11 Consistency and Asymptotic Normality of OLSE

Regression model: $y = X\beta + u, \qquad u \sim (0, \sigma^2 I_n).$

Consistency:

1. Let $\hat{\beta}_n = (X'X)^{-1}X'y$ be the OLS with sample size n.

   Consistency: As n is large, $\hat{\beta}_n$ converges to β.

2. Assume the stationarity condition for X, i.e.,

$$\frac{1}{n}X'X \longrightarrow M_{xx}.$$

   and no correlation between X and u, i.e.,

$$\frac{1}{n}X'u \longrightarrow 0.$$

3. Note that $\frac{1}{n}X'X \longrightarrow M_{xx}$ results in $(\frac{1}{n}X'X)^{-1} \longrightarrow M_{xx}^{-1}$.

$\Longrightarrow$ Slutsky's Theorem

(*) **Slutsky's Theorem** $\quad g(\hat{\theta}) \longrightarrow g(\theta)$, when $\hat{\theta} \longrightarrow \theta$.

4. OLS is given by:

$$\hat{\beta}_n = \beta + (X'X)^{-1}X'u = \beta + (\frac{1}{n}X'X)^{-1}(\frac{1}{n}X'u).$$

Therefore,

$$\hat{\beta}_n \longrightarrow \beta + M_{xx}^{-1} \times 0 = \beta$$

Thus, OLSE is a consistent estimator.

Asymptotic Normality:

1. Asymptotic Normality of OLSE

$$\sqrt{n}(\hat{\beta}_n - \beta) \longrightarrow N(0.\sigma^2 M_{xx}^{-1}), \quad \text{when } n \longrightarrow \infty.$$

2. **Central Limit Theorem:** Greenberg and Webster (1983)

Z_1, Z_2, $\cdots$, Z_n are mutually independent. Z_i is distributed with mean μ and variance Σ_i for $i = 1, 2, \cdots, n$.

Then, we have the following result:

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Z_i - \mu) \longrightarrow N(0, \Sigma),$$

where

$$\Sigma = \lim_{n\to\infty} \left(\frac{1}{n} \sum_{i=1}^{n} \Sigma_i \right).$$

Note that the distribution of Z_i is not assumed.

3. Define $Z_i = x_i' u_i$. Then, $\Sigma_i = \mathrm{V}(Z_i) = \sigma^2 x_i' x_i$.

4. Σ is defined as:

$$\Sigma = \lim_{n\to\infty} \left(\frac{1}{n} \sum_{i=1}^{n} \sigma^2 x_i' x_i \right) = \sigma^2 \lim_{n\to\infty} \left(\frac{1}{n} X'X \right) = \sigma^2 M_{xx},$$

where

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

5. Applying Central Limit Theorem (Greenberg and Webster (1983), we obtain the following:

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i' u_i = \frac{1}{\sqrt{n}} X'u \longrightarrow N(0, \sigma^2 M_{xx}).$$

On the other hand, from $\hat{\beta}_n = \beta + (X'X)^{-1}X'u$, we can rewrite as:

$$\sqrt{n}(\hat{\beta} - \beta) = \Big(\frac{1}{n}X'X\Big)^{-1}\frac{1}{\sqrt{n}}X'u.$$

$$\begin{aligned}
\mathrm{V}\bigg(\Big(\frac{1}{n}X'X\Big)^{-1}\frac{1}{\sqrt{n}}X'u\bigg) &= \mathrm{E}\bigg(\Big(\frac{1}{n}X'X\Big)^{-1}\frac{1}{\sqrt{n}}X'u\Big(\Big(\frac{1}{n}X'X\Big)^{-1}\frac{1}{\sqrt{n}}X'u\Big)'\bigg) \\
&= \Big(\frac{1}{n}X'X\Big)^{-1}\Big(\frac{1}{n}X'\mathrm{E}(uu')X\Big)\Big(\frac{1}{n}X'X\Big)^{-1} \\
&= \sigma^2\Big(\frac{1}{n}X'X\Big)^{-1}\Big(\frac{1}{n}X'X\Big)\Big(\frac{1}{n}X'X\Big)^{-1} \\
&\longrightarrow \ \sigma^2 M_{xx}^{-1}M_{xx}M_{xx}^{-1} = \sigma^2 M_{xx}^{-1}.
\end{aligned}$$

Therefore,

$$\sqrt{n}(\hat{\beta} - \beta) \ \longrightarrow \ N(0, \sigma^2 M_{xx}^{-1})$$

$\Longrightarrow$ Asymptotic normality (漸近的正規性) of OLSE

The distribution of u_i is not assumed.

12 Instrumental Variable (操作変数法)

12.1 Measurement Error (測定誤差)

Errors in Variables

1. True regression model:

$$y = \tilde{X}\beta + u$$

2. Observed variable:

$$X = \tilde{X} + V$$

   V: is called the **measurement error (測定誤差 or 観測誤差)**.

