POLS571 - Longitudinal Data Analysis
October 4, 2001

Cointegration

1 Introduction

Consider the series:

$$Y_t = \rho Y_{t-1} + u_t \tag{1}$$

where $u_t \sim i.i.d.(0, \sigma^2)$. We know that if $|\rho| < 1$, the series is *stationary* (that is, I(0)), while if $\rho = 1$ then the series is *integrated* (I(0)). There are a few other relevant properties of I(0) and I(1) series that will be useful to us today:

1. If Y_t is I(0), then a linear transformation of Y_t (e.g., any transformation along the lines of $a + bY_t$) will also be I(0).
2. The same is true for I(1) series: if Y_t is I(1), then $a + bY_t$ is also I(1).
3. If two series X_t and Y_t are I(0), then any linear combination $aX_t + bY_t$ is also I(0).
4. If X_t is I(0) and Y_t is I(1), then the combination $aX_t + bY_t$ is I(1); that is, *integration dominates stationarity.*

Finally,

- If both X_t and Y_t are I(1), then $aX_t + bY_t$ is ***generally*** I(1) as well.

however...

Suppose we have the following situation:

$$
\begin{aligned}
X_t &= W_t + u_{Xt} \\
Y_t &= AW_t + u_{Yt}, \\
& W_t \sim I(1), \\
& u_{Xt}, u_{Yt} \sim I(0)
\end{aligned}
$$

Under these circumstances, by (4) above, both X_t and Y_t will be I(1), BUT:

$$
\begin{aligned}
Z_t &= Y_t - AX_t \\
&= [AW_t + u_{Yt}] - A[W_t + u_{Xt}] \\
&= AW_t + u_{Yt} - AW_t - Au_{Xt} \\
&= u_{Yt} - Au_{Xt} \qquad (2)
\end{aligned}
$$

which is I(0) by (3), above.

The reason that the combination of the two nonstationary series yields a stationary product is because the nonstationarity arises from a *common component* (W_t) which is I(1). More generally,

*If there exist two series I(1) X_t, Y_t such that $Z_t = \mu + aX_t + bY_t$ is I(0), then X_t and Y_t are said to be **cointegrated**.*

1.1 Intuition: Cointegration as an Attractor

The fact that cointegration implies a common factor leads to one interpretation of cointegration: an *attractor*. In the equation above, the line:

$$X_t = AY_t \qquad (3)$$

is an attractor, in the sense that $Z_t = X_t - AY_t$ will be mean-zero white noise around it. Thus, Z_t will never be "far" from $X_t = AY_t$.

Two examples:

(a) Murray (1994) talks about drunks and dogs. Both drunks and puppies engage in "random walks", in that their behavior is random and its variance is increasing over time. But, the two will never be far from *one another*; that is, the *distance between them*, while random, is stationary around some mean.

(b) Consider prices for the same commodity in different parts of the country. Principles of supply and demand, along with the possibility of arbitrage, means that, while the prices may fluctuate more-or-less randomly, the distance between them will, in equilibrium, be relatively constant (typically around zero).

1.2 Error Correction Models

Recall that if both X_t and Y_t are I(1), then ΔX_t and ΔY_t are I(0). Suppose we write:

$$\begin{aligned} \Delta X_t &= \alpha_1 + \rho_X Z_{t-1} + \gamma_{1X}\Delta X_{t-1} + \gamma_{2X}\Delta Y_{t-1} + \epsilon_{Xt} \\ \Delta Y_t &= \alpha_2 + \rho_Y Z_{t-1} + \gamma_{1Y}\Delta X_{t-1} + \gamma_{2Y}\Delta Y_{t-1} + \epsilon_{Xt} \end{aligned} \tag{4}$$

with:

$$\begin{gathered} \epsilon_X, \epsilon_Y \sim IN(0, \sigma^2) \\ Z_t = X_t - AY_t, \text{ and} \\ \text{At least one } \rho \neq 0 \end{gathered}$$

This is what is known as an *error-correction model* (ECM). In this arrangement,

- If X_t and Y_T are cointegrated, then Z_t is I(0) and all the elements of (4) are stationary.
- If X_t and Y_t are not cointegrated, then Z_t is I(1). This means that Z_t can't explain either ΔX_t or ΔY_t (which are I(0)), which means that $\rho_X = \rho_Y = 0$, which is excluded by assumption.

This means that cointegration implies an error-correction model (in fact, cointegration is a necessary condition for an ECM to hold). We'll discuss ECMs more on Tuesday.

2 Cointegration in practice

2.1 How to do it

Cointegration requires that both variables in question be I(1), but that a linear combination of them be I(0). This means that the first step is to figure out if the series themselves are I(1), using unit root tests. *If one or both are not I(1), cointegration is not an option.*

Next, consider that we can rewrite:

$$Z_t = X_t - \alpha Y_t \tag{5}$$

as:

$$X_t = \alpha Y_t + Z_t \tag{6}$$

which looks strikingly like a regression equation (in fact, its called the *cointegrating regression*). This means that one alternative for the next step is to estimate (6), and generate predicted values $\hat{Z}_t$. If the two series are cointegrated,

$$\begin{gathered} Z_t \sim \text{I}(0), \text{ and} \\ \text{Var}(Z_t) < \infty. \end{gathered}$$

This suggests a general strategy for examining cointegrated series:

1. Determine that the two series are I(1),
2. Estimate $X_t = \hat{\alpha} Y_t + Z_t$ using OLS,
3. Examine Z_t for stationarity, using
    - Durbin-Watson test ("CRDW")
    - Standard unit root tests (DF, ADF, KPSS)

If the $\hat{Z}_t$ do not have a unit root, that is evidence that the two series are cointegrated.

