Macroeconometrics

Topic 2: Univariate Time Series

by Vanessa Berenguer-Rico

University of Oxford

Michaelmas 2016

Univariate Time Series

- ARMA processes
- Non-stationary Time Series (i): Deterministic Trends
- Non-stationary Time Series (ii): Stochastic Trends

ARMA processes

ARMA processes

- Definition
- Properties: Stationarity & Invertibility
- Examples
- Wold Decomposition: $(MA(\infty))$
- Estimation and Inference
- Model Selection

ARMA processes

- We have seen AR(1)

$$X_t = \mu + \phi X_{t-1} + \varepsilon_t,$$

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$

- We can generalize this to an AR(p)

$$X_t = \mu + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t,$$

ARMA processes

- We have also seen a MA(1)

$$X_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1},$$

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$

- We can generalize this to an MA(q)

$$X_t = \mu + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + ... + \theta_q\varepsilon_{t-q},$$

ARMA processes

- We could then combine and get ARMA(p,q)

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q},$$

where $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$

- For instance: ARMA(1,1)

$$X_t - \phi_1 X_{t-1} = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

ARMA processes

- A useful tool in time series: the **Lag Operator** (L)
- Definition: $X_t L = X_{t-1}$
- Properties:
  - $X_t L^j = X_{t-j}$
  - $(aX_t)L = aX_{t-1}$
  - $(X_t + Y_t)L = X_{t-1} + Y_{t-1}$
- Example: AR(1). If $X_t = \mu + \phi X_{t-1} + \varepsilon_t$, then the model can be written as

$$(1 - \phi L)X_t = \mu + \varepsilon_t$$

ARMA processes

- For an ARMA(p,q)

$$X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q},$$

- The model can be written as

$$(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p) X_t = \mu + (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q) \varepsilon_t$$

ARMA processes

- The model can be written as

$$(1-\phi_1 L-\phi_2 L^2-\ldots-\phi_p L^p)X_t=\mu+(1+\theta_1 L+\theta_2 L^2+\ldots+\theta_q L^q)\varepsilon_t$$

- More compactly:

$$\Phi_p(L)X_t=\mu+\Theta_q(L)\varepsilon_t,$$

where

$$\begin{aligned}\Phi_p(L) &= 1-\phi_1 L-\phi_2 L^2-\ldots-\phi_p L^p\\ \Theta_q(L) &= 1+\theta_1 L+\theta_2 L^2+\ldots+\theta_q L^q\end{aligned}$$

are the autoregressive and moving average polynomials (in L), respectively

ARMA processes

- For an ARMA(1,1)

$$X_t - \phi_1 X_{t-1} = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

- The model can be written as

$$(1 - \phi_1 L) X_t = \mu + (1 + \theta_1 L) \varepsilon_t$$

- More compactly:

$$\Phi_1(L) X_t = \mu + \Theta_1(L) \varepsilon_t,$$

  where $\Phi_p(L) = 1 - \phi_1 L$ and $\Theta_q(L) = 1 + \theta_1 L$ are the autoregressive and moving average polynomials (in L), respectively

ARMA processes

- Properties of ARMA processes:

$$\Phi_p(L)X_t = \mu + \Theta_q(L)\varepsilon_t,$$

- Stability and Invertibility

- **Definition**: The ARMA process X_t is **stable** (causal) if the roots of the autoregressive polynomial, $\Phi_p(L)$, are outside the unit circle

- **Definition**: The ARMA process X_t is **invertible** if the roots of the moving average polynomial, $\Theta_q(L)$, are outside the unit circle

ARMA processes

- For an AR(1)

$$X_t - \phi_1 X_{t-1} = \mu + \varepsilon_t$$

- Is the model stable (causal)?
  Check: The roots of $\Phi_1(L) = 0$ have to be outside the unit circle ($|L| > 1$)

$$1 - \phi_1 L = 0 \text{ so that } L = 1/\phi_1$$

  Hence, the model is stationary if $|\phi_1| < 1$

- Is the model invertible? Yes, by construction

ARMA processes

- For a MA(1)

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

- Is the model stable? Yes, by construction

- Is the model invertible?
  Check: The roots of $\Theta_1(L) = 0$ have to be outside the unit circle ($|L| > 1$)

$$1 - \theta_1 L = 0 \text{ so that } L = -1/\theta_1$$

  Hence, the model is invertible if $|\theta_1| < 1$

ARMA processes

- For an ARMA(1,1)

$$X_t - \phi_1 X_{t-1} = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

- Is the model stable?
  Check: The roots of $\Phi_1(L) = 0$ have to be outside the unit circle ($|L| > 1$)

$$1 - \phi_1 L = 0 \text{ so that } L = 1/\phi_1$$

  Hence, the model is stationary if $|\phi_1| < 1$

- Is the model invertible?
  Check: The roots of $\Theta_1(L) = 0$ have to be outside the unit circle ($|L| > 1$)

$$1 + \theta_1 L = 0 \text{ so that } L = -1/\theta_1$$

  Hence, the model is invertible if $|\theta_1| < 1$

ARMA processes

- Representations of **stable and invertible** ARMA processes:

$$\Phi_p(L)X_t = \Theta_q(L)\varepsilon_t$$

- MA representation

$$X_t = \Phi_p(L)^{-1}\Theta_q(L)\varepsilon_t$$

- AR representation

$$\Theta_q(L)^{-1}\Phi_p(L)X_t = \varepsilon_t$$

ARMA processes

- **The Wold Decomposition**: If X_t is a stationary and non-deterministic process, then

$$X_t = \sum_{j=0}^{\infty} \Psi_j u_{t-j} + Z_t = \Psi(L) u_t + Z_t$$

where:

- $\Psi_0 = 1$ and $\sum_{j=0}^{\infty} \Psi_j^2 < \infty$
- u_t is $WN(0, \sigma_u^2)$ with $\sigma_u^2 > 0$
- Z_t is deterministic
- $Cov(u_t.Z_t) = 0$ for all s and t
- Ψ_j and u_t are unique

