Econ 712, Fall 2018
Problem Set 2 (Due Sept 20)

For this problem set use the same data we used in the first problem set, from the Griliches-Mairesse paper. The notation below follows the notation from the first problem set.

1. In this question we will explore various estimators suggested in the dynamic panel literature.

   (a) Estimate the following model

   $$ldsal_{it} = \beta_1 lemp_{it} + \beta_2 ldnpt_{it} + \beta_3 ldrst_{it} + \gamma_t + \delta_t d357 + \alpha_i + \omega_{it} + \varepsilon_{it}$$

   where ω_{it} is not serially correlated but is "transmitted" (i.e., correlated with the current choice of labor and future choice of the other inputs), and α_i is a firm effect. First-difference the data (to get rid of the firm effect) and use lagged values of the inputs as instruments.

   (b) Estimate the same model without the firm (fixed) effect, i.e.

   $$ldsal_{it} = \beta_1 lemp_{it} + \beta_2 ldnpt_{it} + \beta_3 ldrst_{it} + \gamma_t + \delta_t d357 + \omega_{it} + \varepsilon_{it}$$

   but now let $\omega_{it} = \rho\omega_{it-1} + v_{it}$ and continue to assume that it is "transmitted". Quasi- difference the data and use lagged values of the inputs and output as instruments. Note, that you might need to do some of your own programming here. The built in STATA Arellano-Bond command, *xtabond*, does not allow for serial correlation. The command *xtdpd* does, but is a bit more complicated. (Other related commands are *xtdpdsys* and *xtabond2*, which you need to download and instal). You can either program this outside of STATA or you can write a loop that for each value of ρ computes an IV regression, takes the residual and interacts it with lagged output. You then choose the ρ that sets this last moment to zero (or at least as close as possible).

   (c) Add a fixed effect to the above model

   $$ldsal_{it} = \beta_1 lemp_{it} + \beta_2 ldnpt_{it} + \beta_3 ldrst_{it} + \gamma_t + \delta_t d357 + \alpha_i + \omega_{it} + \varepsilon_{it}$$

   where $\omega_{it} = \rho\omega_{it-1} + v_{it}$. Difference the quasi-differences and then use lagged values of the inputs as instruments.

   (d) Comparing the results from a-c what do you conclude about the importance of the firm (fixed) effect? About serial correlation in ω_{it}?

2. We will now explore Olley-Pakes type estimators.

   (a) Compute an Olley-Pakes like estimator by computing the following steps.

      i. Regress $ldsal_{it}$ on $lemp_{it}$, the dummy variables and a 2nd order polynomial in $ldnpt_{it}$, $ldrst_{it}$, and $ldinv_{it}$. Report the coefficients on $lemp_{it}$ and the dummy variables.

      ii. To compute the remaining coefficients define $\widehat{h}_{it} \equiv \widehat{\pi}_{it} - \beta_2 ldnpt_{it} - \beta_3 ldrst_{it}$, where $\widehat{\pi}_{it}$ denotes the polynomial estimated in (a). Use NLLS to minimize the sum of squares of the following residuals:

      $$\xi_{it} = \widehat{\pi}_{it} - \beta_2 ldnpt_{it} - \beta_3 ldrst_{it} - b_1\widehat{h}_{it-1} - b_2\widehat{h}^2_{it-1}$$

      Report the estimates of β_2 and β_3. Note that $\widehat{h}_{it-1}$ is a function of β_2 and β_3, and therefore you need to use non-linear-least-squares and not just OLS.

   (b) Use a Probit model to estimate the probability that the firm exists in t as a function of $ldnpt_{i,t-1}$, $ldrst_{i,t-1}$, and $ldinv_{it}$. Denote by $\widehat{P}_{it}$ the predicted probability from this model. Repeat (ii) but now *instead* of $\widehat{h}_{it-1}$ and $\widehat{h}^2_{it-1}$ include $\widehat{P}_{it-1}$ and $\widehat{P}^2_{it-1}$.

   (c) Repeat (ii) but include a 2nd order polynomial in *both* $\widehat{P}_{it-1}$ and $\widehat{h}_{it-1}$.

(d) Comparing the results from a-c and from question 1, what do you conclude about the importance of the firm (fixed) effect? About allowing for a more general Markov process in ω_{it}?

Note:
(1) Working in groups on these numerical problem sets is fine, and encouraged. All members of a group should ultimately do the calculations and hand them in individually.
(2) When asked to report results present the answer in a table. Nothing fancy but don't simply attach a printout of the statistical program you used. You should attach the code you used to generate the results as an appendix.