MACROECONOMETRÍA

Tema 2: Series temporales

Segundo Cuatrimestre (curso 2006/07), Depto. de Economía, UC3M

Profesor: Genaro Sucarrat

(Coordinador: Juan J. Dolado)

Conceptos fundamentales:

- Esperanzas condicionales vs. esperanzas incondicionales

- **Definicón: Autocovarianza de retardo** k. La autocovarianza ("autocovariance") de retardo k de un proceso estocástico, denotado $Cov(x_t, x_{t-k})$, se define como

$$E[(x_t - \mu_t)(x_{t-k} - \mu_{t-k})],$$

donde $\mu_t = E(x_t)$ y $\mu_{t-k} = E(x_{t-k})$

Conceptos fundamentales (cont.):

• **Definicón: Estacionariedad débil**. Un proceso estocástico es estacionario débil ("weakly stationary") o estacionario en covarianza ("covariance stationary") si:

$$E(x_t) = \mu \text{ para todo } t$$

$$E[(x_t - \mu)(x_{t-k} - \mu)] = \gamma_k \text{ para todo } t$$

Nota 1: Efectivamente esto significa que ni la media $E(x_t)$ ni la covarianza $E[(x_t - \mu)(x_{t-k} - \mu)]$ dependen del tiempo t.

Nota 2: Estacionaridad débil es una propiedad *no condicional*. Por ejemplo, $E(x)$ es la media no condicional y $E(x|y)$ es la media condicional en y, y $Cov(x, y)$ es la covarianza no condicional y $Cov(x, y|z)$ es la covarianza condicional en z.

Nota 3: Cuando $k = 0$ entonces $E[(x_t - \mu)(x_{t-k} - \mu)]$ es igual a la varianza $E(x_t - \mu)^2 = Var(x_t)$

Conceptos fundamentales (cont.):

- **Definición: Ruido blanco**. Se dice que el proceso $\{\epsilon_t\}$ es ruido blanco ("white noise") si:

$$\begin{array}{rrcrcl} & & & E(\epsilon_t) & = & 0 \\ & Var(\epsilon_t) & = & E(\epsilon_t^2) & = & \sigma^2 \\ \text{Para todo } i \neq j: & Cov(\epsilon_i \epsilon_j) & = & E(\epsilon_i \epsilon_j) & = & 0 \end{array}$$

- Notación: $\epsilon_t \sim WN$

- Ruido blanco Gaussiano: Para todo t, $\epsilon_t \sim N(0, \sigma^2)$

Modelos ARMA:

• **Definición: Modelo ARMA**. Un modelo autoregresivo-media móvil ("autoregressive moving average"—ARMA) de una serie estacionaria débil $\{y_t\}$ tiene la forma:

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=0}^{q} \theta_j \epsilon_{t-j},$$

donde el proceso $\{\epsilon_t\}$ es *ruido blanco*.

• Este modelo se denota como ARMA(p, q), y normalmente se normaliza θ_0 a 1

• *Nota*: si hay que diferenciar la serie para que sea estacionaria entonces a veces se usa la terminología ARIMA(p, d, q), donde d es el orden de integración

Modelos ARMA (cont.):

- Con $p = 0$ obtendremos el modelo MA(q):

$$y_t = \phi_0 + \sum_{j=0}^{q} \theta_j \epsilon_{t-j} = \phi_0 + (1 + \theta_1 L + \cdots + \theta_q L^q)\epsilon_t$$

donde $\{\epsilon_t\}$ es ruido blanco y donde $\theta_0 = 1$

$\Rightarrow \mu = E(y_t) = \phi_0$

$\Rightarrow \gamma_0 = Var(y_t) = (1 + \theta_1^2 + \cdots + \theta_q^2)\sigma^2$

$\Rightarrow \gamma_k = Cov(y_t, y_{t-k}) =$

$$\begin{cases} (\theta_k + \theta_{k+1}\theta_1 + \cdots + \theta_q \theta_{q-1})\sigma^2, & k = 1, \ldots, q \\ 0 \text{ para } k > q \end{cases}$$

- ¿Es el modelo MA(q) estacionario? Si

Modelos ARMA (cont.):

