

中国主要宏观变量的非参数稳定性检验^①

蔡楠¹, 蔡宗武^{1,2,3,4}, 方颖^{1,3,4}

(1.厦门大学王亚南经济研究院, 福建厦门 361005, 2. 美国北卡莱娜大学夏洛特校区, 3. 教育部计量经济学重点实验室(厦门大学), 4.福建省统计科学重点实验室)

摘要: 基于非参数广义 F 检验与 Wild Bootstrap 的方法, 本文对从 1997 年 1 月到 2010 年 12 月共 92 个中国主要月度宏观经济变量的双变量关系进行稳定性检验。检验结果表明, 在所有 8370 组双变量关系中, 有超过 70% 的比例是显著不稳定的。

关键词: 非参数时变系数模型; 稳定性检验; LASSO 型的模型选择; Wild Bootstrap

0 引言

在宏观经济分析与预测中, 一个首先需要考虑的问题是检验所研究的宏观经济变量之间的关系是否具有稳定性, 即是否存在时变性的结构变化。各种经济性的和非经济性的因素都可能导致宏观经济模型表现出时变性的特征。比如, 消费者偏好的改变会导致消费需求、储蓄行为和劳动投入的改变, 进而影响厂商生产、投资等行为的变化。制度的变迁、政策的更替和技术进步也会对经济产生深远影响, 进而改变经济结构。而一些重大的政治事件和严重自然灾害的爆发也会使经济偏离原来发展的轨迹。特别是在时间跨度较长的情况下, 这种变化更为显著。很多研究也证明这种经济结构的不稳定性确实是普遍存在的。Stock 和 Watson^[1]以美国战后 76 个具有代表性的月度宏观时间序列数据为样本, 采用了多种不同的稳定性检验方法, 发现在 5700 组宏观经济变量关系中, 有半数以上的关系是显著不稳定的。Ben-David 和 Papell^[2]研究了 74 个国家的战后 GDP 数据, 发现 54 个国家的经济增长率出现断点。McConnell 和 Perez-Quiros^[3]则检验了美国 GDP 增长率数据波动率的稳定性, 发现在 1984 年左右波动率大幅下降。Hansen^[4]在美国制造业劳动生产率数据中发现了存在多个结构断点的经验证据。

中国是世界上最大的经济转型体, 宏观经济政策和宏观经济调控手段的变化以及不断深化的内部市场一体化和国际化进程都会导致宏观经济结构发生深刻的变化。但是, 在利用各种中国宏观经济变量的实证研究中, 经济模型的时变性检验却往往不受到重视。比如说, 向量自回归模型和结构向量自回归模型以及相应的格兰杰因果检验(Granger causality test)和脉

^①基金项目: 国家自然科学基金项目(70971113, 70871003)、国家自然科学基金重点项目(71131008)、中央高校基本科研业务费(2010221092, 2011221015)。

通讯作者: 方颖(1973-), 男, 上海市人, 博士, 助理教授. Email: yifst1@xmu.edu.cn。

冲分析在宏观经济分析与预测的实证研究中得到了越来越广泛的应用。但是，正确使用上述方法的一个前提是经济模型的结构稳定性。忽视宏观经济模型中存在的时变性特点会导致模型误设，从而在政策分析和宏观预测中得出误导性的结论。因此，在建立宏观经济模型时，我们首先要做的工作就是检验宏观经济变量之间相互关系是否存在时变性的结构变化。

本文的主要工作是利用最新发展起来的非参数统计检验方法，系统地检验从 1997 年 1 月至 2010 年 12 月 92 个中国主要宏观经济变量之间相互关系的稳定性，从而为进一步的宏观经济分析、实证模型选择和宏观经济预测提供基础性的研究工作。方颖和郭萌萌^[5]对 1997 年 1 月至 2006 年 11 月间包括产出、消费、投资、价格指数、汇率、短期利率、货币供给、金融市场、进出口贸易和政府财政等十个系列共计 83 个中国主要宏观变量进行了稳定性检验。与方颖和郭萌萌的工作相比，本文的研究包含了更完整的宏观经济变量和样本时间段。更重要的是，方颖和郭萌萌的稳定性检验是建立在单变量基础上的，换言之，即检验某一宏观时间序列本身的自相关结构是否稳定，而本文所提出的稳定性检验是建立在双变量基础上的，即检验宏观变量两两之间的相互关系是否具有时变性。Stock 和 Watson 认为，双变量关系在稳定性检验中具有更一般的意义，如果双变量关系呈现出不稳定的性质，那么更高阶的多元时间序列模型也是不稳定的。从单变量检验扩展到双变量检验，原假设模型中所包含的时滞变量会成倍增加，Stock 和 Watson、方颖和郭萌萌等通过反复使用 AIC(Akaike Information Criterion)或 BIC(Bayesian Information Criterion)模型选择准则来选择时滞变量的方法会带来较大的模型选择误差 (data snooping problem)。在本文中，我们使用最新发展起来的 LASSO (Least Absolute Shrinkage and Selection Operator; 见 Tibshirani^[6])的方法构建原假设下的双变量线性自回归模型，同时进行模型选择和模型估计，避免了传统模型选择方法所可能带来的不一致性，从而提高了稳定性检验的效率和可靠性。

