

央行货币政策调控的影响因素识别

——基于机器学习方法的实证研究

张旭, 周潇

南京信息工程大学管理工程学院, 江苏 南京

收稿日期: 2024年4月9日; 录用日期: 2024年4月23日; 发布日期: 2024年5月31日

摘要

本文使用文本分析方法并结合现有文献, 选取了14个可能影响央行货币政策调控的经济变量, 并通过随机森林模型和SHAP值法量化了这些变量对2003~2022年央行货币政策调控的影响程度。研究结果表明, 在2003~2022年这段时期, 以通胀缺口为主的泰勒规则和以中美10年期国债利差为代表的汇率因素在我国央行货币政策调控的参考指标中占据重要地位。有鉴于此, 本文建议应通过重视通胀在宏观经济调控中的作用; 加强对外汇市场的关注程度; 深化利率市场化改革来完善货币政策调控。

关键词

利率, 货币政策, 机器学习, SHAP值

Identifying the Determinants of Central Bank Monetary Policy Control

—An Empirical Study Based on Machine Learning Methods

Xu Zhang, Xiao Zhou

School of Management Science and Engineering, Nanjing University of Information Science & Technology, Nanjing Jiangsu

Received: Apr. 9th, 2024; accepted: Apr. 23rd, 2024; published: May 31st, 2024

Abstract

This paper employs text analysis methods in conjunction with existing literature to select 14 economic variables that may influence the central bank's monetary policy control. Through the Random Forest model and SHAP value method, it quantifies the impact of these variables on the cen-

tral bank's monetary policy control from 2003 to 2022. The results indicate that during this period, the Taylor rule, primarily based on the inflation gap, and exchange rate factors represented by the US-China 10-year government bond yield spread, occupy an important position in the reference indicators for the central bank's monetary policy control in China. In light of this, the paper suggests that attention should be paid to the role of inflation in macroeconomic control; the importance of the foreign exchange market should be enhanced; and interest rate marketization reforms should be deepened to improve monetary policy control.

Keywords

Interest Rates, Monetary Policy, Machine Learning, SHAP Value

Copyright © 2024 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

利率是资金的价格, 是重要的宏观经济变量, 对宏观经济均衡和资源配置有重要导向意义, 因此也成为各国央行在实行货币政策时的关键指标。自 1995 年《中华人民共和国银行法》面世以来, 中国人民银行对货币政策的调控始终遵循着“保持货币币值稳定, 并以此促进经济增长”的多目标制, 因此, 保持物价稳定和增加社会总产出成为央行货币政策调控的重要参考, 这与各国央行普遍遵循的泰勒规则高度相符。然而, 随着中国经济发展进入新常态, 经济总体增速放缓, 内需不足等问题开始浮现, 继续紧盯标准泰勒规则实行货币政策难以顾及到经济发展的各个方面。为此, 已有研究主要围绕泰勒规则的各种扩展形式对中国货币政策的调控模式开展了大量的分析, 尝试增强央行货币政策的调控效果。

本文通过引入目前较为流行的文本分析方法, 统计了 2003~2022 年央行货币政策执行报告中的高频关键词, 并结合现有研究, 选取了 14 个可能影响央行货币政策调控的宏观经济变量, 加入到随机森林 (Random Forest) 这一广泛使用的机器学习模型中, 并通过计算每一个宏观经济变量的 SHAP 值, 量化不同变量对 2003~2022 年央行货币政策利率的影响程度并以可视化的方式展示。本文将机器学习方法引入到对央行货币政策调控的影响因素分析中, 弥补了基于泰勒规则方法下研究变量较少的缺点, 对于完善央行货币政策调控和市场对央行货币政策利率的预测具有一定的理论价值和现实意义。

2. 文献综述

自上世纪 80 年代以来, 世界上主要国家中央银行已经开始逐步放弃货币数量目标制并转向利率调控, 各国的利率决策都隐含地遵循“泰勒规则”, 并以此作为重要的利率决策参考, 我国也不例外。随着我国经济的市场化程度不断提高, 数量目标的局限性开始显现, 仅靠数量型调控已难以满足我国货币政策调控的需要[1]。李宏瑾和苏乃芳也同样认为在利率市场化基本完成和流动性格局逆转的当下, 货币政策价格调控方式转型的必要性和迫切性日益上升[2]。

许多学者已经对中国货币政策调控的参考变量开展了研究。谢平和罗雄运用历史分析法和反应函数法首次将中国货币政策运用于检验泰勒规则, 发现泰勒规则可以很好地衡量中国货币政策[3]。但由于泰勒规则中潜在产出和均衡实际利率都属于无法通过观测得到的潜在自然率变量, 所以此后的大多数文献都集中于研究通过不同的估计变量, 对泰勒规则进行不断的修正。如郑挺国和王霞采用多种退势方法的产出缺口序列, 比较了基于实时数据和最终数据的泰勒规则的差异[4]。陈创练等构建了关于利率、通胀

与产出缺口之间的时变参数结构向量自回归模型, 并通过反推得到时变参数泰勒规则的估计[5]。

随着研究的不断深入, 一些学者测试了在央行货币政策反应函数中纳入其他的重要变量, 认为央行应将关注点扩展到通胀和产出缺口之外, 包括资产和大宗商品价格、汇率、货币供应、金融稳定指标等变量。在商品市场和经济政策方面, Tiwari 的研究结果揭示了能源市场与国际货币政策之间随时间变化的关联性, 以及“石油对货币政策”的风险溢出, 这表明石油价格控制着货币政策, 此外, 经济不确定性对货币政策和能源价格关联性产生积极影响[6]。在汇率和货币层面, Chevapatrakul 将有效汇率和货币供应量等货币指标加入到模型中, 发现能够有效改善央行利率决策的预测, 对预测的边际贡献与其他变量相当[7]。对于财政与货币政策之间的关联, Jurkšas 研究发现欧洲央行可能将财政赤字增加视为欧元区总需求潜在上升的迹象, 从而对货币政策产生影响, 但只有欧洲最大经济体德国的统计效果显著[8]。除了经济指标之外, 一些学者开始尝试加入其他可能的变量来改进央行货币政策反应函数, Jonas and Johannes 认为欧洲央行行长沟通中讨论了许多不同的变量和指标, 每个变量和指标至少隐含地表明了央行行长对每个变量的某种反应, 因此他们使用欧洲央行沟通的文本分析结果作为分析欧洲央行货币政策反应函数的附加工具[9]。国内研究中, 单强等人构建了中国化的泰勒规则, 研究表明将中美两国 10 年期国债利差作为汇率因素, 将房价涨幅偏离经济增长幅度替代股票作为资产因素扩展泰勒规则是合理的[10]。