3. For the elements which do not include measurement errors in X, the corresponding elements in V are zeros.

4. Regression using observed variable:

$$y = X\beta + (u - V\beta)$$

OLS of β is:

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

5. Assumptions:

(a) The measurement error in X is uncorrelated with $\tilde{X}$ in the limit. i.e.,

$$\text{plim}\Big(\frac{1}{n}\tilde{X}'V\Big) = 0.$$

Therefore, we obtain the following:

$$\text{plim}\Big(\frac{1}{n}X'X\Big) = \text{plim}\Big(\frac{1}{n}\tilde{X}'\tilde{X}\Big) + \text{plim}\Big(\frac{1}{n}V'V\Big) = \Sigma + \Omega$$

(b) u is not correlated with V.

u is not correlated with $\tilde{X}$.

That is,

$$\text{plim}\Big(\frac{1}{n}V'u\Big) = 0, \qquad \text{plim}\Big(\frac{1}{n}\tilde{X}'u\Big) = 0.$$

6. OLSE of β is:

$$\hat{\beta} = \beta + (X'X)^{-1}X'(u - V\beta) = \beta + (X'X)^{-1}(\tilde{X} + V)'(u - V\beta).$$

Therefore, we obtain the following:

$$\text{plim}\,\hat{\beta} = \beta - (\Sigma + \Omega)^{-1}\Omega\beta$$

7. **Example: The Case of Two Variables:**

The regression model is given by:

$$y_t = \alpha + \beta\tilde{x}_t + u_t, \qquad x_t = \tilde{x}_t + v_t.$$

Under the above model,

$$\Sigma = \operatorname{plim}\Big(\frac{1}{n}\tilde{X}'\tilde{X}\Big) = \operatorname{plim}\begin{pmatrix} 1 & \frac{1}{n}\sum \tilde{x}_i \\ \frac{1}{n}\sum \tilde{x}_i & \frac{1}{n}\sum \tilde{x}_i^2 \end{pmatrix} = \begin{pmatrix} 1 & \mu \\ \mu & \mu^2 + \sigma^2 \end{pmatrix},$$

where μ and σ^2 represent the mean and variance of $\tilde{x}_i$.

$$\Omega = \operatorname{plim}\Big(\frac{1}{n}V'V\Big) = \operatorname{plim}\begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{n}\sum v_i^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_v^2 \end{pmatrix}.$$

Therefore,

$$\begin{aligned}
\operatorname{plim}\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - \left(\begin{pmatrix} 1 & \mu \\ \mu & \mu^2 + \sigma^2 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \sigma_v^2 \end{pmatrix}\right)^{-1} \begin{pmatrix} 0 & 0 \\ 0 & \sigma_v^2 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\
&= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - \frac{1}{\sigma^2 + \sigma_v^2}\begin{pmatrix} -\mu\sigma_v^2\beta \\ \sigma_v^2\beta \end{pmatrix}
\end{aligned}$$

Now we focus on β.

$\hat{\beta}$ is not consistent. because of:

$$\text{plim}(\hat{\beta}) = \beta - \frac{\sigma_v^2 \beta}{\sigma^2 + \sigma_v^2} = \frac{\beta}{1 + \sigma_v^2/\sigma^2} < \beta$$

12.2 Instrumental Variable (IV) Method (操作変数法 or IV 法)

Instrumental Variable (IV)

1. Consider the regression model: $y = X\beta + u$ and $u \sim N(0, \sigma^2 I_n)$.

   In the case of $\mathrm{E}(X'u) \neq 0$, OLSE of β is inconsistent.

2. **Proof:**

$$\hat{\beta} = \beta + (\frac{1}{n}X'X)^{-1}\frac{1}{n}X'u \longrightarrow \beta + M_{xx}^{-1}M_{xu},$$

   where

$$\frac{1}{n}X'X \longrightarrow M_{xx}, \qquad \frac{1}{n}X'u \longrightarrow M_{xu} \neq 0$$

3. Find the Z which satisfies $\frac{1}{n}Z'u \longrightarrow M_{zu} = 0$.

Multiplying Z' on both sides of the regression model: $y = X\beta + u$,

$$Z'y = Z'X\beta + Z'u$$

Dividing n on both sides of the above equation, we take plim on both sides.

Then, we obtain the following:

$$\operatorname{plim}\left(\frac{1}{n}Z'y\right) = \operatorname{plim}\left(\frac{1}{n}Z'X\right)\beta + \operatorname{plim}\left(\frac{1}{n}Z'u\right) = \operatorname{plim}\left(\frac{1}{n}Z'X\right)\beta.$$

Accordingly, we obtain:

$$\beta = \left(\operatorname{plim}\left(\frac{1}{n}Z'X\right)\right)^{-1} \operatorname{plim}\left(\frac{1}{n}Z'y\right).$$

Therefore, we consider the following estimator:

$$\beta_{IV} = (Z'X)^{-1}Z'y,$$

which is taken as an estimator of β.