"Wait a minute," you say, "how can I use OLS for the cointegrating regression? What about AR problems? Simultaneity? Etc.?"

Recall that if X_t and Y_t are cointegrated, then $\text{Var}(Z_t) < \infty$. Suppose we choose some method for estimating $\hat{\alpha}$ that gets it wrong, and we come up with some incorrect value (call it δ):

$$\tilde{Z}_t = X_t - \delta Y_t \tag{7}$$

In these circumstances, $\tilde{Z}_t$ will be I(1), and its asymptotic variance will be infinite. Moreover, in finite samples (that is, the kind we usually work with...):

$$\text{Var}(\tilde{Z}_t) > \text{Var}(\hat{Z}_t)$$

Now, recall that OLS estimates $\hat{\alpha}$ by minimizing the residual variance. This means that, since α is the coefficient that makes Z_t stationay (and makes $\text{Var}(Z_t)$ finite), OLS will *always* fine the "right" coefficient. That is, in the cointegrating regression context, OLS is unbiased and consistent:

$$\hat{\alpha}_{OLS} \rightarrow \alpha \text{ as } T \rightarrow \infty$$

Moreover, where as, in "standard" OLS, $\hat{\alpha}_{OLS} \rightarrow \alpha$ at a rate of $\frac{1}{\sqrt{T}}$, in a cointegrating regression this converges at a rate of $\frac{1}{T}$; sometimes this is referred to as being "superconsistent".

2.2 An Example

We'll again look at the Democratic House and Senate percentages data, 1789-1990. Its not unreasonable to believe that (a) the two series share a common component, but (b) each will also have some idiosynchratic variability as well. That is, we might expect the two series to be cointegrated.

Since we've already determined that the series are both I(1), we can proceed to the cointegrating regression:

```
.  reg dhpct dspct
```

which yields:

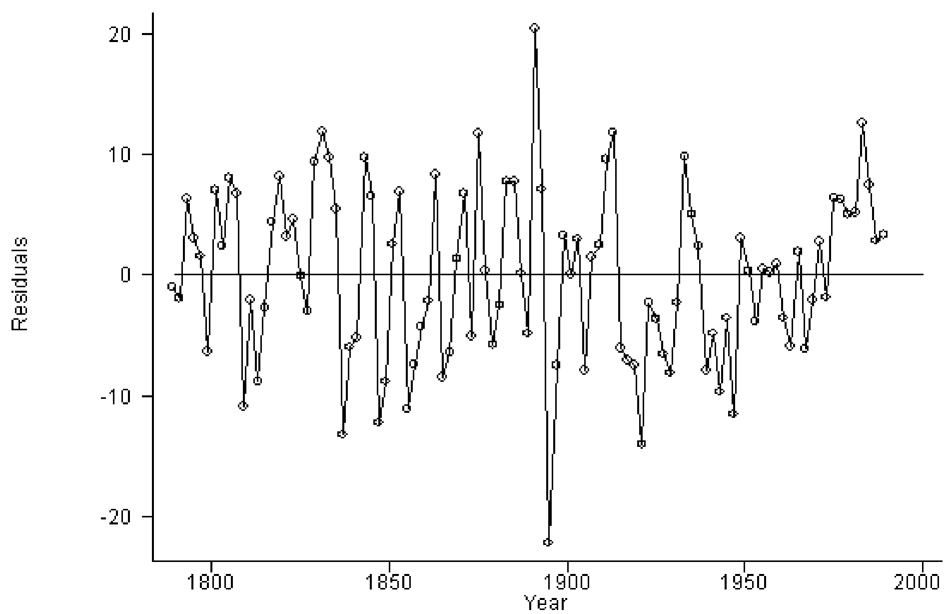
$$
\begin{gathered}
\text{House Percentage} = 16.461 + 0.728(\text{Senate Percentage}) + z_t \\
R^2 = 0.69 \\
F = 222.83 \ (p < .001) \\
N = 101
\end{gathered}
$$

We can then generate residuals:

```
. predict zhat, resid
```

which are graphed in Figure 1.

Figure 1: Residuals from the Cointegrating Regression



We can then do standard unit root tests on the residual series $\hat{Z}_t$, e.g.:

```
. unitroot zhat
. kpss zhat
```

where we find that we can clearly reject the null of no integration (in the first instance) but cannot reject the null of stationarity (in the KPSS test). As an aside, the same thing as what we just did, minus the KPSS test, can be accomplished by one command:

```
. coint dhpct dspct
```

which estimates the cointegrating regression and does the unit root tests all in one.

Next time: Error Correction Models...