

基于深度学习的国债利率期限结构的预测与分析

宋辉鹏, 张金良*

河南科技大学数学与统计学院, 河南 洛阳

收稿日期: 2025年1月26日; 录用日期: 2025年2月19日; 发布日期: 2025年2月28日

摘要

利率期限结构及相关问题一直是金融学的研究热点。文章借助于动态Svensson模型、深度学习, 构建了国债利率期限结构的深度学习预测模型。选取2012年1月~2024年5月银行间零息国债收益率数据, 对其进行了实证分析, 并与DS模型、加入宏观经济变量的DS模型、不加入宏观经济变量的深度学习模型的预测结果进行了对比, 结果显示文章提出的深度学习预测模型效果显著提升, 鲁棒性更好。

关键词

利率期限结构, 动态Svensson模型, 深度学习, 鲁棒性

Prediction and Analysis of the Term Structure of Government Bond Interest Rate Based on Deep Learning

Huipeng Song, Jinliang Zhang*

School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang Henan

Received: Jan. 26th, 2025; accepted: Feb. 19th, 2025; published: Feb. 28th, 2025

Abstract

The term structure of interest rates and related issues has always been a research hotspot in finance. With the help of the dynamic Svensson model and deep learning, this paper constructs a deep learning prediction model for the term structure of treasury bond interest rates. This paper selects the yield data of inter-bank zero-interest treasury bond bonds from January 2012 to April 2024 to

*通讯作者。

文章引用: 宋辉鹏, 张金良. 基于深度学习的国债利率期限结构的预测与分析[J]. 应用数学进展, 2025, 14(2): 448-462.
DOI: 10.12677/aam.2025.142084

conduct an empirical analysis. The results show that compared with the prediction results of the DS model, the DS model with macroeconomic variables, and the deep learning model without macroeconomic variables, the deep learning prediction model proposed in this paper has significantly improved its effectiveness and has better robustness.

Keywords

Term Structure of Interest Rate, Dynamic Svensson Model, Deep Learning, Robustness

Copyright © 2025 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

为了使国民经济快速、均衡、稳健、健康地发展,近几年来,全国各级政府发行了适量的政府债券,所发行的政府债券品种也越来越丰富。特别是中国国债的管理、中国国债利率期限结构、利率期限结构一直是金融学的研究热点[1]-[3]。现代利率期限结构模型按照其模型因子的状态,可分为静态利率期限结构模型、动态利率期限结构模型。静态利率期限结构模型常见的有:1971年 McCulloch [4]提出的多项式样条法;1980年 Powell [5]给出的 B-样条法;1987年, Nelson 和 Siegel [6]构建的包含水平、斜率、曲度三因子的即期利率函数表达式,即 NS 模型;随后在 1994 年, Svensson [7]在 NS 模型的基础上增加了一个因子,得到四因子模型,即 NSS 模型。NS 模型和 NSS 模型均具有拟合能力强、参数估计方法简单和普适性优良等优点。2003 年,赵宇龄[8]对多种利率期限结构预测模型进行了对比,发现 Nelson-Siegel 模型、扩展的 Svensson 模型在拟合中国国债利率期限结构时较为准确。2006 年,郭济敏[9]不仅将 NS 模型应用于中国国债利率期限结构的预测,还详细讨论了不同参数的经济含义。

动态利率期限结构模型能够从时间角度分析收益率,适应经济环境频繁更新的利率数据。2006 年, Diebold [10]首次尝试在静态 Nelson-Siegel 模型中加入三个动态因子,得到了动态 Nelson-Siegel 模型(DNS 模型),基于 DNS 模型,对美国国债收益率进行了预测,对比发现, DNS 预测准确度较静态 NS 模型高。2015 年, Nyholm [11]在 DNS 模型的基础上,提出了 RDNS 模型,并借助此模型,对欧洲中央银行收益率进行了预测,预测效果良好。2014 年,何晓群等[12]基于 DNS 模型,对中国国债利率期限结构进行了拟合,拟合效果良好,这说明了 DNS 模型在拟合中国国债收益率时,具有良好的适用性。

近年来,很多学者在对利率期限结构的研究中加入人工智能方法,如神经网络、遗传算法等,其中神经网络模型具有非线性、非局限性、非常定性等特点,具有较强的学习能力,对非线性的、蕴含模糊信息的关系研究具有较好的适用性。2007 年,姜德峰[13]对于利率期限结构模型建造,则以支持向量机的方式建立相应利率模型。也可以将传统预测方法与人工智能方法结合起来,2010 年,董皓舒[14]将指数函数模型与 BP 神经网络结合起来,构建组合模型对利率期限结构进行预测,研究发现这种方法的预测效果优于两者分别单独用于预测的效果。2018 年,李佳等[15]使用上证指数及沪深 300 指数,对 ARMA、RNN、CNN 以及 LSTM 神经网络等方法的预测效果进行对比分析,分析结果显示, LSTM 神经网络方法的预测误差最低,预测最为精准。LSTM 神经网络预测能够弥补 BP 神经网络时序欠缺、梯度容易消失等方面的缺陷。2018 年,任君[16]将四种不同的 LSTM 模型用于对道琼斯指数的预测,结果显示,弹性网正则化 LSTM 的预测效果最好,该模型的泛化能力强于 RNN 以及传统的 BP 神经网络模型。2022 年, Kauffmann 等[17]基于巴西利率期限结构实证分析,提出一种利用神经网络的收益率曲线预测模型,

样本的预测效果优于传统方法, 尤其是在较长期限中。2023 年, Delucchi 等[18]认为使用机器学习的方法能够更好地说明疫情危机、俄乌战争等社会危机, 以及导致的收益率曲线异常现象。

2003 年, Ang 等[19]提出了含宏观经济变量的利率期限结构模型, 使用 VAR 模型, 研究了宏观经济变量和利率期限结构因子之间的动态关系。2006 年, Diebold 等[10]通过检验通货膨胀等宏观经济变量与利率期限结构三因子之间的动态联系, 得出了宏观经济影响未来利率期限结构, 利率期限结构曲线对未来宏观经济影响较弱的结论。2016 年, Levant 等[20]采用 DNS 模型, 拟合预测英国国债利率期限结构, 并研究了宏观经济变量对利率期限结构的影响, 结果发现, 宏观经济变量对利率期限结构的影响会随着剩余期限的变化而变化。国内学者袁靖等[21]、巴曙松等[22]以及李雪等[23]均发现利率期限结构与宏观经济变量之间存在联系。2014 年, 金雯雯等[24]研究利率期限结构时, 考虑了宏观经济变量, 建立了时变参数向量自回归模型(TVP-VAR 模型), 结果表明, 利率期限结构与宏观经济之间的关系存在时变性和滞后性。2022 年, 李依航等[25]从宏观经济变量中提取主成分信息, 将主成分信息纳入 NS 模型中, 对利率期限结构进行预测, 为制定经济政策和投资策略提供参考。