ARMA processes

The Wold Decomposition and the $MA(\infty)$ representation

- Wold Decomposition of a stationary process

$$X_t = \sum_{j=0}^{\infty} \Psi_j u_{t-j} + Z_t = \Psi(L) u_t + Z_t$$

- For a stationary ARMA process

$$\Phi_p(L) X_t = \mu + \Theta_q(L) \varepsilon_t$$

the MA representation is

$$X_t = \Phi_p(L)^{-1} \Theta_q(L) \varepsilon_t + \Phi_p(L)^{-1} \mu$$

ARMA processes

- Let
$$X_t - \phi X_{t-1} = \mu + \varepsilon_t$$
with $|\phi| < 1$ (so the process is stationary)
- The MA representation is
$$\begin{aligned} X_t &= (1 - \phi L)^{-1}\varepsilon_t + (1 - \phi L)^{-1}\mu \\ &= \frac{1}{1 - \phi L}\varepsilon_t + \frac{1}{1 - \phi L}\mu \end{aligned}$$
- Notice that
$$\frac{1}{1 - \phi L} = (1 + \phi L + \phi^2 L^2 + \phi^3 L^3 + ...)$$
- Hence
$$X_t = \sum_{j=0}^{\infty} \phi^j \varepsilon_{t-j} + \frac{\mu}{1 - \phi}$$

ARMA processes

- Let
$$X_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1}$$
with $|\theta| < 1$ (so the process is invertible)
- The AR representation is
$$(1+\theta L)^{-1}X_t - (1+\theta L)^{-1}\mu = \varepsilon_t$$
- Notice that
$$\frac{1}{1+\theta L} = \frac{1}{1-(-\theta L)} = (1+(-\theta L)+(-\theta L)^2+(-\theta L)^3+\ldots)$$
- Hence
$$X_t = \frac{\mu}{1+\theta} + \sum_{j=1}^{\infty}(-1)^{j+1}\theta^j X_{t-j} + \varepsilon_t$$

ARMA processes

- Autocorrelation function of an ARMA(1,1) process

$$(1-\phi L)X_t = (1-\theta L)\varepsilon_t$$

with $|\phi| < 1$ and $|\theta| < 1$ (so the process is stable and invertible)

- Write

$$X_t X_{t-h} = \phi X_{t-1} X_{t-h} + \varepsilon_t X_{t-h} - \theta \varepsilon_{t-1} X_{t-h}$$

- Take expectations

$$\gamma_X(h) = \phi \gamma_X(h-1) + E(\varepsilon_t X_{t-h}) - \theta E(\varepsilon_{t-1} X_{t-h})$$

ARMA processes

$$\gamma_X(h) = \phi\gamma_X(h-1) + E(\varepsilon_t X_{t-h}) - \theta E(\varepsilon_{t-1} X_{t-h})$$

- This gives as this system of equations:
  $h = 0 : \gamma_X(0) = \phi\gamma_X(1) + \sigma_\varepsilon^2 - \theta(\phi - \theta)\sigma_\varepsilon^2$
  $h = 1 : \gamma_X(1) = \phi\gamma_X(0) - \theta\sigma_\varepsilon^2$
  $h \geq 2 : \gamma_X(h) = \phi\gamma_X(h-1)$

- Hence

$$\rho_X(h) = \begin{cases} 1 & h = 0 \\ \frac{(\phi-\theta)(1-\phi\theta)}{1+\theta^2-2\phi\theta} & h = 1 \\ \phi\rho_X(h-1) & h = 2 \end{cases}$$

ARMA processes: Estimation and Inference

- ARMA process (stable and invertible)

$$\Phi_p(L)X_t = \Theta_q(L)\varepsilon_t$$

with $\varepsilon_t \sim i.i.d.(0, \sigma_\varepsilon^2)$

- MA representation of the ARMA process

$$X_t = \Psi(L)\varepsilon_t$$

where $\Psi(L) = \Phi_p(L)^{-1}\Theta_q(L)$

ARMA processes: Estimation and Inference

- **Law of Large Numbers for the mean**: Let X_t be a covariance-stationary process with $E(X_t) = \mu$ and absolutely summable autocovariances, that is$\sum_{h=-\infty}^{\infty} |\gamma_h| < \infty$. Then

$$\frac{1}{n}\sum_{t=1}^{n} X_t \xrightarrow{p} \mu$$

  as $n \to \infty$. (see Hamilton p. 188 or Brockwell and Davis, p.58 or Hamilton p. 401)

- **Note**: Any covariance-stationary ARMA(p,q) process (roots of the autoregressive polynomial outside the unit circle) satisfies the conditions of this theorem.

ARMA processes: Estimation and Inference

- Recall: MA representation of the ARMA process

$$X_t = \mu + \Psi(L)\varepsilon_t$$

- **Central Limit Theorem for the mean**: If $X_t = \mu + \Psi(L)\varepsilon_t$ where $\varepsilon_t \sim i.i.d.(0, \sigma_\varepsilon^2)$ and $\sum_{j=0}^{\infty} |\psi_j| < \infty$, then as $n \to \infty$,

$$\sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow{d} N\left(0, \sum_{h=-\infty}^{\infty} \gamma_h\right)$$

  where $\sum_{j=-\infty}^{\infty} \gamma_j = \sigma_\varepsilon^2 \Psi^2(1)$ is the long run variance. (See Hamilton, p.195 or Brockwell and Davis, p.58 or Hamilton p. 402)

ARMA processes: Estimation and Inference

- Example AR(1): (see Hamilton p.215)

$$X_t = \phi X_{t-1} + \varepsilon_t$$

where $|\phi| < 1$ and $\varepsilon_t \sim i.i.d.(0, \sigma_\varepsilon^2)$

- Then

$$\sqrt{n}\left(\frac{1}{n}\sum_{t=1}^{n} X_t\right) \xrightarrow{d} N\left(0, \frac{\sigma_\varepsilon^2}{(1-\phi)^2}\right)$$

where $\sigma_\varepsilon^2 \Psi^2(1)$ is the long run variance

ARMA processes: Estimation and Inference

- Estimation of an AR(1)

$$X_t = \phi X_{t-1} + \varepsilon_t$$

where $|\phi| < 1$ and $\varepsilon_t \sim i.i.d.(0, \sigma_\varepsilon^2)$

- **OLS estimation**:

$$\hat{\phi}_n = \frac{\sum_{t=2}^n X_{t-1} X_t}{\sum_{t=2}^n X_{t-1}^2} = \phi + \frac{\sum_{t=2}^n X_{t-1} \varepsilon_t}{\sum_{t=2}^n X_{t-1}^2}$$

