第四章
经典单方程计量经济学模型：放宽基本假定的模型
Relaxing the Assumptions of the Classical Model

§4.1 异方差性*

§4.2 序列相关性***

§4.3 多重共线性*

§4.4 随机解释变量问题**

本章说明

- 基本假定违背主要 包括：
  - 随机误差项序列存在异方差性；
  - 随机误差项序列存在序列相关性；
  - 解释变量之间存在多重共线性；
  - 解释变量是随机变量且与随机误差项相关的随机解释变量问题；
    - ✓模型设定有偏误；
    - ✓解释变量的方差不随样本容量的增而收敛。

- 计量经济检验：对模型基本假定的检验
- 本章主要讨论前4类

为什么不讨论正态性假设?

William H. Greene(2003), Econometric Analysis

- *In most cases, the zero mean assumption is **not restrictive**.*

- *In view of our description of the source of the disturbances, the conditions of the **central limit theorem** will generally apply, at least approximately, and the normality assumption will be reasonable in most settings.*

  Except in those cases in which some alternative distribution is assumed, the normality assumption is probably quite reasonable.

实际中：正态性假设的违背情况

- 当存在模型关系误差时，如果解释变量是随机的，随机误差项的正态性将得不到保证。
- 当模型遗漏了显著的变量，如果遗漏的变量非正态，随机误差项将不具有正态性。
- 如果待估计的模型是原模型经过函数变换得到的，随机误差项将不再服从正态分布。
- 当模型存在被解释变量的观测误差，如果观测误差相对于随机误差项的标准差特别大、样本长度又特别小，随机误差项的正态性假设会导致显著性水平产生一定程度的扭曲。
- 当模型存在解释变量观测误差时，一般情况下，随机误差项的正态性假设都是不能成立的；只有在回归函数是线性的、且观测误差分布是正态的特殊情形下，随机误差项的正态性才成立。

§4.1　异方差性
Heteroscedasticity

一、异方差的概念*

二、异方差性的后果

三、异方差性的检验

四、异方差的修正**

五、例题

一、异方差的概念

1*、异方差

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \mu_i \qquad i = 1, 2, \cdots, n$$

$$Var(\mu_i) = \sigma^2$$

Homoscedasticity

$$Var(\mu_i) = \sigma_i^2$$

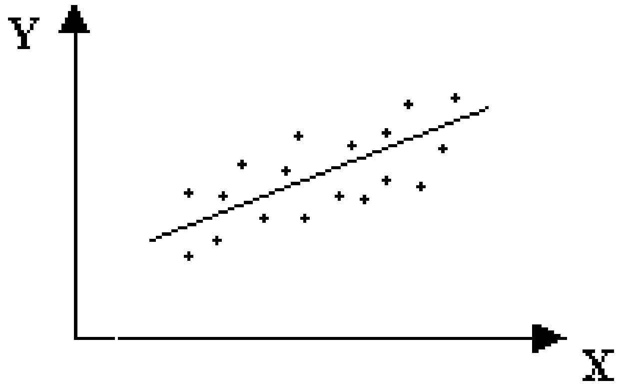
即对不同的样本点 ***i***，随机误差项的方差不再是常数，而互不相同，则认为出现了异方差性(**Heteroskedasticity**)。

2、异方差的类型

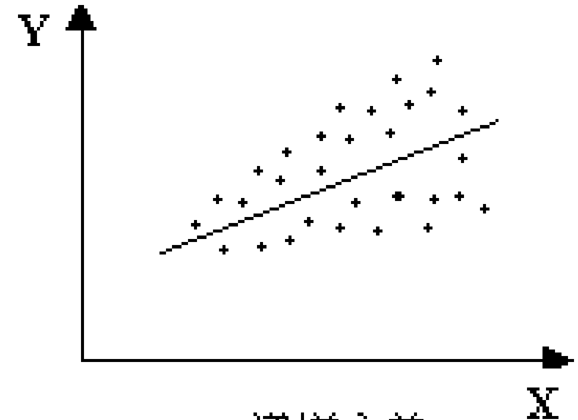
- 同方差：σ_i^2 = 常数，与解释变量观测值X_i无关；

  异方差：$\sigma_i^2 = f(X_i)$，与解释变量观测值X_i有关。

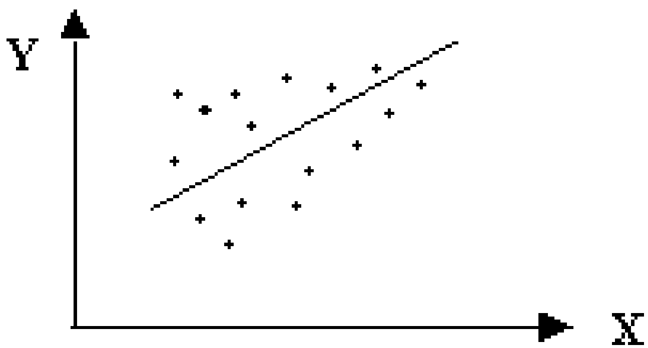
- 异方差一般可归结为三种类型：
  - 单调递增型：σ_i^2随X的增大而增大
  - 单调递减型：σ_i^2随X的增大而减小
  - 复　杂　型：σ_i^2与X的变化呈复杂形式



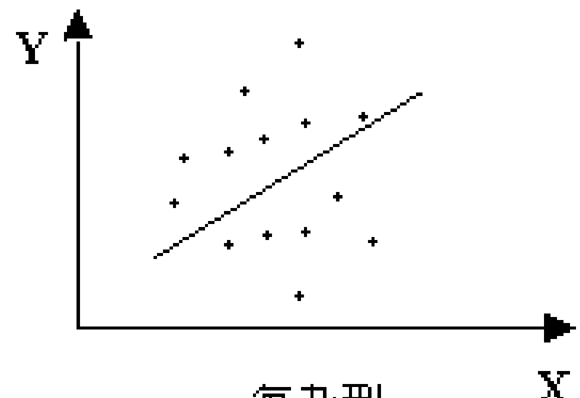
同方差



递增方差



递减方差



复杂型

3、实际经济问题中的异方差性

例4.1.1：截面资料下研究居民家庭的储蓄行为

$$Y_i=\beta_0+\beta_1 X_i+\mu_i$$

Y_i:第i个家庭的储蓄额 X_i:第i个家庭的可支配收入。

高收入家庭：储蓄的差异较大；

低收入家庭：储蓄则更有规律性，差异较小。

μ_i的方差呈现<u>单调递增</u>型变化

例4.1.2: 以绝对收入假设为理论假设、以截面数据为样本建立居民消费函数:

$$C_i=\beta_0+\beta_1 Y_i+\mu_i$$

将居民按照收入等距离分成n组，取组平均数为样本观测值。

- **一般情况下，居民收入服从正态分布**：中等收入组人数多，两端收入组人数少；人数多的组平均数的误差小，人数少的组平均数的误差大。

样本观测值的观测误差 随着解释变量观测值的不同而不同，往往引起随机项的异方差性，且**呈U形**。

例4.1.3：以某一行业的企业为样本建立企业生产函数模型

$$Y_i=A_i^{\beta 1} K_i^{\beta 2} L_i^{\beta 3} e^{\mu i}$$

被解释变量：产出量Y，

解释变量：资本K、劳动L、技术A。

- 每个企业所处的**外部环境**对产出量的影响被包含在随机误差项中。

对不同的企业，它们对产出量的影响程度不同，造成了随机误差项的异方差性。

随机误差项的方差并不随某一个解释变量观测值的变化而呈**规律性变化**，呈现**复杂型**。

二、异方差性的后果

Consequences of Using OLS in the Presence of Heteroskedasticity

1、参数估计量非有效*

- OLS估计量仍然具有无偏性，但不具有有效性。

- 因为在 有效性证明中 利用了: $E(\mu\mu')=\sigma^2 I$

- 而且，在大样本情况下，尽管参数估计量具有一致性，但仍然不具有渐近有效性。

2、变量的显著性检验失去意义

- 变量的显著性检验中，构造了t统计量

$$t=\hat{\beta}_i/S_{\hat{\beta}_i}$$

它是建立在σ^2不变而正确估计了参数方差$S_{\hat{\beta}_i}$的基础之上的。

如果出现了异方差性，估计的$S_{\hat{\beta}_i}$出现偏误（偏大或偏小），t检验失去意义。

- **其他检验，也需要基于σ^2的估计。**

3、模型的预测失效

一方面，由于上述后果，使得模型不具有良好的统计性质；

另一方面，在预测值的置信区间中也包含有参数方差的估计量 $S_{\hat{\beta}_i}$。

所以，当模型出现异方差性时，参数OLS估计值的变异程度增大，从而**造成对Y的预测误差变大，降低预测精度，预测功能失效**。

三、异方差性的检验

Detection of Heteroscedasticity

1、检验思路*

- 检验方法很多
- **Graphical Method**
- **Formal Metrods**
  - Park Test
  - Glejser Test
  - Spearman's Rank Correlation Test
  - Goldfeld-Quandt Test
  - Breusch-Pagan-Godfrey Test
  - White's General Heteroscedasticity Test
  - Koenker-Bassett Test

共同的思路：

- 由于**异方差性**是相对于不同的解释变量观测值，随机误差项具有不同的方差；那么**检验异方差性，也就是检验“随机误差项的方差”与“解释变量观测值”之间的相关性、及其相关的“形式”。**

- **问题在于用什么来表示随机误差项的方差？**

一般的处理方法：首先采用OLS估计，得到残差的近似估计值，用它的平方近似随机误差项的方差。

$$\boxed{\tilde{e}_i = y_i - (\hat{y}_i)_{0ls}} \qquad \boxed{Var(\mu_i) = E(\mu_i^2) \approx \tilde{e}_i^2}$$

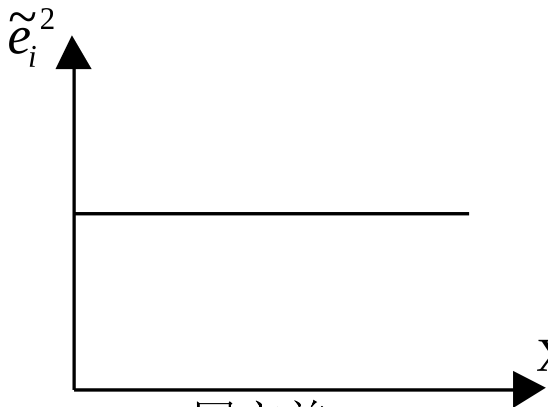
2、图示法

（1）用X-Y的散点图进行判断

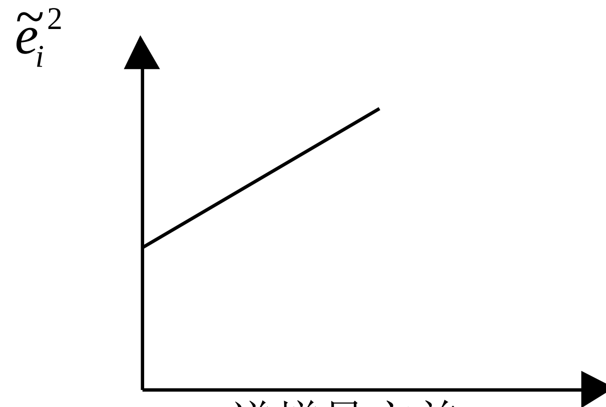
看是否存在明显的**散点扩大**、**缩小**或**复杂型趋势**（即不在一个固定的带型域中）。

(2) X-$\tilde{e}_i^2$的散点图进行判断

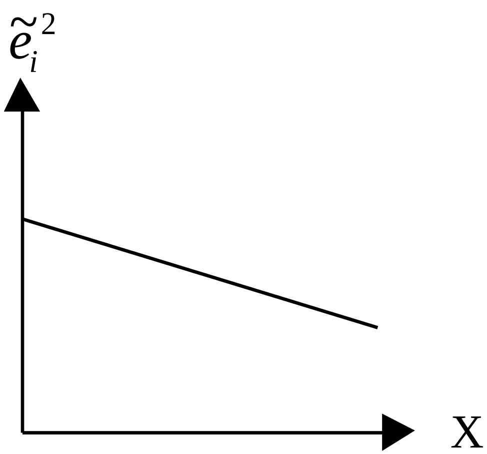
看是否形成一斜率为零的直线。



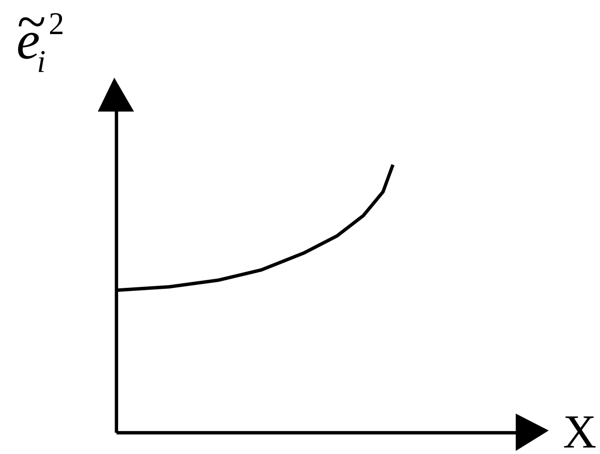
同方差



递增异方差



递减异方差



复杂型异方差

3、帕克(Park)检验与戈里瑟(Gleiser)检验

- **基本思想:** 偿试建立方程:

Park $\tilde{e}_i^2 = f(X_{ji}) + \varepsilon_i$

$|\tilde{e}_i| = f(X_{ji}) + \varepsilon_i$ Gleiser

选择关于变量X的不同的函数形式，对方程进行估计并进行显著性检验，如果存在某一种函数形式，使得方程显著成立，则说明原模型存在异方差性。

- 帕克检验常用的函数形式:

$$f(X_{ji}) = \sigma^2 X_{ji}^{\alpha} e^{\varepsilon_i}$$

$$\ln(\tilde{e}_i^2) = \ln\sigma^2 + \alpha \ln X_{ji} + \varepsilon_i$$

若α在统计上是显著的，表明存在异方差性。

4、戈德菲尔德-匡特(Goldfeld-Quandt)检验

- **G-Q**检验以**F**检验为基础，适用于样本容量较大、异方差递增或递减的情况。
- 先将样本一分为二，对子样①和子样②分别作回归，然后利用两个子样的残差平方和之比构造统计量进行异方差检验。
- 由于该统计量服从**F**分布，因此假如存在递增的异方差，则**F**远大于**1**；反之就会等于**1**（同方差）或小于**1**（递减方差）。

- G-Q检验的步骤：
  - 将n对样本观察值(X_i,Y_i)按观察值X_i的大小排队;
  - 将序列中间的c=n/4个观察值除去，并将剩下的观察值划分为较小与较大的相同的两个子样本，每个子样样本容量均为(n-c)/2;
  - 对每个子样分别进行OLS回归，并计算各自的残差平方和。
  - 在同方差性假定下，构造如下满足F分布的统计量：

$$\boxed{F=\frac{\sum \tilde{e}_{2i}^2 \Big/ \left(\frac{n-c}{2}-k-1\right)}{\sum \tilde{e}_{1i}^2 \Big/ \left(\frac{n-c}{2}-k-1\right)} \sim F\left(\frac{n-c}{2}-k-1,\frac{n-c}{2}-k-1\right)}$$