本文在李依航、Delucchi 等研究的基础上, 集动态 Svensson 模型与深度学习的优点, 构建了中国国债利率期限结构混合预测模型。

2. 模型构建

基于动态 Svensson 模型, 拟合出动态四因子, 将宏观经济变量分别与 L_t 、 S_t 、 C_{t1} 、 C_{t2} 四个动态因子进行 TVP-VAR 建模, 筛选出与利率期限结构预测相关的宏观经济变量。将筛选出的变量与利率期限结构一起纳入到 CNN-LSTM 深度学习的预测模型中, 见图 1。

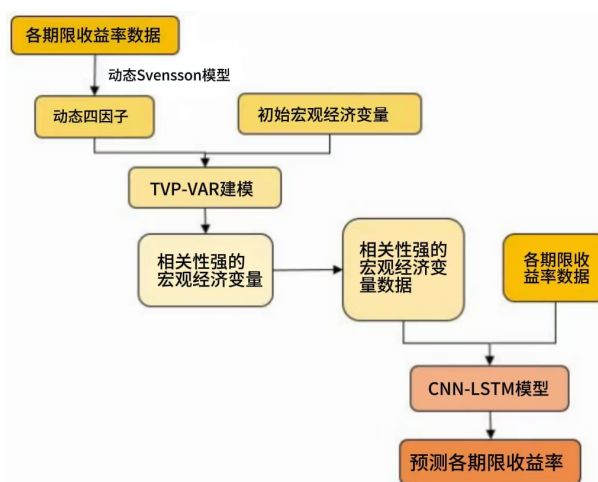


Figure 1. Deep learning prediction model
图 1. 深度学习预测模型

3. 实证分析

3.1. 数据选取

国债市场的交易量和托管量主要来自银行间市场, 本文选取的国债数据为银行间零息国债收益率数据。很多宏观数据的频率都为月度, 因此我们以每个月最后一天的即期收益率作为国债月度数据。数据选取的范围为 2012 年 1 月到 2024 年 5 月, 研究的期限为 1、3、6、12、24、36、60、84、120、180、240、360、600 (单位: 月), 即得到 149×13 个样本数据, 所有数据均来自 Wind 数据库。

3.2. 宏观经济变量选取

本文选取的宏观经济指标是投资者最为关注的且基本覆盖了宏观经济的主要方面。将经济指标分为 5 个大类、13 个分项指标。

经济类：选取价格指标包括消费者价格指数(CPI)、制造业采购经理指数(PMI)、工业增加值同比(IP)。

货币类：选取狭义货币供应量(M1)、广义货币供应量(M2)和社会融资规模(TSF)。

资金类：选取银行间质押式七日回购加权利率(R007)、人民币存款准备金率(RRR)。

宏观景气类：选取宏观经济景气先行指数(MLAI)、宏观经济景气一致指数(MCI)以及宏观经济景气滞后指数(MLEI)。

外部因素类：选取剩余期限为 10 年的美国国债利率期限结构(America10)和美元指数(USDXX)。

3.3. 参数估计

动态 Svensson 模型中的时变参数 L_t 、 S_t 、 C_{t1} 、 C_{t2} 以及 λ_{1t} 、 λ_{2t} 共有 6 个参数，在对样本数据拟合时，要先对参数进行估计，估计参数 λ 采用卡尔曼滤波法，也就是以拟合误差平方均值最小为目标函数优化得出。

$$\text{Min res} = \sum_{i=1}^n [y_i - y_t(\tau)]^2 \quad (1)$$

计算得出 $\lambda_1 = 0.2931$ ，使 C_{t1} 的载荷在剩余期限为 7 年时达到最大， $\lambda_2 = 0.6137$ 使得 C_{t2} 的载荷在剩余期限为 3 年时达到最大。接着在得到 λ 值的基础上，对剩余四个参数 β_1 、 β_2 、 β_3 、 β_4 进行最小二乘法进行估计，得到动态向量 L_t 、 S_t 、 C_{t1} 、 C_{t2} 。

3.4. Pearson 相关性检验

将 Svensson 模型中的参数因子 L_t 、 S_t 、 C_{t1} 、 C_{t2} 与宏观变量之间进行 Pearson 相关性检验，选出与因子相关性高的宏观经济变量，具体结果如表 1 所示：

Table 1. Correlation test between various factors and macroeconomic variables

表 1. 各因子与宏观经济变量的相关性检验

	L_t	S_t	C_{t1}	C_{t2}
CPI	-0.108 (0.283)	0.009 (0.927)	0.138 (0.172)	-0.007 (0.944)
PMI	0.291 (0.003***)	-0.064 (0.524)	-0.107 (0.290)	-0.008 (0.939)
IP	0.385 (0.000***)	0.076 (0.454)	0.083 (0.411)	-0.167 (0.097*)
M1	-0.396 (0.000***)	-0.51 (0.000***)	-0.505 (0.000***)	0.55 (0.000***)
M2	-0.433 (0.000***)	-0.467 (0.000***)	-0.446 (0.000***)	0.537 (0.000***)
TSF	-0.195 (0.052*)	-0.137 (0.174)	-0.067 (0.508)	0.166 (0.099*)
MCI	0.171 (0.089*)	-0.331 (0.001***)	-0.463 (0.000***)	0.239 (0.017**)
MLAI	-0.028 (0.783)	-0.314 (0.001***)	-0.431 (0.000***)	0.252 (0.011**)
MLEI	-0.216 (0.031**)	-0.329 (0.001***)	-0.232 (0.020**)	0.287 (0.004***)
R007	0.608 (0.000***)	0.05 (0.621)	-0.022 (0.826)	-0.388 (0.000***)
RRR	0.552 (0.000***)	0.37 (0.000***)	0.302 (0.002***)	-0.469 (0.000***)
America10	0.599 (0.000***)	-0.369 (0.000***)	-0.441 (0.000***)	0.278 (0.005***)
USDXX	-0.672 (0.000***)	-0.241 (0.016**)	-0.286 (0.004***)	0.34 (0.001***)

注：表中数字为皮尔逊系数，()内数字为 P 值，***、**、* 分别代表 1%、5%、10% 的显著性水平。

从上表结果可以看到：宏观经济变量 M1、M2、MLEI、RRR、America10、USD_X 与四个因子 L_t 、 S_t 、 C_{t1} 、 C_{t2} 的显著性水平平均在 5% 以下。

但 Pearson 相关性检验也存在一些局限性，若变量之间存在非线性关系，那么 Pearson 相关系数可能无法准确反映变量之间的真实关联，需要通过进一步的检验。

3.5. TVP-VAR 建模分析

3.5.1. ADF 检验

在分析各指标的相关性之前，需经过平稳性检验，检验为平稳序列，才能进行时间序列建模分析。借助 EViews 13.0 软件对各组数据进行 ADF 检验。ADF 检验结果为 6 组数据部分为非平稳序列，见表 2，因此对这部分数据进行一阶差分处理，使这 6 组时间序列数据成为平稳序列，见表 3。

