且上涨和下跌存在一定差异。市场大涨情形，波动率指数和偏度指数的系数均在 10% 的显著性水平下显著。市场大跌情形，波动率指数和认沽认购比的系数均在 10% 的显著性水平显著。总的来说，波动率指数和认沽认购比对尾部风险具有稳健的预测作用，偏度指数仅在市场上涨情况下具有稳健的预测作用。

六、理性预期均衡模型与计算实验仿真

前文实证证明了期权交易者是“聪明”的，对市场价格变化和 risk 均有一定预测感知能力，并且结果支持信息发现说。因此，本部分从信息出发，构建了基于信息的两类交易者理性预期模型，可以解释期权市场有预测感知能力的机制，

并通过计算实验仿真进一步证明期权市场的存在,可以提升股票市场价格发现和风险定价能力。

(一) 理论模型

金融市场由信息驱动的,与传统资产定价模型假设不同,现实市场不同交易者拥有不同信息,一些交易者拥有更多信息,信息的精度更高,信息分析能力也更强。非对称信息资产定价模型放松信息对称假设,围绕信息的构建模型。Grossman (1976) 首先建立了一个拥有不同信息的交易者的静态理性预期均衡模型。Wang (1993) 建立了理性交易者和非理性交易者两类交易者的动态模型,解释价格序列自相关、价格过度波动问题。

本部分理论模型借鉴该思路,引入信息构建模型分析基于信息不同的两类交易者的价格形成机制。按照理性预期均衡 (REE) 分析框架,在 Veronesi (1999) 理性均衡的资产定价模型 (引理 1) 的基础上,本文采用非对称信息资产定价理论的经典方法,建立基于两类交易者的理性预期模型,刻画跨期现市场交易者⁵ (以下简称期现交易者) 的存在,提升了股票市场的预测感知能力。

1. 基于股利的资产定价标准模型

假设股利过程遵循均值 θ 的随机布朗运动,本节构建了理性预期均衡的资产定价模型作为本节的基准模型。模型公式表明了证券供给和均衡价格之间的关系。

引理 1: 假设 1, 股利过程遵循随机微分等式

$$dD = \theta dt + \sigma dW \quad (9)$$

θ 为均值漂移项,即平均股利; σ 为波动率; dW 遵循标准维纳过程。

假设 2, 交易者效用函数为常系数绝对风险厌恶函数 (CARA), 证券供给是 X , 无风险资产收益为 r 。

$$U(c, t) = -e^{-\rho t - ac} \quad (10)$$

ρ 为时间偏好系数; a 为绝对风险厌恶系数。

那么, 存在理性预期均衡价格和投资需求分别满足等式

$$P(D, \theta) = \frac{D}{r} + \frac{\theta}{r^2} - \frac{a\sigma^2 X}{r^2} \quad (11)$$

$$X = \frac{rD - r^2 P + \theta}{a\sigma^2} \quad (12)$$

详细证明见 Veronesi (1999), 本文直接引用作为引理。公式表明资产价格由基础价值和风险补偿项组成,前者是当期的股利和未来股利折现值之和,后者

⁵ 本文是指同时交易期权和股票的交易者。

是未来股利方差与风险厌恶系数折现得到，表明股利波动越大，均衡价格越低。

2. 信息的贝叶斯更新

在（9）式基础上引入信息更新，假设影响股利过程的新信息 $\{Y_1, Y_2, \dots, Y_n\}$ 服从正态分布 $N(\theta, \sigma^2)$ ，其中 σ^2 已知。同时，假设股利漂移均值项的先验分布 θ 服从 $N(\mu_0, \sigma_0^2)$ 。接受 n 次信息后， θ 的后验分布遵从正态分布：

$$\theta|Y \sim N\left(-\frac{\mu_0 v_1}{v_1 + v_2} + \frac{\bar{Y} v_2}{v_1 + v_2}, \frac{1}{v_1 + v_2}\right) \quad (13)$$

其中， v_1 为先验均值的精度， v_2 为信息的精度，精度是对应方差的倒数，即 $v_i = 1/\sigma_i^2$ 。其经济含义是，获得新信息后，交易者按贝叶斯更新将信息融入先验值 μ_0 ，其后验值是交易者先验值和新信息均值的加权平均。交易者的信息集方差越小，精度越高，权重越高，新信息对交易者影响就越大。

3. 基于信息的两类交易者理性预期模型

跨期现市场交易者⁶是指同时在期权和股票两个市场交易的交易者。在（13）式基础上，假设期现交易者和现货交易者两类交易者，两者获取信息的信息集不同，信息的分析解读能力也存在差异，因此信息发生后期现交易者将 θ 分别更新为 y_n 和 y_u ，股票交易者的信息集的精度为 v_n ，期现交易者的信息集的精度 v_u ，先验均值的精度为 v 。

定理： 存在期现交易者和现货交易者两类交易者，期现交易者的占比为 ϕ ，股票交易者的占比为 $1-\phi$ 。两类交易者分别从各自信息集获取信息，且信息集 $\{Y_1, Y_2, \dots, Y_n\}$ 均服从正态分布，期现交易者信息集均值为 y_n ，精度为 v_u ，股票交易者信息集均值为 y_u ，精度为 v_n ，信息集先验值一致。那么，理性预期均衡价格为：

$$E(P^*) = \frac{1}{r^2} \left[rD + \bar{\theta} + (1 - \phi) \frac{v_n(\bar{y}_n - \bar{\theta})}{v + v_n} + \phi \frac{v_u(\bar{y}_u - \bar{\theta})}{v + v_u} - \alpha \sigma^2 \xi \right] \quad (14)$$