$\Longrightarrow$ **Instrumental Variable Method (操作変数法 or IV 法)**

4. Assume the followings:

$$\frac{1}{n}Z'X \longrightarrow M_{zx}, \qquad \frac{1}{n}Z'Z \longrightarrow M_{zz}, \qquad \frac{1}{n}Z'u \longrightarrow 0$$

5. **Asymptotic Distribution of β_{IV}:**

$$\beta_{IV} = (Z'X)^{-1}Z'y = (Z'X)^{-1}Z'(X\beta + u) = \beta + (Z'X)^{-1}Z'u,$$

which is rewritten as:

$$\sqrt{n}(\beta_{IV} - \beta) = \Big(\frac{1}{n}Z'X\Big)^{-1}\Big(\frac{1}{\sqrt{n}}Z'u\Big)$$

Applying the Central Limit Theorem to $\Big(\frac{1}{\sqrt{n}}Z'u\Big)$, we have the following result:

$$\frac{1}{\sqrt{n}}Z'u \longrightarrow N(0, \sigma^2 M_{zz}).$$

Therefore,

$$\sqrt{n}(\beta_{IV} - \beta) = \Big(\frac{1}{n}Z'X\Big)^{-1}\Big(\frac{1}{\sqrt{n}}Z'u\Big) \longrightarrow N(0, \sigma^2 M_{zx}^{-1} M_{zz} M_{zx}'^{-1})$$

$\Longrightarrow$ Consistency and Asymptotic Normality

6. The variance of β_{IV} is given by:

$$\mathrm{V}(\beta_{IV}) = s^2 (Z'X)^{-1} Z'Z (X'Z)^{-1},$$

where

$$s^2 = \frac{(y - X\beta_{IV})'(y - X\beta_{IV})}{n-k}.$$

12.3 Two-Stage Least Squares Method (2 段階最小二乗法, 2SLS or TSLS)

1. Regression Model:

$$y = X\beta + u, \qquad u \sim N(0, \sigma^2 I),$$

   In the case of $\mathrm{E}(X'u) \neq 0$, OLSE is not consistent.

2. Find the variable Z which satisfies $\dfrac{1}{n}Z'u \longrightarrow M_{zu} = 0$.

3. Use $Z = \hat{X}$ for the instrumental variable.

   $\hat{X}$ is the predicted value which regresses X on the other exogenous variables, say W.

   That is, consider the following regression model:

$$X = WB + V.$$

Estimate B by OLS.

Then, we obtain the prediction:

$$\hat{X} = W\hat{B},$$

where $\hat{B} = (W'W)^{-1}W'X$.

Or, equivalently,

$$\hat{X} = W(W'W)^{-1}W'X.$$

$\hat{X}$ is used for the instrumental variable of X.

4. The IV method is rewritten as:

$$\beta_{IV} = (\hat{X}'X)^{-1}\hat{X}'y = (X'W(W'W)^{-1}W'X)^{-1}X'W(W'W)^{-1}W'y.$$

Furthermore, β_{IV} is written as follows:

$$\beta_{IV} = \beta + (X'W(W'W)^{-1}W'X)^{-1}X'W(W'W)^{-1}W'u.$$

Therefore, we obtain the following expression:

$$\sqrt{n}(\beta_{IV}-\beta) = \left(\left(\frac{1}{n}X'W\right)\left(\frac{1}{n}W'W\right)^{-1}\left(\frac{1}{n}XW'\right)'\right)^{-1}\left(\frac{1}{n}X'W\right)\left(\frac{1}{n}W'W\right)^{-1}\left(\frac{1}{\sqrt{n}}W'u\right)$$
$$\longrightarrow N\Big(0, \sigma^2(M_{xw}M_{ww}^{-1}M'_{xw})^{-1}\Big).$$

5. Clearly, there is no correlation between W and u at least in the limit, i.e.,

$$\mathrm{plim}\left(\frac{1}{n}W'u\right) = 0.$$

6. **Remark:**

$$\hat{X}'X = X'W(W'W)^{-1}W'X = X'W(W'W)^{-1}W'W(W'W)^{-1}W'X = \hat{X}'\hat{X}.$$

Therefore,

$$\beta_{IV} = (\hat{X}'X)^{-1}\hat{X}'y = (\hat{X}'\hat{X})^{-1}\hat{X}'y,$$

which implies the OLS estimator of β in the regression model: $y = \hat{X}\beta + u$ and $u \sim N(0, \sigma^2 I_n)$.

Example:

$$y_t = \alpha x_t + \beta z_t + u_t, \qquad u_t \sim (0, \sigma^2).$$

Suppose that x_t is correlated with u_t but z_t is not correlated with u_t.

- 1st Step:

Estimate the following regression model:

$$x_t = \gamma w_t + \delta z_t + \cdots + v_t,$$

by OLS. $\Longrightarrow$ Obtain $\hat{x}_t$ through OLS.

- 2nd Step:

Estimate the following regression model:

$$y_t = \alpha \hat{x}_t + \beta z_t + u_t,$$

by OLS. $\Longrightarrow$ α_{iv} and β_{iv}

Note as follows. Estimate the following regression model:

$$z_t = \gamma_2 w_t + \delta_2 z_t + \cdots + v_{2t},$$

by OLS.

$\Longrightarrow$ $\hat{\gamma}_2 = 0$, $\hat{\delta}_2 = 1$, and the other coefficient estimates are zeros. i.e., $\hat{z}_t = z_t$.

