文章编号：1002–1566(2023)01–0109–18
DOI: 10.13860/j.cnki.sltj.20211130–018

中国居民消费价格指数的动态结构研究及中美量化比较

王振中 [1] 陈松蹊 [2,3] 涂云东 [2,3]

(1. 爱荷华州立大学，爱荷华 50011，美国；2. 北京大学光华管理学院，北京 100871；
3. 北京大学统计科学中心，北京 100871)

摘要：本文提供了一个对中国居民消费价格指数 (CPI) 的系统分析，包括 CPI 的动态结构和可预测性，以及中美 CPI 的量化比较。尽管中美 CPI 的构成和动态结构有所不同，但二者都可由 S–ARIMAX 模型很好地刻画，此模型能够准确量化中国 CPI 的春节效应以及其他突发事件对中美 CPI 的影响。我们发现中国 CPI 有显著的季节性和春节效应，且比美国 CPI 有更强的可预测性。此外，中国 CPI 的短期预测可以通过扩散指数 (Diffusion Index) 方法得到进一步提高。

关键词：S–ARIMAX 模型；春节效应；季节性；预测；扩散指数模型
中图分类号：F222, O212 **文献标识码**：A

Analyzing Chinese Consumer Price Index Comparatively with That of United States

WANG Zhen-zhong[1] CHEN Song-xi[2,3] TU Yun-dong[2,3]

(1. Iowa State University, Iowa 50011, USA; 2. Guanghua School of Management, Peking University, Beijing 100871, China; 3. Center for Statistical Science, Peking University, Beijing 100871, China)

Abstract: This paper provides a thorough analysis on the dynamic structures and predictability of Chinese Consumer Price Index (CPI), with a comparison to those of the United States. Despite the differences in the two leading economies, both series can be well modeled by a class of S-ARIMAX models, which can well capture the Spring Festival effects on Chinese CPI and the influences of one-off events on the CPI series. The Chinese CPI series possess regular patterns of dynamics with stable annual cycles and strong Spring Festival effects, with fitting and forecasting errors largely comparable to their US counterparts. Finally, for Chinese CPI, the Diffusion Index (DI) approach offers improved predictions than the S-ARIMAX models.

Key words: S-ARIMAX model; spring festival effect; seasonality; forecasting; diffusion index

0 引言

自 2011 年以来，中国成为了世界第二大经济体，并且日益成为全球经济增长的引擎。随着中国在世界经济舞台上声望和影响力的与日俱增，研究中国宏观经济变量的需求变得更加紧迫。在众多宏观变量中，居民消费价格指数 (CPI) 是一个衡量总体消费价格水平变动的关

收稿日期: 2020 年 4 月 17 日 **收修改稿日期**: 2021 年 3 月 24 日 **通讯作者**: 涂云东，yundong.tu@gsm.pku.edu.cn
基金项目: 自然科学基金项目 (92046021, 12071013, 12026607, 71671002, 71973005, 72073002)；北京大学统计科学中心和数量经济与数理金融教育部重点实验室 (LMEQF)。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

键指标，也是国家宏观经济决策和国民经济核算的重要指标。对于国民经济的宏观调控来说，要想在追求其他经济目标的同时维持价格稳定，CPI 的高低尤为重要。它直接影响了宏观调控措施的出台和力度，例如央行如何调整利息和存款准备金率等。

尽管中国 CPI 十分重要，然而我们对其动态变化和结构的研究却有限，尤其缺乏对 CPI 定基数据的研究。Chow (1987)[1] 代表了对中国物价指数的先驱研究。其他研究请见何启志和范从来 (2001)[2]，栾惠德 (2007)[3]，雷怀英和金勇进 (2008)[4]，王少平和王津港 (2009)[5]，魏璐和钱存华 (2014)[6]，张成思 (2016)[7] 以及 Wang 等 (2016)[8] 等。这种研究不足的一个表现是：由于缺乏对 CPI 定基值的季节性和动态结构的了解，官方公布的 CPI 是没有经过季节调整的。这意味着 CPI 数据可能呈现出诸如季节变动、移动节日效应的特点，从而使得 CPI 的同比和环比有时无法体现真正的价格变动趋势。因此，在对中国 CPI 分析时如果没有考虑这些特征，容易得出有误导性的解读。例如，2015 年 1 月中国 CPI 的同比值降到了 0.8%，这引起了各界对未来通货紧缩的广泛忧虑，因为此月同比值比上一个月降低了几乎 50% (2014 年 12 月的同比是 1.5%)。然而在考察春节时间后会发现，春节的日期在 2014 年是 1 月 31 日而在 2015 年是 2 月 18 日。这次的剧降主要来自春节效应，并不能说明未来有通货紧缩的可能。后文的模型式 (15) 可以量化出此次春节效应的大小，并给出去除春节效应后的同比值。可见如若不能正确量化春节效应，对未来 CPI 的走势可能会出现错误估计和解读。

国际上有学者担心地方政府可能会人为干预中国宏观数据的核算，这是由于对中国宏观数据缺乏了解。同衡量各级政府业绩的关键指标 — GDP 不一样，CPI 从未成为政府的业绩指标。因此，地方政府几乎没有动机去干预 CPI 数据。另外，用于核算 CPI 的消费价格数据是由统计局直接下属的城市调查团队收集的，整个价格指数编算过程越过地方政府而由国家统计局直接汇总并核算和公布。从这两方面可以判断出中国 CPI 数据的人为干预可能性很小。本文的研究结果将佐证中国 CPI 数据的质量。

本文旨在提供一个对中国 CPI 动态结构和可预测性的完善分析，并与美国 CPI 进行比较来揭示中国 CPI 数据的特征。我们的研究有以下几个具体目标。首先，我们利用带解释变量 (X) 的季节性 ARIMA (简称 S–ARIMAX) 模型分析中国 CPI 的基本动态结构，并对春节效应以及主要历史事件引起的异常值进行了量化。研究表明，中国 CPI 序列具有规律的动态结构和稳定的季节性，以及规律性的春节效应。这些都说明中国 CPI 的数据质量是好的。我们采用的模型拟合效果好，其中较大的数据变异可由春节效应以及触发价格波动的一次性事件 (异常值) 来解释。进一步，在模型中加入春节效应和异常值干预分析后，拟合和预测误差都大大减小。最终，我们在总共 11907 个备选模型 (81 种 SARIMA 阶数 ×147 种春节效应) 中选出拟合和预测效果最好的模型，此模型能很好地捕捉到中国 CPI 序列的动态结构、季节性和春节效应。

我们的第二个目标就是比较中国 CPI 和美国 CPI 在 2002 到 2019 年之间的可预测性。为此，我们也应用了线性时间序列模型对于美国 CPI 数据做了类似的分析。通过比较我们发现美国 CPI 数据的拟合和预测误差在很大程度上和中国 CPI 具有可比的均方预测误差。

我们的第三目标是采用 Stock 和 Watson (2002)[9−10] 提出的扩散指数模型 (Diffusion Index model，简称 DI 模型)，通过结合其它宏观经济变量来改进中国 CPI 的预测。我们的分析表明，对于中国 CPI，当预测步长为 1 到 3 个月时，DI 模型的预测结果要比线性时间序列 (S–ARIMAX) 更精确；而对于美国 CPI，DI 模型的表现不如 S–ARIMAX。

本文的结构如下。第 1 节介绍了中美 CPI 数据以及他们的组成。第 2 节概述了单变量线性时间序列模型 — S–ARIMAX 的建模策略，及其在中国 CPI 上的建模结果。第 3 节将 DI 模

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

型运用于中国 CPI 序列，以得到更高的预测精度。第 4 节报告了美国 CPI 的建模结果，其中包括 S–ARIMAX 模型和 DI 模型。第 5 节阐述主要结论并比较中美 CPI 的异同。文中的附件可向作者索取。

1 CPI 数据

在本文中，中美 CPI 数据的时间跨度是从 2002 年 1 月到 2019 年 12 月。选择这段时间来进行研究是基于以下两点考虑。首先，中国于 2001 年 12 月加入 WTO，从那之后中国的经济活动越来越与世界经济接轨。特别是在 2001 年之后，中国的 CPI 构成与美国 CPI 更加相似。其次，美国的 CPI 核算在 1999 年经历过一次很大的修改。从 1999 年之后，美国劳工统计局 (Bureau of Labor Statistics，简称 BLS) 采用几何平均数来计算单个商品的平均价格改变 (BLS, 2007)[11]。所以，2002 年 1 月是对中美 CPI 进行有意义比较研究的最早时间点。

在本文中，我们采用 CPI–U (CPI for All Urban Consumers) 来代表美国的居民消费物价指数。它也是 BLS 最常用的 CPI 序列。中国公布的 CPI 序列有三种，即居民消费价格指数 (总 CPI)、城市居民消费价格指数 (城市 CPI) 和农村居民消费价格指数 (农村 CPI)。本文我们采用总 CPI 来代表中国的居民消费价格指数。中美 CPI 的一个不同之处在于，美国公布的 CPI 有原始数据和季节调整后的数据，而中国公布的 CPI 并没有进行季节调整，只有原始数据。因此，在本文的分析中，为了使中美数据有可比性，我们都采用原始数据 (未经过季节调整) 来比较分析。

首先，中国 CPI 和美国 CPI 在构成上有一定的相似性，都是由八个子类构成。2016 年之前，中国 CPI 的八个子类包括：食品、烟酒及用品、衣着、家庭设备用品及服务、医疗保健及个人用品、交通及通信、娱乐教育文化用品及服务、居住。从 2016 年 1 月开始，中国 CPI 的八大子类从新划分为：食品烟酒、衣着、居住、生活用品及服务、交通和通信、教育文化和娱乐、医疗保健、其他用品和服务。本文我们用 2016 年前的分类方式来比较中美 CPI 的构成。美国的八个子类包括：食品及饮料、衣着、交通、医疗、娱乐、教育及通信、居住、其他商品及服务。在这八大子类中，中美对衣着和居住子类的定义很接近，在其它子类的组成上略有不同。中国将娱乐、教育和文化用品及服务归为一个子类。而美国将娱乐单列为一类，把教育和通信合为一个子类。中国 CPI 中的“家庭设备用品及服务”在美国 CPI 中被归到“其它商品及服务”中。此外，中国 CPI 的“医疗保健及个人用品”子类中的个人用品部分在美国被归到“其它商品及服务”。