目前大多数围绕利率所开展的研究均采用的是传统计量模型, 这些模型因其理论基础牢固、模型解释性强而受到学术界的青睐。但由于研究发现的央行货币政策调控影响因素数量越来越多, 基于传统计量模型进行的研究出现了一定的局限性。近年来, 机器学习技术的快速发展为突破这些限制提供了新的途径。例如传统计量模型如果包含的变量太多, 往往会产生模型求解的复杂性与不稳定性, 而机器学习模型则具有较强的特征选择能力, 能够自动识别和利用对预测目标有重要影响的变量, 从而有效避免维度灾难问题[11]。此外, 与传统计量模型相比, 机器学习模型在处理高维数据、捕捉变量间复杂交互效应方面展现出显著优势, 特别是能够自动从数据中学习变量间的非线性关系, 无需事先对数据结构做出严格假设。

基于上述几点原因, 本文选择机器学习模型中的随机森林模型。随机森林模型作为一种集成学习算法, 自从 Breiman 提出以来, 已广泛应用于各种预测和分类任务中, 王达和周映雪研究了随机森林模型在宏观审慎监管中的应用, 结果表明随机森林模型对系统性风险的识别显著优于传统的逻辑回归模型, 对训练集数据学习后的模型不仅通过了稳健性检验, 而且具有优异的泛化性能[12][13]。此外, 通过机器学习模型与 Lundberg and Lee 提出的 SHAP 解释模型相结合, 可以更直观地了解到各个输入特征对机器学习模型预测的贡献度, 从而使模型决策过程更为透明[14]。研究证明, SHAP 模型能够提高机器学习模型的可解释性[15]。这也是本文研究的重要理论基础。

本文的边际贡献在于: 第一, 本文将优化和解释增强的随机森林模型应用于研究央行货币政策调控影响因素的新场景, 展示了机器学习技术在传统领域的新应用潜力。通过这些实证研究, 本文不仅验证了模型的有效性, 也为相关领域的研究和实践提供了新的思路 and 工具。第二, 相较于过往文献中通常聚焦于少数几个关键变量的研究, 本文的多变量分析提供了更为全面的视角, 能够更全面地捕捉和理解影响央行货币政策调控的各种潜在因素, 为央行货币政策研究提供了新的视角。

3. 研究模型构建

3.1. 货币政策报告词频统计

传统研究均围绕泰勒规则中的变量或加入少量新变量进行研究以改善利率预测性能, 本文参考 Jonas and Johannes 的方法, 希望通过央行沟通来全方位和多角度的衡量央行货币政策调控的潜在影响因素。因

此本文利用目前文本分析方法中流行的 `jieba` 分词器对央行货币政策执行报告文本进行分词, 并统计每个词汇的词频, 最后通过与现有文献的研究变量进行比较和筛选, 设定可能影响央行货币政策调控的变量。但是由于 `jieba` 分词器是一个通用的中文分词工具, 对于一些金融类专业名词的分词效果较差, 因此本文在第一次分词结果的基础上, 对一些大类概括词进行了细分, 例如将“市场”一词细分为“金融市场”、“货币市场”、“汇率市场”、银行间市场、“股票市场”、“债券市场”、“证券市场”等在执行报告中出现过的词汇, 并将其加入 `jieba` 分词词典重新进行分词和统计。经过多次重复操作, 大多数金融类专业名词都能够被有效统计。

3.2. 机器学习模型构建和参数设定

已有探究央行货币政策调控的影响因素的文献通常都是通过研究泰勒规则等经典理论中的不同变量, 或在泰勒规则的基础之上进行拓展, 构建一个新的泰勒规则模型, 并用新的模型对央行货币政策利率进行预测, 最后将结果与之前的模型进行比对以探究新加入的变量对央行货币政策变化的解释程度。但是这些方法普遍存在研究变量较少, 难以量化每一个变量对货币政策影响程度的大小等问题。相比于传统计量模型, 机器学习模型通过正则化技术、集成方法和自动特征选择来克服这些问题, 能有效地处理成较多的变量, 并且预测函数更加灵活, 从而能够更有效地把握各个研究变量对央行货币政策调控的影响。

由于研究变量较多, 需要尽量减少变量之间多重共线性对研究结果的影响, 因此本文在众多机器学习模型中选择了随机森林(Random Forest)和梯度提升树(Gradient Boosting Trees, GBT)模型, 这两种流行的机器学习方法因其处理变量间多重共线性的能力而被广泛使用。本文将使用随机森林进行实证分析, 用梯度提升树模型进行稳健性检验。

随机森林是由 Breiman 提出的一种强大的分类和预测工具, 通过组合多个决策树来提高预测的准确性并控制过拟合现象。在随机森林中, 每棵树独立构建, 整个模型的构建过程中引入了两种形式的随机性: 一是通过自助采样生成不同的训练数据子集; 二是在决策树的每个节点分裂时, 从候选的特征子集中随机选择特征进行最佳分裂。这种随机性的引入显著降低了模型的方差, 提高了对新数据的泛化能力。

