Lecture 13 – Modeling Trends
(Reference – Section 9.1, Hayashi)

We say that a time series y_t has a (time) trend if the s-step ahead forecast, i.e., $E(y_{t+s}|y_t, y_{t-1},...)$ has a time-dependent component that does not go to zero as s gets large.

For instance, if

$$y_t = a + bt + u_t\ ,\ \ b \neq 0$$

where u_t is a zero-mean stationary and ergodic process, then

$$E(y_{t+s}|y_t, y_{t-1},...) = a + b(t+s) + E(u_{t+s} \mid y_t, y_{t-1},\ldots)$$

The last term goes to zero as s goes to infinity, but a+b(t+s) is increasing linearly with s. That is,

$$E(y_{t+s}|y_t, y_{t-1},...) \approx a + b(t+s) \text{ for sufficiently large s.}$$

In this case, we say that y_t has a **deterministic (linear) trend**, which is equal to a+bt..

More general models with deterministic trends can be formulated by replacing a+bt with other functions of t (e.g., higher order polynomials in t, such as $a + bt + ct^2$).

A stochastic process that is stationary around a deterministic trend is called a **trend stationary** process.

Note that one important characteristic of a trend stationary process is that the s-step ahead forecast error variance

$$\text{Var}(y_{t+s} - E(y_{t+s}|y_t, y_{t-1},...)) = \text{Var}(E(u_{t+s} \mid y_t, y_{t-1},\ldots))$$

converges to $\sigma_u^2 = \text{Var}(u_t)$ as s gets large, which is finite provided that u_t is covariance stationary. *The long-run forecast error variance of a trend stationary process is bounded by the variance of the stationary term.*

Part of the appeal of the trend-stationary model is that it is very straightforward to estimate. Part of the appeal is that it is a straightforward model to interpret.

Why might the trend stationary model not be a valid representation of trending time series?

- the trend and cyclical components of the time series might not be determined independently of one another. (For instance, technology shocks might affect both the cyclical and trend behavior of the series.)
- perhaps the long-run forecast error variance of the process is increasing without bound.

An alternative to the trend stationary assumption for a trending time series is the **difference stationary** assumption, which is the assumption that the series needs to be differenced to make it stationary.

Consider, for example, the random walk process:

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

where ε_t is a zero-mean, i.i.d. sequence. We know that y_t will be nonstationary (since the variance of the random walk is increasing with t), but the first difference $y_t - y_{t-1}$ is stationary.

Let y_0 be given. Then –

$$y_t = y_0 + \sum_1^t \varepsilon_t$$

and

$$y_{t+s} = y_0 + \sum_1^t \varepsilon_i + \sum_{t+1}^{t+s} \varepsilon_i$$

so that

$$E(y_{t+s} | y_t, y_{t-1}, ...) = y_0 + \sum_1^t \varepsilon_i \text{ for all } s \geq 0.$$

The random walk has a **stochastic trend,** $y_0 + \sum_1^t \varepsilon_i$.

Note that the long-run forecast error variance of the random walk,

$$\text{Var}(y_{t+s} - E(y_{t+s} | y_t, y_{t-1}, ...)) = \text{Var}\ (\sum_{t+1}^{s} \varepsilon_i) = (s-t)\sigma^2$$

which is increasing without bound as s increases.

The random walk has a stochastic trend and may be a good starting point for describing the way many financial market prices and returns seem to behave. However, realizations of random walks will not usually be characterized by the tendency to grow over time that is so apparent in many macroeconomic time series. That is, the stochastic trend in the random walk is not sufficient to explain the kind of trend behavior we observe in the typical macroeconomic time series.

Suppose we modify the random walk by adding a constant term, resulting in a model called a **random walk with drift**:

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

where ε_t is a zero-mean, i.i.d. sequence. Note that like the pure random walk, the random walk with drift is a difference stationary process because y_t is nonstationary but its first difference $y_t - y_{t-1}$ is stationary.

Let y_0 be given. Then –

$$y_t = y_0 + bt + \sum_1^t \varepsilon_t$$

and

$$y_{t+s} = y_0 + b(t+s) + \sum_1^t \varepsilon_i + \sum_{t+1}^{t+s} \varepsilon_i$$

so that

$$E(y_{t+s} | y_t, y_{t-1}, ...) = y_0 + b(t+s) + \sum_1^t \varepsilon_i \text{ for all s} \geq 0.$$

The random walk with drift has a stochastic trend, which includes a deterministic component that can account for a time series tendency to increase on average over time. Like the pure random walk, it will be characterized by a long-run forecast error variance that is increasing without bound as the forecast horizon gets sufficiently long.