现实生活中，不同商品和服务在居民消费中所占的比重不同，它们的价格变动对 CPI 的影响也就不尽相同。所以必须制定一组权重，以反映居民消费中各个子类的相对重要性。无论是中国还是美国，权重的核算都是来自于居民的消费支出调查。具体来说，美国 CPI 的权重来自于每两年一次的消费者支出调查 (Consumer Expenditure Survey，简称 CES)，中国 CPI 的权重来自于每五年一次的全国城乡居民家庭的消费支出资料。

图 1 展示了中国 CPI (图 1 左) 和美国 CPI (图 1 右) 的八大子类权重。由于中国并没有官方公布的 CPI 权重数据，图 1 (左) 的结果是由 CPI 对其八大子类做线性回归得到的，数据区间是 2011 年 1 月到 2014 年 12 月。为了使中美的权重具有可比性，图 1 (右) 也是由美国 CPI 对其八大子类回归得到。同时考虑到美国 CPI 美每两年调整一次权重，美国 CPI 回归的数据区间是 2013 年 1 月到 2014 年 12 月。此回归所得到权重和 BLS 公布的相对权重 (Relative Importance) 非常接近，由此可见线性回归可以给出对权重一个很好的估计。在中国的 CPI 构

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

成中，食品占了 31% 的比重，居住占比 21%，而其他子类占比均小于 12%。可见中国 CPI 的构成中食品和居住占主导。而对于美国 CPI，居住和交通占主导，权重分别为 43% 和 17%。食品及饮料占比 15%，远低于其在中国 CPI 相应子类的比重。

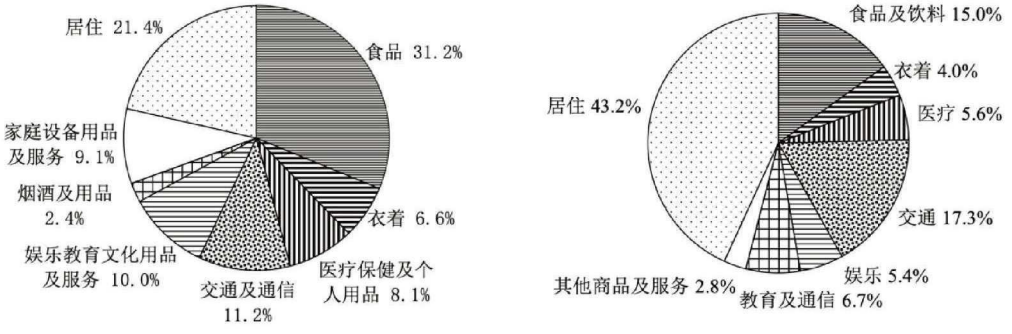

图 1 中国 CPI (左) 和美国 CPI (右) 的构成

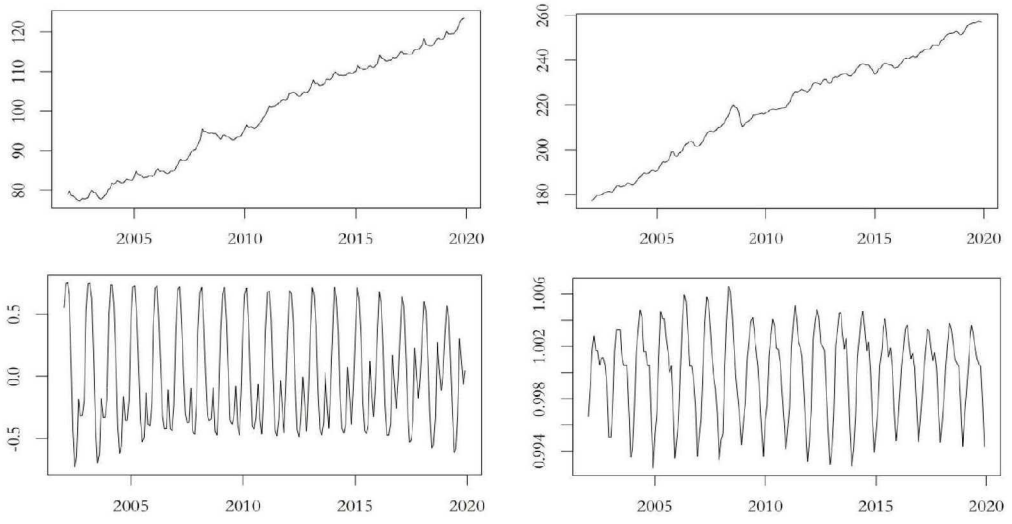

图 2 中国 CPI (左上) 和其季节项 (左下) 以及美国 CPI 定基值 (右上) 和其季节项 (右下)

图 2 (左上) 和 (左下) 分别展示了中美 CPI 的定基值序列。从图中可看出，2008 年的金融危机对美国 CPI 的影响要大于其对中国的影响。这不仅是因为美国是此次金融危机的爆发点，还与中美 CPI 的构成有关。由上面的权重分析可知，美国 CPI 中占主导地位的是居住，而且这次金融危机又是由于美国房地产泡沫破灭引起的。而中国 CPI 中食品所占比重最大，相对于居住，食品的供需受金融危机的影响小很多，价格也就更稳定。

通过使用美国人口普查局 (U.S. Census Bureau) 的季节调整算法 X–13ARIMA–SEATS，并结合下文第三节给出的 S–ARIMAX 模型 (式 (15))，我们对中国 CPI 进行了季节调整。图 2 (左下) 给出了中国 CPI 的季节项。此图展示了中国 CPI 有很强的季节性，每年峰值在 2 月份，7 月份达到谷值。并且此季节性的波动从 2002 年逐年变小，直到 2007 年开始维持在一个新的水平。对于美国 CPI，其季节性 (图 2 右下) 在 5 月左右达到峰值，在 12 月进入谷值。

除了季节性的不同，中美 CPI 的另一个重要区别就是移动节日效应。中国重要的传统节日时间都是基于农历，例如春节、端午和中秋。此类节日的日期在公历中不固定，会在两个月内变动，使得同一月份在不同年份所受的节日效应有所不同。而传统节日对中国的生产和消

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

费活动又有很大影响，特别是春节。具体来说，春节期间居民的消费需求会扩大，而大部分人春节期间放假回家，工厂生产暂停和服务的供给会减少，因此春节期间的物价会有一个明显提升，从而导致 CPI 在春节所在月份出现“异常”波动。

而对于美国 CPI，由于美国节假日时间要么是固定的 (例如圣诞节)，要么只在很小的范围内移动 (如感恩节)，因此美国 CPI 在同一月份不同年份的节日效应基本一致，从而这些节日效应都可被归到季节项中。唯一的例外是复活节，其变动范围是 3 月 22 日到 4 月 25 日之间。而美国劳工统计局 (BLS) 在其官方的季节调整中未考虑复活节效应，这表明复活节对美国 CPI 的影响不显著。我们在第 4.1 节中的分析也证实了这一点。

2 线性时序模型建模

在宏观经济变量的预测研究中，Stock 和 Watson (2002)[9−10] 提出的 DI 模型在很多实际应用中都有良好效果。而对于美国季节调整后的 CPI 序列，MacCracken 和 Ng (2016)[12] 表明就模型拟合和样本外预测而言，传统的线性模型如自回归模型 (Auto–regressive model，简称 AR) 表现效果并不比 DI 模型差。因此，我们首先考虑用传统的线性时间序列模型对中国 CPI 建模，之后再用 DI 模型进行建模，并比较两模型的效果。

2.1 季节性 ARIMA 模型

我们以 $\{Z_t\}_{t=1}^{T}$ 表示中国 CPI 定基值序列，其中 Z_t 是第 t 月的 CPI 的观测值。B 是滞后算子，即 $BZ_t = Z_{t-1}$。由于中国 CPI 有很强的季节性，我们考虑用季节性 Auto-Regressive Integrated Moving Average 模型 (Box 等 2015[13]，简称 S–ARIMA 模型) 进行建模：

$$\phi_p(B)\Phi_P(B^{12})(1-B)^d(1-B^{12})^D Z_t = \theta_q(B)\Theta_Q(B^{12})a_t,$$

其中，$\phi_p(B)$ 和 $\theta_q(B)$ 分别是 B 的 p 阶和 q 阶多项式，分别是模型的自回归项和移动平均项，$\Phi_P(B^{12})$ 和 $\Theta_Q(B^{12})$ 分别是 B^{12} 的 P 阶和 Q 阶多项式，代表模型的季节性自回归项和季节性移动平均项；$\phi_p(\cdot)$、$\theta_q(\cdot)$、$\Phi_P(\cdot)$ 和 $\Theta_Q(\cdot)$ 的根都在单位圆之外。d 表示常规差分的阶数，D 则代表季节性差分的阶数。在绝大多数情况下，d 与 D 之和不会超过 2；$\{a_t\}$ 是期望为 0，方差为 $\sigma_a^2 > 0$ 的白噪声序列。

为了确定模型的阶数 p、q、P、Q、d 和 D，我们先考察 CPI 的样本自相关函数 (sample autocorrelation functions，简称 ACFs) 和偏自相关函数函数 (sample partial autocorrelation functions，简称 SACFs)。图 3 (左上) 和 (右上) 分别给出了 CPI 定基数据的 ACFs 和 PACFs。其中 ACFs 下降缓慢，所以需要对数据进行 1 阶差分。差分之后的 ACFs (图 3 左中) 在 12 的整数阶处显著不为零，而且下降的很缓慢，这表明序列有跨度为 12 的周期性，所以需要再对序列进行 12 阶差分。经过 1 阶和 12 阶差分后 CPI 的 ACFs 变为图 3 (左下) 所示。此时 SACFs 只在 2，10，12 阶处显著不为零，这表明序列已经平稳，不再需要进行差分了。同时，我们也对 1 阶及 12 阶差分后的 CPI 进行了 ADF 平稳性检验：