ARMA processes: Estimation and Inference

- Properties

$$\hat{\phi}_n = \frac{\sum_{t=2}^{n} X_{t-1}X_t}{\sum_{t=2}^{n} X_{t-1}^2} = \phi + \frac{\sum_{t=2}^{n} X_{t-1}\varepsilon_t}{\sum_{t=2}^{n} X_{t-1}^2}$$

- Denominator (LLN for stationary and ergodic processes)

$$\frac{1}{n}\sum_{t=2}^{n} X_{t-1}^2 \xrightarrow{p} E(X_{t-1}^2) = \frac{\sigma_\varepsilon^2}{1-\phi^2}$$

- Numerator (CLT for m.d.s.)

$$\frac{1}{\sqrt{n}}\sum_{t=2}^{n} X_{t-1}\varepsilon_t \xrightarrow{d} N(0, \frac{\sigma_\varepsilon^4}{1-\phi^2})$$

ARMA processes: Estimation and Inference

- Therefore,

$$\sqrt{n}(\hat{\phi}_n - \phi) = \frac{\frac{1}{\sqrt{n}} \sum_{t=2}^{n} X_{t-1}\varepsilon_t}{\frac{1}{n} \sum_{t=2}^{n} X_{t-1}^2} \xrightarrow{d} N(0, 1 - \phi^2).$$

- We can conduct standard inference

- See Hamilton, p. 215 for the estimation of a general AR(p)

ARMA processes: Selection

- AR(p): How should we choose p in practice?
- Different philosophies: General to Specific vs Specific to General
- Tools: significance test, information criteria

ARMA processes: Selection

- Information Criteria: Suppose we have $\bar{k}$ alternative models, $M_1, ..., M_{\bar{k}}$ where $k = 1, ..., \bar{k}$ represents the number of parameters in the models. We choose the model that minimizes the information criteria

$$IC(k) = \ln \hat{\sigma}_k + k\frac{P(n)}{n}$$

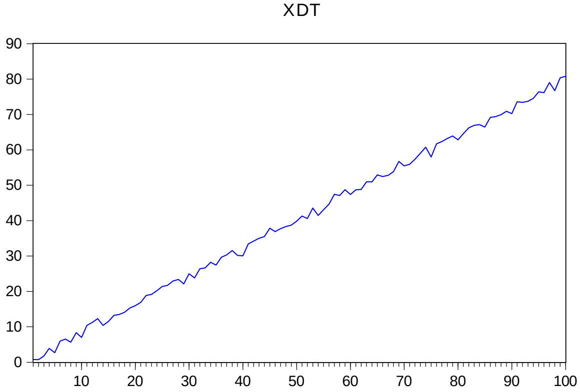
where $\hat{\sigma}_k$ is the variance of the residuals of model M_k, n is the sample size and $P(n)$ is a penalty

- There are different proposals for $P(n)$ in the literature
  - AIC (Akaike): $P(n) = 2$
  - BIC (Bayesina): $P(n) = \ln(n)$
  - HQ (Hannan-Quin): $P(n) = 2\ln(\ln(n))$

Deterministic Trends

Deterministic Trends

$$y_t = \alpha + \delta t + \varepsilon_t; \quad \varepsilon_t \sim i.i.d.\,(0, 1)$$



Deterministic Trends

- Let $\varepsilon_t \sim i.i.d. \left(0, \sigma^2\right)$
- Deterministic trends. Example:

$$y_t = \alpha + \delta t + \varepsilon_t$$

- Stochastic Properties

$$E\left[x_t\right] = \alpha + \delta t$$

$$V\left[x_t\right] = \sigma^2$$

$$Cov\left[x_t, x_s\right] = 0$$

Deterministic Trends: Estimation and Testing

Hamilton: Chapter 16

- $y_t = \alpha + \delta t + \varepsilon_t$: Estimate α and δ by OLS
- Asymptotic theory slightly different from the case of iid regressors
- OLS estimates in general have different rates of convergence
- Nevertheless, usual t and F statistics have the usual asymptotic distributions in this case

Deterministic Trends: Estimation

The model

$$y_t = \alpha + \delta t + \varepsilon_t,$$

can be written in standard regression model form as follows

$$y_t = x_t'\beta + \varepsilon_t,$$

where

$$\underset{(1\times 2)}{x_t'} \equiv \begin{bmatrix} 1 & t \end{bmatrix},$$

and

$$\underset{(2\times 1)}{\beta} \equiv \begin{bmatrix} \alpha \\ \delta \end{bmatrix}.$$

Deterministic Trends: Estimation

Let $\hat{b}_T$ denote the OLS estimate of β based on a sample of size T

$$\hat{b}_T \equiv \begin{bmatrix} \hat{\alpha}_T \\ \hat{\delta}_T \end{bmatrix} = \left[\sum_{t=1}^{T} x_t x_t' \right]^{-1} \left[\sum_{t=1}^{T} x_t y_t \right].$$

It is simple to see that

$$\left(\hat{b}_T - \beta \right) = \left[\sum_{t=1}^{T} x_t x_t' \right]^{-1} \left[\sum_{t=1}^{T} x_t \varepsilon_t \right],$$

or equivalently

$$\begin{bmatrix} \hat{\alpha}_T - \alpha \\ \hat{\delta}_T - \delta \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{T} 1 & \sum_{t=1}^{T} t \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^{T} \varepsilon_t \\ \sum_{t=1}^{T} t\varepsilon_t \end{bmatrix}$$

Deterministic Trends: Estimation

Notice that

$$
\begin{aligned}
\left[\sum_{t=1}^{T} x_t x_t'\right] &= \left[\begin{array}{cc} \sum_{t=1}^{T} 1 & \sum_{t=1}^{T} t \\ \sum_{t=1}^{T} t & \sum_{t=1}^{T} t^2 \end{array}\right] \\
&= \left[\begin{array}{cc} T & T\left(T+1\right)/2 \\ T\left(T+1\right)/2 & T\left(T+1\right)\left(2T+1\right)/6 \end{array}\right]
\end{aligned}
$$