Eviews Command:

```
tsls y x z @ w z ...
```

13 Large Sample Tests

13.1 Wald, LM and LR Tests

Parameter $\theta : k \times 1$, $\qquad h(\theta) : G \times 1$ vector function, $G \leq k$

The null hypothesis $H_0 : h(\theta) = 0 \implies G$ restrictions

$\tilde{\theta} : k \times 1$, restricted maximum likelihood estimate

$\hat{\theta} : k \times 1$, unrestricted maximum likelihood estimate

$I(\theta) : k \times k$, information matrix, i.e., $\quad I(\theta) = -\mathrm{E}\Big(\dfrac{\partial^2 \log L(\theta)}{\partial\theta\partial\theta'}\Big).$

$\log L(\theta)$: log-likelihood function

$$R_\theta = \frac{\partial h(\theta)}{\partial\theta'} : G \times k, \qquad F_\theta = \frac{\partial \log L(\theta)}{\partial\theta} : k \times 1$$

1. **Wald Test (ワルド検定):** $\quad W = h(\hat{\theta})'\Big(R_{\hat{\theta}}(I(\hat{\theta}))^{-1}R'_{\hat{\theta}}\Big)^{-1}h(\hat{\theta})$

   (a) $h(\hat{\theta}) \approx h(\theta) + \dfrac{\partial h(\theta)}{\partial\theta'}(\hat{\theta} - \theta) \quad \Longleftarrow \quad h(\hat{\theta})$ is linearized around $\hat{\theta} = \theta$.

Under the null hypothesis $h(\theta) = 0$,

$$h(\hat{\theta}) \approx \frac{\partial h(\theta)}{\partial \theta'}(\hat{\theta} - \theta) = R_\theta(\hat{\theta} - \theta)$$

(b) $\hat{\theta}$ is MLE.

From the properties of MLE,

$$\sqrt{n}(\hat{\theta} - \theta) \longrightarrow N\Big(0, \lim_{n\to\infty}\Big(\frac{I(\theta)}{n}\Big)^{-1}\Big),$$

That is, approximately, we have the following result:

$$\hat{\theta} - \theta \sim N\Big(0, (I(\theta))^{-1}\Big).$$

(c) The distribution of $h(\hat{\theta})$ is approximately given by:

$$h(\hat{\theta}) \sim N\Big(0, R_\theta(I(\theta))^{-1}R'_\theta\Big)$$

(d) Therefore, the $\chi^2(G)$ distribution is derived as follows:

$$h(\hat\theta)\Big(R_\theta (I(\theta))^{-1} R'_\theta\Big)^{-1} h(\hat\theta)' \longrightarrow \chi^2(G).$$

Furthermore, from the fact that $R_{\hat\theta} \longrightarrow R_\theta$ and $I(\hat\theta) \longrightarrow I(\theta)$ as $n \longrightarrow \infty$ (i.e., convergence in probability, 確率収束), we can replace θ by $\hat\theta$ as follows:

$$h(\hat\theta)\Big(R_{\hat\theta} (I(\hat\theta))^{-1} R'_{\hat\theta}\Big)^{-1} h(\hat\theta)' \longrightarrow \chi^2(G).$$

2. **Lagrange Multiplier Test (ラグランジェ乗数検定):** $LM = F'_{\tilde\theta}(I(\tilde\theta))^{-1} F_{\tilde\theta}$

(a) MLE with the constraint $h(\theta) = 0$:

$$\max_\theta \ \log L(\theta), \quad \text{subject to} \quad h(\theta) = 0$$

The Lagrangian function is: $L = \log L(\theta) + \lambda h(\theta)$.

(b) For maximization, we have the following two equations:

$$\frac{\partial L}{\partial \theta} = \frac{\partial \log L(\theta)}{\partial \theta} + \lambda \frac{\partial h(\theta)}{\partial \theta} = 0, \qquad \frac{\partial L}{\partial \lambda} = h(\theta) = 0.$$

The restricted MLE $\tilde{\theta}$ satisfies $h(\tilde{\theta}) = 0$.

(c) Mean and variance of $\dfrac{\partial \log L(\theta)}{\partial \theta}$ are given by:

$$\mathrm{E}\Big(\frac{\partial \log L(\theta)}{\partial \theta}\Big) = 0, \qquad \mathrm{V}\Big(\frac{\partial \log L(\theta)}{\partial \theta}\Big) = -\mathrm{E}\Big(\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'}\Big) = I(\theta).$$

(d) Therefore, using the central limit theorem,

$$\frac{1}{\sqrt{n}} \frac{\partial \log L(\theta)}{\partial \theta} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \log f(X_i;\theta)}{\partial \theta} \longrightarrow N\Big(0, \lim_{n\to\infty}\Big(\frac{1}{n} I(\theta)\Big)\Big)$$

(e) Therefore, $\dfrac{\partial \log L(\theta)}{\partial \theta} (I(\theta))^{-1} \dfrac{\partial \log L(\theta)}{\partial \theta'} \longrightarrow \chi^2(G)$.

