Unit Roots

- An autoregressive process

$$a(L)y_t = e_t$$

has a unit root if

$$a(1) = 0$$

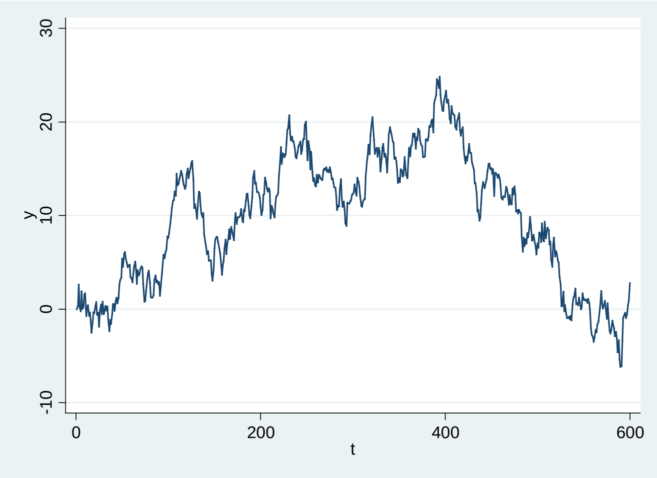
- The simplest case is the AR(1) model

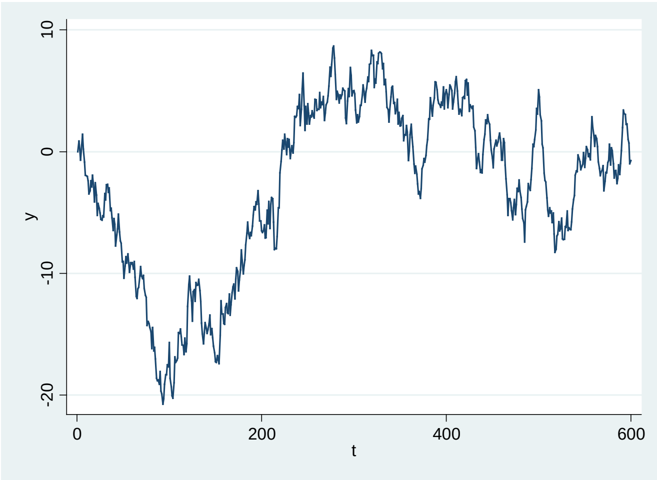
$$(1 - L)y_t = e_t$$

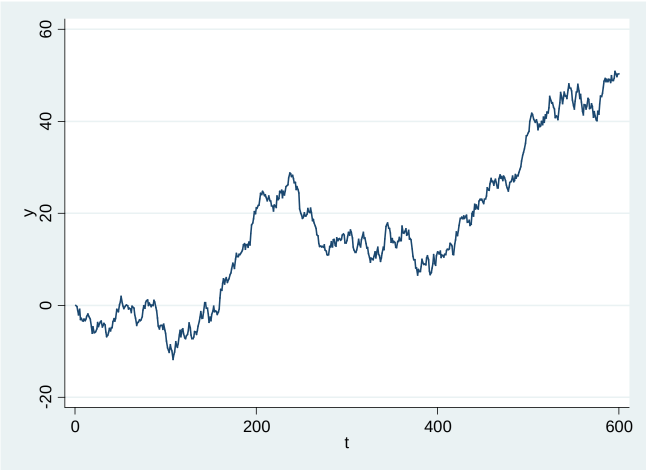
or

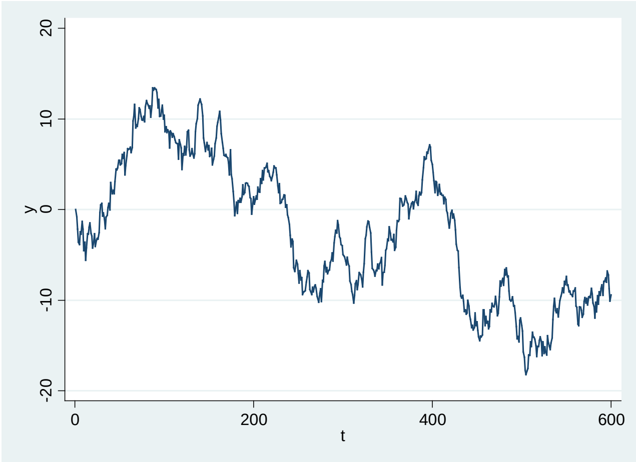
$$y_t = y_{t-1} + e_t$$

