

第7章 扩展的单方程模型

- 经典的单方程计量经济学模型理论与方法：
限于常参数、线性、揭示变量之间因果关系的单方程模型；
被解释变量是连续的随机变量，其抽样是随机和不受限制的；
或者只利用时间序列样本，或者只利用截面数据样本；
主要依靠对经济理论和行为规律的理解确定模型的结构形式。
- 本章中，将讨论几种单方程扩展模型，主要包括：
 1. 将被解释变量抽样由完全随机扩展为受到限制的选择性样本模型；
 2. 将被解释变量是连续的扩展为离散的离散选择模型；
 3. 将单一种类的样本扩展为同时包含截面数据和时间序列数据的平行数据样本（Panel Data）等。

第7章 说明

- 这些模型与方法，无论在计量经济学理论方面还是在实际应用方面，都具有重要意义。
- 但是，这些模型都形成了各自丰富的内容体系，甚至是计量经济学的新分支学科，模型方法的数学过程较为复杂。
- 本章只介绍其中最简单的模型，以了解这些模型理论与方法的概念与思路。

§ 7.1 选择性样本模型

Selective Samples Model

- 一、经济生活中的选择性样本问题：对 y
- 二、“截断”问题的计量经济学模型
- 三、“归并”问题的计量经济学模型*

**The Bank of Sweden Prize in Economic Sciences in
Memory of Alfred Nobel 2000**

**"for his development of theory and methods
for analyzing selective samples"**



James J Heckman

USA

- **“Shadow Prices, Market Wages and Labour Supply”, Econometrica 42 (4), 1974, P679-694**
发现并提出“选择性样本”问题。
- **“Sample Selection Bias as a Specification Error”, Econometrica 47(1), 1979, P153-161**
证明了偏误的存在并提出了Heckman两步修正法。

一、经济生活中的选择性样本问题

1、“截断”（truncation）问题

- 由于条件限制，样本不能随机抽取，即不能从全部个体，而只能从一部分个体中随机抽取被解释变量 y 的样本观测值、这部分个体的观测值都大于或小于某个确定值：“掐头”或“去尾”。
 - 例如消费函数模型：由于抽样原因，被解释变量样本观测值最低200元、最高10000元。
 - 例如农户贷款影响因素分析模型：如果调查了10000户，其中只有6000户在一年内发生了贷款。仅以发生了贷款的6000户的贷款额作为被解释变量观测值，显然是将其它没有发生贷款的4000户“截断”掉了。

2、“归并/审查”(censoring)问题

- 将被解释变量处于某一范围的样本观测值，都用一个相同的值代替。
 - 经常出现在“检查”、“调查”活动中，因此也称为“检查”(censoring)问题。
 - 例如需求函数模型：用实际消费量作为需求量的观测值，如果存在供给限制，就出现“归并”问题。
 - 被解释变量观测值，存在最高和最低的限制；例如考试成绩，最高100、最低0，出现“归并”问题。

二、“截断”问题的计量经济学模型

1、思路

- 如果只能从“掐头”或者“去尾”的连续区间随机抽取“被解释变量 y 的样本观测值”，那么很显然，抽取每一个样本观测值的概率以及抽取一组样本观测值的联合概率，与被解释变量的样本观测值不受限制的情况是不同的。
- 如果能够知道在这种情况下、抽取一组样本观测值的联合概率函数，那么就可以通过该函数极大化求得模型的参数估计量。

2、截断分布

$$f(\xi|\xi > a) = \frac{f(\xi)}{P(\xi > a)}$$

a 为随机变量 ξ
分布范围内的一个常数

$$1) \quad f(\xi|\xi > c) = \frac{f(\xi)}{P(\xi > c)} = \frac{1/(b-a)}{\int_c^b \frac{1}{b-a} d\xi} = \frac{1}{b-c}$$

如果 ξ 服从均匀分布 $U(a, b)$:
但是它只能在 (c, b) 内取得样本观测值,
那么取得每一个样本观测值的概率

$$\begin{aligned}
 2) \quad f(\xi | \xi > a) &= \frac{f(\xi)}{P(\xi > a)} \\
 &= \frac{(2\pi\sigma^2)^{-1/2} e^{-(\xi-\mu)^2/(2\sigma^2)}}{1-\Phi(\alpha)} \\
 &= \frac{\frac{1}{\sigma} \varphi(\frac{\xi-\mu}{\sigma})}{1-\Phi(\alpha)}
 \end{aligned}$$

ξ 服从
正态分布

$$P(\xi > a) = 1 - \Phi\left(\frac{a - \mu}{\sigma}\right) = 1 - \Phi(\cdot)$$

Φ 是标准正态分布
条件概率函数

3、截断被解释变量数据模型的最大似然估计

$$y_i = \mathbf{B}'\mathbf{X}_i + \varepsilon_i$$

$$\varepsilon_i \sim N(0, \sigma^2)$$

$$y_i | \mathbf{X}_i \sim N(\mathbf{B}'\mathbf{X}_i, \sigma^2)$$

$$f(y_i | y_i > \mathbf{a}) = \frac{\frac{1}{\sigma} \varphi((y_i - \mathbf{B}'\mathbf{X}_i) / \sigma)}{1 - \Phi((\mathbf{a} - \mathbf{B}'\mathbf{X}_i) / \sigma)}$$



$$\ln L = -\frac{n}{2} (\ln(2\pi) + \ln \sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mathbf{B}'\mathbf{X}_i)^2 - \sum_{i=1}^n \ln \left(1 - \Phi \left(\frac{a - \mathbf{B}'\mathbf{X}_i}{\sigma} \right) \right)$$



$$\frac{\partial \ln L}{\partial \begin{pmatrix} \mathbf{B} \\ \sigma^2 \end{pmatrix}} = \sum_{i=1}^n \begin{pmatrix} \left(\frac{y_i - \mathbf{B}'\mathbf{X}_i}{\sigma^2} - \frac{\lambda_i}{\sigma} \right) \mathbf{X}_i \\ -\frac{1}{2\sigma^2} + \frac{(y_i - \mathbf{B}'\mathbf{X}_i)^2}{2\sigma^4} - \frac{\alpha_i \cdot \lambda_i}{2\sigma^2} \end{pmatrix} = \sum_{i=1}^n \mathbf{g}_i = \mathbf{0}$$

$$\alpha_i = (a - \mathbf{B}'\mathbf{X}_i) / \sigma$$

$$\lambda_i(\alpha_i) = \phi(\alpha_i) / (1 - \Phi(\alpha_i))$$

- 求解该1阶极值条件，即可以得到模型的参数估计量。
- 由于这是一个复杂的非线性问题，需要采用迭代方法求解，例如牛顿法。

4、例7.1.1:城镇居民消费模型

人均收入	人均消费	人均收入	人均消费	人均收入	人均消费
1120	1020	4640	2900	6090	3900
1310	1150	4750	2980	6200	3950
1300	1145	4800	2970	6330	4000
1430	1230	4810	3050	6450	4030
1500	1275	4990	3200	6570	4080
1670	1385	5070	3100	6700	4130
2100	1660	5130	3175	6840	4000
2370	1840	5210	3200	7010	4200
2530	1950	5300	2450	7170	4160
2790	2110	5390	3230	7350	4210
2980	2240	5450	3310	7500	4325
3200	2380	5500	3500	7670	4385
3460	2550	5570	3510	7840	4450
3630	2660	5630	3590	8000	4500
3880	2700	5690	3600	8190	4865
4040	2730	5770	3650	8350	4880
4210	2720	5860	3720	8500	4890
4390	2850	5930	3850	8690	4920
4520	2800	6000	3800	8830	4970

OLS估计：将样本看为“不受任何限制下”随机抽取的样本

Dependent Variable: Y
 Method: Least Squares
 Date: 01/03/11 Time: 16:26
 Sample: 1 57
 Included observations: 57

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	604.9304	57.65070	10.49303	0.0000
X	0.508307	0.010363	49.04995	0.0000
R-squared	0.977650	Mean dependent var	3228.509	
Adjusted R-squared	0.977244	S.D. dependent var	1076.505	
S.E. of regression	162.3915	Akaike info criterion	13.05235	
Sum squared resid	1450405.	Schwarz criterion	13.12404	
Log likelihood	-369.9921	F-statistic	2405.898	
Durbin-Watson stat	1.011832	Prob(F-statistic)	0.000000	

$$Y_i = 604.93 + 0.5083X_i \quad i = 1, 2, \dots, 57 \quad R^2 = 0.9777$$

ML估计： 将样本看为“在消费水平大于1000元、小于5000元的特定人群”中随机抽取的样本

Equation Estimation

Specification Options

Equation specification
Dependent variable followed by list of
y c x

Distribution
☒ Normal
☐ Logisti
☐ Extreme Value

Dependent variable censoring points
Enter a number, a series, a series expression, or blank
Left 1000
Right 5000

Left & Right points
☒ Actual censoring val
☐ Zero/one indicator of cen
☒ Truncated sampl

Estimation settings
Method: CENSORED - Censored or truncated data (tobit)
Sample: 1 57

确定 取消

截断点
选择

样本类
型选择

估计方法选择

Dependent Variable: Y

Method: ML - Censored Normal (TOBIT) (Quadratic hill climbing)

Date: 01/03/11 Time: 16:49

Sample: 1 57

Included observations: 57

Truncated sample

Left censoring (value) series: 1000

Right censoring (value) series: 5000

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
C	556.7026	63.70923	8.738178	0.0000
X	0.519423	0.011845	43.85160	0.0000
Error Distribution				
SCALE: C(3)	161.6729	15.70998	10.29109	0.0000
R-squared	0.977492	Mean dependent var	3228.509	
Adjusted R-squared	0.976658	S.D. dependent var	1076.505	
S.E. of regression	164.4681	Akaike info criterion	12.96250	
Sum squared resid	1460686.	Schwarz criterion	13.07003	
Log likelihood	-366.4314	Hannan-Quinn criter.	13.00429	
Avg. log likelihood	-6.428621			

$$Y_i = 556.70 + 0.5194X_i \quad i = 1, 2, \dots, 57 \quad R^2 = 0.9775$$

5、为什么截断被解释变量数据模型，不能采用普通最小二乘估计？

- 对于截断被解释变量数据计量经济学模型，如果仍然把它看作为经典的线性模型，采用OLS估计，会产生什么样的结果？
- 因为 y_i 只能在大于 a 的范围内取得观测值，那么 y_i 的条件均值为：

$$\begin{aligned} E(y_i | y_i > a) &= \int_a^{\infty} y_i \phi(y_i | y_i > a) dy_i \\ &= B'X_i + \sigma \frac{\phi((a - B'X_i) / \sigma)}{1 - \Phi((a - B'X_i) / \sigma)} \end{aligned}$$

$$E(y_i | y_i > a) = \mathbf{B}'\mathbf{X}_i + \sigma\lambda(\alpha_i)$$

$$\alpha_i = \frac{a - \mathbf{B}'\mathbf{X}_i}{\sigma}$$

$$\begin{aligned} \frac{\partial E(y_i | y_i > a)}{\partial \mathbf{X}_i} &= \mathbf{B} + \sigma \left(\frac{d\lambda_i}{d\alpha_i} \right) \frac{\partial \alpha_i}{\partial \mathbf{X}_i} \\ &= \mathbf{B} + \sigma(\lambda_i^2 - \alpha_i \lambda_i) \left(\frac{-\mathbf{B}}{\sigma} \right) \\ &= \mathbf{B}(1 - \lambda_i^2 + \alpha_i \lambda_i) \\ &= \mathbf{B}(1 - \delta(\alpha_i)) \end{aligned}$$

$$y_i | y_i > a = E(y_i | y_i > a) + u_i = \mathbf{B}'\mathbf{X}_i + \sigma\lambda(\alpha_i) + u_i$$

$$Var(u_i) = \sigma^2(1 - \lambda_i^2 + \lambda_i \alpha_i) = \sigma^2(1 - \delta_i) \quad \text{略}$$

- 由于被解释变量数据的截断问题，使得原模型变换为包含一个非线性项模型。
- 如果采用OLS直接估计原模型：
 - 实际上忽略了一个非线性项；
 - 忽略了随机误差项的异方差性。
- ◆ 这就造成参数估计量的偏误；而且，如果不了解被解释变量的分布，要估计该偏误的严重性、也是很困难的。

三、“归并/审查”问题的 计量经济学模型：Tobit

1、思路

- 以一种简单的情况为例，讨论“归并”问题的计量经济学模型：

假设原始被解释变量服从正态分布，其样本观测值以0为界：小于0的、都归并为0，大于0的、则取实际值。

- 如果 y^* 以表示原始被解释变量， y 以表示归并后的被解释变量，那么则有：

$$\begin{array}{ll} y = 0 & \text{if } y^* \leq 0 \\ y = y^* & \text{if } y^* > 0 \end{array}$$

$$y^* \sim N(\mu, \sigma^2)$$

- 单方程线性“归并”问题的计量经济学模型为：

$$\begin{cases} y_i^* = X_i B + \mu_i \\ y_i = \max(y_i^*, 0) \end{cases}$$

$$\mu_i \sim N(0, \sigma^2)$$

- 如果能够得到 y_i^* 的概率密度函数，那么就可以方便地采用最大似然法估计模型，这就是研究这类问题的思路。
- 由于该模型是由Tobin于1958年最早提出的，所以也称为Tobit模型。

