Problem set 1
607: Applied Macroeconometrics
Jon Faust
Fall 2016

The following is due at the beginning of next class. You can turn in any paper in my mailbox or in class or by email. You may work in groups; hand in a single submission for the group. The submission should list those who contributed.

1. Review. Review the 'Econometric and Statistical Basics' (linked with the problem set on website). For next week, define the following items from the basics:

   (a) Generalized least squares, feasible GLS

   Answer/comment
   In std. notation, $Y = X\beta + \varepsilon$, $E\varepsilon = 0$, $E\varepsilon\varepsilon' = \Sigma$, the GLS estimator of β is,

   $$\hat{\beta}_{GLS} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$$

   and under standard (e.g., fixed Xs) assumptions have vcov matrix,

   $$(X'\Sigma^{-1}X)^{-1}$$

   Feasible GLS substitutes an estimator $\hat{\Sigma}$ for Σ.

   As a rule of thumb, if we have made the assumptions such that OLS satisfies Gauss-Markov with $\Sigma = I\sigma^2$, then GLS will have that property with general Σ. Sketch of proof: premultiply the DGP equation by Φ s.t. $\Phi\Sigma\Phi' = I$. The new system is

   $$\tilde{Y} = \tilde{X}\beta + \tilde{\varepsilon}$$

   , where, e.g., $\tilde{Y} = \Phi Y$. This system now satisfies the conditions of the Gauss-Markov thm.

   Feasible GLS will also share the asymptotic properties of GLS if the estimate of Σ is good enough.

(b) Instrumental variables estimator (take the just identified case)

Answer/comment

Same model as above. Now presume that some of the x_t's may be correlated with ε_t. An instrumental variable is one correlated with the problematic x_ts, but not ε_t. Let's posit,

$$X = Z\gamma + u$$

and Z follows some well-behaved process unrelated to u and *eps*. Note it is standard (at least in this class) to call an instrument *relevant* if it is correlated with with the regressor in question and *valid* if it is not correlated with the error term. Thus, we'd like valid and relevant instruments.

If the number of Zs and Xs is the same (that is, the no. of columns is same), the IV estimator is,

$$\hat{\beta}_{IV} = (Z'X)^{-1}Z'Y$$

Note, following the discussion above, we can think about this as transforming our equation back to one that is a simple case. In this case, the violation of the classical assumption of OLS is correlation of regressors and errors. How could we fix this? Suppose we pre-multiply the DGP by

$$\Phi = Z(Z'Z)^{-1}Z'$$

Now consider the implied tilde model as in the discussion of GLS:

$\tilde{Y} = \Phi Y$, $\tilde{X} = \Phi X$ and $\tilde{\varepsilon} = \Phi\varepsilon$.

The OLS estimator on the tilde model is,

$$(X'\Phi'\Phi X)^{-1}(X'\Phi'\Phi Y)$$

which through the miracle of cancellation is the IV estimator. The new regressor $\tilde{X}$ is,

$$\tilde{X} = Z[(Z'Z)^{-1}Z'X]$$

You should recognize $\tilde{X}$ as the projected value of X from an OLS regression of X on Z. You should be able to produce the argument why this tilde regressor will (in the limit) be uncorrelated with the $\tilde{\varepsilon}$.

Under standard assumptions the asymptotic vcov matrix of IV will be

$$Q_{ZX}^{-1} Q_{ZZ} Q_{XZ}^{-1} \sigma_\varepsilon^2$$

where

$$Q_{AB} = \text{plim} T^{-1} A'B$$

Note: for each of these, give the textbook formula for the estimator for the slope coefficients and for the variance-covariance matrix of the estimates.

2. The DGP is an AR(1)

$$y_t = \mu + \rho y_{t-1} + \varepsilon_t$$

$\varepsilon_t \sim iidN(0,1)$

(a) Give expressions for the variance of y and the first two autocorrelations in terms of the parameters of the problem.

Answer/comment

First, the autocorrelations are the autocovariances divided by the variance of the process. For the variance, we need the mean.

$$Ey_t = \mu + \rho Ey_{t-1} + E\varepsilon_t$$

Since $E\varepsilon_t = 0$, and under stationarity the mean is constant,

$$Ey_t = \mu + \rho Ey_t$$

or

$$Ey = \mu/(1-\rho)$$

where I drop the t on the LHS to emphasize time invariance. Rememeber, if μ is the constant in the AR(1), it is not the mean of the process.

Now we can write y_t minus its mean as

$$y_t - Ey = \rho(y_{t-1} - Ey) + \varepsilon_t$$

and

$$\text{var}(y_t) = E(y_t - Ey_t)^2 = \rho^2 E(y_{t-1} - Ey)^2 + \sigma_\varepsilon^2$$

plus a twice a covariance term which is zero because $E\varepsilon_t(y_{t-1} - Ey_t) = 0$.