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

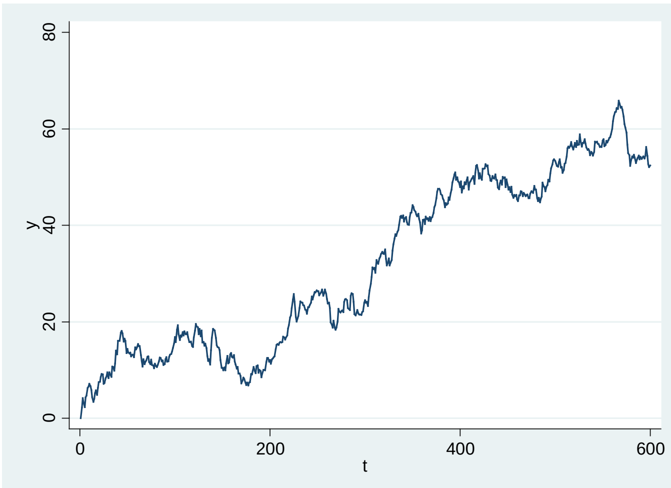
$$y_t = T_t + C_t$$

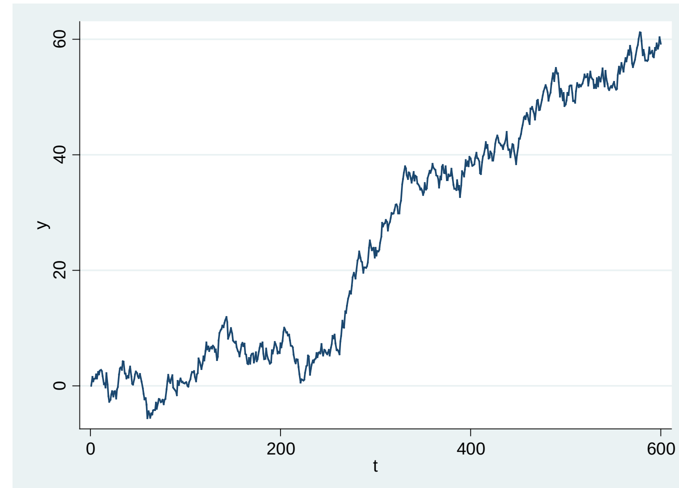
$$T_t = \alpha t$$

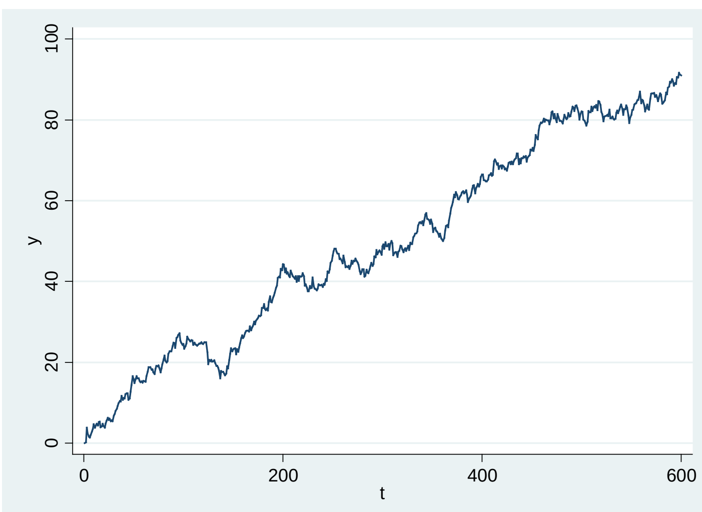
$$C_t = C_{t-1} + e_t$$

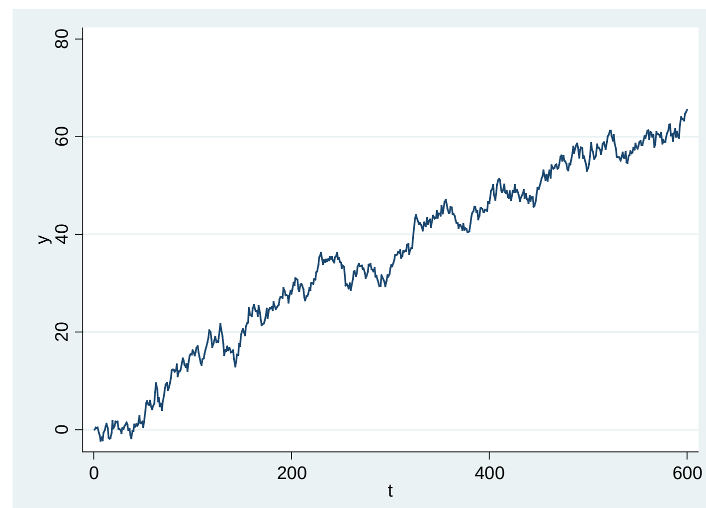
Examples

$$y_t = 0.1 + y_{t-1} + e_t$$
$$e_t \sim N(0,1)$$









