POLS571 - Longitudinal Data Analysis
September 17, 2001

Unit Roots, and Integration, Part I

1 Unit Roots and Stationarity

We noted the other day that an I(1) series is:

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

and that this series had a number of properties that made it different from AR(1) series where $\phi < 1$. To confuse the notation a bit more, I'm now going to replace ϕ with ρ in the AR(1) context, so that we'll now be talking about (1) as the special case of:

$$Y_t = \rho Y_{t-1} + u_t \tag{2}$$

where $\rho = 1$.

We talked briefly at the beginning of the class about the concept of stationarity; in particular, we noted that, for an AR(1) series to be weakly stationary, we require that $\rho < 1$, and that (2) with $\rho = 1$ is nonstationary. There are plenty of other ways to get nonstationary series, however. Consider the equation:

$$Y_t = \beta t + u_t \tag{3}$$

This series is also nonstationary, in that $\bar{Y}$ is increasing over time. Both of these are "first order" nonstationary equations, in that both can be made stationary by differencing. For (2), of course, differencing yields:

$$\begin{aligned} \Delta Y_t \equiv Y_t - Y_{t-1} &= Y_t + u_t - Y_{t-1} & (4) \\ &= u_t & (5) \end{aligned}$$

while differencing (3) gives:

$$
\begin{aligned}
\Delta Y_t \equiv Y_t - Y_{t-1} &= \beta t + u_t - (\beta(t-1) + u_{t-1}) && (6)\\
&= \beta t + u_t - \beta t + \beta - u_{t-1} && (7)\\
&= u_t - u_{t-1} + \beta && (8)
\end{aligned}
$$

which is also stationary.

The problem with these two models is that differencing by itself doesn't tell you *why* your initial series is not stationary (that is, it can't distinguish between (1) and (3)). The way to address this is to examine:

$$Y_t = \rho Y_{t-1} + \beta t + u_t \tag{9}$$

and test for $H_0 : \hat{\beta} = 0$:

- If we cannot reject $\hat{\beta} = 0$, then this is evidence in favor of the series being a "random walk" without a trend;
- if we can reject $\hat{\beta} = 0$, this suggests that the series has a deterministic trend.

More generally, if we consider (2), there are three possibilities:

- $|\rho| > 1$
  - Series is nonstationary / *explosive*
  - Past shocks have a greater impact than current ones
  - Uncommon
- $|\rho| < 1$
  - *Stationary* series
  - Effects of shocks die out exponentially according to ρ
  - Is mean-reverting
- $|\rho| = 1$
  - Nonstationary series

– Shocks persist at full force

– Not mean-reverting; variance increases with t

In most instances, the bigger problem is distinguishing between (2) with $\rho < 1$ and (1). If the former, then we shouldn't difference the series; if the latter, we should. This problem, of distinguishing between (2) and (1), has given rise to the huge literature on *unit root testing*.

2 The Dickey-Fuller Test

At first blush, telling (2) from (1) would seem to be easy: just estimate:

$$Y_t = \rho Y_{t-1} + u_t \tag{10}$$

and test whether $\hat{\rho} = 1$; if so, then difference; if not, then don't.

If that were all there was to it, it would be simple. The problem is that the distribution of $\hat{\rho}$ is nonstandard under the null hypothesis; this means that, while $\hat{\rho} = \frac{\sum Y_t Y_{t-1}}{\sum Y_{t-1}^2}$ is a consistent estimate of ρ, its distribution is not a standard t-distribution.[1] The distribution it follows is known as the "Dickey-Fuller" (D-F) distribution, after the 1979 paper where it was first derived.

The D-F distribution is:

- Right-skewed (so t-statistics will tend to be large and negative).
- This means that we will tend to *overreject* the null hypothesis if we use the standard t-distribution.

Thus, the "Dickey-Fuller test" for a unit root amounts to estimating $\hat{\rho}$ and doing a standard-looking t-test for $H_0 : \hat{\rho} = 1$, but using a non-standard set of critical values.

[1]Specifically, making use of the math of Weiner processes, $T(\hat{\rho} - 1) \rightarrow \frac{\int_0^1 W(r)dW(r)}{\int_0^1 W(r)^2 dr}$, and so the "t-statistic" $t \equiv \frac{\hat{\rho}}{s.e.(\hat{\rho})} = \frac{\frac{1}{2}[W(1)^2 - 1]}{[\int_0^1 W(r)^2 dr]^{\frac{1}{2}}}$, which is the "Dickey-Fuller distribution.

- If t is greater than the critical values indicated, then the series is stationary;
- If t is less than the critical values, this is evidence of nonstationary.
- That is, the null hypothesis is that the series has a unit root.

These were calculated very precisely by MacKinnon (1991). Also note that the D-F test requires that the us be "white noise"; this will become important next class...

The Dickey-Fuller test is sometimes estimated as:

$$\begin{aligned} \Delta Y_t &= (\rho - 1)Y_{t-1} + u_t && (11) \\ &= \delta Y_{t-1} + u_t && (12) \end{aligned}$$

and the test is then for $\hat{\delta} = 0$. Occasionally, you'll see $\hat{\tau}$ in place of either $\hat{\rho}$ or $\hat{\delta}$, since this was Fuller's (1976) and Dickey and Fuller's (1979) original notation.

3 Dickey-Fuller variants

3.1 Drift

Another possibility is that the series has a *drift*:

$$Y_t = \alpha + \rho Y_{t-1} + u_t \tag{13}$$

the other day, we showed that this can be rewritten as:

$$Y_t = Y_0 + \alpha t + \sum_{t=1}^{T} u_t \tag{14}$$

In this series, Y_t has a constant "drift", which manifests itself as a trend (αt) over time. This renders the series nonstationary all by itself; and over time, this drift will come to dominate the series. Moreover, this series "looks" very different from (1), and its tests have different distributional characteristics as well.

