

# 市场状态与股价可预测性：基于深度学习模型的实证

潘含笑

北方工业大学经济管理学院，北京

收稿日期：2026年3月8日；录用日期：2026年4月1日；发布日期：2026年4月10日

## 摘要

金融市场研究表明，股票价格的可预测性具有显著的时变性与状态依赖特征，市场运行机制会随着趋势转折、波动结构变化及外生冲击而发生阶段性调整；然而，现有基于深度学习的价格预测研究更多关注模型整体预测精度，对预测表现在不同市场阶段下的系统性差异缺乏专门分析，从而难以准确刻画预测可靠性的环境约束。基于此，文章以沪深300指数为研究对象，构建融合小波变换的LSTM-CNN混合预测模型，并在样本外框架下比较模型在不同市场阶段中的预测表现。实证结果显示，股票价格预测表现具有显著的阶段依赖性：在趋势相对明确、价格结构较为稳定的阶段预测误差显著较低，而在趋势转折或震荡特征突出的阶段预测可靠性明显下降，且上述差异在不同参数设定下具有稳健性。研究表明，股票价格预测模型的有效性高度依赖市场环境，对预测性能的评估有必要引入市场阶段视角，以更准确界定模型的适用边界与现实价值。

## 关键词

股票价格预测，市场阶段，预测可靠性，深度学习，沪深300指数

# Market Conditions and Predictability of Stock Prices: An Empirical Analysis Based on Deep Learning Models

Hanxiao Pan

School of Economics and Management, North China University of Technology, Beijing

Received: March 8, 2026; accepted: April 1, 2026; published: April 10, 2026

## Abstract

Studies in financial markets suggest that the predictability of stock prices exhibits significant time variation and state dependence. The operating mechanism of financial markets will undergo stage-wise adjustments with bull-bear transitions, changes in volatility structures, and exogenous shocks. However, existing literature on stock price forecasting using deep learning focuses primarily on the overall forecasting accuracy of models, while lacking dedicated analyses of the systematic differences in forecasting performance across distinct market regimes. Consequently, it is difficult to accurately characterize the environmental constraints that affect forecasting reliability. Against this background, this paper takes the CSI 300 Index as the research object and constructs a hybrid LSTM-CNN prediction model integrated with wavelet transform. We then compare the out-of-sample forecasting performance of the model across different market stages. The empirical results show that stock price forecasting performance presents significant regime dependence: the prediction error is significantly lower in periods with relatively clear trends and stable price structures, whereas the forecasting reliability decreases substantially during trend reversals or high-volatility periods. These findings remain robust under different parameter settings. This study indicates that the effectiveness of stock price forecasting models is highly dependent on market conditions. It is necessary to incorporate the perspective of market stages into the evaluation of forecasting performance, so as to better identify the applicable boundaries and practical value of forecasting models.

## Keywords

Stock Price Prediction, Market Regime, Forecasting Reliability, Deep Learning, CSI 300 Index

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

股票市场价格是否具有可预测性一直是金融研究中的核心问题之一。有效市场假说认为，价格已经充分反映所有可获得信息，历史价格不应对未来价格变动具有预测能力；但大量经验研究发现，市场有效性及价格形成机制并非在时间维度上保持稳定，而是随着宏观环境、制度安排与投资者行为变化而不断调整。行为金融学与适应性市场假说的相关研究进一步指出，过度反应、反应不足以及动量效应等现象在特定时期内反复出现，表明股票价格的可预测性具有显著的时变性与状态依赖特征，这为在不同市场环境下讨论价格预测问题提供了理论基础。

在中国股票市场背景下，既有文献从市态划分、牛熊市区分以及上涨与下跌阶段对比等多个角度，系统揭示了价格行为在不同市场阶段下的显著差异：价格波动结构、对信息冲击的反应方式以及统计关系的稳定性均会随着市场状态变化而发生调整。相关研究还指出，若采用静态全样本框架对价格序列进行统一建模，阶段性特征容易被弱化，从而导致对预测关系稳定性的误判。与此同时，随着机器学习与深度学习方法在价格预测中的广泛应用，研究重心逐步从“模型是否优于传统方法”转向“模型预测表现是否具有环境依赖性”，但已有证据显示，同一预测模型在不同市场阶段的样本外表现往往存在显著差异。

