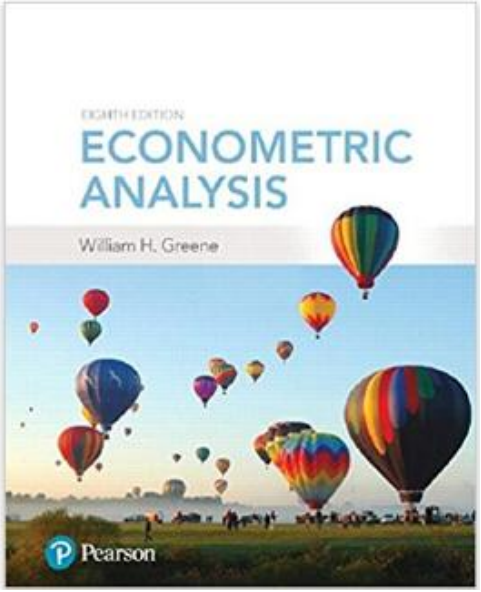
Econometrics I

Professor William Greene
Stern School of Business
Department of Economics



Econometrics I

Part 22 – Time Series

Modeling an Economic Time Series

- Observed y_0, y_1, …, y_t,…
- What is the "sample?" Realization of the entire sequence?
- Random sampling? Not really possible. We are using a different type of statistics.
- The "observation window"

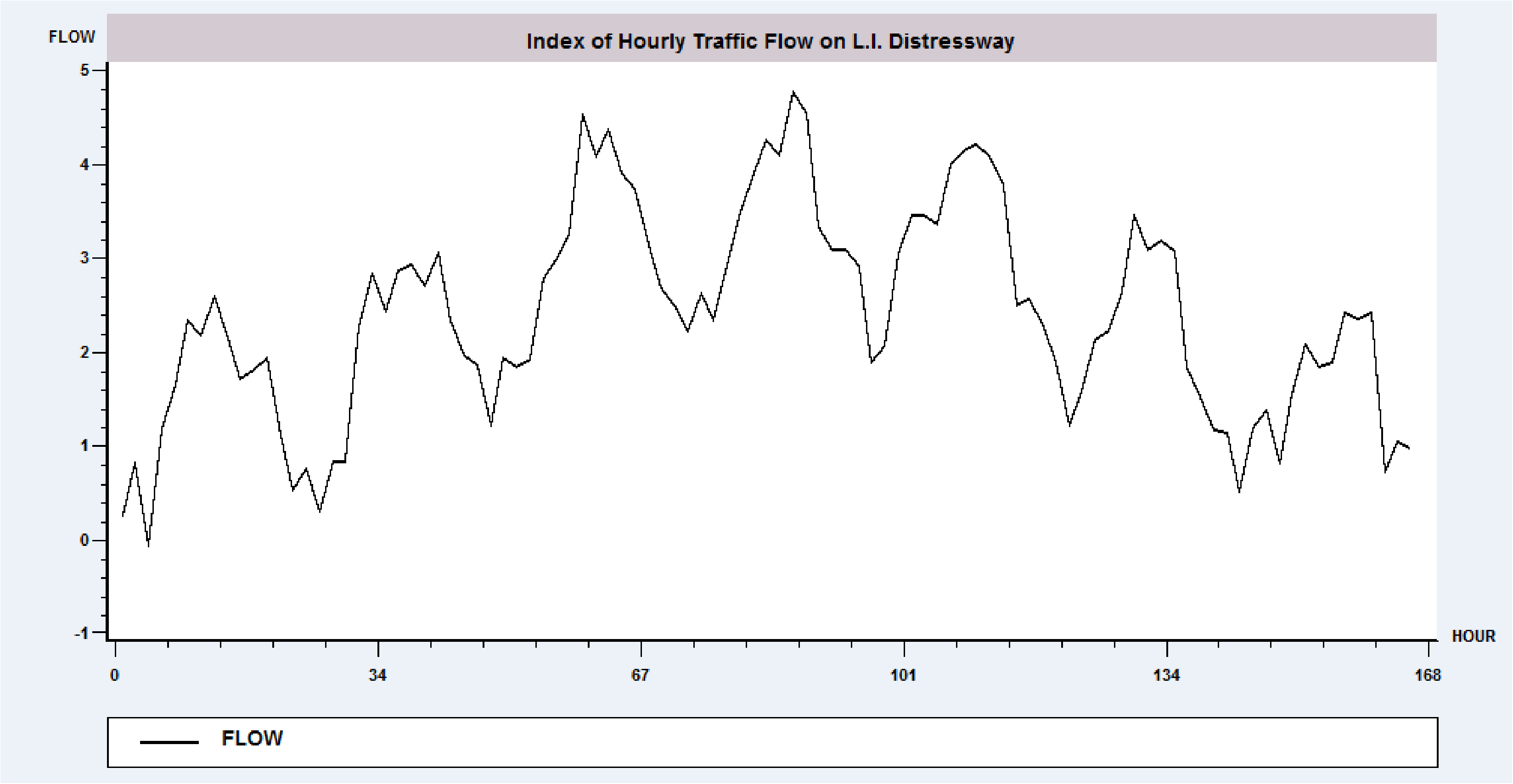
Estimators

- Functions of sums of correlated observations
- Law of large numbers?
  - Non-independent observations
  - What does “increasing sample size” mean?
- Asymptotic properties? (There are no finite sample properties.)