证明：

在均衡价格 P^* 下，两类交易者通过调整各自的投资达到市场均衡，两类交易者的投资分别为：

$$X_n = \frac{rD - r^2 P^* + y_n}{\alpha \sigma^2}$$

$$X_u = \frac{rD - r^2 P^* + y_u}{\alpha \sigma^2}$$

在市场总需求 ξ 固定条件下：

⁶ 从更广泛的含义，期现交易者也可以理解为同时在期现两个市场交易，或者在进行现货交易时候会从期权市场学习或参考的交易者，从这个意义上，该模型适用更普遍的情况。

$$(1 - \phi)X_n + \phi X_u = \xi$$

求解股票价格为：

$$P^* = \frac{1}{r^2} [rD + (1 - \phi)y_n + \phi y_u - \alpha \sigma^2 \xi]$$

则均衡的股票价格可表示为：

$$E(P^*) = \frac{1}{r^2} \left[rD + (1 - \phi) \left(\frac{v}{v+v_n} \bar{\theta} + \frac{v_n}{v+v_n} \bar{y}_n \right) + \phi \left(\frac{v}{v+v_u} \bar{\theta} + \frac{v_u}{v+v_u} \bar{y}_u \right) - \alpha \sigma^2 \xi \right] \quad (15)$$

$$E(P^*) = \frac{1}{r^2} \left[rD + \bar{\theta} + (1 - \phi) \frac{v_n(\bar{y}_n - \bar{\theta})}{v+v_n} + \phi \frac{v_u(\bar{y}_u - \bar{\theta})}{v+v_u} - \alpha \sigma^2 \xi \right] \quad (16)$$

定理的经济含义为：（1）当交易者预测的股利水平 \bar{y} 大于 $\bar{\theta}$ 时，将推动均衡价格上升，反之则推动均衡价格下降。（2）交易者占比越高，该类交易者的信息对均衡价格影响越大。（3）交易者的信息精度对价格具有正向影响作用，精度越高对价格的影响越大。

以上结论（3）成立可以从下式中推出：

$$\frac{v_n}{v+v_n} = \frac{1}{v/v_n + 1}$$

因此具有信息优势交易者的交易行为隐含了更多关于未来价格的信息。同理，具有信息优势的交易者的交易行为隐含了更多关于未来波动的信息，当然更严谨的理论推导有待进一步研究。

推论：假设期现交易者的信息获取和分析能力强于股票交易者，期现交易者获取精度为 v_u 的信息 y_u ，而股票交易者没有获得信息。那么，均衡价格公式为：

$$E(P^*) = \frac{1}{r^2} \left[rD + \bar{\theta} + \phi \frac{v_u(\bar{y}_u - \bar{\theta})}{v+v_u} - \alpha \sigma^2 \xi \right] \quad (17)$$

从推论可以证明，当出现利好消息，期现交易者预测的股利水平 \bar{y}_u 大于 $\bar{\theta}$ 时，将推动均衡价格上升，反之出现利空消息， \bar{y}_u 小于 $\bar{\theta}$ 时，则推动均衡价格下降。这种关系就会体现在期权隐含信息指标领先现货市场价格变化。

（二）计算实验仿真

前文的研究发现，“聪明”期权交易者私有信息获取能力强于股票交易者，对信息的判断也更加精确，因此其存在可以提升股票市场的有效性，降低股票市场波动性（市场风险），也就是说期权市场是聪明的话，通过期现交易者的传导作用会让市场对价格和波动率定价更加准确，即更加接近理论市场。

从上文理论模型可以推出，期现交易者比例越高，信息精度越高，价格更接近理论价格。因此，我们可以设计更加接近现实市场演化特征的计算实验仿真来检验以上结论是否成立，从而验证本文结论和建立的理论模型。相比理论模型推导的理性均衡价格，计算实验仿真能够刻画市场动态演化特征，更精确模拟接近真实市场的交易规则、交易决策和交易者动态博弈特征，从而让结论更具说服力。

1. 仿真市场设计

为进一步深入探究该问题，本文设计一个仿真股票市场，通过引入具有更强私有信息和对信息分析更为精确的期现交易者，其同时在股票市场和期权市场进行交易，通过仿真检验是否会影响股票市场的定价效率和市场波动。仿真研究方法的优势在于，可以排除市场其他干扰因素，从而开展单因素分析，作为实证研究的进一步补充。

计算实验仿真是研究交易者行为和市场机制的新工具，近年来受到跨学科研究者和全球金融监管部门的关注。Farmer 和 Foley（2009）发表在《Nature》的论文指出，经济系统需要科学的计算实验建模。Battiston 等（2016）发表在《Science》的论文也指出，经济政策需要跨学科的网络分析和行为建模。

基于计算实验方法的研究框架，本文首先构建了一个仿真市场，模型中包括 5000 个交易日，仿真周期为 12000 期，相当于仿真 250 个交易日（一年）。当交易者进入市场后，对股票价格进行预测，并根据股票实时报价和自身头寸持仓等信息进行投资决策。

仿真市场基本设置（详细说明见附录 2）如下：一是交易资产为股票资产，初始价格为 10 元。二是微观结构设计方面，本文采用与 A 股市场一致的订单簿模型和交易规则。三是下单决策方面，交易者根据其对股票的预测价格和股价的实际关系进行交易决策。如果交易者的预测价格大于卖出最优报价，则下市价订单买入；如果交易者的预测价格小于买入最优报价，则下市价订单卖出。