Deterministic Trends: Estimation

Now, let

$$\Upsilon_T = \begin{bmatrix} T^{1/2} & 0 \\ 0 & T^{3/2} \end{bmatrix},$$

and notice that

$$\begin{aligned}
& \Upsilon_T^{-1} \left[\sum_{t=1}^{T} x_t x_t' \right] \Upsilon_T^{-1} \\
= \quad & \begin{bmatrix} T^{-1} \sum_{t=1}^{T} 1 & T^{-2} \sum_{t=1}^{T} t \\ T^{-2} \sum_{t=1}^{T} t & T^{-3} \sum_{t=1}^{T} t^2 \end{bmatrix} \\
= \quad & \begin{bmatrix} T^{-1} T & T^{-2} T \left(T+1 \right) / 2 \\ T^{-2} T \left(T+1 \right) / 2 & T^{-3} T \left(T+1 \right) \left(2T+1 \right) / 6 \end{bmatrix} \\
\longrightarrow \quad & \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{bmatrix} \equiv Q.
\end{aligned}$$

Deterministic Trends: Estimation

Recall

$$\Upsilon_T = \begin{bmatrix} T^{1/2} & 0 \\ 0 & T^{3/2} \end{bmatrix}.$$

If $\varepsilon_t \sim i.i.d. \left(0, \sigma^2\right)$ and $E\left[\varepsilon_t^4\right] < \infty$, then

$$\Upsilon_T^{-1}\left[\sum_{t=1}^{T} x_t \varepsilon_t\right] = \begin{bmatrix} \left(1/\sqrt{T}\right) \sum_{t=1}^{T} \varepsilon_t \\ \left(1/\sqrt{T}\right) \sum_{t=1}^{T} \left(t/T\right) \varepsilon_t \end{bmatrix} \xrightarrow{d} N\left(0, \sigma^2 Q\right).$$

Deterministic Trends: Estimation

Therefore,

$$\left[\begin{array}{c} T^{1/2}\left(\hat{\alpha}_T-\alpha\right) \\ T^{3/2}\left(\hat{\delta}_T-\delta\right) \end{array}\right] \xrightarrow{d} N\left(0,\left[Q^{-1}\sigma^2 Q Q^{-1}\right]\right)=N\left(0,\sigma^2 Q^{-1}\right).$$

Deterministic Trends: Estimation

Theorem

Let $y_t = \alpha + \delta t + \varepsilon_t$ where ε_t is i.i.d. $(0, \sigma^2)$ and $E\left[\varepsilon_t^4\right] < \infty$. Then,

$$\left[\begin{array}{c} T^{1/2}\left(\hat{\alpha}_T - \alpha\right) \\ T^{3/2}\left(\hat{\delta}_T - \delta\right) \end{array}\right] \xrightarrow{d} N\left(\left[\begin{array}{c} 0 \\ 0 \end{array}\right], \sigma^2 \left[\begin{array}{cc} 1 & 1/2 \\ 1/2 & 1/3 \end{array}\right]^{-1}\right).$$

Remark: $\hat{\delta}_T$ is superconsistent!

Deterministic Trends: Testing

- Null Hypothesis

$$H_o : \alpha = \alpha_0$$

- Test statistic

$$t_\alpha = \frac{\hat{\alpha}_T - \alpha_0}{\left\{ s_T^2 \left[\begin{array}{cc} 1 & 0 \end{array} \right] \left[\sum_{t=1}^{T} x_t x_t' \right]^{-1} \left[\begin{array}{c} 1 \\ 0 \end{array} \right] \right\}^{1/2}},$$

where

$$s_T^2 = \frac{1}{T-2} \sum_{t=1}^{T} \left(y_t - \hat{\alpha}_T - \hat{\delta}_T t \right)^2$$

Deterministic Trends: Testing

- Asymptotic Distribution

$$t_{\alpha} = \frac{\hat{\alpha}_T - \alpha_0}{\left\{ s_T^2 \begin{bmatrix} 1 & 0 \end{bmatrix} \left[\sum_{t=1}^{T} x_t x_t' \right]^{-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}^{1/2}} \xrightarrow{d} N(0,1)$$

Deterministic Trends: Testing

- Null Hypothesis

$$H_o : \delta = \delta_0$$

- Test statistic

$$t_\delta = \frac{\hat{\delta}_T - \delta_0}{\left\{ s_T^2 \begin{bmatrix} 0 & 1 \end{bmatrix} \left[\sum_{t=1}^{T} x_t x_t' \right]^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}^{1/2}},$$

where again

$$s_T^2 = \frac{1}{T-2} \sum_{t=1}^{T} \left(y_t - \hat{\alpha}_T - \hat{\delta}_T t \right)^2$$

Deterministic Trends: Testing

- Asymptotic Distribution

$$t_\delta = \frac{\hat{\delta}_T - \delta_0}{\left\{ s_T^2 \begin{bmatrix} 0 & 1 \end{bmatrix} \left[\sum_{t=1}^{T} x_t x_t' \right]^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}^{1/2}} \xrightarrow{d} N(0,1)$$

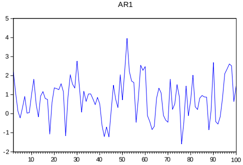
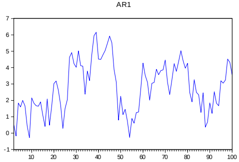
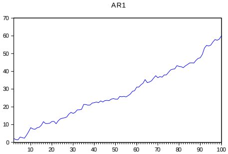
Deterministic Trends: Estimation and Testing

- $\hat{\alpha}_T$ and $\hat{\delta}_T$ converge at different rates
- The corresponding standard errors also incorporate different orders of T
- Hence, the usual OLS t tests are asymptotically valid

Stochastic Trends

Unit Roots

- Let $\varepsilon_t \sim i.i.d.\,(0,1)$, $x_0 = 0$ and consider the following DGPs:

$x_t = 0.5 + 0.3x_{t-1} + \varepsilon_t$;	$x_t = 0.5 + 0.7x_{t-1} + \varepsilon_t$	$x_t = 0.5 + x_{t-1} + \varepsilon_t$
		

Unit Roots

- Recall:

$$x_t = \beta + x_{t-1} + u_t; \quad u_t \sim i.i.d. \left(0, \sigma^2\right)$$

$$x_t = x_0 + \beta t + \sum_{j=1}^{t} u_j$$

- Stochastic Properties

$$E\left[x_t\right] = x_0 + \beta t$$

$$V\left[x_t\right] = \sigma^2 t$$

$$Cov\left[x_t, x_s\right] = \min\left\{t, s\right\} \sigma^2$$

Unit Roots

Some Properties of Unit Root processes:

- First difference is stationary:

$$x_t = \beta + x_{t-1} + u_t \quad vs \quad \Delta x_t = \beta + u_t$$

- Shocks have a permanent effect on the future of the series

$$x_t = x_0 + \beta t + \sum_{j=1}^{t} u_j$$

- Standard inference does not hold...

