Paper 3, Part 2A
Introduction to Econometrics II - Macroeconometrics.
2018-2019
Suggested supervision 1.

1. (Deterministic trend) The STATA file wheatIndia.dta contain annual time series on India's wheat production. Variable `wheat` measures wheat production in 1000 tones.

   (a) Make a time series plot of the variable wheat. How would you model the apparent deterministic trend in the data? What might be economic reason(s) for this trend?

   (b) Estimate a linear and a quadratic trend models. Superimpose your trend estimates on the time series plot of wheat. Use dashed lines for the trends and a solid line for the actual series. What model is preferred by AIC? What model is preferred by BIC? Do these preferences make sense intuitively? Why?

   (c) Using the model preferred by AIC, predict India's wheat production in 2018, and provide a 95% confidence interval for your prediction. What assumptions do you make to justify this forecasting exercise?

   (d) Compute the residuals of the trend model that you have estimated in (b) and used in (c). Make their time series plot. Perform the Durbin-Watson test (at 5% significance level) for no serial correlation in these residuals. What does the result of the test imply about the validity of the analysis in (c)? What may economic reason(s) for the Durbin-Watson test finding be?

2. (Deterministic trend) In class, we have estimated an exponential trend for UK consumption using nonlinear least squares. Another way to estimate an exponential trend is to take the logarithm of UK consumption, estimate a linear trend for this logarithm, and exponentiate the result.

   (a) Using file consumption.dta, estimate the exponential trend by the above two methods. Superimpose the two trends on the time series plot of consumption. Explain the obvious difference in the estimated trends. Which of the estimates you would prefer from a forecasting perspective? Why?

   (b) Why cannot you compare AIC values obtained from the nonlinear least squares and the ordinary least squares that you performed in (a)?

   (c) A popular measure of forecast accuracy is out-of-sample mean squared error, $MSE = E\left(e^2\right)$. Show that MSE is equal to the sum of the variance of the error and the square of the mean error. A forecast is unbiased if the mean forecast error is zero. Why might unbiased forecasts be desirable? Are they necessarily desirable?

(d) Suppose that $(\log y)_{T+h,T}$ is an unbiased forecast of $(\log y)_{T+h}$. Is $\exp\left((\log y)_{T+h,T}\right)$ an unbiased forecast of y_{T+h}? Why?

3. (Seasonality) File flu.dta contains data on the popularity of "flu" Google searchers in UK.

   (a) Make a time series plot of flu variable. What seasonal pattern do you observe? What may be a reason for such a pattern? December 2010 and Januaries 2011 and 2018 look unusually high. Explore what may the reasons be.

   (b) Estimate a seasonal dummy variable model for the flu variable. Include a linear trend and dummies for December 2010, January 2011 and January 2018. Report the estimated seasonal pattern. Which month do you think is the most dangerous for getting a flue? Which month is the least dangerous?

   (c) Perform the Durbin-Watson test (at 5% significance level) on the residuals. What do you conclude?

4. (Covariance stationarity) Consider a stochastic process $...y_{-2}, y_{-1}, y_0, y_1, y_2...$ where y_t and y_s are independent as long as $t \neq s$, and

$$
\begin{aligned}
y_t &= \begin{cases} 3 & \text{with probability } 1/4 \\ -1 & \text{with probability } 3/4 \end{cases} \quad \text{for odd } t, \\
y_t &= \begin{cases} 1 & \text{with probability } 3/4 \\ -3 & \text{with probability } 1/4 \end{cases} \quad \text{for even } t.
\end{aligned}
$$

   (a) Is this process covariance stationary? Explain. What is the limit in distribution of $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_t$ as $T \to \infty$?

   (b) Now, consider process $...x_{-2}, x_{-1}, x_0, x_1, x_2...$ where

$$x_t = y_1 \text{ for odd } t \text{ and } x_t = y_2 \text{ for even } t.$$

   Is this process weakly stationary? Explain.

   (c) Does $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_t$ converge in distribution? Explain.

5. (Sample correlogram) The STATA file BoatRace.dta contains information on the annual Oxford and Cambridge Boat Race for the period 1946-2013. Variable $Winner_t$ equals 1 if the winner of the race during year t is Cambridge, and 0 if it is the other place. Variable $WinTime_t$ is the winning time (in seconds) in the year t race. Use STATA to answer the following questions. Before running any commands, make sure that you declare variable $Year_t$ as time by typing `tsset Year`.

   (a) Make the time series plot of the variable $Winner_t$. Make the sample correlogram of $Winner_t$. What does the correlogram tell us about serial correlation in $Winner_t$?

(b) Use Ljung-Box test (with $m = 4$) to test a hypothesis that the Winner series is white noise.

(c) Make the time plot of the variable $WinTime_t$. Briefly describe the time series behaviour of this variable.

(d) Generate the first differences in $WinTime_t$. Call it DWT_t. Suppose that the data generating process for $WinTime_t$ is

$$WinTime_t = \gamma + \delta \times Year_t + v_t,$$

where v_t are i.i.d. shocks. What are then the population autocorrelations ρ_j, $j = 1, 2, ...$ for DWT_t?

(e) Make the sample correlogram of DWT_t. Comment, relating your result to (d).