5、怀特（White）检验

$$Y_i=\beta_0+\beta_1X_{1i}+\beta_2X_{2i}+\mu_i$$

以二元模型为例

建立**辅助回归**模型

先对该模型作OLS回归，得到 $\tilde{e}_i^2$

$$\tilde{e}_i^2=\alpha_0+\alpha_1X_{1i}+\alpha_2X_{2i}+\alpha_3X_{1i}^2+\alpha_4X_{2i}^2+\alpha_5X_{1i}X_{2i}+\varepsilon_i$$

$$nR^2 \overset{\bullet}{\sim} \chi^2(h)$$

在同方差假设下

辅助回归可决系数

渐近服从

辅助回归解释变量的个数

- **说明：**
- 辅助回归，仍是检验残差与解释变量可能组合的显著性；因此，辅助回归方程中还可引入解释变量的更高次方。
- 如果存在异方差性，则表明确与解释变量的某种组合有显著的相关性，这时往往显示出有较高的可决系数以及某一参数的t检验值较大。
- 在多元回归中，由于辅助回归方程中可能有太多解释变量，从而使自由度减少，有时可去掉交叉项。

四、异方差的修正
—加权最小二乘法

Correcting Heteroscedasticity
—Weighted Least Squares, WLS

1、WLS的思路

- 加权最小二乘法是对原模型加权，使之变成一个新的“不存在异方差性的模型”，然后采用OLS估计其参数。

$$\boxed{\sum W_i e_i^2 = \sum W_i [Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_1 + \cdots + \hat{\beta}_k X_k)]^2}$$

在采用OLS方法时:

对较小的残差平方e_i^2赋予较大的权数;

对较大的残差平方e_i^2赋予较小的权数。

- 例如，对一多元模型

$$\boxed{Var(\mu_i) = E(\mu_i)^2 = \sigma_i^2 = f(X_{ji})\sigma^2}$$

$$\frac{1}{\sqrt{f(X_{ji})}}Y_i = \beta_0\frac{1}{\sqrt{f(X_{ji})}} + \beta_1\frac{1}{\sqrt{f(X_{ji})}}X_{1i} + \beta_2\frac{1}{\sqrt{f(X_{ji})}}X_{2i} + \cdots$$

$$+\beta_k\frac{1}{\sqrt{f(X_{ji})}}X_{ki} + \frac{1}{\sqrt{f(X_{ji})}}\mu_i$$

$$\boxed{Var\left(\frac{1}{\sqrt{f(X_{ji})}}\mu_i\right) = E\left(\frac{1}{\sqrt{f(X_{ji})}}\mu_i\right)^2 = \frac{1}{f(X_{ji})}E(\mu_i)^2 = \sigma^2}$$

加权后的模型、满足同方差性，可用OLS法估计。

- 一般情况下：

$$\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\boldsymbol{\mu}$$

$$E(\mu)=0$$

$$Cov(\mu)=E(\mu\,\mu')=\sigma^2\mathbf{W}$$

$$W=\begin{bmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_n \end{bmatrix}$$

$$\mathbf{D}^{-1}\mathbf{Y}=\mathbf{D}^{-1}\mathbf{X}\beta+\mathbf{D}^{-1}\mu$$

W是一“对称正定矩阵”，存在一可逆矩阵**D**使得

W=DD'

$$\mathbf{Y}_* = \mathbf{X}_*\beta + \mu_*$$

$$E(u\,u^{*\prime})=E(D^{-1}u\cdot u'D^{-1\prime})=D^{-1}E(u\cdot u')D^{-1\prime}$$

$$=D^{-1}\Omega D^{-1\prime}=D^{-1}(\sigma^2 W)D^{-1\prime}=\sigma^2 D^{-1}DD'D^{-1\prime}=\sigma^2 I$$

用OLS法估计新模型，记参数估计量为 $\hat{\beta}_*$ ，则

$$\hat{\beta}_* = (\mathbf{X}_*'\mathbf{X}_*)^{-1}\mathbf{X}_*'\mathbf{Y}_*$$

$$= (\mathbf{X}'\mathbf{D}^{-1'}\mathbf{D}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{D}^{-1'}\mathbf{D}^{-1}\mathbf{Y}$$

$$= (\mathbf{X}'\mathbf{W}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^{-1}\mathbf{Y}$$

这就是原模型$Y=X\beta+\mu$的**加权最小二乘估计量**，是无偏、有效的估计量。

这里**权矩阵为D^{-1}**，它来自于原模型“残差项μ的方差-协方差矩阵$\sigma^2 \mathbf{W}$”。

2、如何得到$\sigma^2\mathbf{W}$ ？

- 一种可行的方法：

对原模型进行OLS估计，得到随机**误差项的近似估计量$\tilde{e}_i$**，以此构成权矩阵的估计量。即

$$\sigma^2\hat{\mathbf{W}} = \begin{bmatrix} \tilde{e}_1^2 & & \\ & \ddots & \\ & & \tilde{e}_n^2 \end{bmatrix}$$

$$\boxed{\mathbf{D}^{-1} = diag\{1/|\tilde{e}_1|, 1/|\tilde{e}_2|, \cdots, 1/|\tilde{e}_n|\}}$$

3、异方差稳健标准误法（Heteroscedasticity-Consistent Variances and Standard Errors）

- 应用软件中推荐的一种选择，适合样本容量足够大的情况。
- 仍然采用OLS，但 对OLS估计量的标准差 进行修正。
- 与不附加选择的OLS估计比较，参数估计量没有变化，但是参数估计量的方差和标准差变化明显。

- 即使存在异方差、仍然采用OLS估计时，变量的显著性检验有效，预测有效。

例3.2.2 地区城镇居民消费模型—OLS估计

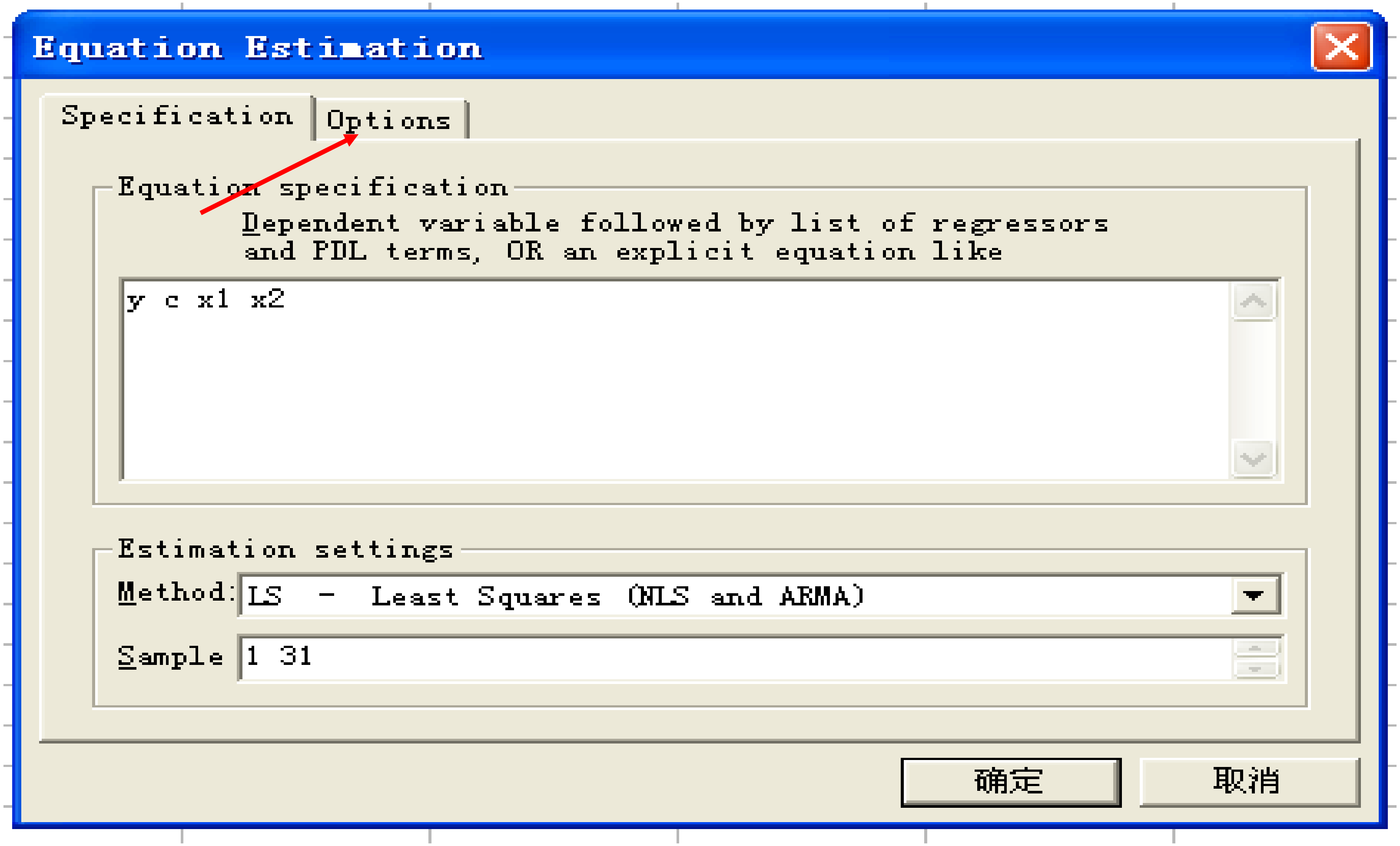
Dependent Variable: Y
Method: Least Squares
Date: 09/10/10 Time: 11:32
Sample: 1 31
Included observations: 31

失效

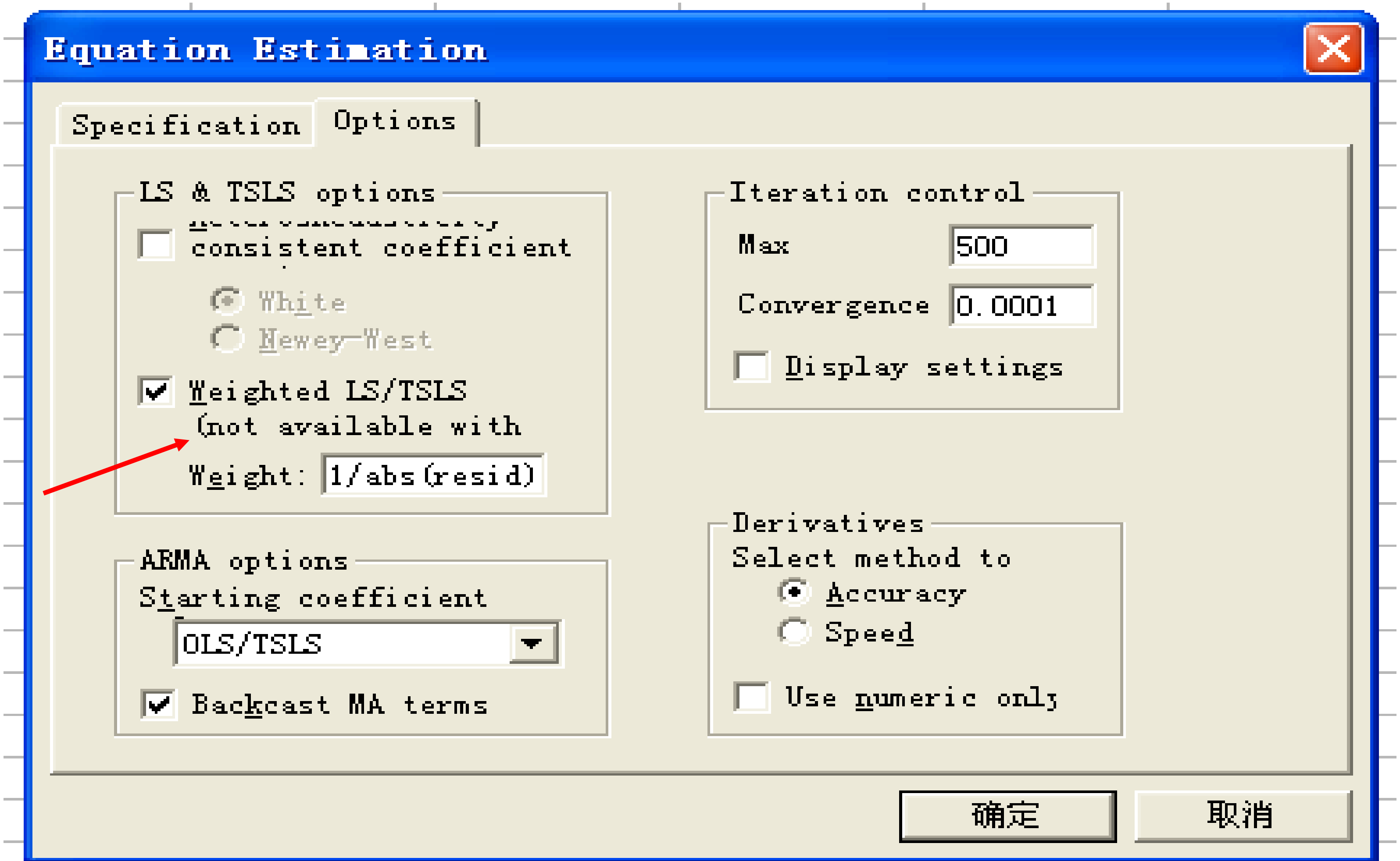
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	143.3265	260.4032	0.550402	0.5864
X1	0.555644	0.075308	7.378320	0.0000
X2	0.250085	0.113634	2.200791	0.0362

R-squared	0.975634	Mean dependent var	8401.468
Adjusted R-squared	0.973893	S.D. dependent var	2388.459
S.E. of regression	385.9169	Akaike info criterion	14.84089
Sum squared resid	4170093.	Schwarz criterion	14.97966
Log likelihood	-227.0337	F-statistic	560.5650
Durbin-Watson stat	1.843488	Prob(F-statistic)	0.000000

地区城镇居民消费模型—WLS



地区城镇居民消费模型—WLS—Weighted



地区城镇居民消费模型—WLS—Weighted

Dependent Variable: Y
Method: Least Squares
Date: 09/10/10 Time: 11:41
Sample: 1 31
Included observations: 31
Weighting series: 1/ABS(RESID)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	206.1846	49.65502	4.152342	0.0003
X1	0.382828	0.067513	5.670425	0.0000
X2	0.501918	0.101364	4.951659	0.0000