仿真市场交易者行为方面，按照 Chiarella、Iori 和 Perello（2009）、崔毅安（2019）的模型设计，结合 A 股市场的实际情况，仿真模型中的交易者采用基本面和技术面的混合异质信念对股票收益率进行预测，预测收益率如公式（18）所示。

$$r_{t+\tau}^{i,j} = \left(x^i \left(\ln \frac{(1+z^i N\sigma) f_t^j}{p_t^j} \right) / \tau_i + y^i \ln \frac{\bar{p}_{\tau_i/4}^j}{\bar{p}_{\tau_i}^j} \right) \quad (18)$$

其中， $\ln(1+z^i N\sigma) f_t^j / p_t^j$ 表示交易者的基本面信念， N 表示股票交易者和期现交易者对私有信息获取能力和对信息分析精确度不同，该值越大说明私有信息越少、信息判断精确度越低， σ 为基本价值的分钟波动率（设为 0.09）。 $N\sigma$ 与理论模型（13）式中的信息精度指标 ν 具有相近含义，信息越精准， $N\sigma$ 越小， ν 值越大，从这个意义上来说，仿真实验可以验证理论模型有关信息对定价影响的结论。 τ_i 表示投资期限， f_t^j 为股票 j 的基本价值， p_t^j 为股票 j 的即时价格， $\ln(\bar{p}_{\tau_i/4}^j / \bar{p}_{\tau_i}^j)$ 表示技术面信念， $\bar{p}_{\tau_i/4}^j$ 为短期均价， $\bar{p}_{\tau_i}^j$ 为长期均价。 x^i 、 y^i 、 z^i 分

别代表技术面、基本面、噪音面三种异质信念的权重。

2. 实验方案设计

从理论模型定理和推论可以看到，信息精度越高，有信息优势的期现交易者比例越高，价格更接近理论价格，因而股票市场定价效率和风险定价能力会更高。

假设期现交易者获取了真实客观的信息，那么意味着：（1）信息精度越高，则价格越接近合理价格，定价效率越高。（2）期现交易者的占比 ϕ 越高，则定价效率越高。

结论（1）可以从推论公式推出：

$$E(P^*) = \frac{1}{r^2} \left[rD + \bar{\theta} + \phi \cdot \frac{1}{v/v_u + 1} \cdot (\bar{y}_u - \bar{\theta}) - \alpha\sigma^2\xi \right]$$

当信息精度 $v_u \rightarrow \infty$ ，期现交易者占比 $\phi \rightarrow 1$ ，那么均衡价格将完全反映理论上的真实价格：

$$P^R = \frac{1}{r^2} [rD + \bar{y}_u - \alpha\sigma^2\xi]$$

那么，信息精度越高，权重 $\left(\frac{1}{v/v_u+1}\right)$ 越大（越接近1），表明其越反映真实信息的合理价格。同理，结论（2）也成立，即期现交易者的占比 ϕ 越高，权重 $\left(\frac{1}{v/v_u+1}\right)$ 越大（越接近1），表明其更接近合理价格。

根据以上分析，实验方案设计中，赋予纯股票市场交易者较大的N值(N=8)，表示其具有较少私有信息、对信息判断准确度较低，同时给予期现交易者赋予较小的N值，表示其比股票市场交易者具有更多私有信息、对信息判断也更为准确，N值越小表示期现交易者聪明程度越高，存在信息获取和分析上的优势。进而设计两组仿真实验，一是引入不同聪明程度(N=6、4、2)的期现交易者，代表聪明交易者的信息优势程度不同；二是引入不同数量（总交易者数量的1%、5%、10%）的期现交易者，代表聪明交易者占比。本文通过对比引入期现交易者后股票市场定价效率和波动性变化，分析期现交易者对股票市场效率和风险的影响。

3. 仿真结果

拥有信息优势的“聪明”期现交易者的引入确实提升了股票市场的定价效率，并且期现交易者聪明程度（信息优势程度）越高（N值越小）、交易者占比越大，股票市场的定价效率 MAE^7 （价格与基本价值偏离程度）越高，市场波动越接近

⁷ 选取股票价格 p_t 与股票基本价值 f_t 的偏离程度作为衡量市场定价效率的指标MAE，MAE指标越小，则

说明市场定价效率越高。其中， $MAE = \frac{1}{T} \sum_{t=1}^T |p_t - f_t|$ 。

股票基本价值波动。例如，引入 1%期现交易者（N=2），股票市场定价效率由 2.13 元优化至 0.89 元，波动性由 3.67 降低至 1.71。期现交易者占比增加，股票市场的定价效率越好，市场波动更加接近股票基本价值波动。例如，期现交易者比例增加到 10%，股票市场定价效率由 2.13 元优化至 0.38 元，波动性由 3.67 降低至 0.60（见表 11）。期现交易者占比进一步增加至 100%，仿真价格已非常接近理论价格（定价效率为 0.11 元，波动率为 0.12）。这都表明，一是引入期现交易者后，股票市场的价格与基本价值偏离幅度减少，市场波动进一步接近基础资产的真实波动，表明期现交易者的存在可以提升市场价格发现和风险定价能力，二是期现交易者的聪明程度越高、数量占比越大，股票市场的定价效率和风险定价能力提升越多。