本文将稳定性检验建立在比较线性时间序列模型和非参数时变模型的基础上的广义 F 统计量，并通过 wild bootstrapping 的方法构建统计量的分布情况。传统的稳定性检验方法大部分都是针对突变式或跳跃式的结构变化，而无法有效地检验缓慢连续的结构变化，但后者在结构变化中往往更为普遍。Hansen 就指出，“结构突变似乎不太可能，而更合理的是结构变化在一段时间内发生”。和传统的稳定性检验方法相比，我们的方法不仅能够检验结构突变，而且能够更有效地捕捉到缓慢的连续性结构变化。不管原假设是否成立，非参数时变模型总能得到一致性的估计量，从而能够有效提高稳定性检验的势能 (power)。由于使用了 bootstrapping 的方法，我们的稳定性检验也具有更好的小样本性质。

本文的第二部分详细论述我们所使用的稳定性检验的方法，包括模型选择、模型估计、

统计量构造；第三部分介绍本文所使用的数据和数据的处理方法以及实证结果；最后则是一个简短的结论。

1 稳定性检验方法

在本文中，我们利用 Cai, Fan 和 Yao^[7]和 Cai 和 Tiwari^[8]提出的广义 F 统计量来检验宏观变量相互关系的稳定性，主要思想是基于比较固定系数线性参数模型和与之相对应的非参数时变系数模型的残差平方和。当原假设（即稳定性）成立时，参数固定系数模型和非参数时变系数模型的估计结果比较接近。当两者的差异较大时，我们拒绝原假设，从而认为这两个变量之间的关系是不稳定的（即时变的）。具体来说，首先估计固定系数参数模型，计算其残差平方和。然后用局部线性回归（local linear estimation）的非参数方法估计与之对应的时变模型，得到其残差平方和。通过用这两个模型得到的残差平方和构造一个广义 F 统计量，并用 wild bootstrap 的方法求得该统计量的样本分布，从而可以计算其 p 值，以此进行统计推断。

1.1 原假设模型的选择

对任意两个时间序列变量 $\{x_t\}_{t=1}^T$ 和 $\{y_t\}_{t=1}^T$ ，我们考虑一个一般化的固定系数线性参数模型：

$$y_t = \mu + \alpha(L)y_{t-1} + \beta(L)x_t + \varepsilon_t. \quad (1)$$

其中 μ 是固定不变的常数，而 $\alpha(L)$ 和 $\beta(L)$ 分别是 p 阶和 q 阶的滞后多项式，即 $\alpha(L) = \alpha_0 + \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_p L^p$ ， $\beta(L) = \beta_0 + \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q$ ， L 代表滞后算子， $\{\varepsilon_t\}$ 是一个序列不相关的随机扰动项。在模型（1）中， $\alpha(L)$ 和 $\beta(L)$ 的所有系数（即 α_i 和 β_j ）都是固定不变的常数，意味着 x 和 y 两个变量之间的关系是稳定的。模型（1）代表了原假设成立时的真实模型。

在构造固定系数模型（1）的过程中，一个重要的问题是如何选择最大滞后阶数 p 和 q ，以及如何选取合适的滞后项。主要困难在于：第一、解释变量的可能滞后项数目较多；第二、滞后项之间可能存在较严重的自相关性。传统的模型选择方法会导致系数的估计量方差过大，从而降低估计的精度和效率。为了解决上述问题，我们采用 Tibshirani 提出的 LASSO

方法来进行模型的选择和估计。LASSO 的估计值可表示为：

$$(\hat{\mu}, \hat{\alpha}_1, \dots, \hat{\alpha}_p, \hat{\beta}_1, \dots, \hat{\beta}_q) = \arg \min \left\{ \sum_{t=1}^T \left(y_t - \mu - \sum_{j=1}^p \alpha_j y_{t-j} - \sum_{l=1}^q \beta_l x_{t-l+1} \right)^2 \right\} \quad (2)$$

$$\text{满足 } \sum_{j=1}^p |\alpha_j| + \sum_{l=1}^q |\beta_l| \leq c$$

其中 $c \geq 0$ 为调节参数 (tuning parameter)，用来控制系数压缩的程度。我们用

$(\hat{\mu}^0, \hat{\alpha}_1^0, \dots, \hat{\alpha}_p^0, \hat{\beta}_1^0, \dots, \hat{\beta}_q^0)$ 来表示无约束的 OLS 的估计量。令 $c_0 = \sum_{j=1}^p |\hat{\alpha}_j^0| + \sum_{l=1}^q |\hat{\beta}_l^0|$ 。当

$c < c_0$ 时，(2) 式系数的估计值会被向 0 压缩。当样本量趋向于无穷大时，不显著的系数估计值就会被压缩到 0，那么它所对应的滞后项就从模型中被自动删除了。c 值越大，模型中的变量就越多；而 c 越小，压缩程度就越大，所选择的变量就越少。因此，LASSO 方法关键在于调节参数 c 的选取。文献中关于调节参数的选取及回归系数估计的算法很多，本文采用最小角度回归法 (Least Angle Regression, 记为 LARS) (Efron, Hastie, Johnstone 和 Tibshirani^[9])。

