安徽省17市增值税流失率测算分析

赵为民

(中国人民大学财政金融学院 北京 100872)

【摘要】本文基于随机前沿分析方法,利用安徽17市5年的面板数据,测算了安徽各市的增值税税收流失率。结果表明,增值税流失率逐年降低,但各市的税收流失率差别较大。本文最后采用税柄法对影响增值税流失的因素做了影响分析,结果表明增值税流失受产业结构、人均收入和外贸部门规模影响,但稽查的力度对增值税流失影响不显著。

【关键词】随机前沿分析 增值税流失 税柄法

一、引言

Aigner,Lovell 和 Schmidt (1977) 及 Meeusen 和 Broeck (1977)在确定性前沿模型基础上,分别独立提出随机前沿分析(SFA)考虑随机因素的影响,把既定条件下投入可实现的最大产出的偏离分解为技术效率和随机扰动。即生产者在生产前沿面上的随机波动由生产效率和随机扰动两种因素导致,随机前沿分析方法成功地把误差项分解为统计误差和代表非有效性的单边误差,从而可以进一步测量代表技术非有效的单边误差。该方法不仅广泛应用于生产效率的评价而且也是公共通讯、医疗健康、教育部门效率评价中的重要方法。

崔兴芳及吕冰洋等首次把用于企业技术效率评价的DEA方法用于税务部门税收行政效率评价,具有开创意义。其后谢滨(2007)、王德祥和李建军(2009)分别用SFA方法测算了税收征管效率。在他们的研究中,生产投入均包括了税务部门的投入(人力、物力),因此其结果可以看成是税务部门效率。而本文在投入中去除了税务部门的投入,即投入中只包括经济的因素,根据经济决定税收的思路,把地区间相关经济因素作为投入,实际税收作为产出,则SFA的结果可以看成是地区间的相对税收流失率。

二、基本模型

随机边界模型的基础模型方程表达式是:

$$y=f(x;\beta)\times\exp\{v-u\} \tag{1}$$

其中,y代表产出标量;x为投入矢量;β是技术参数。

第一个误差项 $v\sim N(0,\sigma^2_v)$ 用以控制统计噪音,第二个误差项 $u\geqslant 0$ 则可以用来衡量技术非有效性。

这样判断生产企业的生产状况是在随机生产边界标准上还是低于这个标准线 $y=f(x;\beta)\times\exp\{v\}$,只要看 $u=0$ 还是 $u>0$。其中对u服从不同分布的假设,可以衍生出不同的具体模型。个别生产者不能达到生产函数前沿,是因为受到上述随机扰动和技术非效率两个因素影响。尽管这两个影响因素都是不可观察的,单身恰当定义的随机扰动仅是一个白噪声,多次观察的均值为零,因而个别生产者的技术效率(Technical efficiency)可以用样本中该生产者产出的期望与随机前沿的期望的比值来确定。即:

$$TE=\frac{E[f(x)\exp(v-u)]}{E[f(x)\exp(v-u)|u=0]}=\exp(-u_i) \tag{2}$$

对u服从不同分布的假设可以得出不同的模型,这里介绍正态—半正态模型,该模型有以下假设:

(i) $v\sim N(0,\sigma^2_v)$

(ii) $u_i\sim N^+(0,\sigma^2_v)$,即为非负的半正态分布

(iii) v和u独立分布,且与自变量无关。

假设 $f(x;\beta)$ 采用柯比—道格拉斯方程的对数线性形式,则随机边界模型可写为:

$$\ln y_i=\beta_0+\sum_n \beta_n \ln x_{ni}+v_i-u_i \tag{3}$$

其中,v_i 是噪音误差项,u_i 是非负的技术无效项。$\varepsilon=v-u$ 是组合误差项。

若独立性假设成立,则u和v的联合密度函数可表示为各自的密度函数的乘积:

$$f(u,v)=\frac{2}{2\pi\sigma_u\sigma_v}\cdot\exp\{-\frac{u^2}{2\sigma_u^2}-\frac{u^2}{2\sigma_v^2}\} \tag{4}$$

因为 $\varepsilon=v-u$,u和ε的联合密度函数可表示为:

$$f(u,\varepsilon)=\frac{2}{2\pi\sigma_u\sigma_v}\cdot\exp\{-\frac{u^2}{2\sigma_u^2}-\frac{(\varepsilon+u)^2}{2\sigma_v^2}\} \tag{5}$$