Unit Roots

Hamilton Chapter 17: Consider the OLS estimation of the AR(1) process,

$$y_t = \rho y_{t-1} + u_t,$$

where $u_t \sim i.i.d.N\left(0, \sigma^2\right)$ and $y_0 = 0$. The OLS estimate of ρ is given by

$$\hat{\rho}_T = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2} = \rho + \frac{\sum_{t=1}^{T} y_{t-1} u_t}{\sum_{t=1}^{T} y_{t-1}^2}.$$

Unit Roots

- If $|\rho| < 1$, then the LLN and the CLT can be applied to obtain the asymptotic distribution of the OLS estimator of ρ. See Hamilton (p. 215)
- LLN:

$$\frac{1}{T}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{p} E\left[y_{t-1}^2\right] = \frac{\sigma^2}{1-\rho^2}.$$

- CLT:

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T} y_{t-1}u_t \xrightarrow{d} N\left(0, E\left[y_{t-1}^2 u_t^2\right]\right) = N\left(0, \frac{\sigma^4}{1-\rho^2}\right)$$

- Therefore,

$$T^{1/2}\left(\hat{\rho}_T - \rho\right) \xrightarrow{d} N\left(0, 1-\rho^2\right)$$

Unit Roots

- If $\rho = 1$, the distribution collapses; that is, $T^{1/2}\left(\hat{\rho}_T - \rho\right) \xrightarrow{p} 0$: not very helpful for hypothesis testing
- To obtain a non-degenerate asymptotic distribution: $T\left(\hat{\rho}_T - \rho\right)$
- Faster than the stationary case $\left(T^{1/2}\right)$ but at slower than the deterministic trend case $\left(T^{3/2}\right)$
- Asymptotic distribution when $\rho = 1$ non-standard. Can be described in terms of functionals of **Brownian motions!**

Brownian Motion

Definition

A Standard Brownian motion $W(.)$ is a continuous-time stochastic process, associating each date $r \in [0, 1]$ with the scalar $W(r)$ such that:

(a) $W(0) = 0$

(b) For any dates $0 \leq r_1 < r_2 < ... < r_k \leq 1$, the changes $[W(r_2) - W(r_1)], [W(r_3) - W(r_2)], ..., [W(r_k) - W(r_{k-1})]$ are independent Gaussian with $[W(s) - W(r)] \sim N(0, s - r)$

(c) For a given realization, $W(r)$ is continuous in r with probability 1

The Functional Central Limit Theorem

- The CLT establishes convergence of random variables, the FCLT establishes conditions for convergence of random functions
- Let ε_t be an *i.i.d.* $(0, \sigma^2)$ sequence
- The CLT considers

$$T^{1/2}\bar{\varepsilon}_T = T^{1/2}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t$$

- The FCLT considers

$$T^{1/2}X_T(r) = T^{1/2}\frac{1}{T}\sum_{t=1}^{[Tr]}\varepsilon_t$$

The Functional Central Limit Theorem

- Consider

$$X_T\left(r\right) = \frac{1}{T}\sum_{t=1}^{[Tr]} \varepsilon_t,$$

where $r \in [0, 1]$, $[Tr]$ denotes the integer part of Tr

- For any given realization, $X_T\left(r\right)$ is a step **function** in r:

$$X_T\left(r\right) = \begin{cases} 0 & 0 \leq r < 1/T \\ \varepsilon_1/T & 1/T \leq r < 2/T \\ \left(\varepsilon_1 + \varepsilon_2\right)/T & 2/T \leq r < 3/T \\ \vdots & \vdots \\ \left(\varepsilon_1 + \varepsilon_2 + ... + \varepsilon_T\right)/T & r = 1 \end{cases}$$

The Functional Central Limit Theorem

The simplest FCLT is known as Donsker's theorem (Donsker, 1951)

Theorem

Let ε_t be a sequence of i.i.d. random variables with mean zero. If $\sigma^2 \equiv var\left(\varepsilon_t\right) < \infty$, $\sigma^2 \neq 0$, then

$$T^{1/2} X_T\left(r\right)/\sigma \xrightarrow{d} W\left(r\right)$$

The Continuous Mapping Theorem

The Continuous Mapping Theorem, CMT, states that if $X_T(.) \xrightarrow{d} X(.)$ and g is a continuos functional, then $g(X_T(.)) \xrightarrow{d} g(X(.))$

The Continuous Mapping Theorem

- Example: $S_T(r) = T^{1/2} X_T(r) \xrightarrow{d} \sigma W(r)$
- Example: $S_T^2(r) = \left[T^{1/2} X_T(r)\right]^2 \xrightarrow{d} \sigma^2 \left[W(r)\right]^2$
- Example: $\int_0^1 S_T(r)\, dr = \int_0^1 T^{1/2} X_T(r)\, dr \xrightarrow{d} \sigma \int_0^1 W(r)\, dr$
- Example:
  $\int_0^1 S_T^2(r)\, dr = \int_0^1 \left[T^{1/2} X_T(r)\right]^2 dr \xrightarrow{d} \sigma^2 \int_0^1 \left[W(r)\right]^2 dr$