Interpreting a Time Series

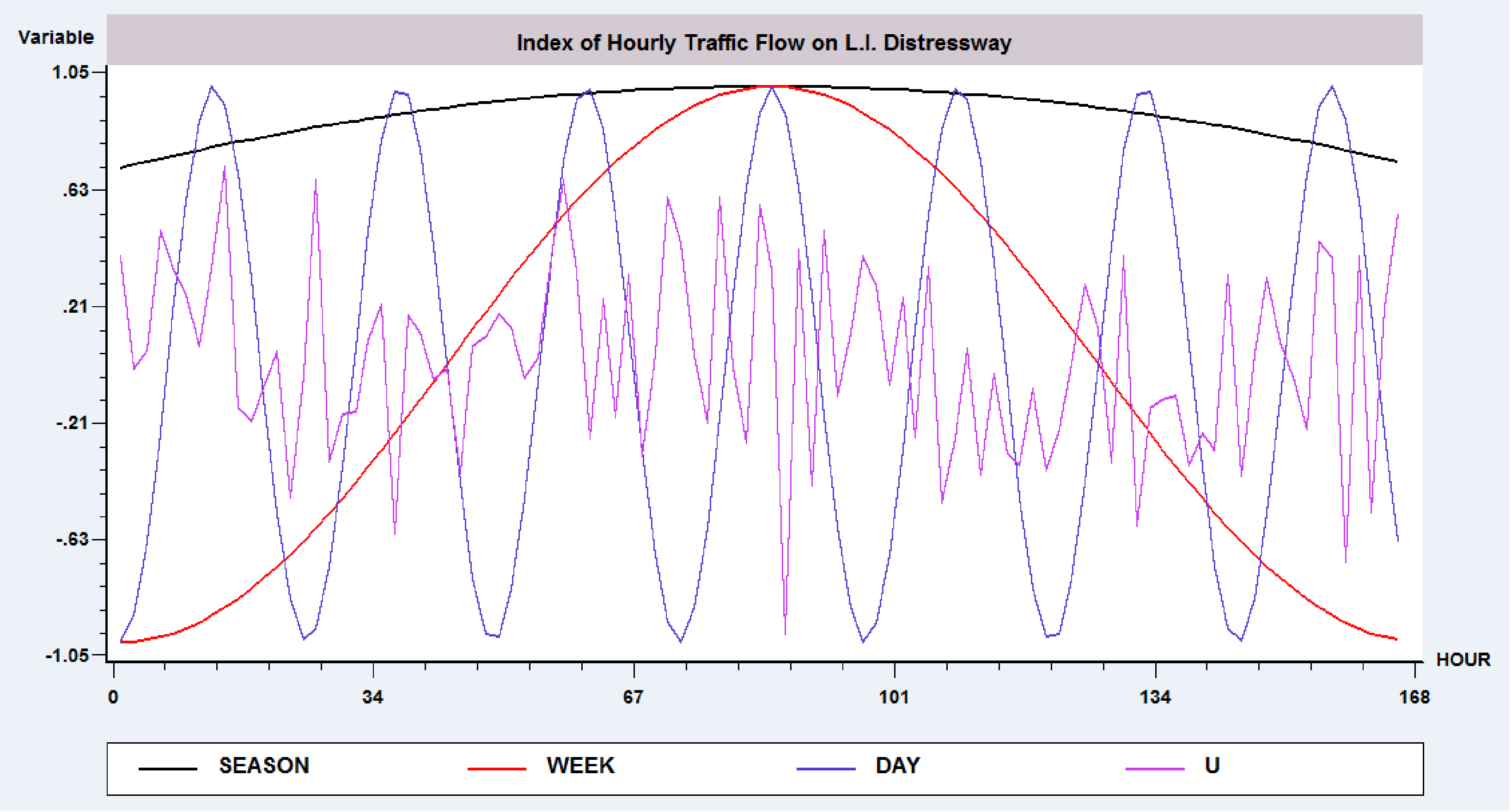
- Time domain: A "process"
  - y(t) = ax(t) + by(t-1) + …
  - Regression like approach/interpretation
- Frequency domain: A sum of terms
  - y(t) = $\sum_j \beta_j Cos(\alpha_j t) + \varepsilon(t)$
  - Contribution of different frequencies to the observed series.
- ("High frequency data and financial econometrics – "frequency" is used slightly differently here.)

For example,…



Index of Hourly Traffic Flow on L.I. Distressway
FLOW
5
4
3
2
1
0
-1
HOUR
0
34
67
101
134
168
FLOW

In parts…



Index of Hourly Traffic Flow on L.I. Distressway
Variable
1.05
.63
.21
-.21
-.63
-1.05
0
34
67
101
134
168
HOUR
SEASON
WEEK
DAY
U