$\varepsilon=v-u$,边际密度函数可由 $f(u,\varepsilon)$ 对u进行积分求得:

$$f(\varepsilon)=\int_0^\infty f(u,\varepsilon)du=\frac{2}{\sigma}\cdot\phi(\frac{\varepsilon}{\sigma})-\Phi(-\frac{\varepsilon\lambda}{\sigma}) \tag{6}$$

其中,$\sigma=(\sigma_u^2+\sigma_v^2)^{1/2}$,$\lambda=\sigma_u/\sigma_v$。$\phi(\cdot)$ 和 $\Phi(\cdot)$ 分别表示标准正态密度函数和累积分布函数。

由式(6),I个样本生产单元的对数似然函数为:

$$\ln L=\text{常数}-I\ln\sigma+\sum_i \ln\Phi(-\frac{\varepsilon_i\lambda}{\sigma})-\frac{1}{2\sigma^2}\sum_i \varepsilon_i^2 \tag{7}$$

通过关于各参数对数似然函数(7)式进行最大化可得到所有参数似然估计值,且当 $I\to+\infty$ 时,这些估计值具有一致性。

在已知ε时的u的条件分布为:

$$f(u|\varepsilon)=\frac{f(u,\varepsilon)}{f(\varepsilon)}=\frac{1}{\sqrt{2\pi}\sigma_*}\cdot\frac{\exp\{-\frac{(u-\mu_*)^2}{2\pi\sigma_*}\}}{1-\Phi(-\frac{\mu_*}{\sigma_*})} \quad (8)$$

其中，$\mu_*=-\varepsilon\sigma_u^2/\sigma^2$，$\sigma_*=\sigma_u^2\sigma_v^2/\sigma^2$。因为 $f(u|\varepsilon)$ 服从 $N^+(\mu_*,\sigma_*^2)$ 分布，所以该分布的均值或众数可作为 u_i 的一个点估计值：

$$E(u_i|\varepsilon_i)=\sigma_*[\frac{\phi(\frac{\varepsilon_i\lambda}{\sigma})}{1-\Phi(-\frac{\varepsilon_i\lambda}{\sigma})}-(\frac{\varepsilon_i\lambda}{\sigma})] \quad (9)$$

一旦求得了 u 的点估计值，各生产单元的技术有效性则可从下式得到：

$$TE_i=\exp\{-\hat{u}_i\} \quad (10)$$

其中 $\hat{u}_i$ 是 $E(u_i|\varepsilon_i)$。

本文采用的模型是 Battese 和 Coelli(1992)年提出针对面板数据的随机前沿模型。模型具体形式如下：

$$Y_{ir}=x_{it}+\beta+(v_{it}+u_{it}) \quad i=1,\cdots,N,t=1,\cdots,T \quad (11)$$

这里 Y_{ir} 代表 i 个单位在第 t 时期的产出，x_{it} 代表 i 单元在第 t 时期投入向量。β 是技术参数。$v_{it}\sim N(0,\sigma^2_v)$ 是随机干扰项。$u_{it}=(u_i\exp(-\eta(r-T)))$，这里 u_i 是非负的随机变量，代表了生产上的技术非效率，服从在零处截断非负的 $N(\mu,\sigma^2_v)$ 分布，v_{it} 和 u_{it} 相互独立。η 表示技术非效率随时间变化的趋势，是待估计量。

采用 Battese 和 Corra（1977）年提出的参数化方法，用 $\sigma^2=\sigma^2_v+\sigma^2_u$ 和 $\gamma=\sigma^2_u/(\sigma^2_v+\sigma^2_u)$ 替代 σ^2_v 和 σ^2_u。使用最大似然估计法估计待估参数。

三、税收流失率估算模型

利用 SFA 测算安徽省增值税税收流失率的思路如下：选择增值税收入作为产出变量，根据增值税税基的角度分析，工业增加值减去出口额应该是一种投入变量，第三产业的增加值剔除住宿餐饮业、金融业、房地产业增加值可以看成投入变量，因为进口也是增值税税基的重要构成，因此在投入变量中应包括进口额。

本文采用 Cobb-Douglas 生产函数，使用安徽省 17 个地市的 2006～2010 年 5 年的面板数据。这 17 个市是：合肥、芜湖、马鞍山、蚌埠、铜陵、淮南、淮北、亳州、黄山、六安、池州、巢湖、宣城、滁州、安庆、阜阳、宿州。具体模型如下：

$$\ln(Vat)_{it}=\beta_0+\beta_1\ln(gy_{it})+\beta_2\ln(jk_{it})+\beta_3\ln(dscy_{it})+(v_{it}+u_{it}) \quad (12)$$

