

中国税收主要序列研究

董瑾杰*

【内容摘要】本文在研究中,运用向量误差修正模型(VECM),以月度数据为基础,对2001年1月至2006年3月期间我国各项主要税收进行分析,并研究增值税、营业税和所得税之间的长期稳定关系,对中国的税收进行短期预测。

【关键词】动态分析 时间序列 向量误差修正模型(VECM)

【中图分类号】F810.422 【文献标识码】A 【文章编号】1000-6052(2006)06-0033-04

一、税种简介和背景回顾

“十五”期间,我国税收收入实现历史性突破。税收年收入2001年、2003年、2004年分别突破15000亿元、20000亿元和25000亿元,2005年超过30000亿元,达到30866亿元。五年共入库税款109217亿元(不包括关税和农业税收,未扣除出口退税),年均增长19.5%,五年翻了一番多。税收增收额2001年、2003年分别超过2500亿元、3400亿元,2004年和2005年连续超过5000亿元。由于税收收入越来越多,而税收间的动态关系也为人们所重视,下面就先对三种税进行简单介绍:

增值税是对销售货物或者提供加工、修理修配劳务以及进口货物的单位和个人就其实现的增值额征收的一个税种。1993年12月13日,国务院发布了《中华人民共和国增值税暂行条例》,12月25日,财政部下发了《中华人民共和国增值税暂行条例实施细则》,于1994年1月1日起施行。实行增值税的优点:第一、有利于贯彻公平税负原则;第二、有利于生产经营结构的合理化;第三、有利于扩大国际贸易往来;第四、有利于国家普遍、及时、稳定地取得财政收入。

营业税是对有偿提供应税劳务、转让无形资产和销售不动产的单位和个人,就其营业收入额征收的一种税。1993年12月13日国务院发布《中华人民共和国营业税暂行条例》,1993年12月27日财政部颁布《中华人民共和国营业税暂行条例实施细则》,并从1994年1月1日起实行。营业税的特点:1、征收范围广,税源普遍;2、税收负担轻、税负均衡,较好地体现了公平税负的原则;3、政策明了,适用性强;

4、计算简单,操作方便,纳税人容易理解。

企业所得税是指对中华人民共和国境内的一切企业(不包括外商投资企业和外国企业),就其来源于中国境内外的生产经营所得和其他所得而征收的一种税。现行的企业所得税法主要是1993年12月23日国务院发布的《中华人民共和国企业所得税暂行条例》和1994年2月4日财政部发布的《中华人民共和国企业所得税暂行条例实施细则》。企业所得税的特点:1、征税对象是所得额;2、应税所得额的计算比较复杂;3、征税以量能负担为原则;4、实行按年计征、分期预缴的征税办法。

二、研究方法 with 实证分析

1.变量的选取与数据处理

本文在研究中以月度数据为基础,对2001年1月至2006年3月期间我国各项主要税收进行分析,并研究增值税、营业税和所得税之间的长期稳定关系,对中国的税收进行短期预测。其中增值税、营业税和所得税分别用各期实际征收额取对数代表。样本数据来自《中国统计年鉴》2001年到2006年相关各期以及中经网。

2.计量模型的建立与分析

基于现代协整理论,本文首先对时间序列变量进行平稳性检验,即是否存在单位根。然后,进行协整检验以确定变量间的长期均衡关系。据协整分析研究,在变量间不存在协整关系时,应用向量自回归模型(VAR)是有效的;如果变量之间存在协整关系,应用向量误差修正模型(VECM)比向量自回归模型更优。最后,根据协整关系的检验,在向量自回归模型(VAR)或向量误差修正模型(VECM)的基础上用预测方差分解方法研究变量间的动态关系。

*董瑾杰:厦门大学 财政系

(1) 时间变量序列的平稳性检验

采用 ADF (Augmented Dickey- Fuller test) 方法分别对我国增值税, (设为 LZZS)、营业税 (设为 LYYS)、所得税 (设为 LSDS) 进行单位根检验。对时序进行回归分析通常要进行单位根检验, 是检验时序是否平稳的方法。若序列是平稳的, 则可进行回归分析; 若序列不是平稳的, 即存在单位根, 就必须对其差分再进行回归。