原假设 H_0：$(1-B)(1-B^{12})Z_t$ 有单位根；备择假设 H_1：$(1-B)(1-B^{12})Z_t$ 是平稳的。

检验的概率 P 值为 0.01，在显著水平为 0.05 的情况下我们拒绝原假设，并接受 $(1-B)(1-B^{12})Z_t$ 是平稳的备择假设。因此我们取 $d = D = 1$：

$$\phi_p(B)\Phi_P(B^{12})(1-B)(1-B^{12})Z_t = \theta_q(B)\Theta_Q(B^{12})a_t. \tag{1}$$

对于 p、q、P 和 Q 的选取，我首先根据 $(1-B)(1-B^{12})Z_t$ 的 SACFs 和 SPACFs 选出一

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

个范围，再在第 2.4 节通过模型选择得出最优模型。图 3 (左下) 刻画了 $(1-B)(1-B^{12})Z_t$ 的 SACFs，其只在 2，10，12 和 23 阶是显著不为零的，这表明 MA 阶数 q 和季节性 MA 阶数 Q 为 1 或者 2。类似的，图 3 (右下) 的 SPACFs 只在 2，10，12 和 24 阶是显著非零的，所以 AR 阶数 p 和季节性 AR 阶数 P 也为 0，1 或者 2。所以我们选定的范围是：$p,q,P,Q\in\{0,1,2\}$，总共 81 种备选阶数，即 81 个备选模型。

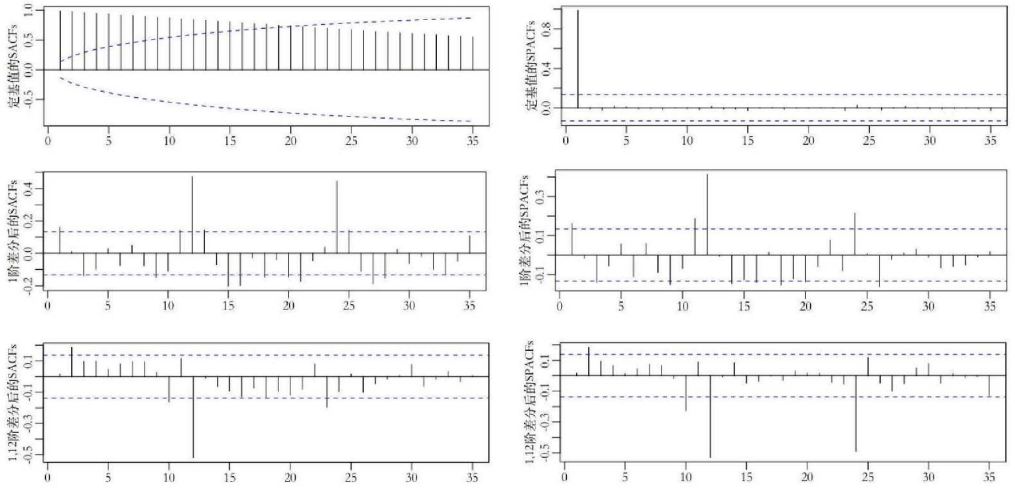

图 3 中国 CPI 定基值及其差分后的的 ACFs 和 PACFs

2.2 异常值检测

在对 81 个备选模型进行选择之前，我们先对 CPI 序列的异常值进行检测。通过异常值检测，我们观察 CPI 序列的异常值是否和历史上的相关重要事件相吻合，从而间接验证中国 CPI 数据的质量。

在本文中，我们采用 Chen 和 Liu (1993)[14] 的方法，对所有 81 个备选模型进行异常值检测。此方法可以检测出四种异常值：可加异常值 (Additive Outlier，简称 AO)；扰动异常值 (Innovative Outlier，简称 IO)；水平位移 (Level Shift，简称 LS)；短暂变化 (Transient Change，简称 TC)。这四种异常值涵盖了绝大多数异常值的情况，在实际中应用十分广泛。首先设 $\Psi(B)=\frac{\theta_q(B)\Theta_Q(B^{12})}{\phi_q(B)\Phi_P(B^{12})(1-B)(1-B^{12})}$，则式 (1) 转化为：

$$Z_t=\Psi(B)a_t.$$

四种异常值分别表示如下

$$\begin{aligned}
\text{AO}: &\quad Z_t=\Psi(B)a_t+\omega P_t^{(t_1)},\\
\text{IO}: &\quad Z_t=\Psi(B)(a_t+\omega P_t^{(t_1)}),\\
\text{LS}: &\quad Z_t=\Psi(B)a_t+\omega S_t^{(t_1)},\\
\text{TC}: &\quad Z_t=\Psi(B)a_t+\frac{\omega}{1-\delta B}P_t^{(t_1)},
\end{aligned}$$

其中 $P_t^{(t_1)}=I(t=t_1)$ 是在 $t=t_1$ 的脉冲函数，I 是示性函数，ω 表示异常值效应的强度。$S_t^{(t_1)}=I(t\geq t_1)$ 是在 t_1 开始的位移函数。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

在扰动 a_t 满足正态分布的前提下，模型式 (1) 中的系数可以由极大似然估计得到，继而我们可以得到 $\Psi(B)$ 的估计值，记作 $\hat{\Psi}(B)$，和模型残差 $\hat{e}_t := \frac{Z_t}{\hat{\Psi}(B)}$。而残差中会体现异常值的影响。具体来说，设 $\hat{\pi}(B) = \frac{1}{\hat{\Psi}(B)}$，则对于每一类异常值我们都有：

$$\hat{e}_t = \omega x_t + a_t,$$

其中，对于 AO 我们有 $x_t = \hat{\pi}(B)P_t^{(t_1)}$；对于 IO 有 $x_t = P_t^{(t_1)}$；对于 LS 有 $x_t = \frac{1}{1-B}\hat{\pi}(B)P_t^{(t_1)}$；对于 TC 有 $x_t = \frac{1}{1-\delta B}\hat{\pi}(B)P_t^{(t_1)}$。Chen 和 Liu (1993)[14] 建议取 $\delta = 0.8$。然后 $\hat{e}_t$ 对 x_t 做回归得到 ω 的估计值，通过对 ω 的显著性作 t– 检验来判断在 t_1 处是否有异常值。

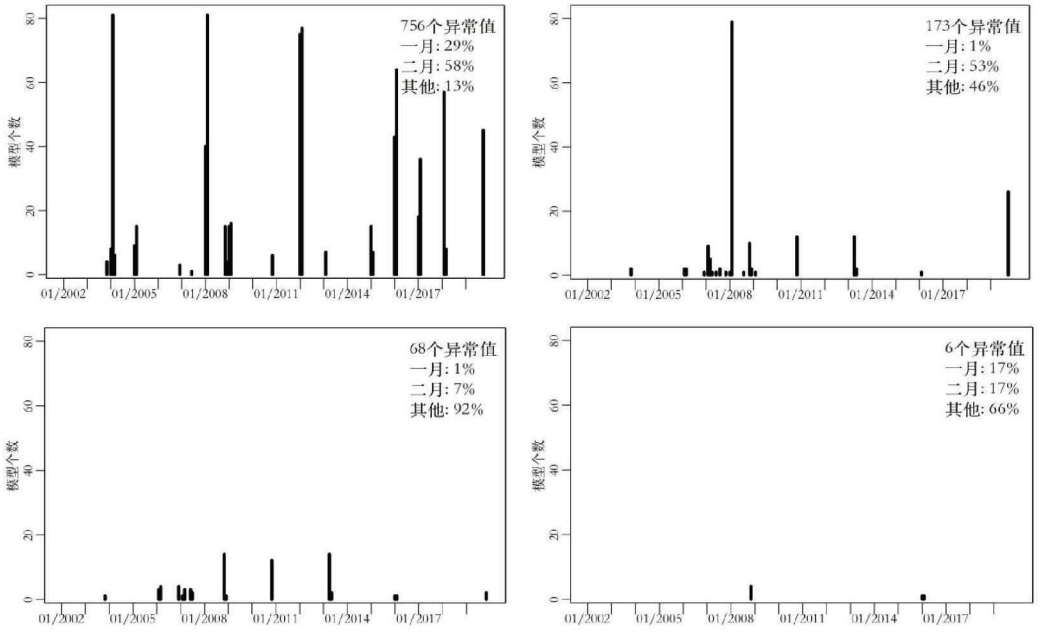

图 4 81 个备选模型在任一给定月份 (2002 年 1 月 – 2019 年 12 月) 检测出异常值的个数：左上未考虑春节效应和 2 异常值 (式 (1))；右上只考虑春节效应 (式 (2))；左下考虑春节效应、2008 年雪灾和 2019 年 10 月猪肉价格上涨 (式 (3))；右下考虑春节效应、2008 年雪灾和 2019 年 10 月猪肉价格上涨 (式 (3)) 的最优 44 个阶数选择。

图 4 展示了异常值检测结果在考虑春节效应和其他异常值前后的变化。图 4 (左上) 展示了式 (1) 所含的 81 个备选模型 (未考虑春节效应和其他异常值) 中，在任一给定月份有多少模型检测出了异常值。其中，2004 年 2 月的值为 81，表明 81 个备选模型都在此月检测出异常值。在 81 个模型总共检测出了 756 个异常值，其中 29% 出现在一月，58% 出现在二月，其他月份占 13%。我们看到 87% 的异常值出现在一月和二月，和春节所在月份相吻合。这也表明春节效应对中国 CPI 而言是一个很重要的影响因素，正确的量化春节效应是非常必要的。而当我们在模型中考虑春节效应之后，异常值的个数由 756 减少到 173，如此巨大的减少也进一步佐证了我们量化春节效应的方法是恰当且有效的。下一节将详细介绍我们量化春节效应的方法。