Table 2. ADF test results of various factors and macroeconomic variable sequences
表 2. 各因子和宏观经济变量序列的 ADF 检验结果

参数	T 统计量	概率(P)	平稳性结论
L_t	-1.519444	0.5196	不平稳
S_t	-5.150992	0.0003	平稳
C_{t1}	-5.240649	0.0002	平稳
C_{t2}	-6.900654	0.0000	平稳
M1	-2.067011	0.5570	不平稳
M2	-1.442562	0.8410	不平稳
MLEI	1.254534	0.9450	不平稳
RRR	-3.321261	0.0011	平稳
America10	-2.177863	0.2157	不平稳
USD _X	-1.357658	0.5997	不平稳

注：概率 P 大于 0.05 不能拒绝原假设，为不平稳序列。

Table 3. ADF test results of various factors and macroeconomic variable sequences after first-order differencing
表 3. 一阶差分后各因子和宏观经济变量序列的 ADF 检验结果

参数	T 统计量	概率(P)	平稳性结论
$D(L_t)$	-9.293497	0.0000	平稳
$D(M1)$	-12.43583	0.0000	平稳
$D(M2)$	-8.614201	0.0000	平稳
$D(MLEI)$	-14.04135	0.0000	平稳
$D(America10)$	-9.950297	0.0000	平稳
$D(USD_X)$	-10.01119	0.0017	平稳

注：概率 P 大于 0.05 不能拒绝原假设，为不平稳序列。

3.5.2. 确定滞后阶数

借助 EViews 软件中的似然比(LR)检验法，分别从 AIC、SC、HQ、LR、FPE 五个方面来考察。表中的“*”表示最佳滞后阶数，带星号最多的滞后阶数，就是我们选择的最佳滞后阶数。因此最佳滞后阶数为 1，见表 4。

Table 4. Estimation results of lag factor P**表 4.** 滞后因子 P 的估计结果

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-1789.892	NA	0.069528	25.71274	25.92285	25.79812
1	-1216.050	1057.508*	8.01e-05*	18.94357*	21.25487*	19.88281*
2	-1151.789	109.2447	0.000136	19.45412	23.86659	21.24722
3	-1089.780	96.55673	0.000247	19.99685	26.51049	22.64380
4	-1004.286	120.9120	0.000336	20.20409	28.81890	23.70489
5	-930.3314	94.02845	0.000580	20.57616	31.29215	24.93082
6	-843.8699	97.57797	0.000926	20.76957	33.58673	25.97808
7	-731.9046	110.3658	0.001176	20.59864	35.51697	26.66100
8	-614.5220	98.93680	0.001672	20.35031	37.36982	27.26653

3.5.3. MCMC 参数估计

根据得到的最佳滞后阶数，借助 OxMetrics 6 软件对参数采用 MCMC (马氏链 - 蒙特卡罗)方法进行估计，设定抽样迭代的次数是 10,000 次。从统计结果我们可以看到 Geweke 统计量均大于 5%，拒绝原假设，马尔科夫链趋于集中。表中无效因子最大为 61.63，也就是连续抽样 10,000 次能获得 $\frac{61.63}{10000}$ 个无效样本，因此上述的 TVP-VAR 模型的模拟效果好。估计结果如表 5 所示。

Table 5. Estimation results of TVP-VAR model**表 5.** TVP-VAR 模型估计结果

Parameter	Estimation Result					
	Mean	STDEV	95% L	95% U	Geweke	Inef.
Sb1	0.0326	0.0070	0.0222	0.0492	0.135	4.84
Sb2	0.0203	0.0020	0.0168	0.0247	0.153	3.23
Sa1	0.0347	0.0080	0.0226	0.0539	0.326	27.09
Sa2	0.1527	0.3624	0.0411	1.0009	0.018	54.34
Sh1	0.0835	0.0310	0.0429	0.1662	0.919	75.44
Sh2	0.0973	0.0545	0.0428	0.2504	0.083	86.53

3.5.4. 脉冲响应分析

在 TVP-VAR 模型估计完成后，将初步筛选的宏观经济变量对四因子做脉冲响应分析，观察响应变量在不同冲击下的动态变化是否呈现非线性特征，例如，某些变量对冲击的响应可能在不同时间阶段表现出不同的强度或方向。并且 TVP-VAR 模型本身允许系数随时间变化，这种时变性就反映了变量之间关系的非线性特征，通过分析脉冲响应函数在不同时期的变化，可以做进一步动态分析。

图 2 中 M1 对水平因子 L 多为负向冲击，2021 年之后开始变为正向冲击；M2 对 L 的正向冲击和负向冲击时间差不多；MLEI 对 L 的正向冲击较多，且冲击强度大；RRR 对 L 的冲击强度较小，但从 2023 年之后，负向冲击急剧增大；America10 对 L 的冲击在 2022 年之前多为负向冲击，但在 2023 年正向冲击急剧增大；USDx 对 L 的冲击从 2023 年开始，正向冲击急剧增大。

