

通货膨胀影响因素的多元线性模型分析

陈旺龙

云南财经大学统计与数学学院, 云南 昆明
Email: wallon_lst@163.com

收稿日期: 2020年11月17日; 录用日期: 2020年12月2日; 发布日期: 2020年12月9日

摘要

通货膨胀是现代市场经济发展到一定阶段之后的必然产物, 我国的经济发展和人民群众的日常生活也会受到其一定程度的影响。在我国, 有很多因素都会对通货膨胀产生影响, 找到这些有效的因素将有助于通货膨胀控制相关方面的研究。居民消费价格指数在很大程度上能够反映通货膨胀的程度。本文基于2008年01月至2019年05月的月度数据, 使用逐步判别法来对变量进行筛选, 然后再使用多元线性回归模型来对数据进行拟合, 进而得到模型。模型显示, 工业生产者出厂价格指数、货币和准货币供应量的期末值、月内人民币的平均汇率都对居民消费价格指数有显著影响。

关键词

通货膨胀, 多元线性回归分析, 影响因素

Multivariate Linear Model Analysis of Inflation Influencing Factors

Wanglong Chen

School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming Yunnan
Email: wallon_lst@163.com

Received: Nov. 17th, 2020; accepted: Dec. 2nd, 2020; published: Dec. 9th, 2020

Abstract

Inflation is an inevitable with the development of modern market economy, and the development of economic and the daily life of the people will be affected in some way. There are many factors that have an impact on inflation in China. Finding these effective factors will advance the research how to control inflation. The consumer price index can reflect the degree of inflation. We will filter variables by using the stepwise discrimination and generate the model by using multiple linear

regression model based on the monthly data from January 2008 to May 2019. The conclusion indicates that the index of producer price, M2 and the monthly average exchange rate of RMB have a significant impact on the consumer price index.

Keywords

Inflation, Multiple Linear Regression Models, Influencing Factors

Copyright © 2020 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. 简介

通货膨胀是世界各国广泛关注的一个复杂经济问题，也是一个国家宏观经济是否健康运行和发展的重要衡量指标。国际上通行的做法是使用 CPI (居民消费价格指数)的增长幅度来衡量一个国家或地区的通货膨胀程度。在我国，居民消费价格指数(Consumer Price Index, 简称 CPI)是度量居民生活消费品和服务价格水平随着时间变动的相对数，综合反映居民购买的生活消费品和服务价格水平的变动情况；涵盖了全国城乡居民生活消费的食品烟酒、衣着、生活用品及服务、交通和通信、教育文化和娱乐、医疗保健、其他用品和服务等 8 大类、262 个基本分类的商品与服务价格。自从改革开放之后，我国的整体经济水平明显上升到了一个新的水平，而在金融危机期间，我国实行了进一步扩大内需、促进经济平稳较快增长的十项措施之后，我国的 CPI 增长幅度屡次超过 3% 的国际警戒线，出现过轮的通货膨胀，CPI 表现为剧烈的振荡以及高位稳定状态。因此，在此背景下，研究通货膨胀的显著影响因素，对指导我国经济的健康有序发展具有重要意义。

2. 文献综述

在通货膨胀影响因素方面的研究，国外的理论研究要早于我国，有学者认为导致通货膨胀的原因是货币供应量的增加速度大于所能购买的商品和服务的增加速度[1]。研究货币制度的内生性和市场经济属性之间的关系来探讨从货币发出到通货膨胀的传导机制的问题[2]。利用 2001 年第 1 季度到 2005 年第 2 季度的数据，经过分析得知，自 2001 年开始，中国人民银行采用货币冲销政策在总体上缓解了中国外汇储备增加所导致通货膨胀的增加程度[3]。在国内，王智勇在之前学者的误差修正模型上使用了协整检验的方法，然后分别用季度数据和年度数据研究了我国经济增长率与通货膨胀之间的关系，结果表明：无论是季度数据还是年度数据，国内生产总值都表现出对通货膨胀率较为显著的影响[4]。黄益平等则通过向量误差修正模型和结构向量自回归模型研究了月度同比和月度环比数据，结论阐明，过剩流动性、房价和股价都对通货膨胀产生了正向影响，过剩流动性在其中扮演了重要的角色[5]。韩学红等则探讨了实际股票收益率与通货膨胀之间的关系，研究发现，当通过膨胀上升是由于攻击侧冲击的话，那么两者之间是负相关；如果是由于需求端，那么就会表现出正相关的关系，也就是说，需求端或者供给侧的相对重要程度将决定这其中的正负相关性关系[6]。在回归分布滞后(ARDL)模型下研究了人民币汇率的影响因素，并对汇率变动下的国内通货膨胀影响因素(以消费者价格指数 CPI 为主要指标)进行了进一步的研究[7]。张成思对之前学者的模型进行了修正，将内生性货币供给理论引入模型中，以 1978~2011 年的数据为基准，实证分析了货币供应量是否对通货膨胀的具有影响，结果显示，无论是长期还是短期，广义货币供