本文利用随机搜索方法对随机森林模型的关键参数进行了精细调优, 针对模型参数如树的数量、树的最大深度、内部节点再划分所需的最小样本数, 以及叶节点的最小样本数进行了优化。最终确定的随机森林模型最优参数配置如下: 模型包括 170 棵树, 最大深度设为 9 层, 可以在充分学习同时避免过深的树结构导致的过拟合问题, 每个内部节点再划分所需的最小样本数设定为 6, 以及每个叶节点的最小样本数为 2。这一参数设定能够最大化模型的学习能力, 同时避免过拟合, 确保模型能够捕捉到数据中的复杂模式和结构。

3.3. 基于 SHAP 值解释性方法的货币政策调控影响因素识别

SHAP 值解释性方法由 Lundberg and Lee 提出, 用来解释机器学习模型中每个影响因素对目标预测值的贡献大小, 其中每一个影响因素被称为一个特征。对于给定的预测模型 f 和包含 N 个特征的输入样本 x , 模型的预测结果 $f(x)$ 可以被解构为所有特征的 SHAP 值之和加上一个基准值 ϕ , 其中 ϕ 通常是模型预测的平均值。对于研究中需要特别关注的每一个特征 i 对特定样本预测结果的贡献量, 其 SHAP 值表示为 ϕ_i , 公式如下:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!} \times [f(S \cup \{i\}) - f(S)] \quad (1)$$

其中, N 为模型中所有特征构成的全集, 包含了参与预测的所有影响因素。 S 为一个子集, 表示从全集 N 中移除特定特征 i 后可能形成的所有特征组合。它代表了所有不包括特征 i 的特征组合的集合。 $|S|$ 和 $|N|$ 分别表示集合 S 和 N 中的特征数量。 $f(S)$ 代表仅考虑特征集 S 时模型的预测输出, 而 $f(S \cup \{i\})$ 是在集合 S 的基础上加入特征 i 后的模型预测输出。该公式中的差异项 $f(S \cup \{i\}) - f(S)$ 量化了在特定的特征组合 S 下, 引入特征 i 对模型预测结果的边际贡献。系数 $\frac{|S|!(|N|-|S|-1)!}{|N|!}$ 则确保了在 SHAP 值的计算过程

中, 不同大小的特征组合 S 对于特征 i 的 SHAP 值贡献得到适当的加权, 从而使得计算出的 ϕ_i 能够准确反映该特征对模型预测结果的平均贡献。

SHAP 值提供了一种量化每个特征对模型预测影响的方式, 考虑了所有可能的特征交互效应。通过分析特征的 SHAP 值, 可以了解哪些特征对模型预测最为重要, 以及它们是如何影响预测结果的。此外, 利率本身的变化是根据经济状况调整的, 本文研究的重点在于影响央行货币政策调控的因素, 而不是这些因素导致利率升高或降低。基于这个原因, 本文将侧重于研究 SHAP 值的绝对值 $|\phi_i|$, 以量化每个特征对央行货币政策利率调控的影响程度, 而忽略这些影响的方向。

为了更直观地对比展示各个变量对央行货币政策调控的影响程度, 本文还借鉴了肖争艳的做法, 运用 VAR 模型方差分解的思想, 计算每个特征在整个时间序列上的 SHAP 绝对值之和, 然后确定每个特征的 SHAP 绝对值之和在所有特征 SHAP 绝对值总和中的比例, 称为贡献率, 以此来评估每个特征在货币政策利率预测中的相对重要性[16]。公式如下:

$$ContributionRate_i = \frac{\sum_t |\phi_i^t|}{\sum_{i \in I} \sum_t |\phi_i^t|} \quad (2)$$

对于每个特征 i , 其在整个时间序列上的 SHAP 绝对值之和可以表示为 $\sum_t |\phi_i^t|$, 其中 $|\phi_i^t|$ 是特征 i 在时间点 t 的 SHAP 绝对值。所有特征在整个时间序列上的 SHAP 绝对值总和可以表示为 $\sum_{i \in I} \sum_t |\phi_i^t|$, 这里 I 代表所有特征集合。采用基于 SHAP 绝对值的总和贡献率分析, 本文旨在提供一个更加全面和客观的视角, 以理解经济变量如何影响货币政策调控。

4. 指标选择与描述性统计

4.1. 货币政策利率的选择与设定

要想研究央行货币政策调控的影响因素, 首先需要选择一个能准确反应央行货币政策意图的央行货币政策利率。经过近 30 年来持续推进利率市场化改革, 目前我国已基本形成了市场化的利率形成和传导机制, 以及较为完整的市场化利率体系。央行主要通过调控多种货币政策工具的利率, 进而调节金融机构流动性, 向市场释放利率调控的信号, 并进而通过利率走廊的辅助, 引导市场基准利率以政策利率为中枢运行, 最后由各个金融机构基于市场基准利率设定贷款和存款利率报价, 完成整个利率的传导。

目前在所有央行的货币政策工具中, 最常用且能反映央行货币政策调控意图的是每日开展的央行逆回购操作和每月开展的中期借贷便利(MLF)操作[17]。因此, 本文选择频率较高的央行 7 天逆回购利率作为央行货币政策利率基准。但是, 由于央行 7 天逆回购操作在 2013 年之前的使用次数较少, 因此本文参考现有研究, 用央行票据发行利率作为 2013 年之前反映央行货币政策的利率[18]。由于这两个利率之间存在一定的利差, 因此本文以 2013 年为时间节点, 将 2013 年及之前的央行 3 个月期票据发行利率整体平移, 使得其调整后的数值在 2013 年与央行 7 天逆回购利率一致, 从而构建一个完整的 2003~2022 年央行货币政策利率时间序列日度数据。最后将日度数据的均值作为月度数据, 构建月度时间序列, 样本区