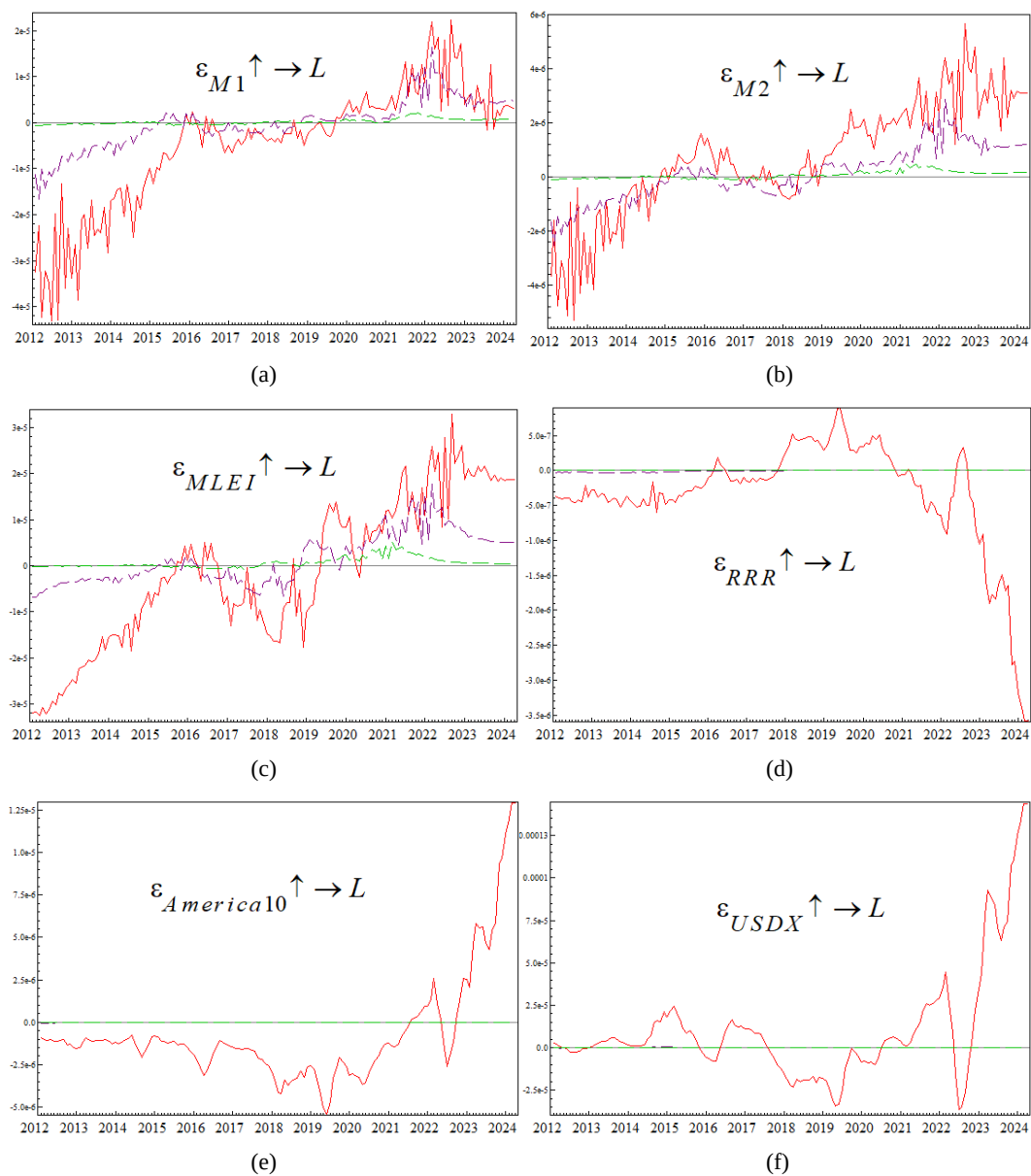
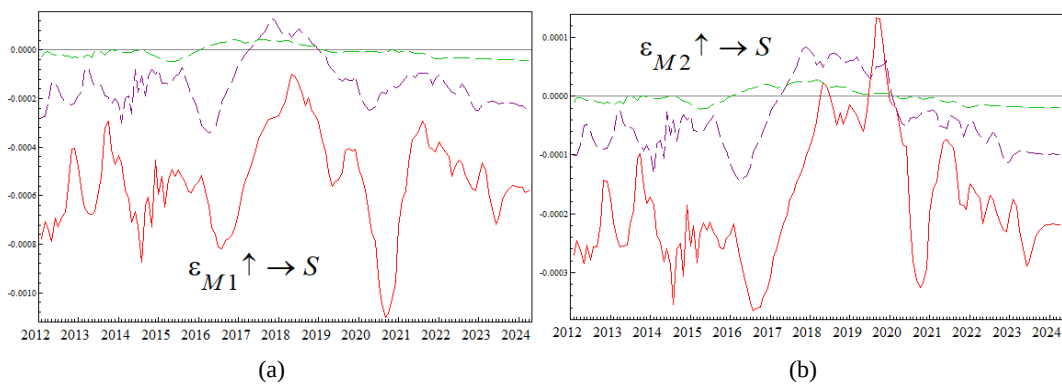


Figure 2. Pulse response diagram of macroeconomic variables to horizontal factor L

图 2. 宏观经济变量对水平因子 L 的脉冲响应图



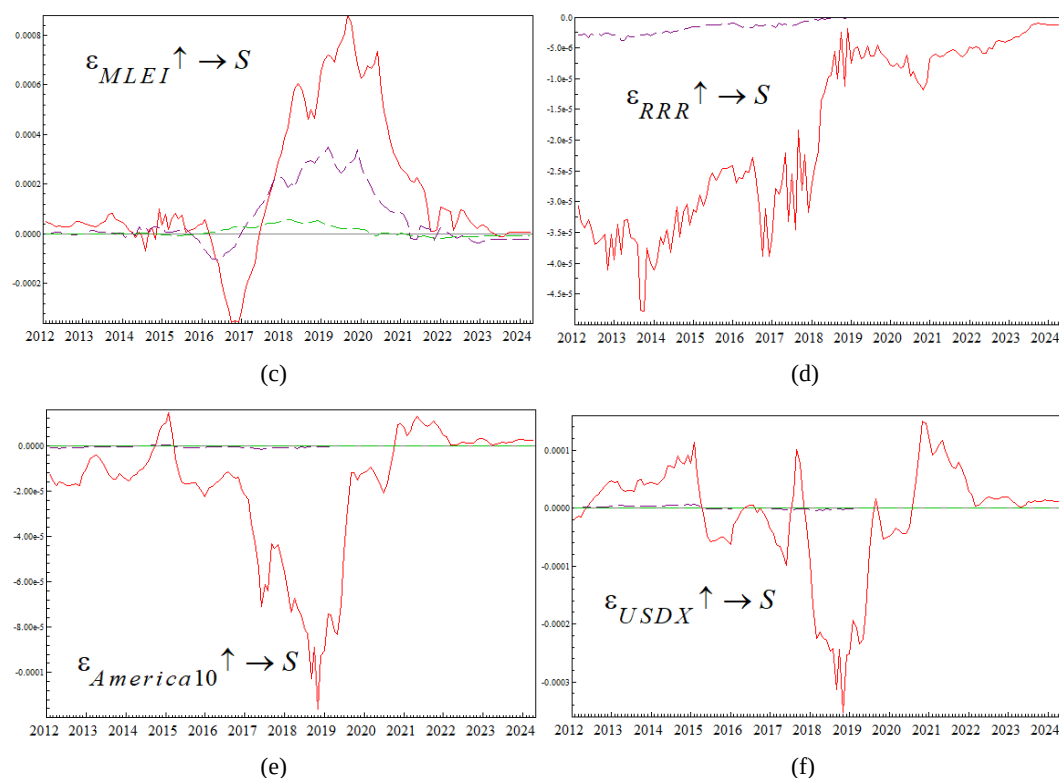
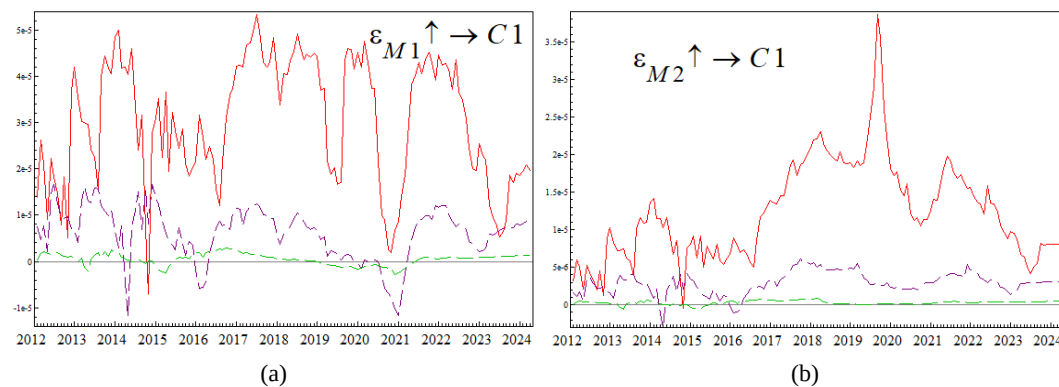


