

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CC. ECONÓMICAS Y EMPRESARIALES
Departamento de Fundamentos de Análisis Económico I



**CONTRASTES DE COINTEGRACIÓN Y RAÍCES
UNITARIAS Y COINTEGRACIÓN CON DATOS
ATÍPICOS Y CAMBIOS ESTRUCTURALES:
SOLUCIONES EN MUESTRAS FINITAS**

**MEMORIA PRESENTADA PARA OPTAR AL GRADO DE
DOCTOR POR Miguel Ángel Arranz Cuesta**

Bajo la dirección del Doctor:
Álvaro Escribano Sáez

Madrid, 2002

ISBN: 84-669-2243-1

**CONTRASTES DE RAÍCES UNITARIAS Y COINTEGRACIÓN
CON DATOS ATÍPICOS Y CAMBIOS ESTRUCTURALES:
SOLUCIONES EN MUESTRAS FINITAS**

MIGUEL A. ARRANZ CUESTA

Tesis para la obtención del Grado de Doctor en Economía
Departamento de Fundamentos de Análisis Económico I
Universidad Complutense de Madrid
Septiembre 2001

Director: Prof. Dr. Alvaro Escribano Sáez

**Contrastes de raíces unitarias y cointegración
con datos atípicos y cambios estructurales:
Soluciones en muestras finitas**

Miguel A. Arranz

Índice General

Introducción	vii
Síntesis	vii
Resúmenes de capítulos individuales	ix
Agenda de investigación	xi
Parte 1. Raíces unitarias	1
Capítulo 1. Efectos de Filtros Lineales y No lineales en Contrastes de Raíces Unitarias en Series con Atípicos Aditivos	3
1. Introducción	3
2. Raíces Unitarias y Atípicos Aditivos	5
3. Filtros y Extracción de Señales	7
3.1. El filtro Hodrick-Prescott.	8
3.2. El filtro de Baxter y King	9
3.3. Filtros no lineales	11
4. Filtrado de Procesos con Raíz Unitaria	12
5. Evidencia Empírica	15
6. El Enfoque Bootstrap	26
7. Aplicación Empírica	28
8. Conclusiones	35
9. Demostración del Teorema 1.4.1	36
Parte 2. Cointegración	39
Capítulo 2. Contrastes de Cointegración con Cambios Estructurales: Un Modelo de Corrección del Error Extendido Robusto	41
1. Introducción	41
2. Modelos de Corrección del Error con y sin Co-ruptura simultánea	43
2.1. Modelos de Corrección del Error con Co-ruptura Simultánea	46
2.2. Modelos de Corrección del Error sin Co-ruptura Simultánea.	51

3. Experimento de simulación de Monte Carlo	53
3.1. Experimento de Monte Carlo: MCE con Co-ruptura Simultánea	56
3.2. Experimento de Simulación de Monte Carlo: Contraste MCE sin Co-ruptura simultánea.	60
4. Modelos de Corrección del Error Extendidos	63
5. Conclusiones	70
6. Co-rupturas en niveles y en diferencias.	72
 Capítulo 3. Contrastes Bootstrap de Cointegración con Cambios Estructurales: Un Contraste MCE Extendido Robusto.	 75
1. Introduction	75
2. Modelos de Corrección del Error con Co-rupturas en Niveles y en Diferencias	76
3. Bootstrap de Modelos de Corrección del Error	79
4. Experimento de Monte Carlo I. No saltos.	81
5. Experimento de Simulación de Monte Carlo II: Co-rupturas	86
5.1. MCE con Co-ruptura simultánea	86
5.2. MCE con Co-rupturas Parciales	89
6. Conclusiones	96
 Capítulo 4. Contrastes de Cointegración MCE Robustos a Atípicos Basados en el Componente Tendencial	 99
1. Introducción	99
2. Modelos de Corrección del Error con y sin Co-ruptura Simultánea	101
2.1. Modelos de Corrección del Error con Co-ruptura Simultánea	103
2.2. Modelos de Corrección del Error sin Co-ruptura Simultánea.	104
3. Filtros y Extracción de Señales	105
3.1. El filtro de Baxter y King (BK)	105
3.2. Filtro de Hodrick y Prescott (HP)	107
3.3. Filtro de la mediana	108
4. Contraste MCE basado en Componentes Tendenciales	108
5. Experimento de Simulación de Monte Carlo	109
5.1. El modelo	110
5.2. Contraste MCE basado en componentes tendenciales: No atípicos.	111
5.3. Contraste MCE con atípicos aditivos	112

5.4. Robustez de los valores críticos a la presencia de atípicos	120
6. El Enfoque Bootstrap	122
7. Ejemplo Empírico	130
8. Conclusiones	134
Parte 3. Estimación de relaciones de largo plazo	137
Capítulo 5. Un Enfoque No-paramétrico de la Cointegración de Pares de Variables $I(d)$	139
1. Introducción	139
2. Definiciones de memoria en series temporales	142
3. Cointegración lineal en el dominio temporal	143
4. Experimento de simulación de Monte Carlo	148
5. Cointegración lineal en el dominio de frecuencias	152
6. Relación entre cointegración en el dominio temporal y en el dominio de frecuencias	155
7. Conclusiones y extensiones	157
8. Demostraciones	158
Demostración del Teorema 5.3.1	158
Demostración del Corolario 5.3.1	159
Demostración del Teorema 5.4.1	160
Demostración del Teorema 5.6.1	161
Demostración del Corolario 5.6.1	161
9. Tablas	162
Bibliografía	167

Introducción

Síntesis

Uno de los principales problemas a los que tenemos que enfrentarnos al hacer análisis empírico con series temporales económicas es que los datos incorporan anomalías tales como observaciones atípicas y cambios estructurales. El efecto de tales anomalías es bien conocido en datos de sección cruzada, pero no tanto en series temporales, a pesar de que en este último caso los efectos son incluso más dramáticos.

El propósito de esta Tesis Doctoral es el estudio y análisis en muestras pequeñas del comportamiento de ciertos contrastes de raíces unitarias y cointegración cuando las series temporales están contaminadas por observaciones atípicas de tipo aditivo o cambios estructurales. Los efectos asintóticos de tales anomalías ha sido analizado extensamente en la literatura, siendo la conclusión fundamental que su presencia hará que los procedimientos de inferencia estadística habituales darán lugar a resultados erróneos, incluso asintóticamente. En esta Tesis Doctoral vamos a seguir un enfoque aplicado a muestras pequeñas o finitas, basado en distintas alternativas de modelización.

La Tesis Doctoral se divide en 3 partes, que se resumen como sigue:

Parte I: Raíces unitarias y atípicos aditivos.

- La primera contribución de esta Tesis es la aplicación de los contrastes de raíces unitarias a los datos filtrados en lugar de a los datos originales. El análisis se centra en algunos filtros lineales y no-lineales: el filtro de Hodrick–Prescott, el filtro de Baxter–King y el filtro de la mediana. Se comprueba que los valores críticos de los contrastes de raíces unitarias aplicados a los componentes tendenciales son muy robustos a la presencia de observaciones atípicas aditivas y presentan potencia alta incluso en muestras de tamaño moderado.
- Con el fin de aumentar la potencia de los contrastes basados en el componente tendencias, se aplican técnicas de sieve bootstrap, y se

comprueba que su potencia es mayor que la de test propuesto por Ng y Perron (1998) en presencia de atípicos aditivos.

- Se realiza una ilustración de esta metodología aplicándola a los datos previamente analizados por Perron y Vogelsang (1992), Franses y Haldrup (1994) y Vogelsang (1999)

Parte II: Contrastes de cointegración

- El concepto de co-ruptura, introducido por Clements y Hendry (1999), se extiende para incorporar los casos de co-rupturas parciales. El estudio se centra en los efectos de co-rupturas parciales sobre el contraste de cointegración del modelo de corrección del error (MCE). Se demuestra que el uso de un modelos de corrección del error extendido proporciona valores críticos que no son sensibles a parámetros molestos en el caso de co-ruptura parcial. Esto no es cierto en el caso del contraste MCE habitual.
- Los problemas de la dependencia de parámetros molestos del contraste MCE en el caso de co-rupturas parciales se solucionan mediante la aplicación de técnicas bootstrap al contraste de cointegración basado en el MCE extendido, proporcionando una potencia excelente incluso en muestras muy pequeñas.
- En el caso de presencia de atípicos aditivos, la solución pasa por la combinación de la aplicación de los contrastes de cointegración a las series filtradas y de técnicas bootstrap. El análisis demuestra que el tamaño y la potencia del contraste son adecuados en muestras pequeñas, pero que el bootstrap requiere muestras de tamaño algo mayor.

Parte III: Estimación de relaciones de largo plazo.

- La última parte de esta Tesis se centra en el análisis de un nuevo estimador de vectores de cointegración basado en el principio de variables instrumentales, VI , aplicable en los casos de cointegración entera y de cointegración fraccional. Se comprueba que el nuevo estimador VI mejora el comportamiento en muestras pequeñas del estimador MCO habitual.

Resúmenes de capítulos individuales

Capítulo 1: Efectos de la Aplicación de Filtros Lineales y No lineales en los Contrastes de Raíces Unitarias en Series con Atípicos Aditivos.

Los contrastes univariantes de Dickey–Fuller convencionales tienden a encontrar estacionariedad espuria cuando existen observaciones atípicas de tipo aditivo en las series temporales. El método habitual para obtener valores críticos correctos es la inclusión de variables dicotómicas (dummy) en la regresión de Dickey–Fuller. Esto constituye una propiedad interesante desde el punto de vista teórico, pero no es muy atractiva desde el punto de vista empírico, pues se puede obtener cualquier resultado mediante la inclusión de las variables dummy convenientes. En este Capítulo se introduce un nuevo procedimiento que consiste en aplicar el contraste de Dickey–Fuller a los componentes tendenciales de las series temporales en lugar de a las series observadas, así como técnicas bootstrap. Se proporcionan justificaciones tanto a nivel de muestras pequeñas como asintóticas. La puesta en práctica de la nueva metodología se ilustra mediante el estudio de un ejemplo empírico basado en el tipo de cambio real entre las monedas de Estados Unidos y Finlandia.

Capítulo 2: Contrastes de Cointegración con Cambios Estructurales: Un Modelo de Corrección del Error Extendido Robusto.

Es bien sabido que los contrastes de raíces unitarias y de cointegración depende de elementos determinísticos tales como: constantes, tendencias, rupturas, atípicos, tendencias segmentadas, etc., que están presentes en la hipótesis nula y puede que también en la alternativa. Esto representa un serio inconveniente para el análisis empírico, pues se puede influir de manera arbitraria en el resultado de un estudio mediante la selección de los elementos deterministas que se incluyen en la regresión que sirve de soporte al contraste. En este Capítulo se analiza si esos problemas se pueden solventar, o al menos reducir, mediante la aplicación del contraste de cointegración en un modelo de corrección del error extendido. El análisis se realiza basándose en experimentos de simulación de Monte Carlo incluyendo distintos tipos de cambios estructurales en el proceso generador de datos.

Capítulo 3: Contrastes Bootstrap de Cointegración con Cambios Estructurales: Un Contraste MCE Extendido Robusto.

El propósito de este Capítulo es el análisis de los contrastes bootstrap de cointegración MCE con cambios estructurales. Los contrastes MCE habituales dependen de parámetros molestos, lo que es una característica no deseable desde el punto de vista empírico. Este problema se elimina mediante la aplicación del contraste bootstrap MCE, que presenta excelentes propiedades de tamaño y potencia en el caso de que no haya cambios estructurales. En este Capítulo se estudian las propiedades en muestras pequeñas de distintas alternativas bootstrap de contrastes MCE en los casos de distintas situaciones de co-rupturas. Los estadísticos para el contraste MCE se hacen robustos a diferentes situaciones de co-rupturas parciales utilizando modelos de corrección del error extendidos o imponiendo la restricción de factores comunes.

Capítulo 4: Contrastes de Cointegración MCE Robustos a Atípicos basados en los Componentes Tendenciales.

El propósito fundamental de este Capítulo es el análisis del comportamiento del contraste de cointegración MCE cuando las series temporales presentan atípicos aditivos en diferentes situaciones de co-rupturas. Se comprueba que el contraste MCE habitual no es robusto a la presencia de shocks transitorios y se propone un procedimiento basado en métodos de extracción de señales para solucionar este problema. Este nuevo procedimiento proporciona un contraste MCE cuya cola izquierda de la distribución bajo la hipótesis nula es robusta a la presencia de atípicos aditivos en las series. Se estudian los valores críticos y la potencia en muestras pequeñas mediante experimentos de Monte Carlo para diferentes filtros pasa-baja. La metodología propuesta se aplica al tipo de cambio real (basado en el IPC) entre Estados Unidos y Finlandia.

Capítulo 5: Una Aproximación No-paramétrica a la Cointegración para Pares de Variables $I(d)$.

En este Capítulo se proponen algunos estadísticos independientes de modelos (no-paramétricos) de dependencia de las series que son útiles para la caracterización de las propiedades de memoria corta y larga de las series en los dominios temporal y de las frecuencias. Se introducen con la ayuda de estos estadísticos condiciones de memoria conjunta de las series, tales como la cointegración. Se muestra la relación existente entre el concepto habitual de cointegración y el no-paramétrico en los dominios temporal y

de frecuencias. El nuevo concepto de cointegración, basado en el ratio de la covarianzas cruzadas de las series, tiene una interpretación inmediata como estimador de variables instrumentales. Se demuestra que el comportamiento en muestras finitas del estimador de variables instrumentales es mejor que el de mínimos cuadrados ordinarios empleado habitualmente. Finalmente, la caracterización propuesta permite la discriminación entre cointegración entera y fraccional.

Agenda de investigación

Basándonos en los resultados de esta Tesis Doctoral, hay toda una serie de temas que están abiertos a futuras investigaciones:

- Estimación de relaciones de cointegración no lineales basada en métodos paramétricos y no paramétricos de variables instrumentales con aplicaciones financieras.
- Extensión a casos de relaciones de cointegración múltiples, con posibles aplicaciones a modelos de oferta y demanda de dinero.
- Contrastes de cointegración Engle–Granger y MCE basados en el nuevo estimador de variables instrumentales.

Parte 1

Raíces unitarias

Efectos de Filtros Lineales y No lineales en Contrastes de Raíces Unitarias en Series con Atípicos Aditivos

1. Introducción

Los procedimientos de contrastes de raíces unitarias univariantes y multivariantes son sensibles a la presencia de sucesos anómalos tales como cambios estructurales y observaciones atípicas. Desde el punto de vista univariante, los efectos de la presencia de tales efectos se ha estudiado profusamente. Perron (1989) demuestra que los contrastes de raíces unitarias dejan de ser útiles (pues encuentran demasiadas raíces unitarias) si hay un cambio estructural en el proceso generador de datos, como puede ser un cambio de nivel. La razón intuitiva es que la hipótesis de raíz unitaria está íntimamente asociada a shocks con efectos permanentes. Un cambio estructural es básicamente un shock con efecto permanente en la serie (ver Perron y Vogelsang, 1992). Por tanto, si no hemos tomado explícitamente en cuenta dicho shock, los contrastes de raíces unitarias habituales confundirán dicho cambio estructural con una raíz unitaria. Por otra parte, si el cambio estructural acontece al principio de la serie, los contrastes de raíces unitarias habituales darán lugar a un excesivo número de rechazos de la hipótesis de raíz unitaria (ver Leybourne et al., 1998).

Una posible solución a este problema pasa por la utilización de variables dicotómicas tipo dummy y tendencias segmentadas. Por ejemplo, Perron (1989) incluye una lista de regresores deterministas para permitir que la hipótesis alternativa tenga un cambio de tendencia en un momento conocido. El mayor problema que tiene este procedimiento es que, incluso si sabemos cuándo tiene lugar el cambio estructural, los valores críticos dependen del tamaño del mismo, así como del momento en que tiene lugar. Por otra parte, las distribuciones asintóticas de los contrastes de raíces unitarias dependen de si la localización del cambio estructural se conoce exógenamente o no (ver, por ejemplo Christiano, 1992).

Otro tipo de suceso anómalo son las observaciones atípicas aditivas transitorias. Estos son sucesos con un efecto (posiblemente) grande, pero temporal en la serie. En el caso de que estos atípicos aditivos (AO's) sean suficientemente grandes o suficientemente frecuentes, dicho efecto puede dominar el resto de la información contenida en la serie y sesgar la inferencia resultante de los contrastes de raíces unitarias hacia el rechazo de la hipótesis de raíz unitaria, tal y como indican Martin y Yohai (1986), Franses y Haldrup (1994), Lucas (1995*a,b*), Shin et al. (1996) y Vogelsang (1999), entre otros.

En principio, es posible incluir variables dummy para estos shocks temporales de la misma forma que se hace habitualmente con los cambios estructurales y calcular los valores críticos correspondientes, ver Franses y Haldrup (1994), Shin et al. (1996) y Vogelsang (1999). Un aspecto crucial en este enfoque es, no obstante, la detección de las observaciones atípicas. Tsay (1986), Chang et al. (1988), Chen y Liu (1993), Gómez y Maravall (1996) y Kaiser (1998) proporcionan toda una serie de instrumentos que constituyen los procedimientos habituales de detección de atípicos aislados. No obstante, desde un punto de vista práctico, se requiere un alto grado de experiencia para la identificación de los distintos atípicos y la interpretación de los diferentes estadísticos. Alternativamente, se podría intentar evitar la utilización de las variables dummy y considerar en su lugar técnicas de estimación robustas, siguiendo el enfoque de Lucas (1995*a,b*). La principal diferencia de ambos enfoques es que, en tanto que la detección de atípicos es para eliminarlos de la serie, la estimación robusta no los elimina. Ver Maddala y Yin (1996) y Yin y Maddala (1997) para más detalles.

Siguiendo el enfoque robusto, en este Capítulo se explora una ruta diferente para contrastar la presencia de raíces unitarias cuando los datos incluyen observaciones atípicas. Supongamos que la serie de interés, z_t , se subdivide en dos componentes no observados, $z_t = z_t^g + z_t^c$, donde z_t^g denota el componente permanente o componente tendencial, y z_t^c representa el componente transitorio o cíclico. Lo que se propone es aplicar el contraste de raíces unitarias a z_t^g en lugar de a la serie observada. La idea es que, puesto que los AO's son transitorios, pueden incorporarse, en una gran parte, al componente cíclico z_t^c y no a la tendencia. Para estimar el componente tendencial, z_t^g , se sugiere la utilización de filtros de tipo pasa-baja.

La estructura del Capítulo es la siguiente: en la Sección 2 se introduce el modelo de interés y se presentan algunos resultados analíticos acerca de los efectos de la frecuencia y el tamaño de los AO's en el estadístico del contraste de Dickey-Fuller. La Sección 3 presenta algunos resultados básicos de filtrado y extracción de señales.

Los efectos de la aplicación de los distintos filtros en los contrastes de Dickey–Fuller se discuten en la Sección 4. En la Sección 5 se hace un estudio de simulación basado en experimentos de Monte Carlo para ilustrar las implicaciones en muestras finitas de los resultados analíticos. La Sección 7 incluye una aplicación empírica y la Sección 8 presenta las conclusiones finales. La prueba del Teorema 1.4.1 se recoge en la Sección 9.

2. Raíces Unitarias y Atípicos Aditivos

Sea y_t generada por un paseo aleatorio con $y_0 = 0$,

$$y_t = y_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots, T, \quad (2.1)$$

donde ε_t son variables aleatorias idéntica e independientemente distribuidas (i.i.d.) $N(0, \sigma_\varepsilon^2)$, y supongamos que atípicos aditivos de tamaño $\pm\theta$ pueden ocurrir con probabilidad π , de forma que la serie temporal que observamos es

$$x_t = y_t + \theta\delta_t, \quad (2.2)$$

donde δ_t son variables aleatorias i.i.d. que siguen un proceso de Bernoulli tal que $P(\delta_t = 1) = \frac{1}{2}\pi$, $P(\delta_t = -1) = \frac{1}{2}\pi$ y $P(\delta_t = 0) = 1 - \pi$, $\delta_0 = 0$.

Considereremos la regresión de raíz unitaria

$$\Delta y_t = \rho y_{t-1} + \varepsilon_t. \quad (2.3)$$

Como es bien sabido, el contraste de raíz unitaria de Dickey–Fuller toma como hipótesis nula $H_0 : \rho = 0$, raíz unitaria, frente a $H_1 : \rho < 0$, raíz estable. El contraste se efectúa por medio de estimador mínimo cuadrático de ρ ,

$$\hat{\rho} = \left(\sum_{t=1}^T y_{t-1} \Delta y_t \right) / \left(\sum_{t=1}^T y_{t-1}^2 \right), \quad (2.4)$$

y el estadístico t correspondiente a contraste de significatividad de dicho coeficiente,

$$t_{\hat{\rho}} = \hat{\rho} / \widehat{\sigma}_y \left(\sum_{t=1}^T y_{t-1}^2 \right)^{-1/2}, \quad (2.5)$$

donde $\widehat{\sigma}_y^2 = T^{-1} \sum_{t=1}^T (\Delta y_t - \hat{\rho} y_{t-1})^2$ es la varianza residual, $\Delta = (1 - B)$ and B es el operador retardo, definido tal que $B^n y_t = y_{t-n}$ para valores positivos y negativos de n .

Se deduce de la ecuación (2.2) que, en presencia de AO's, la regresión (2.3) pasa a ser

$$\Delta x_t = \rho x_{t-1} + u_t, \quad (2.6)$$

donde $u_t = \varepsilon_t + \theta \Delta \delta_t$ con los estadísticos mínimo cuadráticos correspondientes

$$\hat{\rho}_{AO} = \left(\sum_{t=1}^T x_{t-1} \Delta x_t \right) / \left(\sum_{t=1}^T x_{t-1}^2 \right), \quad (2.7)$$

y

$$t_{\hat{\rho}_{AO}} = \hat{\rho}_{AO} / \hat{\sigma}_x \left(\sum_{t=1}^T x_{t-1}^2 \right)^{-1/2}, \quad (2.8)$$

siendo $\hat{\sigma}_x^2 = T^{-1} \sum_{t=1}^T (\Delta x_t - \hat{\rho}_{AO} x_{t-1})^2$. Por tanto, bajo H_0 , la serie observada es un proceso $I(1)$ con innovaciones $MA(1)$, dado que $E(u_t u_{t-1}) = -\pi \theta^2$, $E(u_t u_{t-j}) = 0$ para $j > 1$. Por consiguiente, u_t satisface las condiciones de mixing generales introducidas por Phillips (1987a), y se puede aplicar un teorema central del límite funcional a las sumas parciales de u_t , de manera que se pueden encontrar las distribuciones asintóticas de (2.7) y (2.8). En particular, se puede probar que, cuando $T \rightarrow \infty$,

$$\begin{aligned} T \hat{\rho}_{AO} &\Rightarrow \left(\int_0^1 W(r) dW(r) \right) / \left(\int_0^1 W^2(r) dr \right) \\ &\quad - (\theta/\sigma_\varepsilon)^2 \pi \left(\int_0^1 W^2(r) dr \right)^{-1} \end{aligned} \quad (2.9)$$

y

$$\begin{aligned} t_{\hat{\rho}_{AO}} &\Rightarrow \left(1 + 2(\theta/\sigma_\varepsilon)^2 \pi \right)^{-1/2} \\ &\quad \times \left(\int_0^1 W(r) dW(r) \right) / \left(\int_0^1 W^2(r) dr \right)^{1/2} \\ &\quad - (\theta/\sigma_\varepsilon)^2 \pi \left(\int_0^1 W^2(r) dr \right)^{-1/2}, \end{aligned} \quad (2.10)$$

donde \Rightarrow indica convergencia en sentido débil y $W(r)$ es un movimiento Browniano standard definido en el intervalo inoitario $r \in [0, 1]$.

Las ecuaciones (2.9) y (2.10) aparecen por vez primera en Franses y Haldrup (1994). En estas ecuaciones se observa que si $\pi > 0$, ρ se sigue estimando de forma superconsistente, pero la distribución asintótica de los estadísticos de Dickey–Fuller se desplaza a la izquierda, lo que origina que el tamaño del contraste sea superior al nominal y, por tanto, a que se rechace en exceso la hipótesis de raíz unitaria en favor

de la alternativa de estacionariedad. Obsérvese que un AO positivo y un AO negativo pero de la misma magnitud tienen el mismo efecto en la distribución límite. Nótese también que las distribuciones son las mismas para las diferentes combinaciones de θ , π y σ_ε^2 que dan lugar al mismo valor de $(\theta/\sigma_\varepsilon)^2 \pi$. En otras palabras, el efecto de shocks grandes pero poco probables es el mismo que el de shocks pequeños pero muy probables. Los contrastes de Dickey–Fuller no pueden distinguir entre shocks grandes infrecuentes y shocks pequeños frecuentes. Por otra parte, los resultados de Franses y Haldrup se pueden extender a modelos que tengan también componentes deterministas, ver Yin y Maddala (1997).

3. Filtros y Extracción de Señales

Las técnicas de filtrado se han usado durante mucho tiempo para estimar los estados de los sistemas estocásticos dinámicos o para extraer información de observaciones con ruido. Estas técnicas se han adoptado en muchos campos de la Economía, especialmente en Macroeconomía. El propósito primordial en este último caso es separar de las series temporales macroeconómicas el componente de ciclo de negocios de la tendencia secular y de los más volátiles componentes estacional e irregular.

Por simplicidad, supongamos que estamos en separar una serie temporal observada, z_t , en dos componentes no observados,

$$z_t = z_t^g + z_t^c, \quad (3.1)$$

donde z_t^g es el componente crecimiento o tendencia y z_t^c es el componente cíclico. Obsérvese que la ecuación (3.1) puede reescribirse como

$$z_t = z_t^g + (z_t - z_t^g), \quad (3.2)$$

o

$$z_t = (z_t - z_t^c) + z_t^c. \quad (3.3)$$

La ecuación (3.2) requiere la definición del componente tendencial, en tanto que la ecuación (3.3) precisa la definición del componente cíclico. Burns y Mitchell (1946) constituye la referencia seminal para la medición de los ciclos de negocios. En la misma se propone una descomposición hoy en desuso. En su lugar, la Macroeconomía empírica moderna utiliza toda una variedad de técnicas de filtrado y alisado de series temporales para realizar las descomposiciones tendencia–ciclo. Ejemplos de tales

técnicas son la aplicación de medias móviles bi-direccionales, extracción de tendencias polinómicas, lineales o cuadráticas, y la aplicación del filtro de Hodrick y Prescott (1997).

3.1. El filtro Hodrick-Prescott. El filtro de Hodrick-Prescott (HP), se utiliza extensamente en modelos de ciclos reales de negocios para extraer la tendencia de las series y concentrarse en los hechos estilizados de una economía a lo largo del ciclo de negocios. Ver, por ejemplo, Harvey y Jaeger (1993), King y Rebelo (1993), Cogley y Nason (1995), Guay y St-Amant (1997), Canova (1998), Ehglen (1998), Baxter y King (1999), y Kaiser y Maravall (1999).

La base del filtro es la siguiente: partiendo de la ecuación (3.2), se define el componente permanente de la serie, z_t^g , como la solución del problema de optimización

$$\min_{z_t^g} \sum_{t=1}^T \left[(z_t - z_t^g)^2 + \lambda (\Delta^2 z_{t+1}^g)^2 \right]. \quad (3.4)$$

El primer término de (3.4) puede interpretarse como una medida de la bondad del ajuste del componente tendencial a la serie observada, en tanto que el segundo término impone una penalización con el fin de obtener un componente tendencial suave. King y Rebelo (1993) demuestra que la versión de muestra infinita del filtro HP define el componente cíclico de z_t como $z_t^c = H(B)z_t$, donde

$$H(B) = \frac{\lambda(1-B)^2(1-B^{-1})^2}{1 + \lambda(1-B)^2(1-B^{-1})^2}, \quad (3.5)$$

y es un filtro lineal óptimo en el sentido de error cuadrático medio mínimo.

Por otra parte, haciendo $F(B) = 1 - H(B)$, de (3.5) se tiene que el filtro de la tendencia HP tiene la forma

$$F(B) = \left[1 + \lambda(1-B^{-1})^2(1-B^2) \right]^{-1} \times \left[\lambda B^2 - 4\lambda B + (1+6\lambda) - 4\lambda B^{-1} + \lambda B^{-2} \right]^{-1}, \quad (3.6)$$

que resulta ser un filtro de tipo pasa-baja simétrico que necesita aproximarse en la práctica. Con este fin, una posibilidad puede ser truncar los pesos en un retardo fijo n . Sin embargo, en la práctica se utiliza el siguiente procedimiento alternativo: En primer lugar, se coloca los datos en un vector columna X . Segundo, definir una matriz Γ que relaciona el correspondiente vector de componentes tendenciales, X^g , a los datos: $X = \Gamma X^g$. La matriz Γ viene implícitamente definida por la ecuación (3.6), esto es,

$$x_t = \lambda x_{t+2}^g - 4\lambda x_{t+1}^g + (1+6\lambda)x_t^g - 4\lambda x_{t-1}^g + \lambda x_{t-2}^g, \quad (3.7)$$

con algunas modificaciones alrededor del inicio y el final de la serie:

$$x_1 = (1 + \lambda)x_1^g - 2\lambda x_2^g + (1 + \lambda)x_3^g,$$

y

$$x_2 = -2\lambda x_1^g + (1 + 5\lambda)x_2^g - 4\lambda x_3^g + \lambda x_4^g.$$

Se realizan transformaciones análogas al final de la serie. Finalmente, el componente tendencial definido por el filtro HP viene dado por $X^g = \Gamma^{-1}X$. Este método alternativo tiene la atractiva propiedad de que no perdemos observaciones como consecuencia del proceso de filtrado.

Los valores del parámetro de alisado, λ , sugeridos por Kydland y Prescott (1990) son $\lambda = 1600$ para datos trimestrales y $\lambda = 400$ para datos anuales, puesto que éstos son los ratios de la volatilidad del componente cíclico relativa a la volatilidad del componente tendencial en la economía norteamericana. no obstante, y por razones que se comentarán más adelante, analizaremos otros valores de λ .

3.2. El filtro de Baxter y King. La aplicación del filtro de Hodrick Y Precott es frecuentemente ad hoc en el sentido de que se busca simplemente un componente de ciclo de negocios estacionario sin especificar explícitamente las características estadísticas de los ciclos de negocios. Por el contrario, Baxter y King (1995, 1999) desarrollan una metodología para medir los ciclos de negocio que requiere que el investigador comience por una especificación previa de dichas características. Técnicamente, lo que hacen es desarrollar filtros aproximados de tipo *pasa-banda*, es decir, filtros que permiten solo el paso de frecuencias entre $\underline{\omega}$ y $\bar{\omega}$ del espectro correspondiente (ver, por ejemplo Christiano y Fitzgerald (1999)). Estos filtros de tipo *pasa-banda* son esencialmente medias móviles de orden infinito. En realidad, la mayoría de los filtros utilizados con series temporales macroeconómicas son medias móviles bi-direccionales, del tipo $h(B) = \sum_{j=-\infty}^{\infty} h_j B^j$, que, por razones prácticas, tienen que aproximarse por una media móvil bi-direccional de orden p , $h_p(B) = \sum_{j=-n}^n h_j B^j$. De hecho, nos centraremos en análisis de medias móviles simétricas, donde $h_j = h_{-j}$ para todo j .

Las implicaciones de estos filtros bi-direccionales se aprecian claramente observando en el dominio de frecuencias las correspondientes *funciones de respuesta de frecuencias*. La función de respuesta de frecuencias de la ecuación $MA(\infty)$ bi-direccional se define como

$$\beta(\omega) = \sum_{j=-\infty}^{\infty} b_j e^{-i\omega j}, \quad (3.8)$$

en tanto que para la $MA(n)$ finita bi-direccional, la función de respuesta de frecuencias es

$$\alpha(\omega) = \sum_{j=-n}^n a_j e^{-i\omega j}. \quad (3.9)$$

Baxter y King (1995) demuestra que el filtro aproximado óptimo (en el sentido de minimizar el error cuadrático medio de la discrepancia $\delta(\omega) = \beta(\omega) - \alpha(\omega)$) para un número de retardos máximo n , se construye simplemente haciendo $a_j = b_j$ para $j = 0, 1, \dots, n$, y $a_j = 0$ para $j \geq n + 1$. Además, demuestran que los filtros de medias móviles simétricos con pesos que suman cero, es decir, $\sum_{j=-n}^n a_j = 0$, tienen propiedades de eliminación de tendencia. En realidad, estas medias móviles convierten en estacionarias series que contienen tendencias determinísticas cuadráticas o incluso procesos $I(1)$ e $I(2)$. Obsérvese que $\sum_{j=-n}^n a_j = 0$ si y sólo si $\alpha(0) = 0$.

Estos filtros que eliminan la tendencia reciben el nombre de filtros de tipo *pasa-alta* porque permiten el paso de los componentes de los datos con frecuencia mayor que un valor predeterminado $\bar{\omega}$ próximo a cero, esto es, $\beta(\omega) = 0$ para $|\omega| < \bar{\omega}$ y $\beta(\omega) = 1$ para $|\omega| \geq \bar{\omega}$. Por el contrario, los filtros denominados de tipo *pasa-baja*, mantienen únicamente los componentes de los datos con frecuencias más bajas (periodicidades mayores). Un filtro pasa-baja simétrico ideal, permite el paso únicamente de las frecuencias $-\underline{\omega} \leq \omega \leq \underline{\omega}$, teniendo como función de respuesta de frecuencias $\beta(\omega) = 1$ para $|\omega| \leq \underline{\omega}$ y $\beta(\omega) = 0$ para $|\omega| > \underline{\omega}$. Por tanto, las frecuencias bajas (movimientos con periodicidad alta) permanecen inalteradas, en tanto que el resto de frecuencias se eliminan. En términos de filtro $MA(n)$ finito simétrico, esto quiere decir que los filtros de tipo pasa-baja deben satisfacer la condición $\sum_{j=-n}^n a_j = 1$.

Si $b(L) = \sum_{j=-\infty}^{\infty} b_j L^j$ es la representación en el dominio temporal de filtro pasa-baja simétrico ideal, Baxter y King (1995, 1999) demuestran que los pesos del filtro b_j pueden calcularse a partir de la transformación inversa de Fourier de la función de respuesta de frecuencias, esto es,

$$b_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} \beta(\omega) e^{i\omega j} d\omega,$$

dando como resultado $b_0 = \underline{\omega}/\pi$ y $b_j = \sin(j\underline{\omega})/j\pi$, $j = 1, 2, \dots$. El filtro pasa-alta complementario tiene como coeficientes $(1 - b_0)$ en $j = 0$ en $-b_j$ para $j = 1, 2, \dots$. En tanto que los pesos tienden a cero a medida que j crece, nótese que se necesita una media móvil de orden infinito para construir el filtro ideal. Esto nos lleva a considerar la aproximación del filtro ideal a partir de una media móvil finita $a(B) = \sum_{j=-n}^n a_j B^j$.

Habitualmente, la definición de los ciclos de negocios se asocia con la duración de los ciclos del NBER definida por Burns y Mitchell (1946), en la que $\underline{\omega}$ corresponde a 32 trimestres (ocho años) and $\bar{\omega}$ a seis trimestres (dieciocho meses). Así pues, el componente tendencia se obtiene a partir de un filtro pasa-baja con frecuencias de ocho años o menos. Este es el filtro que denominamos BK en nuestros experimentos posteriores.

3.3. Filtros no lineales. Los filtros HO y BK son ejemplos de filtros lineales. Para mayor exhaustividad, también se considera una clase de filtros no lineales denominada *filtros de la mediana* (Wen y Zeng, 1999), que en tiempos recientes se ha utilizado con una gran utilidad en procesamiento de señales en distintos campos de Ingeniería. los filtros de la mediana tienen dos propiedades muy interesantes: preservan el filo (cambios bruscos de tendencia) y atenúan el ruido de forma robusta frente a innovaciones de tipo impulso. Ninguna de estas propiedades puede obtenerse mediante los filtros lineales tradicionales. Para calcular el filtro de la mediana, se toma una ventana de datos consecutivos de la serie observada y se ordena para obtener el valor central o mediana, que es el output del filtro. Si la longitud (de la ventana) del filtro es $2n + 1$, el procedimiento de filtrado se denota como

$$\text{med}\{x_{t-n}, x_{t-n+1}, \dots, x_t, \dots, x_{t+n}\}. \quad (3.10)$$

Para filtrar también los extremos de las series, en los que parte de las ventanas de los filtros están fuera de la muestra, se suelen añadir los elementos extremos de la serie para obtener el tamaño requerido.

En el caso del filtro de la mediana carece de sentido el análisis del dominio de frecuencias y la función de respuesta, puesto que la función de respuesta a impulsos del filtro de la mediana es cero siempre. No obstante, una propiedad muy importante del filtro de la mediana es la denominada *convergencia en raíz*, es decir, que cualquier serie temporal finita contiene una serie de señales que son invariantes a filtro de la mediana. Desde un punto de vista económico esto es muy importante, puesto que si se produce alguna ruptura en los fundamentos económicos de la serie, puede que se vea perfectamente tras realizar la operación de filtrado. Ver Wen y Zeng (1999) para más detalles.

4. Filtrado de Procesos con Raíz Unitaria

Tal y como se vio en la Sección 2, los contrastes de raíces unitarias en presencia de AO's sistemáticos es similar a contrastar la presencia de raíces unitarias con perturbaciones que siguen procesos MA . Esto puede sugerirnos el uso de contrastes más generales, como el ADF o PP, o bien estadísticos de raíces unitarias robustos a errores MA , como pueden ser las modificaciones de los estadísticos PP propuestas por Perron y Ng (1996). En efecto, Vogelsang (1999) muestra evidencia experimental que apoya las buenas propiedades de tamaño y potencia de estos estadísticos PP modificados en el caso de presencia de atípicos. En ese aspecto, su artículo y el presente Capítulo deben entenderse como complementarios y no excluyentes.

En este caso nos vamos a centrar en el estudio de dos posibles alternativas para realizar contrastes de raíces unitarias en presencia de AO's. En primer lugar, estamos interesados en tratar de aproximar los AO's temporales añadiendo retardos de la serie contaminada, x_t en la ecuación (2.2), y, por tanto, estimar el modelo

$$\Delta x_t = c + \rho x_{t-1} + \sum_{i=1}^p \phi_i \Delta x_{t-i} + \eta_t, \quad (4.1)$$

donde el número de retardos, p , que se puede elegir por medio de criterio de selección de retardos, es tal que η_t es un proceso de ruido blanco. Evidentemente, el estadístico t correspondiente al coeficiente ρ en la ecuación (4.1) para contrastar la hipótesis nula $\rho = 0$ es el contraste ADF convencional.

En segundo lugar, queremos analizar el impacto que tiene aplicar el contraste de raíz unitaria en el componente tendencial, x_t^g , de la serie observada x_t . El procedimiento es el siguiente: Separamos x_t en los componentes no observados x_t^g y x_t^c como en la ecuación (3.1). Al sustituir en (2.6), se tiene

$$\Delta x_t^g = \rho x_{t-1}^g + \xi_t, \quad (4.2)$$

donde $\xi_t = u_t - \Delta x_t^c = \varepsilon_t - \Delta (x_t^c - \theta \delta_t)$ bajo la hipótesis nula $\rho = 0$. Ahora bien, dado que los AO's, generados por δ_t , son transitorios, deberían formar parte, en gran medida, del componente cíclico, no del tendencial. Consecuentemente, $(x_t^c - \theta \delta_t)$ será un proceso $I(0)$ libre de AO's que puede aproximarse bien incluyendo términos autorregresivos de Δx_t^g en la ecuación (4.2).

Para obtener x_t^g se puede utilizar un filtro pasa-baja aproximado, esto es,

$$x_t^g = a(B) x_t = \sum_{j=-n}^n a_j x_{t-j}, \quad (4.3)$$

siendo $\sum_{j=-n}^n a_j = 1$. No hay una elección 'óptima' de n . Al aumentar n se consigue una mejor aproximación al filtro ideal de dimensión infinita, pero al precio de perder más observaciones al inicio y final de la muestra. Por tanto, en última instancia, la elección de n dependerá del tamaño de la muestra y de la importancia otorgada a la precisión de la aproximación al filtro ideal. No obstante, tras un proceso de experimentación, Baxter y King (1999) recomiendan trabajar con los datos de los tres años precedentes y los tres años posteriores, además de la observación corriente, tanto si trabajamos con datos anuales o trimestrales. Además, siguiendo sus recomendaciones, se requiere que el filtro pasa-baja sea *simétrico* ($a_j = a_{-j}$, $j = 1, 2, \dots, n$) e *invariante en el tiempo*, con coeficientes que no dependan del punto de la muestra en que nos encontremos. Estas condiciones evitan que los filtros introduzcan *desfases* en las series (alteración de las relaciones de sincronía para alguna frecuencia) y dependencias de la longitud del período muestral.