尽管上述研究为理解市场状态与价格行为提供了重要启示，但在价格水平预测层面仍存在不足。一方面，部分研究主要以收益序列为研究对象，其结论未必能够直接外推至具有误差累积特征的价格预测

问题；另一方面，在深度学习预测框架下，现有研究更多关注模型结构改进与整体预测精度提升，对预测误差在不同市场阶段下的系统性比较相对有限，从而难以明确判断模型有效性是否具有稳定的阶段排序特征。这使得“价格预测在何种市场阶段更可靠”这一问题仍缺乏直接的经验证据。

从理论与实践角度看，回应上述问题具有重要意义。在适应性市场假说视角下，预测关系本身应被视为随市场环境变化而不断调整的结果，对预测模型的评价若忽略市场阶段差异，容易高估其通用性与稳定性；在实际应用中，若能够识别预测误差在不同阶段的系统性变化，有助于更审慎地使用预测模型并界定其适用边界。因此，在统一模型与预测框架下，对价格预测表现进行分阶段评估，有助于更准确地理解深度学习模型在复杂市场环境中的真实有效性。

基于此，本文以沪深 300 指数为研究对象，构建融合小波变换的 LSTM-CNN 混合股票价格预测模型，并在严格的样本外框架下开展实证分析。具体而言，本文利用小波变换对原始价格特征进行多尺度分解，以缓解非平稳性影响，并通过 LSTM 与 CNN 双分支结构分别刻画长期依赖与局部波动特征；在此基础上，将测试样本划分为牛市尾声、熊市深跌、弱复苏与震荡整理四个市场阶段，系统比较模型在不同阶段下的预测误差表现，从而检验股票价格预测是否具有显著的市场阶段依赖性。

## 2. 文献综述

### 2.1. 股票市场可预测性的时变性与理论基础

股票市场价格是否具有可预测性是金融研究中的经典问题。有效市场假说认为，资产价格已经充分反映所有可获得的信息，历史价格变动不应对未来价格产生预测作用[1]。然而，自 20 世纪 80 年代以来，大量金融市场异象的发现使得有效市场假说在解释现实市场行为时面临挑战。

行为金融理论从投资者非理性行为出发，指出认知偏差与情绪波动会导致价格偏离基本价值，并呈现出可预测的动态特征。其中，过度反应与反应不足被视为最具代表性的机制：前者导致价格短期偏离并在后续阶段出现均值回归[2]，后者则使信息被逐步吸收，从而形成价格动量效应[3]。适应性市场假说在此基础上进一步提出，市场效率与可预测性并非静态属性，而是随投资者结构、制度环境及市场条件变化而动态调整，因而具有显著的环境依赖性。

在中国股票市场背景下，相关研究为上述理论提供了经验证据。基于主要股票指数的分析发现，市场有效性指标及风险 - 回报关系随时间显著波动，并在金融危机或重大政策调整阶段发生剧烈变化，反映出价格形成机制具有明显的阶段性特征[4][5]。从结构变化视角出发的研究进一步指出，经济变量与市场表现之间的关系依赖于潜在市场状态，相关参数在不同状态下可能发生显著变化甚至反转，忽略状态切换是预测模型样本外不稳定的重要原因[6][7]。上述研究共同表明，股票市场的可预测性具有显著的时变性与状态依赖特征，为从市场阶段视角重新审视价格预测问题提供了理论与经验基础。

### 2.2. 市场阶段差异与预测表现的环境依赖性

大量经验研究表明，中国股票市场的价格行为在不同市场阶段下呈现出系统性差异。基于市态划分的分析发现，市场状态变化会引致价格相关统计关系的显著调整，不同市态下价格对同类因素的反应方向与作用强度并不一致，全样本统一建模容易掩盖这些结构性差异[8]。从牛熊市视角与上涨 - 下跌阶段的对比研究同样发现，不同市场阶段中价格调整过程与波动结构具有明显非对称性[9]，价格波动在下跌阶段更为集中和剧烈，表明价格运行环境随市场阶段发生系统性变化[10]。