表 11 引入期权交易者的计算实验仿真结果

结果 1：引入不同聪明程度的期权交易者后股票市场效率和波动变化				
评价指标	仅有现货交易者 (N=8)	引入 1%期权交 易者 (N=6)	引入 1%期权交 易者 (N=4)	引入 1%期权交 易者 (N=2)
定价效率 (MAE)	2.13	1.94	1.81	0.89
波动率 (方差)	3.67	3.56	2.32	1.71
结果 2：引入不同数量的期权交易者后股票市场效率和波动变化				
评价指标	引入 1%期权交 易者 (N=4)	引入 5%期权交 易者 (N=4)	引入 10%期权交 易者 (N=4)	引入 100%期权 交易者 (N=4)
定价效率 (MAE)	1.81	0.66	0.38	0.11
波动率 (方差)	2.32	1.36	0.60	0.12

七、结论与启示

期权是全球重要的金融工具，具有价格发现、风险管理等功能。期权交易者是否是聪明交易者，他们能否领先股票市场预测感知市场价格变化和未来市场波动，是理论研究和金融市场建设关注的问题。本文基于境内市场期权数据，通过提取期权隐含信息，实证检验了期权交易者是否“聪明”等问题，并在理性预期均衡框架下从信息角度讨论了期权市场有预测感知能力的机制。同时，本研究还进一步利用计算实验方法，检验了“聪明”期现交易者对现货市场效应的影响。

本文研究发现，期权交易者是“聪明”交易者，表现在期权隐含信息蕴含了市场价格变化的信息，且对市场风险有较强的预测能力。第一，期权交易者对股票市场未来收益率有一定预测能力。第二，期权交易者会对定期发布的重要宏观数据提前定价，且在“积极”和“负面”数据发布时呈现明显不对称特征。第三，实证证据更倾向于支持信息发现论，即期权隐含信息对指数收益率的预测能力，主要体现在期权交易者对新信息的获取、分析以及定价上具有优势。第四，实证结果强烈支持期权市场核心功能是风险预测和定价能力，期权隐含信息不仅较为精确的预测未来收益率波动，并且对市场未来的“黑天鹅”风险有较好的预测能

力。第五，理性预期均衡模型解释了期权隐含信息对股票市场有预测感知能力的机制，期现交易者的信息越准确，对未来价格与风险的感知越准确。第六，计算实验结果表明，“聪明”的期现交易者的引入，提高了股票市场价格发现和风险定价能力。

本研究发现期权隐含信息蕴含了未来市场价格变化的信息，且对已实现波动和尾部风险具有较强的预测能力，表明股票市场自身发挥价格形成作用，期权市场主要发挥风险管理与波动率定价功能，期现货市场共同构成了价格及时发现、风险有效管理的市场体系。本文在研究期权交易者是否“聪明”及其具体体现等问题之外，也为下一步从市场微观结构角度研究期权交易者“聪明”的来源奠定了较好的认识论基础。当前，我国期权市场正处于蓬勃发展阶段，建设一个功能充分发挥的期权市场，对于提升市场效率，健全资本市场功能，具有十分重要的意义。为此，本研究认为，有必要进一步增加期权品种，引入更多类型和数量的交易者参与期、现市场交易，为股票市场波动风险提供更好的定价基准，也有利于我国资本市场的长期稳定发展。

参考文献

- [1] An B J, Ang A, Bali T G, et al. The joint cross section of stocks and options[J]. *The Journal of Finance*, 2014, 69(5): 2279-2337.
- [2] Andersen T G, Bollerslev T, Diebold F X, et al. Modeling and forecasting realized volatility[J]. *Econometrica*, 2003, 71(2): 579-625.
- [3] Atilgan Y. Volatility spreads and earnings announcement returns[J]. *Journal of Banking & Finance*, 2014, 38: 205-215.
- [4] Bakshi G, Kapadia N, Madan D. Stock return characteristics, skew laws, and the differential pricing of individual equity options[J]. *The Review of Financial Studies*, 2003, 16(1): 101-143.
- [5] Banerjee P S, Doran J S, Peterson D R. Implied volatility and future portfolio returns[J]. *Journal of Banking & Finance*, 2007, 31(10): 3183-3199.
- [6] Bates D S. The crash of '87: was it expected? The evidence from options markets[J]. *The journal of finance*, 1991, 46(3): 1009-1044.
- [7] Battiston S, Farmer J D, Flache A, et al. Complexity theory and financial regulation[J]. *Science*, 2016, 351(6275): 818-819.
- [8] Biais B, Hillion P. Insider and liquidity trading in stock and options markets[J]. *The Review of Financial Studies*, 1994, 7(4): 743-780.
- [9] Chakravarty S, Gulen H, Mayhew S. Informed trading in stock and option markets[J]. *The Journal of Finance*, 2004, 59(3): 1235-1257.
- [10] Chiarella C, Iori G, Perelló J. The impact of heterogeneous trading rules on the limit order book and order flows[J]. *Journal of Economic Dynamics and Control*, 2009, 33(3): 525-537.
- [11] Corsi F. A simple approximate long-memory model of realized volatility[J]. *Journal of Financial Econometrics*, 2009, 7(2): 174-196.
- [12] Coudert V, Gex M. Does risk aversion drive financial crises? Testing the predictive power of empirical indicators[J]. *Journal of Empirical Finance*, 2008, 15(2): 167-184.
- [13] Easley D, O'hara M, Srinivas P S. Option volume and stock prices: Evidence on where informed traders trade[J]. *The Journal of Finance*, 1998, 53(2): 431-465.
- [14] Farmer J D, Foley D. The economy needs agent-based modelling[J]. *Nature*, 2009, 460(7256): 685-686.
- [15] Finucane T J. Put-call parity and expected returns[J]. *Journal of Financial and Quantitative Analysis*, 1991, 26(4): 445-457.
- [16] Fleming J, Ostdiek B, Whaley R E. Trading costs and the relative rates of price discovery in stock, futures, and option markets[J]. *The Journal of Futures Markets (1986-1998)*, 1996, 16(4): 353.
- [17] Giot P. Relationships between implied volatility indices and stock index returns[J]. *Journal of Portfolio Management*, 2005, 31(3): 92-100.
- [18] Guo B, Han Q, Ryu D. Is the KOSPI 200 options market efficient? Parametric and nonparametric tests of the martingale restriction[J]. *Journal of Futures Markets*, 2013, 33(7): 629-652.
- [19] Grossman S J, Stiglitz J E. On the impossibility of informationally efficient markets[J]. *The American economic review*, 1980, 70(3): 393-408.