2.3 春节效应和异常值的干预分析

正如前文所述，由于在春节期间居民对商品和服务的需求增大，春节又是居民回家团聚的日子，因此商品的生产和服务的供给有所减少，导致商品和服务的价格在春节期间有一个明

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

显的提高。量化春节效应的难点在于春节日期在 1 月 21 日到 2 月 20 日之间变动。春节日期的逐年变动，导致了每年的一、二、三月份受春节影响的天数也逐年变动。若某一月份在某一年受春节影响的天数越多，则它的春节效应就越强，反之春节效应就越弱。

前人对如春节这样的移动节日效应早有研究，特别是在季节调整领域。Bell 和 Hillmer (1983)[15] 曾指出在时间序列分析中不应忽略交易日效应和移动节日效应，并提出了如何量化复活节效应。Lin 和 Liu (2003)[16] 探讨了关于春节效应在季节调整过程中的量化，并将他们的方法应用到中国台湾的宏观经济数据中。这些工作大体都是根据每个月份受节日影响的天数来量化该月的节日效应。本文采用 Lin 和 Liu (2003)[16] 的方法来量化中国 CPI 的春节效应。

具体来说，商品价格一般在春节之前某一天抬升，在春节期间维持在一个较高水平，在春节假期之后开始回落，居民的消费活动也在春节前、中、后期有所不同。所以我们将春节效应分成三部分：节前效应、节中效应、节后效应。持续天数分别是 τ_1、τ_2 和 τ_3。这三个参数要通过数据来选取。具体来说，假设春节发生在某月的第 d_0 天，则第 $d_0-\tau_1$ 天到 d_0-1 天是节前效应区间，第 d_0 天到 $d_0+\tau_2-1$ 天是节中效应区间，第 $d_0+\tau_2$ 天到 $d_0+\tau_2+\tau_3-1$ 天是节后效应区间。对任一给定的月份 t 和给定的参数 τ_1、τ_2、τ_3，我们可以得到此月中分别有多少天落入节前、节中和节后效应区间，各自记为 τ_{1t}、τ_{2t} 和 τ_{3t}，然后我们定义如下三个自变量来量化月份 t 所受到的春节效应：

$$H_{it}(\tau_i)=\frac{\tau_{it}}{\tau_i},\qquad i=1,2,3.$$

这里我们假设 t 月的节前、节中、节后效应的大小分别正比于此月落入对应区间的天数。这种量化方法和 Bell 和 Hillmer (1983)[15] 以及 Lin 和 Liu (2003)[16] 一致。将此三个自变量加入模型式 (1)，我们得到：

$$\phi_p(B)\Phi_P(B^{12})(1-B)(1-B^{12})\Big(Z_t-\sum_{i=1}^{3}\beta_i H_{it}(\tau_i)\Big)=\theta_q(B)\Theta_Q(B^{12})a_t, \tag{2}$$

其中，τ_1、τ_2、τ_3 是未知参数，我们将在下一节通过模型拟合和预测效果对 τ_1、τ_2、τ_3 以及阶数 p、q、P、Q 进行选择。对比与 SARIMA 模型式 (1)，此模型多了三个自变量。此类在 SARIMA 模型基础上加入自变量后产生的模型被称为 SARIMAX 模型。

图 4 (右上) 展示了模型式 (2) 的异常值检测结果，我们可以看出加入春节效应异常值有一个极大的减少 (从 756 减少到 173)，这表明我们的春节效应量化方法是有效的。不过，在 2008 年 2 月仍有 79 个模型检测出了异常值，并且 2019 年 10 月有 26 个模型检测出异常值。而查看历史事件我们知道，2008 年中国发生了严重雪灾，从一月中旬持续到二月末，波及范围极广，尤其是中国南部受灾严重。此次雪灾导致当时不少地区食品供应紧张、交通受阻，从而使得物价特别是食品价格大幅上涨。随着灾情得到控制，物价逐渐回落到正常水平。由于这次雪灾是一次突发事件，对物价的影响持续时间不长，所以我们认为雪灾对 CPI 的影响是可加异常值，可以通过加入自变量 $\omega P_t^{(2008-02)}$ 来刻画。而 2019 年 10 月 CPI 的突然上升主要来自猪肉价格的暴涨，而导致猪肉价格的暴涨的主要因素是非洲猪瘟。根据异常值检测结果，其对 CPI 的影响可以用 IO 刻画。加入这两个异常值后，模型改进为：

$$\begin{aligned}&\phi_p(B)\Phi_P(B^{12})(1-B)(1-B^{12})\Big(Z_t-\omega_1 P_t^{(2008-02)}-\sum_{i=1}^{3}\beta_i H_{it}(\tau_i)\Big)\\&\quad=\theta_q(B)\Theta_Q(B^{12})(a_t-\omega_2 P_t^{(2019-10)}).\end{aligned} \tag{3}$$

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

当模型中加入春节效应和 2008 年雪灾效应之后，81 个模型的异常值总个数减少到 68 个，并且最好的 44 个模型仅 6 个异常值 (图 4 (左下) 和 (右下))。至此，我们完成春节效应和异常值的处理。

2.4 参数选择

在模型式 (3) 中还有两类未确定的参数，模型阶数 p,q,P,Q (81 种) 和春节效应参数 τ_1,τ_2,τ_3。为了减少计算量，我们在如下范围内选择 (τ_1,τ_2,τ_3)：$\tau_1,\tau_3\in\{0,4,8,12,16,20,24\}$，$\tau_2\in\{0,4,8\}$，总共 $3\times 7^2=147$ 种选择。因此，我们的模型式 (3) 总共有 81×147 种参数选择。我们将根据每种选择的模型表现来选择最优参数。设 $\boldsymbol{\tau}=(\tau_1,\tau_2,\tau_3)'$。假定 $\{a_t\}$ 服从独立同分布的正态分布 $N(0,\sigma_a^2)$。对于任一组 $(\boldsymbol{\tau};p,q,P,Q)$，我们可以得到 σ_a 在模型式 (3) 中的极大似然估计，记为 $\hat{\sigma}_a(\boldsymbol{\tau};p,q,P,Q)$。我们用以下指标刻画参数组合 $(\boldsymbol{\tau};p,q,P,Q)$ 对应的拟合效果：

$$C_{Fit}(\boldsymbol{\tau};p,q,P,Q)=\frac{1}{\hat{\sigma}_0}\hat{\sigma}_a(\boldsymbol{\tau};p,q,P,Q),$$

其中 $\hat{\sigma}_0$ 是平稳序列 $(1-B)(1-B^{12})Z_t$ 的标准差，其表示了 CPI 序列的基准波动。中国 CPI 的 $\hat{\sigma}_0$ 值是 0.7244。在本节第 1 小节中我们已经检验了 $(1-B)(1-B^{12})Z_t$ 为平稳序列。因此 $\hat{\sigma}_0$ 表示了将 CPI 原序列通过差分得到平稳序列后但未建立 S–ARIMAX 模型时 CPI 序列的标准差。因此 $\frac{1}{\hat{\sigma}_0}\hat{\sigma}_a(\boldsymbol{\tau};p,q,P,Q)$ 表示了通过 S-ARIMAX 模型对序列标准差的相对改变。我们还考虑了用贝叶斯准则作为模型拟合和模型复杂度的指标：

$$BIC(\boldsymbol{\tau};p,q,P,Q)=-2\log\big(L(\boldsymbol{\tau};p,q,P,Q)\big)+k\log n,$$

其中 $L(\boldsymbol{\tau};p,q,P,Q)$ 是参数组 $(\boldsymbol{\tau};p,q,P,Q)$ 所对应的极大似然，$n=T-13$，$k=|\boldsymbol{\tau}|+p+q+P+Q+1$ 是模型式 (3) 中除 σ_a 外的系数个数，$|\boldsymbol{\tau}|$ 表示 $\boldsymbol{\tau}$ 中非零元素的个数。

对任一组 $(\boldsymbol{\tau};p,q,P,Q)$，我们如下定义模型的预测效果。首先将数据分成训练集 (2002 年 1 月至 2008 年 12 月) 和验证集 (2009 年 1 月至 2019 年 12 月)，在训练集估计模型式 (3) 后对验证集进行一步样本外预测，从而得到均方根预测误差 (RMSE)：$RMSE(\boldsymbol{\tau};p,q,P,Q)=(\sum_{t=85}^{216}(\widehat{CPI}_t-CPI_t)^2)^{\frac{1}{2}}$。最后模型的预测效果由如下指标表示：

$$C_{FC}(\boldsymbol{\tau};p,q,P,Q)=\frac{1}{\hat{\sigma}_0}RMSE(\boldsymbol{\tau};p,q,P,Q).$$

此指标代表了通过 S–ARIMAX 建模对序列预测误差的相对改进，其揭示了此序列的模型可预测性。

定义完这三个模型指标后，我们通过以下两阶段程序来选择参数 $(\boldsymbol{\tau};p,q,P,Q)$。在第一阶段，我们先选出对每种 (p,q,P,Q) 而言最优的 $\boldsymbol{\tau}$。对任一给定的阶数 (p,q,P,Q)，我们遍历 147 种 $\boldsymbol{\tau}$ 的选择，并计算三个指标 $C_{Fit}(\boldsymbol{\tau};p,q,P,Q)$，$BIC(\boldsymbol{\tau};p,q,P,Q)$ 和 $C_{FC}(\boldsymbol{\tau};p,q,P,Q)$，并给出他们的名次 (1 到 147)。然后将三个指标的名次加总得到每个 $\boldsymbol{\tau}$ 的总排名，根据这个总排名我们选出最优的 $\boldsymbol{\tau}$，记作 $\boldsymbol{\tau}^*(p,q,P,Q)$。这样对于任一组 (p,q,P,Q)，我们得到了其最优的 $\boldsymbol{\tau}^*$ 以及在此 $\boldsymbol{\tau}^*$ 下三个指标的值，分别记为 $C_{Fit}(p,q,P,Q)=C_{Fit}(\boldsymbol{\tau}^*;p,q,P,Q)$，$BIC(p,q,P,Q)=BIC(\boldsymbol{\tau}^*;p,q,P,Q)$ 和 $C_{FC}(p,q,P,Q)=C_{FC}(\boldsymbol{\tau}^*;p,q,P,Q)$。在第二阶段，我们先对任一组 (p,q,P,Q) 给出 $C_{Fit}(p,q,P,Q)$，$BIC(p,q,P,Q)$ 和 $C_{FC}(p,q,P,Q)$ 的排名 (1 到 81)，再将三个排名之和作为此 (p,q,P,Q) 的最终总排名，来确定最优的 (p,q,P,Q)。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

