Applied Time Series Econometrics

Johannes W. Fedderke
ERSA and University of Cape Town

Second Semester 2006

1. What is Autocorrelation?

1. The violation of the assumption that the error terms of different observations are not related to each other, i.e. $cov\left(u_i, u_j\right) \neq 0$.

2. Reasons for Autocorrelation

1. Inertia.
2. Cobweb Phenomena.
3. Data Manipulation.
4. Spurious inference of AR:
   (a) Estimation by means of OLS in the presence of non-stationary data: the presence of a trend in the variance of the data will generate a trend in the mean - which appears as AR.
   (b) ARCH and GARCH processes

5. Model Misspecification:

(a) Omitted variables.

(b) Dynamics: Suppose that we have:

$$\begin{aligned} y_t &= \beta x_t + u_t \\ u_t &= \rho u_{t-1} + \varepsilon_t \\ \varepsilon_t &\sim iid\left(0, \sigma^2\right) \end{aligned}$$

then since $u_{t-1} = y_{t-1} - \beta x_{t-1}$, by simple substitution it follows that:

$$y_t = \rho y_{t-1} + \beta x_t - \rho\beta x_{t-1} + \varepsilon_t \qquad (1)$$

Now note that estimation of:

$$y_t = \beta_1 y_{t-1} + \beta_2 x_t + \beta_3 x_{t-1} + \varepsilon_t \qquad (2)$$

would give 1 if $\beta_1\beta_2 + \beta_3 = 0$.

Testing the restriction, and where we *reject* the null that $\beta_1\beta_2 + \beta_3 = 0$, we can conclude that we do *not* have AR, but that instead we have a case of misspecified dynamics instead. *Only* where the null is accepted, can we then proceed further in order to test the additional restriction of $\beta_1 = 0$, rejection of which would confirm the presence of autocorrelation.

Since the test for $\beta_1\beta_2 + \beta_3 = 0$ is non-linear in the β's, standard t-tests do not apply. Instead we need to use a Wald test, or a LR or LM test instead.

3. What are the Consequences of Autocorrelation?Model Selection and Validation

What are the consequences of autocorrelation?

These are the same as in the case of heteroscedasticity:

- The OLS estimators of the coefficients remain *consistent*.
- The OLS estimators of the coefficients remain *unbiased*.
- The OLS estimators of the coefficients are, however, no longer the minimum variance estimators, i.e. they are no longer *efficient*. The efficient estimators are the *GLS estimators*.

- The OLS estimators of the standard errors are *biased* and *inconsistent*.
- Because the estimator of σ, $\widehat{\sigma}^2$ is biased, the R^2 of the regression will also prove to be biased.

Concretely this means that if we are interested only in estimating the coefficients, OLS remains a viable method of estimation. Any t-tests or F-tests which we might wish to conduct become unreliable, however.

4. Detecting Autocorrelation

A number of methods are available for detecting autocorrelation:

1. *Plotting residuals*

2. *The Runs/Geary test:* This test examines the number of "runs" in the residuals, defined as the number of sequences of consecutive positive or negative errors to occur from the estimation, and compares them to the number of runs that would be purely stochastic.

In particular, we define:

$$\begin{aligned} n_1 &\equiv number\ of\ +\ symbols \\ n_2 &\equiv number\ of\ -\ symbols \\ n &\equiv n_1 + n_2 \\ k &\equiv number\ of\ runs \\ mean\ :\ \ & E(k) = \frac{2n_1n_2}{n_1+n_2} + 1 \\ variance\ :\ \ & \sigma_k^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)} \end{aligned}$$

Reject the null of randomness if $\widehat{E(k) + 1.96\sigma_k} > k > \widehat{E(k) + 1.96}\sigma_k$.

3. *The Durbin-Watson d-test:* Defined as:

$$\begin{aligned} d &= \frac{\sum_{t=2}^{t=n} \left(\widehat{u}_t - \widehat{u}_{t-1}\right)^2}{\sum_{t=2}^{t=n} \widehat{u}_t^2} \\ &\simeq 2\left(1 - \frac{\sum \widehat{u}_t \widehat{u}_{t-1}}{\sum \widehat{u}_t^2}\right) \\ &\simeq 2\left(1 - \widehat{\rho}\right) \end{aligned}$$

where $\simeq$ means "approximately equal to," $\widehat{u}_t$ the sample estimate of the residuals, and $\widehat{\rho}$ denotes the estimate of the sample first order coefficient of autocorrelation.