Sean $\hat{\rho}_{AO}^g$ y $t_{\hat{\rho}_{AO}^g}$ los estadísticos de Dickey-Fuller aplicados a la ecuación (4.2), esto es,

$$\hat{\rho}_{AO}^g = \left(\sum_{t=1}^T x_{t-1}^g \Delta x_t^g \right) / \left(\sum_{t=1}^T (x_{t-1}^g)^2 \right), \quad (4.4)$$

y

$$t_{\hat{\rho}_{AO}^g} = \hat{\rho}_{AO}^g / \hat{\sigma}_x^g \left(\sum_{t=1}^T (x_{t-1}^g)^2 \right)^{-1/2}, \quad (4.5)$$

con $\hat{\sigma}_x^g = \left(T^{-1} \sum_{t=1}^T (\Delta x_t^g - \hat{\rho}_{AO}^g x_{t-1}^g)^2 \right)^{1/2}$. Además, y sin pérdida de generalidad, supongamos la condición inicial $x_0^g = 0$. El comportamiento en el límite de $\hat{\rho}_{AO}^g$ y su t ratio vienen dados por el siguiente Teorema.

TEOREMA 1.4.1. *Sea $T > 2n + 1$ para n fijo. Entonces, a medida que el tamaño muestral aumenta, tenemos que*

$$T \hat{\rho}_{AO}^g \Rightarrow \left(\int_0^1 W(r) dW(r) + \Phi \right) \left(\int_0^1 W^2(r) dr \right)^{-1} \quad (4.6)$$

$$t_{\hat{\rho}_{AO}^g} \Rightarrow \Psi \left(\int_0^1 W(r) dW(r) + \Phi \right) \left(\int_0^1 W^2(r) dr \right)^{-1/2} \quad (4.7)$$

donde

$$\Phi = \frac{1}{2} \sum_{j=-n}^n a_j(1 - a_j) - (\theta/\sigma_\varepsilon)^2 \pi \sum_{j=-n}^n a_j(a_j - a_{j-1})$$

$$\Psi = \left[\sum_{j=-n}^n a_j^2 + 2(\theta/\sigma_\varepsilon)^2 \pi \sum_{j=-n}^n a_j(a_j - a_{j-1}) \right]^{-1/2}.$$

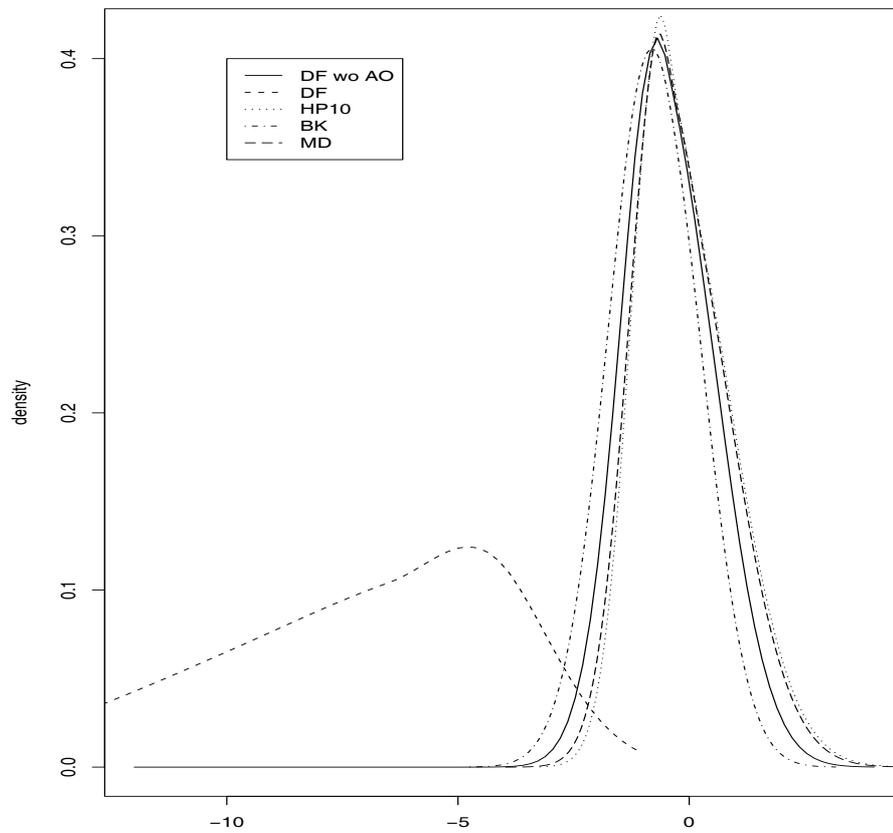


FIGURA 1.1. Densidad estimada del estadístico t del contraste DF bajo la hipótesis nula de raíz unitaria, basada en 50000 réplicas, $T = 1000$. $\pi = 0.1$ y $\theta = 16$. DF wo AO y DF se refieren a las distribuciones del estadístico t en las expresiones (2.5) y (2.8), respectivamente; HP10, BK y MD corresponden a la distribución del estadístico t en la expresión (4.5), utilizando los filtros de Hodrick–Prescott ($\lambda = 10$), Baxter–King y mediana.

Por tanto, al igual que en el caso no filtrado, ρ se estima de forma (super-consistente, y las distribuciones límite tienen parámetros molestos que dependen de

$\theta, \pi, \sigma_\varepsilon^2, n$ y las ponderaciones $\{a_j\}_{j=-n}^n$. Cuando $n = 0, a_0 = 1$, se vuelven a obtener las expresiones (2.9) y (2.10). Obsérvese que si elegimos una sucesión de pesos tales que $\sum_{j=-n}^n a_j^2 \approx \sum_{j=-n}^n a_j a_{j-1} \approx 1$, entonces $\Phi \approx 0$ y $\Psi \approx 1$, independientemente de los AO's y para todo n . En el caso particular en que $a_j = 1/(2n+1)$ para $j = -n, \dots, n$, Φ y Ψ son

$$\Phi = \frac{n}{2n+1}, \quad (4.8)$$

$$\Psi = \left[\frac{1}{2n+1} \right]^{-1/2}. \quad (4.9)$$

En la realidad, este filtro pasa-baja *uniforme* o de *media móvil*

$$x_t^g = \frac{1}{2n+1} \sum_{j=-n}^n x_{t-j}, \quad (4.10)$$

en el que el componente de crecimiento o tendencial se define como una media móvil bi-direccional o centrada, se utiliza muy frecuentemente para extraer la tendencia de series temporales económicas.

En la Figura 1.1 se representa la densidad empírica del estadístico t del contraste de Dickey-Fuller (sin constante ni retardos), utilizando la ecuación (2.5) y la ecuación (4.5) con los filtros HP ($\lambda = 10$), BK y de la mediana filters, para un tamaño muestral $T = 1000$ y para $\pi = 0.1, \theta = 16$, es decir, con la presencia de AO's grandes y frecuentes. Para comparar, también se incluye el estadístico de Dickey-Fuller en el caso de que no haya AO's. Se observa claramente el gran desplazamiento a la izquierda de la distribución de Dickey-Fuller en presencia de AO's, así como los beneficios de utilizar las versiones filtradas de la distribución de Dickey-Fuller.

5. Evidencia Empírica

En esta Sección se muestra la evidencia obtenida mediante experimentos de Monte Carlo de las implicaciones numéricas de los resultados analíticos. Los datos se generan de acuerdo al siguiente proceso generador de datos (DGP):

$$\Delta y_t = \rho y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, 1),$$

$$x_t = y_t + \theta \delta_t,$$

con $\pi = 0, 0.01, 0.05, 0.10$, $\theta = 1, 6, 16$ y $T = 100, 200, 500, 1000$. Se considera que, como máximo, el 10% de la serie y_t está contaminada, lo que es habitual en este tipo de estudios, véase, por ejemplo, Franses y Haldrup (1994). Sin embargo, lo habitual es considerar saltos de tamaños 3, 4, y 5, en tanto que aquí vamos a analizar saltos

mucho más abruptos. El diseño de nuestros experimentos implica que tenemos 40 casos en nuestro experimento de Monte Carlo. Para obtener los valores críticos se simula el DGP bajo $H_0 : \rho = 0$, y tomamos el cuantil 5% de la cola inferior de los estadísticos t de Dickey–Fuller ordenados, $t_{\hat{\rho}_{AO}}$ y $t_{\hat{\rho}_{AO}^g}$ como los valores críticos de interés. La potencia se obtiene simulando el DGP bajo $H_1 : \rho = -0.2$, y calculando el porcentaje de rechazos de la hipótesis nula utilizando los valores críticos calculados anteriormente. Para todos los experimentos se utilizan 10.000 réplicas del DGP.

En nuestros experimentos, HP indica el filtro de Hodrick–Prescott, BK el filtro de Baxter y King, U es el filtro uniforme o de media móvil centrada, y MD es el filtro de la mediana. Con respecto al filtro HP, tomamos distintos valores del parámetro λ , en concreto, $\lambda = 10, 100, 400$ y 1600 . El parámetro de la ventana de los filtros BK, U y MD se toma $n = 3$, lo que hace que las ventanas tengan un tamaño de $2n + 1 = 7$ elementos. El número de retardos de los contrastes tipo ADF se eligen mediante el criterio de información SIC. Los resultados se reflejan en las Tablas 1.1 a 1.5.

π	θ	DF				ADF			
		$T=100$	$T=200$	$T=500$	$T=1000$	$T=100$	$T=200$	$T=500$	$T=1000$
0		-2.899	-2.879	-2.866	-2.850	-2.920	-2.889	-2.865	-2.850
	1	-2.927	-2.887	-2.870	-2.873	-2.942	-2.891	-2.870	-2.873
0.01	6	-3.546	-3.227	-3.136	-3.642	-3.429	-3.111	-3.002	-2.975
	16	-5.814	-4.870	-4.586	-6.902	-5.318	-3.388	-3.123	-3.168
	1	-3.057	-2.979	-2.941	-2.952	-3.053	-2.965	-2.933	-2.933
0.05	6	-5.949	-5.539	-5.201	-5.781	-5.501	-3.892	-3.390	-3.081
	16	-8.909	-9.815	-10.886	-13.024	-8.909	-9.571	-3.584	-3.451
	1	-3.139	-3.074	-3.066	-3.069	-3.118	-3.052	-2.994	-2.910
0.10	6	-6.603	-6.845	-7.479	-7.852	-6.570	-4.704	-3.404	-3.334
	16	-8.819	-10.988	-14.975	-17.319	-8.819	-10.988	-5.915	-3.751

TABLA 1.1. Valores críticos al 5% de los contrastes DF y ADF con AO's. Modelos estimados (5.1)–(5.2).

Consideremos en primer lugar los efectos de los AO's en los contrastes de Dickey–Fuller habituales. Las regresiones estimadas son:

$$\Delta x_t = c + \rho x_{t-1} + \eta_t, \quad (5.1)$$

y

$$\Delta x_t = c + \rho x_{t-1} + \sum_{i=1}^p \phi_i \Delta x_{t-i} + \eta_t. \quad (5.2)$$

Tal y como se aprecia en la Tabla 1.1, los valores críticos de $t_{\hat{\rho}_{AO}}$ son muy inestables en el caso del modelo (5.1), pasando de -2.85 ($\pi = 0$) a -17.319 ($\pi = 0.1$) para $T = 1000$ y $\theta = 16$. Con respecto al modelo (5.2), en la Tabla 1.1 se observa que, en el

π	θ	DF				ADF			
		$T=100$	$T=200$	$T=500$	$T=1000$	$T=100$	$T=200$	$T=500$	$T=1000$
0		87.55	99.99	100.00	100.00	86.35	99.95	100.0	100.0
0.01	1	95.58	95.63	95.57	95.96	74.14	99.85	100.0	100.0
	6	98.74	97.83	97.45	99.36	76.85	99.61	100.0	100.0
	16	100.00	99.98	99.85	100.00	76.48	99.70	100.0	100.0
0.05	1	96.58	96.49	96.33	96.59	76.57	99.83	100.0	100.0
	6	100.00	100.00	100.00	100.00	76.56	99.84	100.0	100.0
	16	100.00	100.00	100.00	100.00	60.92	99.92	100.0	100.0
0.10	1	97.16	97.10	97.05	97.24	78.04	99.75	100.0	100.0
	6	100.00	100.00	100.00	100.00	80.89	99.80	100.0	100.0
	16	100.00	100.00	100.00	100.00	52.38	99.83	100.0	100.0

TABLA 1.2. Potencia ajustada de los contrastes DF y ADF con AO's.
 $\rho = -0.2$. Modelos estimados (5.1)–(5.2).

caso en que no hay AO's ($\pi = 0$), los valores críticos son prácticamente los mismos que cuando no utilizamos retardos en el modelo, lo que indica que el criterio de selección de retardos funciona razonablemente bien. Sin embargo, los valores críticos siguen siendo muy inestables, pasando de -2.9 a -11, para $T < 500$, y algo más estables para $T > 500$, pasando de -2.85 a -6. Por otra parte, la potencia ajustada de $t_{\hat{\rho}_{AO}}$ frente a $H_1 : \rho = -0.2$ es alta para $T > 100$ en ambos casos, como se desprende la Tabla 1.2.

Consideremos ahora el comportamiento de $t_{\hat{\rho}_{AO}}^g$. El modelo estimado es

$$\Delta x_t^g = c + \rho x_{t-1}^g + \sum_{i=1}^p \phi_i \Delta x_{t-i}^g + \xi_t. \quad (5.3)$$

La Tabla 1.3 indica que los valores críticos son muy estables incluso para tamaños muestrales tales como $T = 100$ en el caso del filtro HP o $T = 200$ en el caso del filtro BK. Por ejemplo, los valores críticos se encuentran entre -2.8 y -3.22 en el último caso. En cuanto a los valores críticos obtenidos con el filtro HP, se observa que no hay apenas variación al modificar el parámetro λ , especialmente en muestras de tamaño moderado y grande. Los valores críticos son también muy estables respecto a π , θ y T al utilizar los filtros U y MD. Globalmente, los valores críticos más estables se obtienen al utilizar el filtro MD. Es interesante resaltar que, en el caso de que no haya observaciones atípicas, ni el procedimiento de filtrado utilizado ni el valor de λ afectan mucho al valor crítico. De hecho, los valores críticos obtenidos con la serie filtrada son muy similares a los obtenidos con la serie observada. Por otra parte, la potencia (ajustada) es alta en todos los casos cuando los tamaños son moderado o grande, ver Tabla 1.4. No son tan altas como las obtenidas con las series observadas

Filtro BK										
T	$\pi = 0$	$\pi = 0.01$			$\pi = 0.05$			$\pi = 0.1$		
		$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$
100	-3.023	-2.958	-2.938	-3.344	-2.975	-3.158	-3.295	-2.951	-3.385	-5.630
200	-2.928	-2.945	-2.890	-2.950	-2.917	-2.920	-3.247	-2.923	-2.899	-3.072
500	-2.855	-2.852	-2.818	-2.808	-2.844	-2.789	-2.862	-2.833	-2.823	-3.293
1000	-2.855	-2.870	-2.817	-2.844	-2.861	-2.840	-2.933	-2.847	-2.842	-3.072
Filtro HP10										
100	-3.195	-3.211	-3.081	-2.897	-3.201	-2.963	-2.958	-3.173	-2.967	-3.220
200	-3.038	-3.040	-3.092	-2.932	-3.079	-3.019	-3.106	-3.079	-3.022	-3.416
500	-2.859	-2.877	-2.993	-2.970	-2.912	-2.971	-2.956	-2.995	-2.831	-3.143
1000	-2.858	-2.872	-3.013	-2.931	-2.891	-2.926	-3.041	-2.916	-2.885	-3.350
Filtro HP100										
100	-3.171	-3.188	-3.535	-3.068	-3.329	-3.082	-2.894	-3.367	-3.082	-3.022
200	-2.907	-2.918	-3.030	-3.201	-2.933	-3.181	-3.108	-2.938	-3.108	-3.213
500	-2.861	-2.871	-2.940	-3.137	-2.887	-3.136	-3.160	-2.931	-3.070	-3.036
1000	-2.847	-2.870	-2.965	-3.039	-2.887	-3.102	-3.054	-2.914	-3.013	-3.111
Filtro HP400										
100	-3.232	-3.215	-3.752	-3.385	-3.311	-3.314	-2.818	-3.344	-3.255	-2.864
200	-2.900	-2.943	-3.033	-3.455	-2.963	-3.304	-3.167	-2.977	-3.185	-3.121
500	-2.876	-2.883	-2.963	-3.112	-2.904	-3.021	-3.132	-2.950	-3.193	-3.078
1000	-2.853	-2.867	-2.970	-3.103	-2.896	-2.987	-3.013	-2.918	-3.106	-3.053
Filtro HP1600										
100	-3.389	-3.419	-3.865	-3.990	-3.471	-3.878	-3.024	-3.502	-3.721	-3.145
200	-2.982	-3.008	-3.130	-3.503	-3.043	-3.363	-3.074	-3.071	-3.364	-3.088
500	-2.914	-2.905	-3.000	-3.147	-2.934	-3.052	-3.090	-2.984	-3.223	-3.053
1000	-2.869	-2.878	-2.951	-3.171	-2.913	-2.982	-2.972	-2.932	-3.060	-2.919
Filtro U										
100	-3.562	-3.55	-3.22	-2.91	-3.48	-3.03	-2.62	-3.42	-3.01	-3.11
200	-3.234	-3.22	-3.12	-2.90	-3.21	-2.87	-2.69	-3.17	-2.99	-2.88
500	-3.229	-3.22	-3.13	-2.96	-3.21	-2.94	-2.80	-3.15	-3.07	-2.96
1000	-3.157	-3.15	-3.03	-2.98	-3.14	-3.00	-2.90	-3.13	-2.98	-3.05
Filtro MD										
100	-2.885	-2.902	-2.899	-2.899	-2.900	-2.860	-2.859	-2.914	-2.831	-2.831
200	-2.935	-2.939	-2.935	-2.935	-2.925	-2.885	-2.884	-2.916	-2.836	-2.836
500	-2.973	-2.973	-2.972	-2.972	-2.968	-2.940	-2.940	-2.969	-2.868	-2.668
1000	-2.994	-3.009	-2.997	-2.997	-3.002	-2.963	-2.964	-3.008	-2.917	-2.847

TABLA 1.3. Valores críticos al 5% de los contrastes de raíces unitarias filtrados con AO's. El modelo estimado es (5.3)

(sin filtrar) para tamaños muestrales pequeños, pero éste resultado era de esperar (cf. Ghysels y Perron, 1993).

Por último, es importante analizar la robustez de los valores críticos obtenidos en las Tablas 1.1 y 1.3. En particular, en la Tabla 1.5 se presentan algunos resultados obtenidos con los diferentes contrastes al utilizar los valores críticos correspondientes a $\pi = 0$. Son evidentes las distorsiones en el tamaño del contraste ADF en el modelo (5.2) especialmente para $\theta = 16$, y el buen comportamiento del filtro MD para los

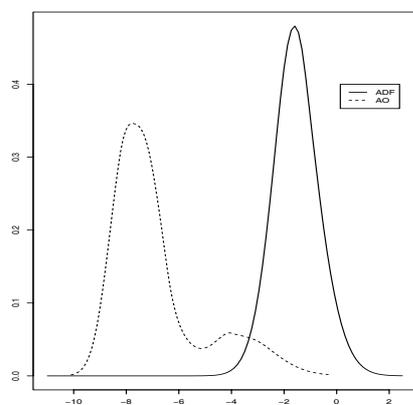


Figura 1.2.1: ADF

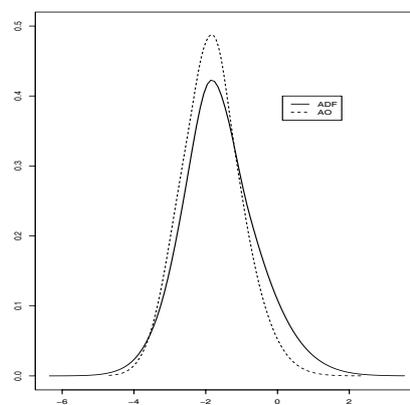


Figura 1.2.2: HP10

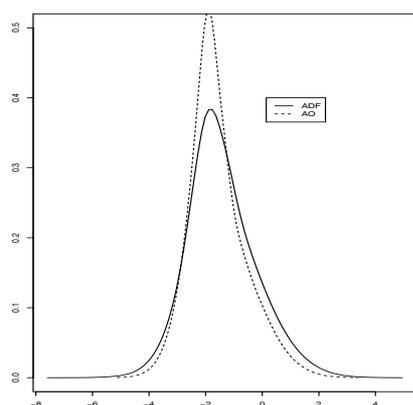


Figura 1.2.3: HP100

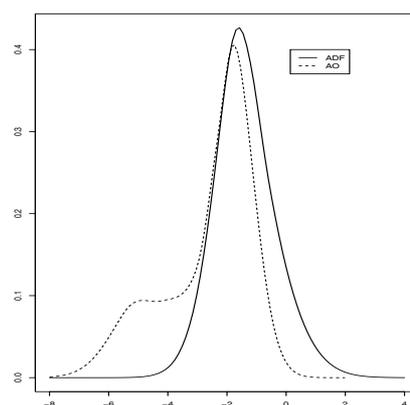


Figura 1.2.4: BK

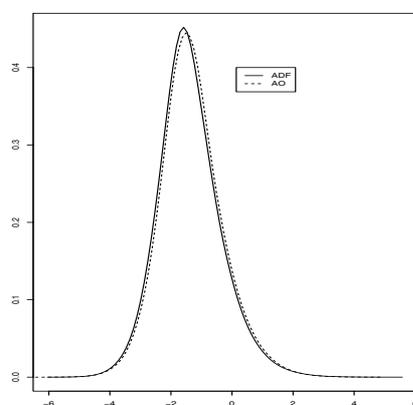


Figura 1.2.5: MD

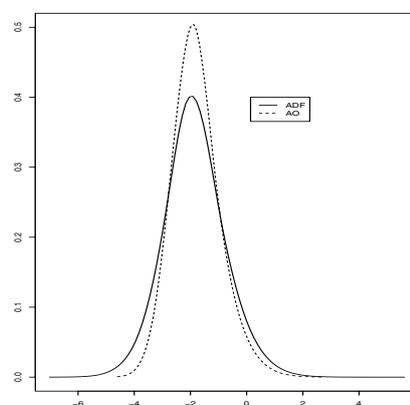


Figura 1.2.6: U

FIGURA 1.2. Densidades estimadas de los contrastes de raíces unitarias ADF y filtrados con y sin AO's, $T = 100$, $\pi = 0.1$, $\theta = 16$.

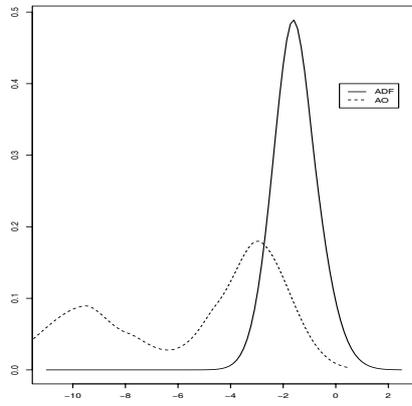


Figura 1.3.1: ADF

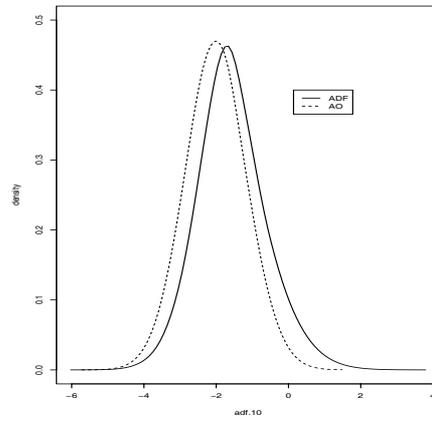


Figura 1.3.2: HP10

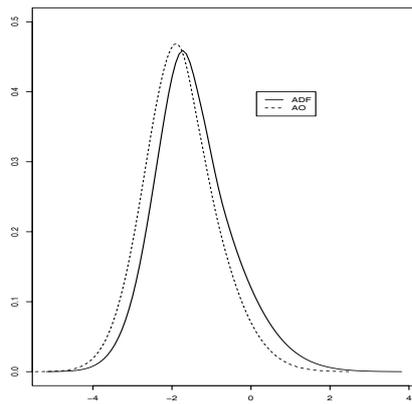


Figura 1.3.3: HP100

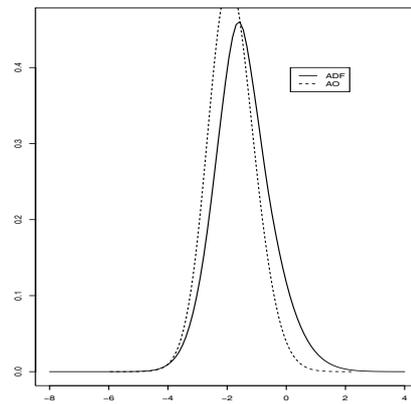


Figura 1.3.4: BK

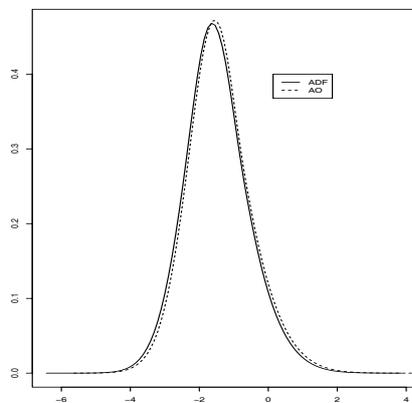


Figura 1.3.5: MD

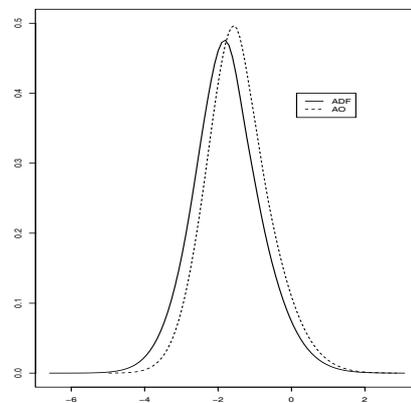


Figura 1.3.6: U

FIGURA 1.3. Densidades estimadas de los contrastes de raíces unitarias ADF y filtrados con y sin AO's, $T = 200$, $\pi = 0.1$, $\theta = 16$.

Filtro BK										
T	$\pi = 0$	$\pi = 0.01$			$\pi = 0.05$			$\pi = 0.1$		
		$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$
100	21.85	23.93	32.43	47.08	25.42	39.28	44.67	26.54	50.23	55.81
200	69.04	68.41	76.42	79.69	71.42	81.57	90.36	73.84	86.21	99.56
500	99.99	99.99	100.00	100.00	99.99	100.00	100.00	99.99	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Filtro HP10										
100	30.42	12.32	14.41	21.70	12.53	24.91	40.53	13.14	25.99	40.77
200	74.81	59.43	59.25	72.10	58.38	76.11	92.52	58.96	83.19	96.04
500	100.00	99.95	99.93	99.99	99.93	100.00	100.00	99.91	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Filtro HP100										
100	15.43	11.89	5.97	11.77	9.33	13.87	20.37	8.59	13.08	16.78
200	44.36	35.16	30.35	23.07	35.17	26.96	38.58	35.01	35.65	50.11
500	98.67	96.45	95.71	93.51	96.37	95.60	99.50	95.87	97.51	99.94
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Filtro HP400										
100	7.69	20.80	12.50	15.11	18.99	19.08	27.25	18.37	17.19	18.06
200	28.99	30.47	26.35	12.02	29.45	17.44	29.32	28.86	22.35	31.56
500	91.80	84.59	81.85	77.22	84.16	83.16	93.78	82.25	78.21	93.69
1000	99.97	100.00	100.00	99.99	99.99	99.98	100.00	99.99	99.99	100.00
Filtro HP1600										
100	4.78	28.34	20.35	17.75	27.37	21.23	38.01	26.56	22.88	29.77
200	15.75	32.04	29.04	19.78	31.67	26.02	44.72	30.84	23.78	39.01
500	72.66	66.09	62.45	57.80	64.55	59.63	70.84	62.15	47.82	62.28
1000	99.01	99.15	99.11	98.45	99.21	99.13	99.29	99.15	98.99	99.95
Filtro U										
100	28.78	29.07	29.36	23.45	28.93	33.92	21.31	28.33	27.25	38.88
200	60.44	61.33	60.30	54.85	59.92	61.16	64.72	60.60	58.70	54.09
500	99.95	99.96	99.96	99.96	99.94	99.94	99.98	99.95	99.99	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Filtro MD										
100	33.81	32.64	32.29	32.31	32.33	30.24	30.26	31.15	29.68	29.71
200	86.70	86.59	86.53	86.53	86.50	86.35	86.37	87.02	86.57	86.59
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

TABLA I.4. Potencia ajustada de los contrastes de raíces unitarias filtrados con AO's. $\rho = -0.2$. Modelo estimado (5.3).

diferentes valores de π , θ y T incluidos en nuestro diseño del experimento de Monte Carlo.

Las Figuras 1.2 y 1.3 nos proporcionan una evidencia gráfica de los comentarios realizados sobre la estabilidad de los valores críticos. Obsérvese en particular que la distribución de estadístico del contraste ADF con la serie filtrada con el filtro MD es prácticamente la misma con y sin AO's (Figuras 1.2.5 and 1.3.5), contrastando con la versión sin filtrar del contraste ADF (Figuras 1.2.1 y 1.3.1). Por otra parte, los problemas del tamaño de la versión filtrada con BK se hacen evidentes al observar

ADF									
T	$\pi = 0.01$			$\pi = 0.05$			$\pi = 0.1$		
	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$
100	5.26	11.78	23.67	6.64	27.91	88.53	7.66	46.36	97.00
200	5.01	7.82	12.30	5.89	20.81	41.52	7.08	23.57	67.50
500	5.05	6.94	8.72	5.81	12.12	12.11	6.79	13.21	34.97
1000	5.35	6.76	9.83	6.09	8.36	15.37	5.86	12.97	20.30
Filtro BK									
100	4.98	4.71	8.67	5.00	6.99	8.76	4.91	9.81	31.67
200	4.96	4.39	5.18	4.62	4.80	8.91	4.70	4.66	7.43
500	4.97	4.39	4.40	4.83	4.06	5.11	4.62	4.47	12.30
1000	4.82	4.24	4.41	4.70	4.48	5.72	4.47	4.60	8.22
Filtro HP10									
100	5.00	3.69	2.31	4.92	2.76	2.39	4.64	2.86	5.44
200	5.12	5.58	3.85	5.58	4.88	6.05	5.64	4.83	11.22
500	5.03	6.60	6.22	5.44	6.36	6.35	6.71	4.64	9.51
1000	5.10	7.07	5.93	5.31	5.80	7.43	5.73	5.17	13.53
Filtro HP100									
100	5.24	8.95	3.71	6.21	4.07	2.30	6.63	4.07	3.73
200	5.00	6.38	9.04	5.15	8.56	8.21	5.20	7.94	10.25
500	4.97	5.95	9.02	5.24	9.14	9.53	5.80	8.10	7.74
1000	5.06	6.30	7.40	5.30	8.53	7.87	5.64	7.30	9.19
Filtro HP400									
100	5.06	9.57	6.31	5.72	5.80	1.51	6.11	5.35	1.82
200	5.01	6.15	12.64	5.30	10.43	8.95	5.53	8.66	8.26
500	5.02	6.27	8.41	5.35	6.88	8.67	5.98	9.71	7.93
1000	5.14	6.62	8.67	5.48	6.83	7.12	5.88	8.79	7.82
Filtro HP1600									
100	5.13	8.81	11.49	5.68	10.29	1.84	5.96	8.69	2.42
200	5.12	6.45	11.80	5.49	9.76	6.14	5.63	9.33	6.32
500	5.07	6.37	8.55	5.41	6.90	7.40	6.05	9.34	6.84
1000	5.15	6.28	9.75	5.62	6.73	6.46	5.85	7.87	5.63
Filtro U									
100	4.83	2.71	1.08	4.36	1.42	0.16	3.72	1.50	1.03
200	4.83	3.85	2.25	4.77	1.89	0.89	4.45	2.74	1.81
500	4.93	3.97	2.64	4.72	2.60	1.51	4.11	3.46	2.52
1000	4.94	3.68	3.26	4.87	3.17	2.49	4.65	3.17	3.76
Filtro MD									
100	5.13	5.06	5.05	5.21	4.54	4.53	5.28	4.42	4.42
200	4.89	4.84	4.84	4.82	4.21	4.20	4.73	3.84	3.83
500	4.95	4.97	4.97	4.92	4.48	4.46	4.84	3.79	2.06
1000	4.96	4.88	4.88	4.97	4.42	4.43	5.01	3.87	3.23

TABLA 1.5. Tamaño de los contrastes de raíces unitarias ADF y filtrados con AO's. Los valores críticos son los obtenidos en las Tablas 1.1 y 1.3 con $\pi = 0$

las diferentes colas inferiores con $T = 100$ (Figura 1.2.4) y $T = 200$ (Figura 1.3.4). Todos estos gráficos apoyan los resultados obtenidos en la Tabla 1.5.

En las Figuras 1.4 y 1.5 se reflejan las potencias (ajustadas) de los contrastes ADF para los datos observados y las series filtradas con AO's ($\pi = 0.1, \theta = 16$) y sin atípicos, y para $T = 100$ (Figura 1.4) y $T = 200$ (Figura 1.5). Con el fin de realizar

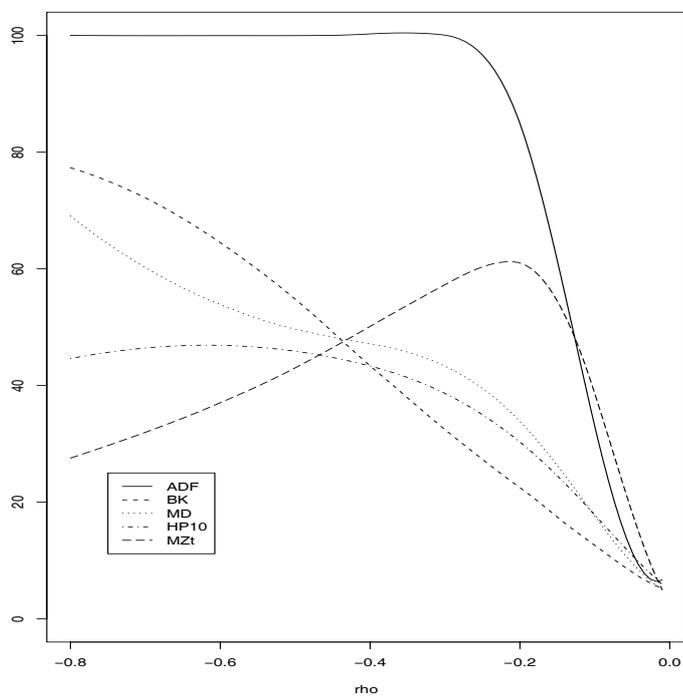


Figura 1.4.1: No atípicos

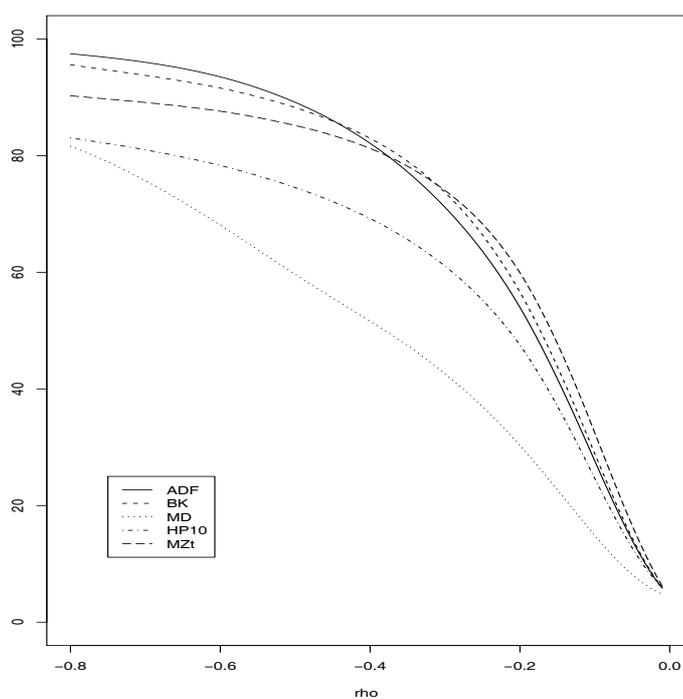
Figura 1.4.2: AO's, $\pi = 0.1$, $\theta = 16$

FIGURA 1.4. Potencia ajustada de los contrastes de raíces unitarias ADF, MZ_t y filtrados con y sin atípicos, $T = 100$, $\pi = 0.1$ y $\theta = 16$

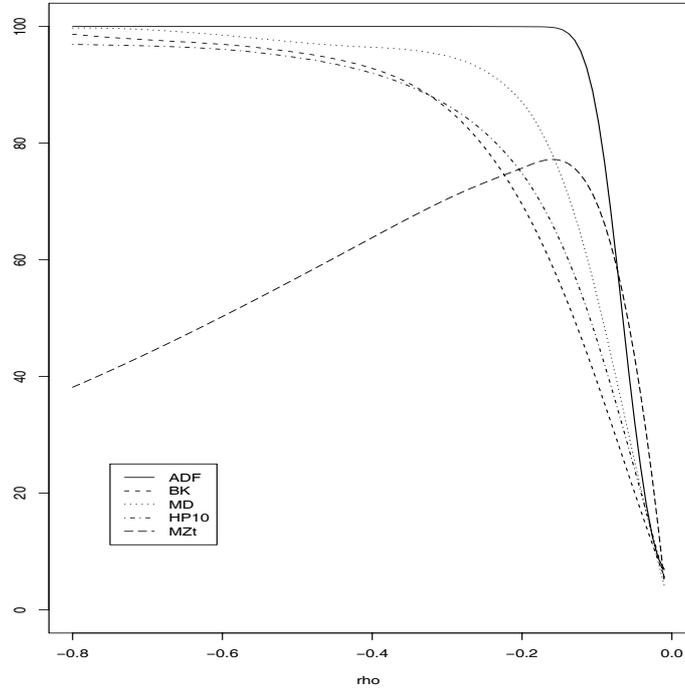


Figura 1.5.1: No atípicos

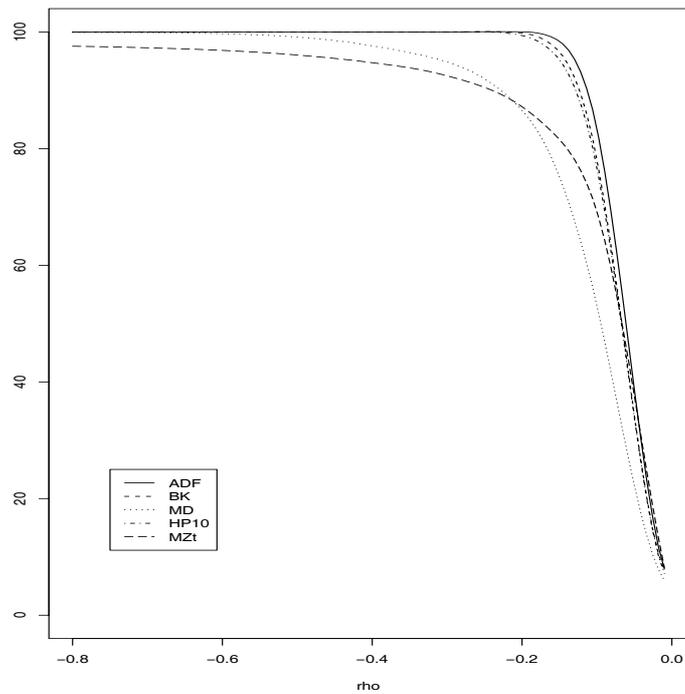
Figura 1.5.2: AO, $\pi = 0.1$, $s = 16$

FIGURA 1.5. Potencia ajustada de los contrastes de raíces unitarias ADF, MZ_t y filtrados con y sin atípicos, $T = 100$, $\pi = 0.1$ y $\theta = 16$

un estudio lo más completo posible, también se incluye el estadístico MZ_t propuesto por Perron y Ng (1996), y definido como

$$MZ_\alpha = \frac{T^{-1}\tilde{x}_T^2 - s^2}{2T^{-2}\sum_{t=1}^T\tilde{x}_t^2} \quad (5.4)$$

$$MSB = \sqrt{T^{-2}\frac{\sum_{t=1}^T\tilde{x}_{t-1}^2}{s^2}} \quad (5.5)$$

$$MZ_t = MSB \times MZ_\alpha \quad (5.6)$$

donde \tilde{x}_t es la serie en desviaciones respecto a la media (GLS), y s_{AR}^2 es el estimador de s^2 (ver Ng y Perron, 1998, Vogelsang, 1999).

$$s_{AR}^2 = \frac{s_{ek}^2}{(1 - \hat{\phi}(1))^2}, \quad (5.7)$$

y $s_{ek}^2 = T^{-1}\sum_{t=k+1}^T\hat{e}_{tk}^2$, $\hat{\phi}(1) = \sum_{j=1}^k\hat{\phi}_j$, donde $\hat{\phi}_j$ y $\{\hat{e}_{jk}\}$ se obtienen de la auto-regresión

$$\Delta\tilde{x}_t = b\tilde{x}_{t-1} + \sum_{j=1}^k\phi_j\Delta\tilde{x}_{t-j} + e_{jk}.$$

El orden de esta auto-regresión, k , se eligen por medio del criterio de información MIC, propuesto por Ng y Perron (1998).