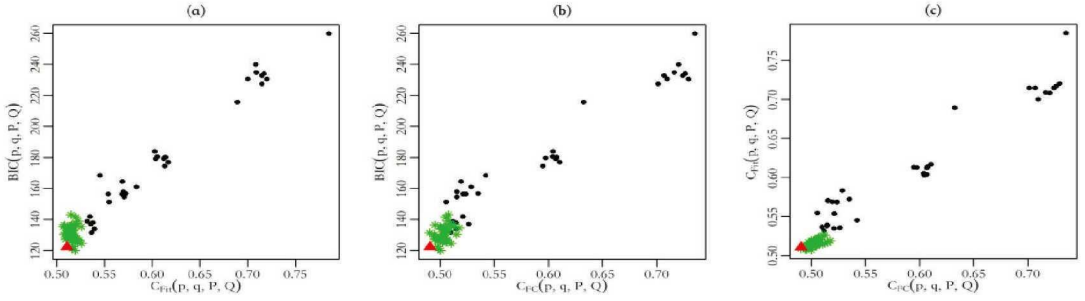

图 5 每种 (p,q,P,Q) 对应的 $C_{Fit}(p,q,P,Q)$，$BIC(p,q,P,Q)$ 和 $C_{FC}(p,q,P,Q)$。红色三角对应最终总排名最好的 (p,q,P,Q)。绿色星状和红色三角对应最终排名为前 44 名的 (p,q,P,Q)。

图 5 给出了 81 种阶数选择 (p,q,P,Q) 所对应的三种指标值的散点图。可见有一簇选择 (绿色星状和红色三角) 在三个指标上明显胜过了其他选择。这一簇包含了 44 种阶数选择，它们的三种指标都满足：$C_{Fit}(p,q,P,Q) \le 0.53$，$BIC(p,q,P,Q) \le 150$，$C_{FC}(p,q,P,Q) \le 0.53$。附件表 A.1 列出了排名前八的模型阶数的三种指标结果和对应排名。这八个模型的拟合和预测效果很接近，没有一个模型在三个指标上同时达到最优。因此，我们将三个指标的排名之和作为总排名，来选择最优的模型。最终，最优模型为 S–ARIMAX$(0,1,1)\times(2,1,1)_{12}$ 以及 $(\tau_1,\tau_2,\tau_3)=(4,0,12)$：

$$(1-\phi B)(1-\Phi_1 B^{12}-\Phi_2 B^{24})(1-B)(1-B^{12})\Big(Z_t-\omega_1 P_t^{(2008-02)}-\sum_{i=1}^{3}\beta_i H_{it}(\tau_i)\Big) \\ =(1-\Theta B^{12})(a_t-\omega_2 P_t^{(2019-10)}). \tag{4}$$

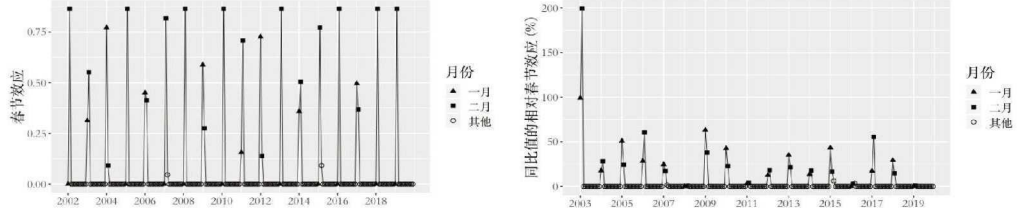

图 6 中国 CPI 定基值的春节效应 (左) 和 CPI 同比值的相对春节效应 (右)

在此模型下没有再检测出任何异常值。图 6 (左) 展示了 CPI 定基值的春节效应估计值 $\sum_{i=1}^{3}\hat{\beta}_i H_{it}(\tau_i)$。此外，CPI 同比值广受受业界关注，对 CPI 趋势的解读也大都集中在同比值，这是因为同比值体现了 CPI 的年度增长率，并且一定程度上消除了季节因素的影响。为了直接刻画春节效应对同比值变化的贡献大小，图 6 (右) 还给出了 CPI 同比值的相对春节效应，其定义如下：

$$\text{同比值的相对春节效应 } =\left|\frac{R_t-\tilde{R}_t}{R_t}\right|\times 100\%,$$

其中，R_t 是 CPI 的同比值，$\tilde{R}_t$ 是 CPI 去掉春节效应之后 (即 $CPI_t-\sum_{i=1}^{3}\hat{\beta}_i H_{it}(\tau_i)$) 的同比值。由图 6 (b) 可以看出，同比值的相对春节效应在每年的一、二月份都不为零，取值从 3% 到 191% 不等，且大都集中在 25% 到 75% 之间。因此在 1，2 月份，春节效应对 CPI 同比的影响能达到 25%–75%，有时甚至高达 191%。可见春节对 CPI 同比值影响非常显著，也说明了在公布 1，2 月份的 CPI 同比时，对原始数据作调整以去掉春节效应是很有必要的。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 1 列出了中美 CPI 最优模型在考虑春节效应和异常值前后的拟合和样本外预测效果。我们用全部数据 (2002 年 1 月至 2019 年 12 月) 考察模型的拟合效果，用 2009 年 1 月到 2019 年 12 月数据考察模型的样本外预测效果，其中我们考虑了 1 步 (1 个月) 到 12 步 (12 个月) 样本外预测。美国 CPI 的 S–ARIMAX 模型请见 4.1 小节。以模型式 (4) 为例。为了计算拟合误差，我们使用 2002 年 1 月至 2019 年 12 月的数据拟合模型式 (4) 并得到每月 CPI 的拟合值 $\widehat{CPI}_t$，拟合误差则为 CPI 的拟合值与真实值之差，即 $\widehat{CPI}_t - CPI_t$。在计算样本外预测误差时，我们使用扩展窗口 (Expanding Window) 来模拟真实的预测情景。例如在 1 步预测 2009 年 1 月 CPI 时，我们先用 2002 年 1 月至 2008 年 12 月的数据估计模型式 (4)，再通过估计后的模型预测 2009 年 1 月 CPI。在 1 步预测 2009 年 2 月 CPI 时，我们用 2002 年 1 月至 2009 年 1 月的数据重新估计模型，再通过估计后的模型预测 2009 年 2 月 CPI。我们可以看到，对于中国 CPI，在考虑春节效应、2008 年雪灾和 2019 年猪肉价格陡增后，模型的拟合和预测效果得到了显著提升，拟合误差 ($\frac{\hat{\sigma}_a}{\hat{\sigma}_0}$) 减少了 26.5%，一步预测误差 ($\frac{1}{\hat{\sigma}_0}RMSE$) 减少了 26.9%。图 7 也体现了这种提升：在未考虑春节效应和任何异常值之前，模型的拟合误差和 1 步预测误差在某些年的 1,2 月份异常的大 (图 7 虚线)，而考虑这些效应之后这些较大误差都消失了。表 1 也说明了对美国 CPI 进行异常值干预分析也能显著改善模型效果。

表 1 中美 CPI 最优模型的拟合和 h– 步样本外预测结果

		模型拟合的 $\frac{\hat{\sigma}_a}{\hat{\sigma}_0}$	h– 步样本外预测 $\frac{1}{\hat{\sigma}_0}RMSE$					
			1 步	2 步	3 步	6 步	9 步	12 步
中国 CPI	考虑春节效应和异常值之前	0.68	0.67	0.91	1.21	1.96	2.69	3.32
	考虑春节效应和异常值之后	0.50	0.49	0.77	1.04	1.77	2.47	2.93
美国 CPI	异常值干预分析前	0.59	0.59	0.89	1.27	2.30	2.95	3.18
	异常值干预分析后	0.45	0.52	0.89	1.25	2.24	2.83	3.06

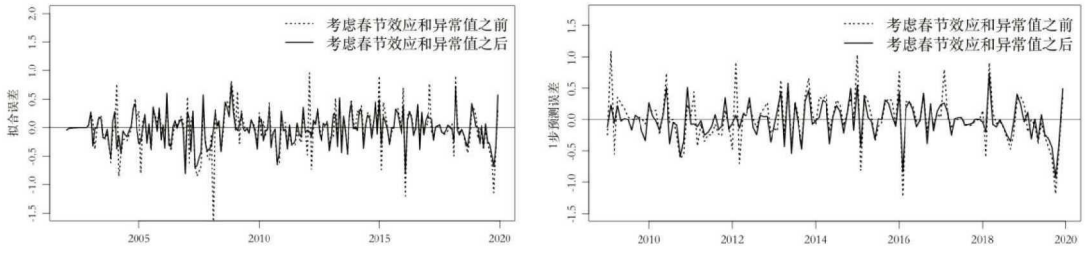

图 7 模型拟合误差 (左) 和 1 步预测误差 (右) 在加入春节效应和 2008 年雪灾效应的前后比较

3 扩散指数 (DI) 方法

到目前为止，我们使用了单变量的 S–ARIMAX 模型对中国 CPI 进行建模。这种单变量模型忽略了其他经济变量中对 CPI 预测有用的信息。为了提高预测效果，我们收集了其他主要经济变量，并通过主成分分析来提取这些变量中的信息来用作扩散指数 (Diffusion Index，简称 DI)。然后我们用估计的 DI 来预测季节调整后的 CPI，其中 DI 的数量根据贝叶斯准则来确定。

