

## 小白学统计|面板数据分析与Stata应用笔记（四）

**\*\*这一章的笔记是基础理论的内容，可能看上去觉得没有什么用，但是和之后的内容联系很紧密\*\***

### 机制识别方法

一篇高质量的经验研究论文必须回答核心变量以什么样的方式影响了被解释变量，要揭示核心解释变量影响被解释变量背后的机制。所以，机制识别方法在论文研究中是十分重要的。

#### 一、联立方程法

联立方程法是一种重要的机制识别方法。

**联立方程模型：**经济理论常常推导出一组相互联系的方程，其中一个方程的解释变量是另一方程的被解释变量，这就是联立方程组。

方红生教授以印发在国际顶级期刊经济与统计学评论上的论文：《民主、波动和经济发展》为例介绍联立方程法。

Ahmed Mushfiq Mobarak, 2005. "Democracy, Volatility, and Economic Development,

"The Review of Economics and Statistics, MIT Press, vol. 87(2), pages 348-361, May

论文《民主、波动和经济发展》构造了一个联立方程来试图解释民主是否通过影响了波动来影响经济发展。

$$Growth_{it} = \alpha_0 Volatility_{it} + \alpha_1 Democracy_{it} + \alpha_2 X_{it} + \mu_i + \varepsilon_{it}$$

$$Volatility_{it} = \beta_0 Democracy_{it} + \beta_1 X_{it} + \mu_i + \varepsilon_{it}$$

在第一个方程中，民主对经济发展的直接效应是  $\alpha_1$ ，而民主通过影响波动来影响经济发展的间接效应是  $\alpha_0 \cdot \beta_0$ 。

#### 二、联立方程的估计方法

在使用单一方程估计法时，由于忽略了各方程之间的联系（包括各方程扰动项之间的联系），所以不如将所有方程作为一个整体进行估计（即**系统估计法**）更有效率。

**系统估计法的缺点：**如果其中的某个方程估计的不准确，则可能会影响系统中其它方程的估计。

#### 三阶段最小二乘法

三阶段最小二乘法是最常见的系统估计法。

**(1)** 对于一个多方程系统，如果各方程中都不**包含内生解释变量**，则对每个方程进行OLS估计是一致的，但却不是最有效的，因为单一方程OLS忽略了不同方程的扰动项之间可能存在的相关性。

此时，利用**SUR** (Seemingly Unrelated Regression Estimation) 对整个方程系统进行估计是有效率的。

**SUR的Stata实现命令为：**

xi:reg3(第一个方程的被解释变量 第一个方程的解释变量 i.code i.year)(第二个方程的被解释变量 第二个方程的解释变量 i.code i.year), sure

**(2)** 对于一个多方程系统，如果方程中**包含内生解释变量**，则对每个方程进行2SLS估计是一致的，但却不是最有效的，因为单一方程2SLS忽略了不同方程的扰动项之间可能存在的相关性。

此时，利用**3SLS**对整个联立方程系统同时进行估计是最有效率的。

**3SLS的Stata实现命令为：**

xi:reg3(第一个方程的被解释变量 第一个方程的解释变量 i.code i.year)(第二个方程的被解释变量 第二个方程的解释变量 i.code i.year), endog(varlist) exog(varlist)

其中，endog()选项代表系统内除了被解释变量外的内生变量，exog()选项代表从系统外寻找的工具变量。

#需要注意的是：论文《民主、波动和经济发展》中，虽然波动在第一个方程中是内生变量，但是不要放入endog()选项中，因为波动在第二个方程中是作为被解释变量的。此外，如果系统内没有除了被解释变量外的内生变量，那么我们就可以删除endog()选项和exog()选项。