在此基础上，预测模型在不同市场环境下的表现差异逐渐成为研究关注的重点。随着机器学习与深度学习方法在价格预测中的应用，相关研究普遍发现，尽管复杂模型在刻画非线性特征方面具有优势，但其预测效果并不在不同市场阶段保持一致，同一模型在结构相对平稳阶段通常表现较好，而在高波动

或结构频繁变化阶段，预测误差显著扩大，样本外稳定性明显下降[11]-[13]。在突发公共事件或重大外生冲击背景下，上述预测退化现象尤为突出，价格波动放大与结构变化往往导致既有预测模型普遍失效[14][15]。

综合来看，既有研究一致表明，股票价格预测效果对市场环境高度敏感，预测误差的变化更多反映价格动态结构在不同市场阶段的调整，而非预测模型本身的优劣差异。这意味着，若忽略市场阶段差异而采用全样本统一评估，容易得到被阶段性特征“平均化”的结论，从而低估预测不稳定性的现实程度。

### 2.3. 不确定性环境与预测退化的经济机制

现有研究普遍认为，不确定性上升是预测误差阶段性放大的重要经济机制。经济政策不确定性与金融市场波动高度相关，当市场参与者难以形成稳定预期时，其资产配置行为的调整会直接反映在价格波动的显著放大上，从而削弱基于历史关系构建的预测模型稳定性[16]。进一步的研究表明，不确定性通过放大预期分歧与情绪成分，提高价格变动中噪声占比，使可被模型有效识别和外推的系统性价格信号相对减少[17]。

动态分析还发现，不确定性冲击与市场波动之间的联动关系具有显著的时变特征，并往往在重大事件发生时集中释放，导致价格结构在短时间内发生阶段性重塑[18]。因此，在高不确定性或冲击频发的阶段，预测误差的系统性上升更可能反映价格运行结构的变化，而非预测模型设定或参数选择的失效。这一研究脉络为从不确定性视角解释不同市场阶段下预测表现差异提供了重要的经济学依据。

### 2.4. 文献述评

综合前述研究可以看出，股票市场价格行为及其预测关系具有显著的阶段依赖特征，预测模型的样本外表现随市场环境变化而系统性调整。尽管相关文献已从不同角度揭示了市场阶段差异与预测不稳定性的存在，但在统一预测框架下，对价格水平预测误差进行系统的分阶段比较仍相对有限，且不确定性环境在其中所起的作用有待进一步明确。基于此，本文从市场阶段划分的视角出发，在样本外预测框架下考察股票指数价格预测表现的阶段差异，并进一步分析不确定性环境对预测稳定性的影响。

## 3. 研究设计

### 3.1. 样本选择与数据来源

本文以沪深 300 指数作为分析对象，验证基于小波分析的深度学习模型股票价格预测效果。沪深 300 指数精选了中国股市中市值最大、最具影响力的 300 家公司，这些公司横跨众多行业与领域，因而该指数的成分股展现了中国股市的全貌。

本文使用 Python 爬虫技术从 Tushare 财经接口端批量获取沪深 300 指数的日度交易数据，时间跨度为 2010 年 1 月 4 日至 2023 年 12 月 29 日，共计 3320 个交易日，涵盖了 3 轮完整的牛熊周期转换，部分数据样例如表 1 所示。数据集包含 9 个核心特征变量：开盘价、最高价、最低价、收盘价、昨收价、换手率、涨跌幅、成交量和成交额。

为确保模型评估的科学性，本文采用时间序列数据的标准划分方法，将数据集按时间顺序划分为训练集、验证集和测试集三部分，并在验证集与测试集之间设置空白隔离带以避免信息泄露。具体划分如表 2 所示。