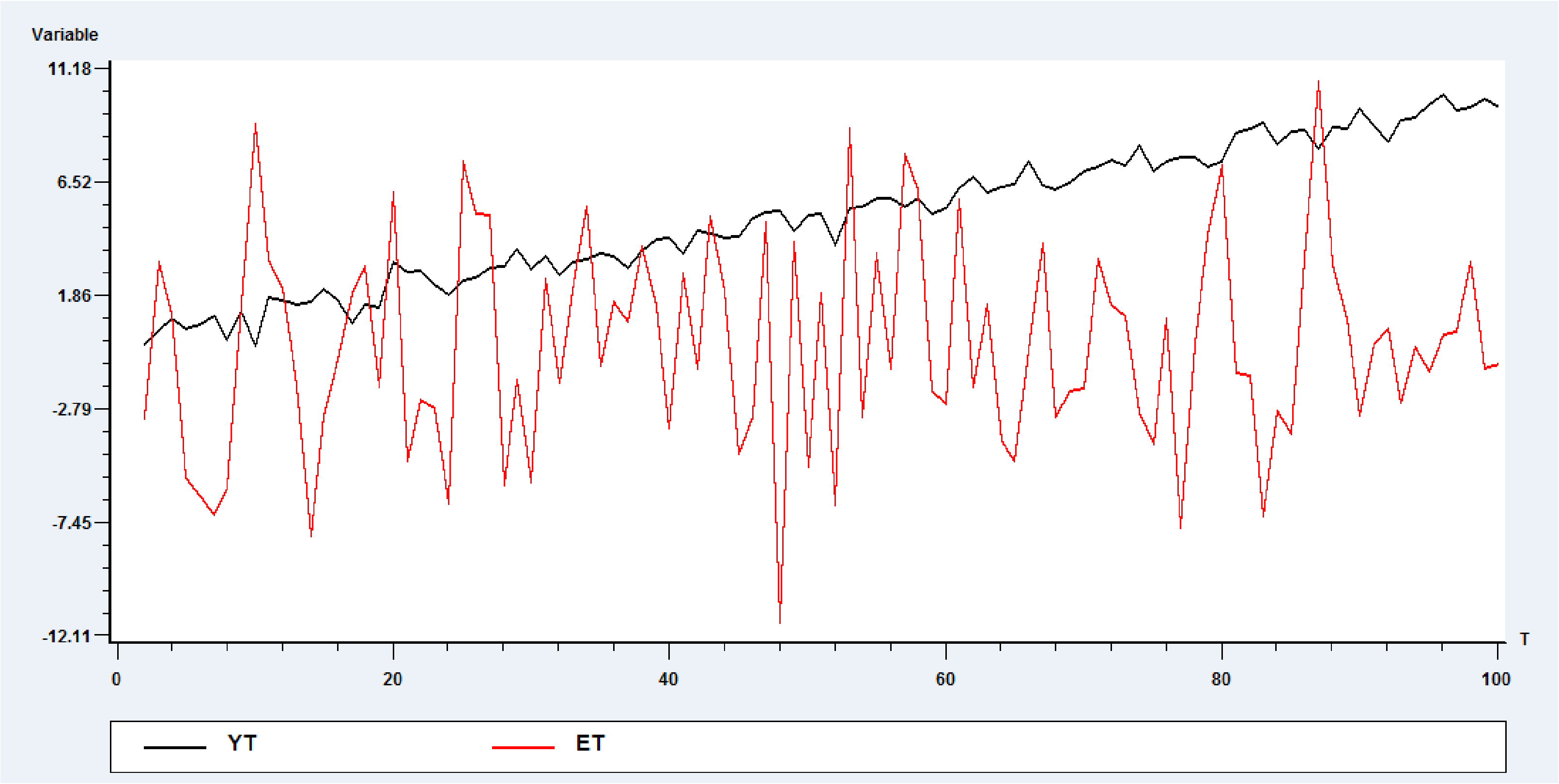
Studying the Frequency Domain

- Cannot identify the number of terms
- Cannot identify frequencies from the time series
- Deconstructing the variance, autocovariances and autocorrelations
  - Contributions at different frequencies
  - Apparent large weights at different frequencies
  - Using Fourier transforms of the data
  - Does this provide “new” information about the series?

Stationary Time Series

- $y_t = b_1y_{t-1} + b_2y_{t-2} + \ldots + b_Py_{t-P} + e_t$
- Autocovariance: $\gamma_k = \text{Cov}[y_t,y_{t-k}]$
- Autocorrelation: $\rho_k = \gamma_k / \gamma_0$
- Stationary series: γ_k depends only on k, not on t
  - Weak stationarity: $E[y_t]$ is not a function of t, $E[y_t * y_{t-s}]$ is not a function of t or s, only of |t-s|
  - Strong stationarity: The joint distribution of $[y_t,y_{t-1},\ldots,y_{t-s}]$ for any window of length s periods, is not a function of t or s.
- A condition for weak stationarity: The smallest root of the characteristic polynomial: $1 - b_1z^1 - b_2z^2 - \ldots - b_Pz^P = 0$, is greater than one.
  - The unit circle
  - Complex roots
  - Example: $y_t = \rho y_{t-1} + e_e$, $1 - \rho z = 0$ has root $z = 1/\rho$, $| z | > 1 \Rightarrow | \rho | < 1$.

Stationary vs. Nonstationary Series



The Lag Operator

- $Lx_t = x_{t-1}$
- $L^2 x_t = x_{t-2}$
- $L^P x_t + L^Q x_t = x_{t-P} + x_{t-Q}$
- Polynomials in L: $y_t = B(L)y_t + e_t$
- $A(L) y_t = e_t$
- Invertibility: $y_t = [A(L)]^{-1} e_t$

Inverting a Stationary Series

- $y_t = \rho y_{t-1} + e_t \Rightarrow (1- \rho L)y_t = e_t$
- $y_t = [1- \rho L]^{-1} e_t = e_t + \rho e_{t-1} + \rho^2 e_{t-2} + \ldots$

$$\frac{1}{1-\rho L} = 1 + (\rho L) + (\rho L)^2 + (\rho L)^3 + \ldots$$

  - Stationary series can be inverted
  - Autoregressive vs. moving average form of series

Autocorrelation

- How does it arise?
- What does it mean?
- Modeling approaches
  - Classical – direct: **corrective**
    - Estimation that accounts for autocorrelation
    - Inference in the presence of autocorrelation
  - Contemporary – **structural**
    - Model the source
    - Incorporate the time series aspect in the model