间为 2003 年 1 月至 2022 年 12 月, 如图 1 所示, 利率数据均来自 Wind 数据库。

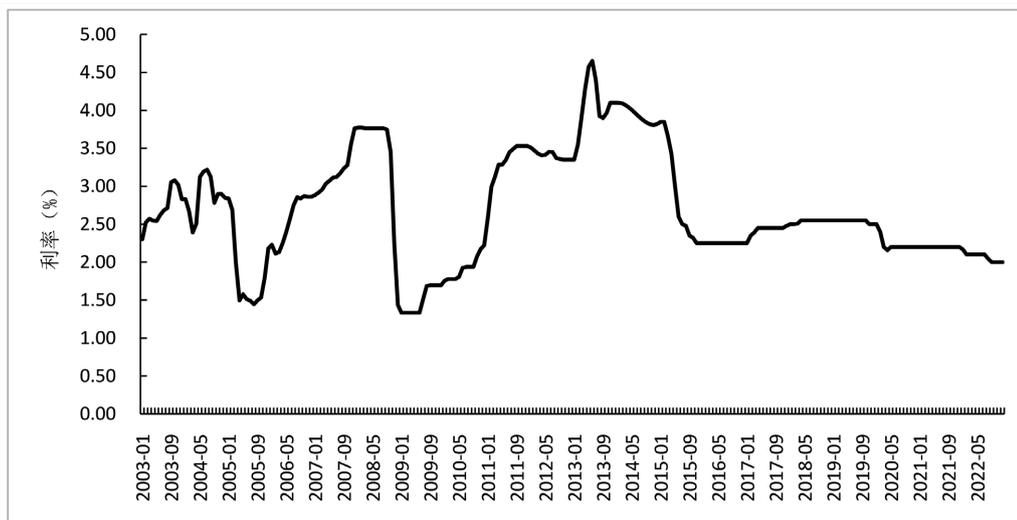


Figure 1. Trends in central bank monetary policy rates from 2003 to 2022

图 1. 2003~2022 年央行货币政策利率走势

4.2. 变量选取与数据说明

首先, 本文选取中国人民银行每季度公布一次的《货币政策执行报告》作为文本分析对象, 该报告是人民银行用于向社会各界介绍货币信贷情况、分析货币政策效果、回顾金融市场情况、阐述宏观经济走势、点明货币政策方向的重要书面沟通手段[19]。第一期报告(2001 年第一季度报告)于 2001 年 5 月公布于人民银行官方网站上, 此后每隔 3 个月发布新的报告。本文试图通过研究央行自身发布的报告内容, 探究可能对央行货币政策调控产生影响的因素, 使用的报告从 2003 年第一季度起, 至 2022 年第四季度止, 共 80 篇文本。针对得到的 jieba 分词结果, 本文删除了仅出现过一次的词语后选取词频统计排名前 5% 的词语(词频数在 316 次及以上)进行研究, 最后在删除无实际意义的形容词等词语后, 共得到包含 104 个词语的央行货币政策执行报告高频词数据。

除此之外, 本文尽可能的考虑了文献综述中所涉及各个变量, 通过结合央行货币政策执行报告高频词统计结果, 归纳整理了七大类别共 14 个变量, 七大类别分别是标准泰勒规则、资产价格、汇率、商品市场、货币、财政和金融稳定指标, 具体变量情况见表 1:

Table 1. Variable details

表 1. 变量情况

变量类别	变量名称	变量符号	货币政策报告中相关高频词汇
标准泰勒规则	通胀缺口	inflation_gap	通货膨胀、居民消费价格等
	伪产出缺口	GDP_gap	供给、需求、生产、产能等
资产价格	股票价格	stock_index	股市、股票、股票市场等
	房地产	house_price-GDP	房地产、房屋、住房等
汇率	美元兑人民币	USD_CNY	美元、汇率、美联储等
	欧元兑人民币	EUR_CNY	欧元、汇率、外汇市场等
	中美 10 年期国债利差	US_CN_10yr	国债、外债、债券市场等

续表

商品市场	原油价格	oil_price	物价、居民消费价格等
	食品价格	CPI_food	食品、消费、物价等
货币	货币供应量 M2	M2_growth	货币供应量、货币市场等
	社会融资规模	social_financing	融资、社会融资、人民币贷款等
财政	国际收支(经常账户余额)	import_export	进出口、贸易、国际收支等
	预计财政赤字	fiscal_deficit	财政、国债、杠杆等
政策稳定指标	经济政策不确定性指数	uncertainty_index	国民经济、改革等

上述 14 个变量构成了本文的基准指标体系, 这些变量的数据主要来源于国家统计局、中国人民银行和国务院等的官方网站, 以及 Wind, Choice 和中经网数据库。为了便于开展实证研究, 本文对数据进行了几方面预处理。第一, 对于一些变量在某些个别月份数据缺失的问题, 本文采用插值法对其进行填补。第二, 对于 GDP 等按季度进行公布的数据, 本文使用三次样条插值法将季度数据转化为月度数据。第三, 对于一些高频的日度数据, 本文采用均值法将其转化为月度数据。第四, 对于目标通胀率和 GDP 增速控制目标等按年度公布的数据, 本文采用复制法将其带入月度数据的计算中。第五, 由于大多数变量为同比增速数据, 为减少绝对值数据较大的波动对 SHAP 值计算带来偏差, 本文将美元兑人民币, 欧元兑人民币和石油价格三个变量设定为环比增速。第六, 由于本文研究的是央行货币政策调控影响因素, 而央行调控基于的是过去数据, 因此最终加入机器学习的数据是 14 个变量各自的滞后一期值。最终, 各变量的描述性统计见表 2:

Table 2. Descriptive statistics of variables
表 2. 各变量描述性统计

变量名称	选取数据	最小值	最大值	平均值	标准差	样本数
通胀缺口(%)	季度通胀率与目标通胀率之差	-5.81	3.94	-0.778	1.77	240
伪产出缺口(%)	GDP 当期同比增速与 GDP 增速控制目标之差	-13.4	12.3	1.23	2.86	240
股票价格	上证综合指数	1042.18	5824.12	2678	877.17	240
房地产(%)	商品房销售价格当月同比增速与 GDP 当期同比增速之差	-16.4	11.04	-0.83	5.45	240
美元兑人民币(%)	美元兑人民币中间价: 环比增速	-2.51	4.38	-0.07	0.81	240
欧元兑人民币(%)	欧元兑人民币中间价: 环比增速	-6.24	6.57	-0.03	2.09	240
中美 10 年期国债利差(%)	中国 10 年期国债到期收益率与美国 10 年期国债到期收益率之差	-2.1	2.43	0.58	1.06	240
原油价格(%)	OPEC 原油价格: 环比增速	-28.59	47.2	1.37	9.29	240
食品价格(%)	CPI_食品: 同比增速	-5.2	23.25	5.35	5.67	240
货币供应量 M2 (%)	M2: 同比增速	8	29.74	14.43	4.82	240
社会融资规模(%)	社会融资增量: 同比增速	-301.66	817.37	30.55	89.2	240
国际收支(经常账户余额)(%)	进出口金额: 同比增速	-29	67.9	13.7	17.46	240
预计财政赤字(%)	财政预算赤字: 同比增速	-923.33	1951.57	11.66	208.25	240
经济政策不确定性指数	中国经济政策不确定性指数	23.72	661.83	155.81	118.13	240

表 3 是基准指标体系中 14 个变量的 VIF 值, 从结果来看, 仅通胀缺口滞后一期和食品价格滞后一期的数值表明存在一定的共线性。但需要说明的是, 随机森林在构建每棵树时只考虑特征的一个随机子集, 这减少了模型对高共线性特征组合的依赖。其本身的设计使其对数据中的各种不规则性, 包括多重共线性、异常值等具有很好的鲁棒性。即使在存在一定程度的共线性, 随机森林也能提供相对稳定和准确的预测。

Table 3. VIF values of variables
表 3. 各变量 VIF 值

变量名称	VIF 值	变量名称	VIF 值
inflation_gap_lag1	7.27	oil_price_lag1	1.16
GDP_gap_lag1	2.59	CPI_food_lag1	5.43
stock_index_lag1	1.61	M2_growth_lag1	2.97
house_price-GDP_lag1	1.19	social_financing_lag1	1.23
USD_CNY_lag1	1.23	import_export_lag1	2.75
EUR_CNY_lag1	1.19	fiscal_deficit_lag1	1.08
US_CN_10yr_lag1	2	uncertainty_index_lag1	2.26

5. 实证结果分析

5.1. 随机森林模型实证结果

在进行实证分析之前, 首先需要评估随机森林模型的预测性能。参考相关文献的常见做法, 本文选取均方误差(MSE)、平均绝对百分比误差(MAE)和决定系数(R^2)作为模型预测性能的评估指标。运用上文计算的最优参数, 最后评估指标的结果如下: 均方误差(MSE)为 0.117, 表明模型的预测误差较小, 即模型的预测准确性较高; 平均绝对百分比误差(MAE)为 8.168%, 表明模型预测的平均误差相对较小, 进一步验证了模型的预测准确性; 决定系数(R^2)为 0.704, 表明该随机森林模型在解释因变量方差方面表现良好, 具有较高的解释能力和预测准确性。

接下来本文将基于模型结果从两方面阐述各类因素对央行货币政策调控的影响。一是根据 SHAP 绝对值之和的大小, 直接对各类因素进行重要性排序。SHAP 绝对值之和最大的因素排在第 1 位, 表示该因素对货币政策的影响是最大的, 然后依次类推, 得到所有影响因素的相对重要性排序。二是基于所计算的 SHAP 绝对值, 通过式(2)进一步计算得到各影响因素对央行货币政策调控的贡献率, 从而更加直观地得到各种因素的影响。最终结果如表 4 所示。

在 2003~2022 年这段时期中, 对央行货币政策调控影响最大的两个因素分别是汇率类别下的中美 10 年期国债利差和标准泰勒规则类别下的通胀缺口, 两者贡献率差距较小, 分别为 24.19%和 23.31%。其中, 中美十年期国债利差这一影响因素与单强的研究相吻合, 他认为考虑到 2008 年国际金融危机后, 美国等发达经济体非常规货币政策的溢出效应明显提高, 有必要考虑其影响, 从图 2 的中美 10 年期国债利差 SHAP 值走势来看, 2008 年后的几年中, 其 SHAP 值显著高于其他时期, 从侧面印证了这一观点。除此之外, 本文认为还有以下两点原因: 第一, 中美十年期国债利率差异不只是简单的经济指标, 它在全球资本市场中起到了“风向标”的作用, 影响着跨境资本流动和汇率政策的制定。中美作为世界上最大的两个经济体, 其十年期国债利差显著影响着全球资本的流动方向。具体而言, 当美国的长期利率相对于中国上升时, 投资者会更偏好美元资产, 导致资本流向美国, 从而对人民币汇率构成贬值压力。在这种情况下, 中国人民银行在制定货币政策和汇率政策时必须密切关注中美十年期国债利率之间的差异, 同

时, 央行对这一指标的关注也显示了其在制定货币政策时考虑的全球视角和对外部变化的敏感度。第二, 中美十年期国债利差对中国吸引外资和利用外资具有较大的影响。制造业作为中国经济的重要支柱, 外资在其中扮演了关键角色。中美国债利差的变动直接影响外国投资者的资本成本和投资回报预期。当美国国债收益率不断上升, 即利差减小时, 资本倾向于流入美国, 寻求更安全的资产回报, 从而减少了流入中国的制造业投资。相反, 只有保持利差在稳定区间时, 中国更快的经济发展速度才能吸引更多的制造业外资进入。因此, 中国央行需要密切监控中美国债利差, 以调整国内利率政策, 保持中国产业对外资的吸引力。