### #内生变量与外生变量：

对联立方程模型系统而言，已经不能用被解释变量与解释变量来划分变量，而将变量分为内生变量与外生变量两大类。

**内生变量：**内生变量是具有某种概率分布的随机变量，它的参数是联立方程系统估计的元素。内生变量是由模型系统决定的，同时也对模型系统产生影响。（内生变量一般都是经济变量，通常情况下，模型中内生变量与随机变量相关）

**外生变量：**外生变量一般是确定性变量，或者是具有临界概率分布的随机变量，其参数不是模型系统研究的元素。外生变量影响系统，但本身不受系统的影响。（外生变量一般是经济变量、条件变量、政策变量、虚变量，通常情况下，模型中外生变量与随机项不相关）

## 三、机制识别方法

### Acemoglu et al.(2003)方法

Acemoglu等人2003年发表在国际顶级期刊《货币经济学》杂志上的文章：《制度原因，宏观经济症状：波动，危机与增长》中所使用的的机制识别方法被后来的许多文章引用。

Acemoglu, D.J.A.Robinson, Y.Y.Thaicharoen, and S.Johnson, 2003, "Institutional cause s, macroeconomic symptoms: volatility, crises and growth", Journal of Monetary Economics, 50, 49-123.

由于文章比较长，为了理解Acemoglu等人的方法，方红生教授以文章《攫取之手、援助之手与中国税收超GDP增长》为例进行讲解。

《攫取之手、援助之手与中国税收超GDP增长》（方红生、张军），《经济研究》2013年第3期

其中，“攫取之手”是指中央税收占总税收的比重；“援助之手”是指中央转移支付占地方政府财政支出的比重。

文章构建了三个方程，分别如下：

(1)

$$Tshare_{it} = \alpha_1 Tshare_{it-1} + \alpha_2 Rgh_{it} + \alpha_3 Rhh_{it} + \alpha_4 Rgh_{it} * Rhh_{it} + \beta X_{it} + \mu_i + \mu_t + \varepsilon_{it}$$

其中，被解释变量Tshare为税收占GDP的比重，解释变量由被解释变量滞后一期（ $Tshare_{it-1}$ ）、攫取之手（Rgh）、援助之手（Rhh）、两只手的交互项（ $Rgh * Rhh$ ）以及其他控制变量组成。核心解释变量为攫取之手、援助之手和两只手的交互项。

预期，攫取之手与援助之手对GDP占税收的比重具有正向效应，而交互项具有负向效应。

文章试图对税收征管效率和高税行业的发展这两个渠道的重要性进行识别考虑到税收征管效率面临着严重的度量问题，所以只为高税行业的发展这一渠道寻找一个相对合适的代理变量，即，使用非农产业化，所以构造第二个方程。

(2)

$$Pnonagri_{it} = \alpha_1 Pnonagri_{it-1} + \alpha_2 Rgh_{it} + \alpha_3 Rhh_{it} + \alpha_4 Rgh_{it} Rhh_{it} + \beta X_{it} + \mu_i + \mu_t + \varepsilon_{it}$$

其中，非农产业化（Pnonagri）为被解释变量；解释变量由被解释变量滞后一期（ $Tshare_{it-1}$ ）、攫取之手（Rgh）、援助之手（Rhh）、两只手的交互项（ $Rgh * Rhh$ ）以及其他控制变量组成。核心解释变量为攫取之手、援助之手和两只手的交互项。

预期，攫取之手与援助之手对非农产业化具有正向效应，而交互项具有负向效应。

第三个方程为：将非农产业化引入第一个方程构建起来。

(3)

$$Tshare_{it} = \gamma Pnonagri_{it} + \alpha_1 Tshare_{it-1} + \alpha_2 Rgh_{it} + \alpha_3 Rhh_{it} + \alpha_4 Rgh_{it} * Rhh_{it} + \beta X_{it} + \mu_i + \mu_t + \varepsilon_{it}$$