应量(M2)的增长都显著具有加剧通货膨胀的作用[8]。

本文根据之前学者的研究成果,选取工业生产者出厂价格指数(上年同月为100)、货币和准货币供应量(单位:亿元人民币)、累计实际利用外商投资额(单位:百万美元)、月内人民币平均汇率、外汇储备(单位:亿美元)、原油进口量(单位:万吨)等可能影响通货膨胀的因素,利用R软件采用多元线性回归模型对我国2008年01月至2019年05月的数据进行实证分析,研究影响通货膨胀的主要因素。关于数据来源,外汇储备、月内人民币平均汇率来源于中国人民银行的月度数据,其余数据均来自于中国统计局网站的月度数据。在本文中,为了减小异方差(可能存在)对模型拟合的影响,我们对所有的数据都求其自然对数。

3. 模型的建立与检验

3.1. 模型的设定及优化方法

3.1.1. 多元回归模型

为了具体分析各要素对通货膨胀影响的大小,我们将居民消费价格指数(CPI)设定为因变量,即被解释变量;将工业生产者出厂价格指数(PPI)、广义货币供应量(M2)、月内人民币平均汇率(ARE)、累计实际利用外商投资额(UFC)、外汇储备(FER)、原油进口量(COI)作为解释变量。将对模型进行回归分析。采用的模型如公式3.1。

$$CPI = \beta_0 + \beta_1 PPI + \beta_2 M2 + \beta_3 UFC + \beta_4 ARE + \beta_5 FER + \beta_6 COI + \mu \quad (3.1)$$

其中, $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$ 为待估参数。 μ 为误差项,误差项 μ 则表示除 PPI、M2、UFC、ARE、FER、COI 与 CPI 的线性关系外的随机因素对 CPI 的影响,表示不能由 PPI、M2、UFC、ARE、FER、COI 与 CPI 的线性关系所解释的因素。

一般地,对误差项 μ 有三种假定:① μ 是期望为 0 的随机变量,即 $E(\mu) = 0$;② 对 PPI、M2、UFC、ARE、FER、COI 的所有值, μ 的方差 σ^2 都相同;③ μ 服从正态分布,且如果存在多组,那么它们之间相互独立,即 $\mu \sim N(0, \sigma^2)$ 。

由模型的假定,我们可以推出相应的多元回归模型有公式3.2:

$$E(CPI) = \beta_0 + \beta_1 PPI + \beta_2 M2 + \beta_3 UFC + \beta_4 ARE + \beta_5 FER + \beta_6 COI \quad (3.2)$$

回归方程中的待估参数 $\beta_k, k = 0, \dots, 6$ 都是需要使用样本数据来估计。当用样本统计量 $\hat{\beta}_k, k = 0, \dots, 6$ 去替代多元回归方程中的未知参数时,便得到了估计的多元回归方程为公式3.3:

$$\widehat{CPI} = \hat{\beta}_0 + \hat{\beta}_1 PPI + \hat{\beta}_2 M2 + \hat{\beta}_3 UFC + \hat{\beta}_4 ARE + \hat{\beta}_5 FER + \hat{\beta}_6 COI \quad (3.3)$$

其中 $\hat{\beta}_k, k = 0, \dots, 6$ 是参数 $\beta_k, k = 0, \dots, 6$ 的样本估计值, \widehat{CPI} 既是对应 CPI 的估计值。