2、“归并”变量的正态分布

- 由于原始被解释变量 y^* 、服从正态分布，有

$$P(y = 0) = P(y^* \leq 0) = \Phi\left(-\frac{\mu}{\sigma}\right) = 1 - \Phi\left(\frac{\mu}{\sigma}\right)$$

$$P(y) = P(y^*) \quad \text{当 } y^* > 0$$

3、归并被解释变量数据模型的最大似然估计

$$\ln L = \sum_{y_i > 0} -\frac{1}{2} \left(\ln(2\pi) + \ln \sigma^2 + \frac{(y_i - B'X_i)^2}{\sigma^2} \right) + \sum_{y_i = 0} \ln \left(1 - \Phi \left(\frac{B'X_i}{\sigma} \right) \right)$$

- 该似然函数由两部分组成：

一部分对应于没有限制的观测值，是经典回归部分；
一部分对应于受到限制的观测值。

- 这是一个非标准的似然函数，它实际上是离散分布与连续分布的混合。

- 如何理解后一部分？

为什么要求和？

- 如果样本观测值不是以0为界，而是以某一个数值 a 为界，则有

$$\begin{array}{ll} y = a & \text{当 } y^* \leq a \\ y = y^* & \text{当 } y^* > a \end{array}$$

$$y^* \sim N(\mu, \sigma^2)$$

估计原理与方法相同。

4、例7.1.2:城镇居民消费模型

人均收入	人均消费	人均收入	人均消费	人均收入	人均消费
1000	1000	1040	1000	1080	1000
1120	1020	4640	2900	6090	3900
1310	1150	4750	2980	6200	3950
1300	1145	4800	2970	6330	4000
1430	1230	4810	3050	6450	4030
1500	1275	4990	3200	6570	4080
1670	1385	5070	3100	6700	4130
2100	1660	5130	3175	6840	4000
2370	1840	5210	3200	7010	4200
2530	1950	5300	2450	7170	4160
2790	2110	5390	3230	7350	4210
2980	2240	5450	3310	7500	4325
3200	2380	5500	3500	7670	4385
3460	2550	5570	3510	7840	4450
3630	2660	5630	3590	8000	4500
3880	2700	5690	3600	8190	4865
4040	2730	5770	3650	8350	4880
4210	2720	5860	3720	8500	4890
4390	2850	5930	3850	8690	4920
4520	2800	6000	3800	8830	4970

OLS估计：将样本看为不受任何限制下随机抽取的样本

Dependent Variable: Y
Method: Least Squares
Date: 01/03/11 Time: 17:29
Sample: 1 60
Included observations: 60

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	571.8328	50.79816	11.25696	0.0000
X	0.513639	0.009360	54.87669	0.0000
R-squared	0.981104	Mean dependent var	3117.083	
Adjusted R-squared	0.980778	S.D. dependent var	1157.512	
S.E. of regression	160.4800	Akaike info criterion	13.02698	
Sum squared resid	1493723.	Schwarz criterion	13.09679	
Log likelihood	-388.8094	F-statistic	3011.451	
Durbin-Watson stat	0.982967	Prob(F-statistic)	0.000000	

$$Y_i = 571.83 + 0.5136X_i \quad i = 1, 2, \dots, 60 \quad R^2 = 0.9811$$

OLS估计：将样本看为在消费水平为1000元的归并样本

Equation Estimation

Specification Options

Equation specification
Dependent variable followed by list of

y c x

Distribution

☒ Normal
☐ Logisti
☐ Extreme Value

Dependent variable censoring points
Enter a number, a series, a series expression, or blank

Left 1000
Right

Left & Right points

☒ Actual censoring val
☐ Zero/one indicator of cen
☐ Truncated sampl.

Estimation settings

Method: CENSORED - Censored or truncated data (tobit)
Sample: 1 60

确定 取消

选择归并值

选择归并样本

Dependent Variable: Y

Method: ML - Censored Normal (TOBIT) (Quadratic hill climbing)

Date: 01/03/11 Time: 17:37

Sample: 1 60

Included observations: 60

Left censoring (value) series: 1000

Convergence achieved after 4 iterations

Covariance matrix computed using second derivatives

	Coefficient	Std. Error	z-Statistic	Prob.
C	545.9460	53.37070	10.22932	0.0000
X	0.517810	0.009767	53.01764	0.0000
Error Distribution				
SCALE: C(3)	163.6581	15.43580	10.60250	0.0000
R-squared	0.980688	Mean dependent var	3117.083	
Adjusted R-squared	0.980011	S.D. dependent var	1157.512	
S.E. of regression	163.6523	Akaike info criterion	12.57143	
Sum squared resid	1526578.	Schwarz criterion	12.67615	
Log likelihood	-374.1429	Hannan-Quinn criter.	12.61239	
Avg. log likelihood	-6.235715			
Left censored obs	3	Right censored obs	0	
Uncensored obs	57	Total obs	60	

$$Y_i = 545.95 + 0.5178X_i \quad i = 1, 2, \dots, 60$$

5、实际模型中的Truncation与Censored

- 时间序列样本，不考虑。
- 截面上的全部个体作为样本，不考虑Truncation。
- 按照抽样理论、选取截面上的部分个体作为样本，不考虑Truncation。
- 按照特定规则、选取截面上的部分个体作为样本，必须考虑Truncation。
- 截面数据作样本，根据样本观测值的经济背景，决定是否考虑Censored。

§ 7.2 二元选择模型

Binary Choice Model

- 一、二元离散选择模型的经济背景
- 二、二元离散选择模型*
- 三、二元Probit离散选择模型及其参数估计
- 四、二元Logit离散选择模型及其参数估计
- 五、二元离散选择模型的检验***

说明

- 在经典计量经济学模型中，被解释变量 y 通常被假定为连续变量。
- 离散被解释变量数据计量经济学模型 (Models with Discrete Dependent Variables) 和离散选择模型 (DCM, Discrete Choice Model)。
- 二元选择模型 (Binary Choice Model) 和 多元选择模型 (Multiple Choice Model)。
- 本节只介绍二元选择模型。

- 离散选择模型起源于Fechner于1860年进行的动物条件二元反射研究。
- 1962年，Warner首次将它应用于经济研究领域，用以研究公共交通工具和私人交通工具的选择问题。
- 70、80年代，离散选择模型被普遍应用于经济布局、企业定点、交通问题、就业问题、购买决策等经济决策领域的研究。
- 模型的估计方法，主要发展于80年代初期。

一、二元离散选择模型的经济背景

实际经济生活中的二元选择问题

- 研究选择结果与影响因素之间的关系。
- 影响因素包括两部分：决策者的属性和备选方案的属性。
- 对于单个方案的取舍：例如，购买者对某种商品的购买决策问题，求职者对某种职业的选择问题，投票人对某候选人的投票决策，银行对某客户的贷款决策。
- 对于两个方案的选择：例如，两种出行方式的选择，两种商品的选择。

二、二元离散选择模型

1、原始模型

- 对于二元选择问题，可建立如下计量经济学模型：
 Y 为观测值为1和0的决策被解释变量；
 X 为解释变量，包括选择对象的属性和选择主体的属性。

$$Y = XB + N \quad y_i = X_i B + \mu_i$$

$$E(\mu_i) = 0 \quad E(y_i) = X_i B$$

$$p_i = P(y_i = 1) \quad 1 - p_i = P(y_i = 0)$$

$$E(y_i) = 1 \cdot P(y_i = 1) + 0 \cdot P(y_i = 0) = p_i$$

$$E(y_i) = P(y_i = 1) = X_i B$$

左/右端矛盾

$$\mu_i = \begin{cases} 1 - X_i B & \text{当 } y_i = 1, \text{ 其概率为 } X_i B \\ -X_i B & \text{当 } y_i = 0, \text{ 其概率为 } 1 - X_i B \end{cases}$$

具有
异方差性

- 由于存在这两方面的问题，所以原始模型不能作为实际研究二元选择问题的模型。
- 需要将原始模型变换为效用模型，这是离散选择模型的关键。

2、效用模型

$$U_i^1 = X_i B^1 + \varepsilon_i^1$$

第*i*个个体 选择1 的效用

$$U_i^0 = X_i B^0 + \varepsilon_i^0$$

第*i*个个体 选择0 的效用

$$U_i^1 - U_i^0 = X_i (B^1 - B^0) + (\varepsilon_i^1 - \varepsilon_i^0)$$



$$y_i^* = X_i B + \mu_i^*$$

作为研究对象的二元选择模型

$$P(y_i = 1) = P(y_i^* > 0) = P(\mu_i^* > -X_i B)$$

- 注意，在模型中：效用不可观测，能够得到的观测值、仍然是选择结果，即1和0。
- 很显然，如果不可观测的 $U^1 > U^0$ ，即对应于观测值为1——因为该个体选择公共交通工具的效用大于选择私人交通工具的效用，他当然要选择公共交通工具；
- 相反，如果不可观测的 $U^1 \leq U^0$ ，即对应于观测值为0——因为该个体选择公共交通工具的效用小于选择私人交通工具的效用，他当然要选择私人交通工具。

3、最大似然估计

- 欲使得效用模型可以估计，就必须为随机误差项 u^* 选择一种特定的概率分布。
- 两种最常用的分布是标准正态分布和逻辑（logistic）分布，于是形成了两种最常用的二元选择模型——**Probit模型**和**Logit模型**。
- 最大似然函数及其估计过程如下：

$$F(-t) = 1 - F(t)$$

标准正态分布或
逻辑分布的 对称性

$$\begin{aligned} P(y_i = 1) &= P(y_i^* > 0) = P(\mu_i^* > -X_i B) \\ &= 1 - P(\mu_i^* \leq -X_i B) \\ &= 1 - F(-X_i B) = F(X_i B) \end{aligned}$$

$$P(y_1, y_2, \dots, y_n) = \prod_{y_i=0} (1 - F(X_i B)) \prod_{y_i=1} F(X_i B)$$

似然函数

$$L = \prod_{i=1}^n (F(\mathbf{X}_i \mathbf{B}))^{y_i} (1 - F(\mathbf{X}_i \mathbf{B}))^{1-y_i}$$

$$\ln L = \sum_{i=1}^n (y_i \ln F(\mathbf{X}_i \mathbf{B}) + (1 - y_i) \ln(1 - F(\mathbf{X}_i \mathbf{B})))$$

$$\frac{\partial \ln L}{\partial \mathbf{B}} = \sum_{i=1}^n \left[\frac{y_i f_i}{F_i} + (1 - y_i) \frac{-f_i}{(1 - F_i)} \right] \mathbf{X}_i = \mathbf{0}$$

1阶极值条件

- 在样本数据支持下，如果知道概率分布函数和概率密度函数，求解该非线性方程组，可得到模型参数估计量。

三、二元Probit离散选择模型及其参数估计

1、标准正态分布的概率分布函数

$$F(t) = \int_{-\infty}^t (2\pi)^{-1/2} \exp(-x^2/2) dx$$

$$f(x) = (2\pi)^{-1/2} \exp(-x^2/2)$$

2、重复观测值不可以得到情况下二元Probit 离散选择模型的参数估计

$$\begin{aligned}\frac{\partial \ln L}{\partial \mathbf{B}} &= \sum_{y_i=0} \frac{-f_i}{1-F_i} \mathbf{X}_i + \sum_{y_i=1} \frac{f_i}{F_i} \mathbf{X}_i \\ &= \sum_{i=1}^n \left(\frac{q_i f(q_i \mathbf{X}_i \mathbf{B})}{F(q_i \mathbf{X}_i \mathbf{B})} \right) \mathbf{X}_i \\ &= \sum_{i=1}^n \lambda_i \mathbf{X}_i \\ &= \mathbf{0}\end{aligned}$$

$$q_i = 2y_i - 1$$

- 关于参数的非线性函数，不能直接求解：需采用完全信息最大似然法中所采用的迭代方法。
- 应用计量经济学软件。
- 这里所谓“重复观测值不可以得到”，是指对每个决策者、只有一个观测值/一个决策结果。