其中，i 表示安徽省 17 个市国税局，t 表示从 2006 至 2010 年的 5 年样本期。变量具体解释见表 1：

表1　　变 量 解 释

变量	产出/投入	含义及口径解释
Vat	产出变量	增值税收入
gy	投入变量	工业增加值(不包括建筑业)并减去出口额
jk	投入变量	进口额
dscy	投入变量	第三产业增加值并减去住宿餐饮业、金融业、房地产业增加值

利用 frontier4.1 估算的具体结果见表 2：

表2　　方程估计结果

	coefficient	standard-error	t-ratio
β_0	0.462 357 74E+01	0.170 073 51E+01	0.271 857 58E+01 ***
β_1	0.182 684 38E+00	0.109 719 53E+00	0.166 501 24E+01 ***
β_2	-0.501 920 53E-01	0.495 119 62E-01	-0.101 373 59E+01 **
β_3	0.459 612 78E+00	0.187 338 73E+00	0.245 337 84E+01 ***
sigma-squared	0.298 440 21E+00	0.148 668 26E+00	0.200 742 39E+01 ***
gamma	0.951 519 38E+00	0.274 523 40E-01	0.346 607 75E+02 ***
mu	0.858 618 58E+00	0.286 529 04E+00	0.299 661 98E+01 ***
eta	0.712 515 10E-01	0.115 837 67E-01	0.615 097 94E+01 ***
og likelihood function=0.200 985 81E+02			
LR test of the one-sided error=0.777 402 73E+02			

注：*** 表示在1%显著性水平下显著，** 表示在5%显著性水平下显著，* 表示在10%显著性水平下显著，下同。

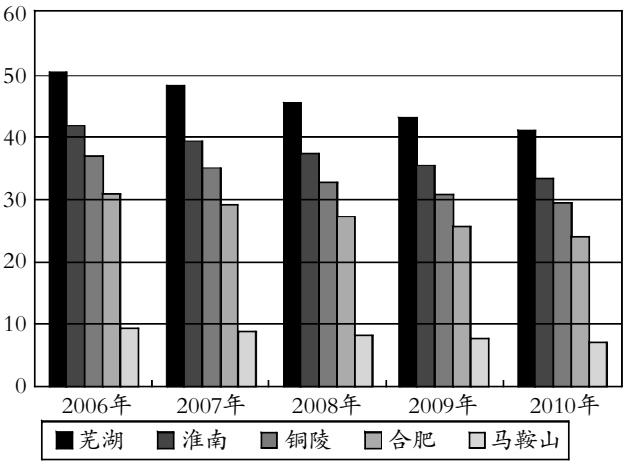
σ^2 的方差为 0.29，γ 接近 1 表明存在很大的技术非效率，并且 T 检验在 1%的显著性水平下显著，说明模型采用的形式是恰当的。eta 表明随着时间的变化技术非效率在降低。极大似然检验表明方程总体显著。

从各系数的值可以看出，β_1 和 β_3 的系数为正，并且极其显著，表明工业和第三产业均对增值税有正向影响，并且第三产业的影响大于工业。进口对增值税的影响为负，在 5%的显著性水平下显著，这与直觉不符，可能原因是进口量大，出口量一般也大，而出口退税会对增值税收入造成负向影响。

表 3 列出了 2006～2010 年安徽 17 市的增值税流失率最低的 5 个城市，如表 3 所示：

表3　2006～2010年安徽增值税流失率最低5个城市的流失率

	2006年	2007年	2008年	2009年	2010年
芜湖	50.1	47.7	45.3	43	40.7
淮南	41.5	39.3	37.2	35.2	33.2
铜陵	36.7	34.7	32.7	30.9	29.1
合肥	30.6	28.8	27.1	25.5	24
马鞍山	9.3	8.7	8.1	7.6	7.1



2006～2010年安徽增值税流失率最低5个城市的流失率

随机前沿分析方法测算出来的是地区间的相对税收流失率,并不表示绝对量,从表3中可以看出马鞍山的增值税相对税收流失率显著低于其他四个市，这是因为马鞍山是我国重要的钢铁基地,马鞍山钢铁公司是特大型国有上市公司,这样的特殊产业结构导致税收流失率显著低于其他城市。但从纵向来看,马鞍山的增值税增长潜力不大,其他四市5年间税收流失率大幅降低,即增值税增长潜力较大。