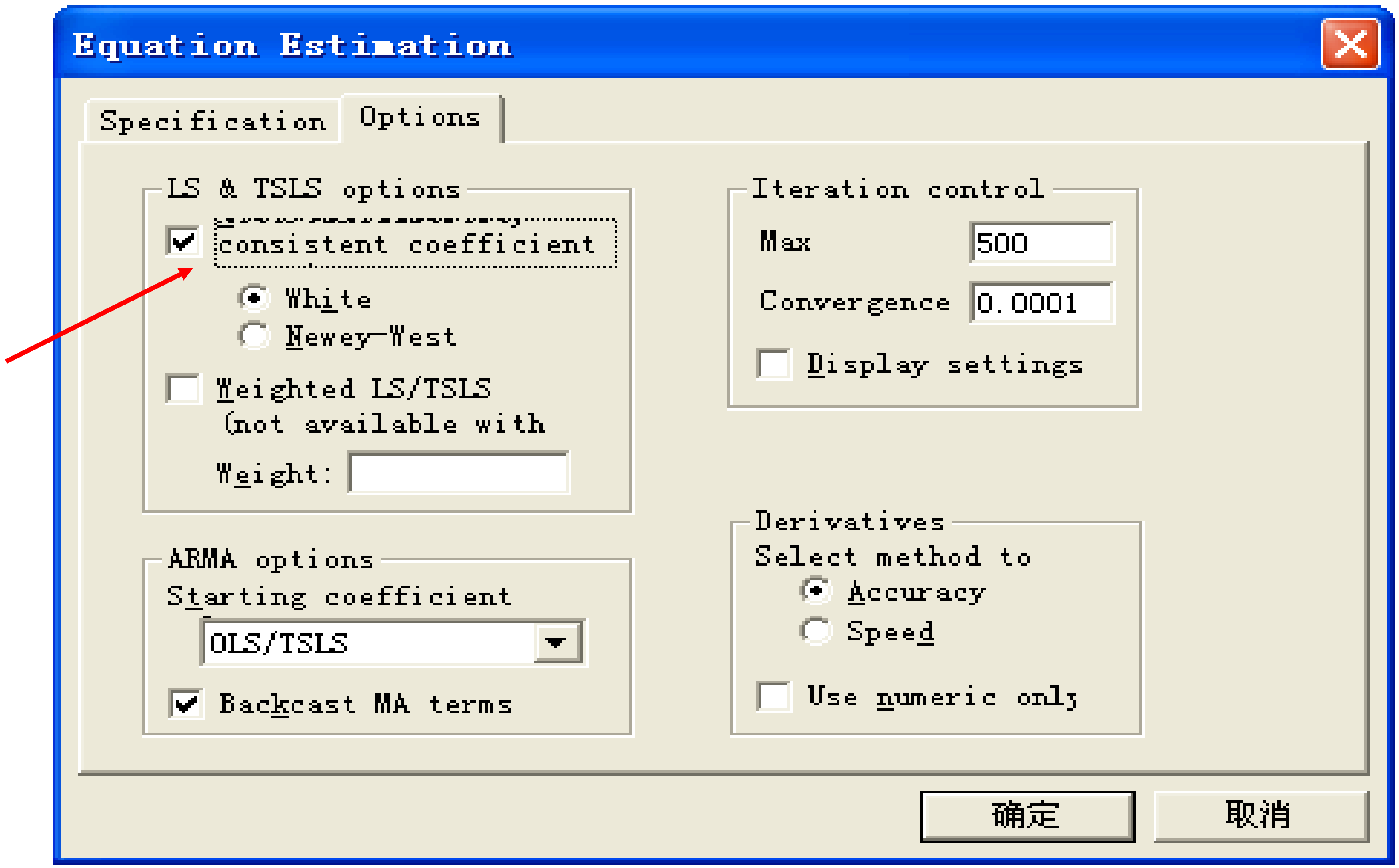
Weighted Statistics

R-squared	0.999991	Mean dependent var	10199.84
Adjusted R-squared	0.999990	S.D. dependent var	29421.06
S.E. of regression	92.37167	Akaike info criterion	11.98128
Sum squared resid	238910.7	Schwarz criterion	12.12006
Log likelihood	-182.7099	F-statistic	33458.21
Durbin-Watson stat	1.769842	Prob(F-statistic)	0.000000

Unweighted Statistics

R-squared	0.970469	Mean dependent var	8401.468
Adjusted R-squared	0.968359	S.D. dependent var	2388.459
S.E. of regression	424.8562	Sum squared resid	5054078.
Durbin-Watson stat	2.120449		

地区城镇居民消费模型—WLS—HCCC



地区城镇居民消费模型—WLS—HCCC

Dependent Variable: Y
Method: Least Squares
Date: 09/10/10 Time: 11:43
Sample: 1 31
Included observations: 31
White Heteroskedasticity-Consistent Standard Errors & Covariance

有效

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	143.3265	202.7027	0.707078	0.4854
X1	0.555644	0.167324	3.320757	0.0025
X2	0.250085	0.250257	0.999314	0.3262

R-squared	0.975634	Mean dependent var	8401.468
Adjusted R-squared	0.973893	S.D. dependent var	2388.459
S.E. of regression	385.9169	Akaike info criterion	14.84089
Sum squared resid	4170093.	Schwarz criterion	14.97966
Log likelihood	-227.0337	F-statistic	560.5650
Durbin-Watson stat	1.843488	Prob(F-statistic)	0.000000

5、在实际操作中通常采用的经验方法

- **采用截面数据作样本时，不对原模型进行异方差性检验，而是直接选择加权最小二乘法。**
  - 如果确实存在异方差，则被有效地消除了；
  - 如果不存在异方差性，则加权最小二乘法等价于普通最小二乘法。

- **采用时序数据作样本时，时常不考虑异方差性检验（ARCH）。**

五、例题--中国农村居民人均消费函数（自学）

例4.1.4 中国农村居民人均消费支出主要由人均纯收入来决定。

农村人均纯收入包括(1)从事农业经营的收入，(2)包括从事其他产业的经营性收入(3)工资性收入、(4)财产收入(4)转移支付收入。

考察**从事农业经营的收入(X_1)**和**其他收入(X_2)**对中国**农村居民消费支出(Y)**增长的影响：

$$\ln Y = \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \mu$$

- 步骤
  - 对模型进行**OLS**估计；
  - 采用散点图检验，表明存在异方差；
  - 采用**G-Q**检验，表明存在异方差；
  - 经试算，寻找适当的权；
  - 采用**WLS**估计模型；
  - 采用稳健标准误方法估计模型。

§4.2　序列相关性***
Serial Correlation

一、序列相关性的概念*

二、序列相关性的后果

三、序列相关性的检验***

四、具有序列相关性模型的估计**

一、序列相关性的概念

1、序列相关性

- 模型随机项之间不存在相关性，称为：**No Autocorrelation**。

以截面数据为样本时，如果模型随机项之间存在相关性，称为：**Spatial Autocorrelation**。

以时序数据为样本时，如果模型随机项之间存在相关性，称为：**Serial Autocorrelation**。

习惯上，统称为序列相关性（**Serial Correlation or Autocorrelation**）。

- 在其他假设仍成立的条件下，随机扰动项序列相关即意味着：

$$\boxed{Cov(\mu_i, \mu_j) = E(\mu_i \mu_j) \neq 0}$$

$$Cov(\boldsymbol{\mu}) = E(\boldsymbol{\mu}\boldsymbol{\mu}') = \begin{pmatrix} \sigma^2 & \cdots & E(\mu_1 \mu_n) \\ \vdots & \ddots & \vdots \\ E(\mu_n \mu_1) & \cdots & \sigma^2 \end{pmatrix}$$

$$= \begin{pmatrix} \sigma^2 & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma^2 \end{pmatrix} = \sigma^2 \Omega \neq \sigma^2 \mathbf{I}$$

- 一阶序列相关，或自相关

$$E(\mu_i \mu_{i+1}) \neq 0 \qquad i = 1, 2, \cdots, n-1$$

$$\mu_i = \rho \mu_{i-1} + \varepsilon_i$$

ρ称为自协方差系数（coefficient of autocovariance）或一阶自相关系数（first-order coefficient of autocorrelation）

2、实际经济问题中的序列相关性

- 没有包含在解释变量中的、被解释经济变量“固有惯性”。
- 模型设定偏误（Specification error）。主要表现在模型中“丢掉了重要的解释变量”或模型“函数形式有误”。
- 数据的“编造”。

- 时序数据作为样本时，一般都存在序列相关性。
- 截面数据作为样本时，一般不考虑序列相关性。

二、序列相关性的后果

Consequences of Using OLS in the Presence of Autocorrelation

- 与异方差性引起的后果相同：
  - 参数估计量非有效
  - 变量的显著性检验失去意义
  - 模型的预测失效

三、序列相关性的检验
Detecting Autocorrelation

1、检验方法的思路

- 序列相关性检验方法，有多种：
  - Graphical Method
  - Regression Method
  - Durbin-Watson Test (D.W. test)
  - **Breusch-Godfrey (BG) Test, (LM test, Lagrange Multiplier)**

- 具有共同的思路。

- **基本思路：**

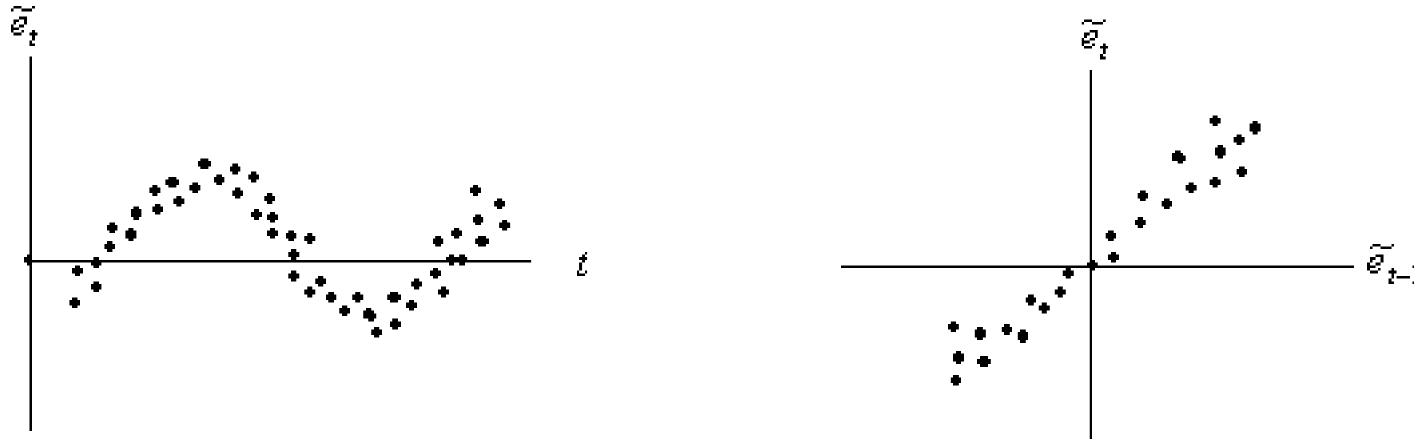
首先，采用 OLS 法估计模型，以求得随机误差项的**“近似估计量”**，用 $\tilde{e}_i$ 表示：

$$\tilde{e}_i = Y_i - (\hat{Y}_i)_{0ls}$$

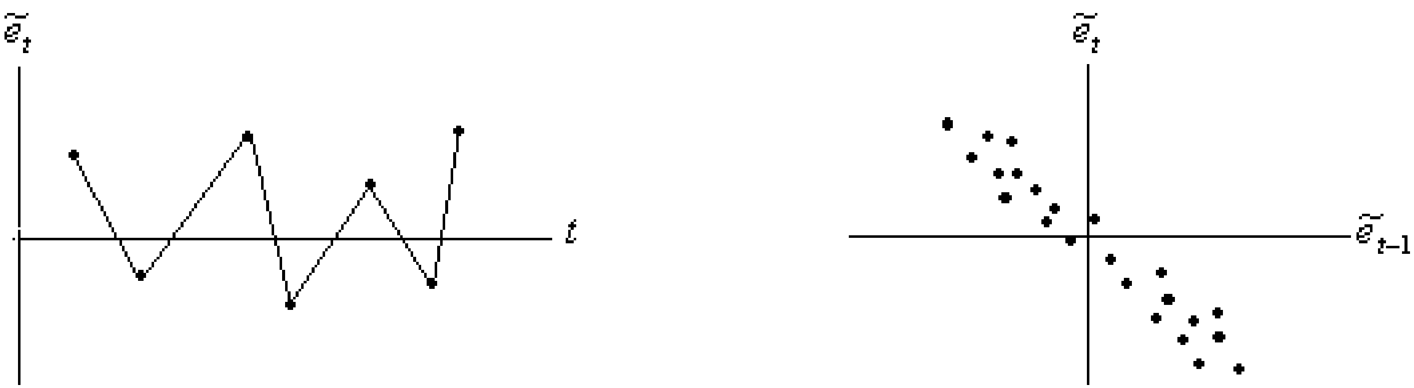
然后，通过分析这些**“近似估计量”**之间的相关性，以判断随机误差项是否具有序列相关性。

2、图示法（不可靠）

用$\tilde{e}_i$的变化图形来判断μ_i的序列相关性：



正序列相关（正自相关）



负序列相关（负自相关）

3、回归检验法

$$\tilde{e}_t = \rho \tilde{e}_{t-1} + \varepsilon_t$$

$$\tilde{e}_t = \rho_1 \tilde{e}_{t-1} + \rho_2 \tilde{e}_{t-2} + \varepsilon_t$$

……

- 如果存在某一种函数形式，使得方程显著成立，则说明原模型存在序列相关性。
- 回归检验法的优点是：
  - 能够确定序列相关的形式；
  - 适用于任何类型序列相关性问题的检验。

4、杜宾–瓦森（Durbin-Watson）检验法

- 杜宾（J.Durbin）和瓦森(G.S. Watson)于1951年提出的一种检验序列自相关的方法。

- 该方法的假定条件是：
  - 解释变量X非随机；
  - 随机误差项μ_i为一阶自回归形式：$\mu_i=\rho\mu_{i-1}+\varepsilon_I$；
  - 回归模型中“不含有滞后应变量”作为解释变量；
  - 回归含有截距项。

- 对原模型进行OLS估计，用残差近似值、构造统计量。

- **D.W. 统计量:**

H$_0$: ρ=0

$$D.W. = \frac{\sum_{t=2}^{n}(\tilde{e}_t - \tilde{e}_{t-1})^2}{\sum_{t=1}^{n}\tilde{e}_t^{\,2}}$$

该统计量的分布 **与出现在给定样本中的X值** 有复杂关系，其精确分布、很难得到。

但是，**他们**成功地导出了临界值的下限d_L和上限d_U，且这些上下限 只与 **"样本的容量*n*"** 和 **"解释变量的个数*k*"** 有关，而与解释变量X的取值无关。

- **D.W检验步骤:**
  - 计算DW值
  - 给定α，由n和k的大小查DW分布表，得临界值d_L和d_U
  - 比较、判断

$0<D.W.<d_L$	存在正自相关
$d_L<D.W.<d_U$	不能确定
$d_U<D.W.<4-d_U$	无自相关
$4-d_U<D.W.<4-d_L$	不能确定
$4-d_L<D.W.<4$	存在负自相关

- **证明：当D.W.值在2左右时，模型不存在一阶自相关。**

$$D.W.=\frac{\sum_{t=2}^{n}\tilde{e}_t^2+\sum_{t=2}^{n}\tilde{e}_{t-1}^2-2\sum_{t=2}^{n}\tilde{e}_t\tilde{e}_{t-1}}{\sum_{t=1}^{n}\tilde{e}_t^2}$$

条件？

$$D.W.\approx 2\left(1-\frac{\sum_{t=2}^{n}\tilde{e}_t\tilde{e}_{t-1}}{\sum_{t=1}^{n}\tilde{e}_t^2}\right)\approx 2(1-\rho)$$

完全一阶正相关，ρ =1，D.W.≈ 0；

完全一阶负相关，ρ = -1, D.W.≈ 4;

完全不相关，ρ=0，D.W.≈2

5、拉格朗日乘数检验***
（Lagrange multiplier, LM）

- 由布劳殊（Breusch）与戈弗雷（Godfrey）于1978年提出的，也被称为**GB检验**。
- 适合于 “**高阶序列相关”以及“模型中存在滞后被解释变量”** 的情形。
- 对原模型进行OLS估计，用“残差近似值的辅助回归模型的可决系数”构造统计量。

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \mu_i$$

$$\mu_t = \rho_1 \mu_{t-1} + \rho_2 \mu_{t-2} \cdots + \rho_p \mu_{t-p} + \varepsilon_t$$

$$Y_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_k X_{kt} + \rho_1 \mu_{t-1} + \cdots + \rho_p \mu_{t-p} + \varepsilon_t$$

H_0: $\rho_1 = \rho_2 = \ldots = \rho_p = 0$

$$LM = (n-p)R^2 \sim \chi^2(p)$$

***n*为样本容量，R²为如下辅助回归的可决系数**

$$\tilde{e}_t = \beta_0 + \beta_1 X_{1t} + \cdots + \beta_k X_{kt} + \rho_1 \tilde{e}_{t-1} + \cdots + \rho_p \tilde{e}_{t-p} + \varepsilon_t$$

- 如何从直观上理解LM统计量？
- 从1阶、2阶、…逐次向更高阶检验。

四、序列相关的补救

—广义最小二乘法

（GLS: Generalized least squares）

—广义差分法

(GD：Generalized Difference)

1、广义最小二乘法（GLS）

- **GLS的原理与WLS相同，只是将<u>权矩阵W</u>、换为更一般的<u>方差-协方差矩阵Ω</u>。**
- **模型的GLS估计量为：**

$$
\mathrm{Cov}(\mu,\mu') = E(\mu\mu') = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{bmatrix} = \sigma^2 \mathbf{\Omega}
$$

$$
\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{Y}
$$

- 如何得到矩阵Ω?