和文献中常用的子集选择法 (AIC 或 BIC) 相比，LASSO 具有以下优点：第一、LASSO 方法是变量选择和参数估计同步进行，因此效率高，计算量小，这一点对高维的情形尤其明显；第二、子集选择法是一个离散的无序过程，变量或者被保留或者被丢弃，因此不能降低整个模型的预测误差。与子集选择法不同，LASSO 的变量选择是一个连续有序的过程，对数据的变化没有子集选择法那么敏感，因此具有良好的稳定性，而且方差较小；第三、对同一数据反复使用 AIC 和 BIC 准则，子集选择法会导致数据挖掘偏误。而 LASSO 方法不存在这个问题。

采用最小角度回归法解 (2) 式，我们就可以得到 LASSO 的估计值 $(\hat{\mu}^*, \hat{\alpha}_1^*, \dots, \hat{\alpha}_p^*, \hat{\beta}_1^*, \dots, \hat{\beta}_q^*)$ ，其中一些估计值为 0，这意味着对应的变量从模型中删除了，因此我们获得了在原假设成立条件下的固定系数模型。在不引起混淆的情况下，为了表述的简单，我们将经 LASSO 选择的模型记为：

$$y_t = \sum_{j=0}^k \gamma_j Z_{jt} + \varepsilon_t \quad (3)$$

其中 y_t 于原模型 (1) 中的 y_t 一致，而变量 $Z_t = (1, Z_{1t}, \dots, Z_{kt})$ 由 $(\hat{\mu}^*, \hat{\alpha}_1^*, \dots, \hat{\alpha}_p^*, \hat{\beta}_1^*, \dots, \hat{\beta}_q^*)$

中参数估计值不为 0 所对应的 $\{x_t\}_{t=1}^T$ 和 $\{y_t\}_{t=1}^T$ 的滞后变量组成, k 为新模型变量的个数, γ_j 为相应的固定系数参数。

1.2 非参数时变模型的估计

在模型 (3) 的基础上, 我们构造对应的时变系数模型:

$$y_t = \sum_{j=0}^k \gamma_j(t) Z_{t,j} + \varepsilon_t. \quad (4)$$

与固定系数模型 (3) 相比, 时变系数模型中所有的参数 γ_j 都是时间 t 的一个连续函数。换言之, 在不同的时间点 t , 模型 (4) 允许有不同的 γ_j 。为了取得渐进一致性的估计, 根据 Cai^[10], 我们需要假设:

A1 $E(\varepsilon_t | Z') = 0$, 其中 $Z' = \{Z_1, Z_2, \dots, Z_k\}$;

A2 $E(\varepsilon_t^2 | Z') = \sigma^2(Z')$, 即容许存在异方差, 且异方差是 Z' 的函数;

A3 $\{(y_t, x_t)\}$ 是严格平稳的 α -混合型序列(strictly stationary α -mixing);

A4 $\{\varepsilon_t\}_{t=1}^T$ 是一个相互独立的时间序列^②;

A5 $\gamma_j(\bullet)$ 是二阶连续可导的。

由于时间 $t = 1, \dots, T$, 是离散序列, 根据 Robinson^[11]和 Robinson^[12]的建议, 在实际估计时, 我们重新定义时间变量 $t_i = i/T$, $i = 1, 2, \dots, T$, 其中 T 是样本量。经过变换后, γ_j 成为 t_i 的函数。当 T 趋向于无穷大时, t_i 在 $[0, 1]$ 区间上的分布变得更加密集, 从而确保了非参数估计的渐进一致性。根据假设 (A5), $\gamma_j(\bullet)$ 具有连续的二阶导数, 那么对任意一个时间点 $s \in [0, 1]$, 我们可以对 $\gamma_j(t_i)$ 在时间点 s 对其泰勒展开, 作为 $\gamma_j(t_i)$ 的近似值, 即:

$$\gamma_j(t_i) = a_j + b_j(t_i - s), \quad 0 \leq j \leq k. \quad (5)$$

局部线性回归的估计量被定义为 $\hat{\gamma}_j(s) = \hat{a}_j$, 而 \hat{b}_j 为其在 s 点的一阶导数的估计量。

$(a_j(s), b_j(s))$ 的估计值为:

^②为简单起见, 我们假设 $\{\varepsilon_t\}_{t=1}^T$ 是相互独立的时间序列。如果 $\{\varepsilon_t\}_{t=1}^T$ 假设为一个严格平稳的混合型序列 (Cai, 2007) 任能满足渐进一致性。

$$(\hat{a}_j(s), \hat{b}_j(s)) = \arg \min \sum_{i=1}^T \left\{ Y_i - \sum_{j=1}^k (a_j + b_j(t_i - s)) Z_{i,j} \right\}^2 K_h(t_i - s). \quad (6)$$

其中 $K_h(t_i - s) = h^{-1} K\left(\frac{t_i - s}{h}\right)$, $K(\bullet)$ 是一个核函数 (kernel function) ^③。 $h > 0$ 代表窗宽 (bandwidth), 它满足以下条件: 当 $T \rightarrow \infty$ 时, $h \rightarrow 0$ 并且 $hT \rightarrow \infty$ 。通过加权局部最小二乘法, 我们可以得到非参数时变模型的参数估计值 $\{\gamma_j(t_i)\}_{i=1}^T$, $j = 1, \dots, k$ 。