DI 方法的前提是预测变量中的信息可以由少量估计因子表示。这种方法避免了变量选择的过程，从而避免错过在样本外具有预测能力但在样本内缺乏显着性的变量。如 Stock 和 Watson (2002)[9−10] 及其中的参考文献所述，这一观点在宏观经济学中有着悠久传统。对于中

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

国 CPI，我们考虑附件表 A.2 中列出的 31 个主要经济变量。对于美国 CPI，应用最广泛且影响最大的就是 Stock–Watson 数据集。但此数据集并没有 2001 年以后的数据。本文采用的是 McCracken 和 Ng (2016)[12] 提出的联储经济数据集 (Federal Reserve Economic Data，简称 FRED)，此数据集已经被证明是 Stock–Watson 数据集的一个合适替代。

3.1 季节调整和数据预处理

用于构建 DI 的协变量也是月度观测值，其具有与 CPI 相同的时间跨度。由于中国数据未经过季节调整，我们需要对每个协变量序列进行季节调整。对于包含春节效应的协变量序列 (记作 Z_t，有的做了对数变换)，首先用 S–ARIMAX 模型拟合，估计出春节效应 $\hat{H}_t = \sum_{i=1}^{3} \hat{\beta}_i H_{it}(\tau_i)$。然后从 Z_t 中减去此春节效应，就得到了去除春节效应的序列 $\tilde{Z}_t = Z_t - \hat{H}_t$。再将 X–13ARIMA–SEATS 应用于序列 $\tilde{Z}_t$ 上，根据加法或乘法模型将其分解为趋势项 T_t，季节项 S_t 和不规则项 I_t，并最终获得季节调整后的序列 saZ_t。其中

$$\begin{aligned} &\text{加法模型：} \quad \tilde{Z}_t = T_t + S_t + I_t, \qquad saZ_t = \tilde{Z}_t - S_t = Z_t - \hat{H}_t - S_t, \\ &\text{乘法模型：} \quad \tilde{Z}_t = T_t S_t I_t, \qquad\quad saZ_t = \frac{\tilde{Z}_t}{S_t} = \frac{Z_t - \hat{H}_t}{S_t}. \end{aligned}$$

对于不同协变量，要根据此变量的性质来选择使用加法模型还是乘法模型。

在季节调整之后，我们对序列 saZ_t 进行适当的变换来得到平稳序列。附件表 A.2 详细地列出了每个协变量的预处理信息。最后，我们将所有变换后的平稳序列标准化，形成多元时间序列 X_t，用于 DI 模型的估计。特别地，我们将季节调整后的 CPI 序列记做 $saCPI_t$。

3.2 DI 模型估计

在数据预处理之后，我们对 n 维多元时间序列 X_t 建立如下因子模型来估计扩散指数 (DI)：

$$X_t = \Lambda F_t + e_t, \tag{5}$$

其中 F_t 是 r 维公因子，也是我们需要估计的 DI，Λ 是 $n \times r$ 维因子载荷矩阵，e_t 是满足 $E(e_t|F_t) = 0$ 和有限二阶矩的误差向量。这里的公因子 F_t 由主成分方法得到。Stock 和 Watson (2002)[9] 证明了此估计是逐点 (对于任意时刻 t) 收敛的，并且在适当条件下，在任意时刻 t 上的均方误差收敛到 0。此外，估计出的公因子 F_t 可以用美国经济研究院 (NBER) 制定的扩散指数来解释，这一指数旨在衡量一系列宏观经济变量中的共同变动 (Stock 和 Watson, 2002)[10]，这也是此方法名称的由来。我们数据集中的一些数据有缺失，如工业生产者出厂价格指数 (PPI) 的轻工业子类。因此，标准的主成分分析算法并不适用。然而，Stock 和 Watson (2002)[10] 的期望最大化算法 (EM algorithm) 可以处理存在缺失值的状况。

3.3 使用 DI 进行预测

因为 $saCPI_t$ 取对数之后是二阶单位根序列，我们参照 Stock 和 Watson (2002)[10] 中的步骤定义：

$$y_{t+h}^{h} := \frac{1200}{h} \ln \frac{saCPI_{t+h}}{saCPI_t} - 1200 \ln \frac{saCPI_t}{saCPI_{t-1}}. \tag{6}$$

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

以及 $y_t = 1200\Delta \ln \frac{saCPI_t}{saCPI_{t-1}}$。DI 方法的一般预测方程是：

$$y_{t+h}^h = \alpha + \sum_{j=0}^{\tilde{p}-1} \gamma_j y_{t-j} + \sum_{j=0}^{m-1} \tilde{\beta}_j' f_{t-j} + \varepsilon_{t+h}. \tag{7}$$

这里 $\tilde{p}$ 和 m 分别是自回归阶数和公因子的滞后阶数，向量 f_t 是 DI 模型式 (5) 估计出的 F_t 的前 k 维子向量 ($k \le r$)。为了在任意给定时间 t_0 预测 h 步之后的 $y_{t_0+h}^h$，我们有三个待确定的调整参数：公因子的数量 k，自回归阶数 $\tilde{p}$ 和公因子的滞后阶数 m。将表示为预测中使用的最大可能因子数记作 $\max.k$，其可能的取值是 1,2,3,4,5。对于任一 $\max.k \in \{1,2,3,4,5\}$，我们基于贝叶斯准则在 $\{1 \le k \le \max.k, 1 \le \tilde{p} \le 6, 1 \le m \le 6\}$ 中选择最优的 $(k, \tilde{p}, m)$；然后使用选择的 $(k, \tilde{p}, m)$ 去拟合模型式 (7) 并预测 $y_{t_0+h}^h$，记预测值为 $\hat{y}_{t_0+h}^h$。在获得 $\hat{y}_{t_0+h}^h$ 后，通过式 (6) 我们可以得到季节调整后 CPI 的 h 步预测 $\widehat{saCPI}_{t_0+h}^h = \exp(\frac{h\hat{y}_{t_0+h}^h}{1200}) \frac{saCPI_{t_0}^{h+1}}{saCPI_{t_0-1}^h}$。然后我们通过 S–ARIMA 模型预测季节因子 S_{t_0+h}，其阶数是根据序列 $\{S_t\}$ 的 SACFs 和 SPACFs 选择的。最后，由式 $\widehat{CPI}_{t_0+h}^h = \widehat{saCPI}_{t_0+h}^h + \hat{S}_{t_0+h} + \hat{H}_{t_0+h}$ 获得最终 $CPI_{t_0+h}^h$ 的预测。

为了比较 DI 方法与 S–ARIMAX 模型的预测效果，我们对 2009 年 1 月 – 2019 年 12 月进行了样本外预测，预测步长分别为 1、2、3、6、9 和 12 个月。这里我们考虑两种预测方案，分别是扩展窗口 (expanding window) 和滚动窗口 (rolling window)。在扩展窗口方案中，训练数据集是从 2002 年 1 月直到待预测月份之前 h 月的区间。在滚动窗口方案中，训练数据集也就是移动窗口，是待预测月份之前的 $(49-h)$ 个月。

3.4 中国 CPI 的 DI 方法预测结果

我们应用上述两节中的程序来预测季节调整后的中国 CPI，并使用 S–ARIMA$(0,0,1) \times (0,1,1)_{12}$ 来对季节项 S_t 建模。表 2 展示了 DI 方法与 S–ARIMAX 方法的样本外预测 RMSE 之比，数值小于 1 的值意味着 DI 方法的 RMSE 比 S–ARIMAX 的更小。可以发现 DI 方法在 1，2 和 3 步预测明显优于 S–ARIMAX 模型。对于滚动窗口，DI 模型分别将预测 RMSE 降低了 7%，6% 和 5%，而对于滚动窗口，DI 模型分别将预测 RMSE 降低了 26%，12% 和 7%。附件图 A.1 的 1–3 步预测误差也反映了这些改进，并可以看出 2010 年之后 DI 方法的预测误差小于 S–ARIMAX 模型的预测误差。

表 2 中国 CPI 的 DI 方法与 S–ARIMAX 方法的样本外预测 RMSE 之比

预测步长	扩展窗口 (Expanding Window)					
	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$
$\max.k=1$	0.932	0.937	0.952	1.126	1.364	1.595
$\max.k=2$	0.932	0.937	0.996	1.119	1.282	1.437
$\max.k=3$	0.932	0.937	0.996	1.123	1.282	1.437
$\max.k=4$	0.932	0.937	0.996	1.123	1.282	1.437
$\max.k=5$	0.932	1.001	1.014	1.121	1.283	1.501
预测步长	滚动窗口 (Rolling Window)					
	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$
$\max.k=1$	0.743	0.874	0.937	1.260	1.480	1.633
$\max.k=2$	0.741	0.876	0.934	1.263	1.432	1.556
$\max.k=3$	0.741	0.876	0.934	1.823	1.443	1.556
$\max.k=4$	0.741	0.876	0.934	1.823	1.443	1.556
$\max.k=5$	0.742	0.882	0.966	1.404	1.494	1.739

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

此外，对于 1，2 和 3 步预测，仅使用第一个因子已经足够，因为从第二个因子到第五个因子对预测都没有明显的改善。这表明第一个因子可以用作中国 CPI 的短期预测指数。附件表 A.3 列出了所有变量的因子负荷。如果只关注因子负荷绝对值大于 0.2 的变量，我们可以看出第一个因子主要由以下协变量组成 (括号内为变量 id，因子负荷):PPI(ppi，0.25)，PPI: 生活材料 (ppi_living，0.34)，PPI: 生活材料: 食物 (ppi_living_food，0.37)，PPI 同比: 轻工业 (ppi_light_rate，0.32)，零售价格指数 (rpi，0.33)，RPI: 食物 (rpi_food,0.29)，CPI: 食物: 畜肉类 (cpi_food_meat，0.3)，CPI: 食物: 猪肉 (cpi_food_pork，0.27)，再贴现利率 (rediscount_rate，0.21)，短期贷款利率 (loan_rate，0.21)，人民币存款准备金率 (reserve_ratio，0.2)。因此，这些协变量具有中国 CPI 的短期预测能力。在这 11 个协变量中，食物占据了 4 个，这反映出中国 CPI 以食品价格为主要驱动。这也和中国 CPI 的构成中食品占主导相一致。