• Con $p = 0$ y $q \to \infty$, obtendremos el modelo MA(∞):

$$y_t = \phi_0 + \sum_{j=0}^{\infty} \theta_j \epsilon_{t-j} = \phi_0 + (1 + \theta_1 L + \theta_2 L^2 + \cdots)\epsilon_t$$

donde $\{\epsilon_t\}$ es ruido blanco y donde $\theta_0 = 1$

• ¿Cómo podemos saber si MA(∞) es un proceso estacionario? Cualquiera de las condiciones siguientes es suficiente:

$$a) \quad \sum_{j=0}^{\infty} \theta_j^2 < \infty$$

$$\Uparrow$$

$$b) \quad \sum_{j=0}^{\infty} |\theta_j| < \infty$$

Modelos ARMA (cont.):

- Esperanzas incondicionales del MA(∞):

$$\Rightarrow \mu = \phi_0$$

$$\Rightarrow \gamma_0 = \sigma^2(\theta_0^2 + \theta_1^2 + \theta_2^2 + \cdots)$$

$$\Rightarrow \gamma_k = \sigma^2(\theta_k\theta_0 + \theta_{k+1}\theta_1 + \theta_{k+2}\theta_2 + \cdots)$$

- Implicación: todos los modelos MA(q) con $q < \infty$ son estacionarios

Modelos ARMA (cont.):

- Con $q = 0$ obtendremos el modelo AR(p):

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t = \phi_0 + (\phi_1 L + \cdots + \phi_p L^p) y_t + \epsilon_t$$

- Si todas las raíces características del polinomio $(1 - \phi_1 L - \cdots - \phi_p L^p)$ están fuera del círculo unidad, entonces:

$\Rightarrow$ El modelo AR(p) se puede escribir/representar como un modelo MA(∞):

$$y_t = \frac{\phi_0}{1-\phi_1-\cdots-\phi_p} + \frac{1}{1-\phi_1 L-\cdots-\phi_p L^p}\epsilon_t$$

$\Rightarrow$ El modelo AR(p) es estable y estacionario

Modelos ARMA (cont.):

- Entonces, tenemos que:

$$\mu = \frac{\phi_0}{1-\phi_1-\cdots-\phi_p}$$

$$\gamma_0 = \phi_1\gamma_1 + \cdots + \phi_p\gamma_p + \sigma^2$$

$$\gamma_k = \phi_k\gamma_{k-1} + \cdots + \phi_p\gamma_{k-p}$$

Modelos ARMA (cont.):

- Más generalmente, un modelo ARMA(p, q):

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=0}^{q} \theta_j \epsilon_{t-j}$$

es estable/estacionario si todas las raíces características del polinomio $(1 - \phi_1 L - \cdots - \phi_p L^p)$ están fuera del círculo unidad

- Además, un modelo ARMA(p, q) estable se puede expresar como un modelo MA(∞):

$$\begin{aligned} y_t &= \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p} + \frac{1 + \theta_1 L + \cdots + \theta_q L^q}{1 - \phi_1 L - \cdots - \phi_p L^p} \epsilon_t \\ &= \psi_0 + \psi(L) \epsilon_t \end{aligned}$$

donde $\psi_0 = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p}$ y $\psi(L) = \frac{1 + \theta_1 L + \cdots + \theta_q L^q}{1 - \phi_1 L - \cdots - \phi_p L^p}$

Modelos ARMA (cont.):

• Si un modelo AR(p) es estable, entonces podemos expresarlo como un MA(∞)

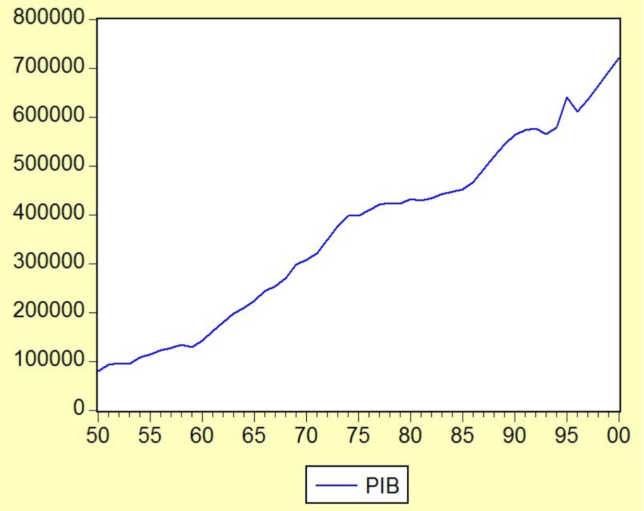
• Si un modelo MA(q) es *invertible*, entonces podemos expresarlo como un AR(∞)