- [20] Jiang G J, Tian Y S. The model-free implied volatility and its information content[J]. *The Review of Financial Studies*, 2005, 18(4): 1305-1342.
- [21] Jin W, Livnat J, Zhang Y. Option prices leading equity prices: Do option traders have an information advantage?[J]. *Journal of Accounting Research*, 2012, 50(2): 401-432.
- [22] Kambouroudis D S, McMillan D G, Tsakou K. Forecasting realized volatility: The role of implied volatility, leverage effect, overnight returns, and volatility of realized volatility[J]. *Journal of Futures Markets*, 2021, 41(10): 1618-1639.
- [23] Lee J, Cheong H Y. Trade size and information-motivated trading in the options and stock markets[J]. *Journal of Financial and Quantitative Analysis*, 2001, 36(4): 485-501.
- [24] Li W X, Chen C C S, French J J. Toward an early warning system of financial crises: What can index futures and options tell us?[J]. *The Quarterly Review of Economics and Finance*, 2015, 55: 87-99.
- [25] Muravyev D, Pearson N D, Broussard J P. Is there price discovery in equity options?[J]. *Journal of Financial Economics*, 2013, 107(2): 259-283.
- [26] Pan J, Poteshman A M. The information in option volume for future stock prices[J]. *The Review of Financial Studies*, 2006, 19(3): 871-908.
- [27] Skinner D J. Options markets and the information content of accounting earnings releases[J]. *Journal of Accounting and Economics*, 1990, 13(3): 191-211.
- [28] Wang H. VIX and volatility forecasting: A new insight[J]. *Physica A: Statistical Mechanics and its Applications*, 2019, 533: 121951.
- [29] Wang J. A model of intertemporal asset prices under asymmetric information[J]. *The Review of Economic Studies*, 1993, 60(2): 249-282.
- [30] Veronesi P. Stock market overreactions to bad news in good times: a rational expectations equilibrium model[J]. *The Review of Financial Studies*, 1999, 12(5): 975-1007.
- [31] Xing Y, Zhang X, Zhao R. What does the individual option volatility smirk tell us about future equity returns?[J]. *Journal of Financial and Quantitative Analysis*, 2010, 45(3): 641-662.
- [32] 陈蓉, 林秀雀. 波动率偏斜与风险中性偏度能预测尾部风险吗[J]. *管理科学学报*, 2016, 19(8): 113-126.
- [33] 崔毅安. 多资产订单簿的计算实验建模与股票期现货市场动态研究[D]. 天津大学, 2019.
- [34] 林辉, 马潇涵, 李铭. 金融衍生品具有股市风险预警功能吗? ——基于机器学习模型的实证检验[J]. *证券市场导报*, 2022(12): 47-56.
- [35] 王琳玉, 倪中新, 郭婧. 上证 50ETF 隐含高阶矩风险对股票收益的预测研究[J]. *统计研究*, 2020, 37(12): 75-90.
- [36] 朱超, 李子若, 李纪鹏. 期权隐含偏度期限结构的股票定价信息含量: 基于中国、美国和日本证据[J]. *国际金融研究*, 2021(04): 77-86.

附表 1 变量描述

变量	变量的涵义
<i>Big_Movement</i>	大涨大跌指标，当出现大涨或者大跌时，赋值 1，否则赋值 0。
<i>VIX</i>	波动率指数，用无模型方法计算，反映期权市场平均的隐含波动率水平。
<i>ΔVIX</i>	波动率指数的变化量，基于 VIX 差分得到，反映期权交易者情绪变化。
<i>RV</i>	已实现波动，基于高频数据（1 分钟）计算得到，反映 1 个交易日指数实际发生的波动水平。
<i>PCR</i>	认沽认购比，由认沽合约持仓除以认购合约持仓计算得到，反映期权交易者多空观点。
<i>VRP</i>	波动率风险溢价，由波动率指数与已实现波动之差计算得到，一定程度反映期权交易者的恐慌情绪。
<i>SKEW</i>	偏度指数，用无模型方法计算，反映尾部风险程度。
<i>CPIV</i>	波动率价差指标，认购期权隐含波动率与认沽期权隐含波动率之差计算得到，反映认购期权与认沽期权供求关系。
<i>BR</i>	隐含基差（升贴水），期权合成期货价格与标的 ETF 价格之差计算得到，反映期权交易者看涨和看跌观点。
<i>PE</i>	指数市盈率，反映指数成分股整体估值水平。
<i>TR</i>	指数换手率，反映指数成分股整体活跃度。
<i>Return</i>	指数收益率，反映 1 个交易日指数涨跌幅度。