如果有多个观测值，也将其看成为多个不同的决策者。

例7.2.2 贷款决策模型

- 分析与建模：

某商业银行从历史贷款客户中随机抽取78个样本，根据设计的指标体系：分别计算它们的“商业信用支持度”（CC）和“市场竞争地位等级”（CM）；

对它们贷款的结果（JG）采用二元离散变量：

1表示贷款成功、0表示贷款失败。

目的是研究JG与CC、CM之间的关系，并为正确贷款决策提供支持。

• 样本观测值

CC=XY
CM=SC

JG	XY	SC	JGF	JG	XY	SC	JGF	JG	XY	SC	JGF
0	125.0	-2	0.0000	0	1500	-2	0.0000	0	54.00	-1	0.0000
0	599.0	-2	0.0000	0	96.00	0	0.0000	1	42.00	2	1.0000
0	100.0	-2	0.0000	1	-8.000	0	1.0000	0	42.00	0	0.0209
0	160.0	-2	0.0000	0	375.0	-2	0.0000	1	18.00	2	1.0000
0	46.00	-2	0.0000	0	42.00	-1	6.5E-13	0	80.00	1	6.4E-12
0	80.00	-2	0.0000	1	5.000	2	1.0000	1	-5.000	0	1.0000
0	133.0	-2	0.0000	0	172.0	-2	0.0000	0	326.0	2	0.0000
0	350.0	-1	0.0000	1	-8.000	0	1.0000	0	261.0	1	0.0000
1	23.00	0	0.9979	0	89.00	-2	0.0000	1	-2.000	-1	0.9999
0	60.00	-2	0.0000	0	128.0	-2	0.0000	0	14.00	-2	3.9E-07
0	70.00	-1	0.0000	1	6.000	0	1.0000	1	22.00	0	0.9991
1	-8.000	0	1.0000	0	150.0	-1	0.0000	0	113.0	1	0.0000
0	400.0	-2	0.0000	1	54.00	2	1.0000	1	42.00	1	0.9987
0	72.00	0	0.0000	0	28.00	-2	0.0000	1	57.00	2	0.9999
0	120.0	-1	0.0000	1	25.00	0	0.9906	0	146.0	0	0.0000
1	40.00	1	0.9998	1	23.00	0	0.9979	1	15.00	0	1.0000
1	35.00	1	0.9999	1	14.00	0	1.0000	0	26.00	-2	4.4E-16
1	26.00	1	1.0000	0	49.00	-1	0.0000	0	89.00	-2	0.0000
1	15.00	-1	0.4472	0	14.00	-1	0.5498	1	5.000	1	1.0000
0	69.00	-1	0.0000	0	61.00	0	2.1E-12	1	-9.000	-1	1.0000
0	107.0	1	0.0000	1	40.00	2	1.0000	1	4.000	1	1.0000
1	29.00	1	1.0000	0	30.00	-2	0.0000	0	54.00	-2	0.0000
1	2.000	1	1.0000	0	112.0	-1	0.0000	1	32.00	1	1.0000
1	37.00	1	0.9999	0	78.00	-2	0.0000	0	54.00	0	1.4E-07
0	53.00	-1	0.0000	1	0.000	0	1.0000	0	131.0	-2	0.0000
0	194.0	0	0.0000	0	131.0	-2	0.0000	1	15.00	0	1.0000

obs	JG	CC	CM		
1	0.000000	125.0000	.2.000000		
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19	1.000000	15.00000	-1.000000		
20					

Equation Specification

Equation specification

Binary dependent variable followed by list of regressors.

jg c cc cm

Binary estimation method:

☒ Probit
☐ Logit
☐ Extreme value

Estimation settings

Method: BINARY - Binary choice (logit, probit, extreme value)

Sample: 1 78

OK

Cancel

Options

View	Procs	Objects	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
------	-------	---------	-------	------	--------	----------	----------	-------	--------

Dependent Variable: JG

Method: ML - Binary Probit (Quadratic hill climbing)

Date: 11/10/05 Time: 11:04

Sample: 1 78

Included observations: 78

Convergence achieved after 13 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	8.797358	7.544067	1.166129	0.2436
CC	-0.257882	0.228894	-1.126642	0.2599
CM	5.061789	4.458482	1.135317	0.2562
Mean dependent var	0.410256	S.D. dependent var	0.495064	
S.E. of regression	0.090067	Akaike info criterion	0.118973	
Sum squared resid	0.608402	Schwarz criterion	0.209616	
Log likelihood	-1.639954	Hannan-Quinn criter.	0.155259	
Restr. log likelihood	-52.80224	Avg. log likelihood	-0.021025	
LR statistic (2 df)	102.3246	McFadden R-squared	0.968942	
Probability(LR stat)	0.000000			
Obs with Dep=0	46	Total obs	78	
Obs with Dep=1	32			

Estimation Command:

=====

BINARY(D=N) JG C CC CM

Estimation Equation:

=====

JG = 1-@CNORM(-(C(1) + C(2)*CC + C(3)*CM))

Substituted Coefficients:

=====

JG = 1-@CNORM(-(8.797358365 - 0.2578816621*CC + 5.061788659*CM))

输出的估计结果

•该方程表示，当CC和CM已知时，代入方程、可计算贷款成功的概率JGF。

例如，将表中第19个样本观测值CC=15、CM=-1代入方程右边，计算{括号内的值为0.1326552}；

查标准正态分布表，对应于0.1326552的累积正态分布为0.5517；于是，JG的预测值JGF=1-0.5517=0.4483——即对应于该客户，贷款成功的概率为0.4483。

模拟预测

obs	JG	CC	CM	JGF	
1	0.000000	125.0000	-2.000000	0.000000	
2	0.000000	599.0000	-2.000000	0.000000	
3	0.000000	100.0000	-2.000000	0.000000	
4	0.000000	160.0000	-2.000000	0.000000	
5	0.000000	46.00000	-2.000000	0.000000	
6	0.000000	80.00000	-2.000000	0.000000	
7	0.000000	133.0000	-2.000000	0.000000	
8	0.000000	350.0000	-1.000000	0.000000	
9	1.000000	23.00000	0.000000	0.997922	
10	0.000000	60.00000	-2.000000	0.000000	
11	0.000000	70.00000	-1.000000	0.000000	
12	1.000000	-8.000000	0.000000	1.000000	
13	0.000000	400.0000	-2.000000	0.000000	
14	0.000000	72.00000	0.000000	0.000000	
15	0.000000	120.0000	-1.000000	0.000000	
16	1.000000	40.00000	1.000000	0.999803	
17	1.000000	35.00000	1.000000	0.999999	
18	1.000000	26.00000	1.000000	1.000000	
19	1.000000	15.00000	-1.000000	0.447233	
20	0.000000	69.00000	-1.000000	0.000000	

- 预测：

如果有一个新客户，根据客户资料、计算其“商业信用支持度”（XY）和“市场竞争地位等级”（SC），代入模型、就可得到贷款成功概率，以此决定、是否给予该客户贷款。

3、重复观测值可得到情况下二元Probit离散选择模型的参数估计（略）

- 思路

- 对每个决策者有多个重复（例如10次左右）观测值。
- 对第 i 个决策者重复观测 n_i 次，选择 $y_i=1$ 的次数比例为 p_i ，那么可将 p_i 作为真实概率 P_i 的一个估计量。
- 建立“概率单位模型”，采用广义最小二乘法估计。
- 实际中并不常用。

- 对第*i*个决策者重复观测*n*次，选择 $y_i=1$ 的次数比例为 p_i ，那么可以将 p_i 作为真实概率 P_i 的一个估计量。

$$p_i = P_i + e_i = F(X_i B) + e_i$$

定义“观测到的”概率单位

$$E(e_i) = 0$$

$$Var(e_i) = p_i(1 - p_i)/n_i$$

$$v_i = F^{-1}(p_i) = F^{-1}(P_i + e_i)$$



$$F^{-1}(P_i + e_i) = F^{-1}(P_i) + \frac{e_i}{f(F^{-1}(P_i))}$$



$$v_i = F^{-1}(P_i) + u_i$$

$$E(u_i) = 0$$

$$Var(u_i) = \frac{P_i(1 - P_i)}{n_i (f(F^{-1}(P_i)))^2}$$

$$F^{-1}(P_i) = X_i B$$

$$P_i = F(X_i B)_i$$

$$v_i = X_i B + u_i$$

$$V = XB + U$$

$$\hat{B} = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} V$$

V 的观测值通过求解标准正态分布的概率分布函数的反函数得到

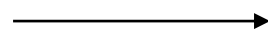
$$p_i = \int_{-\infty}^{v_i} (2\pi)^{-1/2} \exp(-t^2/2) dt$$

实际观测得到的

四、二元Logit离散选择模型及其参数估计

1、逻辑分布的概率分布函数

$$F(t) = \frac{1}{1 + e^{-t}}$$

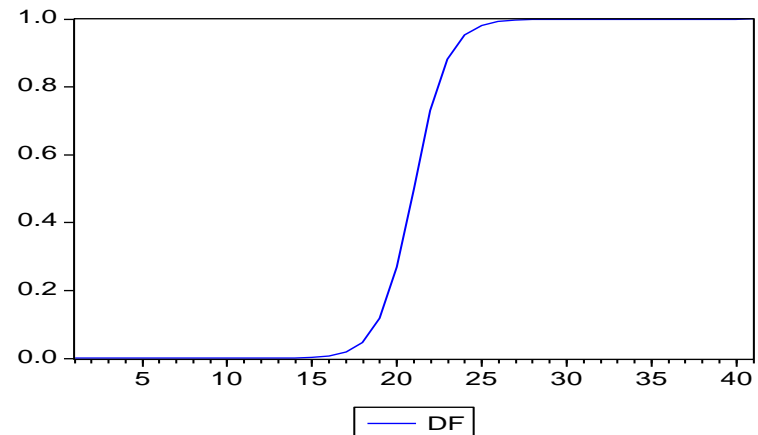
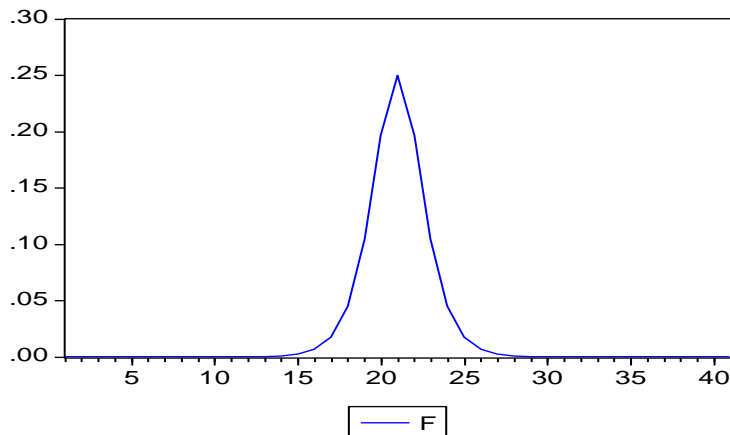


$$F(t) = \frac{e^t}{1 + e^t} = \Lambda(t)$$

$$f(t) = \frac{e^{-t}}{(1 + e^{-t})^2}$$



$$f(t) = \frac{e^t}{(1 + e^t)^2} = \Lambda(t)(1 - \Lambda(t))$$



Börsch-Supan于1987年指出：

- 若按照效用最大化为依据、进行选择/决策，具有极值的逻辑分布是较好的选择；

这种情况下的二元选择模型，应采用Logit模型。

2、重复观测值不可以得到情况下二元logit离散选择模型的参数估计

$$\begin{aligned}\frac{\partial \ln L}{\partial \mathbf{B}} &= \sum_{i=1}^n \left[\frac{y_i f_i}{F_i} + (1 - y_i) \frac{-f_i}{(1 - F_i)} \right] \mathbf{X}_i \\ &= \sum_{i=1}^n (y_i - \Lambda(\mathbf{X}_i \mathbf{B})) \mathbf{X}_i = \mathbf{0}\end{aligned}$$

- 关于 参数B的非线性方程组，不能直接求解：需采用完全信息最大似然法中所采用的迭代方法。
- 应用计量经济学软件。

obs	JG	CC	CM			
1	0.000000	125.0000	-2.000000			
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19	1.000000	15.00000	-1.000000			

Equation Specification

Equation specification

Binary dependent variable followed by list of regressors.

lg c cc cm

Binary estimation method: ☐ Probit ☒ Logit ☐ Extreme value

Estimation settings

Method: BINARY - Binary choice (logit, probit, extreme value)

Sample: 1 78

OK

Cancel

Options

View	Procs	Objects	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
------	-------	---------	-------	------	--------	----------	----------	-------	--------

Dependent Variable: JG

Method: ML - Binary Logit (Newton-Raphson)

Date: 11/10/05 Time: 16:53

Sample: 1 78

Included observations: 78

Convergence achieved after 13 iterations

Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	16.11426	14.56353	1.106481	0.2685
CC	-0.465035	0.431760	-1.077068	0.2814
CM	9.379903	8.712437	1.076611	0.2817
Mean dependent var	0.410256	S.D. dependent var		0.495064
S.E. of regression	0.091187	Akaike info criterion		0.120325
Sum squared resid	0.623629	Schwarz criterion		0.210968
Log likelihood	-1.692674	Hannan-Quinn criter.		0.156611
Restr. log likelihood	-52.80224	Avg. log likelihood		-0.021701
LR statistic (2 df)	102.2191	McFadden R-squared		0.967943
Probability(LR stat)	0.000000			
Obs with Dep=0	46	Total obs		78
Obs with Dep=1	32			

View	Procs	Objects	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
------	-------	---------	-------	------	--------	----------	----------	-------	--------

Estimation Command:

=====

BINARY(D=L,R) JG C CC CM

Estimation Equation:

=====

$JG = 1 - @LOGIT(-(C(1) + C(2)*CC + C(3)*CM))$

Substituted Coefficients:

=====

$JG = 1 - @LOGIT(-(16.11426399 - 0.4650347429*CC + 9.379903458*CM))$

obs	JG	CC	CM	JGFF	
1	0.000000	125.0000	-2.000000	0.000000	
2	0.000000	599.0000	-2.000000	0.000000	
3	0.000000	100.0000	-2.000000	0.000000	
4	0.000000	160.0000	-2.000000	0.000000	
5	0.000000	46.00000	-2.000000	3.64E-11	
6	0.000000	80.00000	-2.000000	0.000000	
7	0.000000	133.0000	-2.000000	0.000000	
8	0.000000	350.0000	-1.000000	0.000000	
9	1.000000	23.00000	0.000000	0.995586	
10	0.000000	60.00000	0.000000	5.41E-14	
11	0.000000	70.00000	0.000000	6.13E-12	
12	1.000000	-8.00000	0.000000	1.000000	
13	0.000000	400.0000	0.000000	0.000000	
14	0.000000	72.00000	0.000000	2.86E-08	
15	0.000000	120.0000	-1.000000	0.000000	
16	1.000000	40.00000	1.000000	0.998986	
17	1.000000	35.00000	1.000000	0.999901	
18	1.000000	26.00000	1.000000	0.999998	
19	1.000000	15.00000	-1.000000	0.440000	
20	0.000000	69.00000	-1.000000	9.76E-12	

结果类似

Probit

0.999999

1.000000

0.447233

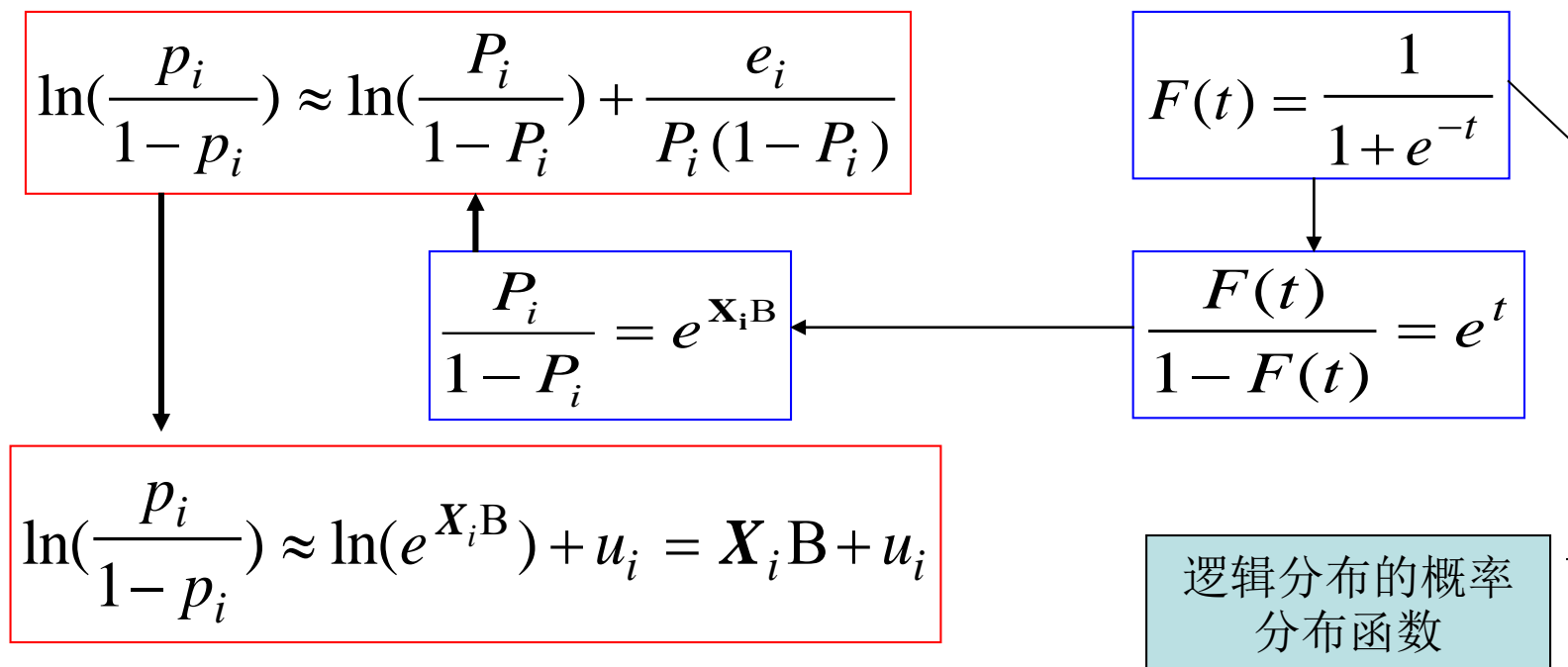
0.000000

3、重复观测值可以得到情况下二元logit离散选择模型的参数估计 (略)

- 思路

- 对每个决策者有多个重复（例如10次左右）观测值。
- 对第 i 个决策者重复观测 n_i 次，选择 $y_i=1$ 的次数比例为 p_i ，那么可以将 p_i 作为真实概率 P_i 的一个估计量。
- 建立“对数成败比例模型”，采用广义最小二乘法估计。
- 实际中并不常用。

- 用样本重复观测得到的 \mathbf{p}_i 构成“成败比例”，取对数并进行台劳展开，有



$$\mathbf{v}_i = \mathbf{X}_i\mathbf{B} + u_i$$

$$\mathbf{V} = \mathbf{X}\mathbf{B} + \mathbf{U}$$

$$\hat{\mathbf{B}} = (\mathbf{X}'\hat{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}^{-1}\mathbf{V}$$

五、二元离散选择模型的检验**

1、计量经济学模型中的两类检验统计量

- 基于LS

- R^2
- 总体显著性F检验
- 约束回归的F检验

- 基于ML

- Wald
- LR (likelihood ratio)
- LM (lagrange multiplier)

- 原理类似 (无约束回归 VS 受约束回归)

2、拟合检验

$$LRI = 1 - \frac{\ln L}{\ln L_0}$$

$$\ln L_0 = n(P \ln P + (1 - P) \ln(1 - P))$$

- P: 样本观测值中, 被解释变量等于1的比例。
L₀: 模型中“所有解释变量的系数都为0”时的似然函数值。
- **LRI=1**, 即L=1, 完全拟合、联合概率100%。
LRI=0, 即L=L₀, 所有解释变量、完全不显著, 完全不拟合。

View	Procs	Objects	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
------	-------	---------	-------	------	--------	----------	----------	-------	--------

Dependent Variable: JG

Method: ML - Binary Probit (Quadratic hill climbing)

Date: 11/10/05 Time: 11:04

Sample: 1 78

Included observations: 78

Convergence achieved after 13 iterations

Covariance matrix computed using second derivatives

$\text{LnL} = -1.639954$

$\text{LnL}_0 = -52.80224$

LRI=0.968942，不是LR

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	8.797358	7.544067	1.166129	0.2436
CC	-0.257882	0.228894	-1.126642	0.2599
CM	5.061789	4.458482	1.135317	0.2562
Mean dependent var	0.410256	S.D. dependent var	0.495064	
S.E. of regression	0.090067	Akaike info criterion	0.118973	
Sum squared resid	0.608402	Schwarz criterion	0.209616	
Log likelihood	-1.639954	Hannan-Quinn criter.	0.155259	
Restr. log likelihood	-52.80224	Avg. log likelihood	-0.021025	
LR statistic (2 df)	102.3246	McFadden R-squared	0.968942	
Probability(LR stat)	0.000000			
Obs with Dep=0	46	Total obs	78	
Obs with Dep=1	32			

3、省略变量检验**

- 经典模型中采用的“变量显著性t检验”，仍然是有效的。
- 如果“省略的变量”与“保留的变量”不是正交的，那么对参数估计量将产生影响，需要进一步检验这种省略是否恰当。

零假设为： $H_0 : \mathbf{Y}^* = \mathbf{X}_1 \mathbf{B}_1 + \mu^*$

备择假设为： $H_1 : \mathbf{Y}^* = \mathbf{X}_1 \mathbf{B}_1 + \mathbf{X}_2 \mathbf{B}_2 + \mu^*$

用于检验的统计量为 Wald 统计量、LR 统计量和 LM 统计量，具体计算方法如下：

$$W = \hat{\mathbf{B}}_2' \mathbf{V}_2^{-1} \hat{\mathbf{B}}_2$$

其中 $\mathbf{V}_2 = \text{AsyVar}(\hat{\mathbf{B}}_2)$ 。

$$LR = -2(\ln \hat{L}_0 - \ln \hat{L}_1)$$

其中 \hat{L}_0 、 \hat{L}_1 分别为 H_0 情形和 H_1 情形下的似然函数值的估计量。

$$LM = \mathbf{g}_0' \mathbf{V}_0^{-1} \mathbf{g}_0$$

其中 \mathbf{g}_0 是 H_1 情形下的对数似然函数对参数估计量的一阶导数向量，

用 H_0 情形下的最大似然参数估计量代入计算； \mathbf{V}_0 是 H_1 情形下参数

最大似然估计量的方差矩阵估计量。

若 X_2 中的变量省略后、对参数估计量没有影响，那么 H_1 和 H_0 情况下的对数最大似然函数值应相差不大：此时 **LR 统计量的值很小**，自然会小于临界值，不拒绝 H_0 。

- 检验步骤：
 - 首先进行“受约束模型”的估计
 - 选择系数检验
 - 引入省略的变量
 - 判断

省略CC、只保留CM，估计模型：

View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: JG

Method: ML - Binary Probit (Quadratic hill climbing)

Date: 03/21/09 Time: 15:56

Sample: 1 78

Included observations: 78

Convergence achieved after 4 iterations

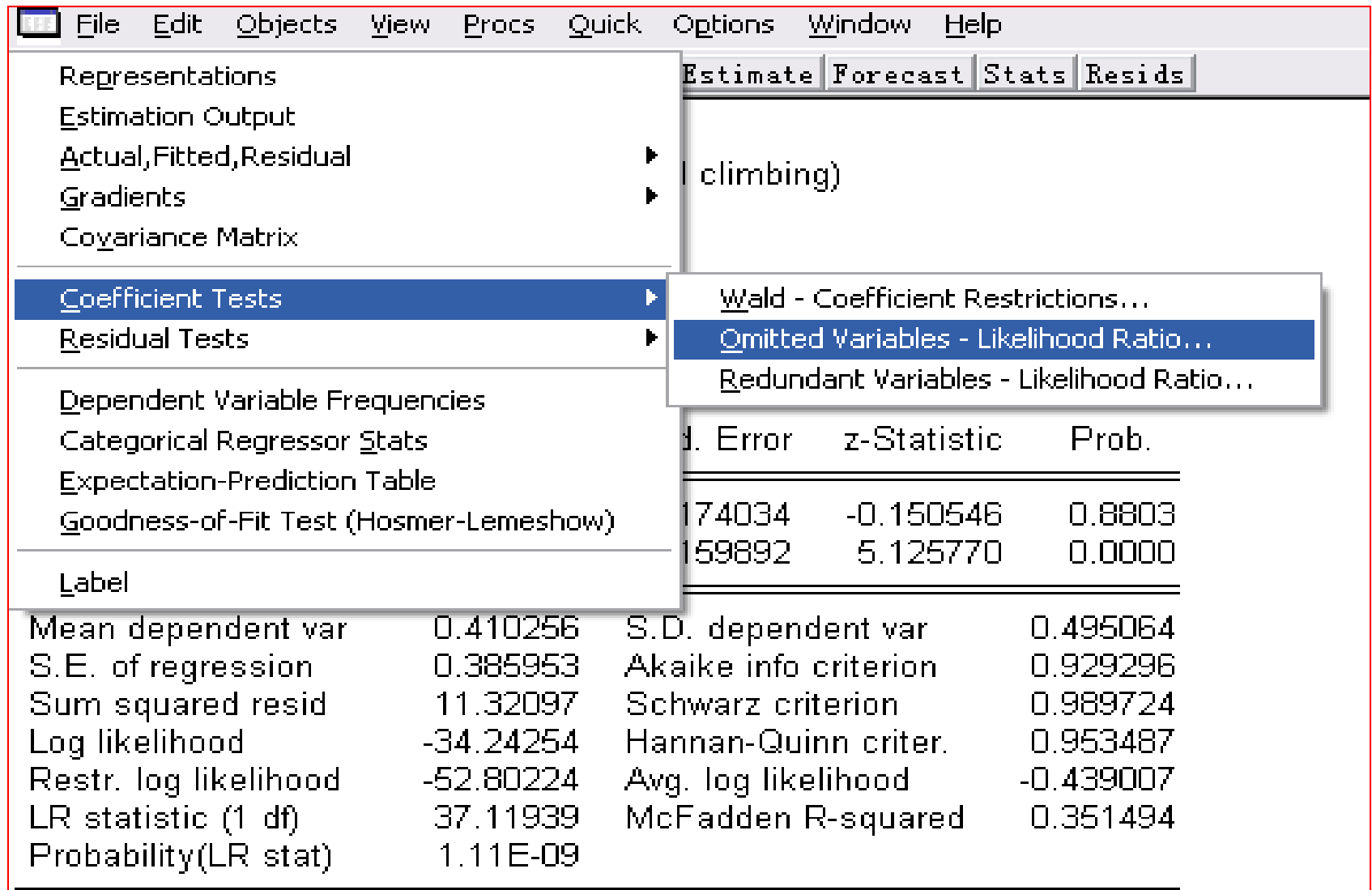
Covariance matrix computed using second derivatives

拟 $R^2 = 0.351494$

有较大下降

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.026200	0.174034	-0.150546	0.8803
CM	0.819570	0.159892	5.125770	0.0000
Mean dependent var	0.410256	S.D. dependent var	0.495064	
S.E. of regression	0.385953	Akaike info criterion	0.929296	
Sum squared resid	11.32097	Schwarz criterion	0.989724	
Log likelihood	-34.24254	Hannan-Quinn criter.	0.953487	
Restr. log likelihood	-52.80224	Avg. log likelihood	-0.439007	
LR statistic (1 df)	37.11939	McFadden R-squared	0.351494	
Probability(LR stat)	1.11E-09			
Obs with Dep=0	46	Total obs	78	
Obs with Dep=1	32			

选择“Omitted Variables-LR Test”



The screenshot shows the Stata software interface. The 'Coefficient Tests' menu is open, and the 'Omitted Variables - Likelihood Ratio...' option is selected. The 'Estimate' tab is active in the background window.