对Ω的形式进行特殊设定后，可得到其估计值。

例如, 设定随机扰动项为一阶序列相关形式

$$\mu_i=\rho\mu_{i-1}+\varepsilon_i$$

$$Var(u_i)=\frac{\sigma_\varepsilon^2}{1-\rho^2}$$

$$\boxed{Cov(\mu,\mu')=\frac{\sigma_\varepsilon^2}{1-\rho^2}\begin{pmatrix}1 & \rho & \cdots & \rho^{n-1}\\ \rho & 1 & \cdots & \rho^{n-2}\\ \cdots & \cdots & \cdots & \cdots\\ \rho^{n-1} & \rho^{n-2} & \cdots & 1\end{pmatrix}=\sigma^2\mathbf{\Omega}}$$

2、广义差分法(GD：LGeneralized Difference)

- 广义差分法是将原模型变换为满足OLS法的差分模型（将u_t剥离为白噪声ε_t），再进行OLS估计。

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + u_t$$

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_l u_{t-l} + \varepsilon_t$$

$$\boxed{Y_t - \rho_1 Y_{t-1} - \cdots - \rho_l Y_{t-l} = \beta_0(1 - \rho_1 - \cdots - \rho_l) + \beta_1(X_{1t} - \rho_1 X_{1t-1} - \cdots - \rho_l X_{1t-l})}$$

$$\boxed{+ \cdots + \beta_k(X_{kt} - \rho_1 X_{kt-1} - \cdots - \rho_l X_{kt-l}) + \varepsilon_t}$$

该模型为广义差分模型，不存在序列相关问题。

3、随机误差项相关系数的估计

- 应用广义最小二乘法或广义差分法，都必须已知随机误差项的自相关系数$\rho_1, \rho_2, \cdots, \rho_L$。
- 实际上，人们并不知道它们的具体数值，所以必须首先对它们进行估计。

- 常用的估计方法有：
  - 科克伦-奥科特（**Cochrane-Orcutt**）迭代法
  - 杜宾（**durbin**）两步法

- 科克伦-奥科特迭代法

采用OLS法估计

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \mu_i$$

得到随机误差项的“近似估计值”，作为**自相关方程**的样本观测值

$$\mu_t = \rho_1 \mu_{t-1} + \rho_2 \mu_{t-2} + \cdots + \rho_l \mu_{t-l} + \varepsilon_t$$

$$\hat{\rho}_1, \hat{\rho}_2, \cdots, \hat{\rho}_p$$

$$Y_t - \rho_1 Y_{t-1} - \cdots - \rho_l Y_{t-l} = \beta_0 (1 - \rho_1 - \cdots - \rho_l) + \beta_1 (X_{1t} - \rho_1 X_{1t-1} - \cdots - \rho_l X_{1t-l})$$

$$+ \cdots + \beta_k (X_{kt} - \rho_1 X_{kt-1} - \cdots - \rho_l X_{kt-l}) + \varepsilon_t$$

$$\hat{\hat{\beta}}_0, \hat{\hat{\beta}}_1, \cdots, \hat{\hat{\beta}}_k$$

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \mu_i$$

再得到u_t的估计值，第二次估计自相关方程

$$\hat{\hat{\rho}}_1, \hat{\hat{\rho}}_2, \cdots, \hat{\hat{\rho}}_p$$

- 类似地，可进行第三次、第四次迭代。
- 两次迭代过程，被称为科克伦-奥科特两步法。

- 杜宾（durbin）两步法

该方法仍是先估计$\rho_1,\rho_2,\ldots,\rho_l$，再对差分模型进行估计。

$$Y_t-\rho_1Y_{t-1}-\cdots-\rho_lY_{t-l}=\beta_0(1-\rho_1-\cdots-\rho_l)+\beta_1(X_{1t}-\rho_1X_{1t-1}-\cdots-\rho_lX_{1t-l})$$
$$+\cdots+\beta_k(X_{kt}-\rho_1X_{kt-1}-\cdots-\rho_lX_{kt-l})+\varepsilon_t$$

$$Step1:\quad Y_t=\rho_1Y_{t-1}+\cdots+\rho_lY_{t-l}+\beta_0(1-\rho_1-\cdots-\rho_l)+\beta_1(X_t-\rho_1X_{t-1}-\cdots-\rho_lX_{t-l})$$
$$+\cdots+\beta_k(X_{kt}-\rho_1X_{kt-1}-\cdots-\rho_lX_{kt-l})+\varepsilon_t$$

$$\hat\rho_1,\hat\rho_2,\cdots,\hat\rho_l$$

$$\hat\beta_0^*,\hat\beta_1^*,\cdots,\hat\beta_k^*\qquad \hat\beta_0=\hat\beta_0^*/(1-\hat\rho_1-\cdots-\hat\rho_l)\qquad \hat\beta_j=\hat\beta_j^*$$

- 应用软件中的广义差分法

在Eview/TSP软件包下，广义差分采用了科克伦-奥科特（Cochrane-Orcutt）迭代法估计ρ。

在解释变量中引入AR(1)、AR(2)、…，即可得到参数和ρ_1、ρ_2、…的估计值。

其中 AR(m) 表示随机误差项的m阶自回归。在估计过程中自动完成了ρ_1、ρ_2、…的迭代。

注意：

- 如果能够找到一种方法，**求得Ω或各序列相关系数 ρ_{ij}** 的估计量，使得GLS能够实现，则称为**可行的广义最小二乘法**（**FGLS, Feasible Generalized Least Squares**）。
- **FGLS**估计量，也称为可行的广义最小二乘估计量（**feasible general least squares estimators**）

- 可行的广义最小二乘估计量不再是无偏的，但却是一致的，而且在科克伦-奥科特迭代法下，估计量也具有渐近有效性。
- 前面提出的方法，就是**FGLS**。

4、稳健标准误法（Newey-West standard errors）

- 应用软件中推荐的一种选择。适合样本容量足够大的情况。
- 仍然采用**OLS**，但对**OLS**估计量的标准差进行修正。
- 与不附加选择的**OLS**估计比较，参数估计量没有变化，但是参数估计量的方差和标准差变化明显。

- 存在异方差和序列相关时，若采用稳健/一致标准误差：即使仍采用**OLS**估计时，变量的显著性检验依然有效。

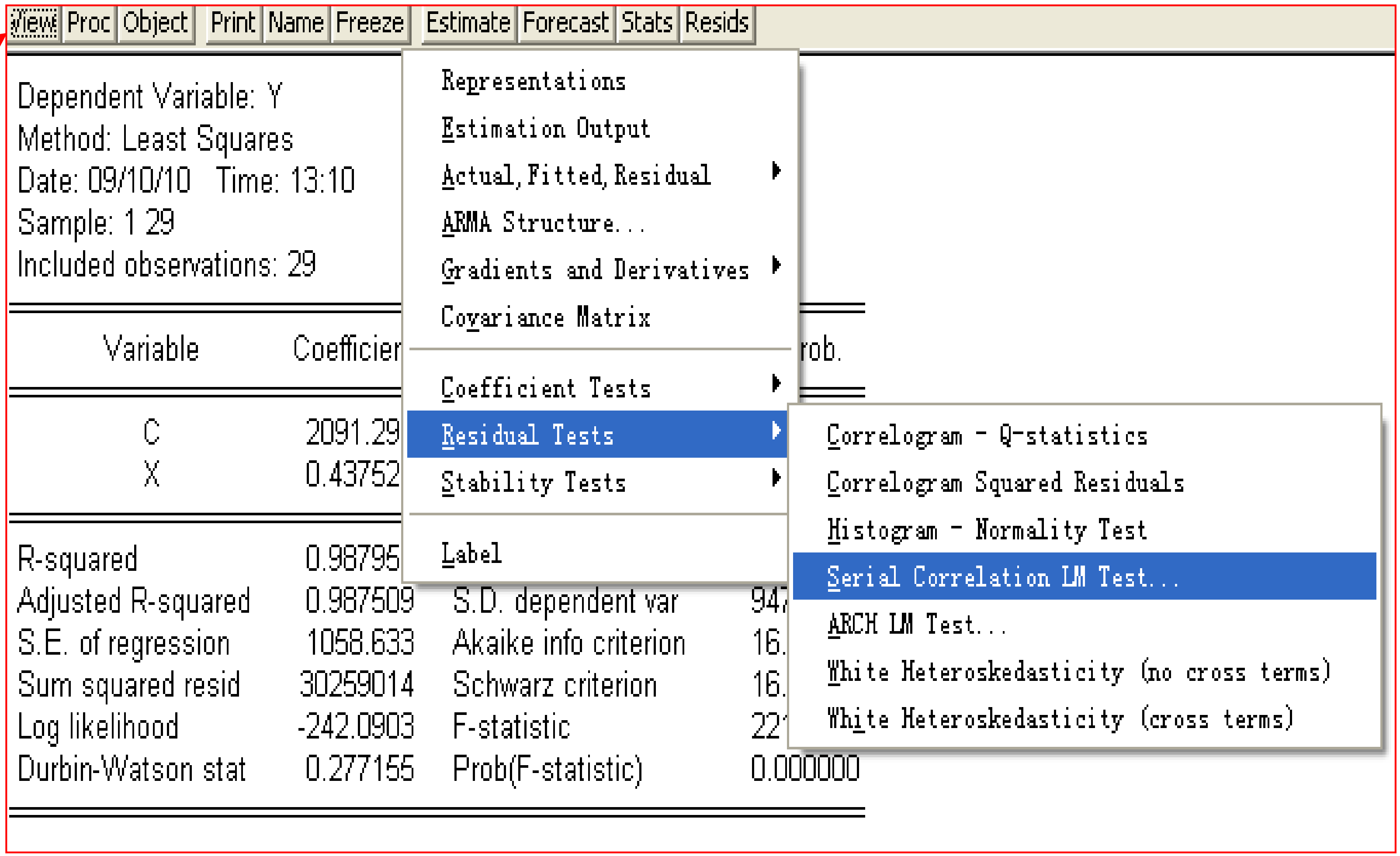
演示：教材例4.2.1(只包含1个解释变量：表2.6.3)

Dependent Variable: Y
Method: Least Squares
Date: 09/10/10 Time: 13:10
Sample: 1 29
Included observations: 29

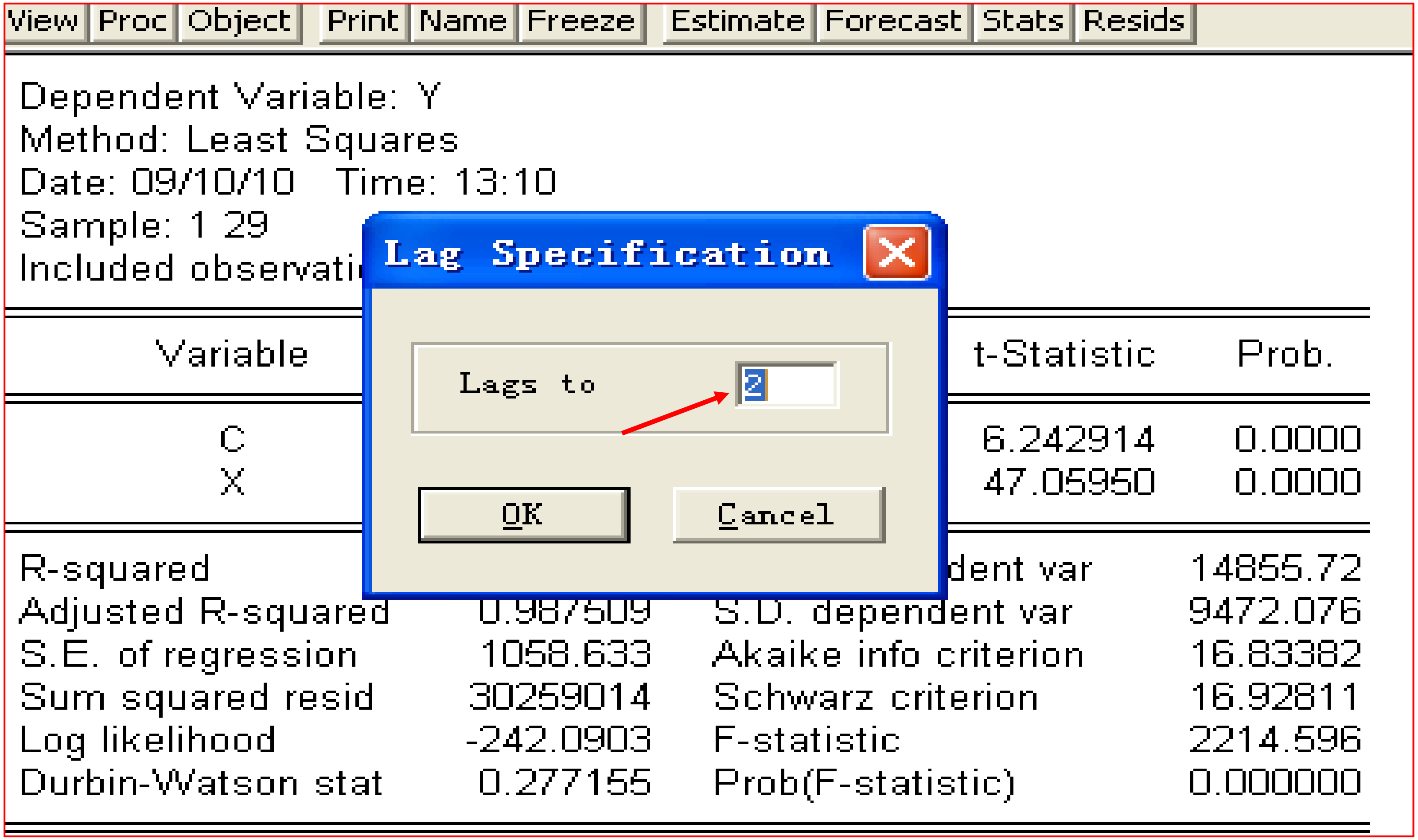
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2091.295	334.9869	6.242914	0.0000
X	0.437527	0.009297	47.05950	0.0000