为降低市场阶段划分的主观性，本文借鉴 Pagan 与 Sossounov (2003)提出的股票市场周期识别方法，基于指数收盘价序列识别波峰与波谷，并结合累计涨跌幅与日均波动率两个量化指标对市场状态进行界定。具体而言，以连续三个月以上指数呈单边运行且累计涨跌幅绝对值超过 15% 作为趋势性行情的判定

标准，以指数在窄幅区间内反复波动、无明显单边趋势且累计涨跌幅绝对值低于 10% 作为震荡行情的判定标准。依据上述方法，将测试集时段划分为以下四个市场阶段，具体见表 3。

**Table 1.** Sample data of the CSI 300 Index

**表 1.** 沪深 300 指数股票数据样本

Date (D)	Opening Price (P)	Closing Price (C)	High Price (H)	Low Price (L)	Volume (V)	Trading Value (TV)	Price Range Amplitude (A)	Daily Return (R)	Price Change ( $\Delta P$ )	Turnover Rate (TR)
2010/1/4	3592.47	3535.23	3597.75	3535.23	66101080	93419838921	1.75	-1.13	-40.45	0.2
2010/1/5	3545.19	3564.04	3577.53	3497.66	85809641	1.28302E+11	2.26	0.81	28.81	0.26
2010/1/6	3558.7	3541.73	3588.83	3541.17	78473125	1.21046E+11	1.34	-0.63	-22.31	0.24

**Table 2.** Dataset partition

**表 2.** 数据集划分

集合	起止日期	交易日数	占比
训练集	2010-01-04~2019-12-31	2440	73.5%
空白带 1	2020-01-02~2020-02-14	33	1.0%
验证集	2020-02-17~2021-06-30	327	9.8%
空白带 2	2021-07-01~2021-08-13	33	1.0%
测试集	2021-08-16~2023-12-29	557	16.8%

**Table 3.** Market regimes

**表 3.** 市场阶段

阶段	时间区间	交易日数	市场特征	累计涨跌幅	日均波动率
阶段 1: 牛市尾声	2021.08~2021.12	105	高市值权重股及前期热门板块集中下跌，指数趋势由涨转跌	-5.20%	1.20%
阶段 2: 熊市深跌	2022.01~2022.10	200	指数持续下行，单日跌幅超过 3% 的交易日占比达 12%，波动显著放大	-23.10%	1.80%
阶段 3: 弱复苏	2022.11~2023.06	140	指数触底后温和反弹，累计上涨 12.30%，日均波动率与成交量均低于历史均值	12.30%	1.10%
阶段 4: 震荡整理	2023.07~2023.12	112	指数在窄幅区间内反复波动，无明显趋势方向	2.10%	1.00%

### 3.2. 评估指标

本文选择了 4 个性能指标评估模型，分别是均方根误差、平均绝对误差、平均绝对百分比误差和决定系数。

MAE 能直观地观察模型预测值与真实值之间的差距大小，可以更准确的评估预测值的误差程度。计算公式如公式(1)所示：

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (1)$$

RMSE 衡量预测值与真实值之间的偏差程度，常用于评估模型的预测精度。计算公式如公式(2)所示：

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

MAPE 以百分比形式直观反映预测误差相对于真实值的相对大小，能有效评估模型在不同价格水平下的预测准确程度。计算公式如公式(3)所示：

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (3)$$

R\_Squared 是表征方程对观测数值拟合程度的指标。计算公式如公式(4)所示：

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

## 4. 模型构建

本章围绕股票价格预测任务构建深度学习模型框架，系统说明各模型模块的设计思路与实现方式。整体而言，本文通过小波变换对原始时间序列进行多尺度特征重构，在此基础上分别构建 LSTM 与 CNN 子模型以刻画价格序列的长期依赖与局部波动特征，并通过联合损失函数将两类模型进行融合，形成兼顾趋势刻画与短期波动捕捉能力的混合预测模型。

