Home Assignment # 3

ARE 256b
Due: 3/10/2023

Instructions. You are welcome to work in groups of a maximum of three students. Each group member has to hand in their individual problem sets and compose their separate Stata files. Please list your group members on your problem set. Present hand-written answers to each question clearly and concisely and use tables to present results wherever possible. Attach a print-out of your Stata do file and upload it on canvas. The problem sets are meant to help you understand the material. If you have any questions, please come to the instructor or TA.

1. **Forecasts and Non-stationary time series.** Let time series y_t satisfy $y_t = \mu + \rho y_{t-1} + \varepsilon_t$, where ε_t is an i.i.d. Gaussian process with zero mean and variance σ_ε^2.
   (a) Assume that the process y_t is stationary, i.e. $|\rho| < 1$. Compute unconditional mean Ey_t.
   (b) Compute conditional expectaiton $E(y_{t+h}|y_t)$ for integer $h > 0$.
   (c) Compute $\lim_{h\to\infty} E(y_{t+h}|y_t)$. Does it depend on y_t?
   (d) Now assume that $\rho = 1$. Compute $\lim_{h\to\infty} E(y_{t+h}|y_t)$. Does it depend on y_t?

2. **Seasonal patterns.**
   In this problem, you use the cement dataset (*cement* and its description *cement_Description* are on Canvas). Note that you have to use *tsset* first to indicate to stata that your data consists of a time series. [1]
   Write down the regression equation for each regression you perform.
   (a) **Seasonal dummies** Generate new variable *seasonal* that is equal to 1 for t corresponding to January and 0 otherwise. Compute regression of *gcem* on $L(0/10)$*seasonal*. Which months has the largest and the smallest average values for *gcem*? Report the values.
   (b) **Breusch-Godfrey Test of Serial Correlation.** Perform the test for serial correlation in the error term for the regression of *gcem* on *grres* for serial correlation of order 1, 3 and 5. Compute the nR^2 test using the procedure outlined below, with the statistic from the ready-made command for the Breusch-Godfrey test, which is *estat bgodfrey lags(p)*, following *regress gcem grres*, where p is the order of serial correlation. Do both methods deliver the same test? What is the asymptotic distribution of the Breusch-Godfrey statistic for serial correlation of order 1, 3, and 5?
   The following steps outline how to perform the Breusch-Godfrey test for first-order serial correlation in the error term:
      i. *Perform the OLS regression* gcem *on* grres.
      ii. *Obtain residuals from that regression.*
      iii. *Generate the lagged residual.*
      iv. *Perform the auxiliary regression of the residual on its own lag and the regressor* grres.
      v. *Compute the Breusch-Godfrey statistic using nR^2 from the above regression.*

[1] To do so, you need to generate a month-year variable using *gen date=ym(year,month)*, then use the date variable in tsset, as in *tsset date*.

(c) **Correcting for Serial Correlation.** In order to correct for serial correlation, you can use Newey-West standard errors, using the *newey* command in stata. Choose the number of lags that is appropriate according to your results in (b). Report the Newey-West standard errors and compare them to the standard errors you get from the *regress* command. What is the difference between the Newey-West standard errors and the robust standard errors you obtain from *regress*.

(d) **Testing for a structural break.** Formulate a statistical hypothesis to test for significant difference in mean between January and February in notation from (a). Test this statistical hypothesis using t-test.

3. **Time trend.**
Demonstrate that the OLS estimator of δ in the model

$$Y_t = \beta_1 + \delta t + \varepsilon_t, t = 1, ..., T$$

is superconsistent (here ε_t is an i.i.d. zero mean r.v.). Show also that it is unbiased in finite samples , despite the fact that Y_t is nonstationary.