Figure 3. Pulse response diagram of macroeconomic variables to slope factor S

图 3. 宏观经济变量对斜率因子 S 的脉冲响应图

图 3 中 $M1$ 对斜率因子 S 多为负向冲击, 且冲击强度较大; $M2$ 对 S 多为负向冲击, 且冲击强度较大; $MLEI$ 对 S 的冲击强度在 2018 年至 2023 年为正向冲击, 且冲击强度较大, 其余时间冲击强度在 0 周围浮动; RRR 对 S 多为负向冲击, 冲击强度越来越小; $America10$ 对 S 在 2015 至 2020 年为负向冲击, 且冲击强度较大, 其余时间冲击强度在 0 周围浮动; $USDx$ 对 S 在 2018 年至 2020 年为负向冲击, 且强度较大。

图 4 中 $M1$ 对曲率因子 $C1$ 为正向冲击, 且冲击的波动性较大; $M2$ 对 $C1$ 为正向冲击, 在 2019 年 10 月冲击强度最大; $MLEI$ 对 $C1$ 的冲击强度在 0 周围浮动, 在 2017 年转为正向冲击, 在 2019 年达到最大值; RRR 对 $C1$ 为负向冲击, 且冲击强度不断减小; $America10$ 对 $C1$ 在 2015 年至 2019 年多为负向冲击; $USDx$ 对 $C1$ 的冲击方向不是很稳定。



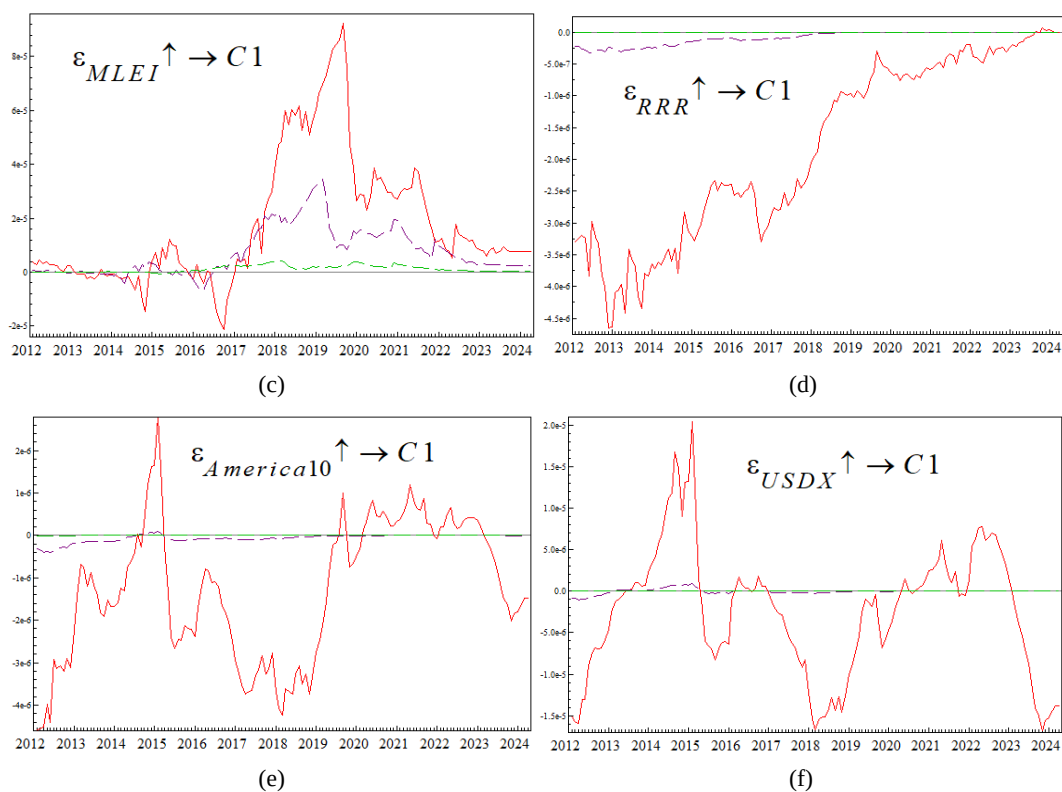
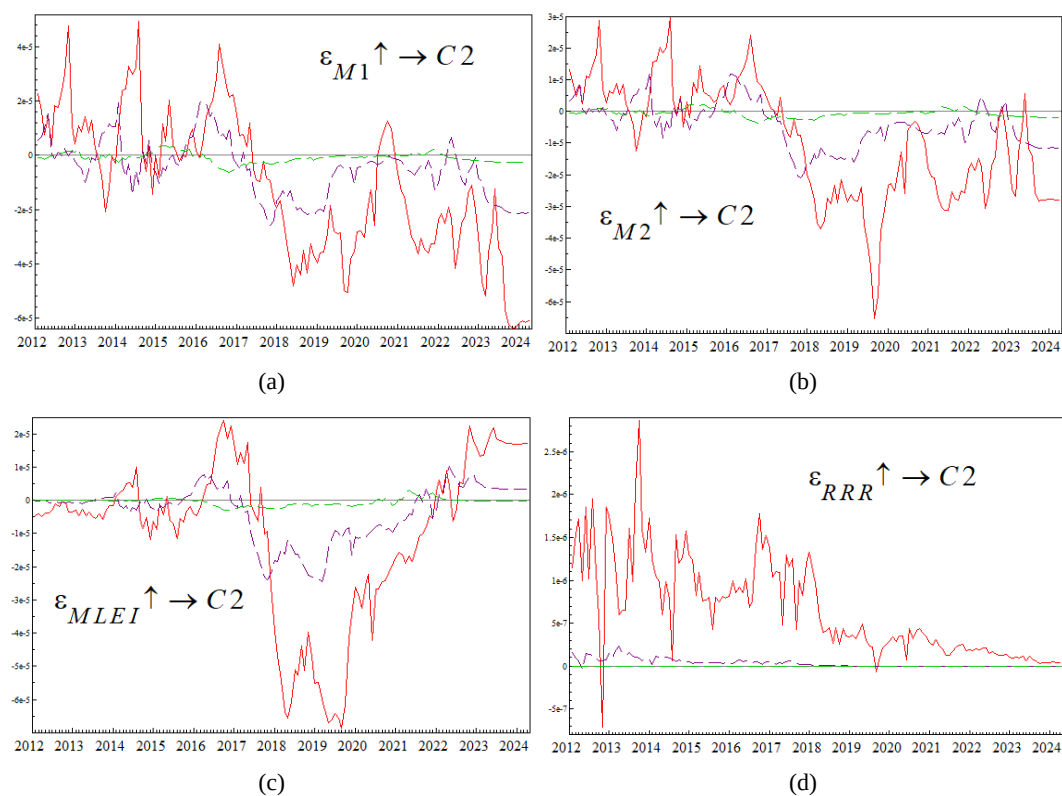