Una primera conclusión de las Figuras 1.4 y 1.5 es la baja potencia del estadístico MZ_t frente a alternativas no locales cuando no hay observaciones atípicas (Figuras 1.4.1 y 1.5.1). En presencia de atípicos y para $T = 100$ (Figura 1.4.2) la potencia de los contrastes filtrados MD y HP10 son uniformemente más bajas que las potencias del resto de los contrastes ADF. la potencia más alta corresponde al contraste MZ_t para alternativas locales y para los contrastes ADF y BK para alternativas no locales. Para $T = 200$ (Figura 1.5.2) la potencia del estadístico MZ_t no es la mayor ni siquiera para alternativas locales; de hecho, tiene la potencia más baja de todas para alternativas no-locales.

A partir del análisis precedente, se deduce que si queremos contrastar la hipótesis de raíces unitarias en presencia de observaciones atípicas de tipo aditivo, la mayor potencia se obtiene utilizando el estadístico MZ_t , junto con las versiones ajustadas de los estadísticos ADF y el contraste ADF-BK. No obstante, debemos ser conscientes de los problemas de tamaño que plantea estos últimos contrastes en muestras de tamaño pequeño o moderado.

T	$\pi = 0$	$\pi = 0.01$			$\pi = 0.05$			$\pi = 0.1$		
		$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$
100	-1.955	-1.794	-1.867	-1.991	-1.840	-1.854	-1.725	-1.849	-1.966	-2.216
200	-1.939	-1.853	-1.907	-1.987	-1.897	-1.859	-1.845	-1.908	-1.972	-1.879
500	-1.951	-1.905	-1.935	-2.002	-1.928	-1.940	-1.925	-1.924	-1.950	-1.938
1000	-1.933	-1.923	-1.948	-1.951	-1.925	-1.970	-2.021	-1.924	-1.989	-2.161

Tabla 1.6.1: Valores críticos de MZ_t con y sin AO's

T	$\pi = 0$	$\pi = 0.01$			$\pi = 0.05$			$\pi = 0.1$		
		$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$
100	60.79	67.19	62.68	53.64	64.35	50.47	69.27	62.98	47.19	70.60
200	75.17	78.02	76.50	73.93	76.35	76.78	89.63	75.65	68.38	90.03
500	91.23	92.09	91.72	91.70	91.85	93.04	96.25	91.94	92.86	97.86
1000	98.33	98.40	98.54	99.25	98.42	99.02	99.83	98.45	99.26	99.98

Tabla 1.6.2: Potencia ajustada del estadístico MZ_t . $\rho = -0.2$.

T	$\pi = 0.01$			$\pi = 0.05$			$\pi = 0.1$		
	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$
100	5.12	6.44	8.55	5.60	5.94	4.13	6.06	7.59	10.90
200	5.10	5.65	6.71	5.60	5.04	4.59	5.82	6.74	5.41
500	5.01	5.43	6.13	5.21	5.42	5.06	5.20	5.44	5.26
1000	5.15	5.42	5.53	5.17	5.84	6.66	5.13	6.18	9.16

Tabla 1.6.3: Tamaño del contraste MZ_t con AO's. Los valores críticos son los obtenidos en la Tabla 1.6.1TABLA 1.6. Propiedades del contraste MZ_t con AO's.

6. El Enfoque Bootstrap

Una de las áreas más productivas en los campos de la Econometría teórica y aplicada es el de los métodos intensivos en computación, tales como el bootstrap ¹. A pesar de que se ha aplicado tradicionalmente en situaciones *iid*, también puede ser útil en el caso de series temporales (ver por ejemplo Li y Maddala, 1996, Kiviet, 1984, Buhlmann, 1999, Berkowitz y Kilian, 2000).

Hay una serie de consideraciones previas acerca de la utilización de la metodología bootstrap para realizar inferencia. El primer punto es resaltar la importancia de los estadísticos (asintóticamente) pivote ² en el bootstrap. Se puede demostrar que el bootstrap proporciona aproximaciones de mayor orden de los valores críticos de estadísticos pivote 'suaves' (Hall, 1988). En este grupo se incluyen los estadísticos de

¹Davidson y Hinkley (1997), Hall (1992), y Efron y Tibshirani (1993) son las referencias básicas sobre el tema.

²La propiedad fundamental de los estadísticos (asintóticamente) pivote es que sus distribuciones (asintóticas) no dependen de parámetros poblacionales desconocidos.

contraste cuyas distribuciones asintóticas son Normales tipificadas o chi-cuadrado. La capacidad del bootstrap de proporcionar refinamientos asintóticos para dichos estadísticos constituye una razón de peso para su utilización en aplicaciones empíricas.

El bootstrap también puede aplicarse a estadísticos que no son asintóticamente pivote, como pueden ser los coeficientes de una regresión, pero en tal caso no proporciona una aproximación de mayor orden de su distribución. Los estimadores bootstrap de las distribuciones de los estadísticos no pivotaes tienen la misma precisión que las aproximaciones asintóticas de primer orden.

Se pueden obtener aproximaciones de mayor orden de las distribuciones de los estadísticos no pivotaes utilizando las técnicas de *prepivoting* o *bootstrap iterado* (Beran, 1987, 1988) o bien método de corrección del sesgo (Efron, 1987). El bootstrap iterado es computacionalmente muy costoso, lo que hace que tenga muy poco interés en los casos en que tenemos un estadístico pivote.

Así pues, una vez vista la importancia de los estadísticos (asintóticamente) pivote para la construcción de intervalos de confianza y los órdenes de aproximación de los diferentes métodos, se ve clara la necesidad de utilizar estadísticos (asintóticamente) pivotaes. Caso contrario, no debemos esperar ninguna mejora respecto a los resultados asintóticos.

Un problema que se plantea cuando aplicamos el bootstrap a series temporales es la dependencia de nuestras observaciones. Una de las soluciones propuestas en la Literatura es dividir los datos (o residuos) en bloques y aplicar las técnicas de remuestreo con reemplazamiento a esos bloques en lugar de utilizar los datos originales. Esta es la base del denominado *moving block bootstrap* (MBB) (Kunsch, 1989, Davidson y Hall, 1993). Podemos hacer que los bloques no se solapen (Carlstein, 1986) o bien permitir su solapamiento (Politis y Romano, 1994). Una opción es tomar la longitud de los bloques como fija o bien, tal y como proponen Politis y Romano (1994), suponer que las longitudes de los bloques son variables aleatorias que siguen una distribución geométrica. Esta técnica de bootstrap en la que la longitud de los bloques es aleatoria es lo que se conoce como *stationary bootstrap*, pues garantiza que los datos remuestreados son estacionarios supuesto que lo sean los datos originales, lo que no está garantizado en el caso de utilizar el MBB con longitudes de los bloques fijas (no aleatorias). Lahiri (1999) analiza detalladamente las propiedades de las diferentes variedades de técnicas MBB. Una alternativa muy reciente es el *threshold bootstrap*, propuesto por Park y Willemain (1999).

Cualquiera que sea el tipo de bloque utilizado, se sabe que su longitud (o su longitud media) debe incrementarse con el tamaño muestral. A pesar de que se han propuesto en la Literatura distintas longitudes ‘óptimas’ de los bloques (Carlstein et al., 1998, Buhlmann y Kunsch, 1999), no está nada claro cuál es la longitud (o longitud media) óptima y, por otra parte, ésta depende de cuál sea la finalidad que se dé al bootstrap.

Estas consideraciones son las que hace que se utilice en su lugar la alternativa del *sieve bootstrap* (Buhlmann, 1997, 1998). La base es la búsqueda de un modelo que nos proporcione residuos independientes de forma que podamos utilizar la metodología bootstrap habitual. Dada la naturaleza de nuestro problema, se considerarán los modelos de la familia AR, que son la base de los contrastes aplicados en las secciones anteriores. Un aspecto fundamental es la necesidad de generar los residuos del modelo restringido (Li y Maddala, 1996, Datta, 1996). Así pues, el procedimiento que seguimos es el siguiente:

- (1) Estimamos el modelo *restringido*

$$\Delta x_t^g = c + \sum_{i=1}^p \phi_i \Delta x_{t-i}^g + \nu_t.$$

- (2) Remuestreamos los residuos ν_t^* y construimos x_t^{g*} a partir del modelo

$$\Delta x_t^{g*} = c + \sum_{i=1}^p \phi_i \Delta x_{t-i}^{g*} + \nu_t^*.$$

El tamaño de la serie reconstruida es $m = T^{3/4}$ para solucionar el problema de la discontinuidad funcional (ver Datta, 1996).

- (3) Estimar el modelo

$$\Delta x_t^{g*} = c + \rho x_{t-1}^{g*} + \sum_{i=1}^p \phi_i \Delta x_{t-i}^{g*} + \xi_t,$$

y obtener $\hat{\rho}^*$ y los estadísticos t habituales, t^* . Repetir los pasos (2) – (3) NB veces, donde NB indica el número de remuestreos del bootstrap. el valor crítico del bootstrap se obtiene como el quantil 5% en la cola inferior de la distribución empírica de t^* .

7. Aplicación Empírica

En esta sección se lleva a cabo el estudio de las serie de tipo de cambio real entre las monedas de Finlandia y Estados Unidos basado en el índice de precios al consumo (IPC), ver Figura 1.6, con los procedimientos de contraste expuestos en las Secciones

ADF										
T	$\pi = 0$	$\pi = 0.01$			$\pi = 0.05$			$\pi = 0.1$		
		$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$
100	5.19	2.55	6.45	14.55	2.75	14.35	58.20	3.35	22.80	72.75
200	5.13	2.20	3.95	7.90	3.30	11.50	32.10	3.80	16.45	42.75
500	5.08	2.70	3.80	6.40	2.40	7.35	12.55	3.55	10.35	34.85
1000	5.01	4.45	6.50	9.25	5.45	10.50	19.50	6.70	14.65	30.15
Filtro BK										
100	5.66	2.80	2.60	4.50	2.60	3.75	5.90	2.95	4.50	16.45
200	4.74	3.05	3.05	3.00	3.90	3.15	6.45	3.95	3.05	5.15
500	5.01	2.75	4.00	1.30	2.60	2.35	3.30	3.20	3.20	5.15
1000	4.82	4.85	4.25	4.60	5.60	5.40	6.65	4.85	5.15	8.05
Filtro HP10										
100	9.15	2.55	6.45	14.55	2.75	14.35	58.20	3.35	22.80	72.75
200	7.35	2.20	3.95	7.90	3.30	11.50	32.10	3.80	16.45	42.75
500	4.80	2.70	3.80	6.40	2.40	7.35	12.55	3.55	10.35	34.85
1000	5.30	4.45	6.50	9.25	5.45	10.50	19.50	6.70	14.65	30.15
Filtro HP100										
100	8.20	4.30	6.70	3.70	4.60	3.90	3.35	4.85	4.40	3.15
200	5.80	2.25	2.85	6.15	3.00	5.25	5.50	2.70	5.60	6.60
500	4.85	3.50	3.40	5.15	3.80	4.90	7.30	4.00	6.05	4.40
1000	5.15	5.65	7.65	8.85	5.15	9.15	10.15	4.90	9.65	8.50
Filtro HP400										
100	9.20	4.45	6.45	9.35	4.65	8.05	3.15	4.40	8.50	3.50
200	6.10	2.95	3.75	7.65	3.00	6.75	5.40	2.80	5.75	5.70
500	5.15	2.35	4.95	4.05	2.90	4.45	6.85	2.95	6.75	5.80
1000	5.30	5.20	6.60	9.45	5.55	6.75	10.45	6.65	9.25	8.50
Filtro MD										
100	4.85	5.15	5.15	5.15	5.45	4.80	5.25	4.70	5.70	5.05
200	3.60	5.65	5.65	5.65	5.50	5.80	5.15	5.65	4.40	5.30
500	4.85	6.65	6.65	6.65	6.70	6.35	5.95	7.15	6.25	4.75
1000	5.20	6.45	6.45	6.45	8.90	6.35	6.15	6.75	6.00	5.15

TABLA 1.7. Tamaño de los contrastes de raíces unitarias bootstrap filtrados en presencia de AO's.

anteriores. Los datos son observaciones anuales entre los años 1900-1988. Esta serie fue inicialmente analizada por Perron y Vogelsang (1992) para ilustrar el contraste de raíces unitarias en presencia de una media no constante en un momento conocido. En su análisis, no obstante, no se considera la existencia de observaciones atípicas en la serie. Al realizar el contraste ADF, con el modelo (5.2) con $p = 1$, se obtiene un t ratio de -5.74 (valor crítico al 5% = -2.89). Por tanto, se rechaza la hipótesis de raíz unitaria para esta serie.

Franses y Haldrup (1994) (abreviado FH94) señalaron que la serie muy bien pudiera contener observaciones atípicas, de forma que los resultados mencionados de Perron y Vogelsang (1992) estuviesen sesgados hacia el rechazo de la hipótesis de raíz unitaria. Al utilizar los métodos de detección de atípicos de Chen y Liu (1993)

ADF										
T	$\pi = 0$	$\pi = 0.01$			$\pi = 0.05$			$\pi = 0.1$		
		$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$	$\theta = 1$	$\theta = 6$	$\theta = 16$
100	85.00	87.55	94.95	99.65	90.95	99.95	100.00	92.40	100.00	100.00
200	99.90	100.00	99.95	100.00	99.95	100.00	100.00	100.00	100.00	100.00
Filtro BK										
100	23.55	30.50	38.10	65.15	29.20	49.20	58.55	33.10	69.90	97.35
200	72.05	76.35	79.75	85.25	76.60	87.65	98.30	79.80	89.85	99.95
Filtro HP10										
100	45.65	46.35	43.05	38.10	43.45	44.60	52.50	42.65	46.95	74.35
200	82.80	82.85	87.15	90.90	85.45	95.05	99.65	85.70	96.70	100.00
Filtro HP100										
100	21.07	23.50	33.05	23.40	26.05	23.25	17.40	28.25	29.05	31.85
200	44.88	45.60	55.25	70.40	48.00	73.45	88.80	49.80	78.90	97.35
500	98.66	98.55	98.80	99.80	98.55	99.65	100.00	98.90	100.00	100.00
Filtro MD										
100	28.15	32.35	32.35	32.35	31.80	27.40	27.35	31.35	24.75	24.60
200	87.03	88.60	88.60	88.60	89.35	86.10	86.05	88.40	85.35	84.85

TABLA 1.8. Potencia ajustada de los contrastes de raíces unitarias bootstrap filtrados en presencia de AO's.

con el programa TRAMO (Gómez y Maravall, 1996), FH94 encontraron varias observaciones atípicas en la serie.

El método que sugieren Franses y Haldrup para eliminar la influencia de las observaciones atípicas es utilizar el modelo

$$\Delta x_t = c + \rho x_{t-1} + \sum_{i=1}^p \phi_i \Delta x_{t-i} + \sum_{i=0}^{p+1} \omega_i D_{t-i}^i + \eta_t, \quad (7.1)$$

para contrastar la hipótesis $\rho = 0$ con el t ratio del contraste ADF. D^i indica variables dummy tipo impulso que toman el valor 1 si hay una observación atípica y 0 en caso contrario.

FH94 señalan la presencia de la siguientes observaciones atípicas (las más significativas). 1917-1919, 1932, 1933, 1945, 1949 y 1950. Es decir, el porcentaje empírico de observaciones atípicas es aproximadamente $\hat{\pi} = 0.09$. El t ratio del contraste ADF en el modelo (7.1) con $p = 2$ es -2.65, por lo que ya no se rechaza la hipótesis de raíz unitaria al nivel de 5%.

Por otra parte, Vogelsang (1999) (V99, abreviado) propone un segundo método para detectar AO's que está relacionado con el enfoque de FH94. Su procedimiento se basa en la siguiente ecuación, estimada mediante MCO:

$$x_t = c + \alpha D_t^i + \eta_t. \quad (7.2)$$

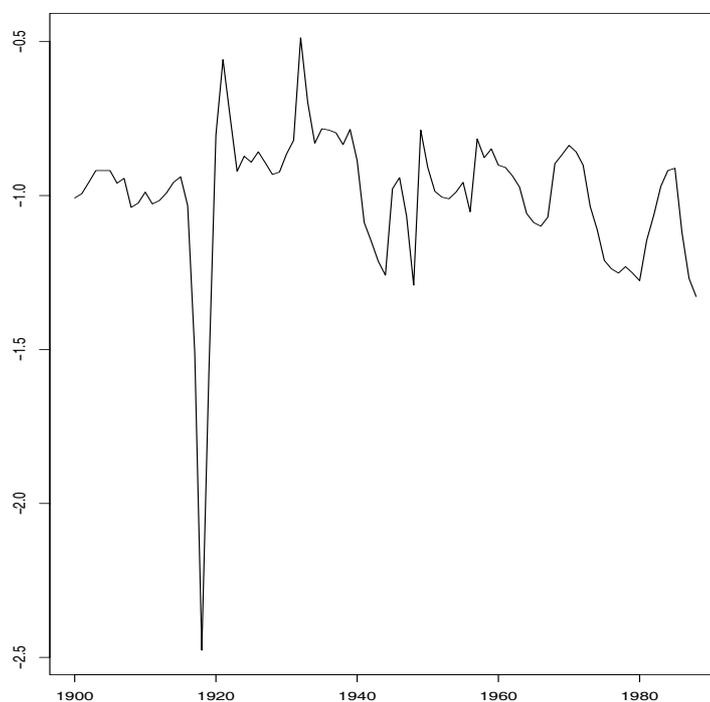


FIGURA 1.6. Logaritmo del tipo de cambio real USA–Finlandia, 1900–88

Sea t_{α}^i es el estadístico t para contrastar $\alpha = 0$ en (7.2). V99 sugiere contrastar la presencia de AO's utilizando $\tau = \sup_{i=1,2,\dots,T} |t_{\alpha}^i|$. La distribución asintótica de τ no es estándar, pero V99 proporciona una tabla de valores asintóticos.

Vogelsang (1999) aplica el estadístico τ de forma iterativa a los datos de tipo de cambio real, detectando los siguientes atípicos: 1917, 1918, 1919, 1921 y 1932. Para eliminar la influencia de los atípicos, utiliza la regresión (7.1) con $p = 0$. El t ratio de ADF es -3.51, por lo que se rechaza la hipótesis de raíz unitaria al 5%, contradiciendo los resultados previos de Franses y Haldrup.

Estas diferencias en las conclusiones de los análisis de la misma serie ilustran claramente lo sensibles que resultan los contrastes de raíces unitarias a la elección de las variables dummy y del número de retardos y mueven a V99 a aplicar los contrastes de Phillips–Perron modificados propuestos por Perron y Ng (1996) a los datos, sin búsqueda previa de observaciones atípicas. Para la serie en cuestión, los estadísticos MZ_{α} y MZ_t , definidos en (5.4)–(5.6), rechazan la hipótesis de raíz unitaria a cualquier nivel de significación razonable.

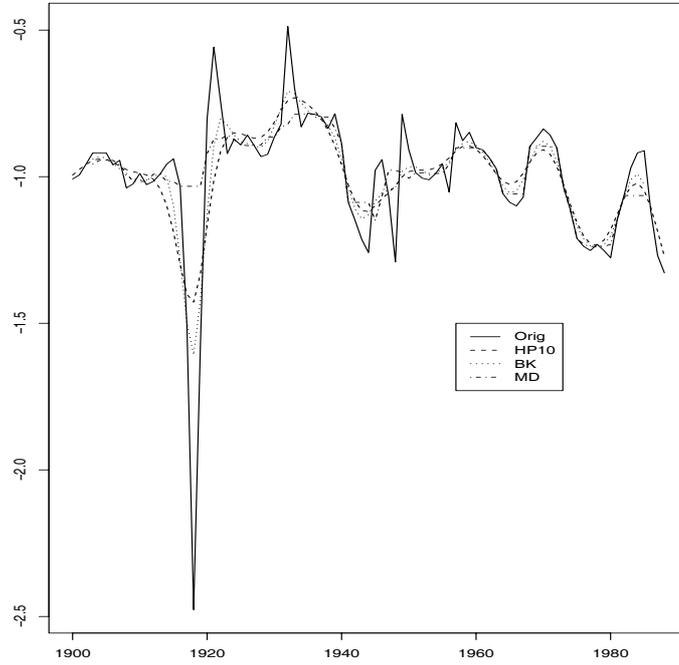


Figura 1.7.1: Tendencia

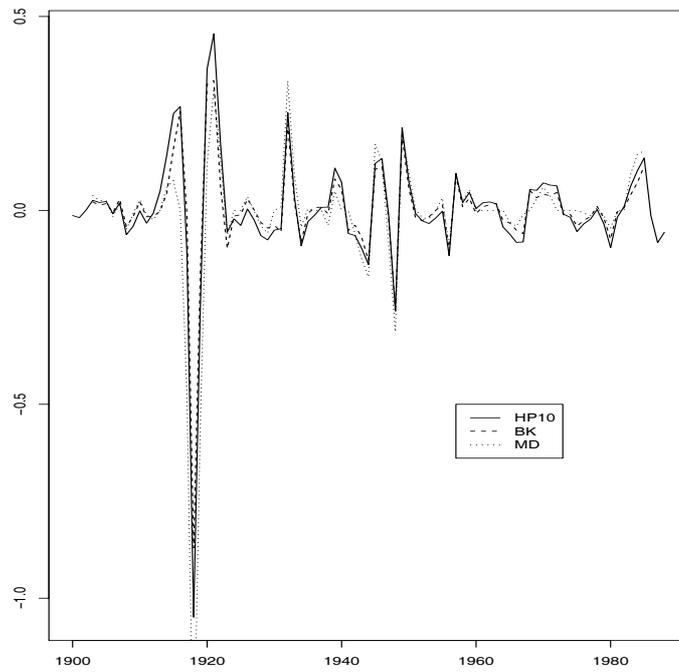


Figura 1.7.2: Ciclo

FIGURA 1.7. Logaritmo del tipo de cambio real USA–Finlandia (IPC). Componentes tendencia y ciclo

Consideremos la aplicación de $t_{\hat{\rho}_{AO}}^g$ en el modelo (5.3) a la serie. A la vista de los resultados de las simulaciones de la Sección 5, se centra el análisis en los filtros BK, HP10 y MD.

	p	$t_{\hat{\rho}_{AO}}^g$	CV_f	CV_f^*	CV_b
HP10	4	-1.4576	-3.2314	-3.1592	-2.8572
MD	1	-2.6381	-2.9376	-2.8983	-2.8924
BK	3	-2.8261	-3.0074	-3.0475	-2.7510

TABLA 1.9. Contrastes ADF filtrados en la serie de logaritmos de tipo de cambio real USA–Finlandia (IPC). CV_f es el valor crítico al 5% de $t_{\hat{\rho}_{AO}}^g$ en el modelo (5.3) suponiendo $\pi = 0$ y tamaño muestral $T = 89$. CV_f^* es el valor crítico al 5% de los tests suponiendo que la serie incluye atípicos en las mismas posiciones y del mismo tamaño que los encontrados por V99. CV_b es el valor crítico al 5% obtenido mediante bootstrap.

En la Tabla 1.9, la columna CV_f indica los valores críticos al 5% de $t_{\hat{\rho}_{AO}}^g$ en el modelo (5.3) suponiendo $\pi = 0$ y tamaño muestral $T = 89$, el número de observaciones de la serie estudiada. Bajo estas condiciones, sólo el filtro BK sufre distorsiones de tamaño, ver Tabla 1.5. CV_f^* es el valor crítico de los contrastes suponiendo que la serie incluye observaciones atípicas en las mismas posiciones y de los mismos tamaños que los indicados en Vogelsang (1999). De la Tabla 1.9 se concluye que con cualquiera de los filtros utilizados el resultado es que no se puede rechazar al nivel del 5% la hipótesis de raíz unitaria, resultado que contradice la conclusión de V99 y concuerda con FH94. No obstante, si nos fijamos más detenidamente, con los filtros BK y MD, los p -valores están alrededor de 0.07 y 0.09 cuando utilizamos los CV_f , y 0.076 y 0.08 si utilizamos CV_f^* , respectivamente.

A la vista de la Tabla 1.4, una posible explicación de las diferencias de las conclusiones puede deberse a la baja potencia del contraste de raíces unitarias cuando se aplica sobre los datos filtrados en muestras pequeñas, por lo que el bootstrap puede proporcionarnos información más adecuada.

Una cuestión abierta es si el contraste bootstrap es válido en una muestra tan pequeña como la que tenemos y con lo shocks que presenta. Para ello realizamos una serie de experimentos que muestran que el contraste bootstrap con el filtro HP100 tiene un tamaño de 5.2%, en tanto que con el filtro BK es 5.8% y 4.05% cuando utilizamos el filtro MD. El tamaño del contraste bootstrap es 8.3% al aplicarlo con el filtro HP10, lo que es excesivamente alto, y superior a 22% cuando no filtramos la serie. Por tanto, tiene sentido aplicar el contraste bootstrap si utilizamos los filtros HP100, BK y MD. En cuanto a la potencia del contraste, tal y como se observa en la

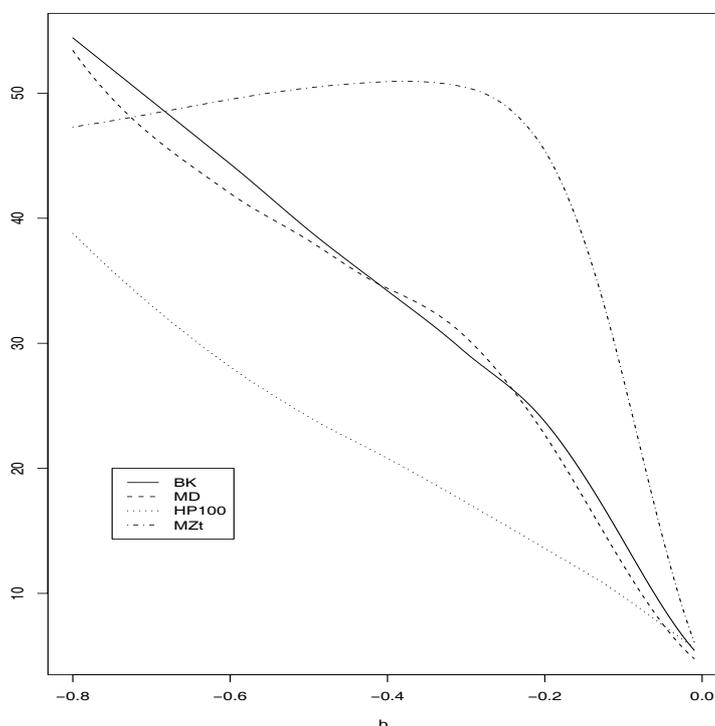


FIGURA 1.8. Potencia del contraste bootstrap en muestras con el mismo tamaño y shocks que el log del tipo de cambio real (basado en IPC) USA—Finlandia.

Figura 1.8, es bastante alta con los filtros BK y MD, aunque no es tan alta como la del contraste MZ_t frente a alternativas locales, como era de esperar.

Los valores críticos basados en el bootstrap ($NB = 10,000$) para la serie de tipo de cambio real USA/Finlandia basado en el IPC con los distintos procedimientos de extracción de tendencia se recogen en la columna CV_b de la Tabla 1.9. Un aspecto interesante es que, como era de esperar, los valores críticos son muy similares a los que obtenemos mediante el experimento de Monte Carlo en la columna CV_f^* . Lo que resulta novedoso es que ahora el contraste ADF sobre el componente tendencial extraído con el filtro BK rechaza la hipótesis nula de raíz unitaria cuando tomamos como nivel de significación 5% (el p -valor es 0.045), lo que es congruente con el resultado de V99. Cuando utilizamos el filtro MD el p -valor del contraste bootstrap 0.0762.

Por último, en la Figura 1.7 se representan los componentes tendenciales de de la serie de logaritmo de tipo de cambio real (basado en el IPC) entre USA y Finlandia con los filtros HP10, Bk y MD. Se observa que los filtros lineales son menos eficientes que el filtro de la mediana para retener la señal de la serie en el sentido de que no

captan toda la intensidad de los bruscos cambios producidos en la tendencia de la serie.

8. Conclusiones

Las series temporales observadas a menudo vienen influidas por elementos que no son intrínsecos a las mismas, como pueden ser huelgas, estallido de guerras, inesperadas crisis políticas o económicas, o incluso errores tipográficos en su registro. Las consecuencias de estos elementos crean observaciones espurias que son inconsistentes con el resto de la serie. Este tipo de observaciones se suele denominar atípicas.

En este Capítulo se han analizado dos métodos alternativos de estudiar si las series tienen raíces unitarias cuando hay observaciones atípicas de tipo aditivo. El primer enfoque consiste en incluir términos dinámicos en la regresión de Dickey–Fuller. Mediante experimentos de Monte Carlo se demuestra que los valores críticos del contraste son más robustos que los obtenidos sin incluir los términos dinámicos adicionales, tal y como sugiere el proceso generador de datos. Sin embargo, la estabilidad (robustez) de los valores críticos es mucho mayor cuando utilizamos el segundo enfoque, consistente en aplicar los contrastes de raíces unitarias en el componente tendencial o permanente de la serie, obtenido mediante algún filtro lineal o no lineal, en lugar de sobre la serie observada. Las simulaciones efectuadas confirman la gran robustez del nuevo método a la presencia de atípicos en términos de tamaño y potencia del contraste. Es especialmente destacable la robustez de filtro de la mediana, de tipo no lineal.

Con el fin de la presencia de raíces unitarias en series que tienen observaciones atípicas de tipo aditivo, los análisis efectuados recomiendan el uso del estadístico MZ_t para alternativas locales, el contraste bootstrap BK–ADF en caso de muestras pequeñas para alternativas globales, y los contrastes BK–ADF y MD–ADF (no necesariamente bootstrap) en el caso de muestras de tamaño moderado o grande para alternativas globales.

Por otra parte, se desprende que el filtro de la mediana puede ser más útil que los filtros lineales para realizar descomposiciones tendencia–ciclo de las series debido a su mayor capacidad de detección de los cambios bruscos de las tendencias estocásticas de las mismas.

9. Demostración del Teorema 1.4.1

Sea $a(B)$ un filtro de tipo pasa-baja de orden n , es decir, un filtro lineal $a(B) = \sum_{j=-n}^n a_j B^j$ con $a(B) = a(B^{-1})$ y $a(1) = 1$. Denotemos mediante $x_t^g = a(B)x_t$ la versión filtrada de x_t en la ecuación (2.2), de modo que $x_t^g = x_{t-1}^g + \xi_t$, $\xi_t = a(B)u_t$ bajo la hipótesis de raíz unitaria, con $u_t = \varepsilon_t + \theta \Delta \delta_t$. Sea $\sigma_u^2 = \lim_{T \rightarrow \infty} T^{-1} E(S_{T,u}^2)$ y $\sigma_\xi^2 = \lim_{T \rightarrow \infty} T^{-1} E(S_{T,\xi}^2)$, donde $S_{t,u} = \sum_{j=1}^t u_j$ y $S_{t,\xi} = \sum_{j=1}^t \xi_j$, $t = 1, 2, \dots, T$. Nótese que u_t y ξ_t satisfacen las condiciones de mixing de Phillips (1987a) para aplicar un teorema funcional del límite a sus sumas parciales, resultando

$$T\hat{\rho}_{AO}^g \Rightarrow \frac{\int_0^1 W(r) dW(r) + \lambda_\xi / \sigma_\xi^2}{\int_0^1 W^2(r) dr} \quad (9.1)$$

y

$$t_{\hat{\rho}_{AO}^g} \Rightarrow \left(\frac{\sigma_\xi}{s_\xi} \right) \frac{\int_0^1 W(r) dW(r) + \lambda_\xi / \sigma_\xi^2}{\left(\int_0^1 W^2(r) dr \right)^{1/2}}, \quad (9.2)$$

donde $\lambda_\xi = 1/2 (\sigma_\xi^2 - s_\xi^2)$ y $s_\xi^2 = \text{var}(\xi_t)$. En Franses y Haldrup (1994) se obtiene que $\sigma_u^2 = \sigma_\varepsilon^2$, lo que, a su vez, implica que $\sigma_\xi^2 = \sigma_u^2$ utilizando el hecho de que $a(1) = 1$. Las varianzas de los procesos u_t y ξ_t son, no obstante, distintas. Franses y Haldrup demuestran que $s_u^2 = \text{var}(u_t) = \sigma_\varepsilon^2 (1 + 2(\theta/\sigma_\varepsilon)^2 \pi)$. Por el contrario,

$$s_\xi^2 = \text{var}(a(B)u_t) = \sum_{j=-n}^n \sum_{i=-n}^n a_j a_i \text{cov}_u(j-i). \quad (9.3)$$

Tal y como se indicó en el texto, u_t es un proceso $MA(1)$ con función de autocovarianzas

$$\text{cov}_u(k) = \begin{cases} \sigma_\varepsilon^2 (1 + 2(\theta/\sigma_\varepsilon)^2 \pi) & k = 0 \\ -\theta^2 \pi & k = 1 \\ 0 & k > 1 \end{cases} \quad (9.4)$$

Por tanto, utilizando (9.4) y las propiedades de $a(B)$, tras algunas manipulaciones algebraicas se puede demostrar que

$$s_\xi^2 = s_u^2 \sum_{j=-n}^n a_j^2 - 2\theta^2 \pi \sum_{j=-n}^n a_j a_{j-1}, \quad (9.5)$$

lo que, a su vez, implica

$$\lambda_\xi/\sigma_\xi^2 = \frac{1}{2} \left[1 - \sum_{j=-n}^n a_j^2 - 2 (\theta/\sigma_\varepsilon)^2 \pi \sum_{j=-n}^n a_j^2 + 2 (\theta/\sigma_\varepsilon)^2 \pi \sum_{j=-n}^n a_j a_{j-1} \right] = \Phi, \quad (9.6)$$

y

$$\left(\frac{\sigma_\xi}{s_\xi} \right)^2 = \left[\sum_{j=-n}^n a_j^2 + 2 (\theta/\sigma_\varepsilon)^2 \pi \sum_{j=-n}^n a_j (a_j - a_{j-1}) \right]^{-1} = \Psi^2. \quad (9.7)$$

El paso final para la demostración del teorema es la sustitución de (9.6) y (9.7) en las expresiones (9.1) y (9.2).

Parte 2

Cointegración

Contrastes de Cointegración con Cambios Estructurales: Un Modelo de Corrección del Error Extendido Robusto

1. Introducción

Las propiedades de los contrastes de cointegración basados en modelos de corrección del error uniecuales (contraste MCE) son bien conocidas. La dependencia de los valores críticos del contraste, así como su potencia, de parámetros molestos aparece reflejada en Banerjee et al. (1986), Engle y Granger (1987), Kremers et al. (1992), Park y Phillips (1988, 1989), y Banerjee et al. (1993).

Desde el punto de vista univariante, los efectos de los cambios estructurales cuando se aplican contrastes de raíces unitarias, como el contraste de Dickey y Fuller (1979), son conocidos. Un buen punto de partida para estudiar dichos efectos es Perron (1989). Siguiendo a Clements y Hendry (1999), podemos definir un cambio estructural como un shock intermitente que tiene un efecto permanente sobre la serie. Si no tomamos en consideración explícitamente dicho shock, los contrastes de raíces unitarias habituales confundirán el cambio estructural con una raíz unitaria. Leybourne et al. (1998) señalan que si el cambio estructural sucede al inicio de la muestra sucederá exactamente lo contrario. Los resultados de Hendry y Neale (1990) y Perron y Vogelsang (1992) también indican que un cambio de tipo escalón en la media no modelizado nos llevará a detectar raíces unitarias espurias. Rappoport y Reichlin (1989) es quizá la primera referencia que analiza los efectos de tener tendencias segmentadas en la hipótesis alternativa del modelo de raíces unitarias, y Andrés et al. (1990) extiende el análisis al caso en que hay más de un punto de ruptura en la tendencia.

El principal problema que nos plantea toda esta Literatura, que ha tenido una rápida expansión, es que su solución pasa por incluir variables de tipo dummy para modelizar los cambios estructurales con el fin de aplicar los contrastes de raíces unitarias. Por tanto, los valores críticos dependen del tamaño y la posición del punto de

ruptura. Nuevamente, ha surgido una vasta literatura con métodos de de detección de puntos de ruptura basada en contrastes recursivos o secuenciales. Ver, por ejemplo, Banerjee et al. (1992), Zivot y Andrews (1992), Andrews (1993), Andrews et al. (1996), Bai (1997), Vogelsang (1997), Bai y Perron (1998), y Banerjee, Lazarova y Urga (1998).

Otra clase de sucesos no habituales la constituyen los atípicos aditivos. Estos son sucesos con efecto grande, pero temporal, sobre las series. En determinados casos, dichos efectos dominan al resto de la información contenida en la serie y sesga la inferencia de raíces unitarias en favor del rechazo de la hipótesis de raíz unitaria incluso si la hipótesis nula de presencia de raíz unitaria es cierta, tal y como señalan Franses y Haldrup (1994) y Lucas (1995*a,b*).

En el caso de series temporales multivariantes, la situación puede ser peor, pues necesitamos tomar una decisión acerca del tipo de modelos que genera las observaciones anomalás (cambios de tendencia, atípicos aditivos, ...) teniendo en cuenta que dichas irregularidades no tienen por qué ocurrir simultáneamente ni en todas las variables. Por tanto, el análisis multivariante es en general más complicado pero, tal y como veremos más adelante, se puede simplificar en algunos casos.

En las aplicaciones empíricas, lo más habitual es incluir variables dummy para obtener modelos MCE con parámetros constantes. Los efectos de la inclusión de variables dummy en modelos MCE para capturar cambios estructurales se analiza en Kremers et al. (1992) y Campos et al. (1996).

El hecho de que los valores críticos (CV) del contraste MCE dependan del tipo particular de variable dummy es extremadamente molesto a la hora de realizar análisis aplicado. No obstante, se puede evitar el uso de variables dummy si utilizamos técnicas de estimación robustas. Este es el enfoque que proponen Lucas (1995*a,b*) en el caso univariante y Franses y Lucas (1997*a,b*) en el marco multivariante.

En este Capítulo seguimos una ruta diferente. El objetivo es encontrar un método de modelización en un contexto MCE que nos permita contrastar cointegración de forma robusta a la presencia de cambios estructurales. En lugar de incluir variables dummy en los modelos MCE, lo que se hace es aproximar dichas rupturas añadiendo términos dinámicos, por medio del criterio de información SBIC, o incluyendo retardos adicionales del mecanismo de corrección del error (modelo MCE extendido). En particular, nos vamos a centrar en el estudio de los valores críticos obtenidos con los modelos sobreparametrizados, la potencia de los mismos y su robustez frente a cambios estructurales por medio de simulaciones y experimentos de Monte Carlo.