The random walk with drift is still not quite enough for us, because it assumes that the first differences in y_t are serially uncorrelated. A more general stochastic trend model that better serves our purposes is obtained by replacing the i.i.d. or white noise process ε_t with the stationary process u_t, i.e.,:

$y_t = b + y_{t-1} + u_t$

where u_t is a zero-mean, stationary and ergodic proces.

We can rewrite y_t as:

$$y_t = y_0 + bt + \sum_{1}^{t} u_i$$

and

$$E(y_{t+s} | y_t, y_{t-1}, \ldots) = y_0 + b(t+s) + \sum_{1}^{t} u_i + \sum_{t+1}^{t+s} E(u_i | y_t, y_{t-1}, \ldots)$$

In this case –

- y_t has a stochastic trend
- the average behavior of y_t in the long-run will be determined by the parameter b (which is the (unconditional) expected change in y_t)
- the long-run forecast error variance will be increasing without bound

So now we have we have two approaches to modeling trends in time series. We know that they different implications about the nature of the trend and the long-run behavior of the time series.

Which one should we choose?

- It will not be obvious just by looking at the data.
- Does one or the other seem more plausible based on the economic theory (if there is any) that underlies the econometric model?
- Apply formal tests to help select the appropriate form of the model – **unit root tests** (A problem with this approach will be that tests for a stochastic trend typically have low power against trend stationary alternatives when the stationary component under the alternative has a root that is "close to" one, which seems to be likely with many trending economic time series.)

Digression – Integrated Processes and the Long-Run Variance Condition

Isn't the trend stationary process also a difference stationary process? That is, suppose

$y_t = a + bt + \varepsilon_t$ where, WLOG, we assume ε_t is white noise.

So, y_t is nonstationary (because its mean is increasing with time.) However,

$\Delta y_t = (a+bt+\varepsilon_t)-(a+b(t-1)+\varepsilon_{t-1}) = b + \varepsilon_t - \varepsilon_{t-1}$, is stationary since

$y_t - y_{t-1}$ has an MA(1) form.

The definition of a difference stationary process actually involves a bit more than simply the condition that the first difference of the process is stationary.

Definition – (Difference Stationary, or DS, Process)

A stochastic process y_t is difference stationary if Δy_t is stationary, and the **long-run variance** of Δy_t is strictly positive.

Definition – Long-Run Variance

The long-run variance of a covariance stationary process u_t, is defined to be

$$\lim_{T\to\infty} Var(\sqrt{T}\bar{u})$$

If the MA form of u_t is absolutely summable then

$$\lim_{T\to\infty} Var(\sqrt{T}\bar{u}) = \sum_{j=-\infty}^{\infty} \gamma_j \quad \text{, where } \gamma_j = Cov(u_t,u_{t-j})$$

So, a stochastic process y_t is DS if Δy_t is (covariance) stationary and the sum of autocovariances is finite and nonzero. (The sum of autocovariances cannot be negative, but it can be zero.)

Note, for example, that a white noise process is not difference stationary even though ε_t-ε_{t-1} is stationary because the long run variance of $\varepsilon_t - \varepsilon_{t-1}$ is equal to zero (see below).

Similarly, the trend stationary process is not difference stationary even though the first difference of the trend stationary process is stationary, because the long-run variance of the first difference of the process is zero.

Consider the first difference of the TS process $y_t = a + bt + \varepsilon_t$, ε_t white noise:

$$\Delta y_t = b + \varepsilon_t - \varepsilon_{t-1}$$

Since this is an MA(1) process (with a unit root in the MA polynomial):

$$\gamma_0 = 2\sigma^2,\ \gamma_1 = \gamma_{-1} = -\sigma^2,\ \gamma_j = 0 \text{ for } |j| > 1$$

and

$$\text{long-run variance of } \Delta y_t = -\sigma^2 + 2\sigma^2 - \sigma^2 = 0.$$

Notes –

- Any covariance stationary process whose MA polynomial has a unit root will have a long-run variance equal to 0.
- The assumption that the ε's in the TS process was not important. Differencing any stationary process will introduce a unit root into that process and, given the previous note, will imply that the resulting process has a long-run varaince equal to 0. (Differencing a stationary process is called "overdifferencing.")
- A stationary process with a nonzero long-run variance is called an "integrated of order zero" or "I(0)" process. A nonstationary process that is stationary with a nonzero long-run variance after first differencing is called an "integrated of order one" or "I(1)" process. A nonstationary process that is stationary with a nonzero long-run variance after differencing d times is called an "integrated of order d" or "I(d)" process.
- The following all mean the same thing: unit root process, I(1) process, stochastic trend process, difference stationary process.

What is the relevance of the non-zero long-run variance condition, other than helping us define integrated processes and distinguish between TS and DS processes? It turns out, for example, that this condition plays an important role in developing asymptotic theory for time series regressions that allow for more temporal dependence and heterogeneity than the regression theory we developed allows for.