• **Definición: Invertibilidad de MA(q)**. Un modelo MA(q) se puede expresar como $y_t - \phi_0 = (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)\epsilon_t$. Si el MA($q$) se puede expresar como un modelo AR(∞) utilizando la inversa de $(1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)$, entonces se dice que MA(q) es invertible.

• Condición suficiente para la invertibilidad: Que todas las raíces del polinomio $(1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q) = 0$ estén fuera del círculo unidad

Tendencias:

- Consideramos la serie siguiente (PIB de España, 1950-2000):



Tendencias (cont.):

• Se dice que esta serie tiene una *tendencia* ("trend"), porque crece durante toda la muestra

• Problema de regresión/correlación espúrea: regresiones/ correlaciones sugieren una relación entre variables con tendencias aunque no hay ninguna relación

• Dos soluciones comunes a este problema son:

→ 1) tratar la serie como si tuviera una tendencia deterministica ("deterministic trend", "trend stationarity")

→ 2) tratar la serie como si tuviera una tendencia estocástica ("stochastic trend", "difference stationarity")

Tendencias (cont.):

• Dicho de otra manera, dos representaciones comunes de series económicas con tendencias son:

$\rightarrow$ 1) Modelos con tendencia deterministica: $y_t = \phi_0 + \delta \cdot t + e_t$

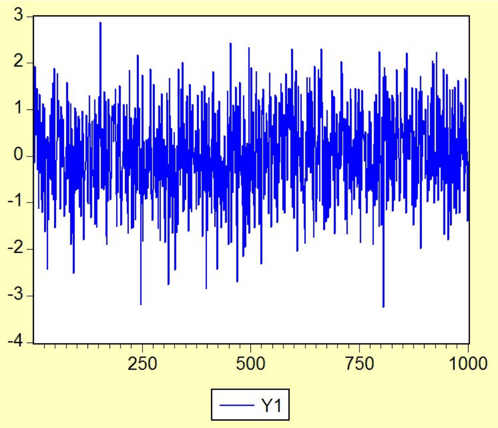
$\rightarrow$ 2) Modelos con tendencia estocástica: $y_t = \phi_0 + \phi_1 y_{t-1} + e_t$

donde t es el tiempo, e_t es una componente estacionaria, y donde $|\phi_1| \geq 1$

• *Nota*: e_t puede tener una estructura ARMA(p, q), por ejemplo, $e_t = \epsilon_t + \theta_1 \epsilon_{t-1}$ (MA(1))

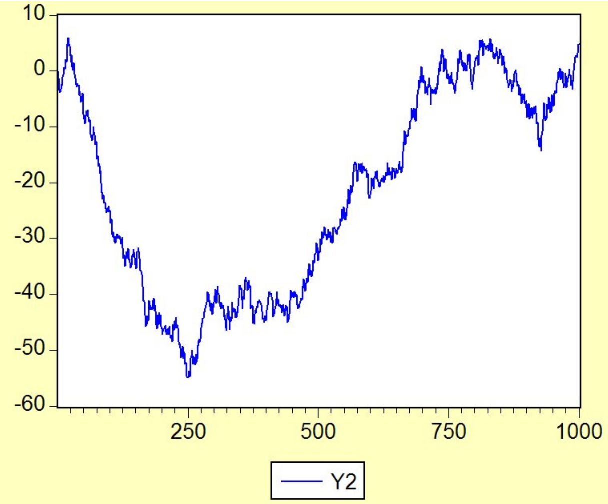
Tendencias (cont.):

- ¿Cómo reconocemos las series con tendencia?
- Por ejemplo, ¿tiene esta serie una tendencia?:



Tendencias (cont.):

- ¿Y esta?:



Tendencias (cont.):

• "Trend stationarity" (TS) denota la propiedad de que la serie $\{y_t - \textit{tendencia deterministica}\}$ es estacionaria (recuerda: $\{e_t\}$ es estacionaria e igual a $\{y_t - \textit{tendencia deterministica}\}$)

