

Economics 312  
Project #6 Assignment

Spring 2014  
Due: Midnight, Wednesday, April 9

*Partner assignments*

Kai Addae	Cole Sprague
Prakher Bajpai	Will Schmid
Emmanuel Enemchukwu	Alec Recinos
Julian Haft	Stephanie Radoslovich
Paapa hMensa	Mat Olson
Julia Hofmann	John Mills
Daniel Hope	Dylan McKenna
Mark Jarrett	Theo Landsman
James LaBelle	Dean Young
Timothy Tyree	Chris Weber
Austin Weisgrau	Colleen Werkheiser

This assignment is based loosely on the textbook's Exercises 9.14 and 9.15 and uses the HGL dataset `bangla.dta`, containing 34 years of data on sugar-cane prices (relative to the alternative crop jute) and the area planted in sugar cane in a region of Bangladesh. These are the only two variables in the data set.

HGL suggest a log-log model and we are going to assume that current and lagged price are exogenous with respect to planting, so the only specification decisions are the lag structure of the equation and the appropriate time-series specification for the error term.

You are to find and estimate an appropriate dynamic specification for the relationship between area planted and price, providing detailed justification for your choices. In coming to this specification, you should explore (and report on) the suggestions of Exercise 9.15 about the error structure and Exercise 9.14 about the lag distribution.

In addition to the tasks described in these two exercises, you should consider the following in your exploration:

- When comparing the performance of models using information criteria or other standards, always make sure that the regressions being compared use identical samples. By default Stata will use all of the available observations for every regression, so shortening lags will increase the number of observations used and make the numbers incomparable unless you manually set a fixed sample. Use the "if" or "in" clause in your estimation command to specify the sample.
- When estimating models, test whether the error term is autocorrelated using a Breusch-Godfrey test or a Box-Ljung  $Q$  test (or both). If appropriate, use Newey-West (HAC) standard errors or GLS in your analysis to correct for autocorrelation. (Note: If you use the `prais` command in Stata and then use `predict`, `resid` to get

residuals, the residuals will be the *untransformed* residuals: in terms of our class notation, they will be  $\hat{u}_t$ , not  $\hat{u}_t^* = \hat{u}_t - \hat{\rho}\hat{u}_{t-1}$ . These residuals *will* have autocorrelation even if the transformed ones do not, so you must transform them before you test them for autocorrelation. There does not appear to be an option in the predict command to allow you to get the transformed residuals.)

- When you find the point estimates for the restricted  $\beta_s$  parameters in the linear restricted lag model, find and discuss their standard errors and  $t$ -statistics as well. You may also want to explore some of the other restricted lag models discussed in the instructor's Chapter 3. Whenever you estimate a restricted lag model, test the hypothesis that the various restrictions on the lag-coefficients hold.
- If you find that you need the Newey-West standard errors or an AR(1) error, then in order to test hypotheses you should use the test command after running the newey or prais command in order to use the appropriate estimates and SEs in the test. Do not use the familiar test based on  $SSE_r - SSE_u$  because it assumes that there is no heteroskedasticity or autocorrelation.
- If there are alternative specifications that suggest themselves, go ahead and explore them.
- Note that you will need to create a "year" variable in Stata in order to define the dataset as time series. Since we don't know the year numbers, just use `gen year = _n` to create one starting at 1. Once you have done this, use `tsset year` to declare year to be your time variable.

Your report should be a short paper describing your results, not a set of answers to specific questions. Provide a narrative of the specifications you tried (including the results) and why you chose to go in the directions you did. Once you have a preferred model, examine the dynamic effects of price on area planted in detail, presenting point and interval estimates for the dynamic marginal effects and the cumulative effects. Summarize your results in a conclusion with an assessment of their validity.