通常采用 ADF 检验, 设

$$Y_t = \alpha + \beta Y_{t-1} + \epsilon_t \quad (a)$$

假设 ϵ_t 为白噪音。当 $\beta = 1$ 时 Y_t 为单位根过程; $\beta < 1$ 时 Y_t 为平稳过程。我们把上式两边减 Y_{t-1} 得到:

$$\Delta Y_t = \alpha + \beta Y_{t-1} + \epsilon_t \quad (b)$$

其中 $\beta = -1$ 。因此检验 Y 是否为单位根过程只需检验 α 是否为 0。然而传统的 t 统计量此时不是适用, 因为在 $H_0: \alpha = 0$ 的零假设下, 方程右边的 Y_{t-1} 是单位根过程, 从而 t 统计量不再服从 t 分布, 而是 DF 分布。通常在进行 ADF 检验时还会加入 Y_t 的滞后项, 即

$$\Delta Y_t = \alpha + \beta Y_{t-1} + \sum_{i=1}^p \gamma_i \Delta Y_{t-i} + \epsilon_t \quad (c)$$

$$H_0: (\alpha = 0 \text{ (或 } \beta = 1))$$

$$H_1: (\alpha < 0 \text{ (或 } \beta < 1))$$

检验统计量 Y_{t-1} 的 t 统计量, 但不服从 t 分布, 而是服从 DF 分布若 $t < t^*$, 则拒绝零假设 (说明是平稳的); 否则无法拒绝零假设 (说明存在单位根)。在表一中显示了对变量 LZZS、LYYS、LSDS 进行单位根检验的结果:

表一 对变量单位根的 ADF 检验

变量	ADF 统计量	关键值 (5%)	关键值 (1%)
LSDS	-2.444087	-2.910860	-3.544063
D(LSDS)	-8.876548*	-2.919952	-3.565430
LYYS	0.160561	-2.913549	-3.550396
D(LYYS)	-13.04282*	-2.919952	-3.565430
LZZS	0.899311	-2.913549	-3.550396
D(LZZS)	-7.230494*	-2.919952	-3.565430

注: *表示在 5% 和 1% 水平下均为统计显著

平稳性检验结果显示, LZZS、LYYS、LSDS 均只有一个单位根, 都为 I (a) 过程, 即一阶平稳的。原变量都是非平稳的, 而一阶差分均为平稳的, 满足预测方差分解对变量一阶差分平稳性要求。

(2) 协整性关系检验 (Cointegration)

在单位根检验的基础上, 本文采用 Johansen 极大

似然估计法进行协整检验。协整检验的目的是决定一组非稳定序列是否是协整的。考虑阶数为 p 的 VAR 模型:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + Bx_t + \epsilon_t \quad (d)$$

其中, y_t 是一个含有非平稳的 I (a) 变量的 k 维向量; x_t 是一个确定的 d 维的向量, ϵ_t 是扰动向量。我们可把 VAR 重写为以下形式:

$$\Delta y_t = \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + Bx_t + \epsilon_t \quad (e)$$

其中: $\Pi = \sum_{i=1}^p A_i - I^k$, $\Gamma_i = -\sum_{j=i+1}^p A_j$

Granger 定理指出: 如果系数矩阵 Π 的秩 $r < k$, 那么存在 $k \times r$ 阶矩阵 α 和 β , 它们的秩都是 r , 使得 $\alpha\beta' = \Pi$, 并且 $\beta'y_t$ 是稳定的。其中 r 是协整关系的数量 (协整秩) 并且 $\beta'y_t$ 的每列是协整向量。正如下面解释, α 中的元素是向量误差修正模型 VEC 中的调整参数。Johansen 方法是在无约束 VAR 的形式下估计 Π 矩阵, 然后求出 α 和 β , 从而检验出协整秩, (秩 ($r < k$)), 得出协整向量。

Johansen 协整检验结果的解释: 对于每一个检验结果, 第一列显示了在原假设成立条件下的协整关系数; 第二列是 (d) 式中矩阵按由大到小排序的特征值; 第三列是最大特征值统计量; 最后两列分别是在 5% 水平下的临界值和 P 值。

最大特征值统计量的检验结果表明, 它所检验的原假设是有 r 个协整关系, 反之, 有 $r+1$ 个协整关系。统计量是按下面的方法计算的:

$$LR_{\max}(r | r+1) = -T \log(1 - \lambda_{r+1})$$

$$= LR_r(r | k) - LR_r(r+1 | k) \quad r = 0, 1, \dots, k-1 \quad (f)$$

表二 变量间协整关系的 Johansen 检验结果

假设的协整关系数	最大特征值	最大特征值统计量	5% 临界值	P 值
0*	0.300913	21.47885	21.13182	0.0447
最多 1	0.185862	12.33752	14.26460	0.0986
最多 2	0.000869	0.052155	3.841466	0.8193