Menu Structure:

- Representations
- Estimation Output
 - Actual, Fitted, Residual
 - Gradients
 - Covariance Matrix
- Coefficient Tests** (selected)
 - Wald - Coefficient Restrictions...
 - Omitted Variables - Likelihood Ratio...** (selected)
 - Redundant Variables - Likelihood Ratio...
- Residual Tests
- Dependent Variable Frequencies
- Categorical Regressor Stats
- Expectation-Prediction Table
- Goodness-of-Fit Test (Hosmer-Lemeshow)
- Label

Background Window (Estimate Tab):

	d. Error	z-Statistic	Prob.
	174034	-0.150546	0.8803
	159892	5.125770	0.0000

Bottom Panel:

Mean dependent var	0.410256	S.D. dependent var	0.495064
S.E. of regression	0.385953	Akaike info criterion	0.929296
Sum squared resid	11.32097	Schwarz criterion	0.989724
Log likelihood	-34.24254	Hannan-Quinn criter.	0.953487
Restr. log likelihood	-52.80224	Avg. log likelihood	-0.439007
LR statistic (1 df)	37.11939	McFadden R-squared	0.351494
Probability(LR stat)	1.11E-09		

引入CC

View | Procs | Objects | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: JG
 Method: ML - Binary Probit (Quadratic hill climbing)
 Date: 03/21/2016
 Sample: 1
 Included observations: 78
 Convergence: 1
 Covariance matrix: Positive definite

Variable	Mean	Std. Dev.	Minimum	Maximum	Prob.
cc	0.46	0.50	0.00	1.00	0.8803
cc	0.46	0.50	0.00	1.00	0.0000

Mean dependent variable = 0.495064
 S.E. of regression = 0.385953
 Sum squared resid = 11.32097
 Log likelihood = -34.24254
 Restr. log likelihood = -52.80224
 LR statistic (1 df) = 37.11939
 Probability(LR stat) = 1.11E-09

Criterion	Value
Akaike info criterion	0.929296
Schwarz criterion	0.989724
Hannan-Quinn criter.	0.953487
Avg. log likelihood	-0.439007
McFadden R-squared	0.351494

Obs with Dep=0	46	Total obs	78
Obs with Dep=1	32		

Omitted-Redundant Variable Test

One or more test series

cc

OK Cancel

View	Procs	Objects	Print	Name	Freeze	Estimate	Forecast	Stats	Resids
------	-------	---------	-------	------	--------	----------	----------	-------	--------

Omitted Variables: CC

F-statistic	1320.579	Probability	0.000000
Log likelihood ratio	65.20518	Probability	0.000000

Test Equation:

Dependent Variable: JG

Method: ML - Binary Probit (Quadratic hill climbing)

Date: 03/21/09 Time: 16:04

Sample: 1 78

Included observations: 78

Convergence achieved after 13 iterations

Covariance matrix computed using second derivatives

拒绝CC系数为0、
的 H_0 假设

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	8.797358	7.544067	1.166129	0.2436
CM	5.061789	4.458482	1.135317	0.2562
CC	-0.257882	0.228894	-1.126642	0.2599
Mean dependent var	0.410256	S.D. dependent var	0.495064	
S.E. of regression	0.090067	Akaike info criterion	0.118973	
Sum squared resid	0.608402	Schwarz criterion	0.209616	
Log likelihood	-1.639954	Hannan-Quinn criter.	0.155259	
Restr. log likelihood	-52.80224	Avg. log likelihood	-0.021025	
LR statistic (2 df)	102.3246	McFadden R-squared	0.968942	
Probability(LR stat)	0.000000			

4、异方差性检验*

- 截面数据样本，容易存在异方差性。
- 假定异方差结构为：

$$\text{Var}(\varepsilon|\mathbf{X}, \mathbf{Z}) = (\exp(\mathbf{Z}'\boldsymbol{\gamma}))^2$$

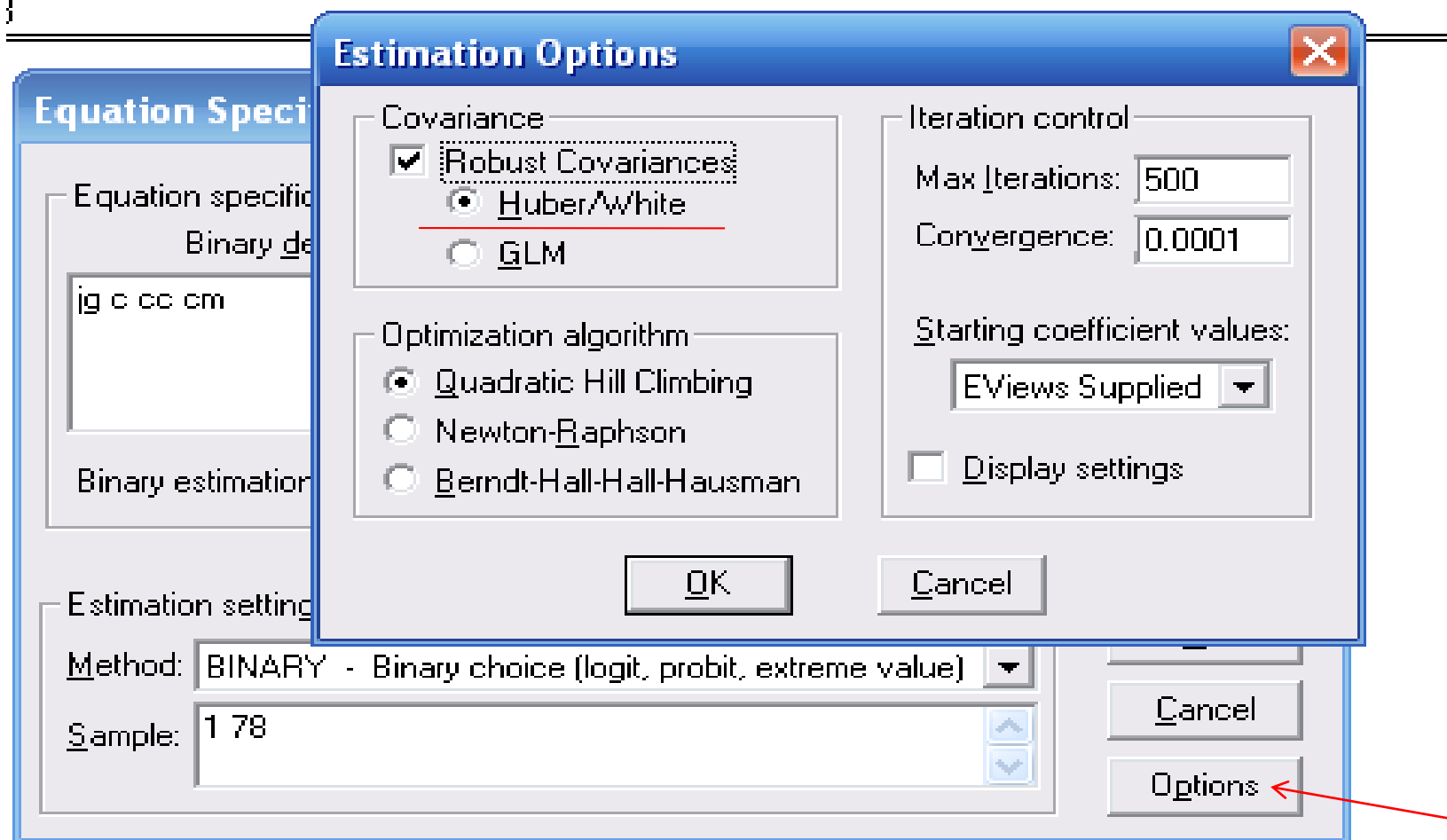
$$H_0 : \boldsymbol{\gamma} = \mathbf{0}$$

- 采用LM检验：

将解释变量、分为两类：
Z为只与个体特征有关的变量；
显然，异方差与这些变量相关。

将异方差检验问题，变
为一个约束检验问题

- 一般都存在异方差;
- 不检验, 直接采用 White/Huber 修正、进行估计



View Procs Objects Print Name Freeze E

Dependent Variable: JG

Method: ML - Binary Probit (Quadratic hill)

Date: 03/21/09 Time: 17:03

Sample: 1 78

Included observations: 78

Convergence achieved after 13 iterations

QML (Huber/White) standard errors & covariance

Coefficient	Std. Error	z-Statistic	Prob.
8.797358	7.544067	1.166129	0.2436
-0.257882	0.228894	-1.126642	0.2599
5.061789	4.458482	1.135317	0.2562

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	8.797358	1.350230	6.515451	0.0000
CC	-0.257882	0.044167	-5.838766	0.0000
CM	5.061789	1.005360	5.034801	0.0000
Mean dependent var	0.410256	S.D. dependent var	0.495064	
S.E. of regression	0.090067	Akaike info criterion	0.118973	
Sum squared resid	0.608402	Schwarz criterion	0.209616	
Log likelihood	-1.639954	Hannan-Quinn criter.	0.155259	
Restr. log likelihood	-52.80224	Avg. log likelihood	-0.021025	
LR statistic (2 df)	102.3246	McFadden R-squared	0.968942	
Probability(LR stat)	0.000000			
Obs with Dep=0	46	Total obs	78	
Obs with Dep=1	32			

5、分布检验 (略)

- 检验关于分布的假设 (probit、logit) 。
- 一般不进行该项检验，具体见相关教科书：
(Greene, P682) 。

- β : 模型1的参数;
- γ : 模型2的参数。

- 组合模型的似然函数 (略) :

$$L_{\rho} = \frac{[(1-\alpha)L_1(y|X, \beta)^{\rho} + \alpha L_2(y|X, \gamma)^{\rho}]^{\frac{1}{\rho}}}{\int_z [(1-\alpha)L_1(z|X, \beta)^{\rho} + \alpha L_2(z|X, \gamma)^{\rho}]^{\frac{1}{\rho}} dz}$$

$$H_0 : \alpha = 0$$

$$H_1 : \alpha > 0$$

构造LM统计量，如果不拒绝 H_0 假设，表明模型1是适当的。

6、回代/预测检验

- 当二元离散选择模型被估计后，将所有样本的解释变量观测值代入模型，计算“每个样本被解释变量、选择1的概率”，再与被解释变量的实际观测值进行比较——以判断模型的预测（回代）效果，是一种实际有效的模型检验方法。
- 概率阈值的3种选择：
 - 朴素选择： $p=0.5$ (1、0的样本相当时)
 - 先验选择： $p=$ （选1的样本数/全部样本）（全样本时）
 - 最优阈值：犯第一类错误最小原则

例7.2.2

- 朴素选择，即以0.5为阈值：除了2个样本外，所有样本都通过了回代检验。
- 先验选择，即以选择1的样本比例0.41、为阈值：除了1个样本外，所有样本都通过了回代检验。

实例1—财务欺诈识别模型

我国上市公司财务欺诈识别模型

- 样本：

年度报告审计意见为“无法发表意见”或者“证监会立案调查”等公司属于财务欺诈样本；

年度报告审计意见为“标准无保留意见”和财务报表满足“ $\text{利润} \times \text{现金流量} > 0$ ”的公司属于配对样本。

- 解释变量：

开始选择11个财务指标；

通过t检验，确定6个指标：资产负债率、资产毛利率、资产周转率、营运资金比率、应收账款周转率、经营活动现金流量/资产额。

- 样本：财务欺诈公司30，非财务欺诈公司30
- 采用犯第一类错误最小原则，确定最优阈值为0.68
 - 欺诈样本中， $p < 0.68$ ，26个，占86.7%
 - 非欺诈样本中， $p > 0.68$ ，25个，占83.3%

实例2—上市公司并购

- 被解释变量：当年发生并购行为为1，反之为0。
- 解释变量：净利润率、.....，全流通虚变量。
- 试图研究全流通对并购的影响。
- 样本：1994-2008上市公司，并购样本731、非并购样本9835。
- 采用先验原则， $P=5\%$
- 模拟结果：
 - 并购样本中： $p>5\%$ 占53%
 - 非并购样本中： $p<5\%$ 占72%