### 4.1. 小波变换模型实现

为缓解股票价格序列非平稳性对深度学习模型建模能力的影响，本文在模型构建阶段引入离散小波变换(Discrete Wavelet Transform, DWT)，对原始时间序列进行多尺度特征分解。具体而言，选用 Daubechies4 (db4)小波对各特征序列实施两层分解。通过该过程，原始信号首先被分解为近似系数与细节系数，分别对应低频趋势成分与高频波动成分；进一步分解后，可得到由长期趋势、中期波动与短期扰动共同构成的多尺度表示，从而在结构上形成由粗到细的层级特征体系。

在小波基函数选择方面，Daubechies 小波因其正交性、紧支撑性及良好的时频局部化特征，被广泛应用于时间序列分析。其消失矩特性可表示为： $\int t^k \psi(t) dt = 0, k = 0, 1, \dots, p-1$ 。

在特征构建过程中，小波变换以滑动窗口为基本处理单元，对每个历史窗口内的特征序列独立实施分解。以长度为 30 个交易日的输入窗口为例，每一维原始特征经两层小波分解后生成三个系数序列(a<sub>2</sub>、d<sub>2</sub>、d<sub>1</sub>)，分别对应不同频率尺度下的信息。随后，将各尺度系数按特征维度进行拼接，形成模型的最终输入特征向量。通过该过程，原始的 9 维特征被映射为 27 维小波系数特征，在不引入外来信息的前提下显著增强了输入数据的表达能力。

### 4.2. LSTM-CNN 混合模型实现

单一深度学习模型在股票价格预测任务中往往只能刻画价格序列的某一侧面特征。LSTM 模型在建模长期时间依赖关系方面具有优势，而 CNN 更擅长捕捉局部波动模式。为同时刻画价格变化的整体趋势与局部结构，本文构建 LSTM-CNN 混合模型，通过联合训练实现两类模型的优势互补。

混合模型的构建包括三个步骤。首先，构建 CNN 分支，以小波变换后的多尺度特征序列为输入，提取价格序列中的局部波动模式，并输出未来第 5 个交易日的预测值；其次，构建 LSTM 分支，使用相同的输入特征以建模价格序列的长期依赖结构，并生成对应的预测结果；最后，通过加权联合两个模型的损失函数，对模型进行协同训练。

各分支模型的损失函数定义如下：

$$L_{\text{CNN}} = \frac{1}{N} \sum_{i=1}^N (R_{\text{CNN},i,t+5} - R_{i,t+5})^2$$

$$L_{\text{LSTM}} = \frac{1}{N} \sum_{i=1}^N (R_{\text{LSTM},i,t+5} - R_{i,t+5})^2$$

$$L_{\text{Total}} = \alpha L_{\text{CNN}} + \beta L_{\text{LSTM}}$$

其中， $N$  为数据点个数，为真实值， $\alpha$  和  $\beta$  为权重参数。这两个参数反映了各模型损失在总损失中的贡献程度。在实验中发现，参数  $\alpha$  值升高、 $\beta$  值下降会使模型的预测表现向 CNN 模型靠近，模型预测值的局部表现更加突出；反之，则会使模型向 LSTM 靠近，模型预测值的整体趋势更加突出。而且，参数  $\alpha$  值不宜过大， $\beta$  值对模型的预测精度影响更大，但当  $\beta$  值达到一定程度时再增大反而会影响整个模型的预测精度，因此两个值需要在合理的区间内取值。经过反复实验，将参数  $\alpha$  和  $\beta$  值分别设为 0.1 和 0.9，以反映联合训练模型中 LSTM 的损失权重大于 CNN 模型，使所提模型的预测效果更优。

联合模型的最终预测值需要按照  $\alpha$  和  $\beta$  的值加权求和，以达到两种模型互补的效果：

$$R_{\text{final}} = \alpha R_{\text{CNN}} + \beta R_{\text{LSTM}}$$

通过联合 RMSE 损失进行训练，CNN 与 LSTM 分支在优化过程中相互约束，在一定程度上起到正则化作用，有助于缓解单一模型可能出现的过拟合问题。CNN 分支侧重于捕捉价格的局部波动特征，LSTM 分支则用于刻画价格的整体趋势，两者的结合使得混合模型在保持模型复杂度可控的同时，实现了对不同时间尺度价格特征的协同建模，从而提升了股票价格预测的稳健性。