Note that:

(a) The d-test is valid only if the regression includes an intercept term.
(b) The d-test is valid only if the explanatory variables are non-stochastic (fixed in repeated sampling).
(c) The d-test is valid only for frst order autoregressive structures: $\varepsilon_t = \rho\varepsilon_{t-1} + \nu_t$.
(d) The d-test is valid only if the model does not include lagged values of the dependent variable, i.e. we do not have an autoregressive model.
(e) The d-test is valid only if there are no missing values in the data.

4. *The Durbin h-test*: For models that do include a lagged dependent variable, for instance such that:

$$y_t = \alpha y_{t-1} + \beta x_t + \varepsilon_t \tag{3}$$

$$\varepsilon_t = \rho \varepsilon_{t-1} + \nu_t \tag{4}$$

Durbin suggests:

$$h = \widehat{\rho}\sqrt{\frac{n}{1 - n\widehat{\sigma}^2_\alpha}}$$

where $\widehat{\rho}$ denotes the estimate of the sample first order coefficient of autocorrelation, n sample size, and $\widehat{\sigma}^2_\alpha$ the estimated variance of the OLS α-estimate. Note that the h-test remains a test for first order autocorrelation.

Under $n\widehat{\sigma}_{\alpha}^{2} > 1$ the test does not apply, and then Durbin suggests the alternative test:

5. *Durbin's alternative test*: From the OLS estimation of 3 compute the residuals $\widehat{\varepsilon}_t$, and then test for the significance of ρ by estimating $\widehat{\varepsilon}_t = \rho\widehat{\varepsilon}_{t-1} + \nu_t$. Note that the alternative test remains a test for first order autocorrelation.

6. *Asymptotic, large-sample test:* Under the null that $\rho = 0$, and for $n \to \infty$, it can be shown that:

$$\widehat{\rho}\sqrt{n} \sim N\left(0, 1\right)$$

7. *The Breusch-Godfrey (BG) test of higher-order autocorrelation:* Suppose we have disturbances, u_t, generated by the p'th-order process:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + ... + \rho_p u_{t-p} + \varepsilon_t \quad (5)$$
$$\varepsilon_t \sim iid\left(0, \sigma^2\right)$$

We can test the null that $\rho_1 = \rho_2 = ... = \rho_p = 0$, by:

(a) Estimating the underlying model to obtain $\widehat{u}_t$.

(b) Estimate 5 to obtain an R^2.

(c) Breusch and Godfrey have shown that:

$$(n - p)\, R^2 \sim \chi_p^2$$

i.e. is distributed chi-square with p degrees of freedom. Where $(n - p)\, R^2$ >the critical χ_p^2, we can reject the $\rho_1 = \rho_2 = ... = \rho_p = 0$ null and accept that at least one $\rho_i \neq 0$.

Note:

(d) The BG test is valid even in the presence of a lagged dependent variable in the underlying model.

(e) The BG test is valid even if the error term follows a p'th order MA process, i.e. if:

$$u_t = \varepsilon_t + \lambda_1 \varepsilon_{t-1} + \lambda_2 \varepsilon_{t-2} + ... + \lambda_p \varepsilon_{t-p}$$

(f) Where p=1, we have an AR(1) process, and the BG test is known as the Durbin m-test.

(g) The p-order of the test cannot be known a priori, and must be established by experimentation.

8. *The Berenblutt-Webb g-test of the hypothesis that* $\rho = 1$: specified as:

$$g = \frac{\sum_{t=2}^{n} \widehat{e}_t^2}{\sum_{t=1}^{n} \widehat{u}_t^2}$$

where the $\widehat{u}$ are the residuals from the original model, and $\widehat{e}$ the residuals from the model in first difference format.

5. Solutions to Autocorrelation

5.1 Where ρ is known (Generalized Least Squares)

There are many different ways in which one might specify the nature of the autocorrelation (i.e. the order of the AR process). We restrict ourselves to the case of a simple AR(1) autoregressive process, i.e. in which the error terms follow the pattern

$$u_t = \rho u_{t-1} + \epsilon_t \tag{6}$$

and where the ϵ_t errors are assumed to be uncorrelated.