§ 7.3 Panel Data 模型

一、概述

二、模型的设定——F检验***

三、固定影响变截距模型**

四、固定影响变系数模型

不讨论

一、Panel Data 模型概述

1、关于Panel Data Model

- 独立的计量经济学分支
 - 较多地用于宏观经济分析——统计数据
 - 也可以用于微观经济分析——调查数据
- 几种翻译
 - 面板数据模型
 - 综列数据模型
 - 平行数据模型
 - 时空数据模型

- 常用Panel Data 模型

- 变截矩模型(Variable-Intercept Models)

- 固定影响(Fixed-Effects)

- 随机影响(Random-Effects)

- 变系数模型(Variable-Coefficient Models)

- 固定影响

- 随机影响

- 动态变截矩模型(Dynamic Models with Variable Intercepts)

- 固定影响

- 随机影响

- 其它Panel Data 模型 (略)
 - 联立方程Panel Data模型
 - 离散数据Panel Data模型
 - 选择性样本Panel Data模型
 - Panel Data单位根检验和协整检验

2、计量经济分析中的Panel Data问题

- 研究目的的需要
 - 建立计量经济学模型进行经济分析，经常发现，只利用截面数据或只利用时间序列数据、不能满足分析目的的需要。
 - 例如，如果分析生产成本问题（分企业）。
 - 例如，分析目前我国的结构性问题（分行业）。
- 数据信息的充分利用
 - 在计量经济分析中，利用信息越多越有效。
 - 利用Panel Data比仅利用截面数据或时间序列数据更有效。

二、Panel Data 模型的设定

——F检验***

1.单方程Panel Data模型常见的三种情形

$$y_{it} = \alpha_i + x_{it}\beta_i + u_{it}$$

$$i = 1, \dots, n$$

$$t = 1, \dots, T$$

- 情形1：变系数模型。

在横截面上、对不同观测点 i ：既有个体影响，又有经济结构变化——因此结构参数 β_i ，在不同横截面单位上是不同的。

$$y_{it} = \alpha_i + x_{it}\beta + u_{it}$$

$$i = 1, \dots, n$$

$$t = 1, \dots, T$$

- 情形2：变截距模型。

在横截面上、对不同观测点：仅有个体影响 α_i ，但无结构变化——个体影响、表现为模型中被忽略的“反映个体差异的变量”的影响，又分为固定影响和随机影响两种情况。

$$y_{it} = \alpha + x_{it}\beta + u_{it}$$

$$i = 1, \dots, n$$

$$t = 1, \dots, T$$

- 情形3:

在横截面上、对不同观测点：既无个体影响，又无结构变化。

此时普通OLS可给出了一致有效估计，相当于
“将多个时期的截面数据、放在一起”作为样本数据。

$$y_{it} = \alpha + x_{it}\beta_i + u_{it}$$

$$i = 1, \dots, n$$

$$t = 1, \dots, T$$

- 情形4：不讨论（相当于特殊的变系数模型）。

Because it is seldom meaningful to ask if the intercepts are the same when the slopes are unequal, we shall ignore the type of restrictions.

$$y_{it} = \alpha_t + x_{it}\beta_t + u_{it}$$

$$i = 1, \dots, n$$

$$t = 1, \dots, T$$

- 情形5:

在时间序列上、对不同观测时刻 t : 既有个体影响, 又有结构变化。

在理论方法方面与情形1相同, 不讨论。

$$y_{it} = \alpha_i + \gamma_t + x_{it}\beta + u_{it}$$

$$i = 1, \dots, n$$

$$t = 1, \dots, T$$

- 情形6:

对不同观测点 i 、对不同观测时刻 t : 截距均存在个体效应;

有实践, 无理论。

2.F检验

- F检验的思路：将情形2、3，分别视为对情形1施加了参数约束。
- 经典模型中的约束检验：

$$F = \frac{(RSS_R - RSS_U)/(k_U - k_R)}{RSS_U/(n - k_U - 1)} \sim F(k_U - k_R, n - k_U - 1)$$

- 假设1：斜率在不同的横截面样本点上和时间上都相同，但截距不相同。（情形2）
- 假设2：截距和斜率在不同的横截面样本点和时间上都相同。（情形3）
- 若接收了假设2，则没有必要进行进一步的检验；若拒绝了假设2，就应该检验假设1，判断是否斜率都相等——此时，若假设1被拒绝，就应该采用情形1的模型；若假设1被接受，则采用情形2。
- 主要工作是计算3种情形的残差平方和。
下面列出了计算过程，可忽略。因为在实际应用中，分别对3种情形的模型进行估计，以得到残差平方和。

• F统计量的计算方法

$$\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$$

$$\bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}$$

$$W_{xx,i} = \sum_{t=1}^T (x_{it} - \bar{x}_i)'(x_{it} - \bar{x}_i)$$

$$W_{xy,i} = \sum_{t=1}^T (x_{it} - \bar{x}_i)'(y_{it} - \bar{y}_i)$$

$$W_{yy,i} = \sum_{t=1}^T (y_{it} - \bar{y}_i)^2$$

第i群的残差平方和

$$y_{it} = \alpha_i + x_{it}\beta_i + u_{it}$$

的残差平方和

$$RSS_i = W_{yy,i} - W_{xy,i}' W_{xx,i}^{-1} W_{xy,i}$$

$$S_1 = \sum_{i=1}^n RSS_i$$

$$W_{xx} = \sum_{i=1}^n W_{xx,i}$$

$$W_{xy} = \sum_{i=1}^n W_{xy,i}$$

$$W_{yy} = \sum_{i=1}^n W_{yy,i}$$

$$S_2 = W_{yy} - W_{xy}' W_{xx}^{-1} W_{xy}$$

$y_{it} = \alpha_i + x_{it} \beta + u_{it}$
 的残差平方和

$$\bar{y} = \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T y_{it}$$

$$\bar{x} = \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T x_{it}$$

$$T_{xx} = \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x})'(x_{it} - \bar{x})$$

$$T_{xy} = \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x})'(y_{it} - \bar{y})$$

$$T_{yy} = \sum_{i=1}^n \sum_{t=1}^T (y_{it} - \bar{y})^2$$

$$y_{it} = \alpha + x_{it}\beta + u_{it}$$

的残差平方和

$$S_3 = T_{yy} - T'_{xy} T_{xx}^{-1} T_{xy}$$

- 检验假设2的F统计量

$$F_2 = \frac{(S_3 - S_1) / [(n-1)(K+1)]}{S_1 / [nT - n(K+1)]} \sim F[(n-1)(K+1), n(T-K-1)]$$

从直观上看，如 $S_3 - S_1$ 很小， F_2 则很小，低于临界值，接受 H_2 。 S_3 为截距、系数都不变的模型的残差平方和， S_1 为截距、系数都变化的模型的残差平方和。

- 检验假设1的F统计量

$$F_1 = \frac{(S_2 - S_1)/[(n-1)K]}{S_1/[nT - n(K+1)]} \sim F[(n-1)K, n(T-K-1)]$$

从直观上看，如 $S_2 - S_1$ 很小， F_1 则很小，低于临界值，接受 H_1 。 S_2 为截距变化、系数不变的模型的残差平方和， S_1 为截距、系数都变化的模型的残差平方和。

- Eviews 不能自动进行F检验，需要单独进行检验。
- 从理论上讲，模型设定检验是不可缺少的。
- 在实际应用中，最容易被忽视。

二、固定影响变截距模型

1. LSDV模型及其参数估计

$$y_{it} = \alpha_i + x_{it}\beta + u_{it}$$

T阶
向量

$$y_i = e\alpha_i + X_i\beta + u_i$$

$$e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{T \times 1}$$

(T×n)
阶向量

$$y = [d_1, d_2, \dots, d_n, X] \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + u$$

$$y = D\alpha + X\beta + u$$

$$[d_1, d_2, \dots, d_n] = \begin{pmatrix} e & 0 & \dots & 0 \\ 0 & e & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e \end{pmatrix}_{nT \times n}$$

- 该模型通常被称为最小二乘虚拟变量(LSDV)模型,有时也称之为协方差分析模型(解释变量既有定量的,也有定性的)。
- 如果 n 充分小,此模型可以当作具有 $(n+K)$ 个参数的多元回归,参数可由普通最小二乘进行估计。
- 当 n 很大,甚至成千上万, OLS 计算可能超过任何计算机的存储容量。此时,可用分块回归的方法进行计算。
- 分块回归的思路是:设法消去数量很多的 α_i ,估计 β ,然后利用每个个体的时间序列数据计算 α_i 。

$$e'e = T$$

令 $Q = I_T - \frac{1}{T}ee'$, 因为 $I_T e = \frac{1}{T}ee'e$, 所以 $Qe = 0$ 。则有

$$Qy_i = Qe\alpha_i + QX_i\beta + Qu_i = QX_i\beta + Qu_i$$

于是

这是一个不含高阶的
 Q_i , 只含 β 的模型,
可以估计 β

$$X_i'Qy_i = X_i'QX_i\beta + X_i'Qu_i$$

$$\sum_i X_i'Qy_i = \left(\sum_i X_i'QX_i\right)\beta + \sum_i X_i'Qu_i$$

$$\hat{\beta}_{CV} = \left[\sum_{i=1}^n X_i'QX_i \right]^{-1} \left[\sum_{i=1}^n X_i'Qy_i \right]$$

β 的协方差估计是无偏的，且当n或T趋于无穷大时，为一致估计。它的协方差阵为：

$$Var(\hat{\beta}_{CV}) = \sigma_u^2 \left[\sum_{i=1}^n X_i' Q X_i \right]^{-1}$$

截距的估计是无偏估计，
且仅当T趋于无穷大时
为一致估计。

$$\hat{\alpha}_i = \bar{y}_i - \bar{X}_i \hat{\beta}_{CV}$$

$$Var(\hat{\alpha}_i) = \frac{\sigma_u^2}{T} + \bar{X}_i Var(\hat{\beta}_{CV}) \bar{X}_i'$$

随机项方差
估计量

$$s^2 = \sum_{i=1}^n \sum_{t=1}^T (y_{it} - \hat{\alpha}_i - x_{it} \hat{\beta}_{CV})^2 / (nT - n - K)$$

- 分块估计的思路：
 - 首先构造1个不含 α_i ，只包含 β 的模型，对其进行OLS。
 - 然后分别在每个个体上计算 α_i ，分块的含义体现于此。
- 通过F统计量检验变截距假设 $\alpha_i = \alpha_j$

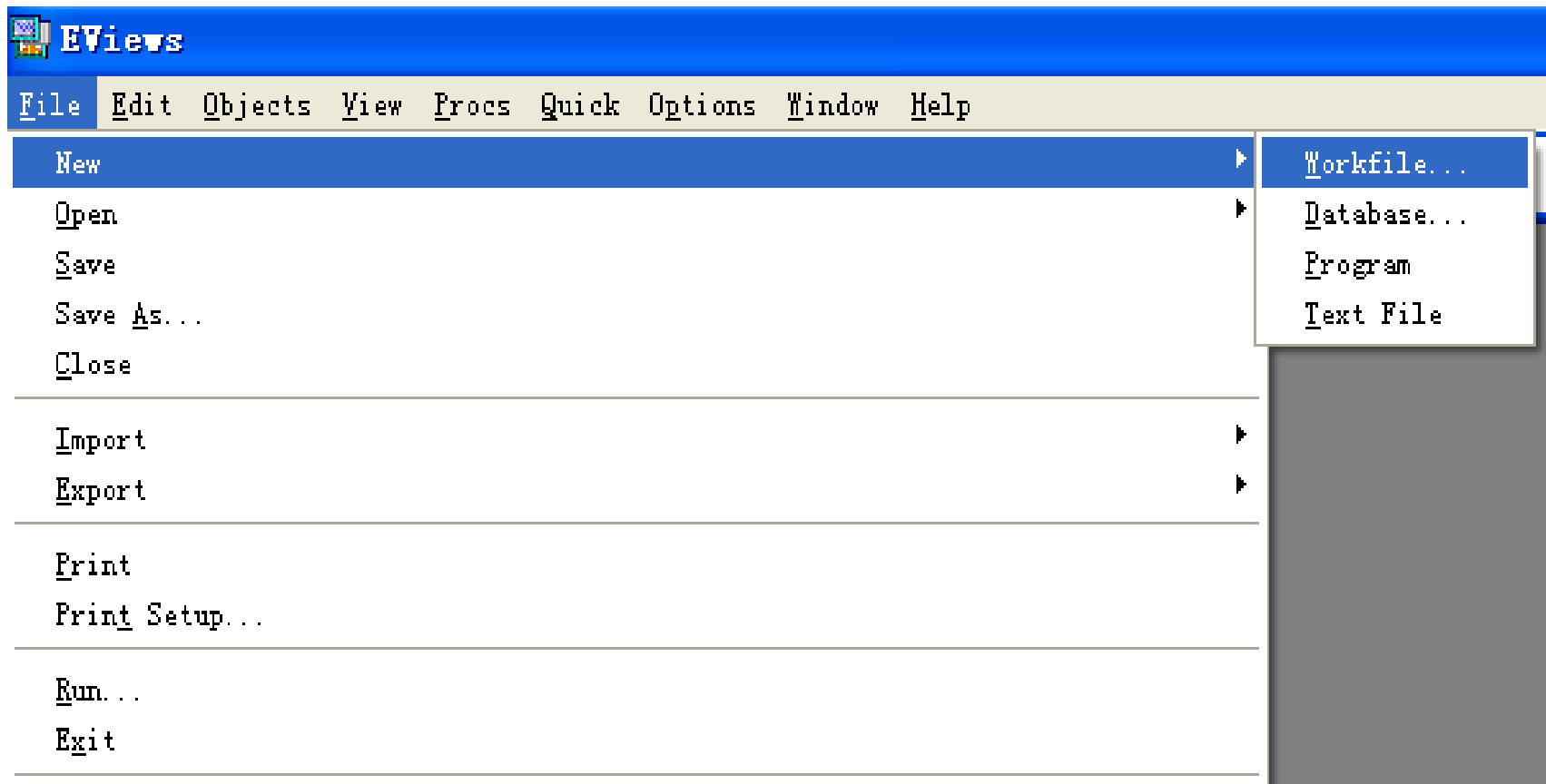
$$F = \frac{(R_u^2 - R_p^2)/(n-1)}{(1 - R_u^2)/(nT - n - K)}$$

