# Serial Correlation in Time Series Regressions

## Lecture 10 / B

# Analyzing Time Series: Serial Correlation.

- Strict exogeneity ⇒OLS is unbiased.

- Contemporaneous exogeneity ⇒ OLS is consistent (provided the time series are weakly dependent).

- Unbiasedness/Consistency did not require assumptions on error autocorrelation.

- There are situations where the nature of the $x_{jt}$ mean that serial correlation in $u_t$ implies that $u_t$ is correlated with $x_{jt}$.

# Analyzing Time Series: Serial Correl. and Heterosced.

- **Testing for serial correlation**
- **Testing for AR(1) serial correlation with strictly exog. regressors**

$$y_t = \beta_0 + \beta_1 x_{t1} + \cdots + \beta_k x_{tk} + u_t$$

AR(1) model for serial correlation (with an i.i.d. series $e_t$)

$$u_t = \rho u_{t-1} + e_t$$

Replace true unobserved errors by estimated residuals

Test $H_0 : \rho = 0$ in $\widehat{u}_t = \rho \widehat{u}_{t-1} + error$

- Large sample justification (unobservable residuals)

# Analyzing Time Series: Serial Correlation.

- **Durbin-Watson test under classical assumptions**
  - Under assumptions TS.1 – TS.6, the Durbin-Watson test is an exact test (whereas the previous t-test is only valid asymptotically).

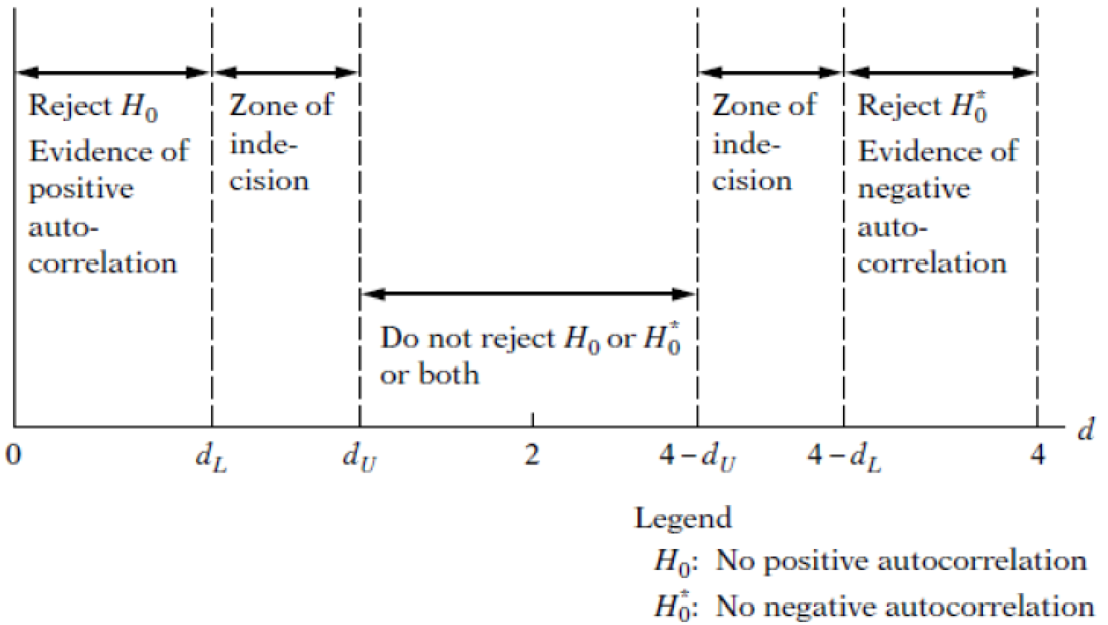
$$DW = \sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2 / \sum_{t=2}^{n} \hat{u}_t^2 \approx 2(1-\hat{\rho})$$

$H_0 : \rho = 0$ vs. $H_1 : \rho > 0$

Reject if $DW < d_L$, "Accept" if $DW > d_U$

Unfortunately, the Durbin-Watson test works with a lower and and an upper bound for the critical value. In the area between the bounds the test result is inconclusive.

# Analyzing Time Series: Serial Correl. and Heterosced.




DURBIN–WATSON *d* TEST: DECISION RULES

| Null hypothesis | Decision | If |
|---|---|---|
| No positive autocorrelation | Reject | $0 < d < d_L$ |
| No positive autocorrelation | No decision | $d_L \le d \le d_U$ |
| No negative correlation | Reject | $4 - d_L < d < 4$ |
| No negative correlation | No decision | $4 - d_U \le d \le 4 - d_L$ |
| No autocorrelation, positive or negative | Do not reject | $d_U < d < 4 - d_U$ |

# Analyzing Time Series: Serial Correlation.

- When strictly exogeneity does not hold, one or more $x_{jt}$ might be correlated with with $u_{t-1}$.
- $t$ −test and DW test are not valid (even asymptotically)
- Example: lagged dependent variables as regressors: $y_{t-1}$ and $u_{t-1}$ are obviously correlated.
- Consider the AR(1) test.