R-squared	0.987955	Mean dependent var	14855.72
Adjusted R-squared	0.987509	S.D. dependent var	9472.076
S.E. of regression	1058.633	Akaike info criterion	16.83382
Sum squared resid	30259014	Schwarz criterion	16.92811
Log likelihood	-242.0903	F-statistic	2214.596
Durbin-Watson stat	0.277155	Prob(F-statistic)	0.000000

LM检验


View Proc Object Print Name Freeze Estimate Forecast Stats Resids
Dependent Variable: Y
Method: Least Squares
Date: 09/10/10 Time: 13:10
Sample: 1 29
Included observations: 29
Variable Coefficier
rob.
C 2091.29
X 0.43752
R-squared 0.98795
Adjusted R-squared 0.987509 S.D. dependent var 94
S.E. of regression 1058.633 Akaike info criterion 16.
Sum squared resid 30259014 Schwarz criterion 16.
Log likelihood -242.0903 F-statistic 22
Durbin-Watson stat 0.277155 Prob(F-statistic) 0.000000
Representations
Estimation Output
Actual, Fitted, Residual
ARMA Structure...
Gradients and Derivatives
Covariance Matrix
Coefficient Tests
Residual Tests
Stability Tests
Label
Correlogram - Q-statistics
Correlogram Squared Residuals
Histogram - Normality Test
Serial Correlation LM Test...
ARCH LM Test...
White Heteroskedasticity (no cross terms)
White Heteroskedasticity (cross terms)

LM检验（2阶相关）


View Proc Object Print Name Freeze Estimate Forecast Stats Resids
Dependent Variable: Y
Method: Least Squares
Date: 09/10/10 Time: 13:10
Sample: 1 29
Included observati
Variable
t-Statistic
Prob.
C
6.242914
0.0000
X
47.05950
0.0000
R-squared
dent var
14855.72
Adjusted R-squared
0.987509
S.D. dependent var
9472.076
S.E. of regression
1058.633
Akaike info criterion
16.83382
Sum squared resid
30259014
Schwarz criterion
16.92811
Log likelihood
-242.0903
F-statistic
2214.596
Durbin-Watson stat
0.277155
Prob(F-statistic)
0.000000
Lag Specification
Lags to
2
OK
Cancel

LM检验（2阶相关）

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	55.32449	Probability	0.000000
Obs*R-squared	23.65532	Probability	0.000007

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 09/10/10 Time: 16:05
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	100.2903	170.9608	0.586628	0.5627
X	-0.005104	0.005482	-0.930993	0.3608
RESID(-1)	1.462964	0.179340	8.157501	0.0000
RESID(-2)	-0.612323	0.224966	-2.721851	0.0117

R-squared	0.815701	Mean dependent var	3.29E-13
Adjusted R-squared	0.793585	S.D. dependent var	1039.557
S.E. of regression	472.3013	Akaike info criterion	15.28055
Sum squared resid	5576712.	Schwarz criterion	15.46915
Log likelihood	-217.5680	F-statistic	36.88299
Durbin-Watson stat	1.946352	Prob(F-statistic)	0.000000

LM检验（3阶相关）

Breusch-Godfrey Serial Correlation LM Test:

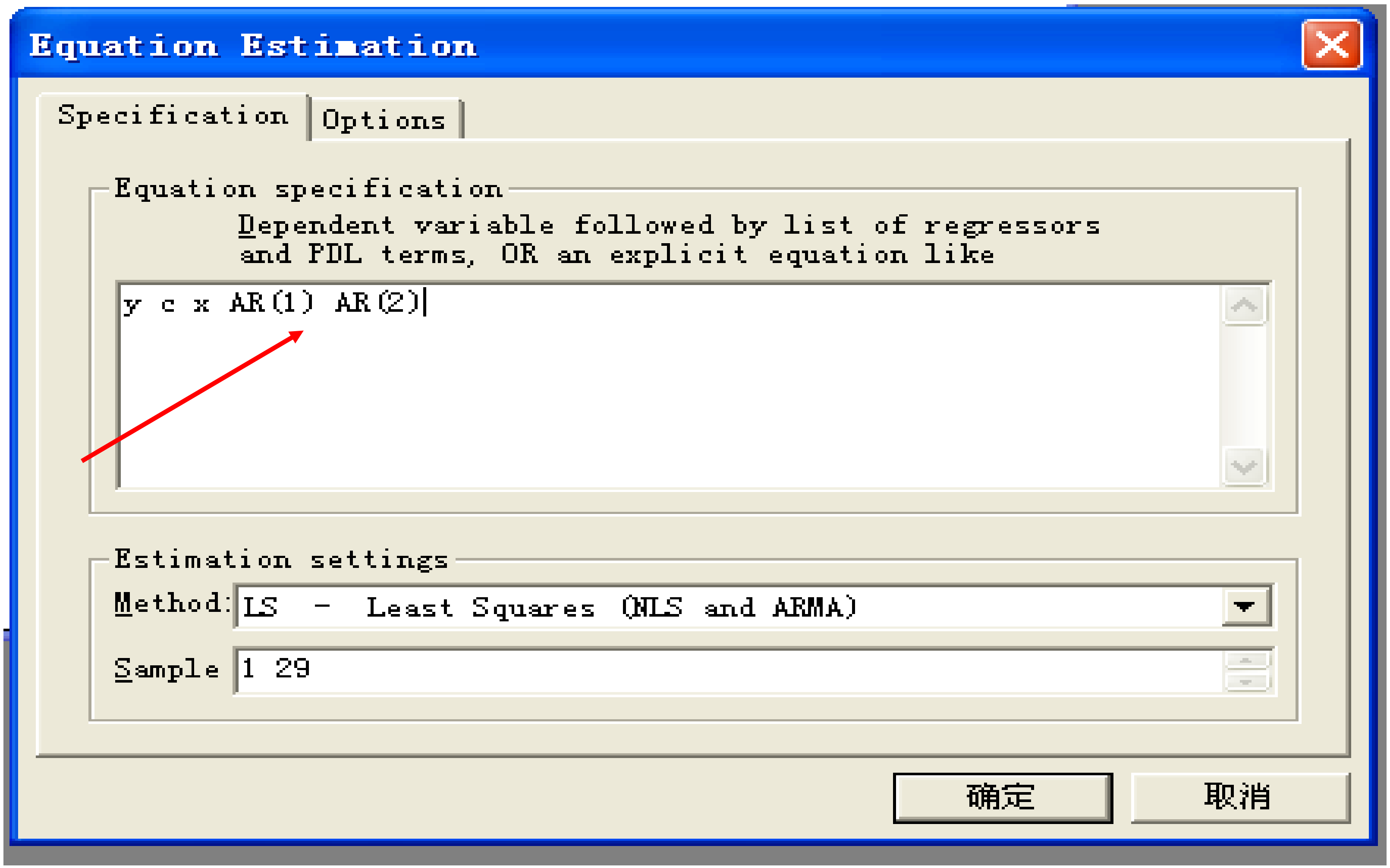
F-statistic	38.02340	Probability	0.000000
Obs*R-squared	23.95909	Probability	0.000025

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 09/10/10 Time: 16:10
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	29.04221	179.5134	0.161783	0.8728
X	-0.002306	0.005911	-0.390083	0.6999
RESID(-1)	1.350088	0.201019	6.716223	0.0000
RESID(-2)	-0.299664	0.342514	-0.874897	0.3903
RESID(-3)	-0.306371	0.254759	-1.202591	0.2409

R-squared	0.826175	Mean dependent var	3.29E-13
Adjusted R-squared	0.797205	S.D. dependent var	1039.557
S.E. of regression	468.1418	Akaike info criterion	15.29101
Sum squared resid	5259762.	Schwarz criterion	15.52675
Log likelihood	-216.7196	F-statistic	28.51755
Durbin-Watson stat	1.864246	Prob(F-statistic)	0.000000

广义差分法（选择2阶差分）



广义差分法（选择2阶差分）

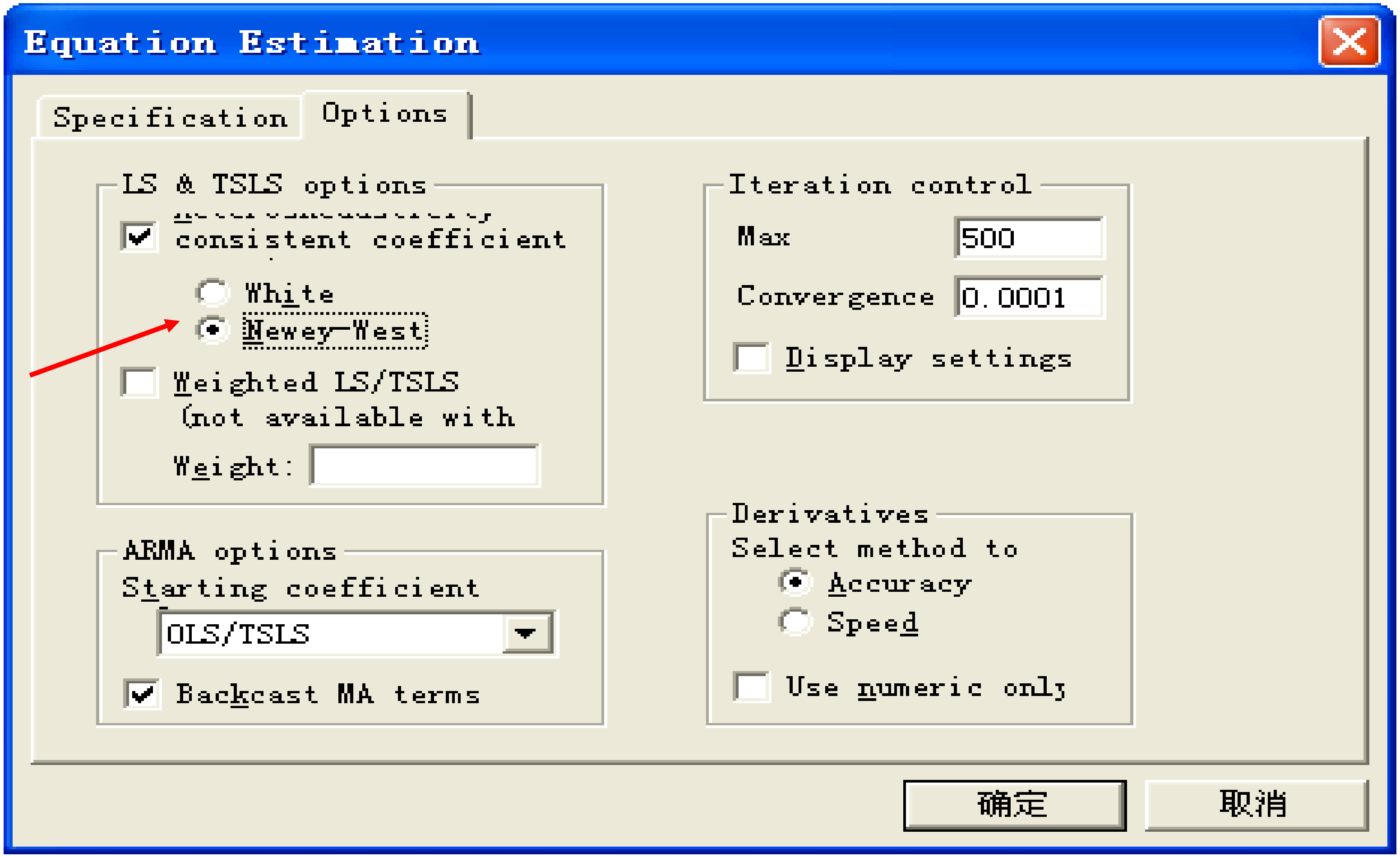
Dependent Variable: Y
Method: Least Squares
Date: 09/10/10 Time: 16:13
Sample (adjusted): 3 29
Included observations: 27 after adjustments
Convergence achieved after 75 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	54612.57	459044.1	0.118970	0.9063
X	0.284508	0.065114	4.369371	0.0002
AR(1)	1.396762	0.212485	6.573462	0.0000
AR(2)	-0.401550	0.233281	-1.721319	0.0986

R-squared	0.998821	Mean dependent var	15656.89
Adjusted R-squared	0.998667	S.D. dependent var	9324.833
S.E. of regression	340.4411	Akaike info criterion	14.63431
Sum squared resid	2665703.	Schwarz criterion	14.82629
Log likelihood	-193.5633	F-statistic	6494.382
Durbin-Watson stat	1.958047	Prob(F-statistic)	0.000000

Inverted AR Roots	.99	.40

Newey-West standard errors



Newey-West standard errors

Dependent Variable: Y
Method: Least Squares
Date: 09/10/10 Time: 16:16
Sample: 1 29
Included observations: 29
Newey-West HAC Standard Errors & Covariance (lag truncation=3)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2091.295	493.4785	4.237864	0.0002
X	0.437527	0.019625	22.29454	0.0000

R-squared	0.987955	Mean dependent var	14855.72
Adjusted R-squared	0.987509	S.D. dependent var	9472.076
S.E. of regression	1058.633	Akaike info criterion	16.83382
Sum squared resid	30259014	Schwarz criterion	16.92811
Log likelihood	-242.0903	F-statistic	2214.596
Durbin-Watson stat	0.277155	Prob(F-statistic)	0.000000

5、虚假序列相关问题

由于随机项的序列相关往往是在模型设定中遗漏了重要解释变量或对模型函数形式设定有误，这种情形可称为**虚假序列相关(false autocorrelation)**，应在**模型设定**中排除。

避免产生虚假序列相关性的措施：是在开始时建立一个“一般”模型，然后逐渐剔除确实不显著变量。

五、案例——中国居民总量消费函数（自学）

- 步骤
  - 对一元模型进行**OLS**估计；
  - 进行序列相关检验，存在正相关；
  - 分析产生序列相关的原因，为了消除虚假相关，引入时间趋势项；
  - 估计新模型，经**D.W.**检验，仍然存在正相关；
  - 进行**LM**检验，判断存在**1**阶序列相关；
  - 采用广义差分法估计模型；
  - 采用稳健标准误方法估计模型。

§4.3 多重共线性*
Multicollinearity

一、多重共线性的概念

二、多重共线性的后果**

三、多重共线性的检验

四、克服多重共线性的方法*

五、例题

六、分部回归与多重共线性***

一、多重共线性的概念

1、多重共线性

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \mu_i \qquad i = 1, 2, \cdots, n$$