Application to Unit Root Processes

- Consider the random walk

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

where $\varepsilon_t \sim i.i.d. \left(0, \sigma^2\right)$, and $y_0 = 0$, so that

$$y_t = \sum_{j=1}^{t} \varepsilon_j$$

Application to Unit Root Processes

$$y_t = \sum_{j=1}^{t} \varepsilon_j$$

- Then, $X_T(r)$ can be consturcted as follows

$$X_T(r) \begin{cases} 0 & 0 \le r < 1/T \\ y_1/T = \varepsilon_1/T & 1/T \le r < 2/T \\ y_2/T = (\varepsilon_1 + \varepsilon_2)/T & 2/T \le r < 3/T \\ \vdots & \vdots \\ y_T/T = (\varepsilon_1 + \varepsilon_2 + ... + \varepsilon_T)/T & r = 1 \end{cases}$$

Application to Unit Root Processes

- Notice that

$$\int_0^1 X_T(r)\,dr = \frac{y_1}{T^2} + \frac{y_2}{T^2} + \ldots + \frac{y_T}{T^2} = \frac{1}{T^2}\sum_{t=1}^{T} y_t$$

- Hence,

$$\frac{1}{T^{3/2}}\sum_{t=1}^{T} y_t = \int_0^1 T^{1/2} X_T(r)\,dr \xrightarrow{d} \sigma \int_0^1 W(r)\,dr$$

- Similarly,

$$\frac{1}{T^2}\sum_{t=1}^{T} y_t^2 = \int_0^1 \left[T^{1/2} X_T(r)\right]^2 dr \xrightarrow{d} \sigma^2 \int_0^1 [W(r)]^2\,dr$$

Application to Unit Root Processes

- Recall, if

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

where $\varepsilon_t \sim i.i.d.\left(0, \sigma^2\right)$, then the OLS estimator of ρ is

$$\hat{\rho}_T = \frac{\sum_{t=1}^{T} y_{t-1} y_t}{\sum_{t=1}^{T} y_{t-1}^2} = \rho + \frac{\sum_{t=1}^{T} y_{t-1} \varepsilon_t}{\sum_{t=1}^{T} y_{t-1}^2}$$

- If $\rho = 1$

$$(\hat{\rho}_T - 1) = \frac{\sum_{t=1}^{T} y_{t-1} \varepsilon_t}{\sum_{t=1}^{T} y_{t-1}^2}.$$

Application to Unit Root Processes

- For the numerator, notice that $y_t^2 = y_{t-1}^2 + \varepsilon_t^2 + 2y_{t-1}\varepsilon_t$. Hence,

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\varepsilon_t &= \frac{1}{2}\left(\frac{1}{T}\sum_{t=1}^{T}\left(y_t^2 - y_{t-1}^2\right) - \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2\right) \\
&= \frac{1}{2}\left(\frac{1}{T}y_T^2 - \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2\right) \\
&\xrightarrow{d} \frac{1}{2}\sigma^2\left(W^2(1) - 1\right)
\end{aligned}$$

- And for the denominator

$$\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{d} \sigma^2 \int_0^1 [W(r)]^2 \, dr$$

Application to Unit Root Processes

- Therefore,

$$T\left(\hat{\rho}_T - 1\right) = \frac{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\varepsilon_t}{\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2} \xrightarrow{d} \frac{\left(W^2\left(1\right) - 1\right)}{2\int_0^1 \left[W\left(r\right)\right]^2 dr}$$

- **Remark 1**: The OLS estimator converges at a rate T: Super-consistent!

- **Remark 2**: The asymptotic distribution is not standard

Testing for Unit Roots

- Unit Root tests: hypotheses testing procedures for a unit root
- This course focuses on: **testing the null of a unit root vs the alternative of trend-stationary** (there are many other types)
- In particular, we will consider one of the most popular test for unit roots: the Dickey-Fuller (DF) test
- For an overview on unit root testing: Phillips and Xiao (1998)

Testing for Unit Roots

Some features of the DF test:

- The asymptotic distribution is not standard
- Deterministic components in true model and/or auxiliary regression affect the asymptotic distribution
- Different tables of critical values have to be used in each case

The Dickey-Fuller test

- The Hypotheses

$$\begin{cases} H_o : y_t \sim I(1) \\ H_a : y_t \sim I(0) \end{cases}$$

- The Auxiliary Regression

$$y_t = \rho y_{t-1} + u_t,$$

  and hence

$$\begin{cases} H_o : y_t \sim I(1) \equiv \rho = 1 \\ H_a : y_t \sim I(0) \equiv \rho < 1 \end{cases}$$

- Equivalently,

$$\Delta y_t = \theta y_{t-1} + u_t,$$

  where $\theta = (\rho - 1)$ and hence

$$\begin{cases} H_o : y_t \sim I(1) \equiv \theta = 0 \\ H_a : y_t \sim I(0) \equiv \theta < 0 \end{cases}$$

The Dickey-Fuller test

- Three possible specifications to consider deterministic components:

- **No Deterministic Components**:

$$(i) \quad \Delta y_t = \theta y_{t-1} + u_t$$

- **Constant Term**:

$$(ii) \quad \Delta y_t = \alpha + \theta y_{t-1} + u_t$$

- **Linear Trend**:

$$(iii) \quad \Delta y_t = \alpha + \beta t + \theta y_{t-1} + u_t$$

The Dickey-Fuller test

- Which specification to use in practice?
- It is convenient to use an auxiliary regression that is able to explain both H_o and H_1
- A graphical simple device:
- If the data looks trended, then (iii) would be a reasonable specification under both hypothesis
- Otherwise, (ii) is recommended

The Dickey-Fuller test

- Auxiliary Regression

$$\Delta y_t = f(t) + \theta y_{t-1} + u_t,$$

where $\theta = (\rho - 1)$

- Two scenarios:

(a) u_t uncorrelated: DF test

(b) u_t correlated: Augmented DF (ADF) test, etc...