• También se dice que $\bar{y}_t = y_t - \textit{tendencia deterministica}$ es "de-trended"

• "Difference stationarity" denota la propiedad de que la serie $\Delta^d y_t$ es estacionaria

• *Nota*: se pueden combinar los dos propiedades. Por ejemplo, puede ser que $\{\Delta y_t\}$ se caracterice por una tendencia determinística

• ¿Es el problema de no-estacionariedad grave? Si, suele dar lugar a *regresiones espúreas*

Tendencias (cont.):

- Intuitivamente, una regresión es espúrea si las pruebas estadísticas sugieren que hay una relación entre las variables cuando no la hay

- Supongamos que estamos considerando estimar la regresión

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t \tag{1}$$

donde $\{y_t\}$ y $\{x_t\}$ son series con tendencias

- ¿Cuál es la probabilidad de una regresión espúrea?

  $\rightarrow$ Muy alta!

Tendencias (cont.):

• *Ejemplo*. Paseos aleatorios independientes (series que "casi" tienen tendencia):

$\longrightarrow y_t = y_{t-1} + \epsilon_{y,t}, \quad \epsilon_{y,t} \sim WN(0, \sigma_y^2)$

$\longrightarrow x_t = x_{t-1} + \epsilon_{x,t}, \quad \epsilon_{x,t} \sim WN(0, \sigma_x^2)$

$\longrightarrow$ $\epsilon_{y,t}$ y $\epsilon_{x,t}$ independientes

$\longrightarrow$ y_t y x_t independientes

• ¿Cuál es la relación empírica entre y_t y x_t? Granger y Newbold (1974) rechazaban la hipótesis de $\beta_1 = 0$ (con un nivel de significación del 5%) en (1) para el 75% de las muestras simuladas

Tendencias (cont.):

• Terminología: Una serie $\{y_t\}$ es *integrada* ("integrated") de orden d si $\Delta^d y_t$ es una serie estacionaria. También se dice que d es el *orden de integración* ("order of integration")

• Notación: $y_t \sim I(d)$ y ARIMA(p, d, q), donde $d \geq 0$

• Si d no es un número entero, por ejemplo, si $d = 0.6$ o si $d = 2.6$, diremos que el orden d es fraccional ("fractional integration")

• *Nota:* En este curso no vamos a ver integración fraccional, siempre vamos a tratar las series como integradas de orden entero, es decir, de orden $0, 1, 2, \ldots$

Tendencias (cont.):

- Ejemplos:

  $\longrightarrow$ Si y_t no es estacionaria, pero Δy_t lo es, entonces denotamos $y_t \sim I(1)$

  $\longrightarrow$ Si ni y_t, ni Δy_t son estacionarias, pero $\Delta^2 y_t$ lo es, entonces denotamos $y_t \sim I(2)$

  $\vdots$

  etc.

  $\longrightarrow$ Si y_t ya es estacionaria, entonces denotamos $y_t \sim I(0)$

Tendencias (cont.):

• Cómo encontramos el orden de integración adecuado?

• Para apoyar o rechazar una hipótesis de orden $d = 1$, por ejemplo, podemos utilizar varios tipos de información:

(a) Inspección visual de gráficos

(b) Propiedades y tests estadísticos

(c) Sentido común

(d) Teoría

• *Nota*: Hay casos en los que los investigadores no están de acuerdo. Por ejemplo, hay casos en que no están de acuerdo si $p_t \sim I(1)$ o si $p_t \sim I(2)$, donde p_t denota un índice de precios en logaritmos

Prueba de Dickey-Fuller:

- Consideramos el modelo

$$y_t = \phi_1 y_{t-1} + \epsilon_t, \qquad \epsilon_t \sim WN(0, \sigma^2) \tag{2}$$

- $|\phi_1| < 1$:

  $\rightarrow \{y_t\}$ estacionario

  $\rightarrow$ prueba de la t es válida

- $\phi_1 = 1$:

  $\rightarrow \{y_t\}$ no es estacionario

  $\rightarrow \{y_t\}$ prueba de la t no es válida

Prueba de Dickey-Fuller (cont.):