如果某两个或多个解释变量之间出现了相关性，则称为**多重共线性(Multicollinearity)**。

$$\boxed{c_1 X_{1i} + c_2 X_{2i} + \cdots + c_k X_{ki} = 0}$$

perfect multicollinearity

$$\boxed{R(\mathbf{X}) < k + 1}$$

$$\boxed{c_1 X_{1i} + c_2 X_{2i} + \cdots + c_k X_{ki} + v_i = 0}$$

approximate multicollinearity

2、实际经济问题中的多重共线性

- 产生多重共线性的主要原因：

（1）经济变量相关的共同趋势

（2）滞后变量的引入

（3）样本资料的限制

二、多重共线性的后果
Consequences of Multicollinearity

1、完全共线性下参数估计量不存在

$$\mathbf{Y} = \mathbf{X}\beta + \mu$$

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

如果存在完全共线性，则$(X'X)^{-1}$不存在，无法得到参数的估计量。

2、近似共线性下OLS估计量非有效

- 近似共线性下，可以得到OLS参数估计量，但参数估计量**方差**的表达式为

$$\boxed{Cov(\hat{\beta}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}}$$

由于$|X'X| \approx 0$，引起$(X'X)^{-1}$主对角线元素较大，使参数估计值的<u>方差增大</u>，**OLS参数估计量非有效。**

- 以二元线性模型 $y=\beta_1 x_1+\beta_2 x_2+\mu$ 为例:

$$\boxed{\operatorname{var}(\hat{\beta}_1)=\sigma^2(X'X)^{-1}_{11}=\frac{\sigma^2\sum x_{2i}^2}{\sum x_{1i}^2\sum x_{2i}^2-(\sum x_{1i}x_{2i})^2}=\frac{\sigma^2/\sum x_{1i}^2}{1-(\sum x_{1i}x_{2i})^2\Big/\sum x_{1i}^2\sum x_{2i}^2}}$$

$$\boxed{=\frac{\sigma^2}{\sum x_{1i}^2}\cdot\frac{1}{1-r^2}}$$

$$\boxed{\frac{(\sum x_{1i}x_{2i})^2}{\sum x_{1i}^2\sum x_{2i}^2}}$$

恰为X_1与X_2的线性相关系数的平方r^2

由于 $r^2\le 1$，故 $1/(1-r^2)\ge 1$。

$$Var\left(\hat{\beta}_j\right)=\frac{\sigma^2}{SST_j\cdot\left(1-R_j^2\right)}$$

当完全不共线时, $r^2=0$ $\boxed{\operatorname{var}(\hat{\beta}_1)=\sigma^2/\sum x_{1i}^2}$

当近似共线时, $0< r^2 <1$ $\boxed{\operatorname{var}(\hat{\beta}_1)=\frac{\sigma^2}{\sum x_{1i}^2}\cdot\frac{1}{1-r^2}>\frac{\sigma^2}{\sum x_{1i}^2}}$

多重共线性使参数估计值的方差增大，$1/(1-r^2)$为**方差膨胀因子(Variance Inflation Factor, VIF)**

表 4.3.1　方差膨胀因子表

相关系数平方	0	0.5	0.8	0.9	0.95	0.96	0.97	0.98	0.99	0.999
方差膨胀因子	1	2	5	10	20	25	33	50	100	1000

当完全共线时， $r^2=1$， $\boxed{\operatorname{var}(\hat{\beta}_1)=\infty}$

3、参数估计量经济含义不合理**

- 如果模型中两个解释变量具有线性相关性，例如 $X_2 = \lambda X_1$，

这时，X_1和X_2前的参数β_1、β_2并不反映各自与被解释变量之间的结构关系，而是反映它们对被解释变量的共同影响。

β_1、β_2已经失去了应有的经济含义，于是经常表现出**反常/反号的现象**：例如β_1本来应该是正的，结果恰是负的。

4、变量的显著性检验失去意义

存在多重共线性时

参数估计值的方差与标准差变大

容易使通过样本计算的t值小于临界值，
误导作出参数为0的推断

可能将重要的解释变量、错误排除在外

5、模型的预测功能失效

- 变大的方差容易使区间预测的“区间”变大，使预测失去意义。

注意：

除非是完全共线性，多重共线性并不意味着任何基本假设的违背；

因此，即使出现较高程度的多重共线性，**OLS估计量仍具有线性性等良好的统计性质。**

问题在于，即使OLS法仍是最好的估计方法，它却不是“完美的”，尤其是**在统计推断上、无法给出真正有用的信息。**

三、多重共线性的检验
Detection of Multillinearity

说明

多重共线性表现为解释变量之间具有相关关系，所以**用于多重共线性的检验方法主要是统计方法：**如**判定系数检验法**、**逐步回归检验法**等。

多重共线性检验的任务是：

（1）检验多重共线性是否存在；

（2）估计多重共线性的范围，即判断哪些变量之间存在共线性。

1、检验多重共线性是否存在

(1) 对两个解释变量的模型，采用简单相关系数法

求出X_1与X_2的简单相关系数r，若|r|接近1，则说明两变量存在较强的多重共线性。

(2) 对多个解释变量的模型，采用综合统计检验法

若在OLS下，“ R^2与F值较大、但t检验值较小”——说明各解释变量对Y的联合线性作用显著，但各解释变量间存在共线性、使得它们对Y的独立作用不能分辨，故t检验不显著。

2、判明存在多重共线性的范围

(1) 判定系数检验法

使模型中每一个解释变量、分别以其余解释变量为解释变量进行**辅助回归**（Auxiliary Regression），并计算相应的拟合优度。

如果某一种回归 $X_{ji}=\alpha_1 X_{1i}+\alpha_2 X_{2i}+\ldots\alpha_L X_{Li}$ 的**判定系数**较大，说明 X_j与其他X 间、存在**共线性**。

可构造F检验：

$$\boxed{F_j=\frac{R_{j.}^2/(k-1)}{(1-R_{j.}^2)/(n-k)}\sim F(k-1,n-k)}$$

(2) 排除变量法/逐步向后回归法
(Stepwise Backward Regression)

在模型中、<u>排除</u>某一个解释变量X_j，估计模型；

如果拟合优度与包含X_j时十分接近，则说明X_j与其它解释变量之间存在共线性。

(3) 逐步向前回归法（Stepwise Forward Regression）

- 以Y为被解释变量，逐个引入解释变量，构成回归模型，进行模型估计。

- 根据拟合优度的变化，决定新引入的变量是否独立。
  - 如果拟合优度变化显著，**则说明新引入的变量是一个独立解释变量；**
  - 如果拟合优度变化很不显著，**则说明新引入的变量与其它变量之间存在共线性关系。**

四、克服多重共线性的方法
Remedial Measures of Multicollinearity

1、第一类方法：排除引起共线性的变量

- 找出引起多重共线性的解释变量，将它排除。
- 以逐步回归法得到最广泛的应用。
- 注意：“剩余解释变量参数”的经济含义和数值，都发生了变化。

2、第二类方法：差分法

- 时间序列数据为样本的线性模型；
- 将原模型变换为差分模型， 可有效消除原模型中的多重共线性。
- 通常，对经济数据：增量之间线性关系远比总量之间线性关系，弱得多。

$$\boxed{\Delta Y_i = \beta_1 \Delta X_{1i} + \beta_2 \Delta X_{2i} + \cdots + \beta_k \Delta X_{ki} + \mu_i - \mu_{i-1}}$$

- 另外一个重要的意义，差分可以将非平稳序列变为平稳序列。在第8章将介绍。

差分模型随机项的序列相关问题

3、第三类方法：减小参数估计量的方差

- 多重共线性的主要后果，是参数估计量具有较大的方差。
- 采取适当方法减小参数估计量的方差，虽然没有消除模型中的多重共线性，但确能消除多重共线性造成的后果。
- 例如，增加样本容量，可使参数估计量的方差减小。
- 例如，岭回归法

*岭回归法（**Ridge Regression**）

- 20世纪70年代发展，**以引入偏误为代价减小参数估计量的方差**。
- 具体方法是：引入矩阵**D**，使参数估计量为

$$\boxed{\hat{\beta} = (\mathbf{X}'\mathbf{X} + \mathbf{D})^{-1}\mathbf{X}'\mathbf{Y}}$$

其中矩阵**D**一般选择为主对角阵，即**D**=a**I**，a为大于0的常数。

- 显然，与未含D的参数B的估计量相比，估计量有较小的方差。

五、案例——中国粮食生产函数(自学)

- 步骤
  - 以粮食产量作为被解释变量，以影响粮食产量的主要因素“农业化肥施用量、粮食播种面积、成灾面积、农业机械总动力、农业劳动力”为解释变量，建立中国粮食生产函数模型；
  - 用OLS法估计模型；
  - 检验简单相关系数；
  - 找出**最简单的回归形式**；
  - 采用逐步回归方法得到最终模型。

六、补充：分部回归与多重共线性***

Partitioned Regression and Multillinearity

1、分部回归法(Partitioned Regression)

$$\boxed{\mathbf{Y} = \mathbf{X}\mathrm{B} + \mathrm{N}}$$

将解释变量分为两部分，对应的参数也分为两部分

$$\boxed{\mathbf{Y} = \mathbf{X}_1\mathrm{B}_1 + \mathbf{X}_2\mathrm{B}_2 + \mathrm{N}}$$

在满足解释变量（X_1，X_2）与随机误差项N不相关的情况下，可写出关于参数估计量的方程组：

$$\begin{bmatrix} \mathbf{X}_1'\mathbf{Y} \\ \mathbf{X}_2'\mathbf{Y} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \hat{\mathrm{B}}_1 \\ \hat{\mathrm{B}}_2 \end{bmatrix}$$

$$\boxed{\begin{aligned}\hat{B}_1 &= (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'(\mathbf{Y}-\mathbf{X}_2\hat{B}_2) \\ &= (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{Y}-(\mathbf{X}_1'\mathbf{X}_1)^{-1}\cdot\mathbf{X}_1'\mathbf{X}_2\cdot\hat{B}_2\end{aligned}}$$

若有： $\mathbf{X}_1'\mathbf{X}_2=\mathbf{0}$

或 $\hat{B}_2=0$

$$\boxed{\hat{B}_1=(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{Y}}$$

这是仅以X_1作为解释变量时的参数估计量。

$$\boxed{\hat{B}_2=(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{Y}}$$

这就是仅以X_2作为解释变量时的参数估计量。

2、由分部回归法得到的启示**

- 如果一个多元线性模型的解释变量之间完全正交，可将该多元模型分为多个一元模型、二元模型、…进行估计，参数估计结果不变；

- 实际模型由于*存在或轻或重的共线性*，若将它们分为多个一元模型、二元模型、…进行估计，参数估计结果将发生变化。

- 当模型变量存在共线性，将某个“不显著共线性变量”去掉：**剩余变量的参数估计结果、将发生变化，且经济含义有发生变化**。

- 严格地说，实际模型由于总存在一定程度共线性，所以每个参数估计量、并不真正反映对应变量与被解释变量之间的结构关系。

§4. 4　随机解释变量问题**
Random Explanatory Variables

一、随机解释变量问题

二、随机解释变量的后果

三、工具变量法*

四、解释变量的内生性检验***

五、例题

一、随机解释变量问题

1、随机解释变量问题

$$\boxed{Y_i=\beta_0+\beta_1 Y_{1i}+\beta_2 X_{2i}+\cdots+\beta_k X_{ki}+\mu_i}$$

- 基本假设:解释变量$X_1,X_2,\ldots,X_k$是确定性变量。

- 如果存在一个或多个随机变量X_j作为解释变量，则称原模型出现**随机解释变量问题**。

假设X_2为随机解释变量。

对随机解释变量问题，分三种不同情况：

(1) 随机解释变量与随机误差项独立(Independence)

$$Cov(X_{2,}\mu) = E(x_2\mu) = E(x_2)E(\mu) = 0$$

(2) 随机解释变量与随机误差项同期无关(contemporaneously uncorrelated)，但异期相关。

$$Cov(X_{2i,}\mu_i) = E(x_{2i}\mu_i) = 0$$

$$Cov(X_{2i,}\mu_{i-s}) = E(x_{2i}\mu_{i-s}) \neq 0 \quad s \neq 0$$

(3) 随机解释变量与随机误差项同期相关(contemporaneously correlated)。

$$Cov(X_{2i,}\mu_i) = E(x_{2i}\mu_i) \neq 0$$

2、实际经济问题中的随机解释变量问题

- 在实际经济问题中，经济变量往往都具有随机性。
- 但在单方程计量经济学模型中，通常外生变量X_j都被认为是确定性的。
- 于是，**随机解释变量问题**多表现于：用**滞后被解释变量y_{t-s}**，作模型解释变量的情况。
- 例如：

（1）耐用品存量调整模型

耐用品存量Q_t，由前一个时期存量Q_{t-1}和当期收入I_t共同决定：

$$Q_t=\beta_0+\beta_1 I_t+\beta_2 Q_{t-1}+\mu_t \qquad t=1,\ldots T$$

如果模型不存在随机误差项的序列相关性，那么随机解释变量 **Q_{t-1}** 只与μ_{t-1}相关，与μ_t不相关。

属于上述第2种异期相关的情况。

(2) 合理预期的消费函数模型

$$\boxed{C_t = \beta_0 + \beta_1 Y_t^e + \mu_t}$$

$$\boxed{Y_t^e = (1-\lambda)Y_t + \lambda Y_{t-1}^e}$$

$$\boxed{C_t = \beta_0 + \beta_1(1-\lambda)Y_t + \beta_1\lambda Y_{t-1}^e + \mu_t}$$

$$\boxed{= \beta_0 + \beta_1(1-\lambda)Y_t + \lambda(C_{t-1} - \beta_0 - \mu_{t-1}) + \mu_t}$$

$$\boxed{= \beta_0(1-\lambda) + \beta_1(1-\lambda)Y_t + \lambda C_{t-1} + \mu_t - \lambda\mu_{t-1}}$$

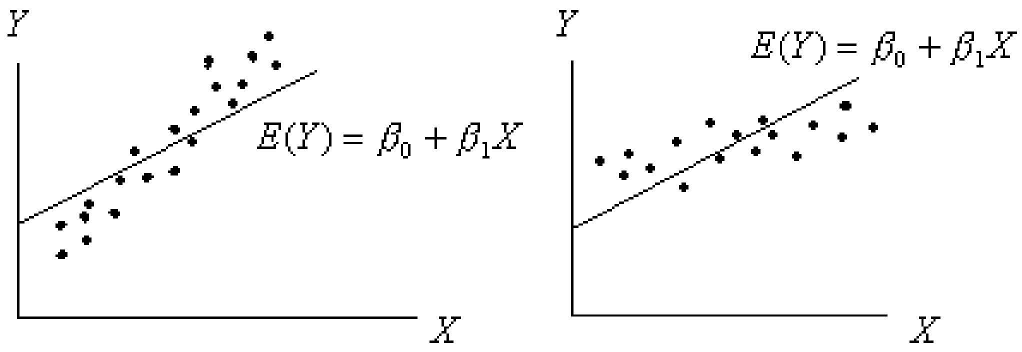
C_{t-1}是一随机解释变量，且与 $(\mu_t-\lambda\mu_{t-1})$高度同期相关（为什么?）：**属于上述第3种情况。**