The Dickey-Fuller test

Consider the following case

$$y_t = y_{t-1} + u_t \;\text{ where }\; u_t = \varepsilon_t \sim i.i.d. \left(0, \sigma^2\right)$$

- Auxiliary Regression

$$\Delta y_t = \theta y_{t-1} + u_t,$$

where $\theta = (\rho - 1)$

- Test statistic under the $H_o : \theta = 0$ is

$$t_\theta = \frac{\hat{\theta}_T}{\hat{\sigma}_{\hat{\theta}_T}} = \frac{\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\varepsilon_t}{\left(\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2\right)^{1/2} s_T},$$

where, given that $\hat{\theta}_T = \hat{\rho}_T - 1$,

$$s_T^2 = \frac{1}{(T-1)}\sum_{t=1}^{T}\left(y_t - \hat{\rho}_T y_{t-1}\right)^2$$

The Dickey-Fuller test

Recall

$$y_t = y_{t-1} + u_t \text{ where } u_t = \varepsilon_t \sim i.i.d.\left(0, \sigma^2\right),$$

Therefore,
(i)

$$s_T^2 \xrightarrow{p} \sigma^2$$

(ii)

$$\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\varepsilon_t \xrightarrow{d} (1/2)\,\sigma^2\left\{[W(1)]^2 - 1\right\}$$

(iii)

$$\frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2 \xrightarrow{d} \sigma^2 \int_0^1 [W(r)]^2\, dr$$

The Dickey-Fuller test

- Hence,
$$t_\theta \xrightarrow{d} \frac{(1/2)\left\{[W(1)]^2 - 1\right\}}{\left(\int_0^1 [W(r)]^2 \, dr\right)^{1/2}}$$
- **Remark**: This distribution is not standard but it can be tabulated. Check DFtest.prg!
- **Remark**: Remember from above that different tables of critical values have to be used in the presence of deterministic components (in the true model and/or the auxiliary regression)

The Dickey-Fuller test

- Assumption $u_t = \varepsilon_t \sim i.i.d.$ typically violated in economic applications
- Relax the iid assumption: $u_t = \psi(L)\,\varepsilon_t$ with $\varepsilon_t \sim i.i.d.$
- If u_t is autocorrelated, then the distribution of the DF test will change
- Several ways of accounting for this fact. Here: **ADF** test

The Augmented Dickey-Fuller test

- ADF test: parametric correction to allow for autocorrelation in the error
- Based on the following "augmented" auxiliary regression

$$\Delta y_t = f(t) + \theta y_{t-1} + \sum_{j=1}^{p-1} \varphi_j \Delta y_{t-j} + \varepsilon_t$$

- Under $H_o : \theta = 0$ the test based on the corresponding t-statistic has the same asymptotic distribution as in the non-autocorrelated case
- As before, the presence of deterministic components will affect the asymptotic distribution

The Augmented Dickey-Fuller test

- ADF auxiliary regression

$$\Delta y_t = f(t) + \theta y_{t-1} + \sum_{j=1}^{p-1} \varphi_j \Delta y_{t-j} + \varepsilon_t$$

- Said and Dickey (1984): if p goes to infinity slowly enough relative to T, $p = T^{1/3}$, then the OLS t-test for $H_o : \theta = 0$ can be carried out using the DF critical values

- How to select the order of the polynomial lags, p, in practice? Information criteria. General to Specific

Nelson and Plosser

- Nelson and Plosser (1982): "**Trends and Random Walks in Macroeconomic Time Series**"
- *"This paper investigates whether macroeconomic time series are better characterized as stationary fluctuations around a deterministic trend or as non-stationary processes that have no tendency to return to a deterministic path."*
- *"Using long historical time series for the U.S. we are unable to reject the hypothesis that these series are non-stationary stochastic processes with no tendency to return to a trend line."*

Nelson and Plosser

Nelson&Plosser data set:

- The U.S. historical time series include measures of output, money, prices, interest rates...
- Annual data. Starting dates varying from 1860 to 1909. All end in 1970
- All series except the bond yield are transformed to natural logs
- Extended version available

Nelson and Plosser

Table 2

Sample autocorrelations of the natural logs of annual data.[a]

Series	Period	Sample autocorrelations: T	r_1	r_2	r_3	r_4	r_5	r_6
Random walk[b]		100	0.95	0.90	0.85	0.81	0.76	0.70
Time aggregated[b] random walk		100	0.96	0.91	0.86	0.82	0.77	0.73
Real GNP	1909–1970	62	0.95	0.90	0.84	0.79	0.74	0.69
Nominal GNP	1909–1970	62	0.95	0.89	0.83	0.77	0.72	0.67
Real per capita GNP	1909–1970	62	0.95	0.88	0.81	0.75	0.70	0.65
Industrial production	1860–1970	111	0.97	0.94	0.90	0.87	0.84	0.81
Employment	1890–1970	81	0.96	0.91	0.86	0.81	0.76	0.71
Unemployment rate	1890–1970	81	0.75	0.47	0.32	0.17	0.04	−0.01
GNP deflator	1889–1970	82	0.96	0.93	0.89	0.84	0.80	0.76
Consumer prices	1860–1970	111	0.96	0.92	0.87	0.84	0.81	0.77
Wages	1900–1970	71	0.96	0.91	0.86	0.82	0.77	0.73
Real wages	1900–1970	71	0.96	0.92	0.88	0.84	0.80	0.75
Money stock	1889–1970	82	0.96	0.92	0.89	0.85	0.81	0.77
Velocity	1869–1970	102	0.96	0.92	0.88	0.85	0.81	0.79
Bond yield	1900–1970	71	0.84	0.72	0.60	0.52	0.46	0.40
Common stock prices	1871–1970	100	0.96	0.90	0.85	0.79	0.75	0.71

[a]The natural logs of all the data are used except for the bond yield. T is the sample size and r_i is the ith order autocorrelation coefficient. The large sample standard error under the null hypothesis of no autocorrelation is $T^{-\frac{1}{2}}$ or roughly 0.11 for series of the length considered here.

[b]Computed by the authors from the approximation due to Wichern (1973).