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

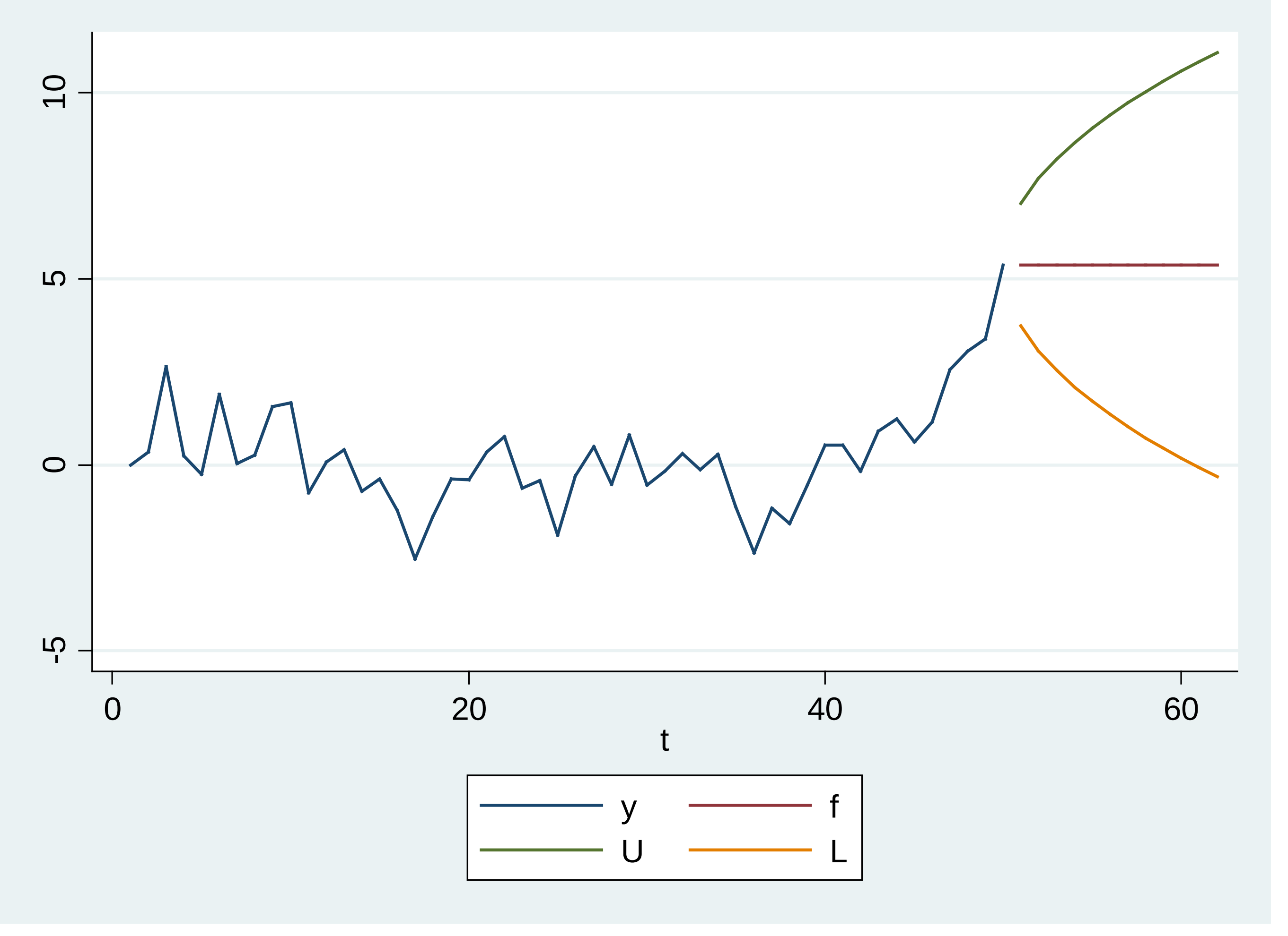
Forecast intervals

- 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

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
y						
L1.	.9240092	.0588153	15.71	0.000	.805688	1.04233
_cons	.0492537	.1419531	0.35	0.730	-.2363192	.3348266