最后，表 2 和附件图 A.1 的 6，9，12 步预测误差都表明，在长期预测上 ($h = 6, 9$ 和 12) DI 方法比 S–ARIMAX 模型具有更大的预测误差。这说明 S–ARIMAX 更利于长期预测。

4 美国 CPI 序列结果

在前两节中，我们应用线性时间序列模型 S–ARIMAX 和 DI 模型来研究了中国 CPI 序列。而对美国 CPI，文献中已有此类的分析研究 (Stock 和 Watson，2002[9−10]，McCracken 和 Ng，2016[12])。一个至关重要的问题是中国 CPI 的动态结构和美国的是否相似？此外，两个序列的可预测性如何？为了回答这些问题，本节将对美国 CPI 进行分析。这里采用的方法与前两节对中国 CPI 的分析类似，故不做赘述，而只报告美国 CPI 的建模结果和发现。我们还将特别讨论我们的结果与已有文献结果的联系。最后，第六节总结了中美 CPI 之间主要异同。

4.1 线性时间序列建模

对于美国 CPI，我们将原始 CPI 序列表示为 Z_t，附件图 A.2 提供了该序列的 SACFs 和 SPACFs，以及一阶和二阶差分后序列 (即 $(1-B)Z_t$ 和 $(1-B)(1B^{12})Z_t$) 的 SACFs 和 SPACFs。与第 3.1 节的分析类似，这些 SACFs 和 SPACFs 揭示了美国 CPI 符合如下模型：

$$\phi_p(B)\Phi_P(B^{12})(1-B)(1-B^{12})Z_t = \theta_q(B)\Theta_Q(B^{12})a_t. \tag{8}$$

这和中国 CPI 的模型类别式 (1) 是相同的。

类似于第三节的建模过程，为了选择模型式 (8) 的阶数 p、q、P、Q，我们先基于 $(1-B)(1-B^{12})Z_t$ 的 SACFs 和 SPACFs 确定一个大致范围。由于附件图 A.2 (左下) 中的 SACF 仅在滞后 1,12,13 阶显著，而附件图 A.2 (右下) 中 SPACF 在滞后 1,2,12,15,24 和 25 阶显著，我们从 $\{0,1,2\}\{0,1,2\}$ 中选择 p、q、P、Q。这和中国 CPI 的模型阶数范围一样，共 81 个备选模型。

附件图 A.3 显示了 81 个候选模型的异常值检测结果。如附件图 A.3 (左上) 所示，81 个模型中绝大多数都在 7 个月份检测到了异常值，分别是 2005 年 9 月，2008 年 10 月和 11 月共有 3 个 IO；2007 年 11 月和 2015 年 1 月的两个 LS，以及 2009 年 6 月的两个 AO 和 2013 年 2 月的一个 TC。2005 年 9 月的 IO 可以归因于 8 月底发生的飓风丽塔 (Rita) 和 2005 年 9 月的飓风卡特里娜 (Katrina)。2015 年 1 月的异常值是由汽油价格大幅下跌引起的，这反映出美国 CPI 受到能源价格的高度影响。2008 年 10 月和 11 月的 IO 可归因于金融危机。最后，2009 年 6 月的 AO 是由于汽车补贴退税制度 (CARS)，这也在 BLS 官方的分析文件 [10] 中得到了确认，文件中说 “$\cdots$ 汽车补贴退税制度导致了二手车供应量减少和二手车价格上涨 $\cdots$”。

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

对这 7 个异常值进行干预分析 (Box 和 Tiao, 1975)[10]，模型式 (8) 被修改为

$$\begin{aligned}&\phi_p(B)\Phi_P(B^{12})(1-B)(1-B^{12})\\&\cdot\Big(Z_t-\omega_1 S_t^{2007-11}-\omega_2 P_t^{2009-06}-\frac{\omega_3}{1-\delta B}\,P_t^{2013-02}-\omega_4 S_t^{2015-01}\Big)\\&=\theta_q(B)\Theta_Q(B^{12})(a_t+\omega_1 P_t^{2005-09}+\omega_2 P_t^{2008-10}+\omega_3 P_t^{2008-11}).\end{aligned}\tag{9}$$

此新模型下的异常值检测结果如附件图 A.3 (右上) 所示，表明有一半以上的候选模型不再有异常值，特别是下文选择出的 20 个最佳模型。

为了在模型式 (9) 中选择正确的阶数 p、q、P、Q，我们计算了 81 个候选模型的三个指标：$C_{Fit}(p,q,P,Q)$，$BIC(p,q,P,Q)$ 和 $C_{FC}(p,q,P,Q)$，如附件图 A.4 所示。我们发现最优的 16 个模型 (附件图 A.4 中标记为绿色星状和红色三角) 都满足下面三个标准：$C_{Fit}(p,q,P,Q)\le 0.5$；$BIC(p,q,P,Q)\le 360$；$C_{FC}(p,q,P,Q)\le 0.55$。附件图 A.3 (左下) 刻画出了这 16 个模型的异常值检测结果，可以发现干预分析后这 16 个模型中的异常值个数大幅减少。起初平均每个模型有 7.46 个异常值，干预分析之后降至平均每个模型 2.47 个异常值。如果进一步缩小到前 16 名模型，此数值进一步降到了 0.75。阶数为 $(2,1,2)\times(0,1,1)_{12}$ 的最优模型在附件图 A.4 中以红色三角形表示，其形式为：

$$\begin{aligned}&(1-\phi_1 B)(1-\phi_2 B^2)(1-B)(1-B^{12})\\&\cdot\Big(Z_t-\omega_1 S_t^{2007-11}-\omega_2 P_t^{2009-06}-\frac{\omega_3}{1-\delta B}P_t^{2013-02}-\omega_4 S_t^{2015-01}\Big)\\&=(1-\theta_1 B)(1-\Theta_2 B^2)(1-\Theta_1 B^{12})(a_t+\omega_1 P_T^{2005-09}+\omega_2 P_t^{2008-10}+\omega_3 P_t^{2008-11}).\end{aligned}\tag{10}$$

表 3 给出了此模型在异常值干预分析前后模型的拟合和预测效果。在干预分析之后，拟合误差 ($\frac{\hat{\sigma}_a}{\hat{\sigma}_0}$) 和 1 步样本外预测误差 ($\frac{1}{\hat{\sigma}_0}RMSE$) 较干预分析之前分别减少了 24% 和 12%。图 8 也显示出，所有异常值导致的较大误差在干预分析之后都被消除了。

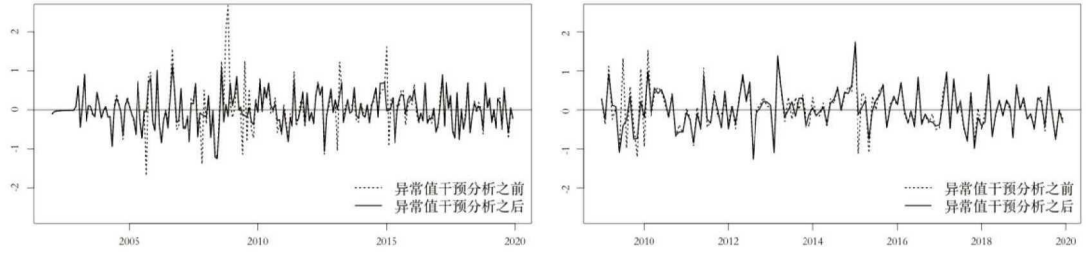

图 8 美国 CPI 模型拟合误差 (左) 和 1 步预测误差 (右) 在异常值干预分析的前后比较

比较中国 CPI 和美国 CPI 各自的前 8 名 S–ARIMAX 模型 (附件表 A.1 和表 A.4)，我们发现中国 CPI 的前 8 名模型中，p 大都取值 1，P 主要取值为 2；而美国 CPI 的前 8 名模型中，p 大都取值为 2，P 主要是 0 或 1。可见美国 CPI 倾向于选择比中国 CPI 更大的常规自回归阶数，以及更小的季节自回归阶数。这种阶数选择的差异意味着，美国 CPI 序列相较于中国 CPI 序列具有更长的月度自相关性和更短的季节自相关性。

4.2 DI 方法结果

前文提到，在对美国 CPI 应用 DI 方法时，我们采用的是 FRED 月度数据集。其中的宏观经济变量包括 CPI 都有官方的季节调整数据。为了使中美 CPI 的量化分析可比，对美国 CPI

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

我们考虑其原始序列 (未经过季节调整)。因为 BLS 采用了乘法模型对 CPI 进行季节调整，我们计算其季节性为 $S_t = \frac{CPI_t}{saCPI_t}$，其中 CPI_t 和 $saCPI_t$ 分别是 BLS 公布的 CPI 原始数据和季节调整之后的数据。对于其他协变量，我们沿用 McCracken 和 Ng (2016)[12] 的数据预处理方法，其中包括季节调整和变量变换的详细过程。然后我们在 $saCPI_t$ 上应用 DI 模型 (参见式 (5) 和 (7))，并使用 S–ARIMA$(2,0,0)\times(2,1,0)_{12}$ 对 $\log S_t$ 进行建模。最后将这两个成分的预测通过 $\widehat{CPI}^h_{t_0+h} = sa\widehat{CPI}^h_{t_0+h}\hat{S}_{t_0+h}$ 合成 CPI 的预测。表 3 中列出了 DI 模型的预测结果。