Our specification of the model is therefore:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + ... + \beta_k x_{kt} + u_t \tag{7}$$

$$u_t = \rho u_{t-1} + \epsilon_t \tag{8}$$

If this model is true for all t, it is certainly true for period t-1, so

$$y_{t-1} = \beta_1 + \beta_2 x_{2t-1} + \beta_3 x_{3t-1} + ... + \beta_k x_{kt-1} + u_{t-1} \tag{9}$$

multiplying by ρ:

$$\rho y_{t-1} = \rho\beta_1 + \rho\beta_2 x_{2t-1} + \rho\beta_3 x_{3t-1} + ... + \rho\beta_k x_{kt-1} + \rho u_{t-1} \tag{10}$$

and we know:

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + ... + \beta_k x_{kt} + \rho u_{t-1} + \epsilon_t \tag{11}$$

Subtracting equation 10 from 11 we obtain:

$$y_t - \rho y_{t-1} = \beta_1 (1 - \rho) + \beta_2 (x_{2t} - \rho x_{2t-1}) + ... + \beta_k (x_{kt} - \rho x_{kt-1}) + \epsilon_t \tag{12}$$

Letting $y_t - \rho y_{t-1} = z_t$, $\beta_1 (1 - \rho) = \alpha$, $(x_{kt} - \rho x_{kt-1}) = w_{kt}$, we obtain

$$z_t = \alpha + \beta_2 w_{2t} + ... + \beta_k w_{kt} + \epsilon_t \qquad (13)$$

In essence what this does is to subtract out the autocorrelated part of the process and leaves a model that is free of autocorrelation. It can therefore be estimated by OLS.

Note that the intercept of the equation 13 does not immediately provide an estimate for the original model. Given ρ however, it is easy to calculate β_1 from α. We have $\beta_1 = \frac{\alpha}{1-\rho}$. The other coefficients are not, however, affected by the differencing process on the data.

5.2 Where ρ is unknown

1. The first-difference method: Where $\rho = 1$, equation 12 reduces to:

$$
\begin{aligned}
y_t - y_{t-1} &= \beta_2 (x_{2t} - x_{2t-1}) + ... + \beta_k (x_{kt} - x_{kt-1}) + \epsilon_t \\
\Delta y_t &= \beta_2 \Delta x_{2t} + ... + \beta_k \Delta x_{kt} + \varepsilon_t
\end{aligned}
$$

Note the absence of an intercept term[1] - but where $\rho = 1$ the first difference method gives reliable estimates of the $\beta_i, i > 1$. Of course, the method applies only where the $\rho = 1$ is satisfied - for which we can employ the Beerenblutt-Webb test.

2. ρ based on d: We have noted that:

$$d \simeq 2(1-\widehat{\rho})$$
$$\therefore \widehat{\rho} \simeq 1-\frac{d}{2}$$

This gives an approximate value for ρ, which allows us to undertake GLS estimation.

2. Cochrane-Orcutt iterative procedure: Consider:

$$y_t = \beta_1 + \beta_2 x_t + u_t \tag{14}$$

$$u_t = \rho u_{t-1} + \varepsilon_t \tag{15}$$

$$\varepsilon_t \sim iid\left(0, \sigma^2\right) \tag{16}$$

Cochrane-Orcutt suggest:

(a) Estimation of 14 to obtain $\widehat{u}_t$.

(b) Employing $\widehat{u}_t$ to estimate 15, to obtain $\widehat{\rho}$.

(c) Employing $\widehat{\rho}$ to estimate:

$$y_t - \widehat{\rho} y_{t-1} = \beta_1 (1 - \widehat{\rho}) + \beta_2 (x_t - \widehat{\rho} x_{t-1}) + \epsilon_t \tag{17}$$

in order to obtain $\widehat{\beta}_1$, $\widehat{\beta}_2$.

(d) Substitute $\widehat{\beta}_1$, $\widehat{\beta}_2$, into 14 to obtain revised $\widehat{u}'_t$.

(e) Repeat from 2, until $\widehat{\rho}$ estimates stable.

3. Cochrane-Orcutt two-step procedure: As above - except we stop after step 3.

4. Durbin's two-step procedure: Suppose again that we have the model as specified in 14 through 16. Then:

$$y_t = \beta_1 (1 - \widehat{\rho}) + \beta_2 (x_t - \widehat{\rho} x_{t-1}) + \rho y_{t-1} + \epsilon_t \tag{18}$$

Estimation gives a direct estimate of $\widehat{\rho}$, from the coeffiecient of y_{t-1}. This then allows direct estimation of 17.

5. Grid-search procedures: Hildreth-Lu: Employ GLS under different values of ρ in the interval $-1 \leq \rho \leq 1$, in stepped increases of 0.1. Choose the ρ with the minimum residual sum of squares. Repeat the grid search in the interval $\rho^* - 0.1 \leq \rho^* \leq \rho^* + 0.1$, in steps of 0.01, where ρ^* denotes the preferred ρ-value.