- N=200

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
y						
L1.	.9737057	.0213262	45.66	0.000	.9316487	1.015763
_cons	.0987149	.076367	1.29	0.198	-.0518868	.2493166

- N=400

y						
L1.	.9899704	.0068761	143.97	0.000	.9764523	1.003489
_cons	.0605234	.0596962	1.01	0.311	-.0568368	.1778837

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 (α,0,1)
- It is important to include the time trend in this case.

Example with simulated data with drift

- N=50

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
t	.0230531	.0159104	1.45	0.154	-.0089728	.055079
y						
L1.	.8814838	.0697116	12.64	0.000	.7411615	1.021806
_cons	.1336359	.2670196	0.50	0.619	-.4038467	.6711185

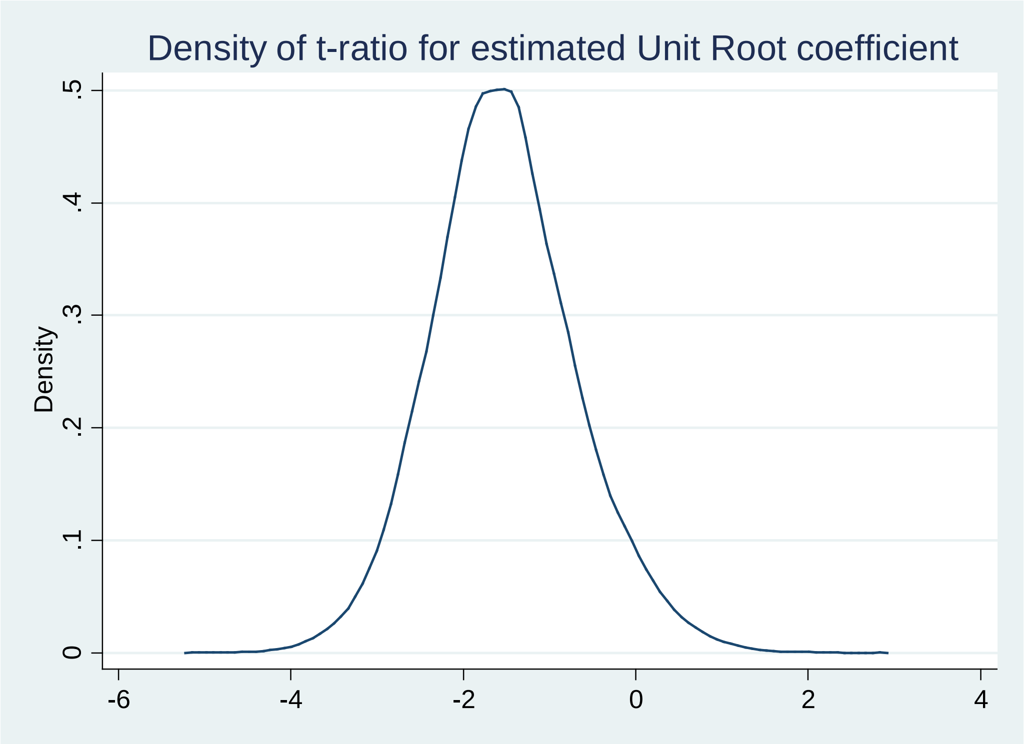
- N=200

y	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
t	.000763	.0015264	0.50	0.618	-.0022472	.0037732
y						
L1.	.9423076	.0187133	50.36	0.000	.9054024	.9792129
_cons	.944347	.2474848	3.82	0.000	.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