附表 2 期权市场指标描述性统计

指标	<i>Big_Movement</i>	<i>VIX</i>	<i>PCR</i>	<i>VRP</i>	<i>SKEW</i>	ΔVIX	<i>Return</i>	<i>RV</i>	<i>BR</i>	<i>CPIV</i>	<i>MD</i>	<i>CMD</i>
Mean	0.3376	20.6654	0.9667	5.8258	102.6436	0.0014	0.0059	14.8396	21.149	-0.0336	0.8738	0.3710
Std	0.4731	4.2601	0.1644	6.6952	3.2282	1.1214	1.2426	8.0072	18.2136	0.0517	0.0766	0.4833
Min	0	13.79	0.5495	-117.933	62.86	-7.23	-7.88	6.357	9.8413	-0.4220	0.6902	0.0000
0.3	0	17.7975	0.8568	4.3665	100.6325	-0.47	-0.67	10.4539	12.0426	-0.0627	0.8346	0.0000
0.5	0	19.91	0.9393	6.6981	102.95	-0.04	0.01	12.9756	15.4936	-0.0275	0.8680	0.0000
0.8	1	22.41	1.0579	8.7656	104.6	0.4225	0.74	16.8174	17.6766	0.0008	0.9606	1.0000
Max	1	39.91	1.7517	17.6243	112.1	5.57	5.67	145.113	123.4221	0.1006	1.0000	1.0000

附表 3 期权市场指标相关性分析

指标	<i>Big_Movement</i>	<i>VIX</i>	<i>PCR</i>	<i>VRP</i>	<i>SKEW</i>	ΔVIX	<i>Return</i>	<i>RV</i>	<i>BR</i>	<i>CPIV</i>
<i>Big_Movement</i>	1.00	0.20	-0.01	-0.22	-0.08	0.02	0.00	0.29	-0.02	-0.05
<i>VIX</i>	0.20	1.00	-0.05	-0.02	-0.29	0.13	-0.06	0.55	-0.28	-0.39
<i>PCR</i>	-0.01	-0.05	1.00	0.22	0.28	-0.02	0.31	-0.21	0.00	-0.08
<i>VRP</i>	-0.22	-0.02	0.22	1.00	0.27	-0.15	0.19	-0.85	0.13	0.12
<i>SKEW</i>	-0.08	-0.29	0.28	0.27	1.00	-0.10	0.25	-0.37	0.06	0.14
ΔVIX	0.02	0.13	-0.02	-0.15	-0.10	1.00	-0.02	0.20	0.00	-0.01
<i>Return</i>	0.00	-0.06	0.31	0.19	0.25	-0.02	1.00	-0.19	0.10	0.00
<i>RV</i>	0.29	0.55	-0.21	-0.85	-0.37	0.20	-0.19	1.00	-0.26	-0.31
<i>BR</i>	-0.02	-0.28	0.00	0.13	0.06	0.00	0.10	-0.26	1.00	0.64
<i>CPIV</i>	-0.05	-0.39	-0.08	0.12	0.14	-0.01	0.00	-0.31	0.64	1.00

附表 4 宏观经济数据发布情况

指标	<i>GDP</i>	<i>PMI</i>	<i>CPI</i>	<i>PPI</i>	工业增加值	社会消费零售总额	固定资产投资
发布频率	季度	月度	月度	月度	月度	月度	月度
发布次数	15 次	44 次	44 次	44 次	44 次	39 次	39 次
“积极”次数	7 次	22 次	13 次	24 次	26 次	15 次	15 次
“负面”次数	8 次	22 次	31 次	20 次	18 次	24 次	24 次

附表 5 稳健性检验（上证 50ETF 期权）

变量	收益率 模型	变量	波动率 模型	变量	尾部风险模 型（上涨）	尾部风险模 型（下跌）
<i>Return</i>	-0.0142 (-3840)	<i>RV_D</i>	0.0996*** (2.3582)	<i>VIX</i>	0.0041* (1.9287)	0.0042* (1.9222)
<i>VIX</i>	0.0033 (0.2566)	<i>RV_W</i>	0.2840*** (3.7823)	ΔVIX	0.0000 (-0.0147)	0.0063 (0.9474)
<i>VRP</i>	-0.4137 (-0.6349)	<i>RV_M</i>	0.0000 (0.0075)	<i>PCR</i>	-0.0456 (0.9658)	-0.0847* (-1.7620)
<i>SKEW</i>	-0.0147 (-0.7812)	<i>VIX</i>	0.3811*** (4.3708)	<i>VRP</i>	0.1399 (1.2879)	0.0800 (0.7235)
<i>CPIV</i>	-1.9945 (-1.0513)	ΔVIX	0.6709*** (3.1687)	<i>SKEW</i>	-0.0135*** (-4.3069)	-0.0027 (-0.8536)
<i>PCR</i>	0.1521 (0.5271)	<i>PCR</i>	-3.8980*** (-2.5459)	<i>CPIV</i>	-0.0343 (-0.1075)	0.1964 (0.6043)
<i>BR</i>	0.0194 (0.0651)	<i>SKEW</i>	-0.1956 (-0.3003)	<i>BR</i>	-0.0791 (-1.5802)	-0.0664 (-1.3038)
<i>PE</i>	-0.0330 (-1.0660)	<i>CPIV</i>	-6.1932 (-0.6177)	<i>PE</i>	-0.0102* (-1.9598)	0.0051 (0.9639)
<i>TR</i>	-0.0400 (-0.0668)	<i>BR</i>	1.5426 (0.9626)	<i>TR</i>	0.1262 (1.2440)	0.1709* (1.6554)
		<i>PE</i>	-0.1733 (-1.0531)			
		<i>TR</i>	11.5189*** (3.3493)			
<i>Adj.R²</i>	-0.0010	<i>Adj.R²</i>	0.2514	<i>Adj.R²</i>	0.0483	0.0146