1.3 稳定性检验

广义 F 检验是建立在比较固定系数模型 (3) 和时变系数模型 (4) 的残差平方和的基础上的。模型 (3) 的残差平方和定义为:

$$RSS_0 = T^{-1} \sum_{i=1}^T \left(y_i - \sum_{j=0}^k \hat{\gamma}_j Z_{i,j} \right)^2 \quad \text{④}。$$

类似地, 模型 (4) 的残差平方和定义为:

$$RSS_1 = T^{-1} \sum_{i=1}^T \left(Y_i - \sum_{j=0}^k \hat{\gamma}_j(t_i) Z_{i,j} \right)^2。$$

广义 F 统计量则定义如下:

$$T_n = (RSS_0 - RSS_1) / RSS_1 = RSS_0 / RSS_1 - 1. \quad (7)$$

原假设定义为:

$$H_0: \gamma_j(t) = \gamma_j, \quad 1 \leq j \leq k。$$

在原假设下, 所有模型的系数都是不随时间变化的固定系数, 即变量之间的关系是稳定的。当 T_n 的值过大时, 拒绝原假设, 认为两变量之间的关系是不稳定的。Chen 和 Hong^[13] 证明在 T 趋向于无穷大时, 经过标准化的 T_n 服从标准正态分布。为了取得较理想的小样本性质而且容许模型 (4) 存在异方差, 我们采用非参数 wild bootstrap 的方法而不是传统的 bootstrap 方法来估计 T_n 的样本分布。具体步骤如下:

③ $K(\bullet)$ 是一个事先给定的对称的概率密度函数, 满足 (i) $\int_{-\infty}^{\infty} K(u) du = 1$, (ii) $\int_{-\infty}^{\infty} K(u) u du = 0$,

(iii) $\int_{-\infty}^{\infty} u^2 K(u) du = C_k < \infty$, (iv) $\int_{-\infty}^{\infty} K^2(u) du = D_k$ 。

④ 此处参数估计值 $\hat{\gamma}_j$ 为 LASSO 的估计值。

B1 估计非参数模型 (4)，得到其残差 $\{\hat{u}_t\}_{t=1}^T$ ， $\hat{u}_t = y_t - \sum_{j=1}^p \hat{\gamma}_{t,j} Z_{t,j}$ 。并将其中心化

$$\tilde{u}_t = \hat{u}_t - \bar{\hat{u}}_t, \text{ 其中 } \bar{\hat{u}}_t = T^{-1} \sum_{t=1}^T \hat{u}_t;$$

B2 构造 wild bootstrap 残差 $\{u_t^*\}_{t=1}^T$ ，其中 $u_t^* = \tilde{u}_t * \eta_t$ ， $\{\eta_t\}_{t=1}^T$ 是一组独立同分布的随机变量，服从标准正态分布^⑤；

B3 利用 (B2) 中得到的 wild bootstrap 残差 $\{u_t^*\}_{t=1}^T$ 来构造 bootstrap 样本 $\{y_t^*\}_{t=1}^T$ 。定义

$$y_t^* = \sum_{j=0}^k \hat{\gamma}_j Z_{t,j} + u_t^*, \text{ (}\hat{\gamma}_0, \hat{\gamma}_1, \dots, \hat{\gamma}_k\text{) 是常系数模型 (3) 的参数估计量}^{⑥};$$

B4 利用 wild bootstrap 所产生的样本 $\{Z_t, y_t^*\}_{t=1}^T$ 重新计算 bootstrap 的检验统计量 T_n^* ；

B5 重复步骤 (B2) 至 (B4) M 次 (比如，M 可取为 1000)，计算事件 $\{T_n^* \geq T_n\}$ 重复的频率，记为 p 。

当 p 值小于 0.01 时，拒绝原假设，认为参数模型是不稳定的，这意味着变量 x 和 y 之间存在着非线性或者非稳定的关系。

和其他的稳定性检验方法相比，本文所采用的方法具有以下优点：第一，这种方法不但可以检验结构性突变，而且可以检验缓慢的连续的结构变化；第二，此方法无需任何关于结构变化方式的先验信息，传统的检验方法需要假定断点的存在，判定断点的位置和个数；第三，这种方法在很弱的假定下具有良好的大样本性质，极限分布是标准正态分布；第四，在 wild bootstrap 的过程中我们使用非参数模型的残差估计值。无论原假设是否成立，我们所使用的残差都具有相合性，从而使我们的检验获得更好的势能 (power)；最后，通过 wild bootstrap 的方法获取 T_n 的样本分布，可以得到较理想的小样本性质，而且容许模型存在异方差；第五，通过对时变模型系数估计值的进一步分析，我们可以判断经济关系结构变化的类型。

^⑤Kreiss, Neumann 和 Yao^[14]建议用上述 wild bootstrap 方法来取代传统的通过 $\{\hat{u}_t\}_{t=1}^T$ 的经验分布来获取 $\{u_t^*\}_{t=1}^T$ 的方法。