四、增值税流失率影响分析

一国的税收流失与该国可衡量的经济特征具有某种规则关系,本文采用罗兹—莫斯、希恩、马斯格雷夫的税柄法并结合税务管理变量进行税收流失影响因素分析。

从理论上来讲，税柄向量应包含全部影响税收流失的因素,但实际操作中只能考察有限的几个经济因素。一般包括三个变量:外贸部门的规模、经济发展阶段、增值部门构成。

外贸部门的规模：通常认为税收流失与外贸部门规模负相关。外贸规模越大,表明货币化、市场化程度越高,从而对外贸征税的管理难度越小;其流失率也就应该越小。经济发展阶段:经济发展伴随着文盲率的降低、货币化程度的提高以及法律制度的规范,这些因素都会降低税收流失率。衡量经济发展的指标很多,但使用最广泛的指标则为人均收入。增值部门构成:一般说来,农业部门产出比例较高,意味着工商业化程度较低,货币化程度低,征税难度大,进而影响税收流失率。

除了经济特征变量以外,本文同时增加了税务管理变量,税务管理的力度越大,税收流失应该越小,代表税务管理的力度的变量也有很多,如稽查比例、处罚比例、人均投入、信息化投入等。

本文采用安徽17个地市的2006~2010年的面板数据进行估计,其具体模型如下:

$$\ln(gap)_{it}=\beta_0+\beta_1\ln(dycy_bl_{it})+\beta_2\ln(jck_bl_{it})+\beta_3\ln(rjcc_{it})+\beta_4\ln(jcsr_bl_{it})+e$$

其中,i表示安徽省17个市国税局,t表示从2006~2010年的5年样本期。变量含义具体解释如表4所示:

表4　　增值税流失影响分析变量解释

变量	应变量/自变量	含义解释
gap	因变量	增值税流失率
dycy_bl	自变量	第一产业占GDP的比例。代表产业部门构成
jck_bl	自变量	进出口总值占GDP的比例。代表外贸部门规模
rjcz	自变量	人均产出。代表经济发展阶段
jcsr_bl	自变量	稽查罚补金额占增值税收入比例。代表税务管理力度

注:dycy_bl,jck_bl,rjcc数据来源于安徽统计年鉴,jcsr_bl数据来源于安徽国税征管数据。

本文采用R软件对上述面板数据模型进行估计,通过F检验,由于混合效应模型的RSS为10.345,固定效应模型的RSS为0.035.得出F值为1239.18.故采用固定效应模型;通过hausman检验,得出p-value为0.256,因此应采用随机效应模型进行模型估计。具体结果如下所示:

表5

	Estimate	Std.Error	t-value	Pr(>|t|)
(Intercept)	4.732 947	0.293 188 04	16.14 3	<2.2e-16 ***
log(dycy_bl)	0.297 709	0.048 469 66	6.142 2	2.980e-08 ***
log(jck_bl)	0.059 794	0.011 042 78	5.414 7	6.267e-07 ***
log(rjcc)	-0.139 83	0.018 150 4	-7.704	3.067e-11 ***
log(jcsr_bl)	-0.000 58	0.004 311 56	0.134 6	0.893 3
Residual Sum of Squares:0.044 747				
R-Squared:0.869 31				
Adj. R-Squared:0.818 18				
F-statistic:133.037 on 4 and 80 DF,p-value:<2.22e-16				

从F检验来看，方程总体十分显著，拟合优度达到了0.82。第一产业占GDP比例变量系数为0.3,并且1%显著性水平下显著,表明第一产业的比例提高1%,增值税流失率会提高0.3%。

人均产值的比例系数为-0.14，并且在1%显著性水平下显著,人均产值提高1%,增值税流失率会降低0.14%。

进出口规模的系数为0.06，并且在1%显著性水平下显著,即进出口规模提高1%,增值税流失会提高0.06%。进出口规模反映一个地区的货币化和市场化程度，进出口规模越过增值税流失率应该越小，但另一方面进出口本身又直接影响增值税收入,进口需要征收进口环节增值税,而出口需要进行出口退税,所以进出口规模对增值税的影响十分复杂,既存在间接影响也存在直接影响,从这里估计的结果来看,进出口所占GDP比例越大增值税流失也越多。

稽查收入所占增值税总收入比例的系数为负，表明稽查的力度越大,增值税的流失越小,这符合常理,但这个系数很小,仅为0.00058,并且统计结果也不显著,因此有理由怀疑稽查的作用还没有完全发挥出来，加大稽查的力度对增值税收入的提高还有很大的潜力。

主要参考文献

1. 谭荣华，梁季. 我国增值税收入能力估测. 涉外税务,2006;1

2. 谢滨税. 收征管效率评估——基于随机前沿模型研究. 税务研究,2007;11

3. 李双杰,范超.随机前沿分析与数据包络分析方法的评析与比较.统计与决策,2009;7