

Lecture 38: ρ unknown: estimating ρ ; the Cochrane–Orcutt procedure

Econometrics ■ Unit III: Autocorrelation and instrumental variables

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

B.Sc. (Hons) Statistics, Semester V
Academic Session 2026–2027

Previous lecture

Lecture 37: ρ known: the generalized difference equation and the Prais–Winsten transformation.

Today

- Solution of autocorrelation
 - When ρ is unknown: estimating ρ
 - The Cochrane–Orcutt iterative procedure

(a) From the Durbin–Watson statistic. $\hat{\rho} = 1 - d/2$.

(b) The Theil–Nagar estimator. In small samples, $1 - d/2$ tends to underestimate ρ . Theil and Nagar (1961) suggested

$$\hat{\rho} = \frac{n^2(1 - d/2) + k^2}{n^2 - k^2},$$

where k is the number of coefficients including the intercept. For large n this is close to $1 - d/2$.

For example, with $n = 20$, $k = 2$ and $d = 1.2$: $1 - d/2 = 0.4$ and

$$\hat{\rho} = \frac{400(0.4) + 4}{400 - 4} = \frac{164}{396} = 0.414.$$

(c) From the OLS residuals. Choose $\hat{\rho}$ to minimise $\sum_{t=2}^n (e_t - \rho e_{t-1})^2$. Differentiating with respect to ρ ,

$$\frac{d}{d\rho} \sum_{t=2}^n (e_t - \rho e_{t-1})^2 = -2 \sum_{t=2}^n e_{t-1} (e_t - \rho e_{t-1}) = 0 \implies \hat{\rho} = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=2}^n e_{t-1}^2}.$$

For the residuals $(2, 1, 0, -1, -2, -1, 0, 1)$:

$\sum e_t e_{t-1} = 2 + 0 + 0 + 2 + 2 + 0 + 0 = 6$ and

$\sum_{t=2}^8 e_{t-1}^2 = 4 + 1 + 0 + 1 + 4 + 1 + 0 = 11$, so $\hat{\rho} = 6/11 = 0.545$.

(d) First differences. If ρ is close to 1, put $\rho = 1$:

$$y_t - y_{t-1} = \beta_1(x_t - x_{t-1}) + (u_t - u_{t-1}).$$

The intercept disappears. This should be used only when ρ is really near 1.

The Cochrane–Orcutt iterative procedure

This procedure is due to Cochrane and Orcutt (1949).

1. Estimate the original model by OLS and obtain the residuals e_t .
2. Estimate ρ by $\hat{\rho} = \sum_{t=2}^n e_t e_{t-1} / \sum_{t=2}^n e_{t-1}^2$.
3. Form $y_t^* = y_t - \hat{\rho} y_{t-1}$ and $x_t^* = x_t - \hat{\rho} x_{t-1}$ for $t = 2, \dots, n$, and estimate $y_t^* = \beta_0^* + \beta_1 x_t^* + v_t$ by OLS. Recover $\hat{\beta}_0 = \hat{\beta}_0^* / (1 - \hat{\rho})$.
4. Compute new residuals from the *original* equation, $e_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t$, and re-estimate ρ .
5. Repeat Steps 3 and 4 until successive estimates of ρ (and of the coefficients) differ by less than a small tolerance, say 0.001.

This is a feasible GLS method. Its estimates are not exactly BLUE, but they are good in large samples. An estimate $\hat{\rho}$ outside $(-1, 1)$ indicates that the AR(1) model is not suitable.

Practice

- Exercise 5.7 (Exercises and Solutions booklet).

Next lecture

Lecture 39: Durbin's two-step method; the Hildreth–Lu method.