La estructura del Capítulo es la siguiente: en la Sección 2 se analizan los efectos de tener elementos deterministas (constante, tendencia determinista, tendencias segmentadas, etc.) en diferentes especificaciones alternativas de los modelos MCE, y, en particular, en los errores de cointegración. Se van a analizar detenidamente tres situaciones determinísticas: co-ruptura simultánea, co-ruptura en niveles (no en diferencias) y co-ruptura en diferencias (no en niveles). Los experimentos de Monte Carlo se presentan en la Sección 3, y los resultados del contraste MCE habitual se analizan detalladamente. En la Sección 4 se introduce el modelo de corrección de error extendido y se muestran sus propiedades mediante el experimento de Monte Carlo. La Sección 5 presenta las conclusiones. Las diferentes interrelaciones que se dan entre los diferentes de co-rupturas se estudian en el Apéndice 6.

2. Modelos de Corrección del Error con y sin Co-ruptura simultánea

Tomemos el siguiente modelo de corrección del error condicional (MCE).

$$\Delta(y_t - \mu_{y,t}) = a\Delta(z_t - \mu_{z,t}) + b[(y_{t-1} - \mu_{y,t-1}) - \alpha(z_{t-1} - \mu_{z,t-1})] + u_{1t} \quad (2.1a)$$

$$\Delta(z_t - \mu_{z,t}) = u_{2t} \quad (2.1b)$$

Supongamos que $\dots, y_{-1}, y_0 = 0$ y $\dots, z_{-1}, z_0 = 0$, sean $\mu_{y,t} = E(y_t)$, $\mu_{z,t} = E(z_t)$ las respectivas medias incondicionales no observadas, basadas en la validez de la parametrización inicial en t_0 . Estas medias incluyen toda la gama de posibles elementos deterministas, tales como: término independiente, tendencias deterministas, variables dummy, tendencias segmentadas, atípicos, etc. Sea B el operador retardo, $B^k y_t = y_{t-k}$, $\Delta = (1 - B)$ el operador primera diferencia, y sea $(1, -\alpha)$ el vector de cointegración. Los errores estocásticos u_{1t} y u_{2t} no presentan ningún tipo de correlación conjunta ni serial, y tienen varianzas constantes $var(u_{1t}) = \sigma_1^2$ y $var(u_{2t}) = \sigma_2^2$. Estas condiciones de los errores pueden relajarse, permitiendo algún tipo de autocorrelación o de dependencia conjunta, siempre que Δz_t sea exógeno en sentido débil para los parámetros del modelo (2.1a). Con el fin de clarificar el análisis, se opta por mantener estas mismas condiciones en el experimento de Monte Carlo.

El modelo (2.1a)–(2.1b) puede expresarse en función de las variables observables, y_t y z_t , como sigue:

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (2.2a)$$

$$\Delta z_t = \Delta \mu_{z,t} + u_{2t} \quad (2.2b)$$

$$c_t \equiv \Delta \mu_{y,t} - a\Delta \mu_{z,t} - b(\mu_{y,t-1} - \alpha \mu_{z,t-1}) \quad (2.2c)$$

Lo que vamos a hacer es investigar los efectos de tener diferentes modelos alternativos para el término independiente c_t en el contraste de no-cointegración ($b = 0$) MCE de (2.2a).

Los errores de cointegración en función de las variables observables se obtienen a partir de (2.2a)–(2.2c)

$$\begin{aligned} y_t - \alpha z_t &= \frac{1}{1 - (b+1)B} c_t + \frac{a - \alpha}{1 - (b+1)B} \Delta \mu_{z,t} \\ &+ \frac{a - \alpha}{1 - (b+1)B} u_{2,t} + \frac{1}{1 - (b+1)B} u_{1,t} \end{aligned} \quad (2.3)$$

y sería necesario incluir toda una serie de regresores determinísticos que vienen dados a partir del término de la derecha de la ecuación (2.3), c_t y $\Delta \mu_{z,t}$. Así pues, la regresión de cointegración basada en las variables observables requiere el siguiente conjunto de regresores

$$y_t = \alpha z_t + \frac{1}{1 - (b+1)B} c_t + \frac{a - \alpha}{1 - (b+1)B} \Delta \mu_{z,t} + v_t. \quad (2.4)$$

Obsérvese que el problema de tener que añadir regresores deterministas de forma arbitraria y con el impacto mencionado sobre el contraste, se reduce, pero no se elimina, al condicionar respecto a $\Delta z_t, \Delta z_{t-1} \dots$

$$y_t = \alpha z_t + \frac{a - \alpha}{1 - (b+1)B} \Delta z_t + \frac{1}{1 - (b+1)B} c_t + w_t \quad (2.5)$$

siendo $w_t = \frac{1}{1 - (b+1)B} u_{1,t}$, y siempre tendremos que incluir los términos que representan los efectos deterministas dinámicos de c_t .

DEFINICIÓN 2.1. Supuestas condiciones iniciales válidas, sean $E(y_t) = \mu_{y,t}$ y $E(z_t) = \mu_{z,t}$, las series y_t y z_t presentan *co-rupturas en diferencias* si $\Delta \mu_{y,t} - a\Delta \mu_{z,t} = c_d$, donde c_d es un parámetro finito constante.

DEFINICIÓN 2.2. Supuestas condiciones iniciales válidas, sean $E(y_t) = \mu_{y,t}$ y $E(z_t) = \mu_{z,t}$, las series y_t y z_t presentan *co-rupturas en niveles* si $\mu_{y,t} - \alpha \mu_{z,t} = c_l$, siendo c_l un parámetro finito constante.

Hay una serie de casos intermedios (co-rupturas parciales) que son de interés en muchas aplicaciones empíricas y se consideran en los experimentos de Monte Carlo.

CASO 2.1. *Co-ruptura en diferencias, pero no en niveles*

Co-break en diferencias: $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$ implica que $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = (a - \alpha)\Delta\mu_{z,t} + c_d$. A partir de sustituciones recursivas $\mu_{y,t} - \alpha\mu_{z,t} = (\mu_{y_0} - \alpha\mu_{z_0}) + c_d t + (a - \alpha)\mu_{z,t}$, y c_t es

$$c_t = c_d - b(\mu_{y,0} - \alpha\mu_{z,0}) - bc_d(t - 1) - b(a - \alpha)\mu_{z,t-1} \quad (2.6)$$

y la ecuación (2.2a) se puede escribir

$$\begin{aligned} \Delta y_t = & c_m + bc_d t - b(a - \alpha)\mu_{z,t-1} \\ & + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \end{aligned} \quad (2.7)$$

donde c_m es una constante igual a $c_m = c_d - b(\mu_{y,0} - \alpha\mu_{z,0}) + bc_d$.

Observación: Suponiendo que $\mu_{y,0} - \alpha\mu_{z,0} = \text{constante}$, co-ruptura en diferencias \Rightarrow co-ruptura en niveles si $a = \alpha$ (COMFAC) y $c_d = 0$.

CASO 2.2. *Co-ruptura en niveles pero no en diferencias.*

Co-ruptura en niveles: $\mu_{y,t} - \alpha\mu_{z,t} = c_l$. Tomando primeras diferencias, se tiene $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = 0$. Pero de la ecuación (2.2c) se obtiene

$$c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - bc_l = (\alpha - a)\Delta\mu_{z,t} - bc_l, \quad (2.8)$$

y la expresión (2.2a) toma la forma

$$\Delta y_t = -bc_l + (\alpha - a)\Delta\mu_{z,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \quad (2.9)$$

Observación: Co-ruptura en niveles \Rightarrow co-ruptura en diferencias si $a = \alpha$ (restricción COMFAC).

DEFINICIÓN 2.3. Supuestas condiciones iniciales válidas, sean $E(y_t) = \mu_{y,t}$ y $E(z_t) = \mu_{z,t}$, las series y_t y z_t presentan *co-ruptura simultánea* si $\Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t} - \alpha\mu_{z,t}) = c_s$, donde c_s es un parámetro finito constante.

Está claro que cuando las series y_t y z_t presentan *co-ruptura en niveles y en diferencias (co-ruptura completa)*, es un caso particular de co-ruptura simultánea. En el caso de de co-ruptura simultánea, el término independiente de (2.2c) c_t , es una constante, $c_t = c$ y el modelo de corrección del error de (2.2a) es el modelo MCE condicional habitual en el que el único regresor determinista es el término constante, c .

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (2.10)$$

Por otra parte, incluso si $c_t = c$, la regresión de cointegración (2.4) tiene un término constante, $\tilde{c} = \frac{1}{1-(b+1)B}c$, así como retardos de $\Delta\mu_{z,t}$

$$y_t = \tilde{c} + az_t + \frac{a - \alpha}{1 - (b + 1)B} \Delta\mu_{z,t} + v_t. \quad (2.11)$$

Por tanto, sería necesario incluir una estructura dinámica muy complicada de las variables dummy en el caso en que $\Delta\mu_{z,t}$ tenga cambios estructurales. Este problema sí se resuelve condicionando en $\Delta z_t, \Delta z_{t-1}, \dots$, ya que (2.11) pasa a ser

$$y_t = \tilde{c} + az_t + \frac{a - \alpha}{1 - (b + 1)B} \Delta z_t + w_t. \quad (2.12)$$

Esta regresión se simplifica si se verifica una restricción de factor común, $a = \alpha$ (restricción COMFAC), puesto que entonces $\Delta\mu_{z,t}$ no tiene efecto alguno sobre la regresión de cointegración (2.11). A partir de la ecuación (2.4), es evidente que el hecho de que se cumpla $a = \alpha$ (restricción COMFAC) no es una solución universal, ya que la regresión de cointegración es de la forma

$$y_t = \alpha z_t + \frac{1}{1 - (b + 1)B} c_t + u_t \quad (2.13)$$

y tenemos una extraña regresión de cointegración con una estructura muy compleja a través de los retardos de los elementos deterministas de c_t .

En general, si no tenemos ninguna co-ruptura parcial en niveles o en diferencias, la representación más parsimoniosa es el modelo MCE condicional (2.1a), y en términos de las variables observables es la representación (2.2a), dado que sólo es preciso añadir los regresores deterministas procedentes de los valores contemporáneos de c_t . Evidentemente, si estamos interesados en estimar los parámetros a , α y b , es más fácil y más parsimonioso estimarlas por procedimientos unietápicos (MCO o MC-NL) en las representaciones MCE (2.1a) o (2.2a) que en cualquiera de las restantes representaciones mencionadas. No obstante, para hacerlo en (2.1a) se necesita conocer o estimar previamente las medias incondicionales $\mu_{y,t}$ y $\mu_{z,t}$, y ésto generalmente conlleva la incorporación arbitraria de información sobre sucesos desconocidos.

2.1. Modelos de Corrección del Error con Co-ruptura Simultánea. A partir de las ecuaciones (2.2a)–(2.2c) y de los análisis de Escribano (1987a) y Andrés et al. (1990), es evidente que cualquier modelo de corrección del error en término de las variables observables debe tener en cuenta los efectos conjuntos de los siguientes elementos: $\Delta\mu_{y,t}$, $\Delta\mu_{z,t}$, $\mu_{y,t-1}$ y $\mu_{z,t-1}$.

Los modelos de corrección del error con co-rupturas ya se han estudiado en parte en los trabajos de Campos et al. (1996) y Clements y Hendry (1999). En esta sección

se consideran tendencias deterministas en y_t y z_t que presentan *co-ruptura simultánea* (ver Definición 2.3), con $c_s = 0$. Las tendencias segmentadas en z_t se generan mediante $\Delta\mu_{z,t} = sD_{j,t}$, donde s es un parámetro que mide el tamaño del salto, y $D_{j,t}$ es una variable dummy que toma el valor 0 antes de que ocurra la ruptura, y el valor 1 en el momento del salto y después del mismo, ver Sección 3.1 para más detalles. En este caso, (2.2a) y (2.2b) se pueden simplificar en

$$\Delta y_t = a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (2.14a)$$

$$\Delta z_t = \Delta\mu_{z,t} + u_{2,t} \quad (2.14b)$$

$$\Delta\mu_{z,t} = sD_{j,t} \quad (2.14c)$$

donde (2.14a) tiene la forma del modelo de corrección del error uniecuacional habitual sin constante, puesto que $c_t = 0$.

Tomemos el PGD dado por las ecuaciones (2.14a) y (2.14b) con $\Delta\mu_{z,t} = 0$. La distribución del t -ratio del parámetro b bajo la hipótesis nula de que $b = 0$ (no cointegración) se analiza en Banerjee et al. (1986), Kremers et al. (1992), Park y Phillips (1988, 1989) y Banerjee et al. (1993).

Suponiendo que α es conocido e igual a 1, $\alpha = 1$ se podrían estimar consistentemente los parámetros de la ecuación (2.14a) mediante MCO

$$\Delta y_t = \hat{a}\Delta z_t + \hat{b}(y_{t-1} - z_{t-1}) + \hat{u}_{1,t} \quad (2.15)$$

y obtener la distribución asintótica del t -ratio de \hat{b} bajo $H_0 : b = 0$. $t_{\hat{b}} = \hat{b}/sdev(\hat{b})$ es el estadístico del contraste MCE y

$$t_{\hat{b}} \Rightarrow \frac{(a-1) \int B_{u_2} dB_{u_1} + r^{-1} \int B_{u_1} dB_{u_1}}{\sqrt{(a-1)^2 \int B_{u_2}^2 + 2(a-1)r^{-1} \int B_{u_2} B_{u_1} + r^{-2} \int B_{u_1}^2}} \quad (2.16)$$

donde ' \Rightarrow ' indica convergencia en sentido débil, B_{u_1} y B_{u_2} son movimientos Brownianos independientes y $r = \sigma_2/\sigma_1$. En términos del 'ratio señal-ruido', $q = -(a-1)r = \frac{-(a-1)\sigma_2}{\sigma_1}$,

$$t_{\hat{b}} \Rightarrow \frac{\int B_{u_2} dB_{u_1} + \frac{1}{q} \int B_{u_1} dB_{u_1}}{\sqrt{\int B_{u_2}^2 + 2\left(\frac{1}{q}\right) \int B_{u_2} B_{u_1} + \left(\frac{1}{q}\right)^2 \int B_{u_1}^2}} \quad (2.17)$$

Nótese que cuando $q = 0$ (o $a = 1$, restricción COMFAC) (2.16) se reduce a

$$t_{\hat{b}} \Rightarrow \frac{\int B_{u_1} dB_{u_1}}{\sqrt{\int B_{u_1}^2}} \equiv DF \quad (2.18)$$

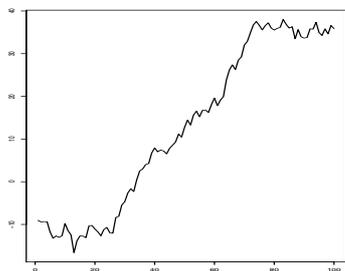
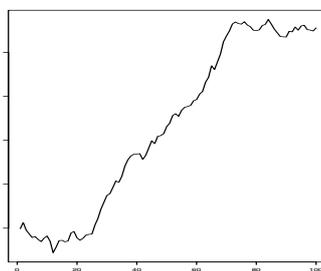
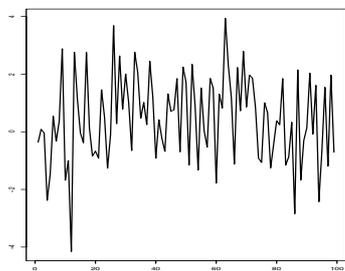
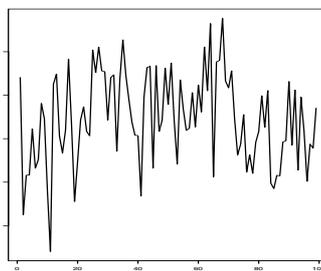
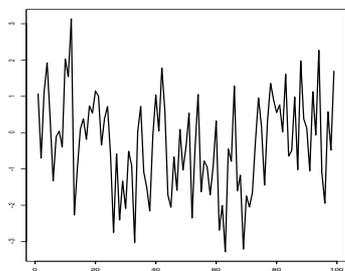
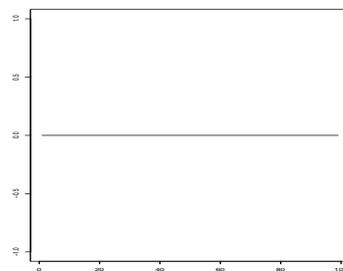
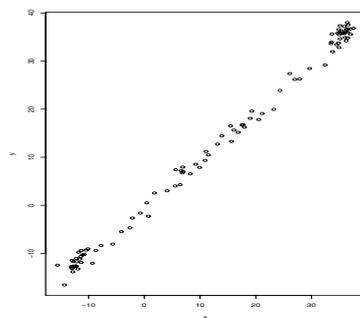
Figura 2.1.1: y Figura 2.1.2: z Figura 2.1.3: Δy Figura 2.1.4: Δz Figura 2.1.5: $y_{t-1} - z_{t-1}$ Figura 2.1.6: c_t 

Figura 2.1.7: Cross-plot

FIGURA 2.1. Co-ruptura simultánea, $T = 100$, $a = 0$, $s = 1$, $b = -0.75$

donde DF es la distribución Dickey–Fuller, (ver Dickey y Fuller, 1979), del t -ratio de \hat{b} de la regresión MCO (2.20), que es el contraste DF de no-cointegración de Engle y Granger (1987) cuando el vector de cointegración es conocido. A partir de (2.17) ya para q grande

$$t_{\hat{b}} \Rightarrow \frac{\int B_{u_2} dB_{u_1}}{\sqrt{\int B_{u_2}^2}} + O_p(q^{-1}) \quad (2.19)$$

Dado que B_{u_1} y B_{u_2} son movimientos Brownianos independendientes, el término principal del lado derecho de la ecuación sigue una distribución Normal estándar (Park y Phillips, 1988).

Cuando $a = \alpha$, y α es conocido, $\alpha = 1$, la ecuación (2.14a) puede expresarse como

$$\Delta(y_t - z_t) = b(y_{t-1} - z_{t-1}) + u_{1t} \quad (2.20)$$

que es la ecuación habitual de Dickey–Fuller. Si se estima la ecuación no restringida (2.14a) con $\alpha = 1$, entonces el t -ratio $t_{\hat{b}} \rightarrow$ distribución DF , ver ecuación (2.13).

Kremers et al. (1992) obtuvo la distribución del t -ratio del parámetro b en (2.15) bajo una hipótesis alternativa local, $b = \frac{h}{T}$, $h < 0$, cuando $\Delta\mu_{z_t} = 0$, siguiendo a Phillips (1987b),

$$t_{\hat{b}} \Rightarrow h(1 + q^2)^{1/2} \left(\int K_e^2 \right)^{1/2} + \frac{(a-1) \int K_{u_2} dB_{u_1} + r^{-1} \int K_{u_1} dB_{u_1}}{\sqrt{(a-1)^2 \int K_{u_2}^2 + 2(a-1)r^{-1} \int K_{u_2} K_{u_1} + r^{-2} \int K_{u_1}^2}} \quad (2.21)$$

donde $e_t = (a-1)\Delta z_t + u_{1t}$ y K_e es un proceso de difusión. Obsérvese que para $h = 0$, $K_i = B_i$ (un movimiento Browniano), y (2.21) coincide con (2.16). Por tanto, la potencia de $t_{\hat{b}}$ debe aumentar con h para un T dado, o aumentar con el valor absoluto de b , dado que la distribución de $t_{\hat{b}}$ bajo H_1 se desplaza a la izquierda de la de $t_{\hat{b}}$ bajo H_0 .

Cuando $a = 1$, restricción COMFAC, el t -ratio de b en (2.15)

$$t_{\hat{b}} \Rightarrow c \left(\int K_e^2 \right)^{1/2} + \frac{\int K_{u_1} dB_{u_1}}{\sqrt{\int K_{u_1}^2}} \quad (2.22)$$

que es la distribución del estadístico DF bajo la alternativa local.

Para $a \neq 1$ y q grande, (2.21) es aproximadamente Normal condicionado a u_{2t} ,

$$t_{\hat{b}} \Rightarrow N \left[\left(c(1 + q^2) \right)^{1/2} \left(\int K_{u_2}^2 \right)^{1/2}, 1 \right] + O_p(q^{-1}) \quad (2.23)$$

La media incondicional de $t_{\hat{b}}$ es aproximadamente $E(t_{\hat{b}}) \approx c(1+q)^2)^{1/2} \frac{1}{\sqrt{2}}$. Por tanto, al aumentar q , dado que c es negativo, la distribución de $t_{\hat{b}}$ bajo la alternativa local puede desplazarse a la izquierda de forma arbitraria y, por tanto, la potencia de contraste puede acercarse de forma arbitraria a 100%.

No obstante, en muestras pequeñas la potencia de $t_{\hat{b}}$ en (2.14a) con $\alpha = 1$ puede ser menor que la potencia de $t_{\hat{b}}$ en (2.20), ya que la estimación del modelo no restringido (2.14a) es menos eficiente, dando lugar a t -ratios menores que los asintóticos en valor absoluto.

A partir de las ecuaciones (2.14a)–(2.14c), es evidente que $y_t \sim I(1)$, y $z_t \sim I(1)$ con tendencias segmentadas, y están cointegradas con vector de cointegración igual a $(1, -\alpha)$. Además, la tendencia segmentada en la variable ‘exógena’, z_t , tiene co-ruptura simultánea ($c_s = 0$) con la variable endógena y_t , y, por tanto, la relación de cointegración puede escribirse de forma explícita como en la ecuación (2.5) con $c_t = 0$. Así pues, a partir del modelo (2.14a)–(2.14c), está claro que si $-2 < b < 0^1$ entonces las variables están cointegradas, es decir, $(y_t - \alpha z_t) \sim I(0)$, pero con errores no estacionarios debido a los cambios estructurales de Δz_t , y si $b = 0$, las variables no están cointegradas, $(y_t - \alpha z_t) \sim I(1)$ con tendencias segmentadas.

Substituyendo las ecuaciones (2.14b) y (2.14c) en (2.3) bajo co-ruptura simultánea con $c_t = 0$, se obtiene una expresión muy interesante de los errores de cointegración,

$$y_t - \alpha z_t = \frac{a - \alpha}{1 - (b + 1)B} sD_{j,t} + \frac{a - \alpha}{1 - (b + 1)B} u_{2t} + \frac{1}{1 - (b + 1)B} u_{1t}. \quad (2.24)$$

Es obvio que para estimar el parámetro de cointegración α en (2.24) es preciso incluir retardos de las variables dummy $D_{j,t}$ a no ser que se satisfaga la restricción de factor común, $a = \alpha$. Por tanto, en general es mejor estimar (2.5) sin utilizar ninguna variable dummy. Estas conclusiones son válidas para la mayoría de los modelos de regresión estáticos que utilicen procedimientos de estimación paramétricos o no paramétricos para estimar α en la regresión de cointegración $y_t = \alpha z_t + \epsilon_t$, tales como MCO (Engle y Granger, 1987), FM-OLS (Phillips y Hansen, 1990) o cointegración canónica (Park, 1992). La solución es sencilla cuando $\Delta \mu_{z,t}$ es una constante o una tendencia determinista, pero no es tan sencilla, ver ecuación (2.24) cuando $\Delta \mu_{z,t}$ presenta cambios de nivel, tendencias segmentadas u otro tipo de cambios estructurales que ocurren en variables económicas.

¹No es infrecuente que la condición de cointegración sea $-1 < b < 0$. Ver, por ejemplo, Kremers et al. (1992) y Campos et al. (1996).

En el caso de co-ruptura simultánea, el modelo dinámico condicional (2.14a) presenta una ventaja evidente, puesto que el vector de cointegración puede estimarse de forma consistente mediante MCO sin tener que utilizar ninguna variable explicativa determinista.

En resumen, los modelos dinámicos condicionales de corrección del error basados en variables económicas que presentan co-ruptura simultánea no precisan el uso de variables dummy. Por otra parte, las regresiones estáticas sí necesitan incorporar explícitamente variables dummy cuando el parámetro de correlación contemporánea de corto plazo difiere del vector de cointegración $a \neq \alpha$ (no COMFAC). Nótese que el caso $a \neq \alpha$ es el más habitual en aplicaciones empíricas.

2.2. Modelos de Corrección del Error sin Co-ruptura Simultánea. En la sección anterior se han discutido las formulaciones del MCE en presencia de co-ruptura simultánea. Ahora se discutirán distintos casos alternativos de interés.

CASO 2.1. *Co-ruptura en diferencias, pero no co-ruptura en niveles.*

A partir de (2.7), (2.14b) y (2.14c), se tiene que

$$\begin{aligned} \Delta y_t &= c_m + bc_d t - b(a - \alpha)\mu_{z,t-1} \\ &+ a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \end{aligned} \quad (2.25a)$$

$$\mu_{z,t} = \mu_{z,0} + s \sum_{i=1}^t D_{j,i}. \quad (2.25b)$$

$$c_m = c_d - b(\mu_{y0} - \alpha\mu_{z0}) + bc_d \quad (2.25c)$$

Por tanto, dependiendo del tipo de variable dummy, $D_{j,t}$, podemos tener tenencias segmentadas con uno o más puntos de ruptura en $\mu_{z,t}$ (Figura 2.2), ver Sección. Bajo $H_0 : b = 0$, la ecuación (2.25a) se convierte en

$$\Delta y_t = c_d + a\Delta z_t + u_{1,t}. \quad (2.26)$$

CASO 2.2. *Co-en niveles, pero no co-ruptura en diferencias.*

A partir de las ecuaciones (2.9), (2.14b) y (2.14c), se tiene

$$\Delta y_t = -bc_l + (\alpha - a)sD_{j,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t}. \quad (2.27)$$

Por tanto, los cambios estructurales en el proceso marginal de Δz_t , (2.14b), afecta al modelo de corrección del error a no ser que se satisfaga la restricción COMFAC ($a = \alpha$). Bajo $H_0 : b = 0$, la ecuación (2.27) es

$$\Delta y_t = (\alpha - a)sD_{j,t} + a\Delta z_t + u_{1,t}. \quad (2.28)$$

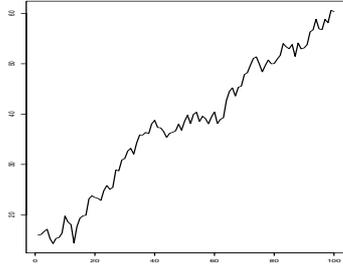
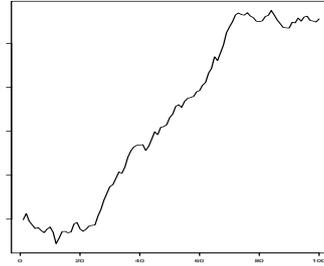
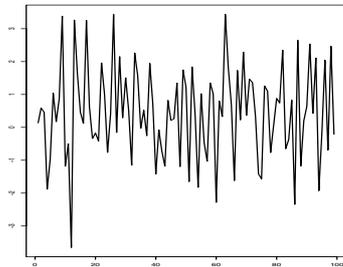
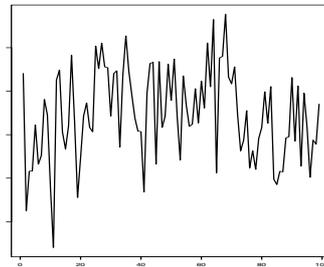
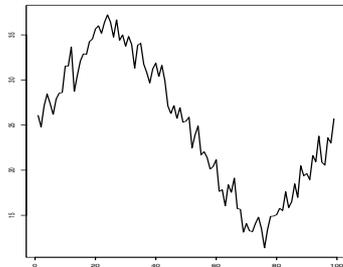
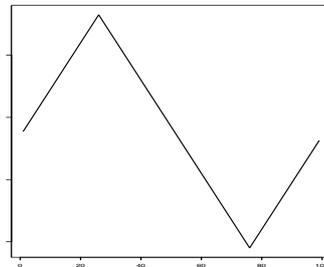
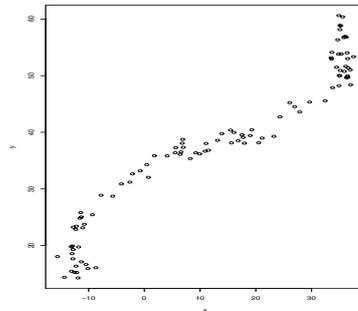
Figura 2.2.1: y Figura 2.2.2: z Figura 2.2.3: Δy Figura 2.2.4: Δz Figura 2.2.5: $y_{t-1} - z_{t-1}$ Figura 2.2.6: c_t 

Figura 2.2.7: Cross-plot

FIGURA 2.2. Co-ruptura en diferencias, no en niveles, $T = 100$, $a = 0$, $c_d = 0.5$, $s = 1$, $b = -0.75$

CASO 2.3. *No co-ruptura en diferencias ni en niveles.* Esta situación empírica, muy probable, es el resultado de unir las ecuaciones (2.27) y (2.25a)–(2.25c), siendo una caso particular de la ecuación general (2.2a).

No obstante, para el propósito de la próxima Sección, es suficiente demostrar que para los casos (2.27) y (2.25a) las situaciones con las que hay que trabajar en la práctica se complican en la presencia de cambios estructurales. En particular, vamos a analizar el impacto de tener diferentes cambios estructurales en términos de los valores críticos (CV's), el tamaño y la potencia del contraste MCE para $b = 0$ obtenido a partir de (2.10).

3. Experimento de simulación de Monte Carlo

El proceso generador de datos (PGD) se basa en diversas extensiones de los utilizados por Kremers et al. (1992) y Campos et al. (1996). Se trata de un proceso vectorial lineal auto-regresivo de primer orden con perturbaciones Normales causalidad en sentido de Granger unidireccional ($z \rightarrow y$), y cambios estructurales en las variables fuertmente exógenas (Δz_t) para los parámetros de interés , a , b , y α .

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (3.1a)$$

$$\Delta z_t = \Delta \mu_{z,t} + u_{2t} \quad (3.1b)$$

$$c_t = \Delta \mu_{y,t} - a\Delta \mu_{z,t} - b(\mu_{y,t-1} - \alpha \mu_{z,t-1}) \quad (3.1c)$$

$$\Delta \mu_{z,t} = sD_{j,t} \quad (3.1d)$$

donde

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim IIN \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right). \quad (3.1e)$$

Permitimos tres tipos de variables dummy ($D_{j,t}$, $j = 1, 2, 3$) para simular tanto un *salto único* en la tendencia determinista (tenencias segmentadas), en dos *puntos de ruptura* diferentes, ($T/4$ o $T/2$, donde T es el tamaño muestral),

$$D_{1t} = \begin{cases} 1 & t \geq T/4 \\ 0 & \text{caso contrario} \end{cases} \quad D_{2t} = \begin{cases} 1 & t \geq T/2 \\ 0 & \text{caso contrario} \end{cases}$$

y un *salto doble* en los puntos $T/4$ y $3T/4$,

$$D_{3t} = \begin{cases} 1 & T/4 \leq t \leq 3T/4 \\ 0 & \text{caso contrario} \end{cases}.$$

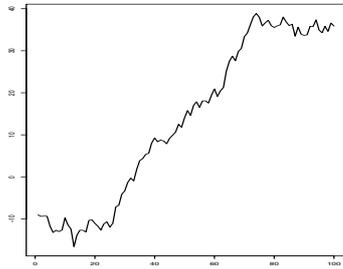
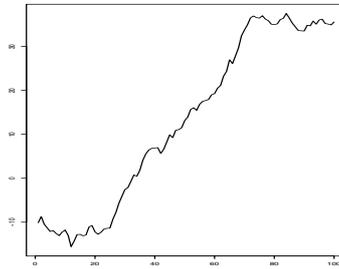
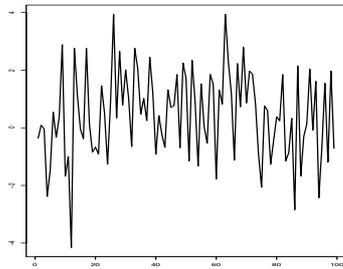
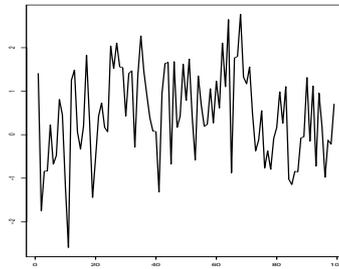
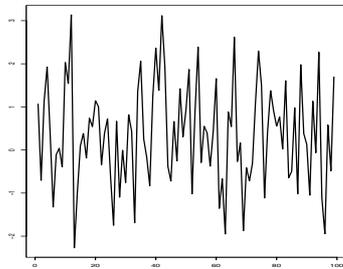
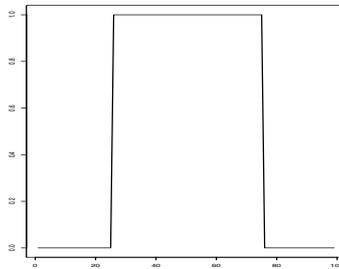
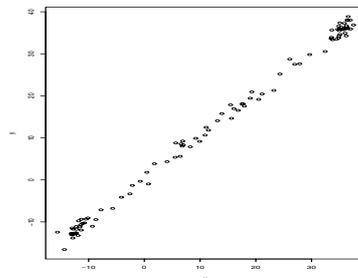
Figura 2.3.1: y Figura 2.3.2: z Figura 2.3.3: Δy Figura 2.3.4: Δz Figura 2.3.5: $y_{t-1} - z_{t-1}$ Figura 2.3.6: c_t 

Figura 2.3.7: Cross-plot

FIGURA 2.3. Co-ruptura en niveles, no en diferencias, $T = 100$, $c_t = 0$, $a = 0$, $s = 1$, $b = -0.75$

Sin pérdida de generalidad, hacemos $\sigma_1^2 = 1$, y $\alpha = 1$. por tanto, las variables del experimento diseñado son los parámetros a, b, s donde $\sigma_2 = s$, y el tamaño muestral, T .

El experimento es un diseño factorial completo con:

$a = 0.0, 0.5, 1$ (correlación contemporánea en primeras diferencias)

$b = 0.0$ (no cointegración), $-0.05, -0.1, -0.25, -0.5, -0.75$ (cointegración)

$s = 1, 6, 16$ (tamaño de los saltos)

$T = 25, 50, 100, 200, 500, 1000$ (tamaño muestral)²

y contemplando las posibilidades de que no haya saltos (NO), un salto único en $T/4$ o $T/2$ (D_1 and D_2) o un salto doble en $T/4$ y $3T/4$ (D_3), donde todos los saltos son en la pendiente de $\mu_{z,t}$ de tamaño s . Esto representa 216 experimentos para cada valor de b . Recordemos que cuando $a = 1$ se cumple la restricción de factor común (COMFAC) en el modelo de corrección del error, puesto que $\alpha = 1$.

Los experimentos de Monte Carlo se basan en 10000 réplicas de cada experimento eliminando las primeras 50 observaciones de cada serie simulada para considerar condiciones iniciales aleatorias.

Para obtener los valores críticos empíricos, se generan las series y_t y z_t a partir del PGD (3.1a)–(3.1e) bajo $H_0 : b = 0$ y se estiman las ecuaciones

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (3.2)$$

$$\phi(B)\Delta y_t = c + \theta(B)\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (3.3)$$

donde se ha impuesto $\alpha = 1$. Los órdenes de los polinomios $\phi(B)$ y $\theta(B)$ se eligen con el criterio SBIC³. El valor crítico considerado es el cuantil 5% de la cola inferior de la distribución empírica del estadístico t -ratio, $t(\hat{b})$, bajo H_0 . El tamaño empírico del contraste se analiza añadiendo un $MA(1)$ a los errores $u_{1,t}$, es decir, $u_{1,t} + \theta u_{1,t-1} = v_t$ con $(\theta = \pm 0.5)$. La potencia empírica se calcula de forma análoga, simulando el PGD (3.1a)–(3.1e) bajo $H_1 : b \neq 0$, y calculando el porcentaje de rechazos obtenido de los modelos (3.2) y (3.3) utilizando los valores críticos empíricos calculados previamente.

Se impone co-ruptura simultánea haciendo $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t} - \alpha\mu_{z,t}) = 0$. Para obtener co-ruptura sólo en diferencias, se toma $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d = 0.5$. Por otra parte, para simular series con co-rupturas sólo en niveles, imponemos $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = 0$, ver Apéndice 6 para más detalles.

²La totalidad de las tablas se encuentra en Arranz y Escribano (1998).

³Para elegir el número de retardos de cada variable en la regresión se permiten entre 0 y 10 retardos de cada variable, y se exploran las 121 posibles combinaciones, excepto en el caso $T = 25$, que no se permiten más de 4 retardos de cada variable.

Hay que resaltar que si hubiésemos tomado $c_d = 0$, los valores críticos obtenidos serían los mismos que para el caso de co-ruptura simultánea, puesto que bajo $H_0 : b = 0$ se tiene co-ruptura simultánea con $c_s = 0$. Además, para $a = \alpha = 1$ la co-ruptura en diferencias implicaría co-ruptura en niveles (co-ruptura completa). Cuando se satisface la restricción COMFAC, co-ruptura en niveles implica co-ruptura en diferencias, ver Apéndice 6.

3.1. Experimento de Monte Carlo: MCE con Co-ruptura Simultánea.

Valores críticos del contraste MCE con co-ruptura simultánea. Los valores críticos al 5% de la cola inferior de la distribución empírica del t -ratio de \hat{b} aparecen recogidos en la Tabla 2.1.1. La Tabla 2.1.1 se genera haciendo σ_2^2 , ver ecuación (3.1e), igual al tamaño del salto (s) para $s = 1, 6, 16$. También se fija la varianza de $u_{2,t}$ ($\sigma_2^2 = 1$) y se modifica $s = 1, 6, y 16$ para hacer el salto en Δz_t más pronunciado, y se obtienen los CV empíricos, que son muy parecidos a los anteriores⁴.

Hay una serie de aspectos que conviene resaltar:

OBSERVACIÓN 3.1. Cuando no hay COMFAC ($a \neq 1$), la distribución del t -ratio ($t_{\hat{b}}$) se desplaza a la derecha con el tamaño del salto. Por tanto, cuanto mayor es el salto en Δz_t , mayor es la probabilidad de aceptar en exceso la hipótesis de raíz unitaria con los CV habituales del contraste de no-cointegración (demasiadas raíces unitarias en los errores de cointegración). Sin embargo, esos saltos en la distribución no son muy pronunciados, puesto que los valores críticos son muy similares para los diferentes tipos de saltos ($D_1, D_2, y D_3$). Por ejemplo, para $T = 100$ y $a = 0.5$ los CV's al 5% están entre -1.7 ($s = 16$) y -2.3 ($s = 1$). Para muestras más grandes ($T = 1000$), los CV's están de forma estable en torno a -1.7. Por tanto, aumentar s es como aumentar q en (2.17) y $t_{\hat{b}}$ se aproxima a la distribución Normal estándar (2.19), ver Figura 2.5.1.

El mayor impacto de la distribución empírica se produce al cambiar el parámetro de corto plazo a ($a=0, 0.5, 1$), ver Figura 2.4.1. Esto no debe sorprendernos, pues es obvio a partir de las ecuaciones (2.16)–(2.18) en las que la distribución límite viene dada por los diferentes parámetros del modelo. La ecuación (2.16) hace explícito que los valores críticos del t -ratio, $t_{\hat{b}}$, dependen del parámetro de corto plazo, a .

OBSERVACIÓN 3.2. Cuando $a = 1$ (restricción COMFAC) $q = -(a - 1)\sigma_1/\sigma_2$ es cero y la distribución límite coincide con la obtenida por, ver ecuación (2.18). Por tanto, cabe esperar que, al pasar de $a = 0$ a $a = 1$, la distribución empírica se desplace a la izquierda, dando como resultado valores críticos con valores absolutos mayores,

⁴Resultados no incluidos, aunque disponibles.