附录 1 期权隐含信息指标说明

假设市场是无套利的，风险中性测度下资产价格过程可表示为一个随机过程：

$$dS_t = r_t S_t dt + \sigma_t S_t dW_t \quad (\text{A-1})$$

S_t 表示 t 时刻标的价格， W_t 是中性测度下的布朗运动， r_t 表示无风险利率， σ_t 表示标的价格波动。

1. 已实现波动

低频时间序列领域有 GARCH 模型体系和 SV 模型体系两类代表性波动率度量方法，但是都需要进行复杂的参数估计。随着高频金融时间序列容易获得，Andersen 等（2003）利用高频时间序列更准确、更方便地估计波动率，并提出了“已实现波动”（RV, Realized Volatility）的波动率度量方法。该方法的具体计算过程如下：

将一个交易日的的时间间隔 n 等分，定义离散时间区间 $[\frac{\tau-1}{n}, \frac{\tau}{n}]$ 的高频收益率为 $\Delta_{\tau}^n p = p_{\tau/n} - p_{\tau-1/n}$ ，则每个交易日的已实现变差 RV 定义如下，当 $n \rightarrow \infty$ 时，RV 依概率收敛于积分波动。

$$RV_t = \sum_{\tau=m+1}^{m+n} |\Delta_{\tau}^n p|^2 \xrightarrow{P} \int_t^{t+1} \sigma_t^2 ds \quad (\text{A-2})$$

$$RV_{t,T} = \frac{1}{T-t} \int_t^T \sigma_t^2 ds \quad (\text{A-3})$$

2. 波动率指数

波动率指数作为预测市场的重要指标，是实践和学术领域广泛应用的指标，由期权价格计算得到。如果期权市场是有效的，则期权价格所反映的隐含波动率能够很好预测风险。常见计算方法有基于 BS 模型和基于无模型的波动率指数，本文采用后者进行估计，学术界普遍认为该方法不受模型假设约束，是更好的编制方法。具体计算公式如下：

$$VIX_{t,T}^2 = \frac{2}{T-t} \sum_i \frac{\Delta K_i}{K_i^2} e^{T-t} O(T, K_i) - \frac{1}{T-t} \left[\frac{F}{K_0} - 1 \right]^2 \quad (\text{A-4})$$

以上公式是基于芝加哥期权交易所公布的方法，并利用期权实际交易数据计算得到。 t 代表当前时刻， T 代表到期日， K 代表执行价格， O 代表期权价格， F 代表隐含的期货价格， K_0 代表平值期权的执行价格。

3. 波动率风险溢价

基于标的资产历史价格的已实现波动率是标的真实波动率的测度，基于期权价格的隐含波动率反映交易者预期波动率水平。理论上预期波动率是真实波动率

的无偏估计。但现实中预期波动率与真实波动率经常互相偏离，导致存在波动率风险溢价，理论上主要有供求关系决定论、风险对冲需求论、风险溢价论三种解释⁸。波动率风险溢价由同一期权标的隐含波动率与已实现波动率之差计算得到：

$$VRP_t = VIX_{t,T} - RV_{t,T} \quad (A-5)$$

4. 偏度指数

偏度指数可以反映大幅上涨或下跌的方向性预期，可以度量市场对极端风险的预期，是对波动率指数的重要补充。Bakshi 等（2003）提出了利用期权价格信息编制偏度指数的方法。该方法从最为基础的偏度公式出发，并对该公式逐步分解，利用期权数据估计各分解项。

$$SKEW = E\left[\frac{(R - \mu)^3}{\sigma^3}\right] = \frac{E[R^3] - 3E[R]E[R^2] + 2E[R]^3}{(E[R^2] - E^2[R])^{3/2}} \quad (A-6)$$

与波动率指数一样，偏度指数也可以由认购期权和认沽期权计算得到。 R 代表收益率， $\mu = E[R]$ ， σ 代表波动率。

5. 认沽认购比

认沽认购比是认沽持仓与认购持仓之比，反映交易者对后市看涨看跌的态度，认沽认购比越高，期权交易者看空后市观点和避险意愿越强烈，反之亦然。原因在于如果交易者预期股票价格上涨，或者希望对冲空头持仓，则买入认购期权，反之则买入认沽期权。具体指标如下：

$$PCR_t = OI_Put_t / OI_Call_t \quad (A-7)$$

其中， OI_Put_t 代表认沽期权持仓量， OI_Call_t 代表认购期权持仓量。

6. 波动率价差

波动率价差（CPIV）是认购期权与认沽期权的隐含波动率的差值，可以衡量期权市场供需的失衡程度。当期权交易者普遍预期后市大涨，同一行权价期权，认购期权可能求大于供，推动隐含波动率上升，认沽期权则可能供大于求，推动隐含波动率下降。因此，波动率价差可以预测未来股市变化。本文选取近月的认沽期权隐含波动率与认购期权隐含波动率之差构建波动率价差指数。

$$Volatility_Spread_t = Vol_{Call,t} - Vol_{Put,t} \quad (A-8)$$

其中， $Vol_{Put,t}$ 代表认沽期权隐含波动率， $Vol_{Call,t}$ 代表认购期权隐含波动率。

7. 期权隐含升贴水

⁸ 三种解释分别为：一是供求关系决定论。金融市场的主要参与者通常是现货多头，他们往往愿意付出额外代价买入期权来对冲现货风险。现货市场的天然多头属性决定了期权供小于求的格局，故而推高了隐含波动率水平。二是风险对冲需求论。交易者更愿意持有波动率多头来对冲大盘的下行风险。利用波动率多头对冲股票极端下行风险的需求带动隐含波动率上涨，导致股票指数的波动率与其对应的股票指数呈负相关，且在市场恐慌性下跌阶段负相关关系更加显著。三是风险溢价论。大量研究发现，股票市场的跳跃现象普遍存在，期权交易者给予股票市场普遍存在的跳跃风险更高风险溢价，导致隐含波动率高于已实现波动率。