标准泰勒规则下的通胀缺口对央行货币政策的调控影响程度排名第二, 表明中国人民银行与世界上大多数央行一样, 依然以泰勒规则作为实行货币政策时的参考。本文认为这一结果可以从以下三个方面进行解释: 第一, 通货膨胀与我国中央银行货币政策多目标制中的“保持物价稳定”有着较强的联系, 结合通胀中的关键因素食品价格在贡献率中排名第 5, 表明我国央行对物价水平高度重视, 将价格稳定视为其货币政策主要目标之一。第二, 通胀缺口也是一个经济信号, 反映了经济活动的热度。通胀率高于目标并持续上升表明经济有过热的风险, 而通胀率持续低于目标则表明经济中的需求疲软。如图 2 所示, 在 2008 年金融危机和 2020 年疫情期间, 由于经济普遍出现问题, 此时通胀缺口的 SHAP 值显著提高, 表明这段时期之内通胀缺口对央行货币政策调控的影响显著增强, 从一定程度上证明了此观点。第三, 通胀缺口不仅反映了当前的价格水平, 更重要的是, 它在引导市场主体通胀预期中扮演了核心角色。当市场参与者观察到一个持续的正通胀缺口时, 他们会形成未来通胀持续上升的预期。这种预期不是孤立存在的, 它会通过工资设定、长期合同定价以及投资决策等渠道影响经济行为, 从而形成预期的“自我实现”。因此, 监控和管理通胀缺口对央行而言至关重要。通过透明的政策沟通和明确的通胀目标, 央行可以有效地引导市场预期, 避免通胀预期脱锚。这种积极的预期管理不仅有助于维持价格稳定, 还能够增强货币政策的前瞻性和适应性, 促进经济的持续健康发展。

Table 4. Ranking and contribution rates of variables affecting central bank monetary policy management

表 4. 各个变量对央行货币政策调控的影响程度排序及贡献率

变量类别	变量名称	SHAP 绝对值之和排名	贡献率
汇率	中美 10 年期国债利差	1	24.19%
标准泰勒规则	通胀缺口	2	23.31%
政策稳定指标	经济政策不确定性指数	3	14.01%
资产价格	股票价格	4	9.38%
商品市场	食品价格	5	9.35%
货币	货币供应量 M2	6	8.73%
标准泰勒规则	伪产出缺口	7	3.53%
资产价格	房地产	8	1.86%
汇率	美元兑人民币	9	1.65%
财政	国际收支(经常账户余额)	10	1.51%
商品市场	原油价格	11	0.68%
汇率	欧元兑人民币	12	0.65%
货币	社会融资规模	13	0.64%
财政	预计财政赤字	14	0.51%

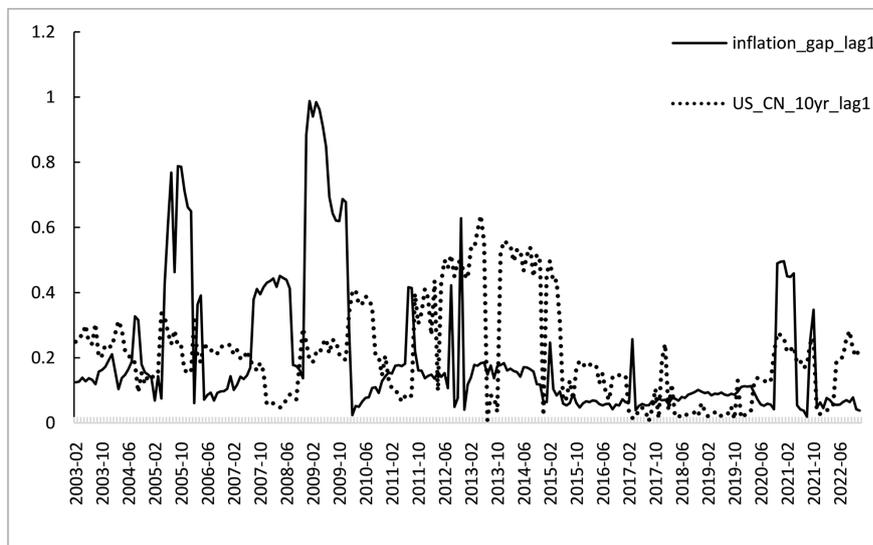


Figure 2. Absolute SHAP values of inflation gap and US-CN 10-year government bond yield spread from 2003 to 2022

图 2. 2003~2022 年通胀缺口和中美 10 年期国债利差 SHAP 绝对值

第二梯队(3~6 名)变量中, 经济政策不确定性指数对央行货币政策调控的影响贡献率达到 14.01%。这与 Jonas and Johannes 的研究结果相符, 他们在研究欧洲央行货币政策的影响因素中加入了欧洲经济政策不确定性指数, 并且统计效果显著。再结合邓伟等的研究结果, 本文认为经济政策的不确定性的增强导致商业银行资产流动性囤积效应, 从而影响实体经济融资和货币政策的有效传导, 最终对央行货币政策的调控产生影响[20]。如图 3 所示, 从时间上来看, 随着 2018 年中美贸易战和近年来中国经济增速放缓, 经济政策不确定性对央行货币政策调控的影响程度在不断提高, 表明经济政策在央行所参考的指标中的地位越来越高。

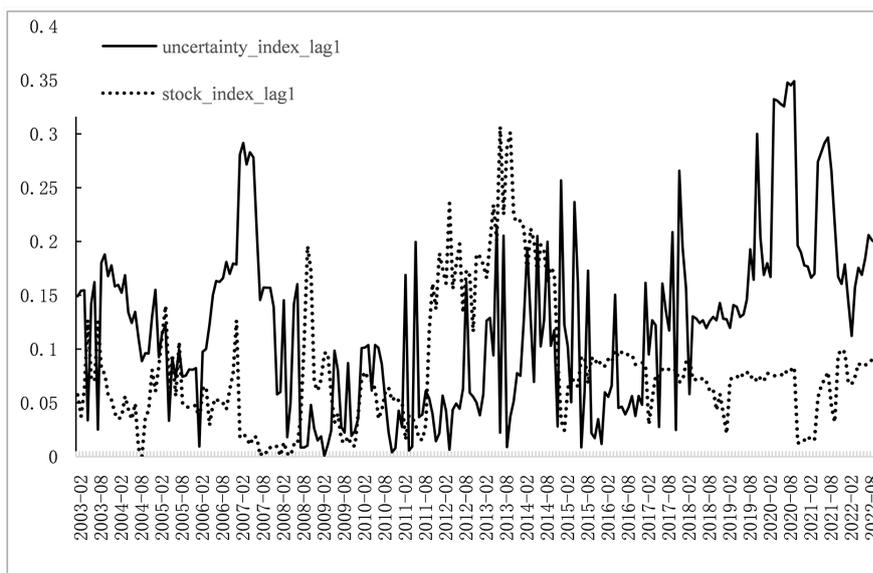


Figure 3. Absolute SHAP values of economic policy uncertainty index and stock prices from 2003 to 2022