T	DUM	a=0			a=0.5			a=1		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	NO	-2.694	-1.979	-1.818	-2.888	-2.210	-1.917	-2.940	-2.940	-2.940
	D1	-2.139	-1.776	-1.732	-2.566	-1.845	-1.756	-3.023	-3.023	-3.023
	D2	-2.338	-1.830	-1.769	-2.758	-1.933	-1.803	-3.079	-3.079	-3.079
	D3	-2.268	-1.816	-1.751	-2.727	-1.905	-1.790	-3.065	-3.065	-3.065
100	NO	-2.610	-1.884	-1.734	-2.790	-2.113	-1.820	-2.852	-2.852	-2.852
	D1	-1.896	-1.724	-1.709	-2.119	-1.756	-1.720	-2.938	-2.938	-2.938
	D2	-1.984	-1.721	-1.688	-2.338	-1.777	-1.708	-2.946	-2.946	-2.946
	D3	-1.927	-1.715	-1.689	-2.219	-1.750	-1.708	-2.944	-2.944	-2.944
1000	NO	-2.589	-1.883	-1.738	-2.766	-2.092	-1.828	-2.860	-2.860	-2.860
	D1	-1.738	-1.678	-1.672	-1.797	-1.687	-1.676	-2.915	-2.915	-2.915
	D2	-1.770	-1.675	-1.661	-1.867	-1.694	-1.668	-2.928	-2.928	-2.928
	D3	-1.767	-1.696	-1.689	-1.839	-1.711	-1.692	-2.922	-2.922	-2.922

Tabla 2.1.1: Valores críticos del contraste MCE. Co-ruptura simultánea ($c_t = 0$)

T	DUM	a=0			a=0.5			a=1		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	NO	-2.428	-1.939	-1.810	-2.518	-2.179	-1.893	-2.494	-2.494	-2.494
	D1	-2.536	-1.780	-1.733	-2.883	-1.880	-1.762	-2.472	-2.472	-2.472
	D2	-2.565	-1.856	-1.770	-2.761	-1.998	-1.816	-2.500	-2.500	-2.500
	D3	-2.604	-1.823	-1.759	-2.831	-1.976	-1.808	-2.474	-2.474	-2.474
100	NO	-2.034	-1.871	-1.736	-2.041	-2.009	-1.830	-2.035	-2.035	-2.035
	D1	-2.129	-1.732	-1.707	-2.692	-1.771	-1.722	-2.026	-2.026	-2.026
	D2	-2.256	-1.731	-1.691	-2.342	-1.810	-1.713	-2.038	-2.038	-2.038
	D3	-2.271	-1.722	-1.689	-2.480	-1.780	-1.707	-2.002	-2.002	-2.002
1000	NO	-1.729	-1.763	-1.696	-1.734	-1.735	-1.767	-1.729	-1.729	-1.729
	D1	-1.808	-1.673	-1.668	-2.114	-1.690	-1.671	-1.760	-1.760	-1.760
	D2	-1.857	-1.680	-1.662	-1.855	-1.711	-1.671	-1.764	-1.764	-1.764
	D3	-1.907	-1.702	-1.691	-1.940	-1.721	-1.697	-1.763	-1.763	-1.763

Tabla 2.1.2: Valores críticos del contraste MCE. Co-ruptura en diferencias, no co-ruptura en niveles. ($\Delta\mu_{y,t} - a\Delta\mu_{z,t} = 0.5$)

T	DUM	a=0			a=0.5			a=1		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	NO	-2.694	-1.979	-1.818	-2.888	-2.210	-1.917	-2.940	-2.940	-2.940
	D1	-3.360	-5.695	-6.328	-3.139	-4.653	-5.988	-3.023	-3.023	-3.023
	D2	-3.690	-7.119	-7.952	-3.266	-5.635	-7.518	-3.079	-3.079	-3.079
	D3	-3.254	-4.490	-4.722	-3.116	-4.043	-4.602	-3.065	-3.065	-3.065
100	NO	-2.610	-1.884	-1.734	-2.790	-2.113	-1.820	-2.852	-2.852	-2.852
	D1	-4.611	-10.856	-12.395	-3.599	-8.322	-11.504	-2.938	-2.938	-2.938
	D2	-5.069	-12.928	-14.883	-3.715	-9.754	-13.837	-2.946	-2.946	-2.946
	D3	-4.237	-7.830	-8.412	-3.413	-6.650	-8.133	-2.944	-2.944	-2.944
1000	NO	-2.589	-1.883	-1.738	-2.766	-2.092	-1.828	-2.860	-2.860	-2.860
	D1	-10.736	-32.393	-37.520	-6.548	-23.927	-34.734	-2.915	-2.915	-2.915
	D2	-12.364	-39.237	-45.199	-7.142	-28.770	-41.999	-2.928	-2.928	-2.928
	D3	-9.479	-22.450	-24.228	-5.917	-18.406	-23.295	-2.922	-2.922	-2.922

Tabla 2.1.3: Valores críticos del contraste MCE. Co-ruptura en niveles, no en diferencias ($\mu_{y,t} - \mu_{z,t} = 0$).

TABLA 2.1. El PGD se genera bajo $H_0 : \Delta y_t = c_t + a\Delta z_t + u_{1,t}$, $\Delta z_t = sD_{j_t} + u_{2,t}$, donde $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t}$, $\sigma_1^2 = \text{var}(u_{1,t}) = 1$, y $\sigma_2^2 = \text{var}(u_{2,t}) = s^2$. El modelo estimado es $\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1,t}$. Se calculan los valores críticos al 5% del contraste MCE para distintos tamaños muestrales ($T = 25, 100, 1000$), valores de correlación de corto plazo ($a = 0.0, 0.5, 1.0$), y tamaños del salto ($s = 1, 6, 16$).

ver Tabla 2.1.1.⁵ Cuando se cumple la restricción COMFAC ($a = 1$), los valores críticos obtenidos son los de la distribución DF cuando el PGD es un paseo aleatorio sin deriva y la regresión se estima con constante. Nótese que los CV no dependen del tamaño del salto ($s = 1, 6, 16$) y que la variación con el tipo de salto (D_1, D_2, D_3) es despreciable. Los CV al 5% están alrededor de -3.0 para $T = 25$, sobre -2.94 para $T = 100$ y para $T = 1000$ está alrededor de -2.92, lo que constituye un buen resultado para realizar aplicaciones empíricas, dado que no se impone la restricción COMFAC para realizar el contraste.

OBSERVACIÓN 3.3. También se obtuvieron los valores críticos permitiendo incertidumbre en la dinámica del modelo, retardos desconocidos. El número de retardos del modelo se estima mediante el criterio SBIC. Las conclusiones son las mismas que antes para tamaños muestrales iguales a 100 y 1000. No obstante, tal y como era de esperar, para $T = 25$ la distribución se deslaza a la izquierda, dejando unos valores críticos al 5% entre -3.2 y -4.9.

Tamaño empírico del contraste MCE con co-ruptura simultánea. Para juzgar la validez de nuestro contraste se analiza el tamaño empírico del mismo utilizando los valores críticos obtenidos bajo la hipótesis nula pero cuando las perturbaciones se generan mediante procesos $MA(1)$. En concreto, se simulan los datos con el siguiente PGD: $\Delta y_t = c_t + a\Delta z_t + u_{1t} + \theta u_{1,t-1}$, donde θ es igual a 0.5 y -0.5. El resultado es una gran distorsión de tamaño del contraste dependiendo del signo de θ cuando no se incluye ninguna dinámica en el modelo estimado, ver modelo (3.2). Para parámetros $MA(1)$ positivos, $\theta = 0.5$, las distorsiones se hacen mayores cuando $a \neq 1$, y para valores del parámetro $MA(1)$ negativos, $\theta = -0.5$, las mayores distorsiones se dan para $a = 1$. Este problema disminuye si incluimos términos dinámicos en el modelo, modelo (3.3), elegidos mediante criterio SBIC, pero el tamaño del contraste sigue dependiendo del signo de θ , ver Arranz y Escribano (1998). Por tanto, los valores críticos obtenidos no son fiables en presencia de errores $MA(1)$, especialmente con $\theta = -0.5$. Podríamos obtener valores críticos más fiables mediante la aplicación de técnicas bootstrap, ver Capítulo 4, pero eso lo dejamos fuera de este Capítulo.

Potencia del contraste MCE con co-ruptura simultánea. La potencia del contraste MCE (t_j) se analiza generando datos con el PGD bajo H_1 para valores de b que cumplen $-2 < b < 0$. Se consideran distintos valores para b , $b = -0.05, -0.1, -0.25, -0.5$, y -0.75 . Recordemos que, a partir de la ecuación (2.4), el error de cointegración

⁵Nótese que cuando $a = 0$ los valores críticos para NO, D_1, D_2, D_3 no coinciden, puesto que en la regresión bajo H_0 se incluye Δz_t , que depende de $D_{j,t}$ y que afecta marginalmente a los valores críticos obtenidos.

T	a	b	NO					D3				
			-0.05	-0.1	-0.25	-0.5	-0.75	-0.05	-0.1	-0.25	-0.5	-0.75
25	0.5	s=1	6.4	8.9	22.5	64.9	94.2	10.4	16.3	40.0	81.1	97.8
		s=6	30.3	63.8	98.0	100.0	100.0	78.5	95.8	99.9	100.0	100.0
		s=16	82.7	98.6	100.0	100.0	100.0	99.4	100.0	100.0	100.0	100.0
	1.0	s=1	6.0	7.5	15.9	48.7	86.2	5.9	7.6	15.7	47.7	84.8
		s=6	6.0	7.5	15.9	48.7	86.2	5.9	7.6	15.7	47.7	84.8
		s=16	6.0	7.5	15.9	48.7	86.2	5.9	7.6	15.7	47.7	84.8
100	0.5	s=1	18.5	49.9	99.4	100.0	100.0	74.2	96.1	100.0	100.0	100.0
		s=6	96.5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
		s=16	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	1.0	s=1	12.7	34.2	97.6	100.0	100.0	12.4	32.6	97.0	100.0	100.0
		s=6	12.7	34.2	97.6	100.0	100.0	12.4	32.6	97.0	100.0	100.0
		s=16	12.7	34.2	97.6	100.0	100.0	12.4	32.6	97.0	100.0	100.0

Tabla 2.2.1: Potencia ajustada del contraste MCE. Co-ruptura simultánea.

T	a	b	NO					D3				
			-0.05	-0.1	-0.25	-0.5	-0.75	-0.05	-0.1	-0.25	-0.5	-0.75
25	0.5	s=1	4.7	3.8	1.8	1.2	1.2	6.2	7.2	10.8	18.2	28.4
		s=6	25.1	47.9	76.0	86.0	90.9	13.2	22.1	31.1	29.6	25.2
		s=16	79.7	97.2	100.0	100.0	100.0	31.0	39.3	40.0	34.0	28.0
	1.0	s=1	4.1	2.9	0.8	0.1	0.1	4.1	2.9	0.7	0.1	0.0
		s=6	4.1	2.9	0.8	0.1	0.1	4.1	2.9	0.7	0.1	0.0
		s=16	4.1	2.9	0.8	0.1	0.1	4.1	2.9	0.7	0.1	0.0
100	0.5	s=1	2.0	0.9	0.4	0.1	0.1	4.2	3.4	2.4	1.9	2.0
		s=6	44.3	52.6	55.8	53.3	50.8	28.0	31.4	28.6	22.0	15.8
		s=16	99.3	99.8	99.8	99.7	99.9	41.8	41.6	35.4	25.9	18.5
	1.0	s=1	1.1	0.1	0.0	0.0	0.0	1.1	0.1	0.0	0.0	0.0
		s=6	1.1	0.1	0.0	0.0	0.0	1.1	0.1	0.0	0.0	0.0
		s=16	1.1	0.1	0.0	0.0	0.0	1.1	0.1	0.0	0.0	0.0
1000	0.5	s=1	0.1	0.1	0.0	0.0	0.0	0.4	0.3	0.1	0.0	0.0
		s=6	37.4	36.3	30.9	23.1	16.9	29.2	27.2	22.1	14.6	9.1
		s=16	69.9	67.9	62.4	55.5	51.6	39.2	35.8	28.1	18.4	11.3
	1.0	s=1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		s=6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
		s=16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tabla 2.2.2: Potencia ajustada del contraste MCE. Co-ruptura en diferencias, no en niveles.

T	a	b	NO					D3				
			-0.05	-0.1	-0.25	-0.5	-0.75	-0.05	-0.1	-0.25	-0.5	-0.75
25	0.5	s=1	6.4	8.9	22.5	64.9	94.2	6.3	8.3	18.6	52.8	87.2
		s=6	30.3	63.8	98.0	100.0	100.0	9.5	15.9	33.7	71.7	96.3
		s=16	82.7	98.6	100.0	100.0	100.0	9.6	15.2	31.3	68.0	96.3
100	0.5	s=1	18.5	49.9	99.4	100.0	100.0	13.1	27.0	85.4	100.0	100.0
		s=6	96.5	100.0	100.0	100.0	100.0	14.4	21.0	54.5	99.2	100.0
		s=16	100.0	100.0	100.0	100.0	100.0	13.4	18.3	45.7	98.5	100.0
1000	0.5	s=1	100.0	100.0	100.0	100.0	100.0	42.7	97.9	100.0	100.0	100.0
		s=6	100.0	100.0	100.0	100.0	100.0	3.5	10.0	98.1	100.0	100.0
		s=16	100.0	100.0	100.0	100.0	100.0	2.3	4.6	88.2	100.0	100.0

Tabla 2.2.3: Potencia ajustada del contraste MCE. Co-ruptura en niveles, no en diferencias

TABLA 2.2. El PGD se genera bajo $H_1 : b < 0$, $\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1,t}$, $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t-1} - \mu_{z,t-1})$, $\Delta z_t = sD_{j_t} + u_{2,t}$, donde $\sigma_1^2 = \text{var}(u_{1,t}) = 1$, y $\sigma_2^2 = \text{var}(u_{2,t}) = s^2$. El modelo estimado es $\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_t$. Los valores críticos del contraste son los de la Tabla 2.1 para diferentes tamaños muestrales ($T=25, 100, 1000$), parámetros de corto plazo ($a=0.0, 0.5, 1.0$), y tamaños del salto ($s=1, 6, 16$).

tiene una representación autorregresiva que depende del parámetro b . Si llamamos ρ_1 al parámetro autorregresivo de primer orden, $\rho_1 = b+1$, por lo que $\rho_1 = 0.95, 0.9, 0.75, 0.5$, y 0.25 , son los valores correspondientes a los valores anteriores el parámetro b . Por tanto, para $b = -0.05$ hemos de esperar que tenga poco poder frente a alternativas estacionarias, ya que el parámetro $AR(1)$, 0.95 , está próximo a la raíz unitaria. Esta intuición se puede explicar a partir de la ecuación (2.21). Cuando b aumenta es como si se incrementara h relativo a T . Por tanto, de la ecuación (2.22) es de esperar una reducción de la potencia cuando se cumple la restricción $COMFAC(a = 1)$.

En general, la potencia ajustada en el caso de co-ruptura simultánea mostrada en la Tabla 2.2.1 es alta para los distintos tamaños de los saltos ($s = 1, 6, 16$) y tamaños muestrales. La potencia más baja del contraste MCE sucede para el valor $a = 1$ (restricción COMFAC), especialmente para tamaños muestrales pequeños ($T = 25$) y valores absolutos pequeños de b (alternativas locales). Recordemos que $a = 1$ corresponde a $q = 0$, ver ecuación (2.18), y que en tal caso la distribución límite del contraste es la distribución de Dickey y Fuller. Esto es lo que lleva a Kremers et al. (1992), Hansen (1995), y Banerjee, Dolado y Mestre (1998) a sugerir la inclusión de variables como Δz_t en la ecuación de la regresión del contraste para aumentar la potencia del contraste de no-cointegración, $b = 0$, o de raíces unitarias en el contexto univariante.

En resumen, bajo co-rupturas estructurales simultáneas, el enfoque basado en el contraste de no-cointegración ($t_{\hat{b}}$) en un modelo de corrección del error es muy robusto cuando no hay COMFAC ($a \neq 1$) y cuando σ_2 es grande respecto a σ_1 . Si $a = 1$ la potencia es baja para $T = 25$ y 100 para valores $-0.5 < b \leq 0$, pero es alta para $T = 1000$. Se obtienen resultados similares para $T = 25$, $a \neq 1$ y $s = 1$. Sin embargo, la potencia aumenta con el tamaño muestral y el tamaño de los saltos, $s = 6, 16$, ver Tabla 2.2.1. El problema está abierto en los casos en que las variables presenta co-ruptura parcial, sólo en niveles o sólo en diferencias, que es el objeto del estudio del apartado siguiente.

3.2. Experimento de Simulación de Monte Carlo: Contraste MCE sin Co-ruptura simultánea.

Valores críticos del contraste MCE. Los valores críticos se obtienen para las diferentes rupturas, suponiendo que se ignora que han tenido lugar y que se aplica el contraste de corrección del error habitual en el que c_t se considera constante (modelo mal especificado).

OBSERVACIÓN 3.4. *Co-ruptura en diferencias, no en niveles:* En este caso el PGD bajo la hipótesis nula de $b = 0$ tiene un término constante, c_d , ver ecuación (2.26), y, por tanto, dado que las ecuaciones del modelo estimado, (3.2)–(3.3) no hay tendencia, el t -ratio, $t_{\hat{b}}$, tiene una distribución límite Normal, ver Tabla 2.1.2. Hay que resaltar que para $T = 1000$, el CV al 5% es estable alrededor de -1.7, y robusto a s , el tipo de salto (D_1, D_2, D_3) , y el valor del parámetro a .

OBSERVACIÓN 3.5. *Co-ruptura en niveles, pero no en diferencias:* Si $a = 1$ co-ruptura en niveles implica co-ruptura en diferencias, y por tanto estamos en el caso de co-ruptura simultánea, concretamente co-ruptura completa. No obstante, si es una co-ruptura parcial en niveles, pero no hay co-ruptura en diferencias, $a \neq 1$, la situación es radicalmente distinta. Bajo $H_0 : b = 0$ las dummies D_1, D_2 , y D_3 afectan al PGD, ver ecuación (2.28), y, por tanto, los CV's son muy inestables. Por ejemplo, para $T = 1000$, $a = 0.5$ y D_3 , el CV es -5.9 para $s = 1$, y -23.3 si $s = 16$. Por tanto, este tipo de error de especificación da lugar a los valores críticos más inestables.

Este resultado complica la aplicación del contraste MCE desde el punto de vista empírico, pues nos obliga a generar los valores críticos correspondientes a cada tipo de salto y para cada tamaño del mismo (s), así como para cada valor del parámetro de corto plazo, a .

OBSERVACIÓN 3.6. Los resultados relativos al CV se mantienen en el caso de que tratemos de aproximar el error de especificación del cambio estructural incluyendo términos dinámicos en el modelo, ver ecuación (3.3). Esta conclusión basada e modelos MCE con error de especificación hace el análisis empírico más dependiente aún si cabe de los valores críticos correctos, que dependen, en definitiva, del tipo de cambio estructural que tiene lugar.

Tamaño Empírico del Contraste MCE sin Co-ruptura Simultánea. El resultado del análisis es que el contraste MCE tiene un tamaño inadecuado en el caso de error de especificación dinámica, especialmente en el caso de co-ruptura parcial en niveles, y si hay co-ruptura parcial en diferencias y el parámetro $MA(1)$ es $\theta = -0.5$, ver Arranz y Escribano (1998). No obstante, estas conclusiones tan negativas se ven atemperadas si incluimos términos dinámicos en el modelo, pues el tamaño del contraste no se ve tan afectado en dicho caso por el signo del parámetro $MA(1)$, θ , aunque siga sin estar alrededor del tamaño nominal del contraste, 5%.

Potencia del contraste MCE sin Co-ruptura Simultánea. La Tabla 2.2.2 presenta los resultados del contraste MCE, $(t_{\hat{b}})$, cuando sólo hay co-ruptura en diferencias, pero no en niveles, y la ignoramos al no incluir variables dummy en el modelo MCE

dinámico. Podríamos analizar situaciones similares si incluyéramos variables dummy en momentos erróneos en el tiempo. Los resultados indican que el contraste MCE basado en la ecuación con error de especificación de los saltos no tiene potencia para ninguno de los tamaños muestrales considerados.

La potencia que se obtiene con el contraste MCE para el otro caso extremo de co-ruptura parcial es mucho mayor. Considerando el caso de co-ruptura en niveles pero no en diferencias⁶, la potencia del contraste es mayor que en el caso de co-ruptura parcial en diferencia, pero menor que en el caso de co-ruptura simultánea si hay salto, $a \neq 1$ y $-0.5 < b < 0$, ver Tabla 2.2.3. En tal caso, la potencia ajustada se ve afectada especialmente por el tamaño del salto. La potencia es razonablemente buena para valores de b mayores o iguales que 0.5 en valor absoluto. Estas conclusiones se mantienen si incluimos términos dinámicos en el modelo, según la ecuación (3.3).

En resumen, estos resultados son desalentadores desde el punto de vista del análisis empírico, dado que nunca tendremos la certeza de si hay o no cointegración debido al error de especificación cometido por una mala modelización de la co-ruptura estructural. Dado que los valores críticos del contraste MCE y su potencia dependen del tipo de cambio estructural considerado, tenemos muchas posibles combinaciones de saltos alternativas que pueden alterar completamente los resultados del contraste de cointegración. En la Sección 4 se abordará la cuestión de si es posible solucionar este problema mediante la aplicación de modelos MCE extendidos.

3.2.1. *Una solución parcial basada en la restricción COMFAC.* Como acabamos de ver, la situación en que hay co-ruptura parcial sólo en niveles pero no en diferencias nos plantea múltiples problemas. En este caso particular de co-ruptura parcial, y sólo en este caso, nuestra recomendación es aplicar el enfoque del contraste de Engle-Granger.

En resumen, partiendo del modelo MCE habitual

$$\Delta y_t = -bc_t + (\alpha - a)sD_{j,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t}$$

se obtiene fácilmente que:

$$\begin{aligned} (\Delta y_t - \alpha \Delta z_t) &= -bc_t + (\alpha - a)sD_{j,t} + (a - \alpha)\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \\ &= -bc_t + (a - \alpha)(\Delta z_t - sD_{j,t}) + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \end{aligned}$$

⁶Recordemos que si $a = 1$, co-ruptura en niveles implica co-ruptura en diferencias, por lo que estaríamos en el caso de co-ruptura simultáneas, razón por la que no se incluye en la Tabla 2.2.3.

La ecuación (2.27) se puede expresar como

$$(\Delta y_t - \alpha \Delta z_t) = -bc_t + b(y_{t-1} - \alpha z_{t-1}) + v_{1,t} \quad (3.4)$$

donde $v_{1,t} = (a - \alpha)(\Delta z_t - sD_{j,t}) + u_{1,t} = (a - \alpha)u_{2,t} + u_{1,t}$ y, por tanto, los saltos desaparecen de la ecuación sobre la que se realiza el contraste.

Así pues, podemos usar la siguiente ecuación para llevar a cabo el contraste MCE basado en el modelo MCE habitual

$$(\Delta y_t - \alpha \Delta z_t) = c + b(y_{t-1} - \alpha z_{t-1}) + \epsilon_t \quad (3.5)$$

Tal y como se aprecia claramente en la Tabla 2.3, la imposición de la restricción COMFAC proporciona valores críticos estables en tres casos: cuando no hay cambios estructurales, cuando se satisface la restricción COMFAC ($a = 1$) y cuando hay co-ruptura sólo en niveles. No obstante, dado que estamos analizando otros casos también, no se puede recomendar este procedimiento como herramienta general para realizar contrastes MCE.

4. Modelos de Corrección del Error Extendidos

En esta Sección se extienden las implicaciones de Toda y Yamamoto (1995) y Dolado y Lutkepohl (1996). Lo que vamos a hacer es añadir un retardo adicional del término de corrección del error para obtener los resultados de inferencia habituales incluso con variables no estacionarias. En el modelo (3.2), simplemente hay que incluir un retardo adicional del mecanismo de corrección del error en la regresión para obtener el siguiente modelo MCE extendido

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + d(y_{t-2} - \alpha z_{t-2}) + u_t. \quad (4.1)$$

El t -ratio del parámetro b en (4.1), t_b no converge a la distribución límite dada en (2.16) o (2.17), sino a la distribución Normal estándar. Este resultado no sólo simplifica en gran medida el análisis empírico, sino que también reduce, como veremos más adelante, la inestabilidad de los valores críticos a los parámetros molestos y la provocada por la presencia de cambios estructurales.

La intuición es la siguiente. Consideremos el modelo de corrección del error introducido en las ecuaciones (2.2a)–(2.2c) y comparemos la formulación con la ecuación (4.1). Ambas son equivalentes si se cumple la condición

$$c_t = c + d(y_{t-2} - \alpha z_{t-2}) \quad (4.2)$$

T	DUM	a=0			a=0.5			a=1		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	NO	-2.977	-3.005	-3.022	-3.012	-2.989	-3.017	-2.989	-2.989	-2.989
	D1	-1.271	-0.687	-0.668	-2.160	-0.755	-0.678	-2.989	-2.989	-2.989
	D2	-1.582	-0.566	-0.530	-2.399	-0.675	-0.541	-2.989	-2.989	-2.989
	D3	-2.130	-1.856	-1.848	-2.553	-1.881	-1.851	-2.989	-2.989	-2.989
100	NO	-2.862	-2.872	-2.890	-2.871	-2.873	-2.883	-2.872	-2.872	-2.872
	D1	0.205	0.983	1.018	-0.707	0.878	1.001	-2.872	-2.872	-2.872
	D2	0.641	1.822	1.870	-0.937	1.680	1.843	-2.872	-2.872	-2.872
	D3	-1.486	-1.382	-1.379	-1.742	-1.404	-1.378	-2.872	-2.872	-2.872
1000	NO	-2.848	-2.850	-2.846	-2.890	-2.849	-2.850	-2.856	-2.856	-2.856
	D1	4.635	6.933	7.030	2.355	6.629	6.981	-2.856	-2.856	-2.856
	D2	6.884	10.036	10.168	3.732	9.622	10.105	-2.856	-2.856	-2.856
	D3	-1.245	-1.156	-1.155	-1.354	-1.173	-1.156	-2.856	-2.856	-2.856

Tabla 2.3.1: Valores críticos del contraste MCE-COMFAC. Co-ruptura simultánea. ($c_t = 0$)

T	DUM	a=0			a=0.5			a=1		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	NO	-2.704	-3.005	-3.006	-2.537	-2.938	-3.002	-2.483	-2.483	-2.483
	D1	-2.547	-0.751	-0.685	-3.264	-0.919	-0.727	-2.483	-2.483	-2.483
	D2	-2.871	-0.869	-0.629	-3.176	-1.394	-0.758	-2.483	-2.483	-2.483
	D3	-2.668	-1.877	-1.854	-2.933	-1.956	-1.863	-2.483	-2.483	-2.483
100	NO	-2.178	-2.837	-2.860	-2.008	-2.680	-2.857	-2.000	-2.000	-2.000
	D1	-1.068	0.937	1.002	-3.596	0.777	0.971	-2.000	-2.000	-2.000
	D2	-2.803	1.593	1.802	-3.527	0.985	1.684	-2.000	-2.000	-2.000
	D3	-1.975	-1.350	-1.367	-2.803	-1.317	-1.359	-2.000	-2.000	-2.000
1000	NO	-1.782	-2.351	-2.746	-1.759	-1.982	-2.541	-1.733	-1.733	-1.733
	D1	3.147	6.795	6.980	-6.670	6.329	6.878	-1.733	-1.733	-1.733
	D2	-2.904	9.527	10.009	-6.917	8.328	9.746	-1.733	-1.733	-1.733
	D3	-0.618	-1.106	-1.136	-2.548	-1.046	-1.119	-1.733	-1.733	-1.733

Tabla 2.3.2: Valores críticos del contraste MCE-COMFAC. Co-ruptura en diferencias, no en niveles, ($\Delta\mu_{y,t} - a\Delta\mu_{y,t} = 0.5$)

T	DUM	a=0			a=0.5			a=1		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	NO	-2.977	-3.005	-3.022	-3.012	-2.989	-3.017	-2.989	-2.989	-2.989
	D1	-2.977	-3.005	-3.022	-3.012	-2.989	-3.017	-2.989	-2.989	-2.989
	D2	-2.977	-3.005	-3.022	-3.012	-2.989	-3.017	-2.989	-2.989	-2.989
	D3	-2.977	-3.005	-3.022	-3.012	-2.989	-3.017	-2.989	-2.989	-2.989
100	NO	-2.862	-2.872	-2.890	-2.871	-2.873	-2.883	-2.872	-2.872	-2.872
	D1	-2.862	-2.872	-2.890	-2.871	-2.873	-2.883	-2.872	-2.872	-2.872
	D2	-2.862	-2.872	-2.890	-2.871	-2.873	-2.883	-2.872	-2.872	-2.872
	D3	-2.862	-2.872	-2.890	-2.871	-2.873	-2.883	-2.872	-2.872	-2.872
1000	NO	-2.848	-2.850	-2.846	-2.890	-2.849	-2.850	-2.856	-2.856	-2.856
	D1	-2.848	-2.850	-2.846	-2.890	-2.849	-2.850	-2.856	-2.856	-2.856
	D2	-2.848	-2.850	-2.846	-2.890	-2.849	-2.850	-2.856	-2.856	-2.856
	D3	-2.848	-2.850	-2.846	-2.890	-2.849	-2.850	-2.856	-2.856	-2.856

Tabla 2.3.3: Valores críticos del contraste MCE-COMFAC. Co-ruptura en niveles, no en diferencias, ($\mu_{y,t} - \mu_{y,t} = 0$).

TABLA 2.3. El PGD se genera bajo $H_0 : \Delta y_t = c_t + a\Delta z_t + u_{1t}$, $\Delta z_t = sD_{j_t} + u_{2t}$, donde $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t}$, $\sigma_1^2 = \text{var}(u_{1t}) = 1$, y $\sigma_2^2 = \text{var}(u_{2t}) = s^2$. El modelo estimado es $\Delta y_t - \Delta z_t = c + b(y_{t-1} - z_{t-1}) + u_{1t}$. Los valores críticos al 5% se calculan para distintos tamaños muestrales, (T= 25, 100, 1000), parámetros de corto plazo ($a= 0.0, 0.5, 1.0$), y tamaños del salto ($s=1, 6, 16$).

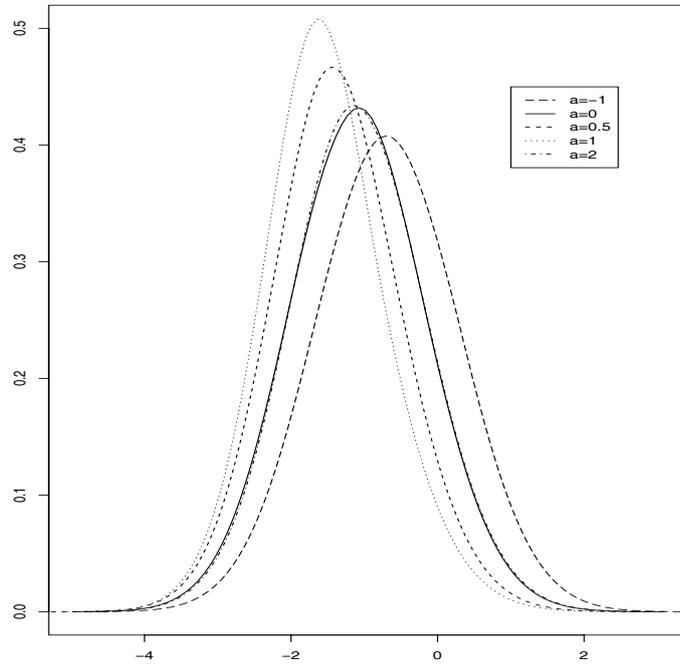


Figura 2.4.1: MCE

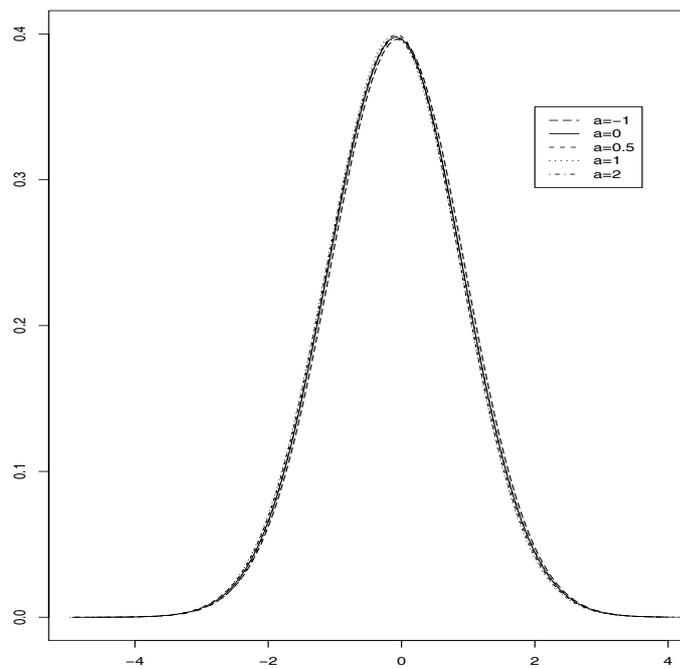


Figura 2.4.2: MCE Extendedido

FIGURA 2.4. Densidad del contraste MCE bajo H_0 , $s = 1$, sin saltos

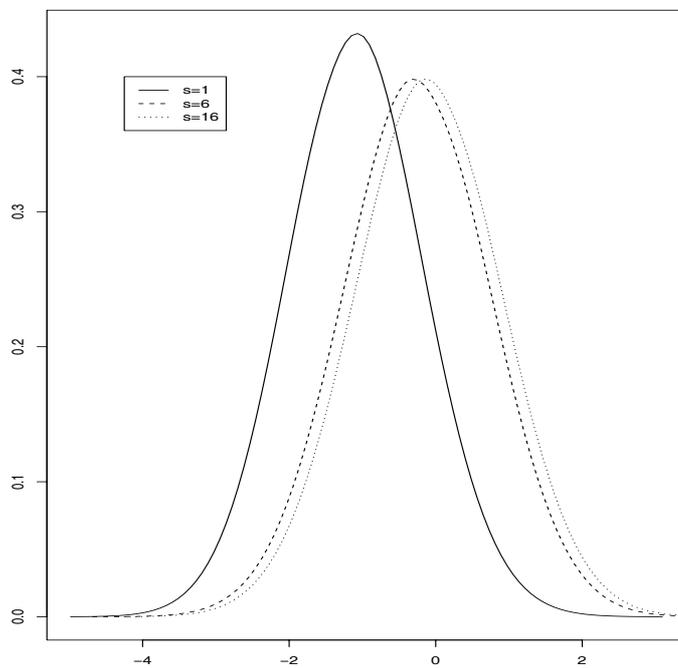


Figura 2.5.1: MCE

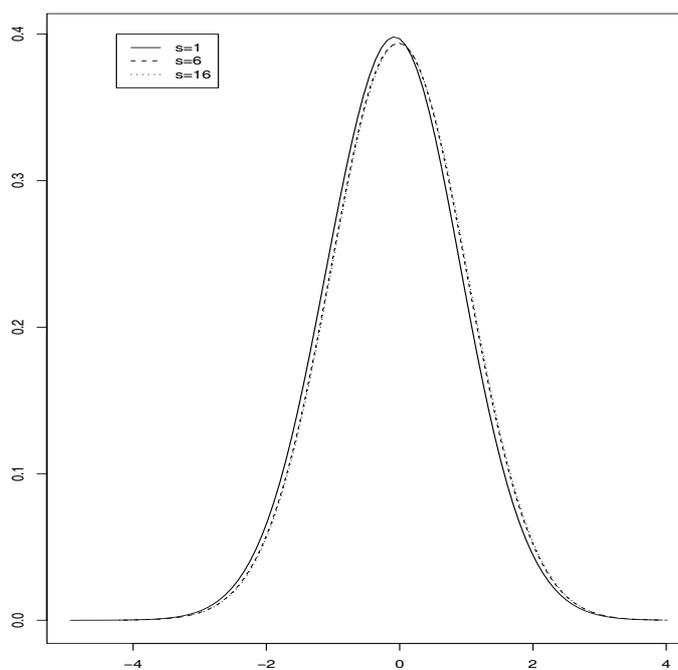


Figura 2.5.2: MCE Extendedido

FIGURA 2.5. Densidad del contraste MCE bajo H_0 , $a = 0$, sin saltos

o si

$$\Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t-1} - \alpha\mu_{z,t-1}) = c + d(y_{t-2} - \alpha z_{t-2}). \quad (4.3)$$

En general, esta condición no se cumple si no hay co-ruptura en niveles ni en diferencias. No obstante, en el caso de que haya co-ruptura parcial en diferencias, el término $(y_{t-2} - \alpha z_{t-2})$ presenta un salto en niveles que aproximará el término omitido $(\mu_{y,t-1} - \alpha\mu_{z,t-1})$. Es decir,

$$(\mu_{y,t-1} - \alpha\mu_{z,t-1}) \approx c + d(y_{t-2} - \alpha z_{t-2}). \quad (4.4)$$

Por otra parte, si hay co-ruptura parcial en niveles, el término $(y_{t-2} - \alpha z_{t-2})$ tiene un salto que puede aproximar el efecto de $(\Delta\mu_{y,t} - a\Delta\mu_{z,t})$. Así pues,

$$\Delta\mu_{y,t} - a\Delta\mu_{z,t} \approx c + d(y_{t-2} - \alpha z_{t-2}). \quad (4.5)$$

En el caso de que haya regresores dinámicos adicionales en el modelo sobreparametrizado, como en (3.3), la ecuación del contraste MCE extendido es

$$\begin{aligned} \phi(B)\Delta y_t = c + \theta(B)\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) \\ + d(y_{t-k-2} - \alpha z_{t-k-2}) + u_t \end{aligned}, \quad (4.6)$$

donde los polinomios en B , $\phi(B)$ y $\theta(B)$ son

$$\begin{aligned} \phi(B) &= 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_{p^*} B^{p^*} \\ \theta(B) &= a - \theta_1 B - \theta_2 B^2 - \dots - \theta_{q^*} B^{q^*} \end{aligned}$$

y donde $k = \max\{p^*, q^*\}$ se elige utilizando el criterio SBIC.

En esta sección se analiza el comportamiento de los valores críticos y la potencia del contraste MCE ($t_{\hat{b}}$) con y sin co-ruptura simultánea cuando utilizamos los modelos MCE extendidos (4.1) y (4.6).

Como se observa en la Tabla 2.4, cuando se estima el Modelo (4.1), se obtienen valores críticos muy estables en los tres casos de co-ruptura analizados. Es importante resaltar que los valores críticos ahora no se aproximan a los obtenidos por Dickey y Fuller, sino que rondan el valor -1.8, que es similar al valor crítico al nivel 5% de una distribución Gaussiana. La mayor diferencia se da cuando hay co-ruptura parcial en niveles, en cuyo caso los valores críticos están entre -1.67 y -2.5.

Como es habitual, los valores críticos de la Tabla 2.4 dependen del tamaño muestral. Sólo para muestras de tamaños $T = 25$ los $T = 100$ los VC's aumentan en valor absoluto cuando se cumple la restricción COMFAC ($a = 1$). Por ejemplo, para

T	DUM	a=0			a=0.5			a=1		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	NO	-2.322	-1.873	-1.779	-2.482	-1.995	-1.840	-2.550	-2.550	-2.550
	D1	-1.899	-1.748	-1.730	-2.098	-1.794	-1.745	-2.587	-2.587	-2.587
	D2	-1.940	-1.739	-1.720	-2.219	-1.776	-1.734	-2.608	-2.608	-2.608
	D3	-2.000	-1.764	-1.720	-2.282	-1.827	-1.744	-2.630	-2.630	-2.630
100	NO	-1.979	-1.754	-1.712	-2.076	-1.820	-1.736	-2.114	-2.114	-2.114
	D1	-1.757	-1.667	-1.650	-1.797	-1.680	-1.655	-2.127	-2.127	-2.127
	D2	-1.727	-1.631	-1.614	-1.795	-1.656	-1.624	-2.136	-2.136	-2.136
	D3	-1.762	-1.657	-1.634	-1.833	-1.696	-1.643	-2.157	-2.157	-2.157
1000	NO	-1.774	-1.708	-1.686	-1.807	-1.726	-1.695	-1.805	-1.805	-1.805
	D1	-1.702	-1.644	-1.629	-1.715	-1.666	-1.643	-1.814	-1.814	-1.814
	D2	-1.696	-1.655	-1.663	-1.708	-1.665	-1.658	-1.821	-1.821	-1.821
	D3	-1.704	-1.638	-1.639	-1.724	-1.645	-1.630	-1.818	-1.818	-1.818

Tabla 2.4.1: Valores críticos del contraste MCE extendido. Co-ruptura simultánea, ($c_t = 0$).