^⑥在求 $\{y_t^*\}$ 的过程中，由于 $\{Z_t\}$ 包含 y_t 的滞后项，我们用 $\{y_t\}$ 的前几个初值来计算。通过对 (3) 式反复迭代，产生一系列 $\{y_t^*\}$ 。

^⑦这里用 LASSO 的估计值。

2 数据处理与检验结果

本文选取了 92 个月度宏观时间序列数据，所有数据均来自中国经济数据库（CEIC）。这些数据可以划分为十类——产出、消费、投资、进出口、价格指数、短期利率、汇率、货币、金融市场和政府财政，时间跨度为 1997 年 1 月到 2010 年 12 月。部分宏观时间序列数据同时包含了实际值和名义值，其中实际值以 1997 年 1 月为基期的 CPI 进行调整。

在进行稳定性检验之前，我们对数据做了以下调整：第一，我们对除了价格指数、上证综指收益率和深证综指收益率以外的所有数据都进行了对数变换；第二，对产出、消费、投资、进出口、货币、政府财政等具有明显季节性因素的时间序列，我们用 X12 对其进行季节调整；第三，我们对所有的数据进行了单位根检验，因为数据的平稳性是非参数时变系数模型渐进理论的基础。我们使用的方法是常用的 ADF 检验（Augmented Dicky-Feller test），遵循 Holdend 和 Permanent^[15]的序贯检验方法来确定模型是否含有截距项和时间趋势。我们采用 BIC 准则来确定 ADF 检验中的自回归滞后项的阶数[®]。此外，为了确保 ADF 检验结果的正确性，对每组时间序列数据我们还使用 KPSS 检验（Kwiatkowski, Phillips, Schmidt 和 Shin^[16]）验证结果。KPSS 检验和 ADF 检验所使用的模型假定完全一致。二者不同之处在于原假设的设定不同，KPSS 检验的原假设是被检验的时间序列是平稳的，而 ADF 检验的原假设是被检验的时间序列是单位根过程。当 KPSS 检验与 ADF 检验的结果不一致时，我们仍把该时间序列数据视为一个单位根过程。对于存在单位根的时间序列数据，对其进行一阶差分使其平稳化。而对仅存在时间趋势的时间序列数据，通过对时间 t 进行回归消除时间趋势。具体的时间序列数据的名称以及所作的调整见表 1。

表 1 变量名称与数据处理

Table 1. The name of variable and data processing

消费			
名义零售总额【1】【2】	实际零售总额【1】【2】		
投资			
名义 FDI【1】【3】	实际 FDI【1】【3】	名义合资【1】【3】	实际合资【1】【3】
名义合作【1】【2】	实际合作【1】【2】	名义外资【1】【3】	实际外资【1】【3】
名义股份【1】【2】	实际股份【1】【2】	名义房产投资【1】【3】	实际房产投资【1】【3】

[®]AIC 通常会高估自回归的阶数，从而降低 ADF 检验的势能（power）。因此，于 AIC 准则相比，SIC 准则是更好的选择。

进出口			
名义进出口总额【1】【2】	实际净出口总额【1】【2】	名义进口总额【1】【2】	实际进口总额【1】【2】
名义出口总额【1】【2】	实际出口总额【1】【2】	名义进出口差额【1】【2】	实际进出口差额【1】【2】
价格指数			
CPI200501【2】	CPI 同比【2】	CPI 食品【2】	CPI 衣着【2】
CPI 家庭【2】	CPI 医保【2】	CPI 交通通讯【2】	CPI 娱乐【2】
CPI 居住【2】	PPI【2】	PPI 轻工业【2】	PPI 重工业【2】
PPI 生产资料【2】	PPI 生活资料【2】		
短期利率			
7 天同业拆借【2】	60 天同业拆借【2】	30 天同业拆借【2】	90 天同业拆借【2】
汇率			
美元【2】	日元【2】	港币【2】	
货币供给			
M0	【1】【2】	M1	【1】【2】
		M2	【1】【2】
金融市场			
上证综指收益率	深证综指收益率	名义国债交易额【1】【2】	实际国债交易额【1】【2】
政府财政			
名义财政收入【1】【3】	实际财政收入【1】【3】	名义财政支出【1】【3】	实际财政支出【1】【3】
名义财政赤字【1】【2】	实际财政赤字【1】【2】	名义税收收入【1】【3】	实际税收收入【1】【3】
产出			
原盐【1】【2】	成品糖【1】【2】	乳制品【1】【2】	纱【1】【3】
生丝【1】【2】	服装【1】【2】	汽油【1】【3】	煤油【1】【3】
柴油【1】【3】	焦炭【1】【2】	化学纤维【1】【2】	塑料制品【1】【3】
水泥【1】【2】	生铁【1】【2】	汽车【1】【2】	船舶【1】【2】
计算机【1】【2】	电视【1】【2】	音响【1】【3】	照相机【1】【3】
发电量【1】【2】	饲料【1】【2】	布【1】【2】	润滑油【1】【3】
硫酸【1】【3】	橡胶【1】【2】	钢材【1】【2】	拖拉机【1】【2】
轿车【1】【2】	摩托车【1】【3】	自行车【1】【2】	电冰箱【1】【2】
洗衣机【1】【2】	空调【1】【3】		