图 3. 2003~2022 年经济政策不确定性指数和股票价格 SHAP 绝对值

以上证指数衡量的股票价格对央行货币政策调控的影响排名第四, 贡献率为 9.38%。通常认为央行货币政策与股票市场联系紧密, 宽松的货币政策有利于股市的上涨。这与王少林等的研究结果相似, 表明我国货币政策与股票市场之间存在非对称的互动关系[21]。

货币供应量 M2 对于央行货币政策调控的影响排名第六, 贡献率为 8.73%, 且根据图 4 来看, SHAP 值的分布较为均匀, 表明我国央行货币政策依然处在由数量规则向利率规则的转型时期。对于剩余的其他变量, 其对于央行货币政策调控的影响贡献度较低, 表明影响程度较小, 本文不再进行分析。

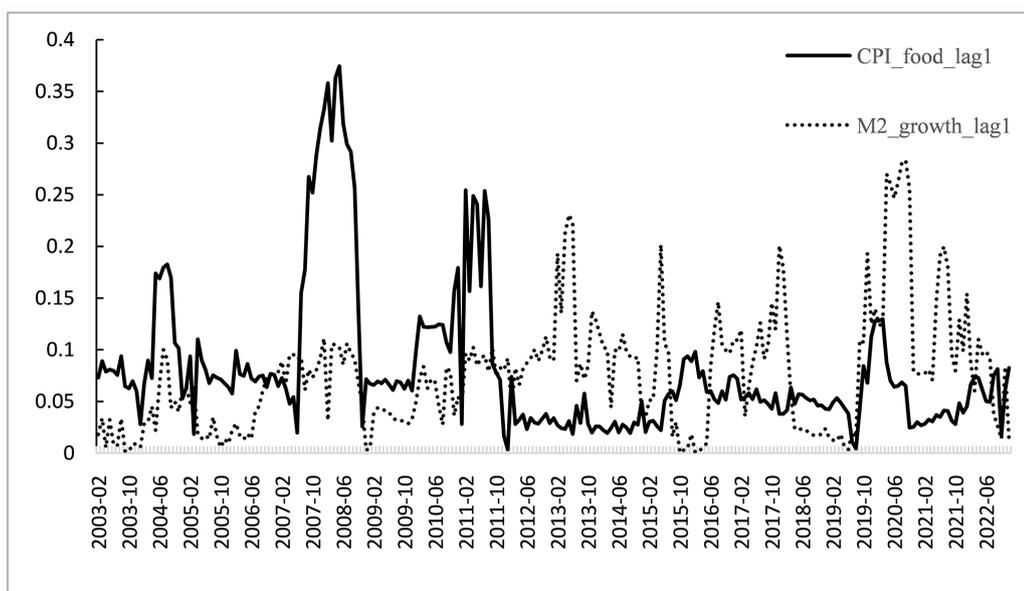


Figure 4. Absolute SHAP values of food price and M2 money supply from 2003 to 2022

图 4. 2003~2022 年食品价格和货币供应量 M2 SHAP 绝对值

5.2. 稳健性检验

为对实验结果进行验证, 本文参考王达和周映雪的方法, 建立梯度提升树模型对随机森林模型的 SHAP 绝对值计算结果进行稳健性检验。梯度提升树(Gradient Boosting Trees, GBT)模型是一种高效的预测工具, 它通过迭代地训练决策树来最小化损失函数。相对于随机森林模型的并行树构建方法, 梯度提升树采用的是顺序建立树的策略, 每棵树学习到的是之前所有树预测结果的残差。这种策略使得梯度提升树在处理各种回归和分类问题时具有出色的准确性, 尤其是在复杂的非线性数据集上。

本文利用时间序列交叉验证和随机搜索的方法对梯度提升模型的超参数进行了优化。最终确定的关键参数如下: 模型由 463 棵树构成, 每棵树的深度为 3。模型的学习率设置为 0.1396, 以确保足够的学习速度而又不过度拟合。为了避免模型过度关注任何单一样本, 内部节点再划分所需的最小样本数被设为 4, 叶节点的最小样本数被设为 10, 从而允许树在形成决策时考虑足够的局部样本信息。在该参数下的模型的评估指标结果如下: 均方误差(MSE)为 0.156, 平均绝对百分比误差(MAE)为 10.41%, 决定系数(R^2)为 0.61。

梯度提升树模型的结果如表 5 所示, 结果显示中美 10 年期国债利差和通胀缺口对央行货币政策调控的影响程度依然排在第一和第二位。而第 3~7 位和第 8~14 位的变量由于相互之间贡献率差距较小, 因此在排名上有变化, 但各个梯度位次分布情况与随机森林模型下的分布差距不大, 表明随机森林模型结果具有一定的稳健性, 本文的主要结论依然成立。

Table 5. Ranking and contribution rates of variables in the GBT model**表 5.** 梯度提升树模型中各个变量的影响程度排序及贡献率

变量名称	随机森林模型贡献率排名	梯度提升树模型贡献率排名
中美 10 年期国债利差	1	1
通胀缺口	2	2
经济政策不确定性指数	3	4
股票价格	4	6
食品价格	5	7
货币供应量 M2	6	3
伪产出缺口	7	5
房地产	8	11
美元兑人民币	9	14
国际收支(经常账户余额)	10	8
原油价格	11	12
欧元兑人民币	12	9
社会融资规模	13	13
预计财政赤字	14	10