T	DUM	a=0			a=0.5			a=1		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	NO	-2.119	-1.849	-1.779	-2.133	-1.965	-1.820	-2.173	-2.173	-2.173
	D1	-2.114	-1.753	-1.728	-2.449	-1.789	-1.745	-2.168	-2.168	-2.168
	D2	-2.148	-1.743	-1.719	-2.362	-1.803	-1.740	-2.203	-2.203	-2.203
	D3	-2.194	-1.769	-1.721	-2.418	-1.838	-1.746	-2.193	-2.193	-2.193
100	NO	-1.836	-1.763	-1.723	-1.846	-1.811	-1.753	-1.869	-1.869	-1.869
	D1	-1.759	-1.665	-1.650	-2.050	-1.677	-1.659	-1.864	-1.864	-1.864
	D2	-1.800	-1.628	-1.616	-1.935	-1.650	-1.626	-1.861	-1.861	-1.861
	D3	-1.802	-1.656	-1.634	-1.941	-1.695	-1.644	-1.867	-1.867	-1.867
1000	NO	-1.719	-1.692	-1.680	-1.754	-1.716	-1.682	-1.734	-1.734	-1.734
	D1	-1.690	-1.645	-1.630	-1.795	-1.660	-1.643	-1.742	-1.742	-1.742
	D2	-1.719	-1.666	-1.665	-1.778	-1.660	-1.670	-1.743	-1.743	-1.743
	D3	-1.697	-1.638	-1.640	-1.741	-1.644	-1.631	-1.747	-1.747	-1.747

Tabla 2.4.2: Valores críticos del contraste MCE extendido. Cor-ruptura en diferencias, no en niveles. ($\Delta\mu_{y,t} - a\Delta\mu_{y,t} = 0.5$).

T	DUM	a=0			a=0.5			a=1		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	NO	-2.322	-1.873	-1.779	-2.482	-1.995	-1.840	-2.550	-2.550	-2.550
	D1	-2.411	-2.645	-2.730	-2.485	-2.476	-2.708	-2.587	-2.587	-2.587
	D2	-2.577	-3.062	-3.229	-2.581	-2.844	-3.135	-2.608	-2.608	-2.608
	D3	-2.506	-2.834	-2.869	-2.538	-2.651	-2.853	-2.630	-2.630	-2.630
100	NO	-1.979	-1.754	-1.712	-2.076	-1.820	-1.736	-2.114	-2.114	-2.114
	D1	-2.068	-2.393	-2.498	-2.109	-2.227	-2.450	-2.127	-2.127	-2.127
	D2	-2.135	-2.644	-2.789	-2.133	-2.421	-2.706	-2.136	-2.136	-2.136
	D3	-2.101	-2.326	-2.392	-2.131	-2.202	-2.341	-2.157	-2.157	-2.157
1000	NO	-1.774	-1.708	-1.686	-1.807	-1.726	-1.695	-1.805	-1.805	-1.805
	D1	-1.870	-2.240	-2.377	-1.810	-2.049	-2.307	-1.814	-1.814	-1.814
	D2	-1.892	-2.366	-2.489	-1.840	-2.167	-2.409	-1.821	-1.821	-1.821
	D3	-1.848	-2.157	-2.227	-1.829	-2.031	-2.188	-1.818	-1.818	-1.818

Tabla 2.4.3: Valores críticos del contraste MCE extendido. Co-ruptura en niveles, no en diferencias. ($\mu_{y,t} - \mu_{y,t} = 0$).

TABLA 2.4. El PGD se simula bajo $H_0 : \Delta y_t = c_t + a\Delta z_t + u_{1t}$, $\Delta z_t = sD_{j_t} + u_{2t}$, donde $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t}$, $\sigma_1^2 = \text{var}(u_{1t}) = 1$, y $\sigma_2^2 = \text{var}(u_{2t}) = s^2$. El modelo estimado es $\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + d(y_{t-2} - z_{t-2}) + u_t$. Los valores críticos al 5% se calculan para distintos tamaños muestrales (T= 25, 100, 1000), parámetros de corto plazo ($a= 0.0, 0.5, 1.0$), y tamaños del salto ($s=1, 6, 16$).

T	a	b	NO					D3				
			-0.05	-0.1	-0.25	-0.5	-0.75	-0.05	-0.1	-0.25	-0.5	-0.75
25	0.5	s= 1	5.8	7.4	18.9	57.8	91.3	7.4	10.5	26.9	70.0	95.4
		s= 6	14.1	32.5	91.2	100.0	100.0	18.6	44.6	96.5	100.0	100.0
		s= 16	48.8	92.1	100.0	100.0	100.0	61.6	97.3	100.0	100.0	100.0
	1.0	s= 1	6.1	7.6	16.6	49.1	84.9	5.8	7.2	15.8	47.6	84.0
		s= 6	6.1	7.6	16.6	49.1	84.9	5.8	7.2	15.8	47.6	84.0
		s= 16	6.1	7.6	16.6	49.1	84.9	5.8	7.2	15.8	47.6	84.0
100	0.5	s= 1	10.8	23.9	81.6	100.0	100.0	13.5	29.1	86.7	100.0	100.0
		s= 6	44.1	91.6	100.0	100.0	100.0	50.7	94.3	100.0	100.0	100.0
		s= 16	98.6	100.0	100.0	100.0	100.0	99.4	100.0	100.0	100.0	100.0
	1.0	s= 1	9.3	18.2	69.3	99.8	100.0	9.3	18.3	69.1	99.8	100.0
		s= 6	9.3	18.2	69.3	99.8	100.0	9.3	18.3	69.1	99.8	100.0
		s= 16	9.3	18.2	69.3	99.8	100.0	9.3	18.3	69.1	99.8	100.0
1000	0.5	s= 1	51.4	96.1	100.0	100.0	100.0	54.5	96.9	100.0	100.0	100.0
		s= 6	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
		s= 16	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	1.0	s= 1	44.1	92.4	100.0	100.0	100.0	43.9	92.3	100.0	100.0	100.0
		s= 6	44.1	92.4	100.0	100.0	100.0	43.9	92.3	100.0	100.0	100.0
		s= 16	44.1	92.4	100.0	100.0	100.0	43.9	92.3	100.0	100.0	100.0

Tabla 2.5.1: Potencia ajustada del contraste MCE extendido. Co-ruptura simultánea.

T	a	b	NO					D3				
			-0.05	-0.1	-0.25	-0.5	-0.75	-0.05	-0.1	-0.25	-0.5	-0.75
25	0.5	s= 1	4.9	5.1	7.9	19.9	43.9	6.0	7.0	12.6	30.3	57.2
		s= 6	13.1	25.6	65.8	94.1	99.1	9.2	16.0	39.9	71.8	90.4
		s= 16	47.0	88.7	100.0	100.0	100.0	22.4	41.9	67.2	85.1	95.3
	1.0	s= 1	4.8	4.9	7.0	16.4	36.9	4.8	4.9	6.8	16.0	36.1
		s= 6	4.8	4.9	7.0	16.4	36.9	4.8	4.9	6.8	16.0	36.1
		s= 16	4.8	4.9	7.0	16.4	36.9	4.8	4.9	6.8	16.0	36.1
100	0.5	s= 1	6.7	10.4	32.5	85.4	99.7	7.0	10.6	32.7	84.7	99.6
		s= 6	19.3	42.6	96.0	100.0	100.0	19.6	38.6	85.0	99.9	100.0
		s= 16	89.4	99.8	100.0	100.0	100.0	49.4	68.7	94.5	100.0	100.0
	1.0	s= 1	6.3	9.0	26.1	76.1	98.6	6.4	9.1	25.9	75.9	98.5
		s= 6	6.3	9.0	26.1	76.1	98.6	6.4	9.1	25.9	75.9	98.5
		s= 16	6.3	9.0	26.1	76.1	98.6	6.4	9.1	25.9	75.9	98.5
1000	0.5	s= 1	19.1	51.1	99.7	100.0	100.0	18.9	50.4	99.6	100.0	100.0
		s= 6	73.9	99.8	100.0	100.0	100.0	64.2	97.0	100.0	100.0	100.0
		s= 16	100.0	100.0	100.0	100.0	100.0	86.7	99.5	100.0	100.0	100.0
	1.0	s= 1	17.6	44.5	99.1	100.0	100.0	17.4	44.3	99.1	100.0	100.0
		s= 6	17.6	44.5	99.1	100.0	100.0	17.4	44.3	99.1	100.0	100.0
		s= 16	17.6	44.5	99.1	100.0	100.0	17.4	44.3	99.1	100.0	100.0

Tabla 2.5.2: Potencia ajustada del contraste MCE extendido. Co-ruptura en diferencias, no en niveles.

T	a	b	NO					D3				
			-0.05	-0.1	-0.25	-0.5	-0.75	-0.05	-0.1	-0.25	-0.5	-0.75
25	0.5	s= 1	5.8	7.4	18.9	57.8	91.3	5.6	7.3	17.6	53.5	88.5
		s= 6	14.1	32.5	91.2	100.0	100.0	8.9	15.0	41.2	89.8	99.7
		s= 16	48.8	92.1	100.0	100.0	100.0	9.9	16.7	44.2	93.0	99.9
100	0.5	s= 1	10.8	23.9	81.6	100.0	100.0	10.0	21.4	76.8	100.0	100.0
		s= 6	44.1	91.6	100.0	100.0	100.0	17.3	40.4	99.1	100.0	100.0
		s= 16	98.6	100.0	100.0	100.0	100.0	18.8	43.1	99.8	100.0	100.0
1000	0.5	s= 1	51.4	96.1	100.0	100.0	100.0	47.8	95.0	100.0	100.0	100.0
		s= 6	99.9	100.0	100.0	100.0	100.0	83.0	100.0	100.0	100.0	100.0
		s= 16	100.0	100.0	100.0	100.0	100.0	89.2	100.0	100.0	100.0	100.0

Tabla 2.5.3: Potencia ajustada del contraste MCE extendido. Co-ruptura en niveles, no en diferencias.

TABLA 2.5. El PGD se simula bajo $H_1 : b < 0, \Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1,t}$, $c_t = \Delta \mu_{y,t} - a\Delta \mu_{z,t} - b(\mu_{y,t-1} - \mu_{z,t-1})$, $\Delta z_t = sD_{j_t} + u_{2,t}$, con $\sigma_1^2 = var(u_{1,t}) = 1$, y $\sigma_2^2 = var(u_{2,t}) = s^2$. El modelo estimado es $\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + d(y_{t-2} - z_{t-2}) + u_t$. Los valores críticos del contraste son los mostrados en la Tabla 2.4 para distintos tamaños muestrales ($T= 25, 100, 1000$), parámetros de corto plazo ($a= 0.0, 0.5, 1.0$), y tamaños del salto ($s=1, 6, 16$).

$T = 100$, $a = 0.5$, $s = 16$, y D_3 , en el caso de co-ruptura simultánea el valor crítico es -1.64, y para $a = 1$ el valor crítico es -2.16.

En resumen, para tamaños muestrales razonables ($T > 100$) no sería descabellado utilizar los valores críticos con una distribución $N(0, 1)$ para cualquier valor del parámetro a , para cualquier tamaño del salto (s), y para cualquier clase de tendencia segmentada (D_1, D_2, D_3). Esto constituye una magnífica solución (robusta) para los problemas cotidianos de aquéllos que se dedican a la econometría aplicada. Además, la potencia del contraste es muy alta en los tres tipos de co-ruptura analizados, ver Tabla 2.5.

OBSERVACIÓN 4.1. Como era de esperar, en la situación en que no hay ningún tipo de co-ruptura parcial, el contraste MCE extendido no es robusto a cambios estructurales. Sin embargo, los valores críticos del contraste MCE extendido son más robustos que los del MCE habitual en esta situación extrema.

OBSERVACIÓN 4.2. El comportamiento el contraste al incluir términos dinámicos adicionales por medio del criterio SBIC es muy similar en términos de los valores críticos. No obstante, la potencia ajustada se reduce, especialmente en el caso de co-ruptura parcial en diferencias, ver Arranz y Escribano (1998).

Probamos otras especificaciones alternativas, tales como

$$\phi(B)\Delta y_t = c + \theta(B)\Delta z_t + b(y_{t-1} - z_{t-1}) + d(y_{t-2} - z_{t-2}) + u_t$$

$$\phi(B)\Delta y_t = c + a\Delta z_t + \theta(B)(y_{t-1} - z_{t-1}) + u_t$$

$$\Delta y_t = c + \theta(B)\Delta z_t + \phi(B)(y_{t-1} - z_{t-1}) + u_t$$

pero los resultados no fueron mejores. Por tanto, con modelos sobreparametrizados como (4.6), tenemos un trade-off entre estabilidad de los valores críticos y potencia del contraste.

Observación final También se ha analizado el contraste MCE sobre los componentes tendenciales de y_t , y z_t , donde los componentes tendenciales se estiman utilizando los filtros de Hodrick y Prescott (1980, 1997), y Baxter y King (1995, 1999) filters, tal y como se hace en el Capítulo 4. Los resultados de los experimentos son similares a los obtenidos con las series observadas, pero la potencia del contraste es baja, ver Arranz y Escribano (1998).

5. Conclusiones

Se han analizado los efectos de diferentes tipos de cambios estructurales sobre el contraste de no-cointegración MCE cuando no se incluyen variables dummy en

el modelo. Se sabe que los valores críticos dependen del tipo de salto y de otros parámetros molestos (constante, tendencias, etc.). En particular, se estudia la dependencia de los valores críticos del momento en que se produce el salto, (D_1, D_2) , el tamaño del mismo, $(s = 1, 6, 16)$, permitiendo incluso la posibilidad de que haya dos saltos (D_3) en la primera diferencia de la media (tendencias segmentadas en niveles).

El hecho de que los valores críticos dependan de parámetros molestos y que los contrastes MCE habituales tengan poca potencia si hay error de especificación del cambio estructural en el nivel de la serie abre la posibilidad de mejorar la robustez de los resultados utilizando modelos MCE extendidos. El más simple de los modelos MCE extendidos es el que incluye un retardo adicional del mecanismo de corrección del error, $(y_{t-2} - \alpha z_{t-2})$, en el modelo MCE habitual. Nuestros experimentos de Monte Carlo indican que los valores críticos de contraste MCE extendido son muy estables (robustos), siguen una distribución estándar y, si no se incluyen términos dinámicos adicionales, la potencia del contraste MCE extendido es excelente para cualquier tipo de co-ruptura parcial. Evidentemente, estos resultados no son válidos si no hay ningún tipo de co-ruptura. Por tanto, los modelos MCE extendidos representan una metodología de contraste de cointegración alternativa muy interesante en el caso de co-rupturas estructurales.

6. Co-rupturas en niveles y en diferencias.

Sea el PGD dado por las ecuaciones (2.2a)–(2.2c). Hay cuatro casos de interés:

- (1) $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$ y $\mu_{y,t-1} - \alpha\mu_{z,t-1} \neq \text{constante}$, co-ruptura en diferencias.
- (2) $\Delta\mu_{y,t} - a\Delta\mu_{z,t} \neq \text{constante}$ y $\mu_{y,t-1} - \alpha\mu_{z,t-1} = c_l$, co-ruptura en niveles.
- (3) $\Delta\mu_{y,t} - a\Delta\mu_{z,t} \neq \text{constante}$ y $\mu_{y,t-1} - \alpha\mu_{z,t-1} \neq \text{constante}$, no co-ruptura.
- (4) $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$ and $\mu_{y,t-1} - \alpha\mu_{z,t-1} = c_l$, co-ruptura en niveles y en diferencias.

OBSERVACIÓN 6.1. Cuando las pendientes de las tendencias deterministas son constantes, $\Delta\mu_z = g_z$ y $\Delta\mu_y = g_y$, entonces

$$\mu_{y,t-1} - \alpha\mu_{z,t-1} = \mu_y^0 - \alpha\mu_z^0 + (g_y - \alpha g_z)(t-1) \quad (6.1a)$$

$$\begin{aligned} c_t &= (g_y - \alpha g_z) - b(\mu_y^0 - \alpha\mu_z^0) \\ &\quad - b(g_y - \alpha g_z)(t-1) \end{aligned} \quad (6.1b)$$

Para co-ruptura en niveles, la condición necesaria es $g_y - \alpha g_z = 0$, que implica que $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = (\alpha - a)g_z$, que es constante, y se cumple la condición de co-ruptura en diferencias. Por tanto, en el caso de pendientes constantes (no saltos), co-ruptura en niveles implica co-ruptura en diferencias (co-ruptura total).

Por otro lado, en el caso de co-ruptura en diferencias, debe ser $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = g_y - \alpha g_z = c_d$ y, por tanto,

$$\mu_{y,t-1} - \alpha\mu_{z,t-1} = \mu_{y,0} - \alpha\mu_{z,0} + c_d(t-1) + g_z(a - \alpha)(t-1).$$

Por tanto, co-ruptura en diferencias implica co-ruptura en niveles sólo cuando $c_d = 0$ y $a = \alpha$ (restricción COMFAC).

OBSERVACIÓN 6.2. Supongamos que hay un salto, de forma que $\Delta\mu_{z,t} = g_z + s_z D_t$ y $\Delta\mu_y = g_y + s_y D_t$, entonces

$$\begin{aligned} \mu_{y,t-1} - \alpha\mu_{z,t-1} &= \mu_y^0 - \alpha\mu_z^0 + (g_y - \alpha g_z)(t-1) \\ &\quad + (s_y - \alpha s_z)(t-1-t_1)D_{t-1} \end{aligned} \quad (6.2a)$$

$$\begin{aligned} c_t &= (g_y - \alpha g_z) + (s_y - \alpha s_z)D_t - b(\mu_y^0 - \alpha\mu_z^0) \\ &\quad - b(g_y - \alpha g_z)(t-1) - b(s_y - \alpha s_z)(t-1-t_1)D_{t-1} \end{aligned} \quad (6.2b)$$

Las condiciones necesarias para que haya co-ruptura en diferencias son $g_y - ag_z = c_d$ y $s_y - as_z = 0$. En tal caso

$$\begin{aligned} \mu_{y,t-1} - \alpha\mu_{z,t-1} &= \mu_y^0 - \alpha\mu_z^0 + c_d(t-1) \\ &+ (a-\alpha)g_z(t-1) + (a-\alpha)s_z(t-1-t_1)D_{t-1} \end{aligned}$$

y co-ruptura en diferencias implica co-ruptura en niveles sólo si $a = \alpha$ (restricción COMFAC) y $c_d = 0$.

Análogamente, las condiciones necesarias para tener co-rupturas en niveles, son, a partir de la ecuación (6.2a), $g_y - ag_z = 0$ y $s_y - \alpha s_z = 0$. Por tanto, cuando $a = \alpha$, estas condiciones son las necesarias para que haya co-ruptura en diferencias tomando $c_d = 0$. En efecto, bajo co-ruptura en niveles

$$\Delta\mu_{y,t} - a\Delta\mu_{z,t} = (a-\alpha)g_z + (a-\alpha)s_zD_t,$$

y en el caso de que $a = \alpha$ (restricción COMFAC), esta expresión toma el valor 0, por lo que tenemos co-ruptura en diferencias.

Contrastes Bootstrap de Cointegración con Cambios Estructurales: Un Contraste MCE Extendido Robusto.

1. Introduction

Los contrastes de cointegración se ven afectados por parámetros molestos originados por los cambios estructurales de las series, el momento del salto, así como el tamaño del mismo (ver, por ejemplo Kremers et al., 1992, Campos et al., 1996). No obstante, lo habitual es considerar la hipótesis de que los cambios estructurales afectan a ambas series de manera que la relación a corto plazo no se ve afectada por ninguno de los saltos, mostrando co-ruptura simultánea, como se ve en Clements y Hendry (1999).

En el Capítulo 2 se realiza un estudio en muestras pequeñas de los efectos de los cambios estructurales sobre los contrastes de cointegración basados en modelos de corrección del error uniecuacionales (contraste de cointegración MCE) en distintas situaciones de co-ruptura. Se aporta en el mismo la evidencia de que los valores críticos dependen del tamaño del salto y de la correlación contemporánea en la relación condicional de corto plazo incluso cuando hay co-ruptura simultánea, y que esa dependencia se hace aún más patente en el caso de que sólo haya co-ruptura parcial. Se demuestra que el contraste MCE se puede hacer robusto (estable) a algunos tipos de cambio estructural si utilizamos modelos de corrección del error extendidos.

Los métodos intensivos en computación, tales como el bootstrap¹, basados en modelos de corrección del error extendidos pueden, por tanto, proporcionarnos contrastes MCE con buenas propiedades en muestras finitas. Las técnicas bootstrap se han aplicado tradicionalmente en situaciones *iid*, pero también pueden aplicarse a series temporales (see Li y Maddala, 1996, Kiviet, 1984, Giersbergen y Kiviet, 1996, Buhlmann, 1999, Berkowitz y Kilian, 2000). Como se ha puesto de manifiesto recientemente (Mantalos y Shukur, 1998), el contraste MCE basado en bootstrap tiene

¹Ver Davidson y Hinkley (1997), Hall (1992), Efron y Tibshirani (1993), Hall (1994), Shao y Tu (1995) y Horowitz (2000).

buenas propiedades de tamaño y potencia. El propósito de este Capítulo es comprobar que dicha conclusión se mantiene en el caso de co-rupturas estructurales. En concreto, nos vamos a centrar en el análisis del contraste de cointegración MCE bajo co-rupturas estructurales cuando los valores críticos del mismo se obtienen aplicando técnicas bootstrap a los residuos del modelo de corto plazo. Se van a comparar diversos esquemas de remuestreo para obtener los valores críticos bootstrap más fiables para contrastar la hipótesis nula de no cointegración frente a la alternativa de cointegración. Se analizarán el tamaño y la potencia del contraste MCE bootstrap bajo diferentes casos de co-rupturas estructurales: co-ruptura simultánea, co-ruptura parcial en niveles y co-ruptura parcial en diferencias.

La estructura del Capítulo es la siguiente: Los conceptos de co-ruptura simultánea, co-ruptura en niveles y co-ruptura en diferencias se revisan brevemente en la Sección 2. La Sección 3 analiza la puesta en práctica de distintas técnicas de remuestreo y estadísticos bootstrap para contrastar hipótesis que se aplican en modelos de regresión dinámica (ver Li y Maddala, 1996, 1997). En la Sección 4 se analizan los resultados de un experimento de simulación de Monte Carlo, sin cambios estructurales, y la Sección 5 incluye las simulaciones de Monte Carlo cuando hay co-rupturas estructurales. Las conclusiones y las direcciones de investigaciones futuras se recogen en la Sección 6.

2. Modelos de Corrección del Error con Co-rupturas en Niveles y en Diferencias

Consideremos el siguiente modelo de corrección del error (MCE) bivalente

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (2.1a)$$

$$\Delta z_t = \Delta \mu_{z,t} + u_{2t} \quad (2.1b)$$

$$c_t \equiv \Delta \mu_{y,t} - a\Delta \mu_{z,t} - b(\mu_{y,t-1} - \alpha \mu_{z,t-1}) \quad (2.1c)$$

$$\Delta \mu_{z,t} = sD_{j,t} \quad (2.1d)$$

donde $\mu_{y,t} = E(y_t)$, $\mu_{z,t} = E(z_t)$ son, dadas condiciones válidas, las medias incondicionales correspondientes² y, por tanto, incluyen todo tipo de componentes deterministas de y_t y de la variable exógena z_t como: términos independientes, tendencias deterministas, variables tipo dummy, tendencias segmentadas, observaciones atípicas,

²En realidad, son las esperanzas condicionales (medias condicionales) siendo las variables que condicionan las condiciones iniciales, y_0 y z_0 , respectivamente. A partir de ahora, diremos que $E(y_t) = \mu_{y,t}$ y $E(z_t) = \mu_{z,t}$ dadas condiciones válidas.

etc. Los errores estocásticos, u_{1t} and u_{2t} no presentan dependencia conjunta ni serial, tienen media cero, son procesos con distribución de ruido blanco Gaussiano con varianzas constantes $var(u_{1t}) = \sigma_1^2$ y $var(u_{2t}) = \sigma_2^2$. En la ecuación (2.1d), s es el valor del parámetro que mide el tamaño del salto, y $D_{j,t}$ es una variable dummy que toma el valor 0 antes del salto y el valor 1 en el momento del salto y con posterioridad al mismo, ver Sección 3.1 para más detalles. En este Capítulo se van a investigar los efectos de diversos cambios estructurales en c_t al aplicar el contraste MCE de no-cointegración ($b = 0$) mediante bootstrap. Se consideran tres posibles situaciones de co-ruptura a partir de las siguientes definiciones:

DEFINICIÓN 2.1. Sean $E(y_t) = \mu_{y,t}$ y $E(z_t) = \mu_{z,t}$ dadas condiciones iniciales válidas, decimos que las series temporales y_t y z_t presentan *co-rupturas en diferencias* si $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$, donde c_d es un parámetro constante finito.

DEFINICIÓN 2.2. Sean $E(y_t) = \mu_{y,t}$ y $E(z_t) = \mu_{z,t}$ dadas condiciones iniciales válidas, decimos que las series temporales y_t y z_t presentan *co-rupturas en niveles* si $\mu_{y,t} - \alpha\mu_{z,t} = c_l$, donde c_l es un parámetro constante finito.

DEFINICIÓN 2.3. Sean $E(y_t) = \mu_{y,t}$ y $E(z_t) = \mu_{z,t}$ dadas condiciones iniciales válidas, decimos *co-ruptura simultánea* si $\Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t} - \alpha\mu_{z,t}) = c_s$, donde c_s es un parámetro constante finito.

Es evidente que cuando y_t y z_t presentan *co-ruptura en niveles y en diferencias* (*co-ruptura completa*), tendremos co-ruptura simultánea.

En el caso de co-ruptura simultánea, el término independiente, c_t de (2.1c) es constante, $c_t = c$, y el modelo de corrección del error de (2.1a) es el MCE condicional estándar en el que el único regresor determinista es el término constante, c .

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (2.2)$$

en nuestro experimento de simulación de Monte Carlo se analizan varios posibles casos intermedios que son interesante desde el punto de vista aplicado.

CASO 2.1. *Co-ruptura en niveles, pero no en diferencias.*

Co-ruptura en niveles implica $\mu_{y,t} - \alpha\mu_{z,t} = c_l$. tomando primeras diferencias, se tiene $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = 0$, y la ecuación (2.1a) se reescribe como

$$\Delta y_t = -bc_l + (\alpha - a)sD_{j,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \quad (2.3)$$

OBSERVACIÓN 2.1. Los saltos en el proceso marginal de Δz_t afectan al modelo de corrección del error y, por tanto, al contraste MCE, bajo las hipótesis nula y la alternativa, a no ser que se satisfaga la restricción de factores comunes (COMFAC) ($a = \alpha$).

OBSERVACIÓN 2.2. Recordemos que co-ruptura en niveles implica co-ruptura en diferencias (y por tanto co-ruptura completa), cuando se satisface la restricción COMFAC ($a - \alpha = 0$) (ver Capítulo 2).

CASO 2.2. *Co-ruptura en diferencias, pero no en niveles.*

Co-ruptura en diferencias implica que $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$. A partir de la ecuación (2.1a)

$$\begin{aligned} \Delta y_t &= c_m + bc_d t - b(a - \alpha)\mu_{z,t-1} \\ &+ a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \end{aligned} \quad (2.4a)$$

$$\mu_{z,t} = \mu_{z,0} + s \sum_{i=1}^t D_{j,i} \quad (2.4b)$$

donde c_m es una constante igual a $c_m = c_d - b(\mu_{y,0} - \alpha\mu_{z,0}) + bc_d$.

OBSERVACIÓN 2.3. El salto en el proceso marginal de z_t afectan a modelo de corrección del error sólo bajo la hipótesis alternativa, $H_1 : b < 0$.

OBSERVACIÓN 2.4. Recordemos que co-ruptura en diferencias implica co-ruptura en niveles en el caso de que se cumpla la restricción COMFAC ($a - \alpha = 0$) y $c_d = 0$ (ver Capítulo 2).

Dependiendo del tipo de variable dummy considerada, $D_{j,t}$, se pueden tener en el modelo (2.1a) tendencias segmentadas con uno o más puntos de ruptura, ver Sección 5 para más detalles.

El modelo de corrección extendido (MCE extendido) es el siguiente:

$$\Delta y_t = c + a\Delta z_t + b[y_{t-1} - \alpha z_{t-1}] + d[y_{t-2} - \alpha z_{t-2}] + \eta_t. \quad (2.5)$$

Una propiedad importante del contraste MCE analizado en el Capítulo 2, es que, al utilizar un modelo de corrección del error extendido, ecuación (2.5), ver (Dolado y Lutkepohl, 1996, Toda y Yamamoto, 1995), los valores críticos se hacen robustos (estables) en el sentido de que no dependen tanto de parámetros molestos, y la potencia del contraste es mayor que con el contraste MCE habitual si hay al menos co-ruptura parcial. Esto es lo que nos lleva a pensar en una técnica de contraste alternativa que se pueda utilizar en muestras pequeñas y que no dependa tanto del tipo proceso de co-ruptura considerado, y el bootstrap es un buen candidato.

3. Bootstrap de Modelos de Corrección del Error

Los dos primeros esquemas de remuestreo fueron inicialmente propuestos por Giersbergen y Kiviet (1993, 1994) al analizar las dos reglas sugeridas por Hall y Wilson (1991) en el contexto del contraste de hipótesis en modelos de regresión dinámicos. Para hacer más clara la exposición de los esquemas de remuestreo, se utilizarán directamente los modelos en el contexto del contraste MCE que se utilizará en las Secciones 4 y 5.

Tomemos el siguiente modelo de regresión dinámico en forma de corrección del error, con $\alpha = 1$,

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + \epsilon_t. \quad (3.1)$$

Nos interesa contrastar la hipótesis nula $H_0 : b = 0$ (no cointegración) vs. $H_1 : b < 0$ (cointegración), con el estadístico t -ratio de b , $t(\hat{b})$. Sean \hat{a} , \hat{b} , y \hat{c} los estimadores MCO de a , b , and c , respectivamente, y sean ϵ_t^* los residuos bootstrap obtenidos al remuestrear los residuos MCO, $\hat{\epsilon}_t$, o los residuos MCO centrados en el caso de que la regresión estimada no incluyera término constante. NB indica el número de remuestreos bootstrap.

Vamos a considerar dos esquemas de remuestreo alternativos S_1 , y S_2 para generar la muestras bootstrap:

$$S_1 : \Delta y_t^* = \hat{c} + \hat{a}\Delta z_t + \hat{b}(y_{t-1}^* - z_{t-1}) + \epsilon_t^* \quad (3.2)$$

$$S_2 : \Delta y_t^* = \hat{c} + \hat{a}\Delta z_t + \epsilon_t^*. \quad (3.3)$$

Ambos procedimientos utilizan ϵ_t^* pero difieren en la forma de generar y_t^* . Para cada procedimiento de remuestreo se consideran los estadísticos t :

$$t_1 : t(\hat{b}) = (\hat{b}^* - \hat{b}) / SE(\hat{b}^*) \quad (3.4)$$

$$t_2 : t(b_0) = \hat{b}^* / SE(\hat{b}^*) \quad (3.5)$$

donde \hat{b}^* es el estimador MCO de b^* en

$$\Delta y_t^* = c^* + a^*\Delta z_t + b^*(y_{t-1}^* - z_{t-1}) + \epsilon_t^* \quad (3.6)$$

y $SE(\hat{b}^*)$, el standard error de \hat{b}^* , se calcula bien a partir de la distribución empírica de \hat{b}^* , es decir, tras realizar los NB remuestreos o utilizando en cada regresión la fórmula habitual de MCO, que es el enfoque seguido en nuestro análisis. El valor crítico del bootstrap se obtiene con el cuantil 5% de la cola inferior de la distribución de $t(\hat{b})$ o $t(b_0)$, al utilizar t_1 o t_2 , respectivamente.

Por tanto, se pueden definir cuatro versiones del estadístico t . Hall y Wilson (1991) sólo considera el esquema S_1 y recomienda utilizar sólo t_1 . No consideran la utilización del esquema S_2 , el apropiado para t_2 . Giersbergen y Kiviet (1993) recomiendan, a partir de un estudio de Monte Carlo de un modelo $AR(1)$, el uso de t_2 bajo S_2 en lugar de t_1 bajo S_1 . Sus conclusiones principales son:

- (1) La inferencia basada en t_2 bajo S_1 tiene poca potencia, y su tamaño es cercano a cero. Análogamente, no debe usarse t_1 bajo S_2 . Por tanto, el esquema de remuestreo debe ser análogo a la distribución del estadístico de bootstrap bajo la nula.
- (2) Usar t_1 bajo S_1 y t_2 bajo S_2 es equivalente en modelos sin dinámica. No obstante, en modelos dinámicos esta equivalencia no es tal en muestras pequeñas. Giersbergen y Kiviet recomiendan el uso de t_2 bajo S_2 .
- (3) Las distribuciones límite de t_1 bajo S_1 y t_2 bajo S_2 son idénticas incluso con modelos dinámicos. La conclusión de que es mejor utilizar t_2 bajo S_2 se obtiene en este contexto a partir de su comportamiento en muestras finitas.

Giersbergen y Kiviet (1994), y Li y Maddala (1996, 1997) proponen un esquema de remuestreo diferente, basado en los residuos MCO del modelo restringido, $\epsilon_0 = \Delta y^* - \hat{c}^r + \hat{a}^r \Delta z_t$, y \hat{a}^r , \hat{b}^r , y \hat{c}^r , son los estimadores MCO del MCE bajo $H_0 : b = 0$,

$$\Delta y_t = c + a \Delta z_t + u_t \quad (3.7)$$

$$S_3 : \Delta y^* = \hat{c}^r + \hat{a}^r \Delta z_t + \epsilon_0^* \quad (3.8)$$

donde ϵ_0^* es la muestra bootstrap de ϵ_0 . Obsérvese que ambos procedimientos, S_1 y S_2 utilizan remuestreos basados en los residuos MCO bajo $H_1 : b < 0$, $\hat{\epsilon}$. Si la hipótesis nula $H_0 : b = 0$ es cierta, pero el estimador MCO de b está lejos de 0, la distribución empírica de los residuos será una mala aproximación de la distribución de los errores bajo la nula. La intuición tras el procedimiento S_3 es la siguiente: si la hipótesis nula es cierta, ϵ_0 tiene exactamente la verdadera distribución de los errores de regresión, y el contraste de hipótesis basado en los mismos tendrá (aproximadamente) el tamaño correcto. Si la hipótesis nula no es cierta, ϵ_0 no tendrá la verdadera distribución de los errores. El contraste de hipótesis basado en t_1 o t_2 tendrá potencia alta dependiendo de lo que se aleje de la hipótesis nula el verdadero valor de b . Por tanto, es de esperar que t_2 bajo S_3 funcione mejor que t_1 bajo S_1 o t_2 bajo S_2 .

Otros esquemas de remuestreo están basados en los anteriores e incorporan estimadores con corrección del sesgo. En concreto, para el caso S_1 se simulan primero H muestras de y^* siendo $c_j, a_j, b_j, j = 1, \dots, H$ los estimadores bootstrap de c, a, b . Los

estimadores con corrección de sesgo son

$$\tilde{c} = \hat{c} + \left(\hat{c} - \frac{\sum_{i=1}^H c_j}{H} \right) \quad \tilde{a} = \hat{a} + \left(\hat{a} - \frac{\sum_{i=1}^H a_j}{H} \right) \quad \tilde{b} = \hat{b} + \left(\hat{b} - \frac{\sum_{i=1}^H b_j}{H} \right) \quad (3.9)$$

y el esquem de remuestreo S_1^* se define

$$S_1^* : \Delta y^* = \tilde{c} + \tilde{a}\Delta z_t + \tilde{b}(y_{t-1}^* - \alpha z_{t-1}) + \epsilon_0^*. \quad (3.10)$$

Los procedimientos de corrección de sesgo S_2^* y S_3^* se definen de forma análoga a partir de las definiciones de S_2 y S_3 .

4. Experimento de Monte Carlo I. No saltos.

Consideremos el siguiente proceso generador de datos (PGD) bivalente,

$$\Delta y_t = a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (4.1a)$$

$$\Delta z_t = u_{2t} \quad (4.1b)$$

donde

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim IIN \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & s^2 \end{bmatrix} \right). \quad (4.1c)$$

Nuestro experimento es una diseño factorial con $a = 0.0, 0.5, 1$ (correlación contemporánea), $b = 0.0$ (no cointegración), $-0.05, -0.1, -0.25, -0.5, -0.75$ (cointegración), $s = 1, 6, 16$, $T = 25, 50, 100, 200, 500, 1000$ (tamaño muestral).

Para obtener el tamaño empírico del contraste, se simulan las series y_t y z_t con el PGD (4.1a)–(4.1c) con $b = 0$ y se estima la ecuación (3.1). La cola inferior al 5% de la distribución bootstrap del estadístico $t(\hat{b})$ es el valor crítico empírico. La potencia empírica del contraste se calcula análogamente, simulando las series con los valores de b bajo la alternativa ($b \neq 0$), a partir del porcentaje de rechazos. Para mayor simplicidad, y sin pérdida de generalidad, se considera el vector de cointegración conocido $(1, -\alpha) = (1, -1)$.

Los resultados del Monte Carlo de las técnicas bootstrap se basan en $M = 10000$ réplicas. En cada réplica de Monte Carlo, los valores críticos bootstrap se calculan a partir de $NB = 200$ muestras bootstrap del estadístico de contraste. En los casos en que se haga corrección del sesgo, se toma $H = 200$ en (3.9). En todos los experimentos se eliminan las primeras 50 observaciones de las series para considerar condiciones iniciales aleatorias.

T	a	S_1, t_1			S_2, t_2			S_3, t_2		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0.0	5.67	5.25	5.34	4.89	5.25	5.36	5.15	5.25	5.47
	0.5	6.36	5.17	5.31	4.96	5.11	5.30	5.08	5.09	5.35
	1.0	6.60	6.60	6.60	5.21	5.21	5.21	5.14	5.14	5.14
50	0.0	5.59	4.96	5.11	5.03	4.98	5.12	4.91	4.94	5.11
	0.5	6.24	5.14	4.99	5.09	4.97	5.10	5.05	4.85	5.03
	1.0	7.03	7.03	7.03	5.42	5.42	5.42	5.35	5.35	5.35
100	0.0	5.80	5.12	5.11	4.88	5.01	5.13	4.92	5.13	5.21
	0.5	6.28	5.14	5.14	4.95	4.88	5.10	4.92	4.94	5.20
	1.0	6.56	6.56	6.56	5.33	5.33	5.33	5.26	5.26	5.26
200	0.0	5.61	4.80	4.68	4.84	4.78	4.70	4.89	4.85	4.78
	0.5	5.87	5.03	4.73	4.55	4.86	4.65	4.51	4.90	4.69
	1.0	6.37	6.37	6.37	4.92	4.92	4.92	4.96	4.96	4.96

TABLA 3.1. Tamaño del contraste bootstrap, t -stat. No saltos. No intercepto. Modelo (4.1a)–(4.1c)

Se comienza estimando el mismo modelo que Mantalos y Shukur (1998), ecuación (4.1a), y los resultados se muestran en la Tabla 3.1. Mantalos y Shukur (1998), utilizando un espacio de parámetros que es un subconjunto del usado es nuestro análisis, concluye que el tamaño del contraste está alrededor del tamaño nominal para muestras con una tamaño igual a 20 observaciones si el estadístico t del bootstrap se calcula en cada regresión con el esquema $S_2 - t_2$, a partir de 10000 réplicas de Monte Carlo y a partir de $NB = 200$ réplicas bootstrap. Como se observa en la Tabla 3.1, nuestra conclusión es congruente para S_2, t_2 , y S_3, t_2 . sin embargo, para S_1, t_1 necesitamos muestras de tamaño mayor para conseguir tamaños próximos al nominal (5%). También encontramos que los resultados no son sensibles al valor de NB . Más aún, no se observan diferencias significativas si usamos las versiones del bootstrap con correcciones del sesgo. Lo que sí es de destacar es que el bootstrap no resulta adecuado si utilizamos la distribución empírica de b^* .