3.用Eviews估计固定影响变截距模型演示

- 利用北京、天津、河北、山西、内蒙5地区**1997-2003**年消费总额与**GDP**的数据，消费总额为被解释变量，以**GDP**为解释变量，建立一元线性**Panel Data**模型。

如何在Eviews中建立Panel Data文件？

打开Eviews，建立新工作文件



输入Panel Data的起止时间（1997、2003）



Workfile Range

Workfile frequency:

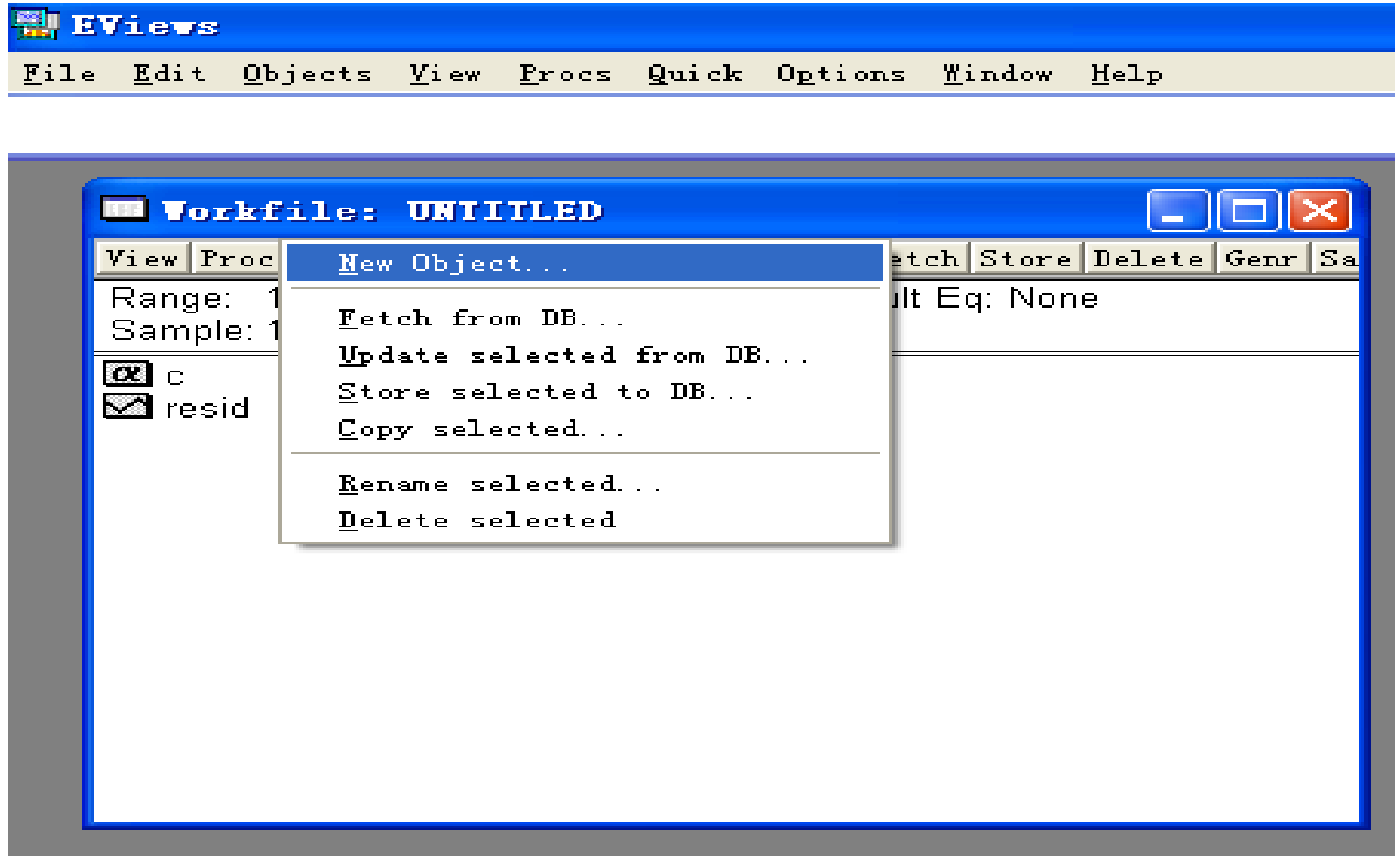
<input checked="" type="checkbox"/> <u>A</u> nnual	<input type="checkbox"/> <u>W</u> eekly
<input type="checkbox"/> <u>S</u> emi-annual	<input type="checkbox"/> <u>D</u> aily [5 day weeks]
<input type="checkbox"/> <u>Q</u> uarterly	<input type="checkbox"/> <u>D</u> aily [7 day weeks]
<input type="checkbox"/> <u>M</u> onthly	<input type="checkbox"/> <u>U</u> ndated or irregular

<u>S</u> tart date	<u>E</u> nd date
<input type="text" value="1997"/>	<input type="text" value="2003"/>