3.1.2. 逐步回归分析

逐步回归分析是通用的选取解释变量的方法之一,而对变量选择的方法分为前进法、后退法和逐步判别法。前进法在最开始时没有用作判别的变量的,然后依次选入一个对判别能力提高最显著的变量,整个过程变量都是只进不出,当待选入的变量的显著要求达不到临界值时,变量的选入过程停止。后退法则与前进法相反,后退法是将所有变量选入,之后每次剔除一个对判别能力提高最不显著的变量,整个过程是只出不进,当余下的变量都达到显著性要求时,后退剔除变量的过程就停止。在本节的逐步回归法则是两者的结合,在变量的选择上是有进也有出。判别分析法在实际应用中颇受欢迎,所以我们接下来就对逐步判别法进行说明(给定显著性水平 α):

第一步：对每一个变量，计算其一元方差分析 F 统计量，设 $F(x_i) = \max_i F(x_i)$ ，即 x_i 有最大的判别能力。若 $F(x_i) < F_\alpha(k-1, n-k)$ ，则表明没有变量可以选入；若 $F(x_i) \geq F_\alpha(k-1, n-k)$ ，则 x_i 选入，并进入下一步。

第二步：在选入 x_i 的前提下，计算偏 F 统计量 $F(x_i | x_1)$ ，在这里，我们不妨设 $F(x_2 | x_1) = \max_{2 \leq i \leq p} F(x_i | x_1)$ ，若 $F(x_2 | x_1) < F_\alpha(k-1, n-k-1)$ ，那么选变量过程结束；若 $F(x_2 | x_1) \geq F_\alpha(k-1, n-k-1)$ ，则选入 x_2 ，然后进入下一步。

在选入 x_1, x_2, \dots, x_r 同时，设 $F(x_{r+1} | x_1, x_2, \dots, x_r) = \max_{r+1 \leq i \leq p} F(x_i | x_1, x_2, \dots, x_r)$ 。若 $F(x_{r+1} | x_1, x_2, \dots, x_r) < F_\alpha(k-1, n-k-r)$ ，则选变量过程结束；如果偏统计量 $F(x_{r+1} | x_1, x_2, \dots, x_r) \geq F_\alpha(k-1, n-k-r)$ ，则 x_{r+1} 选入，并进入下一步。

第三步：在第 $r+1$ 个变量选入后，需要将之前选入的不显著的变量剔除出去。不妨设 $F(x_l | x_1, x_2, \dots, x_{l-1}, x_{l+1}, \dots, x_{r+1}) = \max_{1 \leq i \leq r} F(x_i | x_1, x_2, \dots, x_{l-1}, x_{l+1}, \dots, x_{r+1})$ ，若 $F(x_l | x_1, x_2, \dots, x_{l-1}, x_{l+1}, \dots, x_{r+1}) \geq F_\alpha(k-1, n-k-r)$ ，则变量无需剔除，返回第二步；若 $F(x_l | x_1, x_2, \dots, x_{l-1}, x_{l+1}, \dots, x_{r+1}) < F_\alpha(k-1, n-k-r)$ ，则变量 x_l 需要被剔除，再对其余 $r-1$ 多个变量进行验证是否满足要求，直至所有的变量都经历郭这个过程，然后再回到第二步。

第四步：经过剔除和选入过程的不断重复，最后剩下的所有的变量都不能被选入，也没有变量被剔除，那么变量选择过程停止。

在一般情况下，如果选入变量的临界值 F_1 和剔除变量的临界值 F_2 相同，那么会有很小的可能性导致变量的选入和剔除过程进入一个连续不断、无休止的循环中。剔除变量临界值 F_2 比选入变量临界值 F_1 略微小一点，这种可能性就会被排除。

因为本文中，使用 **R** 程序来进行逐步回归分析，所以判断准则由 F 统计量更改为赤池信息准则(AIC)，但是其原理是相同的，而这个判断的过程也是在检验解释变量的显著性。在本例中，根据样本构造的回归方程在很大可能上不是最优的，所以需要逐步回归分析来对解释变量进行筛选，也就是每次只引入一个解释变量，同时也对已选入的解释变量进行逐个检验，整个过程会反复进行，直到没有显著的解释变量选入方程，也没有不显著的解释变量从回归方程中被剔除为止。