Under $H_0: \ h(\theta) = 0$, replacing θ by $\tilde{\theta}$ we have the result:

$$F'_{\tilde{\theta}} (I(\tilde{\theta}))^{-1} F_{\tilde{\theta}} \longrightarrow \chi^2(G).$$

3. **Likelihood Ratio Test (尤度比検定):** $LR = -2\log\lambda \longrightarrow \chi^2(G)$

$$\lambda = \frac{L(\tilde{\theta})}{L(\hat{\theta})}$$

(a) By Taylor series expansion evaluated at $\theta = \hat{\theta}$, $\log L(\theta)$ is given by:

$$\begin{aligned}\log L(\theta) &= \log L(\hat{\theta}) + \frac{\partial \log L(\hat{\theta})}{\partial \theta}(\theta - \hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})' \frac{\partial^2 \log L(\hat{\theta})}{\partial \theta \partial \theta'}(\theta - \hat{\theta}) + \cdots \\ &= \log L(\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})' \frac{\partial^2 \log L(\hat{\theta})}{\partial \theta \partial \theta'}(\theta - \hat{\theta}) + \cdots\end{aligned}$$

Note that $\dfrac{\partial \log L(\hat{\theta})}{\partial \theta} = 0$ because $\hat{\theta}$ is MLE.

$$\begin{aligned}-2(\log L(\theta) - \log L(\hat{\theta})) &\approx -(\theta - \hat{\theta})'\Big(\frac{\partial^2 \log L(\hat{\theta})}{\partial \theta \partial \theta'}\Big)(\theta - \hat{\theta}) \\ &= \sqrt{n}(\hat{\theta} - \theta)'\Big(-\frac{1}{n}\frac{\partial^2 \log L(\hat{\theta})}{\partial \theta \partial \theta'}\Big)\sqrt{n}(\hat{\theta} - \theta) \\ &\longrightarrow \chi^2(G)\end{aligned}$$

Note:

(1) $\hat{\theta} \longrightarrow \theta$,

(2) $-\dfrac{1}{n}\dfrac{\partial^2 \log L(\hat{\theta})}{\partial\theta\partial\theta'} \longrightarrow -\lim_{n\to\infty}\Big(\dfrac{1}{n}\mathrm{E}\Big(\dfrac{\partial^2 \log L(\hat{\theta})}{\partial\theta\partial\theta'}\Big)\Big) = \lim_{n\to\infty}\Big(\dfrac{1}{n}I(\theta)\Big)$,

(3) $\sqrt{n}(\hat{\theta}-\theta) \longrightarrow N\Big(0, \lim_{n\to\infty}\Big(\dfrac{1}{n}I(\theta)\Big)\Big)$.

(b) Under $H_0: \ h(\theta) = 0$,

$$-2(\log L(\tilde{\theta}) - \log L(\hat{\theta})) \longrightarrow \chi^2(G).$$

Remember that $h(\tilde{\theta}) = 0$ is always satisfied.

For proof, see Theil (1971, p.396).

4. All of W, LM and LR are asymptotically distributed as $\chi^2(G)$ random variables under the null hypothesis $H_0 : h(\theta) = 0$.

5. Under some conditions, we have $W \geq LR \geq LM$. See Engle (1981) "Wald, Likelihood and Lagrange Multiplier Tests in Econometrics," Chap. 13 in *Handbook of Econometrics*, Vol.2, Grilliches and Intriligator eds, North-Holland.

13.2 Example: W, LM and LR Tests

Date file $\Longrightarrow$ cons99.txt (same data as before)

Each column denotes year, nominal household expenditures (家計消費，10 billion yen), household disposable income (家計可処分所得，10 billion yen) and household expenditure deflator (家計消費デフレータ，1990=100) from the left.

1955	5430.1	6135.0	18.1
1956	5974.2	6828.4	18.3
1957	6686.3	7619.5	19.0
1958	7169.7	8153.3	19.1
1959	8019.3	9274.3	19.7
1960	9234.9	10776.5	20.5
1961	10836.2	12869.4	21.8
1962	12430.8	14701.4	23.2
1963	14506.6	17042.7	24.9
1964	16674.9	19709.9	26.0
1965	18820.5	22337.4	27.8
1966	21680.6	25514.5	29.0
1967	24914.0	29012.6	30.1
1968	28452.7	34233.6	31.6
1969	32705.2	39486.3	32.9
1970	37784.1	45913.2	35.2
1971	42571.6	51944.3	37.5
1972	49124.1	60245.4	39.7
1973	59366.1	74924.8	44.1
1974	71782.1	93833.2	53.3
1975	83591.1	108712.8	59.4
1976	94443.7	123540.9	65.2
1977	105397.8	135318.4	70.1
1978	115960.3	147244.2	73.5
1979	127600.9	157071.1	76.0
1980	138585.0	169931.5	81.6
1981	147103.4	181349.2	85.4
1982	157994.0	190611.5	87.7
1983	166631.6	199587.8	89.5
1984	175383.4	209451.9	91.8
1985	185335.1	220655.6	93.9
1986	193069.6	229938.8	94.8
1987	202072.8	235924.0	95.3
1988	212939.9	247159.7	95.8
1989	227122.2	263940.5	97.7
1990	243035.7	280133.0	100.0
1991	255531.8	297512.9	102.5
1992	265701.6	309256.6	104.5
1993	272075.3	317021.6	105.9
1994	279538.7	325655.7	106.7
1995	283245.4	331967.5	106.2
1996	291458.5	340619.1	106.0
1997	298475.2	345522.7	107.3

```
          PROGRAM
LINE  ***********************************************
|          freq a;
|          smpl 1955 1997;
|          read(file='cons99.txt') year cons yd price;
|          rcons=cons/(price/100);
|          ryd=yd/(price/100);
|          lyd=log(ryd);
|          olsq rcons c ryd;
|          olsq @res @res(-1);
|          ar1  rcons c ryd;
|          olsq rcons c lyd;
|          param a1 0 a2 0 a3 1;
|          frml eq rcons=a1+a2*((ryd**a3)-1.)/a3;
|          lsq(tol=0.00001,maxit=100) eq;
|          a3=1.15;
|          rryd=((ryd**a3)-1.)/a3;
|          ar1  rcons c rryd;
|          end;
*****************************************************
```