表 3 美国 CPI 的 DI 方法与 S–ARIMAX 模型式 (10) 样本外预测 RMSE 之比

预测步长	扩展窗口 (Expanding Window)					
	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$
$\max.k=1$	1.23	1.54	1.82	2.12	2.17	2.20
$\max.k=1$	1.17	1.61	1.8	2.12	2.17	2.20
$\max.k=2$	1.20	1.58	1.81	2.26	2.17	2.20
$\max.k=3$	1.20	1.63	1.88	2.38	2.17	2.20
$\max.k=4$	1.20	1.63	1.90	2.74	2.17	2.20
预测步长	滚动窗口 (Rolling Window)					
	$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$
$\max.k=1$	1.12	1.36	1.63	2.02	2.17	2.42
$\max.k=1$	1.11	1.44	1.63	2.01	2.17	2.44
$\max.k=2$	1.14	1.41	1.68	2.13	2.2	2.52
$\max.k=3$	1.14	1.45	1.75	2.39	2.29	2.53
$\max.k=4$	1.14	1.45	1.75	2.6	2.29	2.53

就美国 CPI 而言，DI 方法相对于单变量 S–ARIMAX 模型 (式 (10)) 表现较差。这与中国 CPI 的结果形成鲜明对比，并且似乎与 Stock 和 Watson (2002)[10] 的结论 —“DI 模型较自回归 (AR) 模型预测效果更好”相矛盾。一个原因可能是 2002 年 1 月 –2019 年 12 月的时间跨度较短，不利于进行因子提取和 DI 模型预测。附件图 A.5 所画的预测误差表明，在 2009 年 DI 模型的预测误差要比 S–ARIMAX 大很多，而其他年份两模型的预测误差比较接近，而这段时间恰逢金融危机。可见，相对于 S–ARIMAX 模型，金融危机可能会对 DI 方法的预测产生更严重的影响。同时，对比附件图 A.1 和图 A.5 可以看出金融危机对美国的影响比对中国的影响更为严重。

此外，模型的预测能力可能会随着预测区间的推移而变化，并且会因不同的目标序列而异。在 Stock 和 Watson (2002b)[10] 中，作者就 DI 和 AR 模型在 1970 年 1 月至 1998 年 12 月的 y^h_{t+h} 上进行了比较。但是，我们的预测区间是 2009 年 1 月到 2019 年 12 月，且目标序列是未经季节调整的原始 CPI。为了比较，我们还使用了 Stock 和 Watson (2002)[10] 的相同预测程序来预测 2009 年 1 月至 2019 年 12 月期间的 y^h_{t+h}。在 1、2、3、6、9 和 12 月预测中，DI 的 RMSE 与 AR 的 RMSE 之比分别为 1.1、1.4、1.4、1.3、1.3、1.1，比预测 CPI 定基值时的比值 (表 3) 明显减小。这表明 DI 方法在预测 y^h_{t+h} 上虽然仍不如 AR 模型，但二者差距不大。

此外，我们的结果与 McCracken 和 Ng (2016)[12] 的相一致，该研究比较了 2008 年 1 月至 2012 年 2 月的 y^h_{t+h} 的预测。他们的结果表明，DI 方法在 1 步向前预测上与 AR 模型表现差距不大，但在 12 步和 24 步向前预测上比 AR 模型差。值得注意的是，Stock 和 Watson (2002b)[10] 和 McCracken 和 Ng (2016)[12] 仅使用了没有 MA 成分的 AR 模型，并且使用 Bayes 信息准则来选择模型阶数。本文考虑了更广泛的模型范围，包括 AR 和 MA。我们首先根据 ACFs 和 PACFs 选择潜在的 AR 和 MA 阶数，然后使用三个指标—模型拟合方面的

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

$C_{Fit}(p,q,P,Q)$、$BIC(p,q,P,Q)$ 以及预测方面的 $C_{FC}(p,q,P,Q)$，来选择最优模型。这比仅仅依靠贝叶斯准则来选择更准确，也更适合预测。所以本文中 S–ARIMAX 的预测结果优于 DI 方法也就不足为奇了。总之，不同的预测区间、不同的目标序列和不同的预测程序使得本文的结论与 Stock 和 Watson (2002b)[10] 有所不同。

5 总结和讨论

本文提供了对中国物价指数的系统性分析，包括传统的线形时间序列模型和近 20 年来广泛应用的 DI 模型，同时也对美国 CPI 进行了相对应的量化分析和比较。本文有三个主要发现。首先，中国的 CPI 序列有规律的动态结构，包括季节性和春节效应，它能够由传统的 S–ARIMAX 很好的刻画。序列中显著的春节效应表明了 CPI 数据能够体现中国消费者的消费模式，而其他异常值的可解释性 (例如 2008 年 2 月的可加异常值可由当时的雪灾很好地解释) 也说明了 CPI 数据能够捕捉突发事件对价格波动的影响，这都说明了中国 CPI 数据的可靠性。

其次，虽然中美 CPI 数据在构成和动态结构上有所不同，但他们都能由同一类线形时间序列模型来刻画。二者的不同在于，中国 CPI 是食品子类占主导，居住其次；而美国 CPI 是居住子类占主导，并且中国 CPI 序列相较于美国 CPI 序列有更短的月度自相关性以及更长的季节自相关性。就可预测性而言，根据 S–ARIMAX 模型得到的预测性指标 $\frac{1}{\hat{\sigma}_0}RMSE$ 可以看出 (表 4)，中国 CPI 的预测误差一致地小于美国 CPI 的，尤其是 6 和 9 步预测尤其明显。这主要是因为中国 CPI 更加稳定，异常值更少，并且受金融危机影响更小。因此我们可以得出以下结论：中 CPI 较美国 CPI 具有更强的可预测性。我们发现中国 CPI 的短期预测 (1–3 个月) 可借助其他经济变量通过 DI 模型得到进一步提高。然而就美国 CPI 而言，DI 方法的预测能力不如更简单的 S–ARIMAX 模型。

表 4 中美 CPI 最优的 S–ARIMAX 模型的预测 $RMSE/\hat{\sigma}_0$

预测步长		$h=1$	$h=2$	$h=3$	$h=6$	$h=9$	$h=12$
中国 CPI	扩展窗口	0.49	0.77	1.04	1.77	2.47	2.93
	滚动窗口	0.64	0.86	1.13	1.68	2.30	2.75
美国 CPI	扩展窗口	0.52	0.89	1.25	2.24	2.83	3.06
	滚动窗口	0.58	1.02	1.43	2.40	2.92	3.07

最后，本文将可在以下方面进行拓展和延伸。第一，本文采用多个模型对 CPI 进行预测，而这些预测可以进行组合和平均。结合 Zhang 等 (2020)[19] 提出的简洁模型平均方法，可以考虑如何在时间序列模型中进行模型平均和预测。第二，本文因子模型中所采用的因子，其提取只考虑了解释变量的变异性而忽略了对预测目标的可解释性和可预测性。未来的研究可以参考 Tu 和 Lee (2019)[20] 所考虑的带监督的因子模型，在因子提取中加入预测目标的信息。如何在本文的框架下考虑以上两种拓展具有新的挑战，我们将在未来的研究中报告相关结果。

[参考文献]

[1] Chow G C. Money and price level determination in China [J]. Journal of Comparative Economics, 1987, 11(3): 319–333.

[2] 何启志, 范从来. 中国通货膨胀的动态特征研究 [J]. 经济研究, 2011, 46(7): 91–101.

[3] 栾惠德. 居民消费价格指数的实时监测 — 基于季节调整的方法 [J]. 统计研究, 2007, (2): 59–67.

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

[4] 雷怀英, 金勇进. 不同质量调整法在消费者价格指数 (CPI) 中的应用 [J]. 数理统计与管理, 2008, 27(1): 35–41.
[5] 王少平, 王津港. 中国通货膨胀的惯性变化及其货币政策含义 [J]. 统计研究, 2009, 26(5): 17–24.
[6] 魏璐, 钱存华. 关于对可能影响 CPI 的几个因素的研究 [J]. 数理统计与管理, 2014, 33(1): 122–127.
[7] 张成思. 中国通货膨胀动态形成机制的多重逻辑 [M]. 北京: 中国人民大学出版社, 2016: 38–46.
[8] Wang Y, Tu Y, Chen S X. Improving inflation prediction with the quantity theory [J]. Economics Letters, 2016, 149: 112–115.
[9] Stock J H, Watson M W. Forecasting using principal components from a large number of predictors [J]. Journal of the American Statistical Association, 2002, 97(460): 1167–1179.
[10] Stock J H, Watson M W. Macroeconomic forecasting using diffusion indexes [J]. Journal of Business and Economic Statistics, 2002, 20(2): 147–162.
[11] Bureau of Labor Statistics. BLS Handbook of Methods (Chapter 17): The Consumer Price Index [M]. Washington DC: U.S. Bureau of Labor Statistics, 2007.
[12] McCracken M W, Ng S. FRED-MD: A monthly database for macroeconomic research [J]. Journal of Business and Economic Statistics, 2016, 34(4): 574–589.
[13] Box G E, Jenkins G M, Reinsel G C. Time Series Analysis: Forecasting and Control [M]. New Jersey: John Wiley & Sons, 2015: 305–339.
[14] Chen C, Liu L M. Joint estimation of model parameters and outlier effects in time series [J]. Journal of the American Statistical Association, 1993, 88(421): 284–297.
[15] Bell W R, Hillmer S C. Modeling time series with calendar variation [J]. Journal of the American Statistical Association, 1983, 78(383): 526–534.
[16] Lin J L, Liu T S. Modeling lunar calendar holiday effects in Taiwan [J]. 台湾经济预测与政策, 2003, 33(2): 1–37.
[17] Bureau of Labor Statistics. Intervention analysis in seasonal adjustment of consumer price index [M]. Washington DC: U.S. Bureau of Labor Statistics, 2017.
[18] Box G E P, Tiao G C. Intervention analysis with applications to economic and environmental problems [J]. Journal of the American Statistical Association, 1975, 70(349): 70–79.
[19] Zhang X, Zou G, Liang H, Carroll R J. Parsimonious model averaging with a diverging number of parameters [J]. Journal of the American Statistical Association, 2020, 115(530): 972–984.
[20] Tu Y, Lee T H. Forecasting using supervised factor models [J]. Journal of Management Science and Engineering, 2019, 4: 12–27.

(C)1994-2023 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net