3.1.3. 回归方程的显著性检验

当逐步回归分析对解释变量筛选完成之后，我们就会对回归方程进行拟合并得到与解释变量相对应的系数，随后我们也需要对回归方程的显著性进行检验。对于公式 3.4，

$$y = \beta_0 + x_i \beta_i, i = 1, \dots, p \quad (3.4)$$

对回归方程的显著性检验，就是在检验所有的回归系数是否都等于零，即检验

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0 \quad (3.5)$$

如果原假设被拒绝，也就表明至少有一个 $\beta_i \neq 0$ ，当然也会有所有的 β_i 都不等于零的情况存在。也就是说，我们认为是 y 至少线性依赖于某一个自变量 x_i ，也可能是线性依赖于所有的自变量 $x_i, i = 1, \dots, p$ 。如果是接受了这个原假设，那么我们就认为，相对于误差，所有的自变量对因变量的影响都是不重要的。那么我们就需要构造一个统计量，在多元回归中，我们使用统计量 F ：

$$F = \frac{SSR/P}{SSE/(n-P-1)} \quad (3.6)$$

其中, SSR 是指回归平方和, 即 y 的估计值与实际观察到的值的平均值之间做差, 并对差求平方和, SSE 则是值残差平方和, 即实际观察到的值与其对应的估计值之间做差, 并求差的平方和。在 H_0 的假设下, F 是服从自由度 $(P, n - P - 1)$ 的 F 分布, 然后就可以对回归方程进行显著性检验。

3.2. 模型的拟合与检验

3.2.1. 模型拟合

在 R 中, 使用 `lm()` 来对样本数据进行线性回归方程拟合, 结果见表 1 所示。

Table 1. The result of linear regression

表 1. 线性回归的结果

	估计值	方差	t 值	p 值
$\hat{\beta}_0$	3.5495	0.2212	16.0460	<0.000001
$\hat{\beta}_1$	0.3108	0.0252	12.3130	<0.000001
$\hat{\beta}_2$	0.219	0.0096	2.2740	0.0246
$\hat{\beta}_3$	0.0003	0.0015	0.1800	0.8575
$\hat{\beta}_4$	-0.1186	0.0590	-2.0110	0.0464
$\hat{\beta}_5$	0.0129	0.0153	0.837	0.4040
$\hat{\beta}_6$	0.0047	0.0122	0.3880	0.6989

由表可知, 并不是所有解释变量都能通过显著性检验, 尤其是解释变量 UFC、FER、COI, 这三个变量的 p 值是明显高于其它几个解释变量的, 也就意味着, 自变量对被解释变量 CPI 的影响并不显著, 所以接下来会通过逐步回归分析来对解释变量进行选择, 来进一步优化多元回归模型。

3.2.2. 模型的优化与解释

在 R 软件中, 使用 `step()` 方法对之前的线性回归进行逐步回归分析, 得到的回归分析结果见表 2 所示。

Table 2. Results of stepwise regression analysis

表 2. 逐步回归分析的结果

	估计值	方差	t 值	p 值
$\hat{\beta}_0$	3.6764	0.1103	33.3400	<0.000001
$\hat{\beta}_1$	0.3161	0.0243	13.0250	<0.000001
$\hat{\beta}_2$	0.0151	0.0025	6.1580	<0.000001
$\hat{\beta}_4$	-0.1574	0.2714	-5.8000	<0.000001

在 0.05 的水平下, 变量的显著性明显提高, 再看回归方程显著性检验, 根据逐步回归分析的结果 $F(3, 133) = 69.19$, 对应的 p 值小于 0.000001, 所以在 0.05 的水平下, 通过了回归方程的显著性检验, 再看方程的拟合优度检验, 也就是 R^2 , 在本例中 $R^2 = 0.6095$, 从数值上来看, 在一定程度上反映了被解释变量与解释变量之间的线性关系, 由于存在对被解释变量具有较大影响的解释变量未被考虑进模型中, 所以导致了拟合优度较小。

根据上述，可得最后的模型为公式 3.4

$$\widehat{CPI} = 3.6764 + 0.3161 \text{PPI} + 0.0151 \text{M2} - 0.1574 \text{ARE} \quad (3.4)$$

