

13 Large Sample Tests

13.1 Wald, LM and LR Tests

Parameter $\theta : k \times 1$, $h(\theta) : G \times 1$ 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., $I(\theta) = -E\left(\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'}\right)$.

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

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

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

(a) $h(\hat{\theta}) \approx h(\theta) + \frac{\partial h(\theta)}{\partial \theta'}(\hat{\theta} - \theta) \iff h(\hat{\theta})$ 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,

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

That is, approximately, we have the following result:

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

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

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

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

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

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:

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

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

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

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

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

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

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

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

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

$$E\left(\frac{\partial \log L(\theta)}{\partial \theta}\right) = 0, \quad V\left(\frac{\partial \log L(\theta)}{\partial \theta}\right) = -E\left(\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'}\right) = I(\theta).$$

(d) Therefore, using the central limit theorem,

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

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

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): $LR = -2 \log \lambda \rightarrow \chi^2(G)$

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

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

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

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

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

Note:

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

$$(2) -\frac{1}{n} \frac{\partial^2 \log L(\hat{\theta})}{\partial \theta \partial \theta'} \longrightarrow -\lim_{n \rightarrow \infty} \left(\frac{1}{n} E \left(\frac{\partial^2 \log L(\hat{\theta})}{\partial \theta \partial \theta'} \right) \right) = \lim_{n \rightarrow \infty} \left(\frac{1}{n} I(\theta) \right),$$

$$(3) \sqrt{n}(\hat{\theta} - \theta) \longrightarrow N \left(0, \lim_{n \rightarrow \infty} \left(\frac{1}{n} I(\theta) \right) \right).$$

(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 \implies 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ー, 1990=100) from the left.

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


```

          PROGRAM
LINE *****
|      1  freq a;
|      2  smpl 1955 1997;
|      3  read(file='cons99.txt') year cons yd price;
|      4  rcons=cons/(price/100);
|      5  ryd=yd/(price/100);
|      6  lyd=log(ryd);
|      7  olsq rcons c ryd;
|      8  olsq @res @res(-1);
|      9  ar1 rcons c ryd;
|     10  olsq rcons c lyd;
|     11  param a1 0 a2 0 a3 1;
|     12  frml eq rcons=a1+a2*((ryd**a3)-1.)/a3;
|     13  lsq(tol=0.00001,maxit=100) eq;
|     14  a3=1.15;
|     15  rryd=((ryd**a3)-1.)/a3;
|     16  ar1 rcons c rryd;
|     17  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

Variable	Estimated Coefficient	Standard 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

Parameter	Estimate	Standard 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

Parameter	Estimate	Standard 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:

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

Equation 3 is:

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

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

Restricted MLE \implies Equation 1

Unrestricted MLE \implies Equation 3

The log-likelihood function of Equation 3 is:

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

where

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

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

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

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

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

$$\text{Log of likelihood function} = -431.289$$

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

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

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

$$\text{Log of likelihood function} = -385.419$$

The likelihood ratio test statistic is:

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

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 \rightarrow Lagrange Multiplier Test)

Equation 2 is:

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

where $\text{@RES}_t = \text{RCONS}_t - \hat{\beta}_1 - \hat{\beta}_2 \text{RYD}_t$, 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 \rightarrow Wald Test)

Equation 3 is:

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

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 $a_3 = 1$, we have:

$$\text{RCONS}_t = (a_1 - a_2) + a_2 \text{RYD}_t + u_t,$$

which is equivalent to Equation 1.

The null hypothesis is $H_0 : a_3 = 1$, where $G = 1$.

- MLE with $a_3 = 1$ MLE (Equation 1)

Log of likelihood function = -431.289

- MLE without $a_3 = 1$ (NONLINEAR LEAST SQUARES)

Log of likelihood function = -414.362

The likelihood ratio test statistic is given by:

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

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

$$33.854 > 6.635$$

$H_0 : a_3 = 1$ is rejected.

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.,

$$\text{RCONS}_t = a_1 + a_2 \frac{\text{RYD}_t^{a_3} - 1}{a_3} + u_t,$$

if $a_3 = 0$, we have:

$$\text{RCONS}_t = a_1 + a_2 \log(\text{RYD}_t) + u_t,$$

which is equivalent to Equation 3.

The null hypothesis is $H_0 : a_3 = 0$, where $G = 1$.

- MLE with $a_3 = 0$ (Equation 3)

Log of likelihood function = -495.418

- MLE without $a_3 = 0$ (NONLINEAR LEAST SQUARES)

Log of likelihood function = -414.362

The likelihood ratio test statistic is:

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

Therefore, $H_0 : a_3 = 0$ 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:

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

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

Restricted MLE \implies Equation 1

Unrestricted MLE \implies Equation 4

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

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

The likelihood ratio test statistic is:

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

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

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

$$94.964 > 9.210$$

$H_0 : a_3 = 1, \rho = 0$ is rejected.

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

The null hypothesis is $H_0 : a_3 = 1, \rho = 0$

From Equation 3,

$$\text{Log likelihood} = -385.419$$

From Equation 5,

$$\text{Log likelihood} = -383.807$$

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

$H_0 : a_3 = 1$ is not rejected, given $\rho \neq 0$.

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