• Dickey y Fuller (1979) mostraron que, aunque la prueba de la t no es válida en (2) cuando $\phi_1 = 1$, la prueba de la t es válida para la especificación

$$\Delta y_t = \phi_0 + \pi y_{t-1} + \epsilon_t \tag{3}$$

cuando la hipótesis nula es $\phi_1 = 1$ y donde $\pi = \phi_1 - 1$

• Esta propiedad la utilizaron Dickey y Fuller (1979) para desarrollar una prueba de estacionalidad

• La prueba de "Dickey-Fuller" es una prueba de raíz unitaria ("unit root"), y para determinar los valores críticos para la hípotesis nula de $\phi_1 = 1$ se necesita el método de Monte Carlo (simulación)

Prueba de Dickey-Fuller (cont.):

- La prueba de Dickey-Fuller:

  → Estima $\Delta y_t = \phi_0 + \pi y_{t-1} + \epsilon_t$ con MCO

  → H_0: $\pi = 0$

  → H_1: $\pi < 0$

- Compara el estadístico con el valor de la tabla de Dickey y Fuller (1979) o, mejor, con la tabla de MacKinnon (1996)

- *Nota*: EViews nos da los valores críticos de MacKinnon (1996) automáticamente

Prueba de Dickey-Fuller (cont.):

• Normalmente, las variables requieren una representación AR(p) y no sólo AR(1)

• La prueba de Dickey-Fuller aumentada:

→ Con MCO, estima

$$\Delta y_t = \phi_0 + \pi_0 y_{t-1} + \pi_1 \Delta y_{t-1} + \cdots + \pi_p \Delta y_{t-p} + \epsilon_t \qquad (4)$$

→ H_0: $\pi_0 = 0$

→ H_1: $\pi_0 < 0$

• Compara el estadístico con el valor de la tabla de Dickey y Fuller (1979) o, mejor, con la tabla de MacKinnon (1996)

• *Nota*: EViews nos da los valores críticos de MacKinnon (1996) automáticamente

Prueba de Dickey-Fuller (cont.):

• Problema práctico: elegir p

• Soluciones posibles:

$\rightarrow$ Incluir retardos de Δy_t hasta que el error sea ruido blanco

$\rightarrow$ Incluir retardos de Δy_t significativos

$\rightarrow$ Utilizar AIC, SBC, etc. ("information criteria"), para elegir p

• Sin embargo, la prueba de Dickey-Fuller es frágil $\rightarrow$ siempre hay que utilizar información adicional (inspección visual de la serie, sentido común, historia, teoría) antes de aceptar o rechazar lo que suguiere la prueba de Dickey-Fuller

Prueba de Dickey-Fuller (cont.):

• Tendencia lineal vs. tendencia estocástica: la prueba de Dickey-Fuller nos permite incluir un término de tendencia lineal $\delta \cdot t$

• Es decir, contrastar $\pi_0 = 1$ con la especificación

$$\Delta y_t = \phi_0 + \pi_0 y_{t-1} + \delta t + \pi_1 \Delta y_{t-1} + \cdots + \pi_p \Delta y_{t-p} + \epsilon_t \quad (5)$$

• Problema: si hay variables innecesarias en (5), es más difícil rechazar la hipótesis nula de $\pi_0 = 0$ (raíz unitaria) cuando en realidad $\pi_0 < 0$

• Una solución: hacer tres pruebas

$\rightarrow$ 1. Sin la constante ϕ_0 y sin el término δt

$\rightarrow$ 2. Con la constante ϕ_0 pero sin el término δt

$\rightarrow$ 3. Con la constante ϕ_0 y el término δt

Descomposiciones:

• La descomposición de Beveridge y Nelson (1981) empieza con un modelo ARIMA$(p,1,q)$ de una serie $\{y_t\}$

• Es decir, se supone que $y_t \sim I(1)$, y nosotros vamos a estudiar solamente el caso $p = 0$

• En este caso un modelo ARIMA$(0,1,q)$ de $\{y_t\}$ es igual a

$$\Delta y_t = \phi_0 + e_t, \tag{6}$$

donde $e_t =$MA(q), es decir, $e_t = \sum_{j=0}^{q} \theta_j \epsilon_{t-j}$ con $\theta_0 = 1$