注：【1】表示进行季节调整，【2】表示做一阶差分，【3】表示消除时间趋势。

把我们收集到得 92 组时间序列数据经过上述步骤调整之后，用第二节介绍的方法进行双变量稳定性检验。每一组时间序列数据都可以分别作为解释变量和被解释变量，那么我们一共要进行 $92 \times 91 = 8372$ 次检验。

在估计非参数时变系数模型时，我们采用 Epanechnikov 核函数。Epanechnikov 核函数

在理论上是最优的 (Fan 和 Gijbels^[17]), 所以最常用。该核函数定义如下:

$$K(u) = \frac{3}{4}(1-u)^2 I(|u| \leq 1)$$

其中 $I(\bullet)$ 代表一个示性函数, 当括号里的条件为真时, 其值为 1, 否则为 0。

通过加权局部最小二乘法, 我们可以得到参数的估计值, 具体算法如下:

$$(\hat{\gamma}_0(s), \hat{\gamma}_1(s), \dots, \hat{\gamma}_k(s)) = e(\tilde{X}'W\tilde{X})^{-1} \tilde{X}'WY$$

其中 $e = (\underbrace{1, \dots, 1}_{1+p+q}, \underbrace{0, \dots, 0}_{p+q})$,

$$\tilde{X} = \tilde{X}(t) = \begin{pmatrix} 1 & Z_{1,1} & \dots & Z_{1,k} & Z_{1,1}(t_1-s) & \dots & Z_{1,k}(t_1-s) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & Z_{T,1} & \dots & Z_{T,k} & Z_{T,1}(t_T-s) & \dots & Z_{T,k}(t_T-s) \end{pmatrix},$$

$Y = (Y_1, \dots, Y_T)'$, 并且 $W = W(s) = \text{diag}\{K_h(t_1-s), \dots, K_h(t_T-s)\}$ 。那么, 通过上面的计

算我们可以得到 $\hat{Y} = (\hat{Y}_1, \dots, \hat{Y}_T)'$, 其中 $\hat{Y}_t = \sum_{j=1}^k \hat{\gamma}_{t,j} X_{t,j}$ 。 $\hat{Y} = (\hat{Y}_1, \dots, \hat{Y}_T)'$ 还可以记为

$$\hat{Y} = H^* Y,$$

其中

$$H^* = (H_1, \dots, H_T)', \quad H_t = A(t_i)' Z_t$$

并且

$$A(t_i) = (I_{1+p+q}, 0) (\tilde{X}(t_i)' W(t_i) \tilde{X}(t_i))^{-1} \tilde{X}(t_i)' W(t_i), \quad Z_t = (1, Z_{t,1}, \dots, Z_{t,k})'$$

在文献中 H^* 被称为线性平滑算子 (linear smoothing operator) (Hastie 和 Tibshirani^[18])。

在本文中我们使用偏误修正的 AIC 准则 (bias-corrected AIC, 记为 AIC_C) 来选择窗宽 h (Cai 和 Tiwari):

$$h = \arg \min AIC_C(h) = \log \{RSS\} + \frac{T + \text{tr}(H^*)}{T - [\text{tr}(H^*) + 2]}. \quad (6)$$

其中 $RSS = \sum_{t=1}^T (Y_t - \hat{Y}_t)^2 = \sum_{t=1}^T \left(Y_t - \sum_{j=0}^k \hat{\gamma}_{t,j} Z_{t,j} \right)^2$, 这里的 H^* 前文定义一致。

双变量稳定性检验的结果列于表 A (见附录)。所有数据分为十类, 分别列出。表 A 中的变量表示我们所选取时间序列数据的名称, 百分比下的数值分为两类, 分别表示该变量作为解释变量 (即作为模型 (1) 的 y) 或是被解释变量 (即作为模型 (1) 的 x) 与其它 91

个变量所组成的双变量模型中，所得的 p 值小于等于 0.1 的比重。例如，表 A 中消费类下，名义零售总额后的第一个数 75.82% 表示，名义零售总额作为模型 (1) 的 y ，与其余 91 个变量作为模型 (1) 中 x 所组成的 91 个模型中， p 值小于 0.1 的百分比。由表 A 可以看出，一个变量作为被解释变量还是解释变量是有差别的，具有不对称性。而且对某些变量而言两者的差异还是很大的，比如产出中的船舶和电冰箱，以及投资中的实际合作和价格指数中的食品的 CPI。这意味着，两个变量的关系不是对等的，具有不对称性。哪一个变量做解释变量，哪一个变量做被解释变量对模型的稳定性有重要影响。但是，总体而言，两者基本没有差别。

表 B (见附录) 分别列出了十类数据在不同置信水平下不稳定双变量关系的比重。为了便于比较，我们还列出了在不同置信水平下所有变量不稳定双变量关系的比重，以此作为标准。从表 B 可以看出，在 10% 的置信水平下，我们检验的 8372 组双变量关系中有 72.63% 的比率是不稳定的，消费、财政和利率类数据的比重高于这个平均值，而汇率类的数据稳定性要好些，仅为 41.75% 远低于平均水平。在不同的置信水平下，同一类型的数据并不具有一致性。比如，消费类在 1% 的置信水平下数据高于平均水平，而在 5% 与 10% 的置信水平下却远高于平均水平。进出口、价格指数也存在类似情况。我们从检验结果中还发现，真实值和名义值对稳定性存在一定影响，在使用实际值的情况下稳定性略高。但是从表 C (见附录) 可以看出，在三种置信水平下两者的差别不大。

