Economics 706 Preliminary Examination
June 2018
F.X. Diebold

Do all questions, providing detail and discussion as appropriate. That is, don't just state "answers"; instead, derive and motivate and interpret answers insofar as possible. Note that the exam has TWO pages. WRITE CAREFULLY AND CLEARLY. Good luck!

1. (20 points) State and prove the Granger Representation Theorem in as much generality as you can. Be sure to define all notation, matrix dimensions, etc., and be sure to prove both necessity and sufficiency.

2. Consider the Markov process,

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

$$\varepsilon_t \overset{iid}{\sim} \mathrm{N}(0, \sigma^2).$$

   (a) (20 points) Rewrite the process with Δy_t on the left, and y_{t-1} and Δy_{t-1} on the right. Under what condition(s) is the process $I(1)$? Characterize the distribution of the Dickey-Fuller "studentized-statistic" associated with the estimated coefficient on y_{t-1} in the unit root case.

   ****

   From this point onward through the rest of the exam, assume that the covariance-stationarity condition(s) is(are) satisfied.

   ****

   (b) (20 points) Compute the process' autocovariance function, spectral density function, and Wold representation. What is the relationship between the autocovariance function and the Wold representation? What is the relationship between the spectral density function and the Wold representation? In large samples, how is the variance of the sample mean of a realization of the process related to its spectrum?

(c) (20 points) Do the innovations associated with the process' Wold representation necessarily have constant conditional variance? Is the innovation conditional variance necessarily smaller than the unconditional variance? Why or why not? Are the innovations conditionally Gaussian? Unconditionally Gaussian? Unconditionally leptokurtic? Covariance stationary? Strictly stationary?

(d) (20 points) How would you obtain exact maximum likelihood parameter estimates? How would you obtain exact Bayesian posterior parameter estimates (using quadratic loss and a natural conjugate prior)? In each case, what role does the Kalman filter play? In each case, what role does simulation play? How are the two estimates related in small samples? In large samples?