6. 结论与政策启示

本文通过文本分析方法分析了央行货币政策执行报告,并结合既有文献设定了可能会对央行货币政策调控产生影响的七大类 14 个变量,并通过随机森林模型和 SHAP 值法分析了这些变量对 2003~2022 年央行货币政策调控影响程度及其动态变化。研究结果表明:第一,随机森林模型在分析和预测央行货币政策调控的影响因素方面效果显著,并通过与 SHAP 值法相结合,能够比较清晰地量化不同变量对货币政策调控影响的贡献度,为我国货币政策研究提供了新的视角。第二,2003~2022 年的二十年间,以通胀缺口为主的泰勒规则和以中美 10 年期国债利差为代表的汇率因素在我国央行货币政策调控的参考指标中占据重要地位。

基于以上结论,本文提出以下三点政策建议:第一,重视通胀在宏观经济调控中的作用,加强通胀预期管理。政府应进一步加强对通胀目标的沟通,提高透明度,从而有效地引导市场预期。央行应该继续通过定期发布货币政策执行报告、政策展望以及对通胀风险的评估,帮助市场主体形成合理的通胀预期,降低货币政策调控中的不确定性。第二,强化对国际金融环境的重视,并加强对外汇市场的关注程度。央行应进一步加强对全球金融动态的监测与分析,及时识别和评估国际金融市场特别是汇率市场变化对国内经济可能产生的影响,从而有效地制定应对措施,以此帮助缓解外部冲击,保持货币政策的连续性。第三,深化利率市场化改革,推进价格型货币政策调控。利率市场化改革能够提高政策利率传导机制的有效性,使货币政策调控更加灵活、及时,同时也有助于提升央行货币政策执行效率,更好地应对经济波动。

基金项目

江苏省社会科学基金青年项目(20EYC011);中国博士后科学基金面上资助(2021M691635)。

参考文献

- [1] 易纲. 货币政策回顾与展望[J]. 中国金融, 2018(3): 9-11.

- [2] 李宏瑾, 苏乃芳. 数量规则还是利率规则?——我国转型时期量价混合型货币规则的理论基础[J]. 金融研究, 2020(10): 38-54.
- [3] 谢平, 罗雄. 泰勒规则及其在中国货币政策中的检验[J]. 经济研究, 2002(3): 3-12, 92.
- [4] 郑挺国, 王霞. 泰勒规则的实时分析及其在我国货币政策中的适用性[J]. 金融研究, 2011(8): 31-46.
- [5] 陈创练, 郑挺国, 姚树洁. 时变参数泰勒规则及央行货币政策取向研究[J]. 经济研究, 2016, 51(8): 43-56.
- [6] Tiwari, K.A., Abakah, A.J.E., Abdullah, M., *et al.* (2024) Time-Varying Relationship between International Monetary Policy and Energy Markets. *Energy Economics*, **131**, Article ID: 107339. <https://doi.org/10.1016/j.eneco.2024.107339>
- [7] Chevapatrakul, T., Kim, T. and Mizen, P. (2012) Monetary Information and Monetary Policy Decisions: Evidence from the Euroarea and the UK. *Journal of Macroeconomics*, **34**, 326-341. <https://doi.org/10.1016/j.jmacro.2012.01.002>
- [8] Jurkšas, L. (2024) The Impact of the Heterogeneous Fiscal Policy Stance of Euro-Area Member States on ECB Monetary Policy. *Economic Systems*. <https://doi.org/10.1016/j.ecosys.2024.101216>
- [9] Jonas, G. and Johannes, Z. (2021) What's on the ECB's Mind? Monetary Policy before and after the Global Financial Crisis. *Journal of Macroeconomics*, **68**, Article ID: 103292. <https://doi.org/10.1016/j.jmacro.2021.103292>
- [10] 单强, 吕进中, 王伟斌, 等. 中国化泰勒规则的构建与规则利率的估算——基于考虑金融周期信息的潜在产出与自然利率的再估算[J]. 金融研究, 2020(9): 20-39.
- [11] Jianqing, F., Fang, H. and Han, L. (2014) Challenges of Big Data Analysis. *National Science Review*, **1**, 293-314. <https://doi.org/10.1093/nsr/nwt032>
- [12] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. <https://doi.org/10.1023/A:1010933404324>
- [13] 王达, 周映雪. 随机森林模型在宏观审慎监管中的应用——基于 18 个国家数据的实证研究[J]. 国际金融研究, 2020(11): 45-54.
- [14] Lundberg, S.M. and Lee, S.I. (2017) A Unified Approach to Interpreting Model Predictions. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, 4-9 December 2017, 4768-4777.
- [15] Joseph, A. (2019) Shapley Regressions: A Framework for Statistical Inference on Machine Learning Models. Working paper No. 2019/7.
- [16] 肖争艳, 陈衍, 陈小亮, 等. 通货膨胀影响因素识别——基于机器学习方法的再检验[J]. 统计研究, 2022, 39(6): 132-147.
- [17] 易纲. 中国的利率体系与利率市场化改革[J]. 金融研究, 2021(9): 1-11.
- [18] 钱雪松, 杜立, 马文涛. 中国货币政策利率传导有效性研究: 中介效应和体制内外差异[J]. 管理世界, 2015(11): 11-28, 187.
- [19] 姜富伟, 胡逸驰, 黄楠. 央行货币政策报告文本信息、宏观经济与股票市场[J]. 金融研究, 2021(6): 95-113.
- [20] 邓伟, 宋清华, 杨名. 经济政策不确定性与商业银行资产避险[J]. 经济学(季刊), 2022, 22(1): 217-236.
- [21] 王少林, 林建浩, 杨燊荣. 中国货币政策与股票市场互动关系的测算——基于 FAVAR-BL 方法的分析[J]. 国际金融研究, 2015(5): 15-25.