Problem Set 4

The data set newwagedata.dta has data for 22,124 young men from an administrative data set, with a daily wage in every year from 2002 to 2009. The sample has these features:

1) males only
2) have a job every year
3) experience = [age - mean(education) - 8] is between 1 and 4 in year 1

I averaged reported education over the 8 possible years so meaneducation has some non-integer values.

<u>The variables in the file are as follows:</u>
age = age in 2009
exp1-exp8 = experience in years 1-8 (=2002-2009)
jnum1-jnum8=counter for which job (everyone starts with jnum1=1)
id = id variable

w1=log wage in 2002 …
w8 = log wage in 2009

There are also two variables measuring firm size and mean education at the firm in each year.

For this problem set you will use data on w1-w8, 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 schooling and cubic in experience). These are called r1-r8. It outputs the 36 independent ("LTR" = lower triangular) elements from the covariance matrix of r1-r8, as well as the 36 x 36 matrix of sampling variances/covariances of the second moments. These are placed in 2 .txt files called rescov2.dat and vrescov2.dat. You can replicate this is STATA if you are a stata user, or use the sas code.

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. Fit the model by non-linear least squares (or equally weighted minimum distance), using a variation of simple.prg.

b) Try modifying your model by assuming that

$$r_{it} = \psi_t (\omega_i + u_{it}) + e_{it}$$

where ψ_t is allowed to vary over time. How do the estimates of the ψ_t terms change over time?