注: *代表在 5% 的显著性水平上拒绝零假设

所以表二显示 LZZS、LYYS、LSDS 存在协整关系, 它可表示为:

$$VECM = LZZS - 1.118507LYYS + 0.186745LSDS \quad (g)$$

$$(0.07052) \quad (0.05565)$$

(g) 式反映了三个变量之间的长期均衡关系, 当然我们可以把任何一个变量写在左边, 小括号里的数字为渐进标准误差。从估计的结果看, 增值税和营业税所得税系数符号是有经济意义的, 表明了增值税和营业税都是间接税, 营业税的长期增长也同时说明增值税的增长; 而在长期所得税却对增值税有相反的影响。

$$D(LZZS) = 0.0304315908 * (t-1) - 0.0423145497$$

7*D (LZZS (-1)) -0.3889657524*D (LZZS (-2)) - 0.1390664518*D (LYYS (-1)) +0.07842041082*D (LYYS (-2)) +0.02259752299*D (LSDS (-1)) + 0.008000754526*D (LSDS (-2)) +0.01996946686

协整关系对增值税的上涨起到了正向修正作用。当增值税超出长期均衡约束，则误差修正作用提高当期增值税。

(3) 预测方差分解

方差分解是把内生变量中的变化分解为对 VAR 的分量冲击。因此，方差分解给出对 VAR 中的变量产生影响的每个随机扰动的相对重要性的信息。

方差分解的基本思路：

$$y_{it} = \sum_{j=1}^k (\psi_{0,j} \varepsilon_{jt} + \psi_{1,j} \varepsilon_{j,t-1} + \psi_{2,j} \varepsilon_{j,t-2} + \psi_{3,j} \varepsilon_{j,t-3} + \dots) \quad (h)$$

其中 k 是变量个数。

(h) 式中各括号中的内容是第 j 个扰动项 ε_j 从无限过去到现在时点对第 i 个变量 y_i 影响的总和。求其方差，因为 $\{\varepsilon_{jt}\}$ 无序列相关，故

$$E[(\psi_{0,j} \varepsilon_{jt} + \psi_{1,j} \varepsilon_{j,t-1} + \psi_{2,j} \varepsilon_{j,t-2} + \dots)^2] = \sum_{q=0}^{\infty} (\psi_{q,j})^2 \sigma_{\varepsilon_j} \quad (i)$$

j=1,2,...,k

这是把第 j 个扰动项对第 i 个变量的从无限过去到现在时点的影响，用方差加以评价的结果。此处还假定扰动项向量的协方差矩阵是 Ω 对角矩阵。于是 y_{it} 的方差是上述方差的 k 项简单和

$$\text{var}(y_{it}) = r_{ii}(0) = \sum_{j=1}^k \left\{ \sum_{q=0}^{\infty} (\psi_{q,j})^2 \sigma_{\varepsilon_j} \right\} \quad (j)$$

y_{it} 的方差可以分解成 k 种不相关的影响，因此为了测定各个扰动相对的 y_{it} 方差有多大程度的贡献，定义了 RVC (Relative Variance Contribution) (相对方差贡献率)，根据第 j 个变量基于冲击的方差对 y_{it} 的方差的相对贡献度来作为观测第 j 个变量对第 i 个变量影响的尺度。实际上，不可能用直到 $s=\infty$ 的 $\psi_{k,i,j}$ 来评价，只需有限的 s 项。

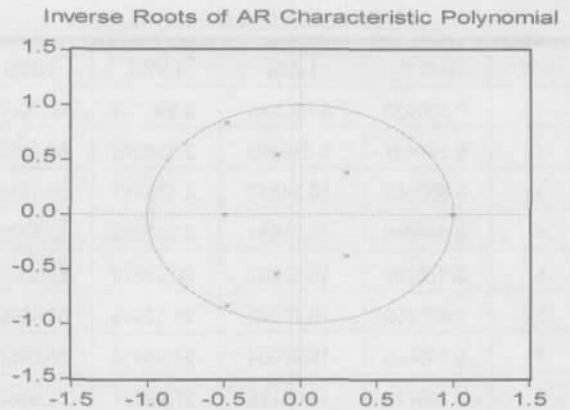
$$RVC_{j \rightarrow i}(s) = \frac{\sum_{q=0}^{s-1} (\psi_{q,j})^2 \sigma_{\varepsilon_j}}{\sum_{j=1}^k \left\{ \sum_{q=0}^{s-1} (\psi_{q,j})^2 \sigma_{\varepsilon_j} \right\}} \quad (k)$$

i,j=1,2,...,k

如果 $RVC_{j \rightarrow i}(s)$ 大时，意味着第 j 个变量对第 i 个变量的影响大，相反地， $RVC_{j \rightarrow i}(s)$ 小时，可以认为第 j 个变量对第 i 个变量的影响小。