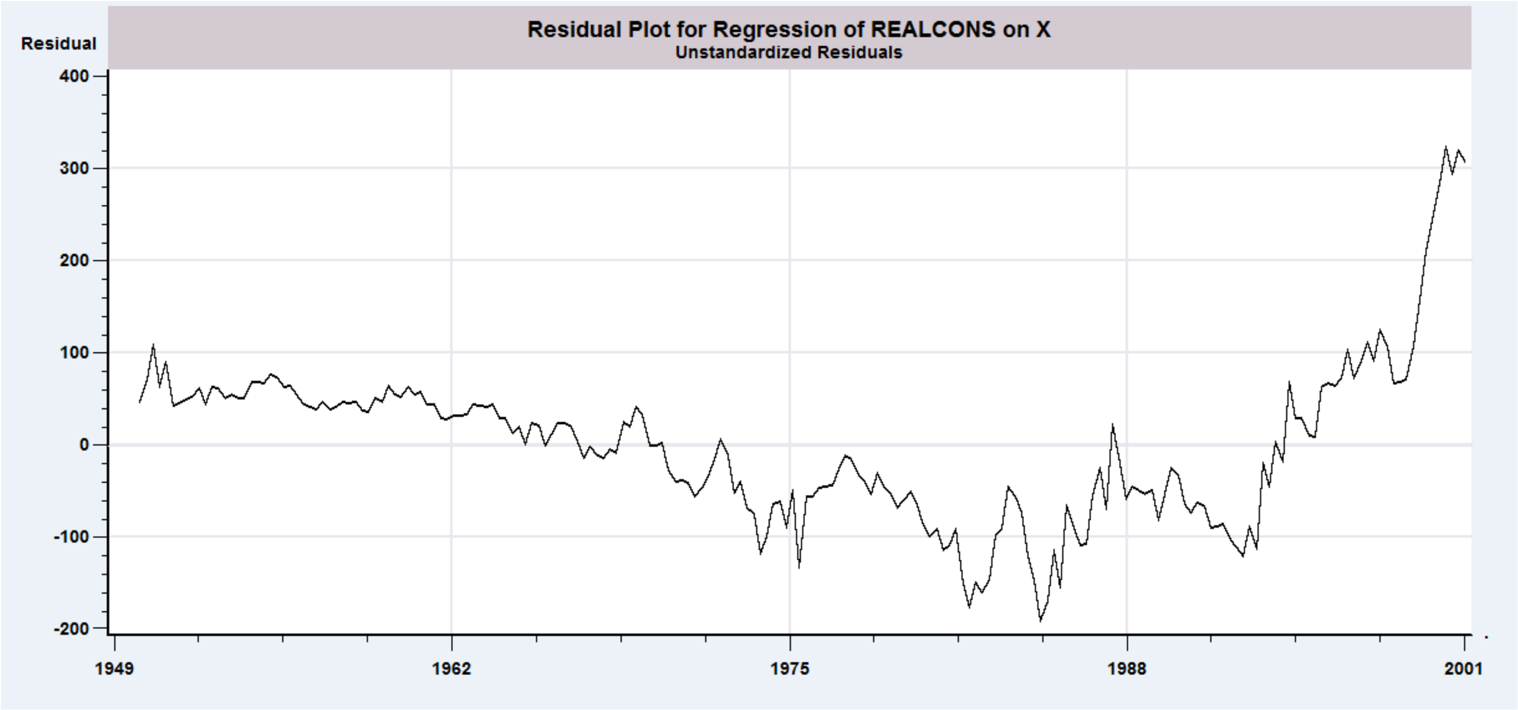
Regression with Autocorrelation

- $y_t = x_t{'}b + e_t$, $e_t = \rho e_{t-1} + u_t$
- $(1-\rho L)e_t = u_t \Rightarrow e_t = (1-\rho L)^{-1}u_t$
  - $E[e_t] = E[(1-\rho L)^{-1}u_t] = (1-\rho L)^{-1}E[u_t] = 0$
  - $Var[e_t] = (1-\rho L)^{-2}Var[u_t] = 1+\rho^2\sigma_u^2 + \ldots = \sigma_u^2/(1-\rho^2)$
  - $Cov[e_t,e_{t-1}] = Cov[\rho e_{t-1} + u_t, e_{t-1}] =$

    $=\rho Cov[e_{t-1},e_{t-1}]+Cov[u_t,e_{t-1}] = \rho\,\sigma_u^2/(1-\rho^2)$

Autocorrelation in Regression

- $Y_t = b'x_t + \varepsilon_t$
- $Cov(\varepsilon_t, \varepsilon_{t-1}) \neq 0$
- Ex. $RealCons_t = a + bRealIncome + \varepsilon_t$ U.S. Data, quarterly, 1950-2000



Generalized Least Squares

Efficient estimation of β and, by implication, the inefficiency of least squares **b**.

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{*\prime}\mathbf{X}^{*})^{-1}\mathbf{X}^{*\prime}\mathbf{y}^{*}$$
$$= (\mathbf{X}'\mathbf{P}'\mathbf{P}\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{P}'\mathbf{P}\mathbf{y}$$
$$= (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}\,\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y}$$

$\hat{\boldsymbol{\beta}} \neq \mathbf{b}$. $\hat{\boldsymbol{\beta}}$ is efficient, so by construction, **b** is not.

Autocorrelation

$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$

('First order autocorrelation.' How does this come about?)

Assume $-1 < \rho < 1$. Why?

u_t = 'nonautocorrelated white noise'

$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ (the autoregressive form)

$= \rho(\rho\varepsilon_{t-2} + u_{t-1}) + u_t$

= ... (continue to substitute)

$= u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \rho^3 u_{t-3} + \ldots$

= (the moving average form)

Autocorrelation

$$\text{Var}[\varepsilon_t] = \text{Var}[u_t + \rho u_{t-1} + \rho^2 u_{t-1} + \ldots]$$

$$= \text{Var}\left[\sum_{i=0}^{\infty} \rho^i u_{t-i}\right]$$

$$= \sum_{i=0}^{\infty} \rho^{2i} \sigma_u^2 = \frac{\sigma_u^2}{1-\rho^2}$$

An easier way: Since $\text{Var}[\varepsilon_t] = \text{Var}[\varepsilon_{t-1}]$ and $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$

$$\text{Var}[\varepsilon_t] = \rho^2 \text{Var}[\varepsilon_{t-1}] + \text{Var}[u_t] + 2\rho \text{Cov}[\varepsilon_{t-1}, u_t]$$

$$= \rho^2 \text{Var}[\varepsilon_t] + \sigma_u^2$$

$$= \frac{\sigma_u^2}{1-\rho^2}$$

Autocovariances

Continuing...

