

Models for non-stationary series - a very brief, albeit useful, intro

A time series  $y_t$  is stationary if, roughly, its features are time invariant.

In particular,

$$E(y_t) = \mu \text{ not a function of time!} \quad (1)$$

$$\text{Var}(y_t) = \sigma^2 \text{ not a function of time} \quad (2)$$

$$\text{Cov}(y_t, y_{t-k}) = \kappa_k \text{ not a function of time} \quad (3)$$

Properties of stationary series

- Stationary series revert in the long run to their mean
- Best long term forecast of series is the mean
- $\lim_{k \rightarrow \infty} \text{Cov}(y_t, y_{t-k}) = 0$
- Series shows many crossings of its mean
- Series looks jagged

Not all series are stationary. Some series are called I(1) (integrated of order 1). These are non-stationary series whose first difference is stationary.

A simple I(1) series is the “random walk,” where

$$y_t = y_{t-1} + \epsilon_t, \epsilon \text{ is stationary} \quad (4)$$

so, by repeated substitution

$$y_t = \epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \dots \quad (5)$$

Note: series is called integrated because it is sum (integration) of all previous values of shocks

Properties of random walks (and integrated series in general)

- $\lim_{t \rightarrow \infty} \text{Var}(y_t) = \infty$
- $\lim_{k \rightarrow \infty} \text{Cov}(y_t, y_{t-k}) \neq 0$
- Best long run forecast is the value of the series today
- Expected time for crossing of mean is infinite

- Series looks very smooth
- Series looks a lot like a trend line, though not straight

Integrated series wreak havoc, since all asymptotics require that variances stay finite and covariances go to zero. Integrated series have non-standard statistics

Testing for stationarity:

Simplest is Dickey-Fuller

Regress  $y_t$  on  $y_{t-1}$  and see how far estimated  $\rho$  is from one. Need to use special tables - see MacKinnon for best. Lots of details on this test, depending on the alternative hypothesis. Note this is an odd test, since we go with non-stationarity if we cannot reject the null that  $\rho = 1$ , which is odd (if not to say backwards).

If we use a non-stationary series in an ADL model, we get an estimate on lagged dependent variable of 1. This means that OLS is no longer correct.

What to do?

Old-fashioned - take first differences to get stationarity. Note that this throws out the long run properties of the series, reducing us to modelling short term effects.

A much nicer idea, “cointegration,” is due to Granger.

Two series may each be integrated, but their difference may be stationary. Loosely speaking, each wander quite far from its mean, but the two series wander very near each other.

If  $x$  and  $y$  are cointegrated, then, by the Granger Representation Theorem, we can model  $y$  and  $x$  as being in an error correcting relationship. This type of model, due to Davidson, Hendry, Srba and Yeo (hence the DHSY) model, is very nice. Basically, it has  $y$  and  $x$  being in an equilibrium relationship, with the short run behavior of  $y$  being a function of the short run behavior of  $x$  and an equilibrating factor.

The EC model is:

$$\Delta y_t = \beta \Delta x_t + \rho(y_{t-1} - \gamma x_{t-1}) + \epsilon_t \quad (6)$$

where  $\epsilon$  is stationary (test for this with Dickey-Fuller).  $\beta$  is the short run effect of  $x$  on  $y$  and  $\rho$  is the speed of equilibration, with  $\gamma$  giving the long run relation between  $y$  and  $x$ .

The EC model is to be highly commended, even if series have long, although not infinite, memories. (A series has a long memory if long run forecasts of it are still a function of current values.)