• El modelo (6) también se puede representar como un paseo aleatorio

$$y_t = \phi_0 + y_{t-1} + e_t,$$

es decir, como un ARMA(1,q) no-estacionario

Descomposiciones (cont.):

• Para obtener la descompocición Beveridge-Nelson, primero hay que calcular la predicción condicional $K > 0$ periodos hacia al futuro

• Tenemos que

$$y_{t+K} = \phi_0 K + y_t + \sum_{k=1}^{K} e_{t+k}$$

$$E_t(y_{t+K}) = \phi_0 K + y_t + E_t(\sum_{k=1}^{K} e_{t+k})$$

donde $E_t(\cdot)$ es la esperanza condicional $E(\cdot|y_t, \epsilon_t, \epsilon_{t-1}, \ldots)$

• El término $\phi_0 K$ es la parte determinística, y la tendencia estocástica se define como

$$\lim_{K\to\infty}[E_t(y_{t+K}) - \phi_0 K]$$

Descomposiciones (cont.):

- *Ejemplo*. Un modelo ARIMA(0,1,1) se puede escribir como:

$$y_t = \phi_0 + y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1}$$

y nos da

$$\begin{aligned} y_{t+K} &= \phi_0 K + y_t + \sum_{k=1}^{K} \epsilon_{t+k} + \theta_1 \sum_{k=1}^{K} \epsilon_{t+k-1} \\ E_t(y_{t+K}) &= \phi_0 K + y_t + \theta_1 \epsilon_t \\ \lim_{K \to \infty} [E_t(y_{t+K}) - \phi_0 K] &= y_t + \theta_1 \epsilon_t \end{aligned}$$

- Si $y_t = 10$, $\theta_1 = 0,1$ y $\epsilon_t = 1$, entonces

$$\lim_{K \to \infty} E_t(y_{t+K}) - \phi_0 K = 10,1$$

Descomposiciones (cont.):

• La descomposición de Hodrick y Prescott (1997) consiste en igualar los valores $\{y_t\}$ de una seria con tendencia a dos componentes μ_t y e_t:

$$y_t = \mu_t + e_t$$

• Intuitivamente μ_t es una tendencia, y $e_t = y_t - \mu_t$ es una componente estacionaria con $E(e_t) = 0$ para todo t

• El método de Hodrick y Prescott consiste en estimar una serie $\mu_1, \ldots, \mu_T$ tal que las diferencias $\mu_{t+1} - \mu_t$ oscilan "suavemente"

Descomposiciones (cont.):

- Más precisamente, el método de Hodrick-Prescott consiste en:

  1. Elegir un valor λ del “coste” de incorporar fluctuaciones en la serie $\mu_1, \ldots, \mu_T$
  2. Minimizar

$$\frac{1}{T}\sum_{t=1}^{T}(y_t - \mu_t)^2 + \frac{\lambda}{T}\sum_{t=2}^{T-1}[(\mu_{t+1} - \mu_t) - (\mu_t - \mu_{t-1})]^2 \qquad (7)$$

- $\lambda = 0$: no hay coste de incorporar fluctuaciones en la serie $\mu_1, \ldots, \mu_T$, y el mínimo de (7) se obtiene cuando $y_t = \mu_t$
- $\lambda \to \infty$: coste alto de incorporar fluctuaciones en la serie $\mu_1, \ldots, \mu_T$, y el mínimo de (7) se obtiene cuando $\mu_{t+1} - \mu_t = \mu_t - \mu_{t-1}$. Es decir, cuando la diferencia $\mu_t - \mu_{t-1}$ es constante. Cuanto más grande es λ, más suave es el cambio $\mu_t - \mu_{t-1}$

Referencias:

Beveridge, S. and C. Nelson (1981). A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycles. *Journal of Monetary Economics 7*, 151–174.

Dickey, D. and W. Fuller (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association 74*, 426–431.

Granger, C. W. and P. Newbold (1974). Spurious Regressions in Econometrics. *Journal of Econometrics 2*, 111–120.

Hodrick, R. J. and E. C. Prescott (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking 29*, 1–16.

MacKinnon, J. G. (1996). Numerical Distribution Functions for Unit Root and Cointegration Tests. *Journal of Applied Econometrics 11*, 601–618.