If the variance is time invariant, then

$$\text{var}(y_t) = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

Alternatively, we could have gotten the variance by solving for the moving average representation:

$$y_t = \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j}$$

Since the LHS is a weighted sum of uncorrelated stuff, we have,

$$\text{var}(y_t) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \rho^{2j}$$

which gives the same answer as before.

This latter clearly relies on $|\rho| < 1$ for convergence of the sum. This is precisely the stationarity constraint.

Now we need the autocovariances, which we will divide by the variance to get the ultimate goal, the autocorr. Autocovariance j: We want

$$\sigma(j) = E(y_t - Ey)(y_{t-j} - Ey) = E\rho(y_t - Ey)(y_{t-j} - Ey)$$

Expand $y_{t-j} - Ey$ as $\rho(y_{t-j} - Ey) + \varepsilon_{t-j}$, multiply by y_t and take expectations,

$$E(y_t - Ey)(y_{t-j} - Ey) = \rho E(y_t - Ey)(y_{t-j} - Ey)$$

where the cross product term with ε has expectation zero. Under cov. stat., we have

$$\sigma(j) = \rho\sigma(j-1)$$

And in particular,

$$\sigma(1) = \rho\sigma(0)$$

where $\sigma(0) = \sigma_y^2$. Thus, autocorrelation j is ρ^j.

(b) Suppose we have the value for an initial condition, $y_0 = \xi$. Write an expression for y_1 in terms of the initial condition and ε. Do the same for y_2, and then for any y_t, $t \geq 1$.

Answer/comment

$$\begin{aligned} y_1 &= \mu + \varepsilon_1 + \rho\xi \\ y_2 &= (1+\rho)\mu + \varepsilon_2 + \rho\varepsilon_1 + \rho^2\xi \\ y_3 &= (1+\rho+\rho^2)\mu + \varepsilon_3 + \rho\varepsilon_2 + \rho^2\varepsilon_1 + \rho^3\xi \end{aligned}$$

And you should see that,

$$y_t = \mu(\sum_{j=0}^{t-1} \rho^j) + \sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j} + \xi \sum_{j=1}^{t} \rho^i$$

(c) You estimate an AR(1) model by OLS:

$$y_t = \alpha + \beta y_{t-1} + u_t$$

and u_t is the residual. Is the OLS estimator of α and β unbiased? Why or why not?

Answer/comment
If we define X as the $(T - 1 \times 2)$ matrix with a column of ones and a column with $y_1, \ldots, y_{t-1}$ and Y as the column vector with $y_2 \ldots y_T$, then the model can be written

$$Y = X\beta + U$$

where $\beta = (\alpha\rho)'$ and

$$\hat{\beta} - \beta = (X'X)^{-1}X'U$$

and the bias is,

$$E\hat{\beta} - \beta = E(X'X)^{-1}X'U$$

Note that as in the standard OLS case $Eu_t|x_t = 0$ so that $EX'U = 0$. If X were fixed or independent of U, then this would deliver unbiased $\hat{\beta}$. But the X matrix includes all the u_ts except the last one. Thus, this expectation is not zero.

(d) The DGP is the AR(1) described above. You estimate the following model by OLS:

$$y_t = \alpha + \beta\tau_t + u_t$$

where $\tau_t = t$, that is τ is a deterministic, linear time trend. And u_t is the residual.

Is the OLS estimator of α and β unbiased?

Answer/comment
Following the logic from above, it is unbiased since the two regressors are both fixed.

---

3. Continuing with the AR(1) expression from last problem.

If we set $\rho = 1$ and $\mu = 0$, the process for y_t is called a random walk (in this case, due to Gaussian innovations, we might more fully say a Gaussian random walk). With $\mu \neq 0$, it is a random walk with drift. For this problem, take the case of the Gaussian random walk without drift.

(a) What is the unconditional expectation of y_t (that is, you don't have y_0).

Answer/comment
See next answer.

(b) What is the expectation of y_t, $t \geq 1$ conditional on $y_0 = \xi$?

Answer/comment
Note that from any starting value, say, y_0,

$$y_t = y_0 + \sum_{j=1}^{t} \varepsilon_t$$

Thus, conditional on $y_0 = \kappa$,

$$Ey_t | y_0 = \kappa$$

As for the unconditional mean, that's undefined here because we have not defined how the process started. We'd need to supply at least an expectation for the y_s at some $s < t$ to answer this or some way to derive that expectation. Or we could specify that the process goes infinitely back in time so that

$$y_t = \sum_{j=0}^{\infty} \varepsilon_{t-j}$$

in which case the mean is zero.

(c) Suppose we have an initial condition, $y_0 = \xi$. Give an expression for the conditional variance of y_t given the initial condition. That is, give

$$E((y_t - \hat{y}_t)^2 | y_0 = \xi),$$

where

$$\hat{y}_t = E(y_t | y_0 = \xi)$$

Answer/comment

The mean for each t is ξ, so using that $y_t = \xi + \sum_{j=1}^{t} \varepsilon_t$

$$E((\xi + \sum_{j=1}^{t} \varepsilon_t) - \xi)^2 = t\sigma^2$$