二、随机解释变量的后果

- 计量经济学模型一旦出现随机解释变量，若仍采用OLS法估计模型参数，<u>不同性质的随机解释变量、会产生不同的后果</u>。

- 下面以一元线性回归模型为例进行说明。

1、随机解释变量与随机误差项相关图



（a）正相关　　（b）负相关

拟合的样本回归线可能低估截距项，而高估斜率项。

拟合的样本回归线可能高估截距项，而低估斜率项。

2、如果X与μ相互独立，OLS参数估计量仍然是无偏、一致估计量。

$$\boxed{Y_t = \beta_0 + \beta_1 X_t + \mu_t}$$

$$\boxed{\hat{\beta}_1 = \frac{\sum x_t y_t}{\sum x_t^2} = \beta_1 + \frac{\sum x_t \mu_t}{\sum x_t^2}}$$

$$\boxed{E(\hat{\beta}_1) = \beta_1}$$

3、如果X与μ同期不相关，异期相关，得到的参数估计量有偏、但却是一致的。

$$E(\hat{\beta}_1) = \beta_1 + E(\sum \frac{x_t}{\sum x_t^2}\mu_t) = \beta_1 + \sum E(k_t \mu_t)$$

k_t的分母中包含不同期的X_{t-s}，k_t与μ_t相关

$$E(\hat{\beta}_1) \neq \beta_1$$

$$P\lim_{n\to\infty}\left(\beta_1 + \frac{\sum x_t \mu_t}{\sum x_t^2}\right) = \beta_1 + \frac{P\lim(\frac{1}{n}\sum x_t \mu_t)}{P\lim(\frac{1}{n}\sum x_t^2)}$$

$$= \beta_1 + Cov(X_t, \mu_t)/Var(X_t) = 0$$

4、如果X与μ同期相关，得到的参数估计量有偏、且非一致。

前面已经证明

三、工具变量法
Instrument variables

1、工具变量的选取

- 工具变量z：模型估计过程中被作为工具使用，以替代模型中与随机误差项相关的随机解释变量。

- 选择为工具变量的变量z，必须满足以下条件：
  - 与所替代的随机解释变量x_j高度相关；
  - 与随机误差项u不相关；
  - 与模型中其它解释变量x_i不相关，以避免出现多重共线性。

2、工具变量的应用

$$
\begin{cases}
\Sigma Y_i - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) = \Sigma \hat{\mu}_i \\
\Sigma Y_i X_{1i} - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{1i} = \Sigma \hat{\mu}_i X_{1i} \\
\Sigma Y_i X_{2i} - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_{2i} X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{2i} = \Sigma \hat{\mu}_i X_{2i} \\
\vdots \\
\Sigma Y_i X_{ki} - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{ki} = \Sigma \hat{\mu}_i X_{ki}
\end{cases}
$$

多元线性模型的正规方程组

$$
\begin{cases}
\Sigma Y_i - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) \approx E(\Sigma \mu_i) = 0 \\
\Sigma Y_i X_{1i} - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{1i} \approx E(\Sigma \mu_i X_{1i}) = 0 \\
\Sigma Y_i X_{2i} - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_{2i} X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{2i} \approx E(\Sigma \mu_i X_{2i}) \neq 0 \\
\vdots \\
\Sigma Y_i X_{ki} - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{ki} \approx E(\Sigma \mu_i X_{ki}) = 0
\end{cases}
$$

X_2为与μ相关的随机变量

$$
\begin{cases}
\Sigma Y_i - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) = E(\Sigma \mu_i) = 0 \\
\Sigma Y_i X_{1i} - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{1i} = E(\Sigma \mu_i X_{1i}) = 0 \\
\Sigma Y_i Z_i - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_{2i} X_{2i} + \cdots + \hat{\beta}_k X_{ki}) Z_i = E(\Sigma \mu_i Z_i) = 0 \\
\vdots \\
\Sigma Y_i X_{ki} - \Sigma(\hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki}) X_{ki} = E(\Sigma \mu_i X_{ki}) = 0
\end{cases}
$$

Z作为X_2的工具变量

- 能否说“用工具变量代替了模型中的随机解释变量”？
- 能否说“其它解释变量、用自己作为工具变量”？
- 能否说“用Z作为X1的工具变量，用X1作为X2的工具变量”（只要Z是选定的，OK！）？

- 这种求模型参数估计量的方法，称为**工具变量法**(**instrumental variable method**)；相应的估计量，称为**工具变量法估计量**（**instrumental variable (IV) estimator**）。

$$\boxed{\mathbf{Y}=\mathbf{X}\beta+\mu} \longrightarrow \boxed{\mathbf{Z}'\mathbf{Y}=\mathbf{Z}'\mathbf{X}\beta} \longrightarrow \boxed{\tilde{\beta}=(\mathbf{Z}'\mathbf{X})^{-1}\mathbf{Z}'\mathbf{Y}}$$

$$\mathbf{Z}'=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{12} & \cdots & X_{1n} \\ Z_1 & Z_2 & \cdots & Z_n \\ \vdots & & & \\ X_{k1} & X_{k2} & \cdots & X_{kn} \end{bmatrix}$$

工具变量矩阵

3、工具变量法估计量是一致估计量

以一元回归为例，工具变量法估计量为：

$$\tilde{\beta}_1 = \frac{\sum z_i(\beta_1 x_i + \mu_i)}{\sum z_i x_i} = \beta_1 + \frac{\sum z_i \mu_i}{\sum z_i x_i}$$

$$P\lim(\tilde{\beta}_1) = \beta_1 + \frac{P\lim \frac{1}{n}\sum z_i \mu_i}{P\lim \frac{1}{n}\sum z_i x_i}$$

$$P\lim \frac{1}{n}\sum z_i \mu_i = \mathrm{cov}(Z_i, \mu_i) = 0$$

$$P\lim \frac{1}{n}\sum z_i x_i = \mathrm{cov}(Z_i, X_i) \neq 0$$

$$\boxed{P\lim(\tilde{\beta}_1) = \beta_1}$$

4、几个重要的概念

(1) 在小样本下，工具变量法估计量仍是<u>有偏的</u>。

$$E\left(\frac{1}{\sum z_i x_i}\sum z_i \mu_i\right) \neq E\left(\frac{1}{\sum z_i x_i}\right)E\left(\sum z_i \mu_i\right) = 0$$

(2) 工具变量、并没有替代模型中的解释变量，只是在估计过程中作为“工具”被使用。

(3) 如果模型中有两个以上的随机解释变量与随机误差项相关，就必须找到两个以上的工具变量。但是，一旦选定工具变量：它们在估计过程<u>被使用的次序</u>，不影响估计结果**(Why？)**。

(4) OLS可以看作工具变量法的一种特殊情况。

(5) 如果1个随机解释变量可以找到多个互相独立的工具变量，人们希望充分利用这些工具变量的信息，就形成了广义矩方法（**Generalized Method of Moments, GMM**）。

在**GMM**中，矩条件个数、大于待估参数的数量，于是如何求解、成为它的核心问题。

工具变量法，又是**GMM**的一个特例。

(6) 要找到与随机扰动项不相关而又与随机解释变量相关的工具变量并不是一件很容易的事

可以用X_{t-1}，作为原解释变量X_t的工具变量。

5、IV演示：居民总消费模型

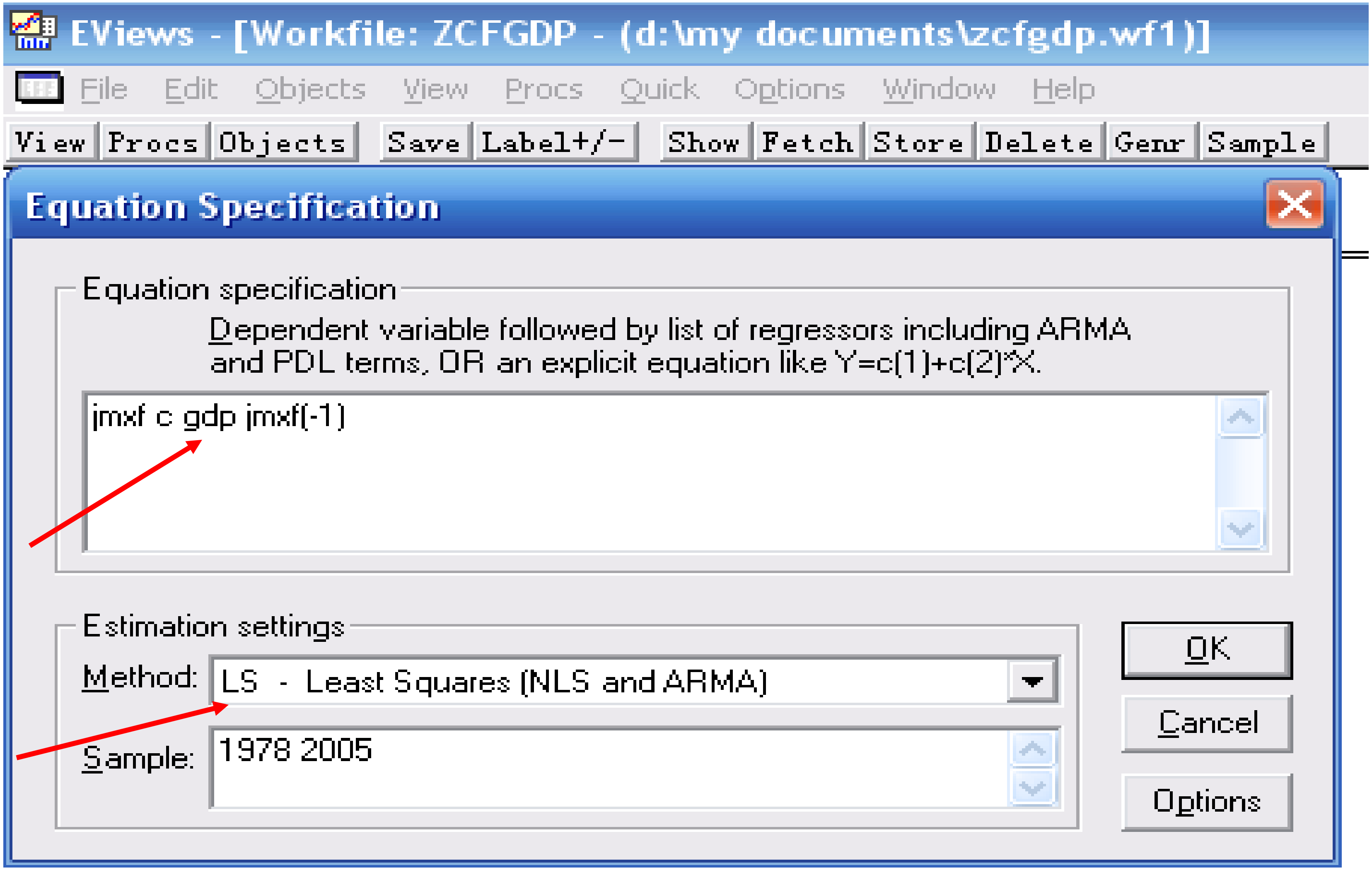
- 以居民消费总额JMXF为被解释变量；
- 以GDP和JMXF(-1)为解释变量；
- 进行OLS估计。

- JMXF(-1)为随机解释变量，且与随机误差项相关：
  - ✓ 以 政府消费ZFXF 作为工具变量，进行IV估计；
  - ✓ 以 政府消费ZFXF和资本形成ZBXC 作为工具变量，进行GMM估计。

数据

obs	JMXF	GDP	JMXF(-1)	ZFXF	ZBXC
1978	1759.100	3605.600	NA	480.0000	1377.900
1979	2011.500	4092.600	1759.100	622.2000	1478.900
1980	2331.200	4592.900	2011.500	676.7000	1599.700
1981	2627.900	5008.800	2331.200	733.6000	1630.200
1982	2902.900	5590.000	2627.900	811.9000	1784.200
1983	3231.100	6216.200	2902.900	895.3000	2039.000
1984	3742.000	7362.700	3231.100	1104.300	2515.100
1985	4687.400	9076.700	3742.000	1298.900	3457.500
1986	5302.100	10508.50	4687.400	1519.700	3941.900
1987	6126.100	12277.40	5302.100	1678.500	4462.000
1988	7868.100	15388.60	6126.100	1971.400	5700.200
1989	8812.600	17311.30	7868.100	2351.600	6332.700
1990	9450.900	19347.80	8812.600	2639.600	6747.000
1991	10730.60	22577.40	9450.900	3361.300	7868.000
1992	13000.10	27565.20	10730.60	4203.200	10086.30
1993	16412.10	36938.10	13000.10	5487.800	15717.70
1994	21844.20	50217.40	16412.10	7398.000	20341.10
1995	28369.70	63216.90	21844.20	8378.500	25470.10
1996	33955.90	74163.60	28369.70	9963.600	28784.90
1997	36921.50	81658.50	33955.90	11219.10	29968.00
1998	39229.30	86531.60	36921.50	12358.90	31314.20
1999	41920.40	90964.10	39229.30	13716.50	32951.50
2000	45854.60	98749.00	41920.40	15661.40	34842.80
2001	49213.20	108972.4	45854.60	17665.10	39769.40
2002	52571.30	120350.3	49213.20	19119.90	45565.00
2003	56834.40	136398.8	52571.30	20615.10	55963.00
2004	63833.50	160280.4	56834.40	23199.40	69168.40
2005	70906.00	186700.9	63833.50	26012.10	79559.80

OLS估计


EViews - [Workfile: ZCFGDP - (d:\my documents\zcfgdp.wf1)]
File Edit Objects View Procs Quick Options Window Help
View Procs Objects Save Label+/- Show Fetch Store Delete Genr Sample
Equation Specification
Equation specification
Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like Y=c(1)+c(2)*X.
jmxf c gdp jmxf(-1)
Estimation settings
Method: LS - Least Squares (NLS and ARMA)
Sample: 1978 2005
OK
Cancel
Options

EViews - [Equation: UNTITLED Workfile: ZCFGDP]