$$\begin{aligned}
\text{Cov}[\varepsilon_t, \varepsilon_{t-1}] &= \text{Cov}[\rho\varepsilon_{t-1} + u_t, \varepsilon_{t-1}] \\
&= \rho\text{Cov}[\varepsilon_{t-1}, \varepsilon_{t-1}] + \text{Cov}[u_t, \varepsilon_{t-1}] \\
&= \rho\text{Var}[\varepsilon_{t-1}] = \rho\text{Var}[\varepsilon_t] \\
&= \frac{\rho\sigma_u^2}{(1-\rho^2)}
\end{aligned}$$

$$\begin{aligned}
\text{Cov}[\varepsilon_t, \varepsilon_{t-2}] &= \text{Cov}[\rho\varepsilon_{t-1} + u_t, \varepsilon_{t-2}] \\
&= \rho\text{Cov}[\varepsilon_{t-1}, \varepsilon_{t-2}] + \text{Cov}[u_t, \varepsilon_{t-2}] \\
&= \rho\text{Cov}[\varepsilon_t, \varepsilon_{t-1}] \\
&= \frac{\rho^2\sigma_u^2}{(1-\rho^2)} \quad \text{and so on.}
\end{aligned}$$

Autocorrelation Matrix

$$\sigma^2\mathbf{\Omega} = \left(\frac{\sigma_u^2}{1-\rho^2}\right)\begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{bmatrix}$$

(Note, trace $\mathbf{\Omega}$ = n as required.)

Generalized Least Squares

$$\mathbf{\Omega}^{-1/2} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \ldots & 0 \\ -\rho & 1 & 0 & \ldots & 0 \\ 0 & -\rho & 1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & -\rho & 0 \end{bmatrix}$$

$$\mathbf{\Omega}^{-1/2}\mathbf{y} = \begin{pmatrix} \left(\sqrt{1-\rho^2}\right)y_1 \\ y_2 - \rho y_1 \\ y_3 - \rho y_2 \\ \ldots \\ y_T - \rho_{T-1} \end{pmatrix}$$

The Autoregressive Transformation

$$y_t = \mathbf{x_t}'\boldsymbol{\beta} + \varepsilon_t \qquad \varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

$$\rho y_{t-1} = \rho\mathbf{x_{t-1}}'\boldsymbol{\beta} + \rho\varepsilon_{t-1}$$

$$- - - - - - - - - - -$$

$$y_t - \rho y_{t-1} = (\mathbf{x_t} - \rho\mathbf{x_{t-1}})'\boldsymbol{\beta} + (\varepsilon_t - \rho\varepsilon_{t-1})$$

$$y_t - \rho y_{t-1} = (\mathbf{x_t} - \rho\mathbf{x_{t-1}})'\boldsymbol{\beta} + u_t$$

(Note the first observation is lost.)

Unknown Ω

- The problem (of course), Ω is unknown. For now, we will consider two methods of estimation:
  - Two step, or feasible estimation. Estimate Ω first, then do GLS. Emphasize - same logic as White and Newey-West. We don't need to estimate Ω. We need to find a matrix that behaves the same as $(1/n)\mathbf{X}'\Omega^{-1}\mathbf{X}$.
  - Properties of the feasible GLS estimator
- Maximum likelihood estimation of β, σ^2, and Ω all at the same time.
  - Joint estimation of all parameters. Fairly rare. Some generalities…
  - We will examine two applications: Harvey's model of heteroscedasticity and Beach-MacKinnon on the first order autocorrelation model

Sorry to bother you again, but an important issue has come up. I am using LIMDEP to produce results for my testimony in a utility rate case. I have a time series sample of 40 years, and am doing simple OLS analysis using a primary independent variable and a dummy. There is serial correlation present. The issue is what is the BEST available AR1 procedure in LIMDEP for a sample of this type?? I have tried Cochrane-Orcott, Prais-Winsten, and the MLE procedure recommended by Beach-MacKinnon, with slight but meaningful differences.

By modern constructions, your best choice if you are comfortable with AR1 is Prais-Winsten. No one has ever shown that iterating it is better or worse than not. Cochrance-Orcutt is inferior because it discards information (the first observation). Beach and MacKinnon would be best, but it assumes normality, and in contemporary treatments, fewer assumptions is better. **If you are not comfortable with AR1, use OLS with Newey-West and 3 or 4 lags.**

OLS vs. GLS

- OLS
  - Unbiased?
  - Consistent: (Except in the presence of a lagged dependent variable)
  - Inefficient
- GLS
  - Consistent and efficient

The Newey-West Estimator Robust to Autocorrelation

Heteroscedasticity Component - Diagonal Elements

$$\mathbf{S}_0 = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \mathbf{x}_i \mathbf{x}_i'$$

Autocorrelation Component - Off Diagonal Elements

$$\mathbf{S}_1 = \frac{1}{n}\sum_{l=1}^{L}\sum_{t=l+1}^{n} w_l e_t e_{t-l}(\mathbf{x}_t \mathbf{x}_{t-l}' + \mathbf{x}_{t-l}\mathbf{x}_t')$$