## 5. 实证结果

### 5.1. 基准样本外预测表现

在整体样本外预测中，完整模型的预测误差保持在较为稳定的水平。以基准设定为例，样本外 RMSE 约为 276 点，MAPE 约为 5%， $R^2$  约为 0.55，表明模型在不区分市场阶段的情况下，能够在整体层面较好地刻画指数价格的运行特征。该结果为分阶段分析提供了基准参考，也表明后续观察到的预测差异并非源于模型整体性能不足。见表 4：

**Table 4.** Overall out-of-sample forecasting performance

**表 4.** 模型整体预测结果

RMSE	MAE	MAPE	$R^2$
276.30	210.47	4.94	0.55

在不区分市场阶段的情况下，模型在测试样本上的预测表现总体稳定。样本外预测误差处于可接受区间，表明模型能够在整体层面较好地刻画指数价格的运行特征。该结果为后续分阶段比较提供了基准参考，即在统一预测框架下，整体预测能力并不存在系统性失效，阶段性差异更可能来源于市场环境变化而非模型本身。

### 5.2. 不同市场阶段下的预测可靠性差异

分阶段样本外预测结果显示(见表 5)，模型预测表现随市场阶段发生系统性变化，不同阶段之间的误差水平不仅存在显著差异，而且呈现出与市场环境特征高度一致的排序结构。整体而言，预测误差在弱复苏阶段最低，在熊市深跌阶段保持在相对较低水平，在震荡整理阶段明显上升，并在牛市尾声阶段达

到最高水平。该结果表明，预测表现并非在不同时间段内随机波动，而是对市场状态变化表现出稳定的阶段依赖性。

**Table 5.** Forecasting performance across market regimes

**表 5.** 模型阶段预测结果

阶段	RMSE	MAE	MAPE	R <sup>2</sup>
阶段 1: 牛市尾声	580.44	575.10	11.70	-82.38
阶段 2: 熊市深跌	262.83	200.91	4.67	0.19
阶段 3: 弱复苏	164.22	132.64	3.31	-0.83
阶段 4: 震荡整理	168.72	153.94	4.29	0.26

从误差量级看，弱复苏阶段的 RMSE 约为 160 点、MAPE 约为 3%，显著低于其他阶段；熊市深跌阶段的 RMSE 约为 260 点、MAPE 约为 4%~5%，尽管价格波动幅度较大，但预测误差仍处于可控区间；震荡整理阶段的 RMSE 上升至约 170 点，MAPE 接近 4%，显示在缺乏明确趋势的情况下，预测精度有所下降；牛市尾声阶段的预测误差显著放大，RMSE 超过 500 点、MAPE 高于 10%，远高于其余阶段，同时该阶段 R<sup>2</sup> 为负值(-82.38)，表明模型预测效果劣于简单使用样本均值作为预测值，反映出价格结构剧烈调整时历史价格信息的可迁移性明显下降。这种清晰的误差分布格局表明，市场阶段本身对预测可靠性具有重要影响。

为进一步剖析牛市尾声阶段模型失效的具体原因，本文对该阶段预测结果进行了逐日回溯分析。从预测误差的构成来看，模型在该阶段的预测偏差主要表现在方向判断错误，在指数由涨转跌的关键拐点附近，模型连续多日预测上涨而实际下跌，方向性预测准确率不足。

从交易策略的角度看，上述预测错误可能带来严重的实际后果。若依据模型预测构建趋势跟踪策略，在牛市尾声阶段将面临“追涨杀跌”的双重损失：在下跌初期因预测上涨而持有仓位，随后因预测滞后而在价格已大幅下跌后才转为看空，导致策略收益显著劣于简单买入持有策略。这表明，在市场结构发生剧烈调整的阶段，基于历史模式构建的预测模型不仅预测精度下降，其预测结果在应用中甚至可能产生误导性信号。