Equation 1
============

Method of estimation = Ordinary Least Squares

Dependent variable: RCONS
Current sample: 1955 to 1997
Number of observations: 43

Mean of dep. var.	= 146270.	LM het. test	= .207443 [.649]
Std. dev. of dep. var.	= 79317.2	Durbin-Watson	= .115101 [.000,.000]
Sum of squared residuals	= .129697E+10	Jarque-Bera test	= 9.47539 [.009]
Variance of residuals	= .316335E+08	Ramsey's RESET2	= 53.6424 [.000]
Std. error of regression	= 5624.36	F (zero slopes)	= 8311.90 [.000]
R-squared	= .995092	Schwarz B.I.C.	= 435.051
Adjusted R-squared	= .994972	Log likelihood	= -431.289

Variable	Estimated Coefficient	Standard Error	t-statistic	P-value
C	-2919.54	1847.55	-1.58022	[.122]
RYD	.852879	.935486E-02	91.1696	[.000]

```
                                   Equation   2
                                   ============
                    Method of estimation = Ordinary Least Squares

Dependent variable: @RES
Current sample:  1956 to 1997
Number of observations:  42

       Mean of dep. var. = -95.5174
  Std. dev. of dep. var. = 5588.52
Sum of squared residuals = .146231E+09
   Variance of residuals = .356662E+07
Std. error of regression = 1888.55
               R-squared = .885884
      Adjusted R-squared = .885884
            LM het. test = .760256 [.383]
           Durbin-Watson = 1.40409 [.023,.023]
              Durbin's h = 1.97732 [.048]
         Durbin's h alt. = 1.91077 [.056]
        Jarque-Bera test = 6.49360 [.039]
         Ramsey's RESET2 = .186107 [.668]
          Schwarz B.I.C. = 377.788
          Log likelihood = -375.919

           Estimated    Standard
Variable   Coefficient    Error       t-statistic   P-value
@RES(-1)   .950693      .053301      17.8362       [.000]
```

Equation 3
============

FIRST-ORDER SERIAL CORRELATION OF THE ERROR
Objective function: Exact ML (keep first obs.)

Dependent variable: RCONS
Current sample: 1955 to 1997
Number of observations: 43

Mean of dep. var.	= 146270.	R-squared	= .999480
Std. dev. of dep. var.	= 79317.2	Adjusted R-squared	= .999454
Sum of squared residuals	= .145826E+09	Durbin-Watson	= 1.38714
Variance of residuals	= .364564E+07	Schwarz B.I.C.	= 391.061
Std. error of regression	= 1909.36	Log likelihood	= -385.419

Parameter	Estimate	Standard Error	t-statistic	P-value
C	1672.42	6587.40	.253881	[.800]
RYD	.840011	.027182	30.9032	[.000]
RHO	.945025	.045843	20.6143	[.000]

Equation 4
============

Method of estimation = Ordinary Least Squares

Dependent variable: RCONS
Current sample: 1955 to 1997
Number of observations: 43

Mean of dep. var.	= 146270.	LM het. test	= 2.21031 [.137]
Std. dev. of dep. var.	= 79317.2	Durbin-Watson	= .029725 [.000,.000]
Sum of squared residuals	= .256040E+11	Jarque-Bera test	= 3.72023 [.156]
Variance of residuals	= .624487E+09	Ramsey's RESET2	= 344.855 [.000]
Std. error of regression	= 24989.7	F (zero slopes)	= 382.117 [.000]
R-squared	= .903100	Schwarz B.I.C.	= 499.179
Adjusted R-squared	= .900737	Log likelihood	= -495.418