Density of t-ratio



- 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

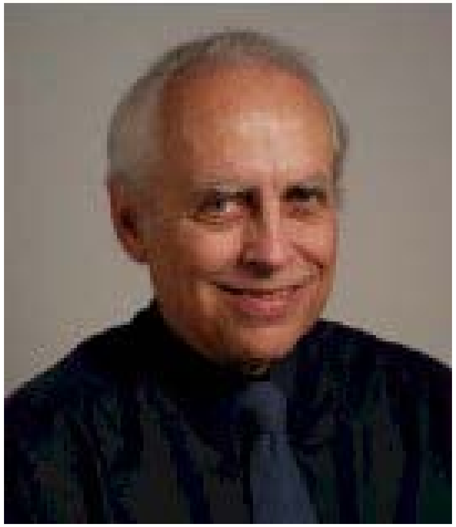
- Estimate

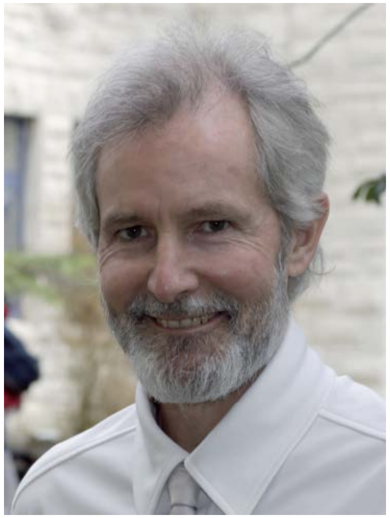
$$\Delta y_t = \hat{\alpha} + \hat{\rho} y_{t-1} + \hat{\beta}_1 \Delta y_{t-1} + \cdots + \hat{\beta}_k \Delta y_{t-k} + \hat{e}_t$$

- Test for ρ=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

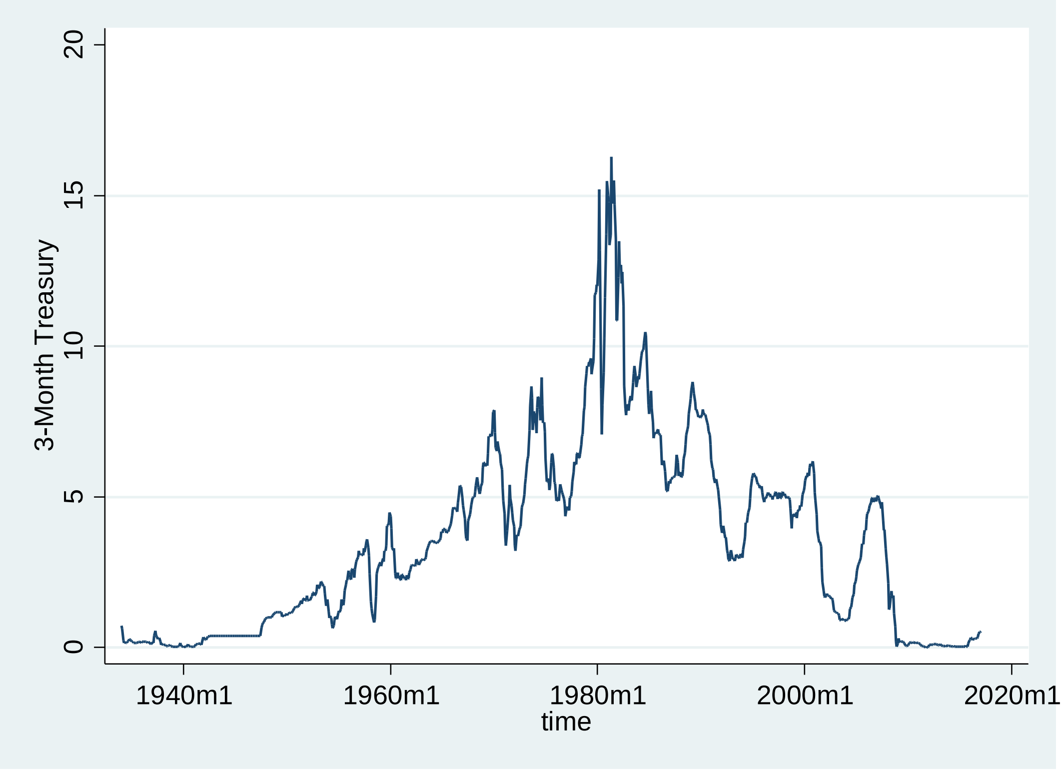




STATA ADF test

- **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   =       985

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -2.002            -3.430            -2.860            -2.570
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.2858
```

- 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