Nelson and Plosser

Table 3

Sample autocorrelations of the first difference of the natural logs of annual data.[a]

Series	Period	T	r_1	r_2	r_3	r_4	r_5	r_6	$s(r)$
		Sample autocorrelations							
Time aggregated random walk[b]			0.25	0.00	0.00	0.00	0.00	0.00	
Real GNP	1909–1970	62	0.34	0.04	−0.18	−0.23	−0.19	0.01	0.13
Nominal GNP	1909–1970	62	0.44	0.08	−0.12	−0.24	−0.07	0.15	0.13
Real per capita GNP	1909–1970	62	0.33	0.04	−0.17	−0.21	−0.18	0.02	0.13
Industrial production	1860–1970	111	0.03	−0.11	−0.00	−0.11	−0.28	0.05	0.09
Employment	1890–1970	81	0.32	−0.05	−0.08	−0.17	−0.20	0.01	0.11
Unemployment rate	1890–1970	81	0.09	−0.29	0.03	−0.03	−0.19	0.01	0.11
GNP deflator	1889–1970	82	0.43	0.20	0.07	−0.06	0.03	0.02	0.11
Consumer prices	1860–1970	111	0.58	0.16	0.02	−0.00	0.05	0.03	0.09
Wages	1900–1970	71	0.46	0.10	−0.03	−0.09	−0.09	0.08	0.12
Real wages	1900–1970	71	0.19	−0.03	−0.07	−0.11	−0.18	−0.15	0.12
Money stock	1889–1970	82	0.62	0.30	0.13	−0.01	−0.07	−0.04	0.11
Velocity	1869–1970	102	0.11	−0.04	−0.16	−0.15	−0.11	0.11	0.10
Bond yield	1900–1970	71	0.18	0.31	0.15	0.04	0.06	0.05	0.12
Common stock prices	1871–1970	100	0.22	−0.13	−0.08	−0.18	−0.23	0.02	0.10

[a]The first differences of the natural logs of all the data are used except for the bond yield. T is the sample size and r_i is the estimated ith order autocorrelation coefficient. The large sample standard error for r is given by $s(r)$ under the null hypothesis of no autocorrelation.

[b]Theoretical autocorrelations as the number of aggregated observations becomes large; result due to Working (1960).

Nelson and Plosser

Table 4

Sample autocorrelations of the deviations from the time trend.[a]

Series	Period	Sample autocorrelations T	r_1	r_2	r_3	r_4	r_5	r_6
Detrended random		61	0.85	0.71	0.58	0.47	0.36	0.27
walk[b]		101	0.91	0.82	0.74	0.66	0.58	0.51
Real GNP	1909–1970	62	0.87	0.66	0.46	0.26	0.19	0.07
Nominal GNP	1909–1970	62	0.93	0.79	0.65	0.52	0.43	0.05
Real per capita GNP	1909–1970	62	0.87	0.65	0.43	0.24	0.11	0.04
Industrial production	1860–1970	111	0.84	0.67	0.53	0.40	0.30	0.28
Employment	1890–1970	81	0.89	0.71	0.55	0.39	0.25	0.17
Unemployment rate	1890–1970	81	0.75	0.46	0.30	0.15	0.03	−0.01
GNP deflator	1889–1970	82	0.92	0.81	0.67	0.54	0.42	0.30
Consumer prices	1860–1970	111	0.97	0.91	0.84	0.78	0.71	0.63
Wages	1900–1970	71	0.93	0.81	0.67	0.54	0.42	0.31
Real wages	1900–1970	71	0.87	0.69	0.52	0.38	0.26	0.19
Money stock	1889–1970	82	0.95	0.83	0.69	0.53	0.37	0.21
Velocity	1869–1970	102	0.91	0.81	0.72	0.65	0.59	0.56
Bond yield	1900–1970	71	0.85	0.73	0.62	0.55	0.49	0.43
Common stock prices	1871–1970	100	0.90	0.76	0.64	0.53	0.46	0.43

[a]The data are residuals from linear least squares regression of the logs of the series (except the bond yield) on time. See footnote for table 3.

[b]Approximate expected sample autocorrelations based on Nelson and Kang (1981).

Nelson and Plosser

Table 5

Tests for autoregressive unit roots[a]

$$z_t = \hat{\mu} + \hat{\gamma}t + \hat{\rho}_1 z_{t-1} + \hat{\rho}_2(z_{t-1} - z_{t-2}) + \cdots + \hat{\rho}_k(z_{t-k+1} - z_{t-k}) + \hat{u}_t.$$

Series	T	k	$\hat{\mu}$	$t(\hat{\mu})$	$\hat{\gamma}$	$t(\hat{\gamma})$	$\hat{\rho}_1$	$\tau(\hat{\rho}_1)$	$s(\hat{u})$	r_1
Real GNP	62	2	0.819	3.03	0.006	3.03	0.825	−2.99	0.058	−0.02
Nominal GNP	62	2	1.06	2.37	0.006	2.34	0.899	−2.32	0.087	0.03
Real per capita GNP	62	2	1.28	3.05	0.004	3.01	0.818	−3.04	0.059	−0.02
Industrial production	111	6	0.103	4.32	0.007	2.44	0.835	−2.53	0.097	0.03
Employment	81	3	1.42	2.68	0.002	2.54	0.861	−2.66	0.035	0.10
Unemployment rate	81	4	0.513	2.81	−0.000	−0.23	0.706	−3.55*	0.407	0.02
GNP deflator	82	2	0.260	2.55	0.002	2.65	0.915	−2.52	0.046	−0.03
Consumer prices	111	4	0.090	1.76	0.001	2.84	0.986	−1.97	0.042	−0.06
Wages	71	3	0.566	2.30	0.004	2.30	0.910	−2.09	0.060	0.00
Real wages	71	2	0.487	3.10	0.004	3.14	0.831	−3.04	0.034	−0.01
Money stock	82	2	0.133	3.52	0.005	3.03	0.916	−3.08	0.047	0.03
Velocity	102	1	0.052	0.99	−0.000	−0.65	0.941	−1.66	0.067	0.11
Interest rate	71	3	−0.186	−0.95	0.003	1.75	1.03	0.686	0.283	−0.02
Common stock prices	100	3	0.481	2.02	0.003	2.37	0.913	−2.05	0.158	0.20

[a] z_t represents the natural logs of annual data except for the bond yield. $t(\hat{\mu})$ and $t(\hat{\gamma})$ are the ratios of the OLS estimates of μ and γ to th respective standard errors. $\tau(\hat{\rho}_1)$ is the ratio of $\hat{\rho}_1 - 1$ to its standard error. $s(\hat{u})$ is the standard error of the regression and r_1 is the first-ord autocorrelation coefficient of the residuals. The values of $\tau(\hat{\rho}_1)$ denoted by an (*) are smaller than the 0.05 one tail critical value of the distributi of $\tau(\hat{\rho}_1)$ and similarly for $\hat{\rho}_1$. It should also be noted that $t(\hat{\mu})$ and $t(\hat{\gamma})$ are not distributed as normal random variables.