Problem Set 3

The data set extract.dta has data for 9,473 young men from an administrative data set, with a daily wage in 9 consecutive years in the period between 1999 and 2009. The sample has these features:

1) males only
2) have a job every year in the 9-year interval

Some of the person-specific intervals start in 1999, others in 2000 and others in 2002. I will call the person specific year index **X**: this runs from 1 to 9 for each person

<u>The variables in the file are as follows:</u>
age1 -- age9 = age in year **X**

year1 - year9 = actual calendar year in year **X**

h1 - h9 = total days worked in year **X**

g1 - g9 = total earnings in year **X** (in real Euros)

schooling = person's schooling

logg1 - logg9 = log earnings in year **X**

logw1 - logw9 = log average earnings per day in year **X**

For this problem set you will use data on logw1-logw9, and the other variables, to estimate a variance components model for wages. You will need to use Matlab or Gauss or MATA. You will also want to read the appendix to Abowd-Card, ECA, 1989, and take a look at the first half of the paper for a general overview.

Look at program newcovs.sas This is sas code that creates residual wages in each year for each person (regressing on year dummies and a quadratic in experience). These are called r1-r9. It outputs the 45 independent ("LTR" = lower triangular) elements from the covariance matrix of r1-r9, as well as the 45 x 45 matrix of sampling variances/covariances of the second moments. These are placed in 2 .txt files called rescov.dat and vrescov.dat. You can replicate these files in STATA if you are a stata user, or use the rescov and vrescov data files directly in MATLAB

QUESTIONS

a) Develop a model for the variance-covariance terms that includes a permanent person effect, a first-order autoregressive transitory error, and a pure measurement error that is uncorrelated over time. I.e., assume:

$$r_{it} = \omega_i + u_{it} + e_{it}$$

where e_{it} is iid and u_{it} is AR(1):

$$u_{it} = \alpha\, u_{it-1} + \varsigma_{it}$$

You can assume that var[ς_{it}] is constant over time but be sure to allow an unrestricted covariance for the pre-sample observation u_{i0} to allow for the non-stationarity in the variances over time. (Note that var[r_{it}] is "U-shaped") .

(i) Fit the model by non-linear least squares (or equally weighted minimum distance).

(ii) Develop a test statistic for the goodness of fit of the model.

b) Take first differences of wages, residualize these, and fit an autocovariance model for the change in log wages. Can you reject a random walk model with non-stationary innovations?