# Analyzing Time Series: Serial Correlation.

1. Regress $y_t$ on a $1, x_{1t}, x_{2t}, \dots, x_{kt}$ and obtain $\hat{u}_t$.
2. Regress $\hat{u}_t$ on $1, x_{1t}, x_{2t}, \dots, x_{kt}$, $\hat{u}_{t-1}$.
3. (heteroscedasticity robust) $t$ –test coefficient on $\hat{u}_{t-1}$

The inclusion of the regressors $x_{1t}, x_{2t}, \dots, x_{kt}$ allows for each $x_{jt}$ to be correlated with $u_t$.

# Analyzing Time Series: Serial Correlation.

- **Testing for AR(1) serial correlation with general regressors**
  - The t-test for autocorrelation can be easily generalized to allow for the possibility that the explanatory variables are not strictly exogenous:

$$\widehat{u}_t = \alpha_0 + \alpha_1 x_{t1} + \cdots + \alpha_k x_{tk} + \boxed{\rho}\,\widehat{u}_{t-1} + error$$

The test now allows for the possibility that the strict exogeneity assumption is violated.

Test for $H_0 : \rho = 0$

  - The test may be carried out in a heteroscedasticity robust way
- **General Breusch-Godfrey test for AR(q) serial correlation**

$$\widehat{u}_t = \alpha_0 + \alpha_1 x_{t1} + \cdots + \alpha_k x_{tk} + \boxed{\rho_1}\,\widehat{u}_{t-1} + \cdots + \boxed{\rho_q}\,\widehat{u}_{t-q} + \ldots$$

Test $H_0 : \rho_1 = \cdots = \rho_q = 0$

# Analyzing Time Series: Serial Correl. and Heterosced.

- **Correcting for serial correlation with strictly exog. regressors**
  - Under the assumption of AR(1) errors, one can transform the model so that it satisfies all GM-assumptions. For this model, OLS is BLUE.

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

Simple case of regression with only one explanatory variable. The general case works analogously.

$$\rho y_{t-1} = \rho\beta_0 + \rho\beta_1 x_{t-1} + \rho u_{t-1}$$

Lag and multiply by $\rho$

$$\Rightarrow \boxed{y_t - \rho y_{t-1}} = \beta_0 \boxed{(1-\rho)} + \beta_1 \boxed{(x_t - \rho x_{t-1})} + \boxed{u_t - \rho u_{t-1}}$$

$$u_t = \rho u_{t-1} + e_t \Leftrightarrow u_t - \rho u_{t-1} = \boxed{e_t}$$

The transformed error satisfies the GM-assumptions.

  - Problem: The AR(1)-coefficient is not known and has to be estimated (FGLS)

# Analyzing Time Series: Serial Correlation

- **Serial correlation-robust inference after OLS**
  - In the presence of positive serial correlation, OLS standard errors overstate statistical significance
  - One can compute serial correlation-robust std. errors after OLS
  - This is useful because FGLS requires strict exogeneity and assumes a very specific form of serial correlation (AR(1) or, generally, AR(q))
  - Serial correlation-robust standard errors:

$$se(\widehat{\beta}_j) = \left[\text{“}se(\widehat{\beta}_j)\text{”}/\widehat{\sigma}\right]^2 \sqrt{\widehat{v}}$$

The usual OLS standard errors are normalized and then „inflated“ by a correction factor.

  - Serial correlation-robust F- and t-tests are also available

# Analyzing Time Series: Serial Correlation

- **Correction factor for serial correlation (Newey-West formula)**

$$\hat{v} = \sum_{t=1}^{n} \hat{a}_t^2 + 2 \sum_{h=1}^{g} [1 - h/(g+1)] \left( \sum_{t=h+1}^{n} \hat{a}_t \hat{a}_{t-h} \right)$$

$$\hat{a}_t = \hat{r}_t \hat{u}_t$$

This term is the product of the residuals and the residuals of a regression of $x_{tj}$ on all other explanatory variables

The integer g controls how much serial correlation is allowed:

g=2: $\hat{v} = \sum_{t=1}^{n} \hat{a}_t^2 + \sum_{t=2}^{n} \hat{a}_t \hat{a}_{t-1}$

The weight of higher order autocorrelations is declining

g=3: $\hat{v} = \sum_{t=1}^{n} \hat{a}_t^2 + (4/3) \sum_{t=2}^{n} \hat{a}_t \hat{a}_{t-1} + (2/3) \sum_{t=3}^{n} \hat{a}_t \hat{a}_{t-2}$

# Analyzing Time Series: Serial Correlation

- **Discussion of serial correlation-robust standard errors**
  - The formulas are also robust to heteroscedasticity; they are therefore called "heteroscedasticity and autocorrelation consistent" (=HAC)
  - For the integer g, values such as g=2 or g=3 are normally sufficient (there are more involved rules of thumb for how to choose g)
  - Serial correlation-robust standard errors are only valid asymptotically; they may be severely biased if the sample size is not large enough
  - The bias is the higher the more autocorrelation there is; if the series are highly correlated, it might be a good idea to difference them first
  - Serial correlation-robust errors should be used if there is serial corr. and strict exogeneity fails (e.g. in the presence of lagged dep. var.)

# Reading

- Chapter 12