Unit Roots

• An autoregressive process

\[ a(L)y_t = e_t \]

has a unit root if

\[ a(1) = 0 \]

• The simplest case is the AR(1) model

\[ (1 - L)y_t = e_t \]

or

\[ y_t = y_{t-1} + e_t \]
Examples of Random Walks
Random Walk with Drift

• AR(1) with non-zero intercept and unit root

\[ y_t = \alpha + y_{t-1} + e_t \]

• This is same as Trend plus random walk

\[ y_t = T_t + C_t \]
\[ T_t = \alpha t \]
\[ C_t = C_{t-1} + e_t \]
Examples

\[ y_t = 0.1 + y_{t-1} + e_t \]

\[ e_t \sim N(0,1) \]
Optimal Forecasts in Levels

• Random Walk

\[ y_{t+1|t} = y_t \]
\[ y_{t+h|t} = y_t \]

• Random Walk with drift

\[ y_{t+h|t} = \alpha + y_t \]
\[ y_{t+h|t} = \alpha h + y_t \]
Optimal Forecasts in Changes

• Take differences (growth rates if $y$ in logs)

$$z_t = \Delta y_t = y_t - y_{t-1}$$

• Optimal forecast: Random walk

$$z_{t+h|t} = 0$$

• Optimal forecast: Random walk with drift

$$z_{t+h|t} = \alpha h$$
Forecast Errors

• By back-substitution

\[ y_t = y_{t-1} + e_t \]

\[ = y_{t-h} + e_{t-h+1} + \cdots + e_{t+1} \]

• So the forecast error from an h-step forecast is

\[ e_{t-h+1} + \cdots + e_{t+1} \]

• Which has variance

\[ \sigma^2 + \cdots + \sigma^2 = h \sigma^2 \]

• Thus the forecast variance is linear in \( h \)
Forecast intervals

• The forecast intervals are proportional to the forecast standard deviation

\[ \sqrt{h \sigma^2} = \sqrt{h \sigma} \]

• Thus the forecast intervals fan out with the square root of the forecast horizon \( h \)
Example: Random Walk
General Case

• If \( y \) has a unit root, transform by differencing

\[
z_t = \Delta y_t = y_t - y_{t-1}
\]

• This eliminates the unit root, so \( z \) is stationary.

\[
a(L)y_t = e_t
\]

\[
a(L) = b(L)(1 - L)
\]

\[
b(L)z_t = e_t
\]

• Make forecasts of \( z \)
  – Forecast growth rates instead of levels
Forecasting levels from growth rates

• If you have a forecast for a growth rate, you also have a forecast for the level
• If the current level is 253, and the forecasted growth is 2.3%, the forecasted level is 259
• If a 90% forecast interval for the growth is [1%, 4%], the 90% interval for the level is [256, 263]
Estimation with Unit Roots

• If a series has a unit root, it is non-stationary, so the mean and variance are changing over time.
• Classical estimation theory does not apply
• However, least-squares estimation is still consistent
Consistent Estimation

• If the true process is

\[ y_t = y_{t-1} + e_t \]

• And you estimate an AR(1)

\[ y_t = \hat{\alpha} + \hat{\beta} y_{t-1} + \hat{e}_t \]

• Then the coefficient estimates will converge in probability to the true values (0 and 1) as T gets large
Example on simulated data

- **N=50**

|       | Coef.   | Std. Err. |    t  |    P>|t| | [95% Conf. Interval] |
|-------|---------|-----------|-------|---------|----------------------|
| y     | 0.9240092 | 0.0588153 | 15.71 | 0.000   | 0.805688 - 1.04233   |
| L1.   | 0.0492537 | 0.1419531 | 0.35  | 0.730   | -0.2363192 - 0.3348266 |
| _cons |         |           |       |         |                      |

- **N=200**

|       | Coef.   | Std. Err. |    t  |    P>|t| | [95% Conf. Interval] |
|-------|---------|-----------|-------|---------|----------------------|
| y     | 0.9737057 | 0.0213262 | 45.66 | 0.000   | 0.9316487 - 1.015763 |
| L1.   | 0.0987149 | 0.076367  | 1.29  | 0.198   | -0.0518868 - 0.2493166 |
| _cons |         |           |       |         |                      |