其中，核心解释变量依然是攫取之手、援助之手和两只手的交互项。

接下来，我们通过Acemoglu等人的判别规则判断“非农产业化”这个机制的重要性。

- 如果与第一个方程进行对比，第三个方程中的核心变量由显著变为不显著，或其显著性和（或）系数有明显的下降，而Pnonagri显著，那么Pnonagri是其作用于税收占GDP比重的一个主要渠道。
- 如果核心变量显著而Pnonagri不显著，那么Pnonagri不是其作用于税收占GDP比重的一个渠道。在此情形下，核心变量只能通过其他渠道（如提高税收征管效率）起作用。
- 如果核心变量和Pnonagri都显著，且前者的显著性和系数并没有明显的下降，那么核心变量作用于税收占GDP比重的主要渠道是税收征管效率的提高而非Pnonagri。

渠道检验的结果如下

表 3 攫取之手、援助之手与中国税收超 GDP 增长：渠道检验

| 被解释变量                        | <i>Pnonagri</i>       |                      | <i>Taxshare</i>     |                       |
|------------------------------|-----------------------|----------------------|---------------------|-----------------------|
| 模型                           | 模型 1                  | 模型 2                 | 模型 3                | 模型 4                  |
| <i>Taxshare_l</i>            |                       |                      | 0.854***<br>(0.000) | 0.585***<br>(0.000)   |
| <i>Pnonagri_l</i>            | 1.097***<br>(0.000)   | 1.112***<br>(0.000)  |                     |                       |
| <i>Pnonagri</i>              |                       |                      |                     | 0.176***<br>(0.003)   |
| <i>Rgh</i>                   | 0.025**<br>(0.021)    | 0.018**<br>(0.013)   | 0.192***<br>(0.000) | 0.198***<br>(0.001)   |
| <i>Rhh</i>                   | 0.015**<br>(0.021)    | 0.0149**<br>(0.026)  | 0.130**<br>(0.017)  | 0.086*<br>(0.053)     |
| <i>Rgh*Rhh</i>               | -0.00056**<br>(0.015) | -0.0005**<br>(0.012) | -0.003**<br>(0.014) | -0.003**<br>(0.014)   |
| <i>Pgdp</i>                  |                       | -0.0001*<br>(0.097)  | 0.0001<br>(0.437)   | -0.0006***<br>(0.001) |
| <i>Popbonus</i>              |                       | 0.029<br>(0.203)     | -3.192**<br>(0.020) | -2.355**<br>(0.012)   |
| <i>Popbonus</i> <sup>2</sup> |                       |                      | 0.022**<br>(0.020)  | 0.017**<br>(0.014)    |
| <i>Privatization</i>         |                       | -0.002<br>(0.671)    | -0.069**<br>(0.039) | -0.053*<br>(0.070)    |
| <i>Urban</i>                 |                       | -0.004<br>(0.486)    | -0.043<br>(0.281)   | -0.025<br>(0.484)     |
| <i>Open</i>                  |                       | 0.007<br>(0.235)     | 0.023<br>(0.143)    | 0.017<br>(0.210)      |
| <i>Inflation</i>             |                       | 0.038<br>(0.558)     | 0.027<br>(0.657)    | -0.00004<br>(0.999)   |
| <i>Cons</i>                  | -1.389***<br>(0.000)  | -2.951<br>(0.141)    | 112.02**<br>(0.024) | 85.855**<br>(0.011)   |



其中，模型2为上文中的第二个方程，可以看到，攫取之手、援助之手和两只手的交互项对Pnonagri产生了预期的影响；模型3为上文中的第一个方程，同样，攫取之手、援助之手和两只手的交互项对税收占GDP的比重产生了预期的影响；模型4即为上文中的第三个方程，可以看到，Pnonagri显著为正，在核心解释变量方面，与模型3相比，只有援助之手的系数和显著性有显著下降，攫取之手和两只手的交互项并未发生显著性的变换。

根据Acemoglu等人的判别规则我们可以得出结论：

(1) 中央政府援助之手是将非农产业化作为其推高税收占GDP比重的一个主要渠道，税收征管效率仅是次要渠道。

(2) 攫取之手及其与援助之手的交互项是将税收征管效率作为其推高税收占GDP比重的一个主要渠道，非农产业化仅是次要渠道。