$$w_l = 1 - \frac{l}{L+1} \quad = \text{"Bartlett weight"}$$

$$\text{Est.Var}[\mathbf{b}] = \frac{1}{n}\left(\frac{\mathbf{X'X}}{n}\right)^{-1}[\mathbf{S}_0 + \mathbf{S}_1]\left(\frac{\mathbf{X'X}}{n}\right)^{-1}$$

```
----------------------------------------------------------------------------
Ordinary     least squares regression ............
LHS=REALCONS Mean                    =     2999.43578
             Standard deviation      =     1459.70669
             Number of observs.      =            204
Model size   Parameters              =              2
             Degrees of freedom      =            202
Residuals    Sum of squares          =    .153632E+07
             Standard error of e     =       87.20983
Fit          R-squared               =         .99645
             Adjusted R-squared      =         .99643
Model test   F[  1,   202] (prob) = 56669.7(.0000)
Autocorrel   Durbin-Watson Stat.     =         .09205
             Rho = cor[e,e(-1)]      =         .95398
--------+-------------------------------------------------------------------
        |                Standard            Prob.       95% Confidence
REALCONS|  Coefficient     Error         t    |t|>T*         Interval
--------+-------------------------------------------------------------------
Constant|   -80.3547***    14.30585    -5.62  .0000   -108.3937  -52.3158
 REALDPI|     .92169***      .00387   238.05  .0000      .91410    .92927
--------+-------------------------------------------------------------------

Robust VC    Newey-West, Periods   =              5
--------+-------------------------------------------------------------------
        |                Standard            Prob.       95% Confidence
REALCONS|  Coefficient     Error         t    |t|>T*         Interval
--------+-------------------------------------------------------------------
Constant|   -80.3547**     33.53918    -2.40  .0175   -146.0903  -14.6192
 REALDPI|     .92169***      .01220    75.54  .0000      .89777    .94560
--------+-------------------------------------------------------------------
```

```
----------------------------------------------------------------------
AR(1) Model:     e(t) = rho * e(t-1) + u(t)
Initial value of rho          =         .95398
Maximum iterations            =            100
Method = Prais - Winsten
Iter=  1, SS= 140665.304, Log-L=   -957.342
Iter=  2, SS= 138257.782, Log-L=   -955.738
Iter=  3, SS= 135050.589, Log-L=   -953.537
Iter=  4, SS= 129017.792, Log-L=   -949.200
Iter=  5, SS= 121595.615, Log-L=   -943.666
Iter=  6, SS= 118367.007, Log-L=   -941.372
Final value of Rho     =               .998782
Iter=  6, SS= 118367.007, Log-L=   -941.372
Durbin-Watson:   e(t) =                .002436
Std. Deviation:  e(t) =             490.567910
Std. Deviation:  u(t) =              24.206926
Durbin-Watson:   u(t) =               1.994957
Autocorrelation: u(t) =                .002521
N[0,1] used for significance levels
--------+-------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
REALCONS|  Coefficient       Error        z    |z|>Z*         Interval
--------+-------------------------------------------------------------
Constant|    1019.33**     411.1772     2.48   .0132      213.43   1825.22
 REALDPI|     .67343***      .03973    16.95   .0000      .59557    .75129
     RHO|     .99878***      .00346   288.39   .0000      .99199   1.00557
--------+-------------------------------------------------------------
```

Detecting Autocorrelation

- Use residuals
  - Durbin-Watson d= $\dfrac{\Sigma_{t=2}^{T}(e_t - e_{t-1})^2}{\Sigma_{t=1}^{T} e_t^2} \approx 2(1-r)$
  - Assumes normally distributed disturbances strictly exogenous regressors
- Variable addition (Godfrey)
  - $y_t = \beta' x_t + \rho\varepsilon_{t-1} + u_t$
  - Use regression residuals e_t and test $\rho = 0$
  - Assumes consistency of b.

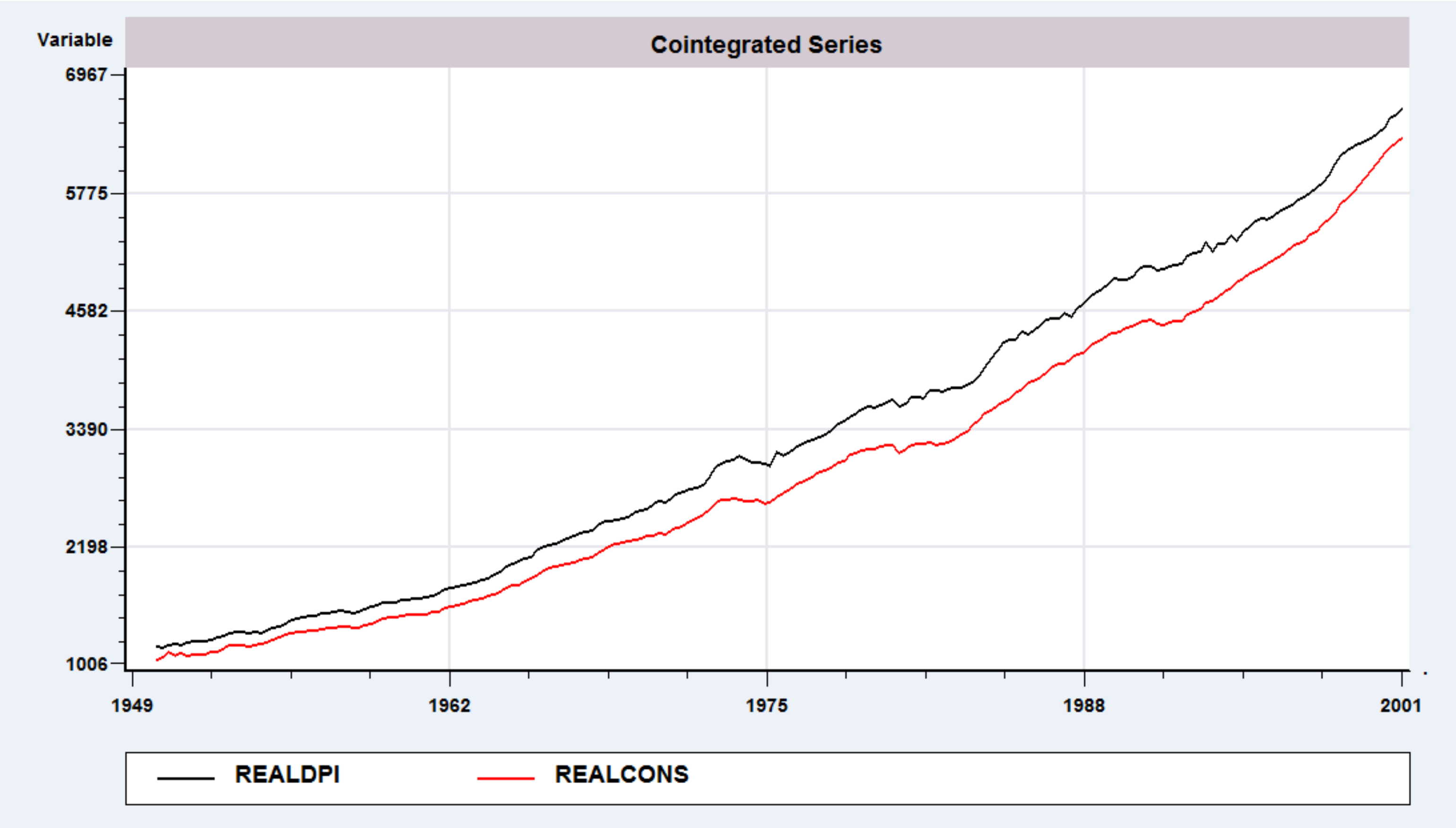
A Unit Root?