Figure 4. Pulse response diagram of macroeconomic variables to curvature factor C1

图 4. 宏观经济变量对曲率因子 C1 的脉冲响应图



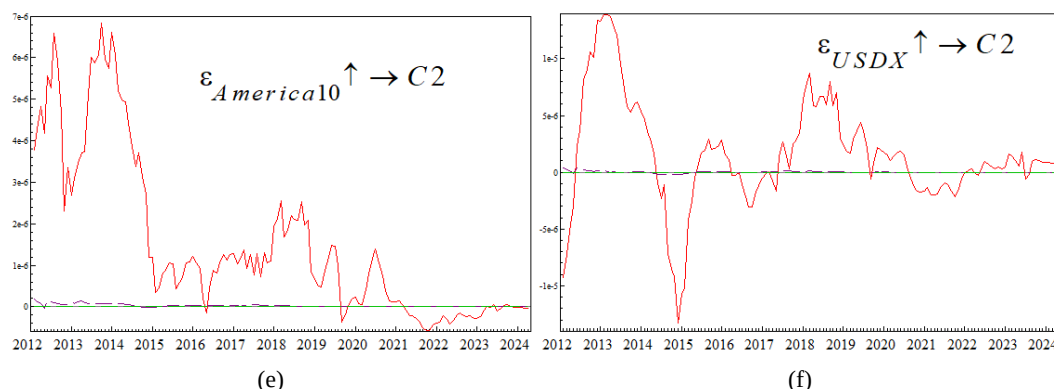


Figure 5. Pulse response diagram of macroeconomic variables to curvature factor C2
图 5. 宏观经济变量对曲率因子 C2 的脉冲响应图

图 5 中 M1 对曲率因子 C2 冲击的方向多变, 在 2018 年变为负向冲击, 且冲击强度不断增大; M2 对 C2 多为负向冲击, 且在 2019 年 9 月负向冲击强度达到最大; MLEI 对 C2 在 2017 年至 2022 年为负向冲击, 且冲击强度较大; RRR 对 C2 为正向冲击, 冲击强度不断减小; America10 对 C2 为正向冲击, 在 2014 年冲击强度最大, 随后不断减小; USDX 对 C2 在 2015 年短暂地出现过负向冲击, 后来冲击强度越来越小。

综上, M1 和 M2 对四因子的冲击图存在高度的相似性, 两者都代表货币量这一经济含义, 但 M1 对四因子的冲击更为敏感; MLEI 作为宏观经济指标, 受社会的经济变动更为明显, 对四因子的冲击也多为正向冲击, 与利率期限结构联系密切; RRR 作为货币政策工具, 对四因子的冲击会受到经济状况的影响而减弱; America10 与 USDX 作为外部因素对四因子的冲击, 存在相似性, 但 America10 对四因子的冲击更为敏感。M1、MLEI、RRR、America10 四个变量与利率期限结构联系更为密切, 更具代表性, 因此, 将 M1、MLEI、RRR、America10 四个宏观经济变量纳入到利率期限结构的预测中去。

3.6. CNN-LSTM 神经网络预测

将筛选出的宏观经济变量 M1、MLEI、RRR、America10 作为新的变量加入到输入集中, 卷积神经网络(CNN)可以提取输入数据的特征, 再经过长短时记忆网络(LSTM)学习这些特征, 对数据集做出预测。

在输入数据之前要先对数据进行归一化处理,

$$N(x_i) = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (2)$$

处理后的数据划分为训练集和测试集, 并通过实验对神经网络各个参数进行设置, 经过实验将 MiniBatchSize 设置为 128, MaxEpochs 设置为 500, 学习率设置为 0.01 等等, 训练混合网络。

模型的预测效果用均方误差(RMSE)来检验, 从下图 6 中我们可以清楚地看到, 剩余期限为 1 月(短期)、60 个月(中期)、600 个月(长期)的 CNN-LSTM 训练模型的效果准确, 其中中期预测最为准确, 短期次之。

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y})^2} \quad (3)$$

从表 6 的对比结果可以发现: 新加入宏观经济变量这一约束条件后, 模型的拟合误差都有不同程度的减小, 拟合精度更高。但对比 DS 模型加入宏观经济变量预测后, 剩余期限 1 月、3 月的预测精度还