期权合成期货升贴水是期权合成期货头寸与指数之间的价差，可以反映交易者看涨和看跌情绪的强烈程度。理论上，利用行权价和到期日相同的一对认购和认沽期权可以合成相应股指期货。该合成股指期货与真实标的资产之间的价差相当于股指期货的基差，可以反映期权交易者看涨和看跌情绪的强烈程度。当市场看涨情绪强烈，期权合成期货升水，反之则贴水。Finucane（1991）实证检验了 S&P100 指数期权合成期货升贴水领先指数价格变动 15 分钟。Muravyev（2013）基于美国个股期权的研究发现，期权隐含价格对未来股票价格没有明显的预测作用。

附录 2: 计算实验设计

一、订单簿市场中的交易规则设计

本文建模了与中国股票市场接近的纯限价订单簿市场模型，目的是让计算实验模型更加符合真实市场的特征：

- (1) 按照时间优先、价格优先原则进行撮合；
- (2) 股票采取 T+1 交易；
- (3) 将真实市场的 5 分钟设定为一个仿真周期。

二、交易者下单行为设计

1. 交易者初始化设计

从运算能力考虑，将交易者总人数设置为 5000 人，并给予首次进入市场的交易者按照均匀分布分配一定数量的股票，初始现金等于初始股票市值。当交易者总财富过低时，被视为破产，引入新的初始化交易者。

2. 交易者价格预测行为建模

模型 (18) 考虑了混合异质信念，其中， x^i 、 y^i 、 z^i 分别代表技术面、基本面、噪音面的信念权重，满足 $x^* + y^* + z^* = 1$ 约束条件。

交易者的交易期限 τ^i 根据技术面信念权重 x^i 和基本面信念权重 y^i 来确定，与技术面信念权重成反比，与基本面信念权重成正比。初始交易期限 $\tau^* = 5day = 240t$ 。

$$\tau^i = \tau^* \frac{1 + y^i}{1 + x^i} \quad (\text{A-9})$$

当技术面信念大于基本面信念时，交易者倾向于采用交易更为频繁的技术分析，平均交易期限小于 1 个交易日。当基本面信念大于技术面信念时，交易者倾向于采用交易频率较低的基本面分析，平均交易期限大于 1 个交易日。基于混合异质信念，交易者 i 对于股票 j 的预测价格 $p_{f,t}^{i,j}$ 如公式 (A-10) 所示：

$$p_{f,t}^{i,j} = p_t^j e^{r_{t+\tau^i}^i} \quad (\text{A-10})$$

3. 交易者下单规则

交易者的交易决策根据预测价格和订单簿订单情况决定。当预测价格大于卖出最优报价，交易者以市价订单买入股票；当预测价格小于买入最优报价，交易者以市价订单卖出股票。

致 谢

1.研究背景

有一次看到金融数学介绍文章时，我被“Option”（期权）吸引了。从爱因斯坦 1905 年用布朗运动来描述花粉扩散现象，到大名鼎鼎的随机游走模型，以及 1997 年诺奖的基于布朗运动模型的 Black-Scholes 期权定价公式，数学、物理和金融三位一体的完美结合，深深吸引了我。因此此次研究我将期权定为研究目标方向，开始了准备工作。之前我在参加厦门大学经济学夏令营认识的韩乾教授，研究主要领域就是金融工程，联系之后教授耐心听取我的想法，鼓励、指导我确定论文的研究目标和主题。选题也非常契合我的数学和计算机编程的特长，我对研究充满了好奇和动力。

韩教授专业的指导在我完成这篇金融论文中起到了至关重要的作用，并且指导都是无偿的。在整个论文写作过程中，对我的研究思路进行了全程的引导和修改，指导我完善研究框架，不断提炼逻辑和观点。韩教授对数据和编程方面给予了关键的指导，鼓励我选取与我擅长的 C++ 高度兼容的 Matlab 作为主要计量分析软件，还无私提供了一些研究中需要用到的分析数据。很顺利我确定了合适方法和数据样本，编写了论文所需的程序，完成实证分析。

2.研究过程

研究从文献整理、指标设计、数据处理、模型检验 4 个方面有序推进，在韩教授指导下，耗时 6 个月最终完成了论文写作。一是文献整理。该领域的代表性论文，我认真研读，认真整理关键核心观点。二是指标设计和数据分析。期权定价涉及复杂的数学知识，在指导下我掌握了主要指标含义，自己动手计算一些计算难度小的指标，也利用了一些韩教授以往研究积累下来的数据。三是模型检验和论文写作。通过 Matlab 开展实证，借鉴成熟方法建立模型，并用韩教授研究团队已建立计算实验平台模拟运行了本文设计的仿真实验。

3.心得体会

论文写作对我是一个不小的挑战和考验，论文写作是一次艰难的学术之旅，也收获了收获的喜悦。我相信这篇论文只是我踏上学术研究的第一步，它让我收获了宝贵的学术经验。当然，限于个人水平，论文难免有欠缺和不足之处，还需要今后进一步改进。非常感谢韩教授的悉心指导，才有了我这次难得的学习机会和意义非凡的学术体验。特别感谢父母，是他们鼓励我自由探索，鼓舞我不断寻找心中的高峰。