3 结论

本文用基于非参数的广义 F 检验对中国从 1997 年 1 月到 2010 年 12 月共 14 年 92 个主要的月度宏观数据，包括消费、价格、汇率、财政、金融和产出等变量，两两之间关系的稳定性进行检验。在我们所检验的 8372 组双变量关系中，在 10% 的置信水平下，有高达 72.36% 的比例存在不稳定性。与 Stock 和 Watson 的研究结果相比，他们用美国从 1959 年 1 月到 1993 年 12 月长达 34 年 76 个月度宏观时间序列数据为样本进行了稳定性检验，发现大约 55% 左右的双变量关系是不稳定的。两者相差 17.36%。如果能够得到更长的数据，差异会更大。与美国相比，中国的宏观时间序列数据的结构不稳定更加普遍，程度更严重。

两者存在较大差异的原因主要是中国经济正在从计划经济向市场经济过渡，处在一个经济的转型期。在这期间，消费者和生产者的行为方式发生深刻变化、宏观经济政策的目标不

断调整、宏观经济调控手段不断创新、国民经济统计方法和统计手段也不断变化，再加上其他一些制度创新因素，宏观经济经历巨大变化。而美国市场经济已经非常成熟，经济政策的制定和经济的调控手段相对规范科学。所以美国的经济比中国要稳定得多。

因此，用线性时间序列模型来拟合中国数据在大多数情况下并不合适。如果忽略中国时间序列数据的这种结构不稳定，而用线性模型来拟合数据，这种情况下所获得的估计值没有意义，统计推断被严重扭曲，预测也失去了准确性，所得到的政策建议也是不合理的。所以在研究中国问题时，各种非线性模型应该是更合适的选择。

参考文献:

- [1] Stock J H, Watson M W. Evidence on structural instability in macroeconomic time series relations [J]. *Journal of Business and Economic Statistics*, 1996, 14(3): 11-30.
- [2] Ben-David D, Papell D H. Slowdowns and meltdowns: postwar growth evidence from 74 countries [J]. *Review of Economics and Statistics*. 1998, 80(4): 561-71.
- [3] McConnell M M, Perez-Quiros G. Output fluctuations in the United States: What has changed since the early 1980s? [J]. *American Economic Review*, 2000, 90(5):1464-1476.
- [4] Hansen B E. The new econometrics of structural change: Dating breaks in U.S. labor productivity [J]. *Journal of Economic Perspectives*, 2001, 15(4): 117-128.
- [5] 方颖, 郭萌萌. 中国主要宏观变量的稳定性检验: 基于非参数估计与 Bootstrapping 的一个方法[J]. *世界经济文汇*, 2009.
Fang Y, Guo M. Testing instability in Chinese macroeconomic time series relation: Using nonparametric time-varying coefficient method [J]. *World Economic Papers*, 2009.
- [6] Tibshirani R J. Regression shrinkage and selection via the lasso [J]. *Journal of the Royal Statistical Society*, 1996, 58(1): 267-288.
- [7] Cai Z W, Fan J Q, Yao Q W. Functional-coefficient regression models for nonlinear time series [J]. *Journal of the American Statistical Association*, 2000, 95(451):941-956.
- [8] Cai Z W , Tiwari R C. Application of a local linear autoregressive model to BOD time series [J]. *Environmetrics*, 2000, 11(2): 341-350.
- [9] Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression [J]. *The Annals Statistic*, 2004, 32(2): 407-499.
- [10] Cai Z W. Trending time-varying coefficient time series models with serially correlated errors [J]. *Journal of Econometrics*, 2007, 136(1): 163-188.
- [11] Robinson P M. Nonparametric estimation of time-varying parameters [M], Berlin : Springer, 1989.
- [12] Robinson P M. Time-varying nonlinear regression [M], Berlin : Springer, 1991.
- [13] Chen B, Hong Y M. Testing for smooth structural changes in time series models via nonparametric regression. Working Paper, Department of Economics, Cornell University, 2007.
- [14] Kreiss J-P, Neumann M H, Yao Q. Bootstrap tests for simple structures in nonparametric time series regression. *Statistics and its Interface*, 2008,1(2):367-380.
- [15] Holdend D, Permanent R. Unit Roots and Cointegration for the Economist [M], New York: Saint Martin Press, 1994.
- [16] Kwiatkowski L G, Phillips P C B, Schmidt P, Shin Y. Testing the null of stationarity against the alternative of

a unit root: How sure are we that economic time series have a unit root? [J]. Journal of Econometrics, 1992, 54: 159–178.

[17] Fan J, Gijbels I. Local Polynomial Fitting and Its Applications [M]. London: Chapman and Hall, 1996.

[18] Hastie T J, Tibshirani R J. Generalized additive models [M]. London: Chapman and Hall, 1990.