- **N=400**

|       | Coef.   | Std. Err. |    t  |    P>|t| | [95% Conf. Interval] |
|-------|---------|-----------|-------|---------|----------------------|
| y     | 0.9899704 | 0.0068761 | 143.97 | 0.000   | 0.9764523 - 1.003489 |
| L1.   | 0.0605234 | 0.0596962 | 1.01  | 0.311   | -0.0568368 - 0.1778837 |
| _cons |         |           |       |         |                      |
Model with drift

• If the truth is

\[ y_t = \alpha + y_{t-1} + e_t \]

• And you estimate an AR(1) with trend

\[ y_t = \hat{\alpha} + \hat{\gamma}t + \hat{\beta}y_{t-1} + \hat{e}_t \]

• Then the coefficient estimates converge in probability to the true values \((\alpha, 0, 1)\)

• It is important to include the time trend in this case.
Example with simulated data with drift

- N=50

|       | Coef.    | Std. Err. | t     | P>|t| | [ 95% Conf. Interval ] |
|-------|----------|-----------|-------|------|------------------------|
| y     | 0.230531 | 0.0159104 | 1.45  | 0.154| -0.0089728 - 0.055079  |
| t     | 0.881484 | 0.0697116 | 12.64 | 0.000| 0.7411615 - 1.021806  |
| L1. y | 0.133635 | 0.2670196 | 0.50  | 0.619| -0.4038467 - 0.6711185 |

- N=200

|       | Coef.    | Std. Err. | t     | P>|t| | [ 95% Conf. Interval ] |
|-------|----------|-----------|-------|------|------------------------|
| y     | 0.000763 | 0.0015264 | 0.50  | 0.618| -0.0022472 - 0.0037732 |
| t     | 0.942307 | 0.0187133 | 50.36 | 0.000| 0.9054024 - 0.9792129  |
| L1. y | 0.944347 | 0.2474848 | 3.82  | 0.000| 0.4562721 - 1.432422   |
Non-Standard Distribution

• A problem is that the sampling distribution of the least-squares estimates and t-ratios are not normal when there is a unit root
• Critical values quite different than conventional
• Non-Normal
• Negative bias
Testing for a Unit Root

- Null hypothesis:
  - There is a unit root
- In AR(1)
  - Coefficient on lagged variable is “1”
- In AR(k)
  - Sum of coefficients is “1”
AR(1) Model

• Estimate

\[ y_t = \hat{\alpha} + \hat{\beta}y_{t-1} + \hat{e}_t \]

• Or equivalently

\[ \Delta y_t = \hat{\alpha} + \hat{\rho}y_{t-1} + \hat{e}_t \]

\[ \hat{\rho} = \hat{\beta} - 1 \]

• Test for \( \beta=1 \) same as test for \( \rho=0 \).

• Test statistic is t-ratio on lagged \( y \)
AR(k+1) model

• Estimate

\[ \Delta y_t = \hat{\alpha} + \hat{\rho} y_{t-1} + \hat{\beta}_1 \Delta y_{t-1} + \cdots + \hat{\beta}_k \Delta y_{t-k} + \hat{e}_t \]

• Test for \( \rho = 0 \)

• Called ADF test
  
  – Augmented Dickey-Fuller
  
  – (Test without extra lags is called Dickey-Fuller, test with extra lags called Augmented Dickey-Fuller)
Theory of Unit Root Testing

• Wayne Fuller (Iowa State)
  – David Dickey (NCSU)
  – Developed DF and ADF test

• Peter Phillips (Yale)
  – Extended the distribution theory
STATA ADF test

• `dfuller t3, lags(12)`

• This implements a ADF test with 12 lags of differenced data

• Equivalent to an AR(13)

• Alternatively

• `reg d.t3 L.t3 L(1/12).d.t3`
Example: 3-month T-bill
Example: 3-month T-bill

. dfuller t3, lags(12)

Augmented Dickey-Fuller test for unit root Number of obs = 961

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-1.964</td>
<td>-3.430</td>
<td>-2.860</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.3027

- The p-value is not significant
- Equivalently, the statistic of -2 is not smaller than the 10% critical value
- Do not reject a unit root for 3-month T-Bill
• James MacKinnon
• Queen’s University
• Computed p-value function
Alternatively

\[ \text{reg d.t3 L.t3 L(1/12).d.t3} \]