- How to test for $\rho = 1$?
- By construction: $\varepsilon_t - \varepsilon_{t-1} = (\rho - 1)\varepsilon_{t-1} + u_t$
  - Test for $\gamma = (\rho - 1) = 0$ using regression?
  - Variance goes to 0 faster than 1/T. Need a new table; can't use standard t tables.
  - Dickey – Fuller tests
- Unit roots in economic data.
  - Nonstationary series
  - Implications for conventional analysis

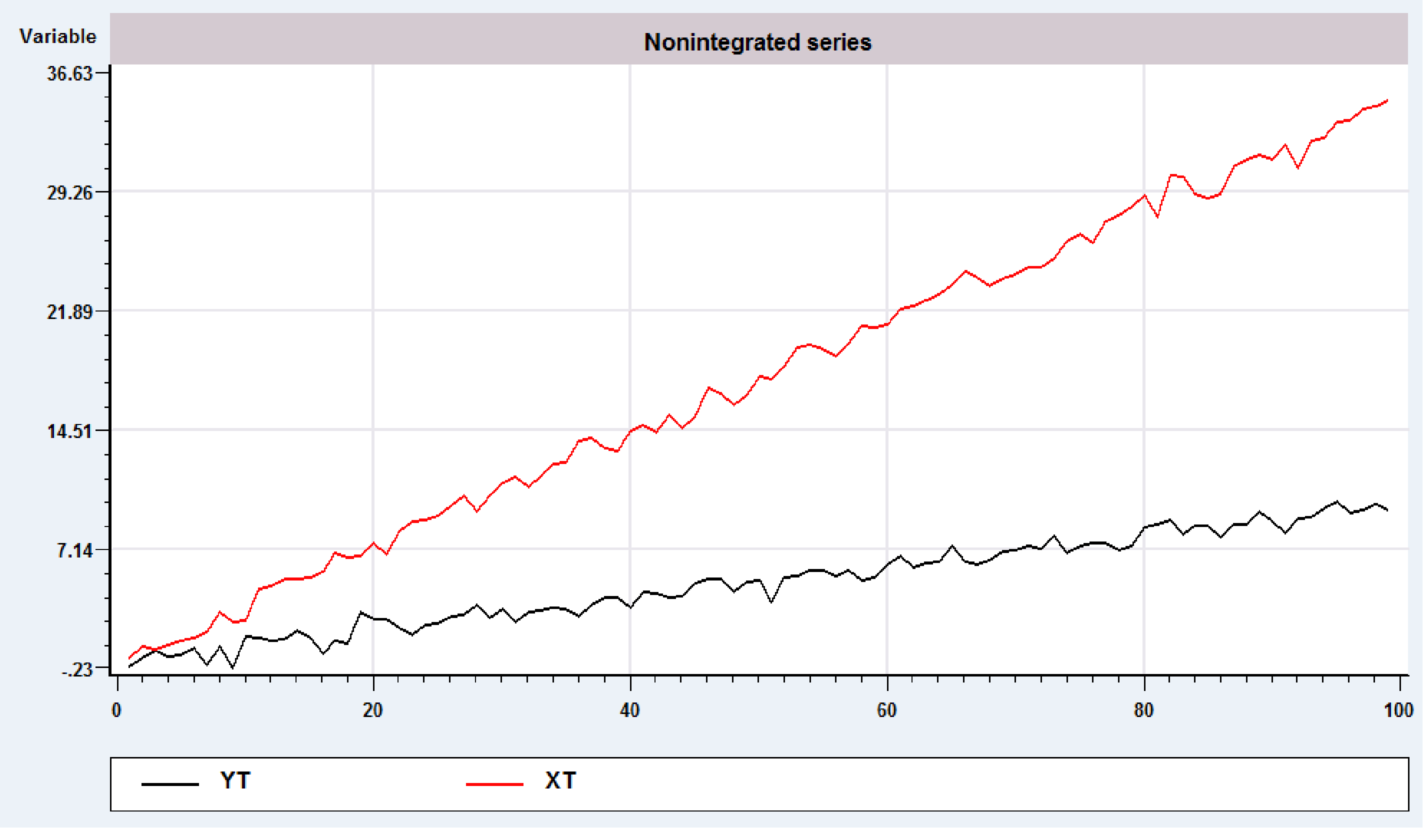
Integrated Processes

- Integration of order (P) when the P'th differenced series is stationary
- Stationary series are I(0)
- Trending series are often I(1). Then $y_t - y_{t-1} = \Delta y_t$ is I(0). [Most macroeconomic data series.]
- Accelerating series might be I(2). Then $(y_t - y_{t-1})\text{-}(y_t - y_{t-1}) = \Delta^2 y_t$ is I(0) [Money stock in hyperinflationary economies. Difficult to find many applications in economics]