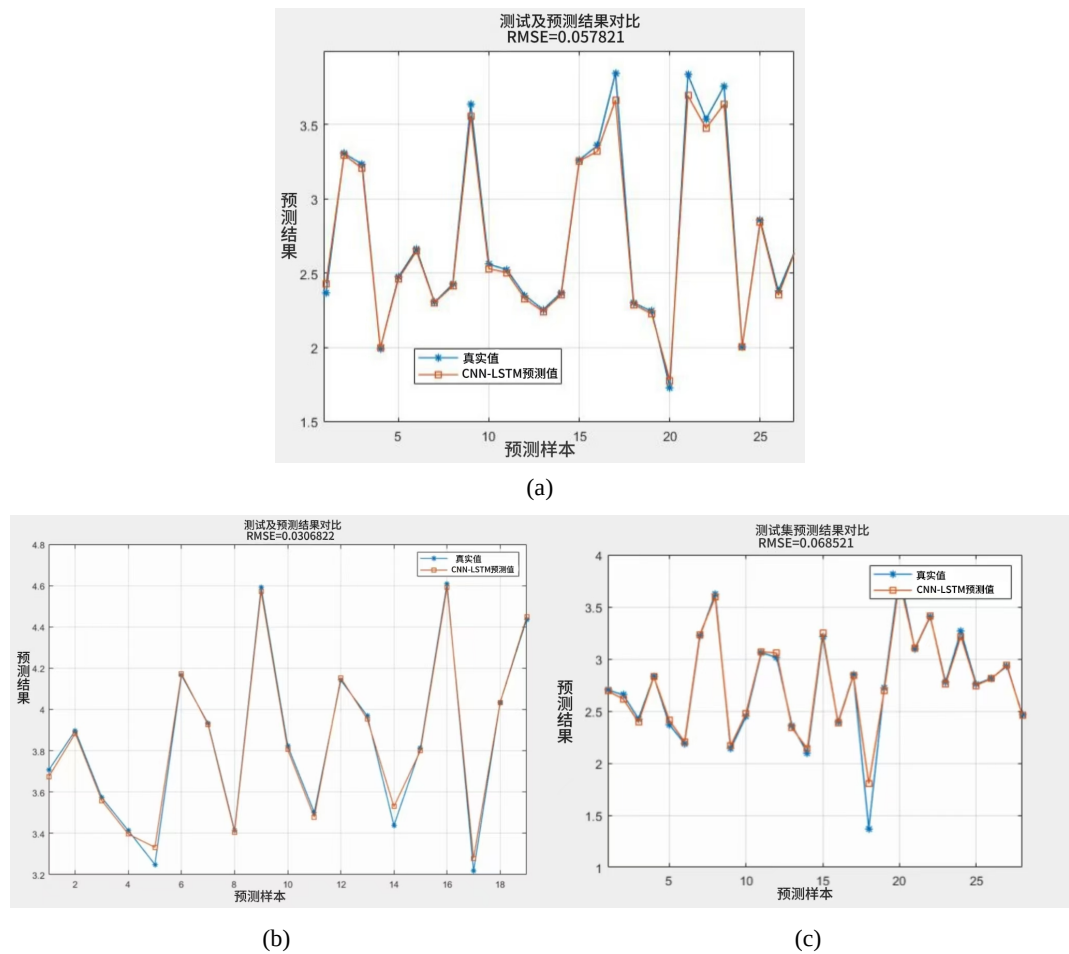


Figure 6. Prediction results of test sets with remaining maturities of 1 month, 60 months, and 600 months
图 6. 剩余期限为 1 月、60 个月、600 个月的测试集预测结果

Table 6. RMSE comparison of various model test sets before and after adding macroeconomic variables
表 6. 加入宏观经济变量前后各模型测试集的 RMSE 对比

Maturities (month)	DS	DS (加入宏观经济变量)	CNN-LSTM	CNN-LSTM (加入宏观经济变量)
1	0.14154	0.24636	0.13451	0.05782
3	0.10225	0.16992	0.12131	0.03421
6	0.10375	0.10347	0.12853	0.03076
12	0.10999	0.10021	0.11068	0.02163
24	0.10021	0.09651	0.08634	0.02109
36	0.07923	0.07097	0.06581	0.01577
60	0.10442	0.07035	0.07782	0.03068
84	0.13163	0.06741	0.05203	0.02041
120	0.26679	0.08152	0.0991	0.02931
180	0.11989	0.07112	0.10985	0.03841

续表

240	0.08541	0.06409	0.09469	0.04465
360	0.12346	0.09615	0.12364	0.05891
600	0.26547	0.10586	0.16457	0.06852

不如 DS 模型,可能是由于短期收益率的影响因素比较多,情况相对复杂,加入宏观经济变量反而没有提高预测能力。从中长期来看,加入宏观经济变量的 DS 模型预测效果还是有所提高。通过 CNN-LSTM 模型拟合的 RMSE 值来看,几乎都是呈现出先下降再上升的趋势,大概在剩余期限为 36 个月至 84 个月间误差最小,在这个期限范围内模型的拟合效果最佳。同时这也与我们前文假设 $\varphi_1 = 0.2931$ 会使得 C_{t1} 的载荷在剩余期限为 7 年时取得最大, $\varphi_2 = 0.6137$ 会使得 C_{t2} 在剩余期限为 3 年时达到最大相契合。综上,通过对比各个模型的 RMSE 值,加入宏观经济变量的 CNN-LSTM 模型的预测效果更好,尤其是中期预测。

补充确定系数(R^2)、平均绝对误差(MAE)两个指标来检验,更全面地评估模型的拟合能力和预测精度。

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

$$\text{MAE}(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}| \quad (5)$$

Table 7. Comparison of R^2 values of various model test sets before and after adding macroeconomic variables

表 7. 加入宏观经济变量前后各模型测试集的 R^2 对比

Maturities (month)	DS	DS (加入宏观经济变量)	CNN-LSTM	CNN-LSTM (加入宏观经济变量)
1	0.94539	0.88325	0.97445	0.98397
3	0.97312	0.94723	0.97347	0.98855
6	0.96119	0.95294	0.98891	0.98354
12	0.98426	0.97461	0.99257	0.99805
24	0.98025	0.96131	0.99759	0.99520
36	0.98618	0.97335	0.99354	0.99849
60	0.98866	0.98316	0.93749	0.99432
84	0.98578	0.90691	0.97146	0.99673
120	0.97062	0.60369	0.95676	0.99582
180	0.95312	0.92375	0.98526	0.98479
240	0.96119	0.97642	0.98865	0.98283
360	0.95143	0.86455	0.97303	0.97841
600	0.95025	0.62071	0.96183	0.96548

从表 7 中对比结果可以看到：CNN-LSTM (加入宏观经济变量)的 R^2 值最接近于 1，说明该模型预测效果最好。从表中可以看到短期、中期的 R^2 值更大，预测效果更好，优于长期。加入宏观经济变量的预测其预测精度都有不同程度的提高，说明本文提出的 CNN-LSTM (加入宏观经济变量)模型对数据的拟合度高。

Table 8. Comparison of MAE of each model test set before and after adding macroeconomic variables
表 8. 加入宏观经济变量前后各模型测试集的 MAE 对比

Maturities (month)	DS	DS (加入宏观经济变量)	CNN-LSTM	CNN-LSTM (加入宏观经济变量)
1	0.15834	0.19353	0.05712	0.04419
3	0.07708	0.13278	0.03898	0.03983
6	0.09091	0.12732	0.02967	0.03342
12	0.09977	0.10504	0.02617	0.01458
24	0.08965	0.09467	0.01789	0.01814
36	0.04232	0.08790	0.02971	0.01538
60	0.09137	0.06481	0.04569	0.03121
84	0.12221	0.06393	0.01901	0.01897
120	0.26191	0.07014	0.04789	0.18699
180	0.10112	0.06103	0.03291	0.04016
240	0.05695	0.05313	0.03915	0.02382
360	0.14656	0.08047	0.03342	0.03427
600	0.25332	0.07982	0.05237	0.02999

MAE 可以提供一个直观的误差量度，帮助理解模型在实际应用中的表现，MAE 数值越低，说明模型的预测值与真实值之间的平均绝对偏差越小，模型的预测精度越高，从表 8 中对比结果看到：在中短期预测中，加入宏观经济变量后的 CNN-LSTM 模型的 MAE 值降低，显示出更好的预测性能，如果目标是进行短期或中期预测，加入宏观经济变量后的 CNN-LSTM 模型可能是更好的选择。对于长期预测，DS 模型的预测也很准确，因此长期预测需要更仔细地考虑宏观经济变量的加入是否有助于提高模型性能，对比选择更加适合的预测模型。

3.7. 稳健性检验

为验证 CNN-LSTM 模型的稳健性，将符合正态分布、均值为 0、方差为 0.05 的噪声添加至国债利率期限结构数据中，加噪后剩余期限为 1 月的预测结果见图 7。