根据已经得到的多元线性回归模型，可知我国的通货膨胀按照其主要影响因素的不同，大体上可以分为三种：成本推动型通货膨胀、流动性过剩型通货膨胀和输入型通货膨胀。

从成本推动方面来说，产品生产链上由于原材料、劳动力成本等上游生产要素的价格上涨导致下游消费品的价格上涨，进而推高了成本推动型通货膨胀的压力。过剩的流动性是引起通货膨胀的重要原因，而在本文分析中，外汇储备并不是主要因素，也就是说我国的过剩流动性并不是通过外汇储备来显现的，而是通过汇率传导机制也就是央行按照汇率发行基础货币造成的过剩流动性，进而引起的通货膨胀。外部冲击同样也是引起我国通货膨胀的重要因素，也就意味着国际大宗商品例如原油的价格并不能直接影响我国的通货膨胀程度，反而更多的是通过人民币的有效汇率变动来对消费价格指数产生间接的负向效应。当国外出现通货膨胀和价格上涨时，会造成外汇相对于本币的贬值，而本币的升值又会导致本国对国外商品的需求增加，在短期内，进口又无法满足这部分需求，就会转而增加对本国商品的消费，最终引起整个社会总需求的增加，进而导致需求拉升型的通货膨胀，其中最需要留意的是，一个国家的汇率安排将会决定这个传导作用的大小，也就是由国家的汇率浮动程度所决定，因此，我国的通货膨胀不能集中的理解为输入型通货膨胀，尽管如此，也不能忽略汇率的作用。

4. 总结

控制流动性最直接、见效最快的手段包括调节法定存款准备金率、公开市场操作(操作对象包括央行票据和有价证券，以央行票据为主)、再贴现等措施。在预期通货膨胀增强的条件下，调节货币供应量并结合高利率的措施，进而达到降低物价总水平的目的，而这正是由本文建立的多元线性回归模型中，消费价格指数与广义货币供应量之间存在正相关的关系得出的结论，通过调节货币供应量和利率对调节物价是有效的。

在应对通货膨胀的外部冲击上，国际大宗原材料的价格变化会是我们的首先关注对象，特别是能源产品价格，通过参与国际大宗商品定价权的竞争，尽可能地减少外部冲击，同时也尽可能在国内实行新型绿色能源来替代严重依赖进口的石油能源；其次，汇率浮动制度是能够在国际收支失衡时发挥调节作用的重中之重，所以坚持有管理的浮动汇率制度也是增强我国抵御输入型通货膨胀的有力手段；再者，我国也要有计划地制定和实行稳定的人民币升值计划，管理汇率预期，严防人民币升值导致的货币替代效应给通货膨胀带来压力，尽可能化解国内的价格上涨。

最后，要对上游生产要素的价格例如原材料、劳动力成本加大监测力度，建立健全关于工业生产者出厂价格相关的预测预警机制，防范因为成本提高而引发的成本推动型通货膨胀的风险。同时针对成本推动型的通货膨胀，调控供给应该是主要的治理政策。

本文在 2008 年 01 月至 2019 年 05 月的月度数据基础上，产生了多元线性回归模型，确定能显著性影响通货膨胀的因素是货币和准货币供应量、工业生产者出厂价格指数和月内人民币平均汇率，并且通过检验，模型是合理有效的，能够为政策制定者的相关决策提供科学的参考依据。

参考文献

- [1] Friedman, M. and Schwartz, A.J. (1963) A Monetary History of the United States, 1867-1960. Princeton University Press, Princeton, New Jersey.
- [2] 易纲. 中国货币供求与通货膨胀[J]. 经济研究, 1995(5): 51-58.
- [3] 方先明, 裴平, 张谊浩. 外汇储备增加的通货膨胀效应和货币冲销政策的有效性[J]. 金融研究, 2006(7): 13-21.

- [4] 王智勇. 中国的经济增长与通货膨胀: 1981-2003[J]. 山东社会科学, 2008(1): 93-98.
- [5] 黄益平, 王勋, 华秀萍. 中国通货膨胀的决定因素[J]. 金融研究, 2010(6): 46-59.
- [6] 韩学红, 郑妍妍, 伍超明. 对我国股票收益率与通货膨胀率关系的解释: 1997-2007[J]. 金融研究, 2008(4): 21-36.
- [7] 刘亚, 李伟平, 杨宇俊. 人民币汇率变动对我国通货膨胀的影响: 汇率传递视角的研究[J]. 金融研究, 2008(3): 28-41.
- [8] 张成思. 中国 CPI 通货膨胀率子成分动态传导机制研究[J]. 世界经济, 2009(7): 3-12.