En la mayoría de nuestras simulaciones de Monte Carlo se opta por incluir un término constante en la regresión para hacer que los resultados no sean sensibles a las condiciones iniciales de las variables. También se incluye un término constante en la mayoría de las aplicaciones empíricas para tener en cuenta los valores medios de las variables de la regresión, y_t and z_t . ¿Influye en los resultados del contraste bootstrap la inclusión del término constante? Para responder a esa pregunta se estima el modelo MCE,

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_t \quad (4.2)$$

y los resultados de las simulaciones se recogen en la Tabla 3.2.1. Como puede verse, la inclusión del intercepto en el modelo crea importantes distorsiones de tamaño,

T	a	S_1, t_1			S_2, t_2			S_3, t_2		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0.0	12.06	5.45	4.92	15.60	5.71	4.97	5.34	5.08	4.89
	0.5	17.25	6.73	5.23	21.55	7.48	5.35	5.50	5.06	4.97
	1.0	21.20	21.20	21.20	26.32	26.32	26.32	6.20	6.20	6.20
50	0.0	11.78	5.40	5.08	14.79	5.64	5.01	5.34	4.96	5.04
	0.5	16.41	6.57	5.26	20.63	7.13	5.31	5.48	4.87	4.98
	1.0	20.41	20.41	20.41	25.47	25.47	25.47	5.93	5.93	5.93
100	0.0	12.09	5.49	4.92	14.02	5.45	4.91	5.00	4.82	4.93
	0.5	16.67	6.70	5.17	20.35	6.79	5.28	5.52	4.91	4.86
	1.0	19.87	19.87	19.87	23.90	23.90	23.90	6.01	6.01	6.01
200	0.0	12.18	5.32	4.66	13.43	5.27	4.63	5.03	4.80	4.64
	0.5	16.34	6.41	5.05	18.89	6.59	4.90	5.40	4.91	4.75
	1.0	20.53	20.53	20.53	23.36	23.36	23.36	5.89	5.89	5.89

Tabla 3.2.1: Modelo estimado (4.2). MCE

T	a	S_1, t_1			S_2, t_2			S_3, t_2		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0.0	8.88	5.08	4.98	8.09	5.04	4.83	5.18	5.14	4.99
	0.5	10.81	5.81	4.21	8.34	5.41	4.07	5.66	4.99	5.04
	1.0	13.48	13.49	13.45	11.10	11.30	11.25	6.10	6.10	6.10
50	0.0	7.09	4.83	4.66	6.90	4.66	4.62	4.85	5.27	5.08
	0.5	9.07	4.97	3.81	8.12	4.93	3.76	5.31	5.08	5.23
	1.0	10.56	10.61	10.63	9.79	9.79	9.81	5.70	5.70	5.70
100	0.0	6.28	4.99	4.94	6.10	4.91	4.97	4.84	4.66	4.75
	0.5	7.59	4.79	4.10	7.11	4.63	4.05	5.19	4.61	4.70
	1.0	8.46	8.48	8.50	8.12	8.21	8.23	5.22	5.22	5.22
200	0.0	5.96	5.20	5.17	5.92	5.18	5.19	4.88	4.78	4.89
	0.5	6.64	4.48	4.17	6.46	4.54	4.07	5.08	4.63	4.80
	1.0	7.13	7.12	7.14	7.03	7.03	7.02	5.19	5.19	5.19

Tabla 3.2.2: Modelo estimado (4.3). MCE Extendido

TABLA 3.2. Tamaño del contraste bootstrap test. Modelo simulado (4.1a)–(4.1c). No saltos.

especialmente para S_1, t_1 , y S_2, t_2 , dando como resultado tamaños del contraste de hasta 25%. En el caso S_3, t_2 , la mayor desviación del valor nominal se produce para el tamaño muestral 25, que vale 6%. Este resultado es importante a favor del uso del método bootstrap basado S_3, t_2 .

Estimamos el modelo MCE model

$$\Delta y_t = c + a\Delta z_t + b[y_{t-1} - z_{t-1}] + d[y_{t-2} - z_{t-2}] + u_t, \quad (4.3)$$

para comprobar si el contraste MCE bootstrap es más robusto a la presencia de la constante en el modelo. Los resultados aparecen en la Tabla 3.2.2 y muestran que con todos los esquemas de remuestreo se consiguen unos tamaños más próximos al

valor nominal. Las mayores distorsiones se reducen de 24% a 11%, y, de nuevo, S_3, t_2 domina claramente. Con S_3, t_2 el tamaño del contraste nunca es superior a 6.1% para muestras con 25 observaciones. Esto constituye un resultado interesante. Hemos demostrado que, incluso si el modelo no tiene cambios estructurales, en el caso de que dudemos al incluir una constante en el modelo, debemos utilizar el modelo MCE extendido (4.3) para asegurarnos de que el tamaño del contraste está próximo al valor nominal.

Por otra parte, en el caso de que se omita el término independiente y el PGD lo incluya, el contraste bootstrap es inadecuado, y la inclusión de un retardo adicional del mecanismo de corrección del error no soluciona el problema del error de especificación. Así pues, nuestra recomendación a la hora de realizar trabajos empíricos aplicados es estimar el modelo (4.3) en lugar de (4.2).

La corrección del sesgo supone una mejora sólo si calculamos el estadístico t con la desviación típica empírica de los estimadores bootstrap en lugar de usar el estadístico t habitual. No obstante, la ganancia es mínima y no vamos a mostrar los resultados (que están disponibles),

Modelo estimado		MCE, eq.(4.2)			Extended MCE, eq. (4.3)		
T	a	s=1	s=6	s=16	s=1	s=6	s=16
25	0.0	100.00	100.00	100.00	23.60	99.45	100.00
	0.5	100.00	100.00	100.00	16.60	77.75	99.95
	1.0	100.00	100.00	100.00	13.50	13.50	13.50
50	0.0	100.00	100.00	100.00	48.95	100.00	100.00
	0.5	100.00	100.00	100.00	32.05	98.30	100.00
	1.0	100.00	100.00	100.00	25.35	25.35	25.35
100	0.0	100.00	100.00	100.00	80.90	100.00	100.00
	0.5	100.00	100.00	100.00	61.10	100.00	100.00
	1.0	100.00	100.00	100.00	50.50	50.50	50.50
200	0.0	100.00	100.00	100.00	98.15	100.00	100.00
	0.5	100.00	100.00	100.00	90.85	100.00	100.00
	1.0	100.00	100.00	100.00	83.50	83.50	83.50

TABLA 3.3. Potencia del contraste bootstrap, S_3, t_2 . Modelo simulado (4.1a)–(4.1c). $b = -0.2$. No saltos.

Dados los resultados obtenidos, nos vamos a centrar en el procedimiento bootstrap S_3, t_2 , que se basa en el remuestreo de los residuos obtenidos bajo $H_0 : b = 0$. Nuestro objetivo es analizar la potencia de contraste para los modelos (4.2) y (4.3). La potencia del contraste se obtiene a partir del porcentaje de rechazos de la hipótesis nula con $M = 5000$ y $NB = 5000$. En la Tabla 3.3 se muestran los resultados de las potencias estimadas cuando no hay cambios estructurales. La mayor potencia se obtiene para el modelo (4.2). cuando no hay factor común ($a \neq 1$), el modelo (4.3) tiene potencia alta para tamaños muestrales moderados ($T > 50$). Por el contrario, cuando se satisface la restricción de factor común, necesitamos muestras con más de 200 observaciones para obtener potencias de 90% o más. Estos resultados son congruentes con los obtenidos

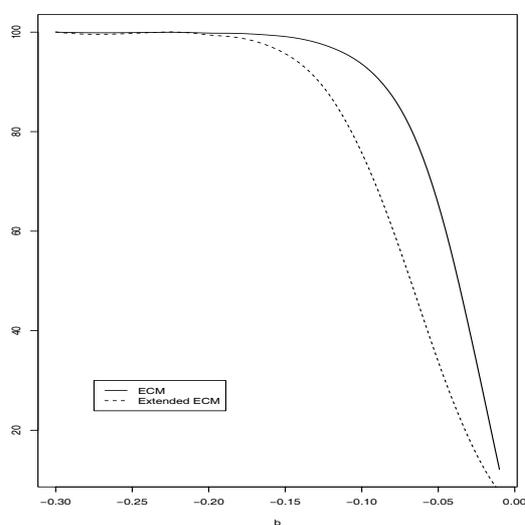


Figura 3.1.1: T=25

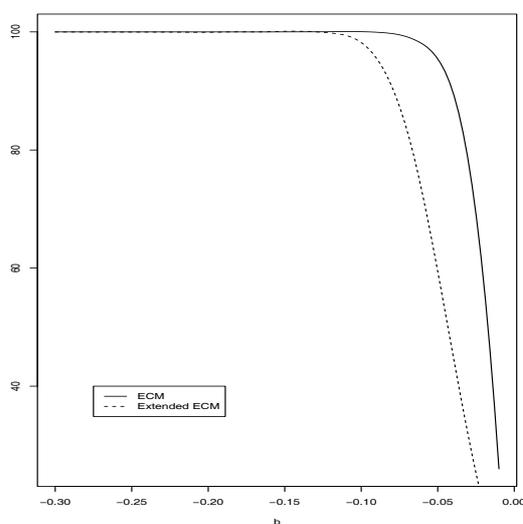


Figura 3.1.2: T=50

FIGURA 3.1. Potencia del contraste bootstrap MCE y MCE Extendido.

por Arranz y Escribano (2000) y Zivot (2000). La Figura 3.1 muestra la potencia del contraste bootstrap MCE, $t(b)$, para el modelo MCE habitual, ecuación (4.2), y el modelo MCE extendido, ecuación (4.3), para $a = 0.0$, $s = 1$, y diferentes valores e b . La potencia del contraste MCE habitual nunca es menor que la de MCE extendido,

pero para valores de b menores que $b = -0.1$, el comportamiento de ambos contrastes es muy similar (potencia próxima a 100%).

5. Experimento de Simulación de Monte Carlo II: Co-rupturas

5.1. MCE con Co-ruptura simultánea. Nuestro proces generador de datos (PGD) se basa en el usado en Capítulo 2, que es una extensión del propuesto por Kremers et al. (1992) y Campos et al. (1996). Se trata de una auto-regresión vectorial lineal de primer orden con perturbaciones Normales, causalidad de Granger unidireccional ($z \rightarrow y$), y varios cambios estructurales en la variable exógena en sentido fuerte, Δz_t , donde los parámetros de interés son los que aparecen en la ecuación (5.1a). Sea el siguiente modelo,

$$\Delta y_t = a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (5.1a)$$

$$\Delta z_t = sD_{jt} + u_{2t} \quad (5.1b)$$

con

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim IIN \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \right). \quad (5.1c)$$

Obsérvese que bajo co-ruptura simultánea, hay saltos en el proceso marginal de Δz_t , ecuación (4.1b), pero no en el proceso condicional, ecuación (4.1a). Se analizan tres tipos de variables dummy para simular un único salto en la tendencia determinista (tendencias segmentadas), en dos puntos diferentes ($T/4$ o $T/2$) siendo T el tamaño de la muestra,

$$D_{1t} = \begin{cases} 1 & t \geq T/4 \\ 0 & \text{caso contrario} \end{cases} \quad D_{2t} = \begin{cases} 1 & t \geq T/2 \\ 0 & \text{caso contrario} \end{cases}$$

y dos saltos en los puntos $T/4$ y $3T/4$.

$$D_{3t} = \begin{cases} 1 & T/4 \leq t \leq 3T/4 \\ 0 & \text{caso contrario} \end{cases}$$

Basándonos en (5.1a)–(5.1c), el tamaño del contraste, ver Tabla 3.4 ($b = 0$), y la potencia ($b \neq 0$) del contraste se obtienen mediante experimentos de simulación de Monte Carlo. Sin pérdida de generalidad, se toma $\sigma_1^2 = 1$ y $\alpha = 1$. Por tanto, las variables del diseño del experimento son los parámetros a, b, s (obsérvese que $\sigma_2 = s$), donde s es el coeficiente de la variable dummy $D_{j,t}$ of (5.1b), y T es el tamaño de la muestra.

T	a	D_1			D_2			D_3		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.32	5.14	5.09	5.12	5.12	5.13	5.06	5.16	5.17
	0.5	5.01	5.26	5.08	5.58	5.05	5.10	4.81	5.20	5.18
	1	6.43	6.43	6.43	6.56	6.56	6.56	6.45	6.45	6.45
50	0	4.97	4.84	4.88	5.13	5.01	4.97	4.80	4.76	4.74
	0.5	4.78	4.85	4.88	4.90	4.99	5.05	4.76	4.70	4.78
	1	5.70	5.70	5.70	5.94	5.94	5.94	5.81	5.81	5.81
100	0	4.36	4.44	4.43	4.55	4.60	4.71	4.64	4.62	4.63
	0.5	4.17	4.39	4.42	4.40	4.51	4.68	4.24	4.71	4.62
	1	5.97	5.97	5.97	6.05	6.05	6.05	5.97	5.97	5.97
200	0	4.88	4.97	4.96	4.74	4.86	4.89	4.73	4.81	4.86
	0.5	4.71	4.95	4.97	4.59	4.79	4.87	4.62	4.93	4.84
	1	5.83	5.83	5.83	5.69	5.69	5.69	5.94	5.94	5.94

Tabla 3.4.1: Modelo estimado (4.2). Modelo MCE.

T	a	D_1			D_2			D_3		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.39	5.06	5.08	5.23	5.23	5.25	5.18	5.05	4.91
	0.5	5.10	5.14	5.14	5.22	5.31	5.20	5.14	5.08	5.04
	1	6.15	6.15	6.15	6.20	6.20	6.20	6.09	6.09	6.09
50	0	4.86	4.89	5.12	4.73	5.22	5.21	4.85	5.01	5.11
	0.5	4.71	4.98	4.94	4.63	5.09	5.23	4.61	4.89	5.03
	1	5.79	5.79	5.79	5.81	5.81	5.81	5.88	5.88	5.88
100	0	4.50	4.60	4.79	4.68	4.90	4.94	4.42	4.65	4.66
	0.5	4.46	4.63	4.70	4.27	4.80	4.86	4.46	4.43	4.60
	1	5.14	5.14	5.14	5.30	5.30	5.30	5.27	5.27	5.27
200	0	4.62	4.77	4.77	4.83	4.69	4.93	4.78	4.74	4.84
	0.5	4.99	4.87	4.82	4.84	4.62	4.78	4.93	4.74	4.72
	1	5.36	5.36	5.36	5.38	5.38	5.38	5.30	5.30	5.30

Tabla 3.4.2: Modelo estimado (4.3). Modelo MCE Extendido

TABLA 3.4. Tamaño empírico, procedimiento bootstrap S_3, t_2 . Co-ruptura simultánea, estadístico t .

Los tamaños de contraste MCE bootstrap cuando las series presentan co-ruptura simultánea y estimamos el modelo MCE habitual (4.2) son altos, especialmente cuando se cumple la restricción COMFAC y para tamaños muestrales de 200 observaciones o menos, ver Tabla 3.4.1. Este problema lo causa la inclusión del término constante en el modelo estimado. Nuevamente, el problema se soluciona en parte si utilizamos el método de remuestreo S_3, t_2 .

Para el modelo MCE extendido (4.3), el tamaño del contraste se muestra en la Tabla 3.4.2. Se observa que los tamaños del contraste están alrededor del valor nominal, es decir, que hay una mejora marginal al utilizar el modelo MCE extendido (4.3) en lugar del modelo MCE habitual (4.2).

T	a	D ₁			D ₂			D ₃		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	61.70	99.95	100.00	60.75	100.00	100.00	56.15	99.95	100.00
	0.5	25.40	99.50	100.00	25.55	99.30	100.00	23.85	98.80	100.00
	1	14.65	14.65	14.65	15.15	15.15	15.15	14.80	14.80	14.80
50	0	96.95	100.00	100.00	97.80	100.00	100.00	96.75	100.00	100.00
	0.5	68.60	100.00	100.00	70.60	100.00	100.00	66.65	100.00	100.00
	1	34.80	34.80	34.80	34.40	34.40	34.40	34.90	34.90	34.90
100	0	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	0.5	99.35	100.00	100.00	99.35	100.00	100.00	99.25	100.00	100.00
	1	86.25	86.25	86.25	85.65	85.65	85.65	85.50	85.50	85.50
200	0	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	0.5	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	1	100.00	100.00	100.00	99.95	99.95	99.95	100.00	100.00	100.00

Tabla 3.5.1: Modelo estimado (4.2). Modelo MCE

T	a	D ₁			D ₂			D ₃		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	26.15	99.95	100.00	27.50	99.85	100.00	26.75	99.90	100.00
	0.5	16.50	83.35	100.00	17.40	84.40	100.00	16.85	85.55	100.00
	1	13.65	13.65	13.65	14.35	14.35	14.35	14.35	14.35	14.35
50	0	53.45	100.00	100.00	55.25	100.00	100.00	54.20	100.00	100.00
	0.5	33.05	99.25	100.00	35.25	99.45	100.00	33.35	99.45	100.00
	1	25.65	25.65	25.65	25.75	25.75	25.75	25.60	25.60	25.60
100	0	84.95	100.00	100.00	86.30	100.00	100.00	86.05	100.00	100.00
	0.5	64.55	100.00	100.00	65.40	100.00	100.00	65.90	100.00	100.00
	1	50.30	50.30	50.30	50.10	50.10	50.10	50.15	50.15	50.15
200	0	98.90	100.00	100.00	99.00	100.00	100.00	99.10	100.00	100.00
	0.5	91.75	100.00	100.00	92.40	100.00	100.00	92.10	100.00	100.00
	1	83.50	83.50	83.50	83.60	83.60	83.60	83.35	83.35	83.35

Tabla 3.5.2: Modelo estimado (4.3). Modelo MCE Extendido.

TABLA 3.5. Potencia del contraste basado en el procedimiento bootstrap S_3, t_2 . Co-ruptura simultánea. $b = -0.2$

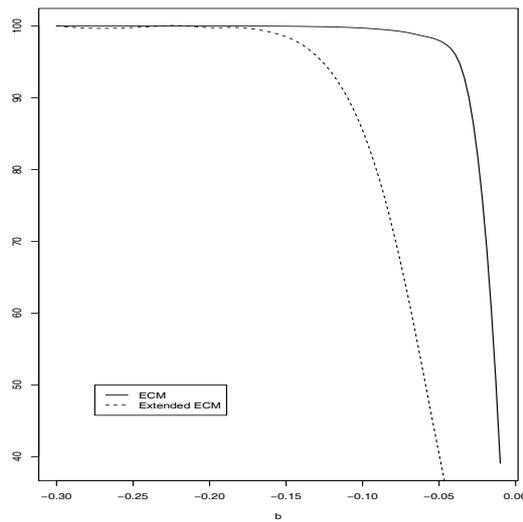


FIGURA 3.2. Potencia del contraste con co-ruptura simultánea, D₃, $T = 25$, $a = 0$, y $s = 6$.

En términos de potencia, en caso de co-ruptura simultánea el modelo (4.2) es mejor que el modelo (4.3), ver Tabla 3.5. Ambos modelos tienen un buen comportamiento en muestras de tamaño moderado, $T \geq 100$. Como era de esperar, la potencia más baja se obtiene con $a = 1$, debido a la restricción de factor común. Recordemos que para $a = 1$ la distribución límite de $t(\hat{b})$ es la distribución Dickey-Fuller y, por tanto, no se ve afectada por el tipo de saltos analizados, (D_1, D_2, D_3) ni por el tamaño de los mismos si tenemos co-ruptura simultánea.

Si $a \neq 1$, la potencia del contraste aumenta con el tamaño del salto (s), puesto que se incrementa el ratio señal-ruido q , donde $q = (1 - a)s/\sigma_1$.

Finalmente, la Figura 3.2 muestra que la potencia de contraste bootstrap es alta (cercana a 100%) para ambos modelos, el MCE, ecuación (4.2), y el MCE extendido, ecuación (4.3), si el parámetro b toma valores menores que $b = -0.15$, y hay co-ruptura simultánea.

5.2. MCE con Co-rupturas Parciales. La co-ruptura parcial se caracteriza por el hecho de tener saltos en el proceso marginal Δz_t , ecuación (2.1d), que afectan al proceso condicional, ecuación (2.1a), a partir de dos fuentes alternativas que afectan a c_t , ecuación (2.1c): bien a través de $\Delta\mu_{y,t} - a\Delta\mu_{z,t}$ (no co-ruptura en diferencias) o a través de $(\mu_{y,t-1} - \alpha\mu_{z,t-1})$ (no co-ruptura en niveles).

El PGD se basa en el siguiente modelo de corrección del error,

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (5.2a)$$

$$\Delta z_t = \Delta\mu_{z,t} + u_{2t} \quad (5.2b)$$

$$c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t-1} - \alpha\mu_{z,t-1}) \quad (5.2c)$$

$$\Delta\mu_{z,t} = sD_{j,t} \quad (5.2d)$$

donde $D_{j,t}$ son las variables dummy utilizadas previamente, ver Sección 5.1. Para generar co-ruptura en diferencias pero no en niveles, hacemos $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d = 0.5$. Si queremos generar series con co-ruptura en niveles pero no en diferencias, imponemos la condición $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = 0$ (recordemos que hacemos $\alpha = 1$).

5.2.1. *Co-ruptura en diferencias, no en niveles.* En este caso, la ecuación (5.2a) puede escribirse como

$$\begin{aligned} \Delta y_t = & c_m + bc_d t - b(a - \alpha)s \sum_{i=1}^t D_{j,i} \\ & + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \end{aligned} \quad (5.3)$$

T	a	D ₁			D ₂			D ₃		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.46	5.18	5.07	5.62	5.33	5.26	5.19	5.10	4.91
	0.5	5.54	5.23	5.13	5.40	5.40	5.27	5.12	5.00	5.06
	1	4.55	4.55	4.55	4.79	4.79	4.79	4.70	4.70	4.70
50	0	4.57	4.81	5.12	4.92	5.30	5.17	4.62	5.05	5.11
	0.5	5.22	5.03	4.97	4.86	5.21	5.20	5.00	5.01	5.03
	1	5.09	5.09	5.09	5.07	5.07	5.07	4.90	4.90	4.90
100	0	4.53	4.58	4.78	4.78	5.03	4.94	4.57	4.66	4.65
	0.5	4.82	4.64	4.69	4.88	4.73	4.89	4.62	4.42	4.63
	1	4.60	4.60	4.60	4.67	4.67	4.67	4.69	4.69	4.69
200	0	4.72	4.77	4.78	4.79	4.72	4.90	4.78	4.73	4.83
	0.5	5.08	4.87	4.80	4.74	4.54	4.78	5.17	4.72	4.74
	1	4.89	4.89	4.89	4.92	4.92	4.92	4.99	4.99	4.99

Tabla 3.6.1: Tamaño empírico del contraste MCE

T	a	D ₁			D ₂			D ₃		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	10.70	39.10	43.30	17.90	46.40	50.35	16.75	37.60	41.75
	0.5	15.60	28.35	40.90	6.60	35.60	48.45	8.80	28.70	40.05
	1	1.95	1.95	1.95	2.45	2.45	2.45	1.55	1.55	1.55
50	0	9.25	39.45	41.95	15.65	46.30	49.15	13.40	39.95	42.20
	0.5	21.30	31.75	40.35	2.50	38.65	47.85	7.40	33.25	41.35
	1	0.40	0.40	0.40	0.85	0.85	0.85	0.30	0.30	0.30
100	0	9.15	33.95	36.30	12.75	39.90	41.95	11.25	34.35	35.85
	0.5	9.25	28.10	34.85	0.85	34.00	41.20	2.50	29.55	35.50
	1	0.10	0.10	0.10	0.20	0.20	0.20	0.00	0.00	0.00
200	0	7.25	34.10	36.15	9.60	38.05	40.35	9.15	34.05	36.05
	0.5	3.60	28.75	35.15	0.80	31.80	39.30	0.95	28.70	34.90
	1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Tabla 3.6.2: Potencia del contraste MCE, $b = -0.2$ TABLA 3.6. Procedimiento Bootstrap S_3, t_2 . Co-ruptura en diferencias, no en niveles. Modelo estimado (4.2), modelo MCE.

El modelo de corrección del error extendido (4.3) aproxima bien la ecuación (5.3) si

$$(y_{t-2} - \alpha z_{t-2}) \cong b_1 t + b_2 s \sum_{i=1}^t D_{j,i}, \quad (5.4)$$

y esto es muy probable cuando hay co-ruptura parcial sólo en diferencias, no en niveles.

Los resultados del experimento de Monte Carlo del contraste MCE bootstrap basado en S_3, t_2 se recogen en la Tabla 3.6. La Tabla 3.6.1 muestra que no hay distorsiones del tamaño, pero la Tabla 3.6.2 muestra importantes reducciones de potencia respecto al caso de co-ruptura simultánea. El peor caso es cuando la potencia es prácticamente nula si se cumple la restricción COMFAC ($a = 1$). Se observa que en algunos casos la potencia del contraste no sólo no aumenta, sino que disminuye, al aumentar el tamaño de la muestra.

T	a	D_1			D_2			D_3		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.46	5.18	5.07	5.62	5.33	5.26	5.19	5.10	4.91
	0.5	5.54	5.23	5.13	5.40	5.40	5.27	5.12	5.00	5.06
	1	4.55	4.55	4.55	4.79	4.79	4.79	4.70	4.70	4.70
50	0	4.57	4.81	5.12	4.92	5.30	5.17	4.62	5.05	5.11
	0.5	5.22	5.03	4.97	4.86	5.21	5.20	5.00	5.01	5.03
	1	5.09	5.09	5.09	5.07	5.07	5.07	4.90	4.90	4.90
100	0	4.53	4.58	4.78	4.78	5.03	4.94	4.57	4.66	4.65
	0.5	4.82	4.64	4.69	4.88	4.73	4.89	4.62	4.42	4.63
	1	4.60	4.60	4.60	4.67	4.67	4.67	4.69	4.69	4.69
200	0	4.72	4.77	4.78	4.79	4.72	4.90	4.78	4.73	4.83
	0.5	5.08	4.87	4.80	4.74	4.54	4.78	5.17	4.72	4.74
	1	4.89	4.89	4.89	4.92	4.92	4.92	4.99	4.99	4.99

Tabla 3.7.1: Tamaño empírico del contraste MCE

T	a	D_1			D_2			D_3		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	10.50	48.10	64.70	15.45	48.45	62.10	14.35	52.35	67.30
	0.5	14.70	25.75	55.60	10.35	27.90	54.70	11.65	29.25	58.75
	1	6.60	6.60	6.60	6.90	6.90	6.90	5.95	5.95	5.95
50	0	16.95	68.80	79.45	21.85	65.75	76.20	19.40	69.45	78.10
	0.5	25.00	45.25	74.70	12.75	45.70	70.85	14.25	49.10	74.00
	1	9.25	9.25	9.25	9.30	9.30	9.30	8.90	8.90	8.90
100	0	31.45	89.75	93.40	33.25	86.25	91.20	33.10	85.75	89.45
	0.5	30.70	74.50	91.50	22.20	69.75	88.90	21.35	72.25	87.50
	1	18.00	18.00	18.00	18.30	18.30	18.30	17.45	17.45	17.45
200	0	56.50	98.80	99.30	57.45	98.25	99.45	56.65	97.00	98.40
	0.5	48.40	93.95	99.15	42.90	92.80	98.90	41.10	92.15	97.45
	1	34.70	34.70	34.70	34.25	34.25	34.25	34.40	34.40	34.40

Tabla 3.7.2: Potencia del contraste MCE, $b = -0.2$ TABLA 3.7. Procedimiento Bootstrap S_3, t_2 . Co-ruptura en diferencias, no en niveles. Modelo estimado (4.3). Modelo MCE extendido.

Como era de esperar, a la vista del Capítulo 2, los resultados mejoran notablemente si utilizamos el contraste basado en el contraste MCE extendido, ver Tabla 3.7. La Tabla 3.7.1 muestra que el tamaño del contraste bootstrap MCE extendido está siempre próximo al tamaño nominal. La Tabla 3.7.2 evidencia un importante aumento de la potencia relativa al contraste MCE habitual. Las potencias más bajas se obtienen de nuevo cuando se satisface la restricción e factor común, COMFAC ($a = 1$), pero la potencia del contraste siempre aumenta al incrementar el tamaño de la muestra.

La Figura 3.3 muestra la potencia de los contrastes bootstrap MCE y MCE extendido para una muestra de 50 observaciones sin cambio estructural y con un cambio doble en la tendencia (variable dummy $D_{3,t}$). En ambos casos, la potencia del contraste basado en el MCE extendido es mayor que la del MCE habitual para valores de b menores que $b = -0.15$. La Figura 3.3.2 muestra que la potencia del contraste

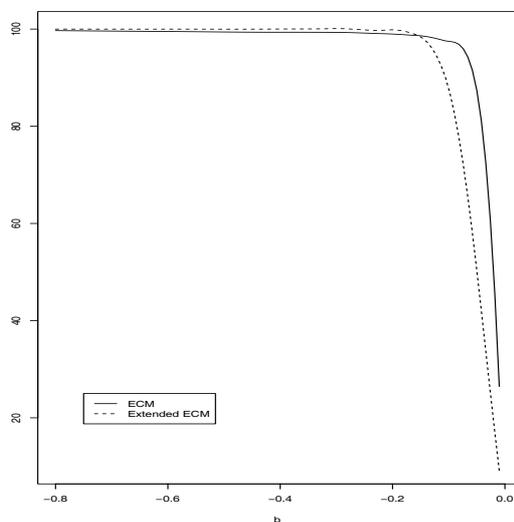


Figura 3.3.1: No salto

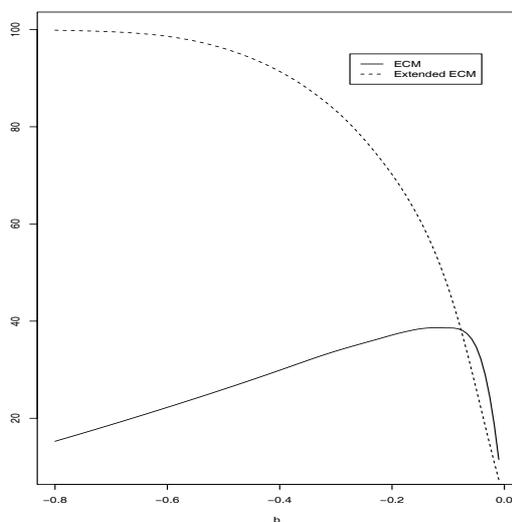


Figura 3.3.2: D3

FIGURA 3.3. Potencia de los contrastes bootstrap MCE y MCE extendido. Co-ruptura en diferencia, no en niveles. $T = 50$.

MCE decrece con $|b|$, en tanto que la potencia del contraste basado en el MCE extendido presenta el comportamiento deseable en el contraste, aumentando con $|b|$. Estos resultados muestran que la aproximación (5.4) funciona bien.

T	a	D ₁			D ₂			D ₃		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	13.13	31.37	32.32	17.99	37.20	37.91	10.78	21.47	22.50
	0.5	8.24	27.52	31.76	9.27	35.09	37.65	7.54	19.33	21.94
	1	6.43	6.43	6.43	6.56	6.56	6.56	6.45	6.45	6.45
50	0	21.46	38.71	39.44	30.03	42.65	43.07	16.04	28.46	29.17
	0.5	11.55	36.01	38.97	13.70	41.85	42.87	9.85	26.30	28.84
	1	5.70	5.70	5.70	5.94	5.94	5.94	5.81	5.81	5.81
100	0	32.37	43.48	44.10	40.44	46.16	46.36	23.39	35.11	36.04
	0.5	18.30	42.01	43.82	23.71	45.24	46.13	14.15	33.02	35.35
	1	5.97	5.97	5.97	6.05	6.05	6.05	5.97	5.97	5.97
200	0	40.13	45.66	45.89	45.59	47.23	47.32	32.27	40.52	41.02
	0.5	28.37	45.25	45.71	36.26	46.87	47.31	21.10	39.53	40.83
	1	5.83	5.83	5.83	5.69	5.69	5.69	5.94	5.94	5.94

TABLA 3.8. Tamaño empírico del contraste MCE. Procedimiento bootstrap. S_3, t_2 . Co-ruptura en niveles, no en diferencias. Modelo estimado (4.2).

5.2.2. *Co-ruptura en niveles, no en diferencias.* En este caso, la ecuación (5.2a) toma la forma

$$\Delta y_t = -bc_l + (\alpha - a)sD_{j,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \quad (5.5)$$

En el caso de co-ruptura en niveles, pero no en diferencias, el modelo de corrección del error extendido no será una buena aproximación, ya que

$$(y_{t-2} - \alpha z_{t-2}) \approx (\alpha - a)sD_{j,t}$$

La Tabla 3.8 muestra las dramáticas distorsiones del contraste MCE bootstrap basado en el modelo (4.2). El tamaño del mismo puede llegar a alcanzar valores de 47%, cuando el valor nominal del mismo es 5%.

Si hacemos el contraste basado en el bootstrap del estadístico t del modelo MCE extendido, ecuación (4.3), la Tabla 3.9.1 muestra que las distorsiones mayores han pasado de 47% a 18%, pero el tamaño sigue siendo inadecuado. Evidentemente, este problema no aparece cuando se satisface la restricción COMFAC ($a = 1$), pues en ese caso concreto tenemos co-ruptura completa. La Tabla 3.9.2 y la Figura 3.4 muestran la potencia del contraste bootstrap basado en el MCE extendido como función del parámetro b . Se observa que para valores de b menores que $b = -0.15$ la potencia está próxima al 100%. También se observa que la potencia aumenta al pasar de 25 a 50 observaciones.

T	α	D_1			D_2			D_3		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	6.45	11.50	12.59	7.68	17.18	18.75	7.34	12.68	13.20
	0.5	6.05	9.29	12.14	6.64	13.63	17.87	6.45	11.08	12.89
	1	6.15	6.15	6.15	6.20	6.20	6.20	6.09	6.09	6.09
50	0	6.30	12.78	14.29	7.88	16.69	18.45	7.47	12.04	12.69
	0.5	5.41	10.29	13.47	6.28	13.41	17.18	6.19	10.55	12.40
	1	5.79	5.79	5.79	5.81	5.81	5.81	5.88	5.88	5.88
100	0	6.80	12.92	14.60	7.76	15.76	17.57	7.19	11.48	12.09
	0.5	5.90	10.19	13.67	6.54	12.39	16.66	6.30	9.91	11.75
	1	5.14	5.14	5.14	5.30	5.30	5.30	5.27	5.27	5.27
200	0	6.97	13.09	14.71	7.78	15.10	16.81	6.96	11.18	11.78
	0.5	6.10	9.85	13.67	6.96	11.93	15.76	6.26	9.86	11.65
	1	5.36	5.36	5.36	5.38	5.38	5.38	5.30	5.30	5.30

Tabla 3.9.1: Tamaño empírico del contraste MCE extendido

T	α	D_1			D_2			D_3		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	23.25	58.80	61.45	24.70	51.25	52.45	23.45	50.40	52.60
	0.5	16.50	48.40	59.65	17.80	45.40	51.95	17.65	44.45	52.10
	1	13.65	13.65	13.65	14.35	14.35	14.35	14.35	14.35	14.35
50	0	46.25	89.25	92.30	47.35	80.05	82.95	45.95	81.15	82.75
	0.5	33.35	80.35	90.80	35.55	74.45	81.85	34.25	73.40	82.05
	1	25.65	25.65	25.65	25.75	25.75	25.75	25.60	25.60	25.60
100	0	77.50	99.80	99.95	77.75	99.20	99.70	76.95	98.75	99.25
	0.5	62.25	98.45	99.90	62.55	96.95	99.50	62.70	96.55	99.05
	1	50.30	50.30	50.30	50.10	50.10	50.10	50.15	50.15	50.15
200	0	96.85	100.00	100.00	96.60	100.00	100.00	97.15	100.00	100.00
	0.5	90.50	100.00	100.00	90.40	100.00	100.00	90.70	100.00	100.00
	1	83.50	83.50	83.50	83.60	83.60	83.60	83.35	83.35	83.35

Tabla 3.9.2: Potencia del contraste MCE extendido, $b = -0.2$

TABLA 3.9. Procedimiento bootstrap S_3, t_2 . Co-ruptura en niveles, no en diferencias. Modelo estimado (4.3). Contraste MCE extendido.

Una solución basada en la restricción COMFAC.. A partir de los resultados del Capítulo 2, la ecuación (5.5) se puede reexpresar como

$$(\Delta y_t - \alpha \Delta z_t) = -bc_l + b(y_{t-1} - \alpha z_{t-1}) + v_{1,t} \quad (5.6)$$

donde $v_{1,t} = (a - \alpha)(\Delta z_t - sD_{j,t}) + u_{1,t} = (a - \alpha)u_{2,t} + u_{1,t}$ y, por tanto, los saltos en diferencias desaparecen de la ecuación del contraste.

Podemos, por tanto, utilizar las siguientes ecuaciones para realizar el contraste MCE habitual o el contraste MCE extendido:

$$(\Delta y_t - \alpha \Delta z_t) = c + b(y_{t-1} - \alpha z_{t-1}) + \epsilon_t \quad (5.7)$$

$$(\Delta y_t - \alpha \Delta z_t) = c + b(y_{t-1} - \alpha z_{t-1}) + d(y_{t-2} - \alpha z_{t-2}) + \epsilon_t \quad (5.8)$$

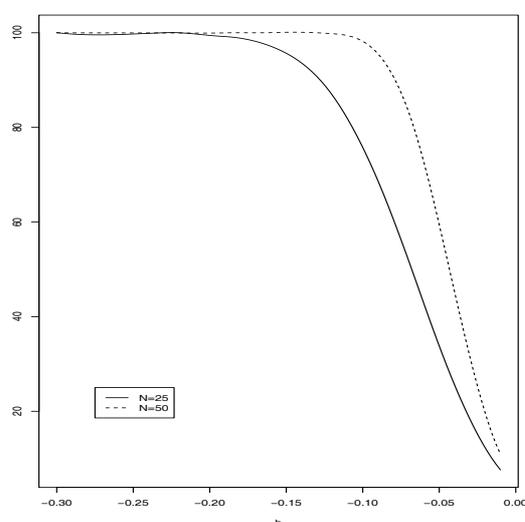


FIGURA 3.4. Potencia del contraste bootstrap MCE extendido. Co-ruptura en niveles, no en diferencias, D3, $a = 0$, $s = 6$. Modelo estimado (4.3).

La Tabla 3.10 resume los principales resultados del contraste MCE en el caso de que tengamos co-ruptura parcial en niveles, no en diferencias, y se imponga la restricción COMFAC. Tal y como puede verse, el tamaño del contraste cuando se usa el modelo (5.7) toma valores superiores a 6% cuando el tamaño de la muestra es de 25 observaciones, e incluso para muestras de tamaño 200 encontramos combinaciones de parámetros que hacen que el contraste bootstrap tenga un tamaño cercano a 6%, ver Tabla 3.10.1.

El tamaño del contraste se aproxima a su valor nominal cuando utilizamos el contraste bootstrap basado en el modelo MCE extendido, ecuación (5.7) Tal y como se observa en la Tabla 3.10.2, el tamaño del contraste sólo es ligeramente superior a 6% cuando el tamaño de la muestra es 25 y se satisface la restricción COMFAC. Al tomar muestras de tamaño igual o superior a 50, el tamaño del contraste bootstrap no difiere significativamente de su valor nominal.

Hay que resaltar que esto sucede sólo en el caso de que hay co-ruptura parcial en niveles. Se comprobaron algunas otras especificaciones, como co-ruptura simultánea, pero la imposición de la restricción COMFAC resulta absolutamente inadecuada en esos casos, lo que es congruente con los resultados del Capítulo 2 (tablas disponibles, pero no incluidas).

T	a	D_1			D_2			D_3		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.26	5.58	5.54	5.26	5.58	5.54	5.26	5.58	5.54
	0.5	5.37	5.57	5.60	5.37	5.57	5.60	5.37	5.57	5.60
	1	6.27	6.27	6.27	6.27	6.27	6.27	6.27	6.27	6.27
50	0	5.71	5.52	5.55	5.71	5.52	5.55	5.71	5.52	5.55
	0.5	6.03	5.17	5.65	6.03	5.17	5.65	6.03	5.17	5.65
	1	5.97	5.97	5.97	5.97	5.97	5.97	5.97	5.97	5.97
100	0	5.58	5.43	5.74	5.58	5.43	5.74	5.58	5.43	5.74
	0.5	5.75	5.46	5.52	5.75	5.46	5.52	5.75	5.46	5.52
	1	5.88	5.88	5.88	5.88	5.88	5.88	5.88	5.88	5.88
200	0	5.95	5.67	5.43	5.95	5.67	5.43	5.95	5.67	5.43
	0.5	5.83	5.80	5.61	5.83	5.80	5.61	5.83	5.80	5.61
	1	5.82	5.82	5.82	5.82	5.82	5.82	5.82	5.82	5.82

Tabla 3.10.1: Modelo MCE-COMFAC.