Variable	Estimated Coefficient	Standard Error	t-statistic	P-value
C	-.115228E+07	66538.5	-17.3175	[.000]
LYD	109305.	5591.69	19.5478	[.000]

```
                           NONLINEAR LEAST SQUARES
                           =======================
CONVERGENCE ACHIEVED AFTER  84 ITERATIONS

Number of observations = 43       Log likelihood = -414.362
        Schwarz B.I.C. = 420.004

                             Standard
Parameter  Estimate            Error       t-statistic   P-value
A1         16544.5         2615.60         6.32530       [.000]
A2         .063304         .024133         2.62307       [.009]
A3         1.21694         .031705         38.3839       [.000]

Standard Errors computed from quadratic form of analytic first derivatives
(Gauss)

Equation: EQ
Dependent variable: RCONS

       Mean of dep. var. = 146270.
  Std. dev. of dep. var. = 79317.2
Sum of squared residuals = .590213E+09
   Variance of residuals = .147553E+08
Std. error of regression = 3841.27
               R-squared = .997766
      Adjusted R-squared = .997655
            LM het. test = .174943 [.676]
           Durbin-Watson = .253234 [.000,.000]
```

```
                              Equation   5
                              ============
                  FIRST-ORDER SERIAL CORRELATION OF THE ERROR
                  Objective function:  Exact ML (keep first obs.)

Dependent variable: RCONS
Current sample:  1955 to 1997
Number of observations:  43

       Mean of dep. var. = 146270.              R-squared = .999470
  Std. dev. of dep. var. = 79317.2     Adjusted R-squared = .999443
Sum of squared residuals = .140391E+09      Durbin-Watson = 1.43657
   Variance of residuals = .350977E+07     Schwarz B.I.C. = 389.449
Std. error of regression = 1873.44         Log likelihood = -383.807

                          Standard
Parameter  Estimate        Error        t-statistic   P-value
C          12034.8        3346.47       3.59628       [.000]
RRYD       .140723        .282614E-02   49.7933       [.000]
RHO        .876924        .068199       12.8583       [.000]
```

1. `Equation 1` vs. `Equation 3` (Test of Serial Correlation)

`Equation 1` is:

$$\texttt{RCONS}_t = \beta_1 + \beta_2 \texttt{RYD}_t + u_t, \qquad \epsilon_t \sim \text{ iid } N(0, \sigma_\epsilon^2)$$

`Equation 3` is:

$$\texttt{RCONS}_t = \beta_1 + \beta_2 \texttt{RYD}_t + u_t, \qquad u_t = \rho u_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{ iid } N(0, \sigma_\epsilon^2)$$

The null hypothesis is $H_0: \ \rho = 0$

Restricted MLE $\Longrightarrow$ `Equation 1`

Unrestricted MLE $\Longrightarrow$ `Equation 3`

The log-likelihood function of `Equation 3` is:

$$
\begin{aligned}
\log L(\beta, \sigma_\epsilon^2, \rho) = & -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma_\epsilon^2) + \frac{1}{2}\log(1-\rho^2) \\
& -\frac{1}{2\sigma_\epsilon^2}\sum_{t=1}^{n}(\mathtt{RCONS}_t^* - \beta_1\mathtt{CONST}_t^* - \beta_2\mathtt{RYD}_t^*)^2,
\end{aligned}
$$

where

$$
\mathtt{RCONS}_t^* = \begin{cases} \sqrt{1-\rho^2}\,\mathtt{RCONS}_t, & \text{for } t = 1, \\ \mathtt{RCONS}_t - \rho\mathtt{RCONS}_{t-1}, & \text{for } t = 2, 3, \cdots, n, \end{cases}
$$

$$
\mathtt{CONST}_t^* = \begin{cases} \sqrt{1-\rho^2}, & \text{for } t = 1, \\ 1-\rho, & \text{for } t = 2, 3, \cdots, n, \end{cases}
$$

$$\texttt{RYD}^*_t = \begin{cases} \sqrt{1-\rho^2}\texttt{RYD}_t, & \text{for } t = 1, \\ \texttt{RYD}_t - \rho\texttt{RYD}_{t-1}, & \text{for } t = 2, 3, \cdots, n. \end{cases}$$

- MLE with the restriction $\rho = 0$ (`Equation 1`) solves:

$$\max_{\beta,\sigma_\epsilon^2} \log L(\beta, \sigma_\epsilon^2, 0)$$

Restricted MLE $\Longrightarrow \tilde{\beta}, \tilde{\sigma}_\epsilon^2$

`Log of likelihood function = -431.289`

- MLE without the restriction $\rho = 0$ (`Equation 3`) solves:

$$\max_{\beta,\sigma_\epsilon^2,\rho} \log L(\beta, \sigma_\epsilon^2, \rho)$$

Unrestricted MLE $\Longrightarrow \hat{\beta}, \hat{\sigma}_\epsilon^2, \hat{\rho}$

`Log of likelihood function = -385.419`

The likelihood ratio test statistic is:

$$
\begin{aligned}
-2\log(\lambda) &= -2\log\Big(\frac{L(\tilde{\beta},\tilde{\sigma}_\epsilon^2,0)}{L(\hat{\beta},\hat{\sigma}_\epsilon^2,\hat{\rho})}\Big) = -2\Big(\log L(\tilde{\beta},\tilde{\sigma}_\epsilon^2,0) - \log L(\hat{\beta},\hat{\sigma}_\epsilon^2,\hat{\rho})\Big) \\
&= -2\Big(-431.289 - (-385.419)\Big) = 91.74.
\end{aligned}
$$

The asymptotic distribution is given by:

$$
-2\log(\lambda) \sim \chi^2(G),
$$

where G is the number of the restrictions, i.e., $G = 1$ in this case.

The 1% upper probability point of $\chi^2(1)$ is 6.635.

$$
91.74 > 6.635
$$

Therefore, $H_0:\ \rho = 0$ is rejected.

There is serial correlation in the error term.