VAR 模型所有根模的倒数小于 1，位于单位圆内，则其是稳定的。

从对 LZZS 的预测方差分解可以看到，LZZS 的



表三关于 LZZS 的预测方差分解

时期	标准差	LZZS	LYYS	LSDS
1	0.064675	100.0000	0.000000	0.000000
2	0.085198	94.12414	2.959869	2.915995
3	0.091816	93.96007	3.472495	2.567431
4	0.105375	95.29872	2.722894	1.978383
5	0.114444	95.56229	2.753718	1.683995
6	0.121721	95.58580	2.670996	1.743200
7	0.131503	96.10904	2.314859	1.576099
8	0.138351	96.22967	2.341571	1.428756
9	0.144703	96.27748	2.278208	1.444312
10	0.152995	96.57997	2.052450	1.367580

表四关于 LYYS 的预测方差分解

时期	标准差	LZZS	LYYS	LSDS
1	0.133110	50.74574	49.25426	0.000000
2	0.165468	65.27214	32.20132	2.526544
3	0.166204	64.89673	32.34693	2.756337
4	0.194369	59.95996	34.15851	5.881532
5	0.210487	65.72043	29.14255	5.137016
6	0.212484	66.16314	28.71262	5.124248
7	0.232831	62.87773	30.66269	6.459578
8	0.245011	66.40074	27.69141	5.907843
9	0.247313	66.82176	27.32211	5.856136
10	0.263759	64.72396	28.51337	6.762670

变化大部分可以由它自己的新息解释，而 LYYS、LSDS 对 LZZS 的解释能力比很小。与此相比，从对 LYYS 预测方差分解可以得出，LYYS 的增长从第 1 期有一半可以由 LYYS 的自身的新生来解释而另一半由 LZZS 来解释，之后 LZZS 对 LYYS 的解释能力越来越强。从对 LSDS 预测方差分解可以得出，一开始 LSDS 大部分变化可以由它自己的新息来解释，而后

表五关于 LSDS 的预测方差分解

时期	标准差	LZZS	LYYS	LSDS
1	0.536023	6.893540	0.996734	92.10973
2	0.547407	8.819405	2.855576	88.32502
3	0.560663	10.34847	4.475411	85.17612
4	0.648846	12.21654	20.23363	67.54983
5	0.673709	16.52207	20.22314	63.25479
6	0.680759	16.30520	21.50488	62.18992
7	0.743699	16.47004	27.96417	55.56579
8	0.766723	19.63376	27.48947	52.87677
9	0.774303	19.46414	28.31685	52.21900
10	0.825111	19.21104	32.44724	48.34172

LSDS大部分变化可以由它自身的新息来解释，而后随着时间的推移，LYYS、LZZS对其都有较大的解释力。

最后，通过方差分解得出结论：增值税在预测的时期内受到营业税和所得税的影响较小，而营业税则受到增值税影响较大，所得税则都有受到其他两税的影响。

三、结论

在实证研究中，我们对这三种税收进行预测时，可以使用时间序列分析的方法，通过序列的观测数据，对变量进行评估预测。由于所建模型是以观测数据序列为依据的，因此，对所建模型还应不断用新的观测数据进行修正，以正确描述该税种的动态变化过程。

(责任编辑：凌霄)

参考文献:

[1] Bas van Aarle Harry Garretsen, Niko Gob-

bin, Monetary and fiscal policy transmission in the Euro- area: evidence from a structural VAR analysis

[2] Hamilton, James D. (1994a) . Time Series Analysis, Princeton University Press

[3] Jhenson, Soren,(1988) , Statistical Analysis of Cointegration Vectors, Journal of Economic Dynamic and Control, 12:231- 254.

[4] Dickey, D.A. and W.A. Fuller (1979) . Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association, 74, 427- 431.

[5] Jhenson, Soren,(1991) , Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, Econometrica, 59:1551- 1580.

[6] MacKinnon, James G. (1991) . Critical Values for Cointegration Tests, Chapter 13 in R. F. Engle and C. W. J. Granger (eds) , Long- run Economic Relationships Readings in Cointegration, Oxford: Oxford University Press

[7] MacKinnon, James G. (1996) . Numerical Distribution Functions for Unit Root and Cointegration Tests, Journal of Applied Econometrics, 11, 601- 618.

[8] MacKinnon, James G., Alfred A. Haug, and Leo Michelis (1999) , Numerical Distribution Functions of Likelihood Ratio Tests For Cointegration, Journal of Applied Econometrics, 14, 563- 577.

A Research into the Revenue Sequence in China

Dong Jingjie

Abstract: Based on the monthly data and VECM,the author made a research into the primary revenues between Jan. 2001&Mar.2006 in China.Further more,the author found the long- term stable connection among value- added tax,sales tax and income tax. At last, the author forecasts some points to China's Revenue in short term.

Key words: Dynamic Analyzing Time Sequence; VECM