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 961</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>33.09023</td>
<td>13</td>
<td>2.5454023</td>
<td>F( 13, 947) = 23.96</td>
</tr>
<tr>
<td>Residual</td>
<td>100.623536</td>
<td>947</td>
<td>.106255054</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Total</td>
<td>133.713766</td>
<td>960</td>
<td>.139285173</td>
<td>R-squared = 0.2475</td>
</tr>
</tbody>
</table>

| D.t3month | Coef.   | Std. Err. | t    | P>|t|  | [95% Conf. Interval] |
|-----------|---------|-----------|------|------|---------------------|
| t3month   |         |           |      |      |                     |
| L1        | -.0067016 | .0034127 | -1.96 | 0.050 | -0.0133988 | -4.30e-06 |
| LD        | .4232105 | .0321545 | 13.16 | 0.000 | .3601083 | .4863128 |
| L2D       | -.198499 | .0348316 | -5.70 | 0.000 | -.266855 | -.1301429 |
| L3D       | .0720256 | .0353008 | 2.04  | 0.042 | .0027487 | .1413025 |
| L4D       | -.0816056 | .0351049 | -2.32 | 0.020 | -.1504979 | -.0127133 |
| L5D       | .1607123 | .0351436 | 4.57  | 0.000 | .0917444 | .2296807 |
| L6D       | -.256786 | .0355267 | -7.23 | 0.000 | -.3265061 | -.1870658 |
| L7D       | .0015034 | .0354747 | 0.04  | 0.966 | -.0681147 | .0711216 |
| L8D       | .0703407 | .0351216 | 2.00  | 0.045 | .0014154 | .1392659 |
| L9D       | .141971 | .0350786 | 4.05  | 0.000 | .0731301 | .2108118 |
| L10D      | -.0841005 | .0353162 | -2.38 | 0.017 | -.1534076 | -.0147934 |
| L11D      | .1028565 | .0347711 | 2.96  | 0.003 | .0346192 | .1710938 |
| L12D      | -.1291987 | .0322077 | -4.01 | 0.000 | -.1924055 | -.065992 |
| _cons     | .0242965 | .0163062 | 1.49  | 0.137 | -.0077039 | .056297 |

- The t for L1.t3 is -2
- Ignore reported p-value, compare with table
Interest Rate Spread
ADF test for Spread

. dfuller spread, lags(12)

Augmented Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-4.644</td>
<td>-3.430</td>
<td>-2.860</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.0001

• The test of -4.6 is smaller than the critical value
• The p-value of .0001 is much smaller than 0.05
• We reject the hypothesis of a unit root
• We find evidence that the spread is stationary
Testing for a unit Root with Trend

- If the series has a trend

\[ \Delta y_t = \hat{\alpha} + \hat{\rho} y_{t-1} + \hat{\gamma} t + \hat{\beta}_1 \Delta y_{t-1} + \cdots + \hat{\beta}_k \Delta y_{t-k} + \hat{\epsilon}_t \]

- Again test for \( \rho = 0 \).

- `dfuller y, trend lags(2)`
Example: Log(RGDP)

- ADF with 2 lags

```
. dfuller ln_rgd, trend lags(2)
```

Augmented Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-1.844</td>
<td>-3.989</td>
<td>-3.429</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.6830

- The p-value is not significant.
- We do not reject the hypothesis of a unit root.
- Consistent with forecasting growth rates, not levels.
Unit Root Tests in Practice

• Examine your data.
  – Is it trended?
  – Does it appear stationary?
• If it may be non-stationary, apply ADF test
  – Include time trend if trended
• If test rejects hypothesis of a unit root
  – The evidence is that the series is stationary
• If the test fails to reject
  – The evidence is not conclusive
  – Many users then treat the series as if it has a unit root
    • Difference the data, forecast changes or growth rates
Spurious Regression

• One problem caused by unit roots is that it can induce \textit{spurious correlation} among time series
  – Clive Granger and Paul Newbold (1974)
    • Observed the phenomenon
    • Paul Newbold a UW PhD (1970)
  – Peter Phillips (1987)
    • Invented the theory
Spurious Regression

• Suppose you have two independent time-series \( y_t \) and \( x_t \)
• Suppose you regress \( y_t \) on \( x_t \)
• Since they are independent, you should expect a zero coefficient on \( x_t \) and an insignificant t-statistic, right?
Example