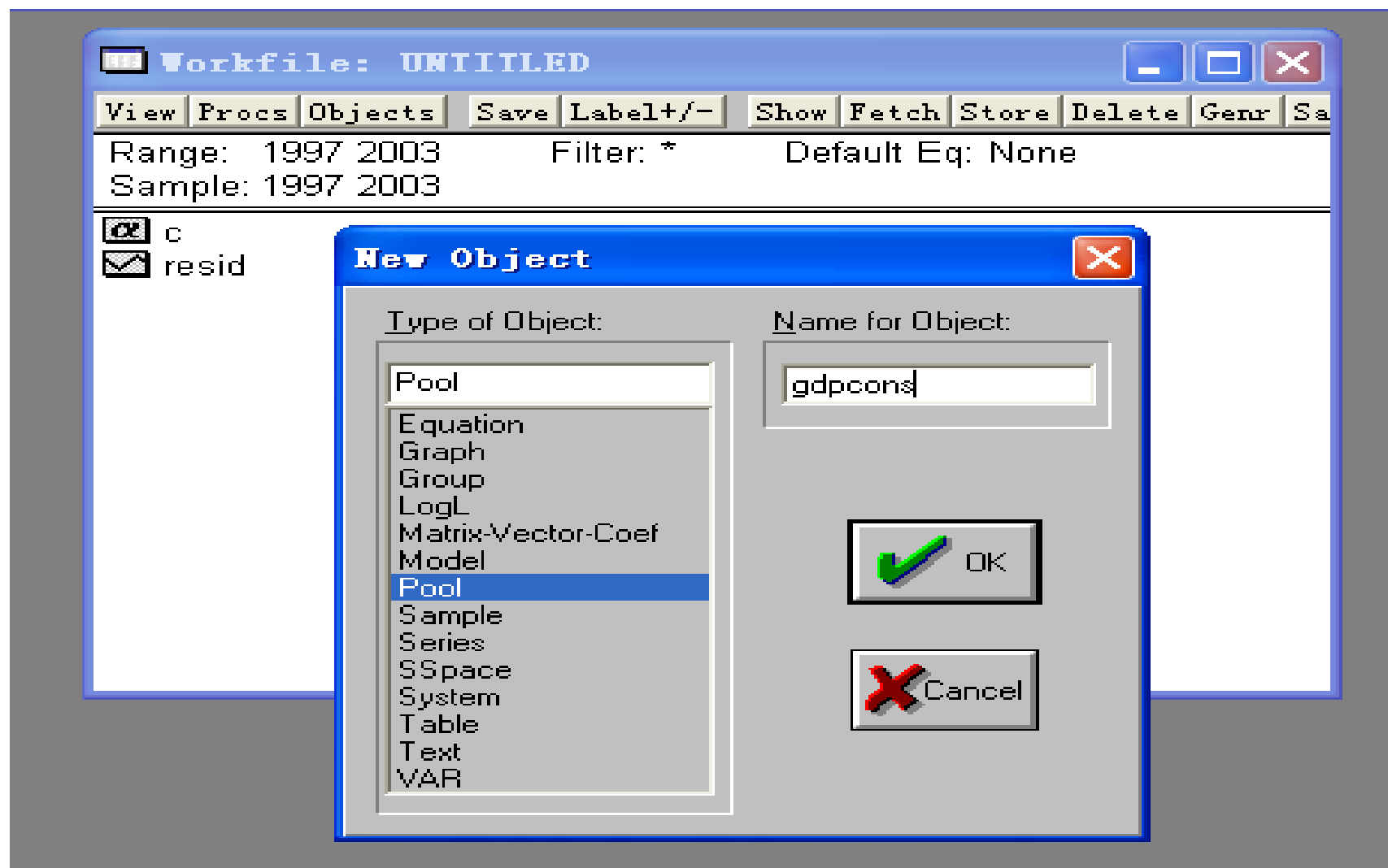
 OK

 Cancel

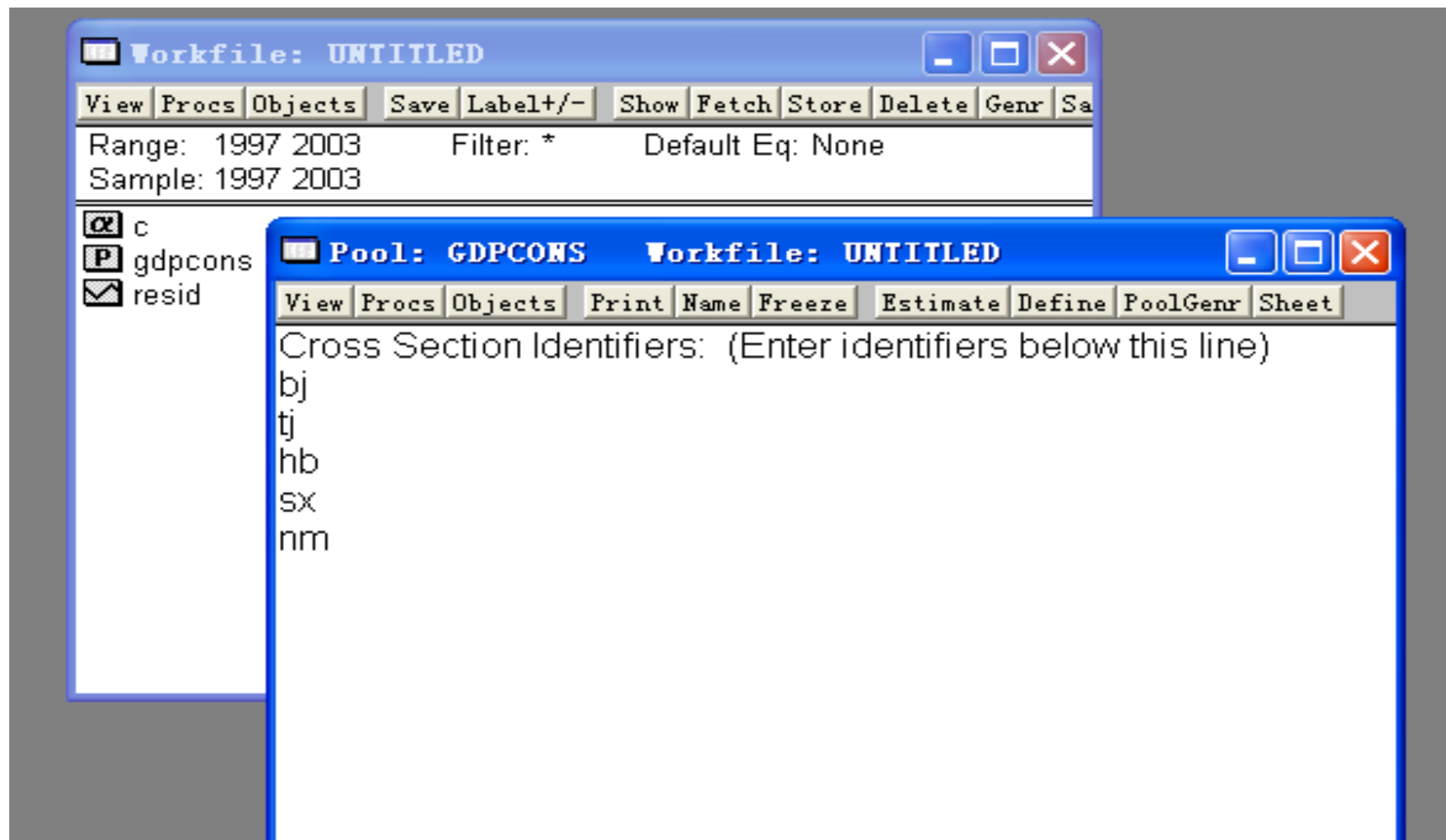
选择建立新的数据文件



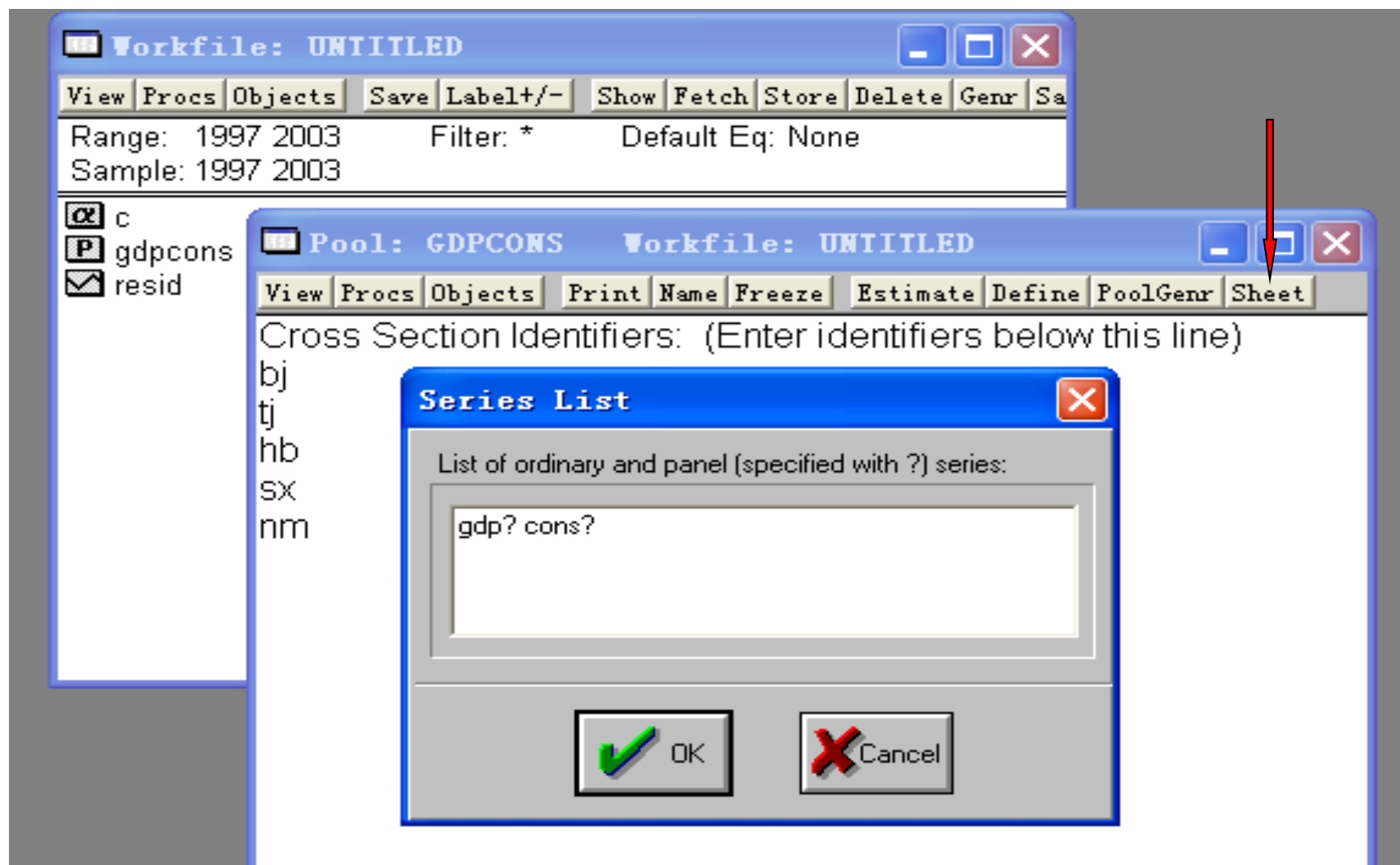
数据类型选择（Pool）和文件命名（gdpcons）



输入截面个体名称（BJ、TJ、HB、SX、NM）



选择**Sheet**，输入变量名（**GDP?**、**CONS?**）



出现数据表

Workfile: UNTITLED

View Procs Objects Save Label+/- Show Fetch Store Delete Genr Sa

Range: 1997 2003 Filter: * Default Eq: None

Sample: 1997 2003

☒ c
☒ consbj
☒ conshb
☒ consnm
☒ conssx
☒ constj
☒ gdpbj
☒ gdpcons
☒ gdphb
☒ gdpnm
☒ gdpsx
☒ gdptj
☒ resid

Pool: GDPCONS Workfile: UNTITLED

View Procs Objects Print Name Freeze Edit+- Order+- Smpl+- Title Estima

obs	GDP?	CONS?			
BJ-1997	NA	NA			
BJ-1998	NA	NA			
BJ-1999	NA	NA			
BJ-2000	NA	NA			
BJ-2001	NA	NA			
BJ-2002	NA	NA			
BJ-2003	NA	NA			
TJ-1997	NA	NA			
TJ-1998	NA	NA			
TJ-1999	NA	NA			
TJ-2000	NA	NA			
TJ-2001	NA	NA			
TJ-2002	NA	NA			
TJ-2003	NA	NA			
HB-1997	NA	NA			
HB-1998	NA	NA			

输入数据、显示数据

obs	GDP?	CONS?	
BJ-1997	1871.000	703.0000	
TJ-1997	2046.000	810.0000	
HB-1997	2174.000	954.0000	
SX-1997	2479.000	1221.000	
NM-1997	2846.000	1468.000	
BJ-1998	3213.000	1700.000	
TJ-1998	3663.000	1968.000	
HB-1998	1240.000	554.0000	
SX-1998	1336.000	623.0000	
NM-1998	1450.000	717.0000	
BJ-1999	1639.000	805.0000	
TJ-1999	1840.000	902.0000	
HB-1999	2051.000	990.0000	
SX-1999	2448.000	1135.000	
NM-1999	3954.000	1740.000	
BJ-2000	4256.000	1819.000	
TJ-2000	4569.000	1984.000	
HB-2000	5089.000	2241.000	
SX-2000	5578.000	2509.000	
NM-2000	6123.000	2820.000	
BJ-2001	7099.000	3260.000	
TJ-2001	1473.000	852.0000	
HB-2001	1596.000	791.0000	
SX-2001	1501.000	857.0000	
NM-2001	1638.000	946.0000	
BJ-2002	1788.000	1046.000	
TJ-2002	2042.000	1184.000	
HB-2002	2516.000	1374.000	
SX-2002	1083.000	640.0000	
NM-2002	1169.000	664.0000	
BJ-2003	1255.000	725.0000	
TJ-2003	1392.000	787.0000	
HB-2003	1545.000	936.0000	

如何在Eviews中估计Panel Data 模型?

在数据表界面上, 选择**Estimate**命令, 进行估计

The screenshot shows the EViews software interface. The top menu bar includes 'File', 'Edit', 'Objects', 'View', 'Procs', 'Quick', 'Options', 'Window', and 'Help'. The 'View' menu is open, showing options like 'Print', 'Name', 'Freeze', 'Edit+-', 'Order+-', 'Smpl+-', 'Title', 'Estimate', 'Define', 'PanelGenr', and 'Sample'. A red arrow points from the 'Estimate' button to the 'Pooled Estimation' dialog box.

The 'Pooled Estimation' dialog box has the following settings:

- Dependent Variable: Sample
- Regressors and AR() terms: Common coefficients
- Intercept: ☒ Common
- Weighting: ☒ No weighting
- Options: ☐ Iterate to convergence

The background data table is as follows:

obs	GDP?	CONS?
BJ-1997	1871.000	703.0000
TJ-1997	2046.000	810.0000
HB-1997	2174.000	954.0000
SX-1997	2479.000	1221.000
NM-1997	2846.000	1468.000
BJ-1998	3213.000	1700.000
TJ-1998	3663.000	1968.000
HB-1998	1240.000	554.0000
SX-1998	1336.000	623.0000
NM-1998	1450.000	717.0000
BJ-1999	1639.000	805.0000
TJ-1999	1840.000	902.0000
HB-1999	2051.000	990.0000
SX-1999	2448.000	1135.000
NM-1999	3954.000	1740.000
BJ-2000	4256.000	1819.000
TJ-2000	4569.000	1984.000
HB-2000	5089.000	2241.000
SX-2000	5578.000	2509.000
NM-2000	6123.000	2820.000
BJ-2001	7099.000	3260.000
TJ-2001	1473.000	852.0000
HB-2001	1596.000	791.0000

估计

obs	COM?	GDP?			
BJ-1997	702.0000	1074.000			
BJ-1998	810.0000	1100.000			
BJ-1999	954.0000	1150.000			
BJ-2000	1220.000	1200.000			
BJ-2001	1480.000	1250.000			
BJ-2002	1700.000	1300.000			
BJ-2003	1900.000	1350.000			
TJ-1997	554.0000	750.000			
TJ-1998	620.0000	780.000			
TJ-1999	710.0000	810.000			
TJ-2000	800.0000	840.000			
TJ-2001	900.0000	870.000			
TJ-2002	990.0000	900.000			
TJ-2003	1100.000	930.000			
HB-1997	174.0000	250.000			
HB-1998	180.0000	260.000			
HB-1999	190.0000	270.000			
HB-2000	224.0000	280.000			
HB-2001	250.0000	290.000			
HB-2002	280.0000	300.000			
HB-2003	320.0000	310.000			
SX-1997	852.0000	1473.000			
SX-1998	704.0000	1500.000			

Pooled Estimation

Dependent Variable: Sample: ☐ Balanced Sample

Regressors and AR() terms:

Common coefficients:

Cross section specific coefficients:

Intercept: ☐ None ☐ Common ☒ Fixed effects ☐ Random effects

Weighting: ☒ No weighting ☐ Cross section weights ☐ SUR ☐ Iterate to convergence

Random effects uses Variance Components model.
Weights, AR() and CS specific coefs not allowed.

输出

EViews - [Pool: CONGDP Workfile: CONGDP]				
File Edit Objects View Procs Quick Options Window Help				
View	Procs	Objects	Print	Name
Freeze	Estimate	Define	PoolGenr	Sheet
Dependent Variable: COM?				
<u>Method: Pooled Least Squares</u>				
Date: 11/05/05 Time: 12:28				
Sample: 1997 2003				
Included observations: 7				
Number of cross-sections used: 5				
Total panel (balanced) observations: 35				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
GDP?	0.550205	0.019529	28.17358	0.0000
Fixed Effects				
BJ--C	-177.1921			
TJ--C	-125.5225			
HB--C	-543.1295			
SX--C	20.39002			
NM--C	50.28222			
R-squared	0.990569	Mean dependent var	1258.143	
Adjusted R-squared	0.988943	S.D. dependent var	666.1618	
S.E. of regression	70.04712	Sum squared resid	142291.4	
Log likelihood	-195.0928	Durbin-Watson stat	0.767855	

结果

Substituted Coefficients:

=====

$$\text{COMBJ} = -177.19207 + 0.5502047064 * \text{GDPBJ}$$

$$\text{COMTJ} = -125.5224709 + 0.5502047064 * \text{GDPTJ}$$

$$\text{COMHB} = -543.1294537 + 0.5502047064 * \text{GDPHB}$$

$$\text{COMSX} = 20.39001648 + 0.5502047064 * \text{GDPSX}$$

$$\text{COMNM} = 50.28222237 + 0.5502047064 * \text{GDPNM}$$

三、固定影响变系数模型

- 将 β_i 视为固定的不同的常数时，可写成：

$$y = X\beta + u$$

将截距项也看作一个虚变量

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{nT \times 1} \quad X = \begin{pmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & X_n \end{pmatrix}_{nT \times nK} \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}_{nK \times 1} \quad u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}_{nT \times 1}$$

- 显然，如果随机干扰项在不同横截面个体之间不相关，上述模型的参数估计极为简单，即以每个截面个体的时间序列数据为样本，采用经典单方程模型的估计方法分别估计其参数。即使采用 **GLS** 估计同时得到的 **GLS** 估计量，也是与在每个横截面个体上的经典单方程估计一样。
- 条件：

$$Eu_i u_j' = 0 \quad i \neq j$$

$$Eu_i u_i' = \sigma_i^2 I$$

- 如果随机项在不同横截面个体之间的协方差不为零，**GLS**估计比每个横截面个体上的经典单方程估计更有效。
- 为什么？

$$\Omega_{ij} = Eu_i u_j'$$

$$V = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \cdots & \Omega_{1n} \\ \Omega_{21} & \Omega_{22} & \cdots & \Omega_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \Omega_{n1} & \Omega_{n2} & \cdots & \Omega_{nn} \end{bmatrix}_{nT \times nT}$$

问题的关键是然后求得V矩阵。
各种文献中提出各种V矩阵的方法，形成了各种**FGLS**估计

$$\hat{\beta}_{GLS} = (X'V^{-1}X)^{-1} X'V^{-1}y$$