```
. reg d.t3 L.t3 L(1/12).d.t3
```

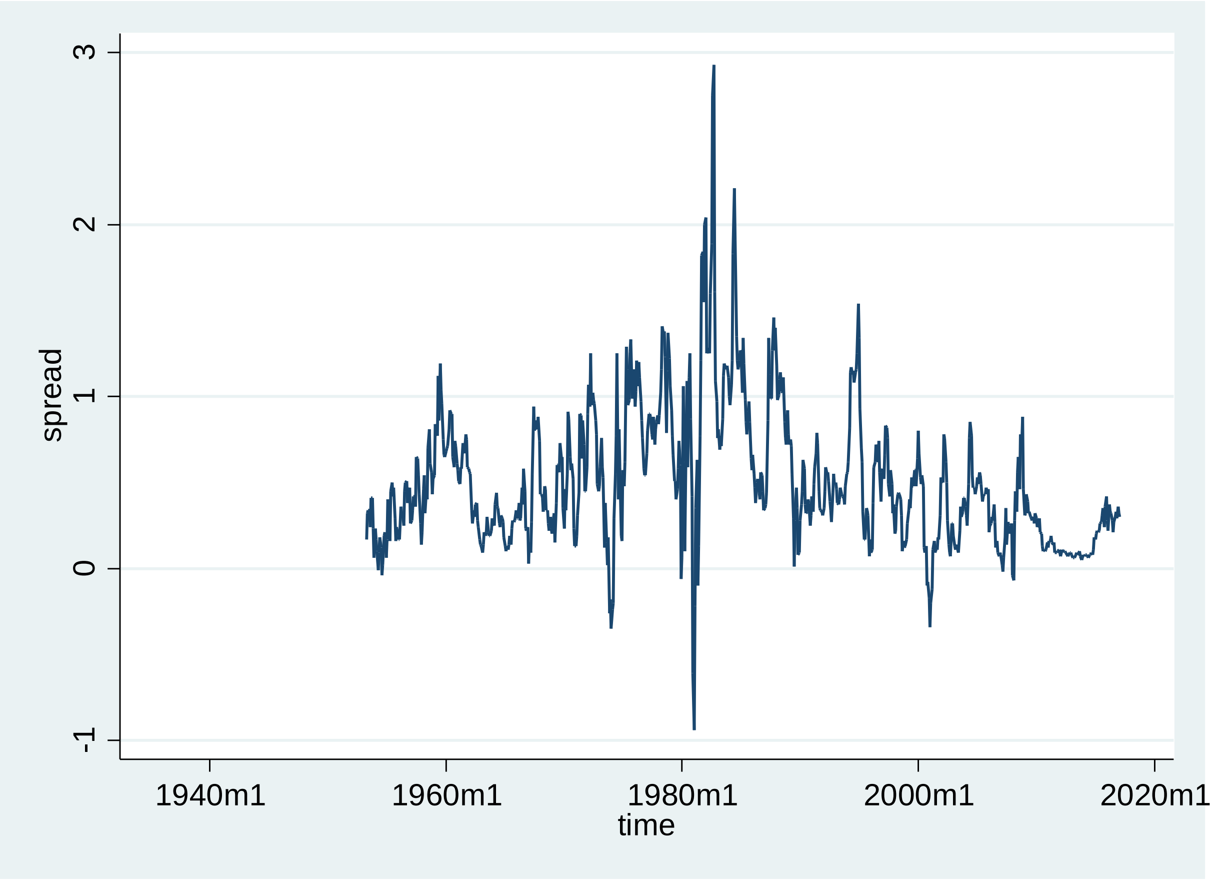
Source	SS	df	MS
Model	33.1025667	13	2.54635128
Residual	100.667329	971	.103673872
Total	133.769896	984	.135945016

Number of obs	=	985
F(13, 971)	=	24.56
Prob > F	=	0.0000
R-squared	=	0.2475
Adj R-squared	=	0.2374
Root MSE	=	.32198

D.t3month	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
t3month						
L1.	-.0066426	.0033185	-2.00	0.046	-.0131547	-.0001304
LD.	.4231827	.0317526	13.33	0.000	.3608711	.4854942
L2D.	-.1986427	.0343953	-5.78	0.000	-.2661404	-.131145
L3D.	.0719919	.0348601	2.07	0.039	.0035821	.1404017
L4D.	-.0817067	.0346674	-2.36	0.019	-.1497384	-.0136751
L5D.	.1606678	.0347055	4.63	0.000	.0925614	.2287743
L6D.	-.2567361	.0350839	-7.32	0.000	-.3255851	-.1878871
L7D.	.0014404	.035034	0.04	0.967	-.0673107	.0701915
L8D.	.0701928	.034685	2.02	0.043	.0021267	.138259
L9D.	.1419992	.0346426	4.10	0.000	.0740163	.2099821
L10D.	-.0842016	.0348785	-2.41	0.016	-.1526476	-.0157556
L11D.	.1029241	.0343411	3.00	0.003	.0355328	.1703153
L12D.	-.1291541	.0318061	-4.06	0.000	-.1915706	-.0667376
_cons	.0239211	.015662	1.53	0.127	-.0068141	.0546563

- 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         Number of obs   =       754

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -4.711            -3.430            -2.860            -2.570
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0001
```