加噪后模型 RMSE 为 0.053206，证实本文模型在加噪数据上也能表现良好，说明模型鲁棒性高，具有更好的稳健性，可以更好地应对现实世界中的不确定性和变化性。

4. 结论

1) 构建了融入金融知识的中国国债利率期限结构深度学习预测模型。与 DS 模型、加入宏观经济变量的 DS 模型、不加入宏观经济变量的 CNN-LSTM 模型的预测结果进行对比，实证结果显示本文提出的融入金融知识的模型预测效果显著提升，鲁棒性更好。

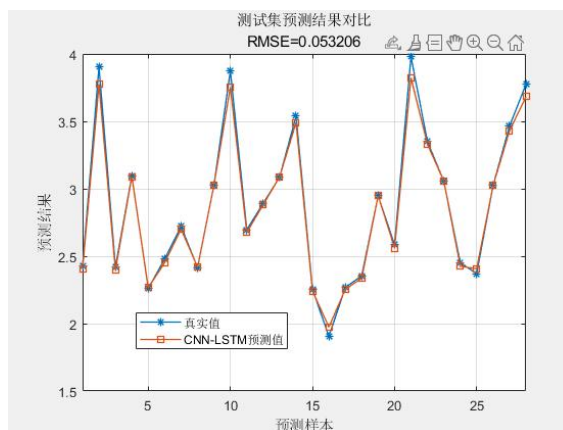


Figure 7. Prediction results of the test set with a remaining period of 1 month after adding noise

图 7. 加噪后剩余期限为 1 月的测试集预测结果

2) 基于动态 Svensson 模型, 利用 Pearson 相关性检验, 初步筛选出与动态 Svensson 模型中四个动态因子 L_t 、 S_t 、 C_{t1} 、 C_{t2} 相关性强的宏观经济变量; 建立了 TVP-VAR 模型, 并从冲击方向和冲击强度两个方面, 分析了宏观经济变量对四个动态因子的影响。本文数据选取的范围为 2012 年 1 月到 2024 年 5 月内, M1、MLEI、RRR、America10 四个变量与利率期限结构联系更为密切, 更具代表性。

基金项目

国家自然科学基金(51675161)资助项目。

参考文献

- [1] Fisher, I. (1896) Appreciation and Interest. *Publications of the American Economic Association*, **11**, 1-98.
- [2] Culbertson, J.M. (1957) The Term Structure of Interest Rates. *The Quarterly Journal of Economics*, **71**, 485-517. <https://doi.org/10.2307/1885708>
- [3] Keynes, J.M. (1936) *The General Theory of Employment, Interest and Money*. Harcourt Brace.
- [4] McCulloch, J.H. (1971) Measuring the Term Structure of Interest Rates. *The Journal of Business*, **44**, 19-31. <https://doi.org/10.1086/295329>
- [5] Powell, M.J.D. (1980) *Approximation Theory and Methods*. Cambridge University Press.
- [6] Nelson, C.R. and Siegel, A.F. (1987) Parsimonious Modeling of Yield Curves. *The Journal of Business*, **60**, 473-489. <https://doi.org/10.1086/296409>
- [7] Svensson, L.E.O. (1994) Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994. *IMF Working Papers*, **1994**, 94-114. <https://doi.org/10.5089/9781451853759.001>
- [8] Zhao, Y.L. (2003) A Comparative Analysis of the Construction of China's Treasury Yield Curve. *Shanghai Finance*, **9**, 29-31.
- [9] 郭济敏, 张嘉为. 基于 Nelson-Siegel 模型预测中债国债收益率曲线形态[J]. 债券, 2016(7): 66-72.
- [10] Diebold, F.X. and Li, C. (2006) Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics*, **130**, 337-364. <https://doi.org/10.1016/j.jeconom.2005.03.005>
- [11] Nyholm, K. (2015) A Rotated Dynamic Nelson-Siegel Model with Macro-Financial Applications. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2666978>
- [12] 何晓群, 王彦飞. 中国利率期限结构与宏观经济运行的关系——基于动态 Nelson-Siegel 模型的研究[J]. 经济理论与经济管理, 2014, 34(8): 69-77.
- [13] 姜德峰. 基于神经网络的利率期限结构研究[J]. 青海金融, 2007(5): 21-23.
- [14] 董皓舒. 基于智能神经网络组合预测的国债利率期限结构建模与实证[J]. 中国市场, 2010(44): 47-49.

- [15] 李佳, 黄之豪, 陈冬兰. 基于 LSTM 等深度学习方法的股指预测研究[J]. 软件导刊, 2019, 18(9): 17-21.
- [16] 任君, 王建华, 王传美, 等. 基于正则化 LSTM 模型的股票指数预测[J]. 计算机应用与软件, 2018, 35(4): 44-48+108.
- [17] Kauffmann, P.C., Takada, H.H., Terada, A.T. and Stern, J.M. (2022) Learning Forecast-Efficient Yield Curve Factor Decompositions with Neural Networks. *Econometrics*, **10**, Article No. 15.
<https://doi.org/10.3390/econometrics10020015>
- [18] Delucchi, A. and Giribone, P.G. (2023) Modeling the Interest Rates Term Structure Using Machine Learning: A Gaussian Process Regression Approach. *Risk Management Magazine*, **18**, 16-35. <https://doi.org/10.47473/2020rmm0131>
- [19] Ang, A. and Piazzesi, M. (2003) A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables. *Journal of Monetary Economics*, **50**, 745-787.
[https://doi.org/10.1016/s0304-3932\(03\)00032-1](https://doi.org/10.1016/s0304-3932(03)00032-1)
- [20] Levant, J. and Ma, J. (2017) A Dynamic Nelson-Siegel Yield Curve Model with Markov Switching. *Economic Modelling*, **67**, 73-87. <https://doi.org/10.1016/j.econmod.2016.10.003>
- [21] 袁靖, 薛伟. 中国利率期限结构与货币政策联合建模的实证研究[J]. 统计研究, 2012, 29(2): 42-47.
- [22] 巴曙松, 袁佳, 廖慧. 通货膨胀水平、股票市值与中国国债利率期限结构[J]. 金融发展研究, 2017(1): 3-10.
- [23] 李雪, 朱超, 易祯. 人口学特征与利率期限结构: 老年社会平缓的收益率曲线[J]. 金融研究, 2020(6): 96-113.
- [24] 金雯雯, 陈亮, 毛德勇. 利率期限结构内含的宏观经济信息——基于 TVP-VAR 模型的时变参数研究[J]. 经济评论, 2014(5): 123-135.
- [25] 李依航, 陈越. 纳入宏观经济因子的利率期限结构预测模型构建——基于主协变量回归(PCovR)和动态 Nelson-Siegel 方法[J]. 金融监管研究, 2022(7): 55-74.