Nonparametric Stability Tests on Chinese Main Macroeconomic Time Series

CAI Nan¹, CAI Zongwu^{1,2,3,4}, FANG Ying^{1,3,4}

1. The Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen, 361005, China;
2. Department of Mathematics and Statistics, University of North Carolina at Charlotte, Charlotte, NC 28223, United States of America;
3. MOE Key Laboratory of Econometrics (Xiamen University);
4. Fujian Provincial Key Laboratory of Statistical Science (Xiamen University).

Abstract: Using the generalized F test and the wild bootstrapping method, this paper tests the stability of bivariate relations among 92 Chinese main monthly macroeconomic time series. The results demonstrate that more than 70% of 8,372 bivariate relations are significantly instable.

Key Words: nonparametric time-varying coefficients model; stability test; LASSO; wild bootstrap

附录:

表格 A 稳定性检验结果

Table A Results of stability tests

变量名称	百分比 (%)		变量名称	百分比 (%)	
	被解释变量	解释变量		被解释变量	解释变量
消费					
名义零售总额	75.8	81.3	实际零售总额	91.2	80.2
投资					
名义 FDI	93.4	78.0	名义外资	89.0	80.2
实际 FDI	83.5	78.0	实际外资	85.7	83.5
名义合资	97.8	78.0	名义股份	80.2	86.8
实际合资	97.8	76.9	实际股份	75.8	84.6
名义合作	34.0	75.8	名义房产投资	52.7	81.3
实际合作	24.1	78.0	实际房产投资	80.2	80.2
进出口					

名义进出口总额	84.6	62.6	名义出口额	86.8	72.5
实际进出口总额	86.8	65.9	实际出口额	80.2	69.2
名义进出口额	58.2	64.8	名义进出口差额	78.0	70.3
实际进出口额	62.6	62.6	实际进出口差额	68.1	68.1

价格指数

CPI20050101	97.8	68.1	CPI 娱乐教育	87.9	68.1
CPI 同比	29.6	69.2	CPI 居住	49.4	85.7
CPI 食品	24.1	63.7	PPI	92.3	64.8
CPI 衣着	93.4	70.3	PPI 轻工业	98.9	64.8
CPI 家庭	84.6	70.3	PPI 重工业	58.2	68.1
CPI 医保	79.1	72.5	PPI 生产资料	74.7	71.4
CPI 交通通讯	72.5	64.8	PPI 生活资料	98.9	63.7

短期利率

7 天同业拆借	100	78.0	60 天同业拆借	96.7	71.4
30 天同业拆借	37.3	69.2	90 天同业拆借	100	70.3

汇率

美元	15.3	74.7	日元	13.1	74.7
港币	96.7	67.0			

货币供给

M0	28.5	75.8	M2	97.8	73.6
M1	95.6	81.3			

金融市场

上证综指收益率	91.2	79.1	名义国债交易额	40.6	64.8
深圳综指收益率	92.3	76.9	实际国债交易额	39.5	70.3

政府财政

名义财政收入	98.9	69.2	名义财政赤字	95.6	81.3
实际财政收入	91.2	69.2	实际财政赤字	96.7	79.1
名义财政支出	52.7	72.5	名义税收收入	98.9	71.4
实际财政支出	24.1	81.3	实际税收收入	62.6	70.3

产出

原盐	87.9	68.1	计算机	97.8	80.2
成品糖	94.5	94.5	电视	39.5	75.8
乳制品	86.8	72.5	音响	92.3	64.8
纱	85.7	64.8	照相机	84.6	62.6
生丝	37.3	76.9	发电量	54.9	76.9
服装	47.2	78.0	饲料	95.6	68.1
汽油	87.9	74.7	布	48.3	72.5

煤油	40.6	68.1	润滑油	54.9	73.6
柴油	52.7	80.2	硫酸	63.7	65.9
焦炭	97.8	70.3	橡胶	89.0	74.7
化学纤维	67.0	65.9	钢材	86.8	75.8
塑料制品	98.9	69.2	拖拉机	95.6	67.0
水泥	41.7	68.1	轿车	57.1	65.9
生铁	80.2	68.1	摩托车	79.1	78.0
汽车	82.4	70.3	自行车	80.2	72.5
船舶	14.2	54.9	电冰箱	10.9	72.5
洗衣机	84.6	71.4	空调	89.0	71.4

注：显著性水平为 10%

表格 B 十类数据的稳定性检验结果

Table C Results of stability tests in ten categories

变量类别	显著水平		
	1%	5%	10%
所有变量	34.6	59.5	72.6
消费	28.5	73.0	83.5
投资	39.9	64.5	74.5
进出口	32.6	59.4	75.6
价格指数	39.6	58.2	74.4
短期利率	50.3	74.4	82.9
汇率	11.3	35.5	41.7
货币供给	45.7	64.4	73.9
金融市场	19.7	54.9	65.9
政府财政	52.6	69.7	77.6
产出	28.3	55.5	70.8

表格 C 区别名义变量与实际变量的稳定性检验结果

Table C Results of stability tests for nominal and real variables

变量类别	显著水平		
	1%	5%	10%
名义变量	43.6	65.2	76.0
实际变量	33.8	60.7	71.9