进一步结合阶段特征可以发现，预测误差较低的弱复苏与熊市深跌阶段，尽管价格运行方向不同，但均具有较强的阶段内一致性：前者表现为价格缓慢回升、波动节奏平稳，后者则表现为持续下行、方向明确。相反，预测误差显著偏高的牛市尾声与震荡整理阶段，通常伴随价格运行结构的不稳定性增强，价格方向频繁调整或处于趋势转折附近，使得历史价格信息在预测中的可迁移性明显下降。

上述结果在阶段层面揭示了预测可靠性与不确定性环境之间的内在联系。具有更强不确定性特征的阶段，其预测误差系统性偏高；而不确定性相对较弱、价格结构更稳定的阶段，预测表现更为可靠。

### 5.3. 稳健性检验

为检验结论对时间窗口设定的敏感性，本文在相同预测步长与训练策略下，将输入窗口由 30 日替换为 20 日与 40 日，并比较完整模型的测试集表现。结果表明，窗口长度会影响总体精度水平，但不改变主要结论。以沪深 300 指数为例，30 日窗口下模型表现相对最优(RMSE 264.77, MAPE 5.29%, R<sup>2</sup> 0.586)；20 日窗口与 40 日窗口下精度有所下降(RMSE 分别为 304.01 与 313.55, MAPE 分别为 5.81% 与 6.17%, R<sup>2</sup> 分别为 0.481 与 0.380)。具体见表 6。

Table 6. Robustness check results

表 6. 稳健性检验结果

窗口大小	数据集	模型	RMSE	MAE	MAPE	R <sup>2</sup>
20	沪深 300	窗口 20 天	304.01	239.42	5.81	0.48
30	沪深 300	窗口 30 天	264.77	212.33	5.29	0.59
40	沪深 300	窗口 40 天	313.55	249.61	6.17	0.38

更为重要的是，不同窗口长度下模型结构贡献与阶段性规律保持一致：小波分解、联合损失与双分支结构均能在不同设定下带来误差降低，其 RMSE 改善幅度与基准窗口结果处于相近区间；分阶段比较亦稳定复现“弱复苏阶段误差最低、牛市尾声阶段误差最高”的特征。上述结果表明，本文关于模型有效性与阶段依赖性的结论具有一定稳健性。

## 6. 结论与启示

本文在统一预测框架与严格样本外设定下，系统分析了股票指数价格预测表现在不同市场阶段中的差异，并对相关结果进行了综合讨论。基于实证结果，可以得到以下结论与启示。

首先，研究表明，股票价格预测表现具有显著的市场阶段依赖性，不同阶段下预测误差呈现出稳定且有序的差异结构。这一发现与适应性市场假说所强调的市场效率与可预测性随环境动态调整的观点相一致，也与既有关于市场状态变化会影响价格行为统计特征的研究结论形成呼应。相较于以往多聚焦收益序列或单一市场状态的研究，本文在价格水平预测框架下直接比较不同阶段的预测误差，使阶段依赖性特征得以更加直观地呈现。

其次，分阶段结果进一步显示，预测可靠性并不取决于价格上涨或下跌方向本身，而更依赖于价格运行结构的稳定程度。弱复苏与熊市深跌阶段尽管价格走势方向不同，但均呈现出较为稳定的价格运行结构，使历史价格信息在预测中的可迁移性保持在较高水平；相反，在牛市尾声与震荡整理阶段，价格方向频繁调整或处于趋势转折附近，预测误差显著放大。该发现在一定程度上补充并修正了仅从牛熊市划分判断预测难度的简化认识。

再次，从综合解释的角度看，不同市场阶段下预测误差的系统性差异，更可能反映市场不确定性环境变化对价格动态结构稳定性的影响。在不确定性较高或冲击频发的阶段，价格运行结构更易发生重塑，历史价格模式在不同时间段之间的可迁移性下降，从而削弱预测关系的稳定性。这一解释与既有关于经济政策不确定性、市场波动及预测稳定性关系的研究结论相一致，并在阶段层面提供了补充证据。