File Edit Objects View Procs Quick Options Window Help

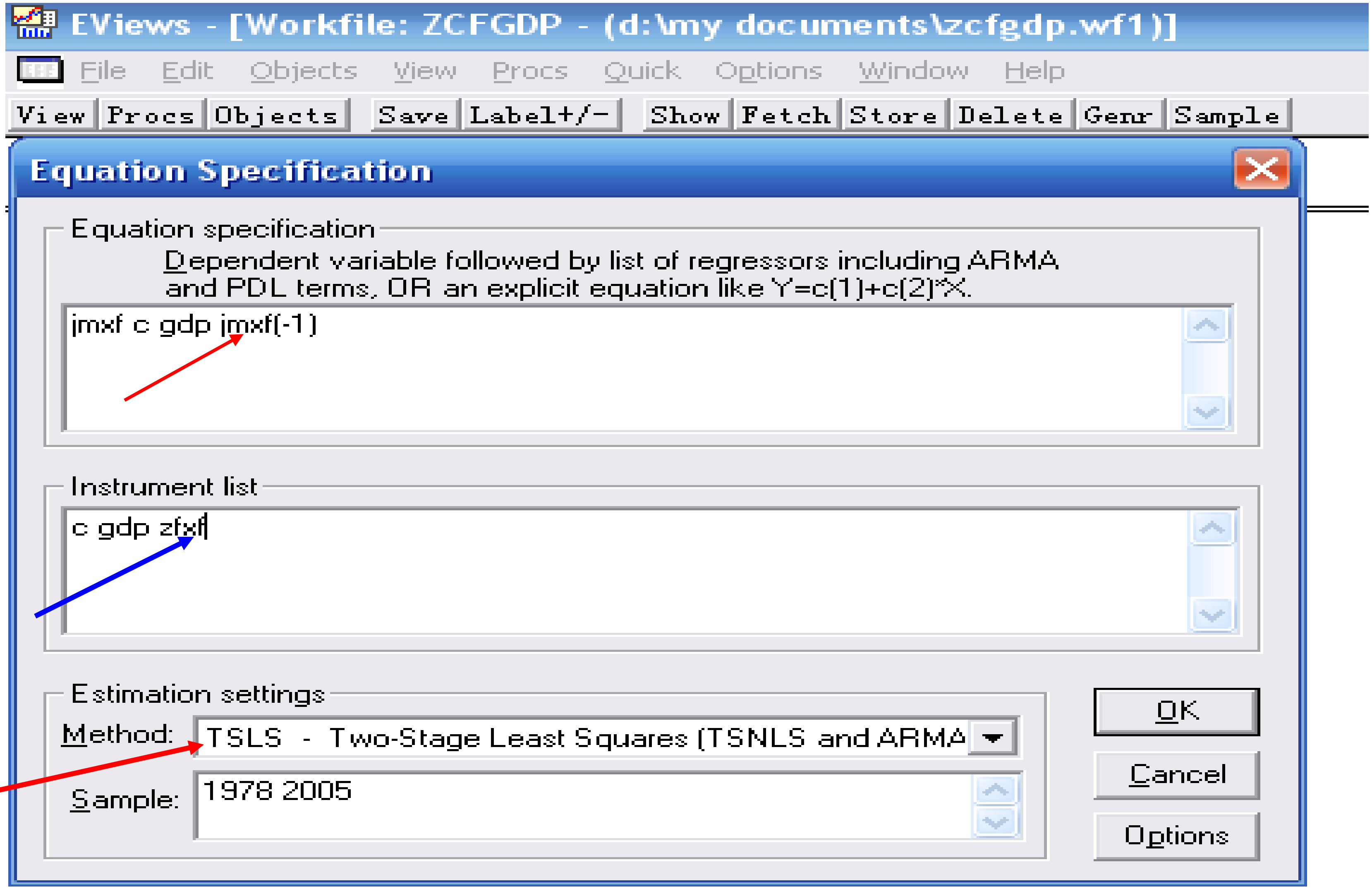
View | Procs | Objects | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: JMXF
Method: Least Squares
Date: 02/14/07 Time: 11:48
Sample(adjusted): 1979 2005
Included observations: 27 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1001.165	315.8403	3.169845	0.0041
GDP	0.136770	0.031572	4.331988	0.0002
JMXF(-1)	0.723818	0.084118	8.604824	0.0000

R-squared	0.997660	Mean dependent var	23729.28
Adjusted R-squared	0.997465	S.D. dependent var	21866.25
S.E. of regression	1101.001	Akaike info criterion	16.95027
Sum squared resid	29092892	Schwarz criterion	17.09425
Log likelihood	-225.8286	F-statistic	5115.625
Durbin-Watson stat	0.494865	Prob(F-statistic)	0.000000

IV估计



EViews - [Workfile: ZCFGDP - (d:\my documents\zcfgdp.wf1)]
File Edit Objects View Procs Quick Options Window Help
View Procs Objects Save Label+/- Show Fetch Store Delete Genr Sample
Equation Specification
Equation specification
Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like Y=c(1)+c(2)*X.
jmxf c gdp jmxf(-1)
Instrument list
c gdp zfxf
Estimation settings
Method: TSLS - Two-Stage Least Squares (TSNLS and ARMA
Sample: 1978 2005
OK
Cancel
Options

EViews - [Equation: UNTITLED Workfile: ZCFGDP]

File Edit Objects View Procs Quick Options Window Help

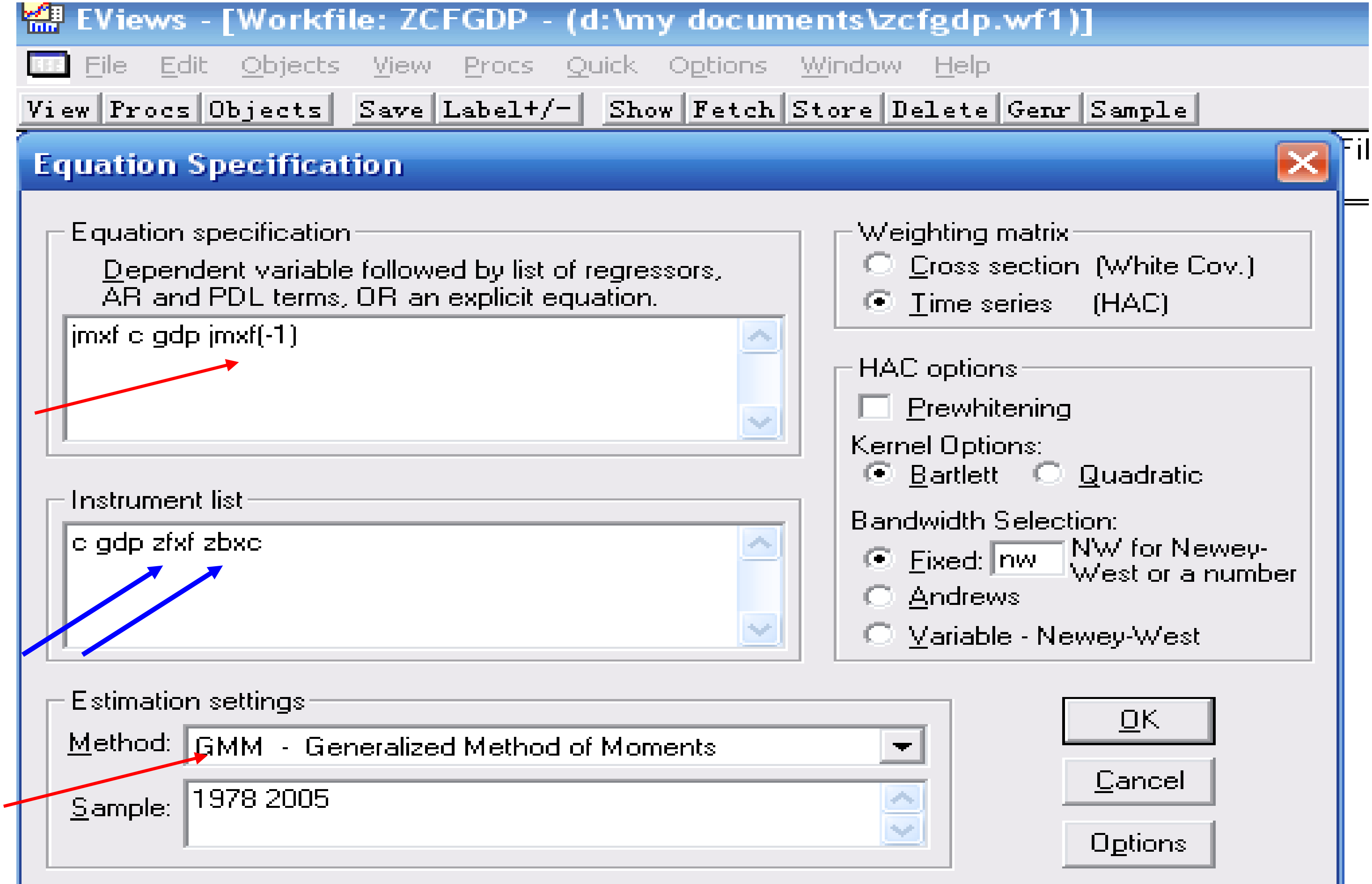
View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: JMXF
Method: Two-Stage Least Squares
~~Date: 02/14/07 Time: 11:51~~
Sample(adjusted): 1979 2005
Included observations: 27 after adjusting endpoints
Instrument list: C GDP ZFXF

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1059.997	334.2714	3.171066	0.0041
GDP	0.158449	0.048716	3.252512	0.0034
JMXF(-1)	0.665581	0.130407	5.103864	0.0000

R-squared	0.997613	Mean dependent var	23729.28
Adjusted R-squared	0.997414	S.D. dependent var	21866.25
S.E. of regression	1111.941	Sum squared resid	29673919
F-statistic	4992.187	Durbin-Watson stat	0.458540
Prob(F-statistic)	0.000000		

GMM估计


EViews - [Workfile: ZCFGDP - (d:\my documents\zcfgdp.wf1)]
File Edit Objects View Procs Quick Options Window Help
View Procs Objects Save Label+/- Show Fetch Store Delete Genr Sample
Equation Specification
Equation specification
Dependent variable followed by list of regressors, AR and PDL terms, OR an explicit equation.
jmxf c gdp jmxf(-1)
Instrument list
c gdp zfxf zbxc
Estimation settings
Method: GMM - Generalized Method of Moments
Sample: 1978 2005
Weighting matrix
Cross section (White Cov.)
Time series (HAC)
HAC options
Prewhitening
Kernel Options:
Bartlett
Quadratic
Bandwidth Selection:
Fixed: nw
NW for Newey-West or a number
Andrews
Variable - Newey-West
OK
Cancel
Options

EViews - [Equation: UNTITLED Workfile: ZCFGDP]

File Edit Objects View Procs Quick Options Window Help

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

Dependent Variable: JMXF
Method: Generalized Method of Moments
Date: 02/14/07 Time: 11:55
Sample(adjusted): 1979 2005
Included observations: 27 after adjusting endpoints
Kernel: Bartlett, Bandwidth: Fixed (2), No prewhitening
Simultaneous weighting matrix & coefficient iteration
Convergence achieved after: 5 weight matrices, 6 total coef iterations
Instrument list: C GDP ZFXF ZBXC

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	738.8737	309.5051	2.387275	0.0252
GDP	0.094350	0.017496	5.392559	0.0000
JMXF(-1)	0.835320	0.047788	17.47955	0.0000

R-squared	0.997397	Mean dependent var	23729.28
Adjusted R-squared	0.997180	S.D. dependent var	21866.25
S.E. of regression	1161.080	Sum squared resid	32354562
Durbin-Watson stat	0.525981	J-statistic	0.034450

估计结果

- OLS:

JMXF = 1001.164757 + 0.1367699684*GDP + 0.7238178139*JMXF(-1)

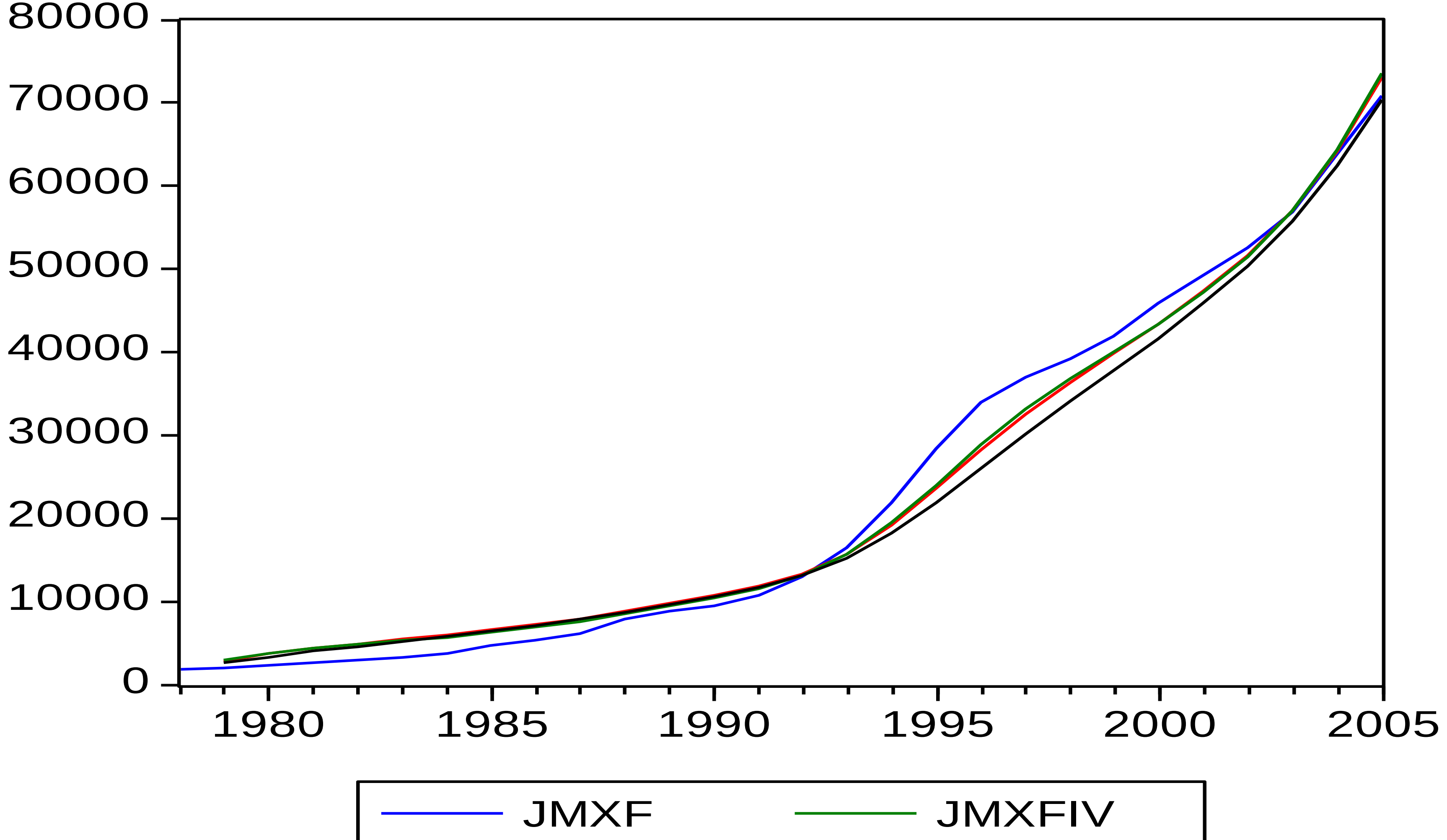
- IV:

JMXF = 1059.996753 + 0.1584492759*GDP + 0.6655810226*JMXF(-1)

- GMM:

JMXF = 738.873724 + 0.09435006141*GDP + 0.8353195568*JMXF(-1)

拟合结果


80000
70000
60000
50000
40000
30000
20000
10000
0
1980
1985
1990
1995
2000
2005
JMXF
JMXFIV
JMXFOLS
JMXFGMM

四、解释变量的内生性检验***

Hausman两步法检验

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_{i1} + \mu_i$$

Z_1外生，与μ不相关

选择Z_2作为X 的工具变量

Step1，辅助回归： $X_i = \alpha_0 + \alpha_1 Z_{i1} + \alpha_2 Z_{i2} + v_i$

$$Step\ 2: \quad Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_{i1} + \delta \cdot \hat{v}_i + \varepsilon_i$$

- **如果δ显著为0：υ与Y同期无关→υ与μ同期无关→ X与μ同期无关→X是同期外生变量；**
- **如果δ显著不为0**：υ与Y同期相关 → υ与μ同期相关→ X与μ同期相关→ **X是同期内生变量**。

五、例：中国城镇居民人均消费函数（自学）

- 步骤
  - 以中国城镇居民人均消费为被解释变量，人均可支配收入和前一年城镇居人均消费支出为解释变量，建立模型；
  - 经分析认为，人均可支配收入可能具有同期内生性；
  - 选择工具变量；
  - 采用 Hausman检验 判断，城镇居民人均可支配收入确实是内生变量；
  - 采用工具变量估计。
  - 为了比较，采用OLS估计。