Cointegration: Real DPI and Real Consumption



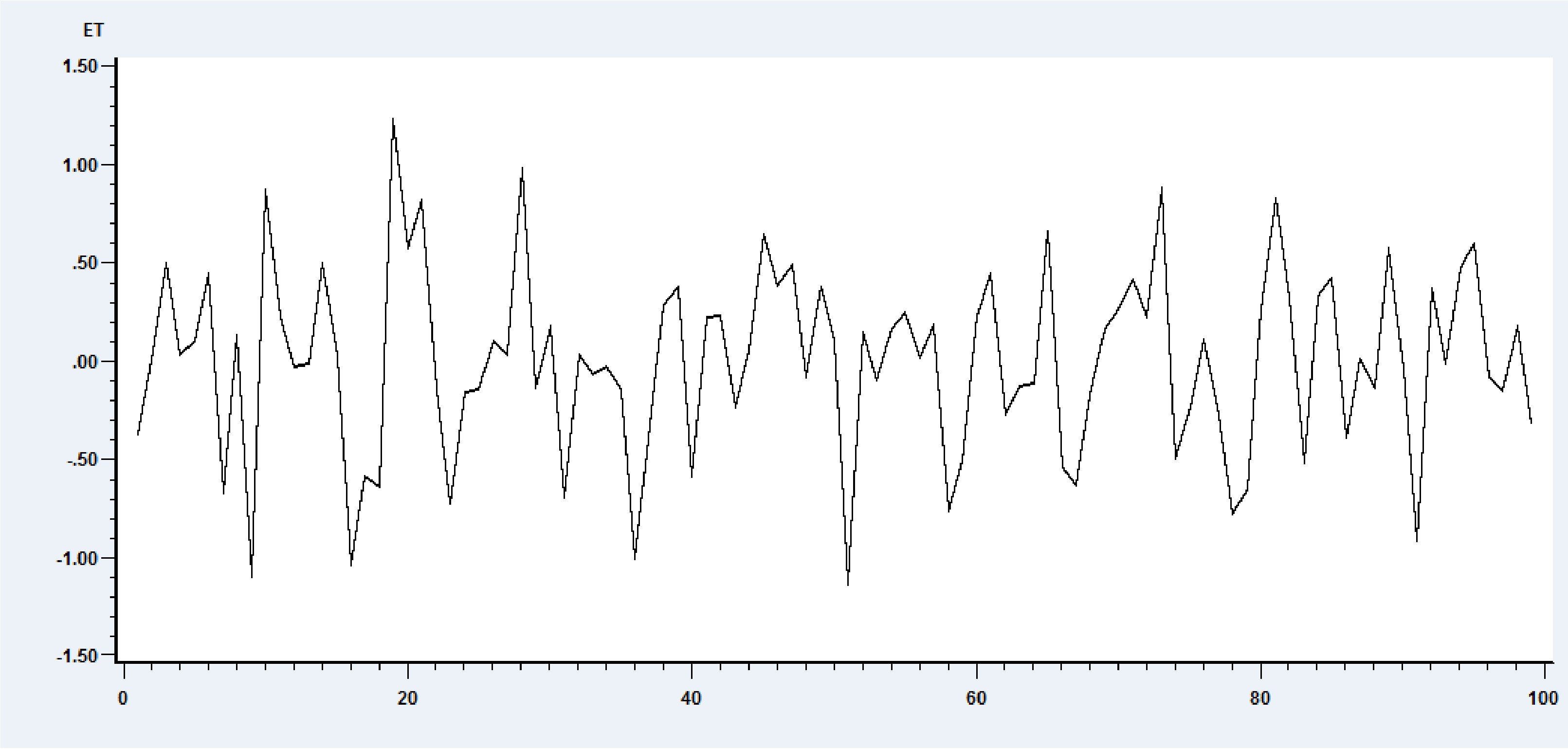
Cointegration – Divergent Series?



Cointegration

- X(t) and y(t) are obviously I(1)
- Looks like any linear combination of x(t) and y(t) will also be I(1)
- Does a model y(t) = bx(t) + u(u) where u(t) is I(0) make any sense? How can u(t) be I(0)?
- In fact, there is a linear combination, [1,-β] that is I(0).
- y(t) = .1*t + noise, x(t) = .2*t + noise
- y(t) and x(t) have a **common trend**
- y(t) and x(t) are cointegrated.

Cointegration and I(0) Residuals



Reinterpreting Autocorrelation

Regression form

$$y_t = \beta' x_t + \varepsilon_t, \ \varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

Error Correction Form

$$y_t - y_{t-1} = \beta'(x_t - x_{t-1}) + \alpha(y_{t-1} - \beta' x_{t-1}) + u_t, (\alpha = \rho - 1)$$

$\beta' x_t =$ the equilibrium

The model describes adjustment of y_t to equilibrium when x_t changes.

```
-----------------------------------------------------------------------------
User Defined Optimization.........................
Nonlinear    least squares regression ............
LHS=DCONS    Mean                 =       26.02069
             Standard deviation   =       24.81678
             Number of observs.   =            203
Model size   Parameters           =              2
             Degrees of freedom   =            201
Residuals    Sum of squares       =        74111.5
             Standard error of e  =       19.10709
Fit          R-squared            =         .40428
             Adjusted R-squared   =         .40721
Model test   F[  1,   201] (prob) =   136.4(.0000)
Not using OLS or no constant. Rsqrd & F may be < 0
--------+--------------------------------------------------------------------
        |                  Standard            Prob.      95% Confidence
UserFunc|  Coefficient       Error       z    |z|>Z*         Interval
--------+--------------------------------------------------------------------
    BETA|      .31932***      .04474     7.14  .0000      .23162     .40701
   ALPHA|      .00926***      .00064    14.53  .0000      .00801     .01051
--------+--------------------------------------------------------------------
```