基于上述结论，本文具有以下启示：在理论层面，研究结果进一步支持了股票价格可预测性具有显著时变性与状态依赖性的观点，强调在评价预测模型有效性时有必要引入市场阶段视角。在方法层面，本文表明，在统一模型框架下进行分阶段样本外评估，有助于更准确地识别预测关系的适用边界。对于实际应用而言，预测模型更适合作为条件性工具，其预测结果的使用应结合所处市场阶段与不确定性环境进行审慎判断。

需要指出的是，本文仍存在一定局限性，例如研究对象仅限于单一指数，市场阶段划分方式亦具有一定主观性。未来研究可在不同市场、不同资产类别或更精细的状态识别框架下，进一步检验预测表现的阶段依赖特征及其经济机制。

## 参考文献

- [1] Fama, E.F. (1970) Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, **25**,

- 383-417. <https://doi.org/10.2307/2325486>
- [2] de Bondt, W.F.M. and Thaler, R. (1985) Does the Stock Market Overreact? *The Journal of Finance*, **40**, 793-805. <https://doi.org/10.1111/j.1540-6261.1985.tb05004.x>
- [3] Jegadeesh, N. and Titman, S. (1993) Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, **48**, 65-91. <https://doi.org/10.1111/j.1540-6261.1993.tb04702.x>
- [4] 李云红, 魏宇, 吴晓雄. 中国股票市场适应性特征的实证研究[J]. 管理工程学报, 2016, 30(1): 72-79.
- [5] 周孝华, 宋庆阳. 时变收益可预测性与适应性市场假说——来自中国股市的证据[J]. 中央财经大学学报, 2015(4): 26-34.
- [6] 周侗, 胡家放, 李仲飞. 基于时变隐马尔可夫机制转换模型的多资产配置研究[J]. 系统工程理论与实践, 2025, 45(10): 3223-3244.
- [7] 包双宝, 孙振, 凤兰. 基于吉布斯抽样的我国股票市场周期性研究[J]. 技术经济与管理研究, 2010(1): 20-21+46.
- [8] 杨阳, 万迪昉. 不同市态下投资者情绪与股市收益、收益波动的异化现象——基于上证股市的实证分析[J]. 系统工程, 2010, 28(1): 19-23.
- [9] 崔海蓉, 李铭, 储小俊. 中国股票市场羊群效应的非对称性研究[J]. 阅江学刊, 2023, 15(2): 123-134+174.
- [10] 顾锋娟. 牛熊市波动非对称性与非预期交易量关系实证研究[J]. 财经问题研究, 2012(6): 49-54.
- [11] 陈煜之, 李心悦, 方毅. 基于频率分解的机器学习模型预测效果比较[J]. 统计与决策, 2024, 40(12): 29-34.
- [12] 许伟河. 上证综合指数波动情况研究——基于滚动窗口的马尔科夫链预测模型[J]. 武汉金融, 2015(5): 22-24.
- [13] 李俊儒, 汪寿阳, 魏云捷. 基于波动率测量误差的波动率预测模型[J]. 系统工程理论与实践, 2018, 38(8): 1905-1918.
- [14] 杨子晖, 陈里璇, 陈雨恬. 经济政策不确定性与系统性金融风险的跨市场传染——基于非线性网络关联的研究[J]. 经济研究, 2020, 55(1): 65-81.
- [15] 方意, 于渤, 王炜. 新冠疫情影响下的中国金融市场风险度量与防控研究[J]. 中央财经大学学报, 2020(8): 116-128.
- [16] 李超杰, 郑凯鑫. 中国经济政策与金融市场波动的内在逻辑——基于大数据挖掘的视角[J]. 现代经济探讨, 2022(9): 49-61.
- [17] 陈国进, 张润泽, 赵向琴. 经济政策不确定性与股票风险特征[J]. 管理科学学报, 2018, 21(4): 1-27.
- [18] 潘长春, 王伟强. 经济政策不确定性与股票市场波动性之间的跨类及跨国关联: 来自中美两国的经验证据[J]. 世界经济研究, 2022(11): 89-105.