In the series with drift, we need to test both $\hat{\rho} = 1$ and $\hat{\alpha} = 0$; we can also test the joint hypothesis (i.e., that *both* $\hat{\rho} = 1$ and $\hat{\alpha} = 0$). The critical values for these tests are all in MacKinnon (1991).

3.2 Trends

Consider:

$$Y_t = \alpha + \beta t + \rho Y_{t-1} + u_t \tag{15}$$

Estimating this model now explicitly considers the possibility of a deterministic trend in Y.

- This means that α is now a "constant".
- This also requires a (slightly) different set of critical values, also in MacKinnon.
- One can also do F-tests on the joint nulls: $\rho = 1$ and $\beta = 0$

Note as well that, in general and not surprisingly, its worse to *omit* a trend from a model when the data generating process has one, than it is to *include* one where the DGP is trendless. In practice, we include a constant 99.99% of the time, and a trend quite often as well.

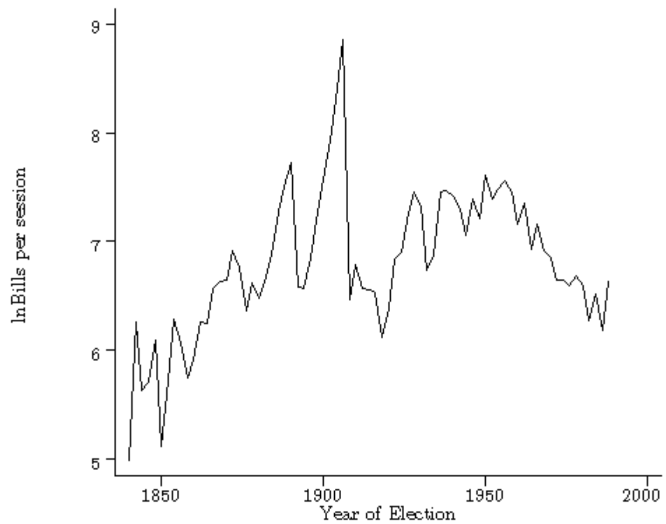
4 An Example: Congressional Activity

We'll examine data on the (logged) number of bills passed by each of the 1st through the 101st Congresses (1789 - 1990, $T = 101$). The series is in Figure 1.

In Stata, there are a couple different ways we can go about doing unit root tests. One way is to "brute force" the test, by creating trends, lagged variables, etc. and estimating the model yourself. So, the most basic unit root test (that is, that based on (11) can be estimated here by:

```
.  sort congress
.  tsset congress
time variable:  congress, 1 to 101
```

Figure 1: Logged Congressional Output of Bills, 1789-1990




```
. lag 1 lnbills
. dif 1 lnbills
. reg D_lnbill L_lnbill, nocons
```

which yields:

$$\begin{aligned} \Delta Y_t &= 0.00000148 Y_{t-1} + u_t \\ & \quad (< 0.00001) \end{aligned} \tag{16}$$

(t-values in parentheses) and for the series with a constant and/or trend as:

```
. gen trend=_n

. reg D_lnbill L_lnbill
```

$$\begin{aligned} \Delta Y_t &= 0.8598 - 0.1313 Y_{t-1} + u_t \\ &= (2.845)(-2.816) \end{aligned} \tag{17}$$

and:

```
. reg D_lnbill L_lnbill trend
```

$$\begin{aligned} \Delta Y_t &= 1.3434 + 0.0049t - 0.2460Y_{t-1} + u_t \\ &= (3.666) \;\; (2.241) \;\; (-3.586) \end{aligned} \tag{18}$$

Compare these results to the table in MacKinnon (1991). There, the critical values for the "no constant", "with constant" and "with trend" models are:

No Constant	$p < .10$	-1.616
	$p < .05$	-1.939
	$p < .01$	-2.566
No Trend	$p < .10$	-2.567
	$p < .05$	-2.862
	$p < .01$	-3.434
With Trend	$p < .10$	-3.128
	$p < .05$	-3.413
	$p < .01$	3.964

This suggests that in the no constant and no trend models we cannot reject the null hypothesis of a unit root in Congressional bills at the $p < .05$ level. However, in the model with trend, we can reject the null; this suggests that the nonstationary we observe in the variable is due to the fact that it is trending, rather than the presence of a unit root.

Alternatively, you can use Stata's built-in time-series commands. There are three:

- `-dickey-` purports to do Dickey-Fuller tests. I've found some bugs in it, though, so don't use it!
- `-dfuller-` does D-F and Augmented D-F tests (more on the latter on Thursday).
- `-unitroot-` also does D-F and ADF tests.

Of these, -dfuller- generates a table for the $Z(t)$ statistic, which is the same as the t-statistic mentioned before. -unitroot- generates a series of values, including the estimate of τ, as well as the $Z(t)$ and $Z(\alpha)$ statistics of Phillips and Perron (more on these later). I generally find -dfuller- to be more flexible, since it allows you to drop the constant term if you care to. In either case, using -dfuller- or -unitroot- yields the same values for $\hat{\tau}$ as the "brute force" method:

```
.  dfuller lnbills, nocon lags(0)
.  dfuller lnbills, lags(0)
.  dfuller lnbills, lags(0) trend
```

or:

```
.  unitroot lnbills, lags(0)
.  unitroot lnbills, lags(0) trend
```

Next time: More on unit roots.