Two independent Random Walks
Regression of $y$ on $x$

- $X$ has an estimated coefficient of 0.6
- A $t$-statistic of 18! Highly significant!
- But $x$ and $y$ are independent!

```
. reg y x

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>21379.9809</td>
<td>1</td>
<td>21379.9809</td>
</tr>
<tr>
<td>Residual</td>
<td>31322.7492</td>
<td>498</td>
<td>62.8970868</td>
</tr>
<tr>
<td>Total</td>
<td>52702.7302</td>
<td>499</td>
<td>105.616694</td>
</tr>
</tbody>
</table>

| y        | Coef.   | Std. Err. | t     | P>|t|  | [95% Conf. Interval] |
|----------|---------|-----------|-------|------|----------------------|
| x        | 0.6096435 | 0.0330664 | 18.44 | 0.000 | 0.5446765 - 0.6746104 |
| _cons    | -1.062265 | 0.7635661 | -1.39 | 0.165 | -2.562473 - 0.437943  |
```

Number of obs = 500
F( 1, 498) = 339.92
Prob > F = 0.0000
R-squared = 0.4057
Adj R-squared = 0.4045
Root MSE = 7.9308

Total

Root MSE = 7.9308
Spurious Regression

• This is not an accident
• It happens whenever you regress a random walk on another.
• Traditional implication:
  – Don’t regress levels on levels
  – First difference your data
• Even better
  – Make sure your dynamic specification is correct
  – Include lags of your dependent variable
Dynamic Regression

• Regress y on lagged y, plus x

. reg y L.y x

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 499</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>52032.1</td>
<td>2</td>
<td>26016.0</td>
<td>F( 2, 496) = 26485.68</td>
</tr>
<tr>
<td>Residual</td>
<td>487.2</td>
<td>496</td>
<td>.982268</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Total</td>
<td>52519</td>
<td>498</td>
<td>105.460</td>
<td>R-squared = 0.9907</td>
</tr>
<tr>
<td></td>
<td>3238</td>
<td></td>
<td>105.460</td>
<td>Adj R-squared = 0.9907</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Root MSE = 0.99109</td>
</tr>
</tbody>
</table>

|          | Coef.  | Std. Err. | t      | P>|t|   | [95% Conf. Interval] |
|----------|--------|-----------|--------|-------|----------------------|
| L1.y     | 0.9917 | 0.00558   | 177.18 | 0.000 | 0.9807974 1.002794  |
| x        | 0.0059 | 0.00537   | 1.11   | 0.267 | -.0045827 0.0165038 |
| _cons    | -.0458 | 0.0960    | -0.48  | 0.634 | -.2345104 0.1428875 |

• Now x has insignificant t-statistic, and much smaller coefficient estimate

• Coefficient estimate on lagged y is close to 1.
Message

- If your data might have a unit root
  - Try an ADF test
  - Consider forecasting differences or growth rates
  - Always include lagged dependent variable when series is highly correlated
Examples of Spurious Regression

• Prepared by Jesus Gonzalo
  – Universidad Carlos III de Madrid
• All examples use annual data
• Estimates, t-statistics, and R squared reported.
1. Egyptian infant mortality rate ($Y$) on Income of USA farmers ($X_1$) and Honduran money supply ($X_2$)

- 1971-1990

$$Y = 180 - 0.30X_1 - 0.04X_2$$

$$R^2 = 0.92$$

$$t_1 = 2.3$$

$$t_2 = 4.3$$
2. USA Export Index (Y) on Australian males’ life expectancy (X)

- 1960-1990

\[ Y = -2943 + 46X \]

\[ t = 18 \]

\[ R^2 = 0.92 \]
3. USA Defense Expenditure (Y) on Population of South Africa (X)

• 1960-1990

\[ Y = -367 + 0.018X \]

\[ t = 17 \]

\[ R^2 = 0.94 \]
4. USA Crime Rate \((Y)\) on Life Expectancy in South Africa \((X)\)

- 1971-1990

\[
Y = -24569 + 629X
\]

\[
t = 9
\]

\[
R^2 = 0.811
\]
5. Population of South Africa (Y) on R&D in USA (X)

• 1971-1990

\[ Y = 21699 + 112X \]

\[ t = 26 \]
\[ R^2 = 0.97 \]
Message

• If your data are trended
  – Do not trust simple levels regression
  – Standard errors much too small, t-ratios misleading large
  – R squared misleadingly large
  – Consider forecasting differences or growth rates
  – Include lagged dependent variable in regression