2. `Equation 1` (Test of Serial Correlation $\longrightarrow$ Lagrange Multiplier Test)

`Equation 2` is:

$$\texttt{@RES}_t = \rho\texttt{@RES}_{t-1} + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2),$$

where $\texttt{@RES}_t = \texttt{RCONS}_t - \hat{\beta}_1 - \hat{\beta}_2\texttt{RYD}_t$, and $\hat{\beta}_1$ and $\hat{\beta}_2$ are OLSEs.

The null hypothesis is $H_0: \ \rho = 0$

```
 @RES(-1)   .950693         .053301         17.8362         [.000]
```

Therefore, the Lagrange multiplier test statistic is $17.8362^2 = 318.13 > 6.635$.

$H_0: \ \rho = 0$ is rejected.

3. `Equation 3` (Test of Serial Correlation $\longrightarrow$ Wald Test)

Equation 3 is:

$$\text{RCONS}_t = \beta_1 + \beta_2 \text{RYD}_t + u_t, \qquad u_t = \rho u_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{ iid } N(0, \sigma_\epsilon^2)$$

The null hypothesis is $H_0 : \ \rho = 0$

```
RHO          .945025        .045843        20.6143          [.000]
```

The Wald test statistics is $20.6143^2 = 424.95$, which is compared with $\chi^2(1)$.

4. Equation 1 vs. NONLINEAR LEAST SQUARES (Choice of Functional Form – linear):

NONLINEAR LEAST SQUARES estimates:

$$\text{RCONS}_t = a1 + a2 \frac{\text{RYD}_t^{a3} - 1}{a3} + u_t.$$

When $a3 = 1$, we have:

$$\mathtt{RCONS}_t = (a1 - a2) + a2\mathtt{RYD}_t + u_t,$$

which is equivalent to `Equation 1`.

The null hypothesis is $H_0 :\ a3 = 1$, where $G = 1$.

- MLE with $a3 = 1$ MLE (`Equation 1`)

  `Log of likelihood function = -431.289`

- MLE without $a3 = 1$ (`NONLINEAR LEAST SQUARES`)

  `Log of likelihood function = -414.362`

The likelihood ratio test statistic is given by:

$$-2\log(\lambda) = -2\Big(-431.289 - (-414.362)\Big) = 33.854.$$

The 1% upper probability point of $\chi^2(1)$ is 6.635.

$$33.854 > 6.635$$

$H_0: \ a3 = 1$ is rejected by the likelihood ratio test.

Therefore, the functional form of the regression model is not linear.

5. `Equation 4` vs. `NONLINEAR LEAST SQUARES` (Choice of Functional Form – log-linear):

   In `NONLINEAR LEAST SQUARES`, i.e.,

$$\mathtt{RCONS}_t = a1 + a2\frac{\mathtt{RYD}_t^{a3} - 1}{a3} + u_t,$$

   if $a3 = 0$, we have:

$$\mathtt{RCONS}_t = a1 + a2\log(\mathtt{RYD}_t) + u_t,$$

which is equivalent to `Equation 3`.

The null hypothesis is $H_0: \ a3 = 0$, where $G = 1$.

- MLE with $a3 = 0$ (`Equation 3`)

  `Log of likelihood function = -495.418`

- MLE without $a3 = 0$ (`NONLINEAR LEAST SQUARES`)

  `Log of likelihood function = -414.362`

The likelihood ratio test statistic is:

$$-2\log(\lambda) = -2\Big(-495.418 - (-414.362)\Big) = 162.112 > 6.635.$$

Therefore, $H_0: \ a3 = 0$ is rejected.

As a result, the functional form of the regression model is not log-linear, either.

6. `Equation 1` vs. `Equation 5` (Simultaneous Test of Serial Correlation and Linear Function):

`Equation 5` is:

$$\mathtt{RCONS}_t = a1 + a2\frac{\mathtt{RYD}_t^{a3} - 1}{a3} + u_t, \qquad u_t = \rho u_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{ iid } N(0, \sigma_\epsilon^2)$$

The null hypothesis is $H_0:\ a3 = 1,\ \rho = 0$

Restricted MLE $\Longrightarrow$ `Equation 1`

Unrestricted MLE $\Longrightarrow$ `Equation 4`

Remark: In Lines 14–16 of `PROGRAM`, we have estimated `Equation 4`, given $a3 = 0.00, 0.01, 0.02, \cdots$.

As a result, $a3 = 1.15$ gives us the maximum log-likelihood.

The likelihood ratio test statistic is:

$$-2\log(\lambda) = -2\Big(-431.289 - (-383.807)\Big) = 94.964.$$

$-2\log(\lambda) \sim \chi^2(2)$ in this case.

The 1% upper probability point of $\chi^2(2)$ is 9.210.

$$94.964 > 9.210$$

$H_0: \; a3 = 1, \rho = 0$ is rejected.

`Equation 3` vs. `Equation 5` vs. (Taking into account serially correlated errors, Choice of Functional Form – linear):

The null hypothesis is $H_0: \ a3 = 1$

From `Equation 3`,

```
Log likelihood = -385.419
```

From `Equation 5`,

```
Log likelihood = -383.807
```

$$2(-383.807 - (-385.419)) = 3.224 < 6.635.$$

$H_0: \ a3 = 1$ is not rejected, given $\rho \neq 0$.

Thus, if serial correlation is taken into account, the regression model is linear.