- The test of -4.7 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

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

- Again test for ρ=0.
- **dfuller y, trend lags(2)**

Example: Log(RGDP)

- ADF with 2 lags

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

Augmented Dickey-Fuller test for unit root         Number of obs   =       277

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -1.625            -3.989            -3.429            -3.130
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.7824
```

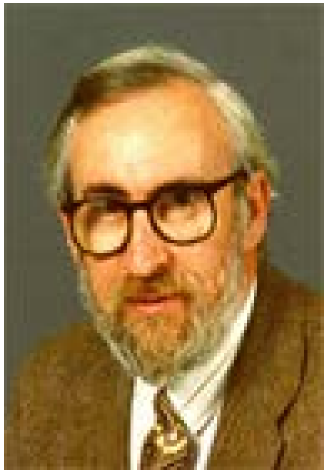
- The p-value is not significant.
- We do not reject the hypothesis of a unit root
- Consistent with forecasting growth rates, not levels.

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 *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

```
. reg y x

      Source |       SS       df       MS              Number of obs =     500
-------------+------------------------------           F(  1,   498) =  339.92
       Model |  21379.9809     1  21379.9809           Prob > F      =  0.0000
    Residual |  31322.7492   498  62.8970868           R-squared     =  0.4057
-------------+------------------------------           Adj R-squared =  0.4045
       Total |  52702.7302   499  105.616694           Root MSE      =  7.9308

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           x |   .6096435   .0330664    18.44   0.000     .5446765    .6746104
       _cons |  -1.062265   .7635661    -1.39   0.165    -2.562473     .437943
------------------------------------------------------------------------------
```

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

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

      Source |       SS       df       MS              Number of obs =     499
-------------+------------------------------           F(  2,   496) =26485.68
       Model |  52032.1184     2  26016.0592           Prob > F      =  0.0000
    Residual |  487.205408   496  .982268967           R-squared     =  0.9907
-------------+------------------------------           Adj R-squared =  0.9907
       Total |  52519.3238   498   105.46049           Root MSE      =  .99109

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           y |
         L1. |   .9917958   .0055978   177.18   0.000     .9807974    1.002794
             |
           x |   .0059606   .0053662     1.11   0.267    -.0045827    .0165038
       _cons |  -.0458114   .0960418    -0.48   0.634    -.2345104    .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

$$Y = 21699 + 112X$$

$$t = 26$$
$$R^2 = 0.97$$

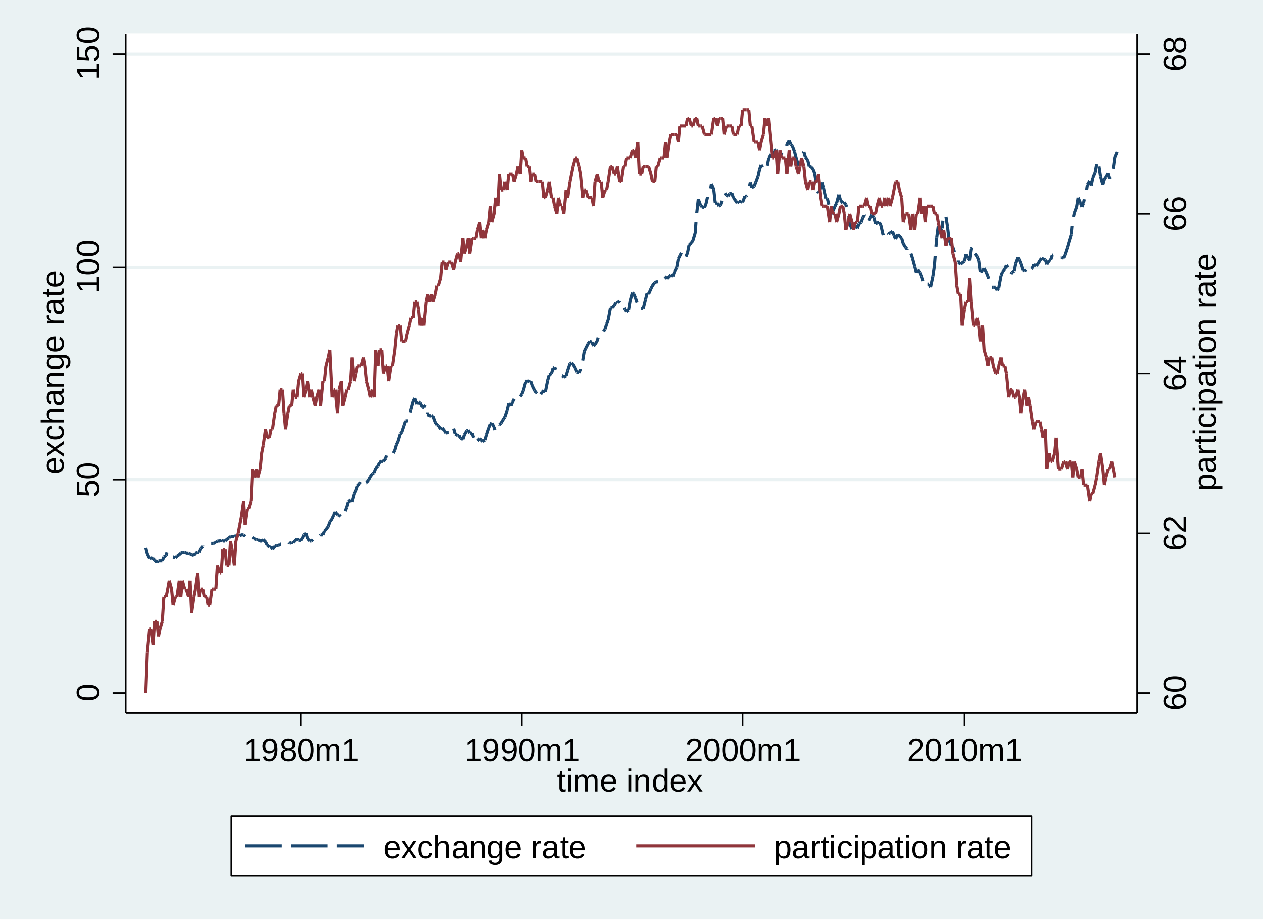
My Example:
Trump-Era Exchange Rate Theory

- Outsourcing causes displaced workers to be discouraged, and the production shift alters the exchange rate
- Monthly data: 1977-2016

$$Exchange\ Rate = 10 Labor\ Force\ Participation$$

$$t = 18$$
$$R^2 = 0.35$$

Exchange and Labor Force Participation Rates



Dynamic Regression

```
. reg exchange L(1/4).exchange labor, r

Linear regression                               Number of obs     =        523
                                                F(5, 517)         >   99999.00
                                                Prob > F          =     0.0000
                                                R-squared         =     0.9990
                                                Root MSE          =     1.0101
```

exchange	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
exchange						
L1.	1.433293	.0505001	28.38	0.000	1.334082	1.532504
L2.	-.5687469	.0824459	-6.90	0.000	-.7307171	-.4067768
L3.	.1829668	.0936117	1.95	0.051	-.0009393	.3668729
L4.	-.0490465	.0503467	-0.97	0.330	-.1479558	.0498629
labor	.0126531	.0316666	0.40	0.690	-.049558	.0748641
_cons	-.5773737	1.94393	-0.30	0.767	-4.396347	3.2416

- Labor insignificant in dynamic regression

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

Assignments

- Read Chapter 10, *The Signal and the Noise*
  - Reading Reflection
  - Thursday (4/13)