T	a	D_1			D_2			D_3		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.24	5.81	5.79	5.24	5.81	5.79	5.24	5.81	5.79
	0.5	5.68	5.72	5.73	5.68	5.72	5.73	5.68	5.72	5.73
	1	6.02	6.02	6.02	6.02	6.02	6.02	6.02	6.02	6.02
50	0	5.26	5.52	5.36	5.26	5.52	5.36	5.26	5.52	5.36
	0.5	5.33	5.45	5.38	5.33	5.45	5.38	5.33	5.45	5.38
	1	5.62	5.62	5.62	5.62	5.62	5.62	5.62	5.62	5.62
100	0	5.11	5.33	5.35	5.11	5.33	5.35	5.11	5.33	5.35
	0.5	5.21	5.06	5.44	5.21	5.06	5.44	5.21	5.06	5.44
	1	5.15	5.15	5.15	5.15	5.15	5.15	5.15	5.15	5.15
200	0	5.09	4.92	5.06	5.09	4.92	5.06	5.09	4.92	5.06
	0.5	5.43	5.12	5.07	5.43	5.12	5.07	5.43	5.12	5.07
	1	5.27	5.27	5.27	5.27	5.27	5.27	5.27	5.27	5.27

Tabla 3.10.2: Modelo MCE extendido-COMFAC

TABLA 3.10. Tamaño empírico de los contrastes MCE-COMFAC, ecuación (5.7), y MCE extendido-COMFAC, ecuación (5.8). Procedimiento bootstrap S_3, t_2 . Co-ruptura en niveles, no en diferencias.

6. Conclusiones

Se ha procedido al análisis de diferentes contrastes MCE bootstrap con cambios estructurales no modelizados con vector de cointegración conocido. En los casos en que no hay cambio estructural o hay co-ruptura simultánea, la conclusión es que la inclusión de un término constante en la regresión del contraste puede originar severas distorsiones del tamaño del contraste basado en el MCE habitual. Este problema se soluciona usando modelos MCE extendidos para realizar el contraste bootstrap. Como contrapartida, el único esquema de remuestreo válido es el denominado S_3, t_2 .

Este procedimiento se basa en el remuestreo de los residuos obtenidos tras estimar el modelo MCE restringido, bajo la hipótesis nula $b = 0$.

Además, en el caso de co-rupturas parciales en diferencias, no en niveles, el contraste basado en el MCE habitual no tiene potencia. Este problema también se soluciona utilizando el modelo MCE extendido.

En el caso de co-ruptura parcial en niveles, no en diferencias, los modelos MCE extendidos hacen que el tamaño del contraste esté más próximo al valor nominal que con los modelos MCE habituales, pero en cualquier caso son excesivamente altos. Este problema desaparece si se impone la restricción de factor común (COMFAC) en los modelos estimados.

Contrastes de Cointegración MCE Robustos a Atípicos Basados en el Componente Tendencial

1. Introducción

Las propiedades de los contrastes de cointegración basados en modelos de corrección del error uniecuacionales (contrastos MCE) son bien conocidas. La dependencia de los valores críticos y la potencia del contraste MCE de parámetros molestos se recoge Banerjee et al. (1986), Engle y Granger (1987), Kremers et al. (1992), Banerjee et al. (1993) ¹.

Los efectos de cambios estructurales al aplicar contrastes de raíces unitarias, tales como el contraste de Dickey y Fuller (1979), etc., también se han analizado, ver Stock (1994). Perron (1989) es un buen punto de partida para analizar tales impactos. Según Clements y Hendry (1999), un cambio estructural equivale esencialmente a un shock intermitente con efecto temporal en la serie. Si este shock no se modeliza explícitamente, los contrastes de raíces unitarias habituales confundirán el cambio estructural con una raíz unitaria. Los resultados de Hendry y Neale (1990) y Perron y Vogelsang (1992) también indican que un cambio de tipo escalón en la media no modelizado nos llevará a detectar raíces unitarias espurias. Rappoport y Reichlin (1989) es quizá la primera referencia que analiza los efectos de tener tendencias segmentadas en la hipótesis alternativa del modelo de raíces unitarias, y Andrés et al. (1990) extiende el análisis al caso en que hay más de un punto de ruptura en la tendencia. Otras referencias sobre el tema de cambios estructurales y raíces unitarias son Banerjee et al. (1992), Zivot y Andrews (1992) y Leybourne et al. (1998).

Desde el punto de vista aplicado, el principal problema que se plantea es que debemos incluir variables tipo dummy para aplicar contrastes de raíces unitarias válidos cuando los valores críticos dependen del tamaño y la situación de los saltos. De nuevo, la selección de las variables dummy resulta fundamental para el correcto resultado del contraste. Hay toda una Literatura que trata de métodos de detección de puntos de ruptura basados en contrastes recursivos o secuenciales (Andrews, 1993, Andrews

¹Ver en el Capítulo 2 un resumen de los principales resultados.

et al., 1996, Bai, 1997, Vogelsang, 1997, Bai y Perron, 1998, Banerjee, Lazarova y Urga, 1998).

Otra clase de sucesos no habituales la constituyen los atípicos aditivos. Estos son sucesos con efecto grande, pero temporal, sobre las series. En determinados casos, dichos efectos dominan al resto de la información contenida en la serie y sesga la inferencia de raíces unitarias en favor del rechazo de la hipótesis de raíz unitaria incluso si la hipótesis nula de presencia de raíz unitaria es cierta, tal y como señalan Franses y Haldrup (1994) y Lucas (1995*a,b*).

En el caso de series temporales multivariantes, la situación puede ser peor, pues necesitamos tomar una decisión acerca del tipo de modelos que genera las observaciones anomalás (cambios de tendencia, atípicos aditivos, ...) teniendo en cuenta que dichas irregularidades no tienen por qué ocurrir simultáneamente ni en todas las variables. Por tanto, el análisis multivariante es en general más complicado. No obstante, se puede simplificar en algunos casos si hay co-ruptura parcial.

En las aplicaciones empíricas, lo más habitual es incluir variables dummy para obtener modelos MCE con parámetros constantes. Los efectos de la inclusión de variables dummy en modelos MCE para capturar cambios estructurales se analiza en Kremers et al. (1992) y Campos et al. (1996). De nuevo, los valores críticos (CV) dependen del tipo concreto de dummy que se incluya en el modelo, lo que constituye un problema a la hora de realizar trabajos empíricos.

Una alternativa para evitar el uso de variables dummy si utilizamos técnicas de estimación robustas. Este es el enfoque que proponen Lucas (1995*a,b*) en el caso univariante y Franses y Lucas (1997*a,b*) en el marco multivariante.

En este Capítulo seguimos una ruta diferente. El objetivo es encontrar un método de modelización en un contexto MCE que nos permita contrastar cointegración de forma robusta a la presencia de cambios estructurales. En lugar de incluir variables dummy en los modelos MCE, lo que se hace es aproximar dichas rupturas añadiendo términos dinámicos, por medio del criterio de información SBIC. En particular, nos vamos a centrar en el estudio de los valores críticos obtenidos con los modelos sobreparametrizados, la potencia de los mismos y su robustez frente a atípicos aditivos por medio de simulaciones y experimentos de Monte Carlo. También se analizará la robustez de los contrastes MCE cuando en lugar de aplicarlos a las series observadas se utilizan los componentes tendenciales obtenidos mediante procedimientos de descomposición tendencia-ciclo, como en el Capítulo 1. En concreto, se aplicarán tres filtros, el filtro Hodrick y Prescott (1980, 1997), el filtro Baxter y King (1995, 1999),

HP y BK respectivamente a partir de ahora, y el filtro de la mediana (ver Wen y Zeng, 1999). Guay y St-Amant (1997) y Baxter y King (1995) proporcionan información acerca de las relaciones entre los filtros HP y BK.

La estructura del Capítulo es la siguiente. En la Sección 2 se analizan los efectos de tener shocks transitorios en las especificaciones alternativas de los modelos MCE, y en particular en la relación de cointegración. Se estudian detalladamente tres tipos de co-ruptura: co-ruptura simultánea, co-ruptura parcial en niveles (no en diferencias) y co-ruptura parcial en diferencias (no en niveles). También se analizan algunos casos en los que no hay ningún tipo de co-ruptura. Los filtros de extracción de señales que se utilizan para estimar el componente tendencial de las series se resumen en la Sección 3, y los modelos MCE basados en el componente tendencial se introducen en la Sección 4. La Sección 5 presenta los resultados principales de los experimentos de simulación de Monte Carlo. La utilidad del enfoque de este Capítulo se ilustra mediante una aplicación empírica en la Sección 7. Por último, la Sección 8 presenta las conclusiones y las direcciones de futuras investigaciones.

2. Modelos de Corrección del Error con y sin Co-ruptura Simultánea

Sea el siguiente modelo de corrección del error condicional (MCE)

$$\Delta(y_t - \mu_{y,t}) = a\Delta(z_t - \mu_{z,t}) + b[(y_{t-1} - \mu_{y,t-1}) - \alpha(z_{t-1} - \mu_{z,t-1})] + u_{1t} \quad (2.1a)$$

$$\Delta(z_t - \mu_{z,t}) = u_{2t} \quad (2.1b)$$

Supongamos que $\dots, y_{-1}, y_0 = 0$ y $\dots, z_{-1}, z_0 = 0$, sean $\mu_{y,t} = \tilde{\mu}_{y,t} + s_y \delta_t^y$, $\mu_{z,t} = \tilde{\mu}_{z,t} + s_z \delta_t^z$, donde δ_t^y y δ_t^z son variables *iid* Bernoulli independientes de $u_{1,t}$ y $u_{2,t}$

$$\begin{cases} \Pr(\delta_t = 1) = \Pr(\delta_t = -1) = \frac{\pi}{2} \\ \Pr(\delta_t = 0) = 1 - \pi. \end{cases}$$

Los términos $\tilde{\mu}_{y,t}$ y $\tilde{\mu}_{z,t}$ incluyen todo tipo de componentes deterministas, tales como constantes, tendencias deterministas, variables dummy, tendencias segmentadas, etc. Sea B el operador back-shift, $B^k y_t = y_{t-k}$, $\Delta = (1 - B)$ es el operador primeras diferencias, y sea $(1, -\alpha)$ el vector de cointegración. Dadas condiciones iniciales válidas, los errores estocásticos u_{1t} y u_{2t} no presentan correlación conjunta ni serial, tienen media cero y varianzas constantes σ_1^2 y σ_2^2 , respectivamente.

El modelo (2.1a)–(2.1b) puede expresarse en función de las variables observables y_t y z_t como sigue

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (2.2a)$$

$$\Delta z_t = \Delta \mu_{z,t} + u_{2t} \quad (2.2b)$$

$$c_t \equiv \Delta \mu_{y,t} - a\Delta \mu_{z,t} - b(\mu_{y,t-1} - \alpha \mu_{z,t-1}) \quad (2.2c)$$

$$\mu_{z,t} = \tilde{\mu}_{z,t} + s_z \delta_t^z \quad (2.2d)$$

$$\mu_{y,t} = \tilde{\mu}_{y,t} + s_y \delta_t^y \quad (2.2e)$$

En este Capítulo se analizan los efectos de tener diferentes modelos alternativos (con co-rupturas parciales, etc.) en el término estocástico c_t de (2.2a) dado por (2.2c) en el contraste MCE de no-cointegración ($b = 0$).

DEFINICIÓN 2.1. Decimos que las series y_t y z_t tienen *co-rupturas en niveles* si $\mu_{y,t} - \alpha \mu_{z,t} = c_t$, donde c_t es un parámetro finito constante.

DEFINICIÓN 2.2. Decimos que las series y_t y z_t tienen *co-rupturas en diferencias* si $\Delta \mu_{y,t} - a\Delta \mu_{z,t} = c_d$, donde c_d es un parámetro finito constante.

DEFINICIÓN 2.3. Decimos que las series y_t y z_t tienen *co-ruptura simultánea* si $\Delta \mu_{y,t} - a\Delta \mu_{z,t} - b(\mu_{y,t} - \alpha \mu_{z,t}) = c_s$, donde c_s es un parámetro finito constante.

De las definiciones 2.1 y 2.2, se deduce que si y_t and z_t presentan *co-ruptura en niveles y en diferencias (co-ruptura completa)*, tenemos una caso particular de co-ruptura simultánea.

En las aplicaciones empíricas hay una serie de posibles casos intermedios e interés que analizaremos en nuestros experimentos.

CASO 2.1. *Co-ruptura en niveles, no en diferencias.* Co-ruptura en niveles: ($\mu_{y,t} - \alpha \mu_{z,t} = c_l$). Al tomar primeras diferencias se tiene $\Delta \mu_{y,t} - \alpha \Delta \mu_{z,t} = 0$. Pero de la ecuación (2.2c)

$$c_t = \Delta \mu_{y,t} - a\Delta \mu_{z,t} - bc_l = (\alpha - a)\Delta \mu_{z,t} - bc_l, \quad (2.3)$$

y (2.2a) se reescribe como

$$\Delta y_t = -bc_l + (\alpha - a)\Delta \mu_{z,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \quad (2.4)$$

OBSERVACIÓN 2.1. Co-ruptura en niveles \Rightarrow co-ruptura en diferencias si $a = \alpha$ (COMFAC), para todo valor de c_l .

CASO 2.2. *Co-ruptura en diferencias, pero no en niveles.* Co-ruptura en diferencias: $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$ implica que $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = (a - \alpha)\Delta\mu_{z,t} + c_d$. A partir de sustituciones recursivas, $\mu_{y,t} - \alpha\mu_{z,t} = (\mu_{y_0} - \alpha\mu_{z_0}) + c_d t + (a - \alpha)\mu_{z,t}$, y c_t es

$$c_t = c_d - b(\mu_{y,0} - \alpha\mu_{z,0}) - bc_d(t - 1) - b(a - \alpha)\mu_{z_{t-1}} \quad (2.5)$$

y la ecuación (2.2a) se convierte en

$$\Delta y_t = c_m + bc_d t - b(a - \alpha)\mu_{z,t-1} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t}, \quad (2.6)$$

donde c_m es una constante igual a $c_m = c_d - b(\mu_{y,0} - \alpha\mu_{z,0}) + bc_d$.

OBSERVACIÓN 2.2. Suponiendo que $\mu_{y,0} - \alpha\mu_{z,0} = \text{constant}$, co-ruptura en diferencias \Rightarrow co-ruptura en niveles si $a = \alpha$ (COMFAC) y $c_d = 0$.

En general, si no tenemos ningún tipo de co-ruptura parcial, la representación más parsimoniosa en función de las variables no observables es el modelo MCE condicional (2.1a). En términos de variables observables, la representación con mayor parsimonia es (2.2a), por que sólo hay que incluir los regresores procedentes de los valores contemporáneos de c_t . Si queremos estimar los valores de los parámetros a , α y b , podemos hacerlo mediante procedimientos monoetápicos (MCO o MCNL) en la representación MCE, ver Capítulo 2 para más detalles. Sin embargo, previamente debemos estimar $\mu_{y,t}$ y $\mu_{z,t}$, y esto puede incorporar hipótesis arbitrarias de sucesos desconocidos (variables dummy, etc.), ver Vogelsang (1999).

En esta Tesis se sostiene que, en un contexto no estacionario, $\mu_{y,t}$ y $\mu_{z,t}$ pueden incluir cualquier combinación posible de shocks transitorios y observaciones atípicas, lo que complica el análisis. Por tanto, se recomienda el uso de procedimientos de extracción de tendencia para eliminar esos elementos transitorios no estacionarios no observados. En concreto, se aconseja especificar los modelos de corrección del error en función de los componentes de crecimiento (tendencia), para obtener contrastes MCE de no cointegración ($b = 0$) robustos a la presencia de atípicos, como quedará claro en las Secciones 4 y 5.

2.1. Modelos de Corrección del Error con Co-ruptura Simultánea. A partir de las ecuaciones (2.2a)–(2.2c) y de los análisis de Escribano (1987a) y Andrés et al. (1990), es evidente que cualquier modelo de corrección del error en término de las variables observables debe tener en cuenta los efectos conjuntos de los siguientes elementos: $\Delta\mu_{y,t}$, $\Delta\mu_{z,t}$, $\mu_{y,t-1}$ y $\mu_{z,t-1}$.

Los modelos de corrección del error con co-rupturas ya se han estudiado en parte en los trabajos de Campos et al. (1996) y Clements y Hendry (1999). En esta Sección

se estudian los modelos con co-ruptura simultánea tales que $c_t = 0$. Bajo co-ruptura simultánea, (2.2a) y (2.2b) pueden simplificarse como

$$\Delta y_t = a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (2.7a)$$

$$\Delta z_t = \Delta \mu_{z,t} + u_{2,t} \quad (2.7b)$$

$$\mu_{z,t} = \tilde{\mu}_{z,t} + s_z \delta_t^z \quad (2.7c)$$

donde (2.7a) toma la forma de un modelo de corrección del error uniecuacional si término independiente, puesto que $c_t = 0$. el parámetro s_z mide el tamaño del salto, y δ_t^z es el atípico aditivo, ver Sección 5 para más detalles.

A partir de las ecuaciones (2.7a) y (2.7b) se deduce que $y_t \sim I(1)$, $z_t \sim I(1)$, y que están cointegradas en el sentido de Engle y Granger (1987) con vector de cointegración igual a $(1, -\alpha)$ para algunos valores del parámetro b ($-2 < b < 0$). Obsérvese que se permiten shocks transitorios en la variable ‘exógena’ z_t que presenta co-ruptura simultánea con la variable endógena y_t y, por tanto, los saltos desaparecen del modelo condicional, ecuación (2.7a).

2.2. Modelos de Corrección del Error sin Co-ruptura Simultánea. En la Sección anterior se ha estudiado el caso en que hay co-ruptura simultánea. Ahora se analiza un serie de casos intermedios, ver Capítulo 2 para una discusión en profundidad de las relaciones entre los mismos.

CASO 2.1. Co-ruptura en niveles, no en diferencias.

De las ecuaciones (2.2a)–(2.2c) se tiene

$$\Delta y_t = -bc_t + (\alpha - a)s_z \Delta \delta_t^z + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \quad (2.8)$$

Por tanto, los saltos en δ_t^z en el proceso marginal de Δz_t afectan al modelo de corrección del error a no ser que se satisfaga la restricción COMFAC ($a = \alpha$). Más adelante se analizarán los efectos de omitir el segundo término del lado derecho de la ecuación (2.8).

CASO 2.2. Co-ruptura en diferencias, no en niveles.

A partir de la ecuación (2.6)

$$\begin{aligned} \Delta y_t = & c_m + bc_{dt} - b(a - \alpha)s_z \delta_{t-1}^z + a\Delta z_t \\ & + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \end{aligned} \quad (2.9)$$

La Sección 5 evidencia los efectos que tiene la omisión del segundo y tercer término del lado derecho de la ecuación (2.9)

CASO 2.3. *Saltos independientes*

La última posibilidad es que no haya co-ruptura ni en niveles ni en diferencias. Esta situación es el resultado de la combinación de los efectos de las ecuaciones (2.8) y (2.9). En nuestras simulaciones, vamos a considerar shocks independientes en ambas variables, y_t and z_t , y también shocks en sólo una de las dos series. Por taanto, el término c_t de (2.2c) será uno de los siguientes:

$$c_t = s_y \Delta \delta_t^y - a s_z \Delta \delta_t^z - b (s_y \delta_{t-1}^y - \alpha s_z \delta_{t-1}^z) \quad (2.10a)$$

$$c_t = -a s_z \Delta \delta_t^z + b \alpha s_z \delta_{t-1}^z \quad (2.10b)$$

$$c_t = s_y \Delta \delta_t^y + b s_y \delta_{t-1}^y \quad (2.10c)$$

y se analizarán los efectos de considerar c_t como un término constante en las diferentes estrategias de modelización.

3. Filtros y Extracción de Señales

El propósito habitual del filtrado es separar de las series temporales macroeconómicas el componente de ciclo de negocios de la tendencia secular y de los más volátiles componentes estacional e irregular. Supongamos que estamos en separar una serie temporal observada, z_t , en dos componentes no observados,

$$y_t = y_t^g + y_t^c \quad (3.1)$$

donde z_t^g es el componente crecimiento o tendencia y z_t^c es el componente cíclico. Se consideran dos posibilidades. En el primer caso, se puede definir el componente tendencial y el ciclo es el residuo

$$y_t = y_t^g + (y_t - y_t^g) \quad (3.2a)$$

En el segundo caso, se define el ciclo y la tendencia es el residuo si las series son integradas $I(d)$, con $d \geq 1$

$$y_t = (y_t - y_t^c) + y_t^c \quad (3.2b)$$

3.1. El filtro de Baxter y King (BK) . La mayoría de los filtros utilizados en Macroeconometría son medias móviles simétricas de orden infinito bidireccionales (hacia adelante y hacia atrás), como señalan King y Rebelo (1993) y Baxter y King (1995, 1999). En la práctica, el filtro tiene que aproximarse mediante una $MA(k)$

bi-direccional

$$y_t^* = a_0 + \sum_{h=1}^k a_h (B^h + B^{-h}) y_t$$

Las implicaciones del filtro se ven más claramente en el dominio de las frecuencias observando la *función de respuesta frecuencial* del filtro. La función de respuesta frecuencial de la $MA(\infty)$ es

$$\beta(\omega) = \sum_{h=-\infty}^{\infty} b_h e^{-i\omega h} \quad (3.3)$$

en tanto que para la $MA(k)$ finita en ambos sentidos viene dada por

$$\alpha(\omega) = \sum_{h=-k}^k a_h e^{-i\omega h}. \quad (3.4)$$

Baxter y King (1995) obtiene filtros aproximados óptimos minimizando el error cuadrático medio de $\delta(\omega) \equiv \beta(\omega) - \alpha(\omega)$ respecto a a_h . El componente y^c se puede obtener aplicando *filtros reductores de tendencia* (filtros pasa-alta) basados en transformaciones $MA(k)$ simétricas. Baxter y King demuestran que cuando $\sum_{h=-k}^k a_h = 0$, y_t^* no tiene tendencia si el componente de crecimiento de y_t viene generado por componentes deterministas (lineales o cuadráticos) o por procesos $I(1)$ o $I(2)$.

Obsérvese que la condición de reducción de tendencia, $\sum_{h=-k}^k a_h = 0$, implica que la función de respuesta frecuencial satisface $\alpha(0) = 0$. Por tanto, el espectro de y_t^c es cero en la frecuencia cero, por lo que y_t^* se asocia con el componente del ciclo (y_t^c) e ($y_t - y_t^c$) con la tendencia (y_t^g). Estos filtros reductores de tendencia se denominan *filtros pasa-alta* al pasar los componentes de los datos con frecuencia mayor que un valor predeterminado, $\underline{\omega}$, próximo a 0. Esto es, $\beta(\omega) = 0$ para $|\omega| < \underline{\omega}$ y $\beta(\omega) = 1$ para $|\omega| \geq \underline{\omega}$.

Por otra parte los, *filtros pasa-baja* se construyen de forma que $\beta(\omega) = 0$ para $|\omega| > \underline{\omega}$ y $\beta(\omega) = 1$ para $|\omega| \leq \underline{\omega}$ y, por tanto, para frecuencias bajas, (movimientos a largo plazo) permanecen inalteradas, en tanto que otras se eliminan. En términos del filtro $MA(k)$ simétrico finito, esto implica que los filtros pasa-baja deben cumplir $\sum_{h=-k}^k a_h = 1$.

Baxter y King (1995, 1999) demostraron que se puede obtener una aproximación a un *filtro pasa-baja* ‘ideal’ si tomamos los coeficiente del filtro $MA(k)$ simétrico bi-direccional iguales a $a_0 = \frac{1}{\pi}\underline{\omega}$ y $a_h = \frac{1}{h\pi} \sin(hw)$ para $h = 1, 2, 3, \dots$. Por tanto, el filtro *pasa-alta* complementario tiene coeficientes $(1 - a_0)$ en $h = 0$ y $-a_h$ para $h = 1, 2, 3, \dots$

Cuando el filtro pasa las frecuencias entre $\underline{\omega}$ y $\bar{\omega}$, del espectro, siendo $0 < |\underline{\omega}| < |\bar{\omega}| < \pi$, tenemos un filtro *pasa-banda* y puede obtenerse, por ejemplo, restando dos filtros *pasa-baja*.

Por lo general, las frecuencias se asocian con la duración de los ciclos del calendario del NBER a partir de la definición de Burns y Mitchell (1946) donde $\underline{\omega}$ corresponde a 32 trimestres (8 años) y $\bar{\omega}$ a 6 trimestres (1.5 o 2 años). Así pues, el componente tendencia se obtiene a partir de un filtro *pasa-baja* con frecuencias de ocho años o menos. Este es el filtro que denominamos BK en nuestros experimentos posteriores.

3.2. Filtro de Hodrick y Prescott (HP). El filtro Hodrick y Prescott (1980, 1997) se utiliza habitualmente en Macroeconomía para eliminar la tendencia de las series para estudiar los 'hechos estilizados' de una economía a lo largo del ciclo de negocios. La base del filtro es la siguiente: a partir de (3.1) se define el componente tendencial como la solución del siguiente problema de optimización:

$$\min_{\{y_t^g\}} \sum_{t=1}^T \left[(y_t - y_t^g)^2 + \lambda (\Delta^2 y_{t+1}^g)^2 \right] \quad (3.5)$$

Podemos interpretar el primer término de (3.5) como una medida de bondad de ajuste de la tendencia a la serie observada, en tanto que el segundo término impone una penalización para conseguir una tendencia suave. Los valores del parámetro λ sugeridos por Kydland y Prescott (1990) son $\lambda = 1600$ para datos trimestrales y $\lambda = 400$ para datos anuales, dado que consideran que son éstos los ratios de la volatilidad de los componentes irregulares relativas a las volatilidades de los componentes tendenciales (el ratio señal-ruido).

Expresando el problema en función del operador backward shift, B , la descomposición es

$$y_t = F(B)^{-1} y_t + C(B) y_t \quad (3.6)$$

con $y_t^g = F(B)^{-1} y_t$ y $y_t^c = C(B) y_t$, donde los polinomios en B (filtros) son

$$F(B) = \lambda B^2 - 4\lambda B + (6\lambda + 1) - 4\lambda B^{-1} + \lambda B^{-2} = \lambda(1 - B)^2(1 - B^{-1})^2 + 1$$

y

$$C(B) = \frac{\lambda(1 - B)^2(1 - B^{-1})^2}{\lambda(1 - B)^2(1 - B^{-1})^2 + 1},$$

que es lo que habitualmente se entiende como filtro HP en Macroeconomía. Nótese que $F(1) = 1$ and $C(1) = 0$.

Las propiedades del filtro HP son objeto de estudio en una serie de artículos como Harvey y Jaeger (1993) y King y Rebelo (1993).

3.3. Filtro de la mediana. Los filtros HO y BK son ejemplos de filtros lineales. Para mayor exhaustividad, también se considera una clase de filtros no lineales denominada *filtros de la mediana* (Wen y Zeng, 1999), que en tiempos recientes se ha utilizado con una gran utilidad en procesamiento de señales en distintos campos de Ingeniería. los filtros de la mediana tienen dos propiedades muy interesantes: preservan el filo (cambios bruscos de tendencia) y atenúan el ruido de forma robusta frente a innovaciones de tipo impulso. Ninguna de estas propiedades puede obtenerse mediante los filtros lineales tradicionales. Para calcular el filtro de la mediana, se toma una ventana de datos consecutivos de la serie observada y se ordena para obtener el valor central o mediana, que es el output del filtro. Si la longitud (de la ventana) del filtro es $2n + 1$, el procedimiento de filtrado se denota como

$$\text{med}\{y_{t-n}, y_{t-n+1}, \dots, y_t, \dots, y_{t+n}\}. \quad (3.7)$$

En el caso del filtro de la mediana carece de sentido el análisis del dominio de frecuencias y la función de respuesta, puesto que la función de respuesta a impulsos del filtro de la mediana es cero siempre. No obstante, una propiedad muy importante del filtro de la mediana es la denominada *convergencia en raíz*, es decir, que cualquier serie temporal finita contiene una serie de señales que son invariantes a filtro de la mediana. Desde un punto de vista económico esto es muy importante, puesto que si se produce alguna ruptura en los fundamentos económicos de la serie, puede que se vea perfectamente tras realizar la operación de filtrado. Ver Wen y Zeng (1999) para más detalles. Esta propiedad es particularmente importante para nuestro propósito, tal y como se comprobará en los experimentos de simulación de Monte Carlo de la Sección 5.

4. Contraste MCE basado en Componentes Tendenciales

Al descomponer la serie z_t como en (3.1) se obtiene

$$z_t = z_t^g + z_t^c \quad (4.1)$$

Si las series reales son $I(1)$ e $(y_t - \alpha z_t)$ es $I(0)$ están cointegradas. En términos de los componentes no observados, se puede expresar

$$y_t - \alpha z_t = (y_t^g - \alpha z_t^g) + (y_t^c - \alpha z_t^c) \quad (4.2)$$

Sea $\bar{b}(B)$ un filtro genérico de medias móviles en el que se impone una serie de restricciones en los coeficientes \bar{b}_k de modo que sea un filtro pasa-baja (ver Sección 3.1

para más detalles), y llamemos $y_t^g = \bar{b}(B)y_t$ al componente tendencial correspondiente. Entonces, si multiplicamos la ecuación (5.1a) por $\bar{b}(B)$ se obtiene

$$\Delta \bar{b}(B)y_t = \bar{b}(B)c_t + a\Delta \bar{b}(B)z_t + b[\bar{b}(B)y_{t-1} - \alpha \bar{b}(B)z_{t-1}] + \bar{b}(B)u_{1,t} \quad (4.3)$$

que es el modelo MCE sobre los componentes tendenciales

$$\Delta y_t^g = c_t^g + a\Delta z_t^g + b[y_{t-1}^g - \alpha z_{t-1}^g] + \bar{b}(B)u_{1,t} \quad (4.4)$$

Puesto que $\bar{b}(B)u_{1,t}$ puede presentar correlación serial, se considera una versión dinámica de MCE sobre los componentes tendenciales dada por

$$\phi_y(B)\Delta y_t^g = c_t^g + a\phi_z(B)\Delta z_t^g + b[y_{t-1}^g - \alpha z_{t-1}^g] + \eta_t \quad (4.5)$$

donde η_t se considera ruido blanco y los retardos de $\phi_y(B)\Delta y_t^g$ y $\phi_z(B)\Delta z_t^g$ se determinan por el criterio SBIC. Es de esperar que, para un nivel suavizado suficiente, podamos aproximar c_t^g mediante una constante o una tendencia lineal.

A partir de la ecuación (5.1a) se puede escribir el modelo MCE para una descomposición genérica tendencia-ciclo como

$$\Delta(y_t^g + y_t^c) = c_t + a\Delta(z_t^g + z_t^c) + b[(y_{t-1}^g + y_{t-1}^c) - \alpha(z_{t-1}^g + z_{t-1}^c)] + u_{1,t} \quad (4.6)$$

y, agrupando términos, las ecuaciones (4.6) y (5.1d) puede expresarse como

$$\Delta y_t^g = c_t^* + a\Delta z_t^g + b(y_{t-1}^g - \alpha z_{t-1}^g) + u_{1,t} \quad (4.7a)$$

$$\begin{aligned} c_t^* = & (\Delta \mu_{y,t} - \Delta y_t^c) - a(\Delta \mu_{z,t} - \Delta z_t^c) \\ & - b[(\mu_{y,t-1} - \alpha \mu_{z,t-1}) - (y_{t-1}^c - \alpha z_{t-1}^c)] \end{aligned} \quad (4.7b)$$

En la práctica, dado que no sabemos donde están los saltos, necesitamos aproximar (4.7a)–(4.7b) mediante

$$\phi_y(B)\Delta y_t^g = c_0^* + a\phi_z(B)\Delta z_t^g + b[y_{t-1}^g - \alpha z_{t-1}^g] + \epsilon_t \quad (4.8)$$

La cuestión es si el contraste MCE basado en el t -ratio (t_b) de la ecuación (4.8) es robusto a la presencia de observaciones atípicas en las series

5. Experimento de Simulación de Monte Carlo

El proceso generador de datos (PGD) se basa en diversas extensiones de los utilizados por Kremers et al. (1992) y Campos et al. (1996). Se trata de un proceso

vectorial lineal auto-regresivo de primer orden con perturbaciones Normales causalidad en sentido de Granger unidireccional ($z \rightarrow y$), y posibles shocks en las variables fuertemente exógenas (z_t) para los parámetros de interés , a , b , y α .

5.1. El modelo. Nuestro PGD se basa en

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (5.1a)$$

$$\Delta z_t = \Delta \mu_{z,t} + u_{2t} \quad (5.1b)$$

$$c_t = \Delta \mu_{y,t} - a\Delta \mu_{z,t} - b(\mu_{y,t-1} - \alpha \mu_{z,t-1}) \quad (5.1c)$$

$$\mu_{z,t} = \mu_{z,o} + s_z \delta_{j,t}^z \quad (5.1d)$$

$$\begin{cases} \Pr(\delta_t^z = 1) = \Pr(\delta_t^z = -1) = \frac{\pi}{2} \\ \Pr(\delta_t^z = 0) = 1 - \pi \end{cases} \quad (5.1e)$$

En el caso de que no haya co-rupturas de ningún tipo, se añade otro shock a μ_y dado por

$$\mu_{y,t} = \mu_{y,o} + s_y \delta_{j,t}^y \quad (5.1f)$$

donde $\delta_{j,t}^y$ sigue un proceso estocástico similar a $\delta_{j,t}^z$ en (5.1e) pero mutuamente independientes.

Sin pérdida de generalidad, hacemos $\sigma_1^2 = 1$, y $\alpha = 1$. por tanto, las variables del experimento diseñado son los parámetros a, b, s donde $\sigma_2 = s$, y el tamaño muestral, T . El experimento es un diseño factorial completo con:

$a = 0.0, 0.5, 1$ (correlación contemporánea en primeras diferencias)

$b = 0.0$ (no cointegración), $b < 0$ (cointegración)

$\pi = 0$ (no atípicos), 0.05, 0.1.

$s = 1, 6, 16$ (tamaño de los saltos)

$T = 100, 200, 500, 1000$ (tamaño muestral)

Esto representa 144 experimentos para cada valor de b . Recordemos que cuando $a = 1$ se cumple la restricción de factor común (COMFAC) en el modelo de corrección del error, puesto que $\alpha = 1$.

Los experimentos de Monte Carlo se basan en 2000 réplicas de cada experimento eliminando las primeras 50 observaciones de cada serie simulada para considerar condiciones iniciales aleatorias.

Para obtener los valores críticos empíricos, se generan las series y_t y z_t a partir del PGD (5.1a)–(5.1e) bajo $H_0 : b = 0$ y se estiman los tres modelos siguientes:

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (\text{Modelo 1})$$

$$\phi(B)\Delta y_t = c + \theta(B)\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (\text{Modelo 2})$$

$$\phi(B)\Delta y_t^g = c + \theta(B)\Delta z_t^g + b(y_{t-1}^g - z_{t-1}^g) + u_{1t} \quad (\text{Modelo 3})$$

donde se ha impuesto $\alpha = 1$ y los órdenes de los polinomios $\phi(B)$ y $\theta(B)$ se eligen con el criterio SBIC². Las variables y_t^g y z_t^g son los componentes tendenciales obtenidos al aplicar los filtros de Hodrick y Prescott (HP10 y HP100), Baxter y King (BK) y de la mediana (MD). Ver Sección 3 para más detalles. El valor crítico considerado es el cuantil 5% de la cola inferior de la distribución empírica del estadístico t -ratio, $t(\hat{b})$, bajo H_0 . La potencia empírica se calcula de forma análoga, simulando el PGD (5.1a)–(5.1e) bajo $H_1 : b \neq 0$, y calculando el porcentaje de rechazos obtenido de los contrastes con los Modelos 1, 2 y 3, utilizando los valores críticos empíricos calculados previamente.

Se impone co-ruptura simultánea haciendo $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t} - \alpha\mu_{z,t}) = 0$. Para obtener co-ruptura sólo en diferencias, se toma $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d = 0.5$. Por otra parte, para simular series con co-rupturas sólo en niveles, imponemos $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = 0$, ver Capítulo 2 para más detalles de este apartado.

Hay que resaltar que si hubiésemos tomado $c_d = 0$, los valores críticos obtenidos serían los mismos que para el caso de co-ruptura simultánea, puesto que bajo $H_0 : b = 0$ se tiene co-ruptura simultánea con $c_s = 0$. Además, para $a = \alpha = 1$ la co-ruptura en diferencias implicaría co-ruptura en niveles (co-ruptura completa). Esto se debe a que, cuando se satisface la restricción COMFAC, co-ruptura en niveles implica co-ruptura en diferencias.

5.2. Contraste MCE basado en componentes tendenciales: No atípicos.

La figura 4.1 representa el estimador kernel de la densidad de la distribución empírica de $t(\hat{b})$ para el Modelo 2 (contraste MCE) y el Modelo 3 (contraste MCE filtrado) usando los filtros HP10, HP100, BK y MD respectivamente. con todos los filtros empleados se obtienen resultados similares. Es importante resaltar que los mejores resultados se obtienen para las colas inferiores de las distribuciones. Esto implica que los valores críticos empleados para contrastar la hipótesis de no cointegración frente

²Para elegir el número de retardos de cada variable en la regresión se permiten entre 0 y 10 retardos de cada variable, y se exploran las 121 posibles combinaciones, excepto en el caso $T = 25$, que no se permiten más de 4 retardos de cada variable.

a la alternativa de cointegración van a ser robustos frente al tipo de filtro empleado, ver Tabla 4.1.

a	T=100	T=200	T=500	T=1000
Modelo 1				
0.0	-2.624	-2.518	-2.654	-2.639
0.5	-2.769	-2.738	-2.758	-2.752
1.0	-2.875	-2.842	-2.878	-2.929
Modelo 2				
0.0	-2.596	-2.495	-2.641	-2.652
0.5	-2.801	-2.787	-2.754	-2.789
1.0	-2.951	-2.868	-2.862	-2.949
Modelo 3. HP10				
0.0	-3.017	-2.670	-2.668	-2.614
0.5	-3.272	-3.071	-2.859	-2.874
1.0	-3.481	-3.164	-2.940	-2.913
Modelo 3. HP100				
0.0	-3.238	-2.726	-2.650	-2.608
0.5	-3.707	-3.071	-2.832	-2.797
1.0	-4.114	-3.085	-2.898	-2.891
Modelo 3. BK				
0.0	-2.889	-2.650	-2.667	-2.592
0.5	-3.069	-2.750	-2.676	-2.752
1.0	-3.021	-2.783	-2.889	-2.958
Modelo 3. MD				
0.0	-2.749	-2.616	-2.654	-2.664
0.5	-2.816	-2.599	-2.561	-2.652
1.0	-2.859	-2.715	-2.736	-2.778

TABLA 4.1. Valores críticos de los contrastes MCE (Modelo 1 y Modelo 2) y MCE filtrado (Modelo 3). No atípicos ($\pi = 0$).

En términos de potencia, se observa una ligera pérdida al utilizar el filtro BK, y la pérdida es aún menor al utilizar el filtro MD, ver Figura 4.2 y Tabla 4.2. Los filtros HP10 y HP100 muestran las potencias más bajas entre las alternativas consideradas.

Para evaluar el tamaño de los contrastes basados en los Modelos 1–3 añadimos estructura $MA(1)$ a la perturbación $u_{1,t}$, es decir $u_{1,t} + \theta u_{1,t-1} = v_t$ con $v_t \sim iid N(0, 1)$, y el parámetro $MA(1)$ igual a $\theta = \pm 0.2$ y $\theta = \pm 0.5$. Los resultados empíricos para un tamaño nominal de 5% se muestran en la Tabla 4.3

Tal y como se esperaba, las mayores distorsiones de tamaño se generan en muestras pequeñas y con $\theta = -0.5$. en el caso del Modelo 2 con $T = 200$, $a = 1$, y $\theta = -0.5$ el 5% nominal se transforma en 20%. Este tamaño empírico se reduce a 6.55% con el filtro HP100, 3.15% con HP10, 6.65% con BK y 6.8% con el filtro MD. De nuevo, los resultados más estables en términos de tamaño se obtienen al utilizar el filtro de la mediana (MD).

5.3. Contraste MCE con atípicos aditivos. Con el fin de evaluar los efectos de la presencia de atípicos aditivos en los contrastes MCE nos vamos a centrar en la situación más favorable. Vamos a permitir retardos adicionales en los modelos MCE para aproximar el efecto de los atípicos y realizar el contraste MCE en el Modelo 2,

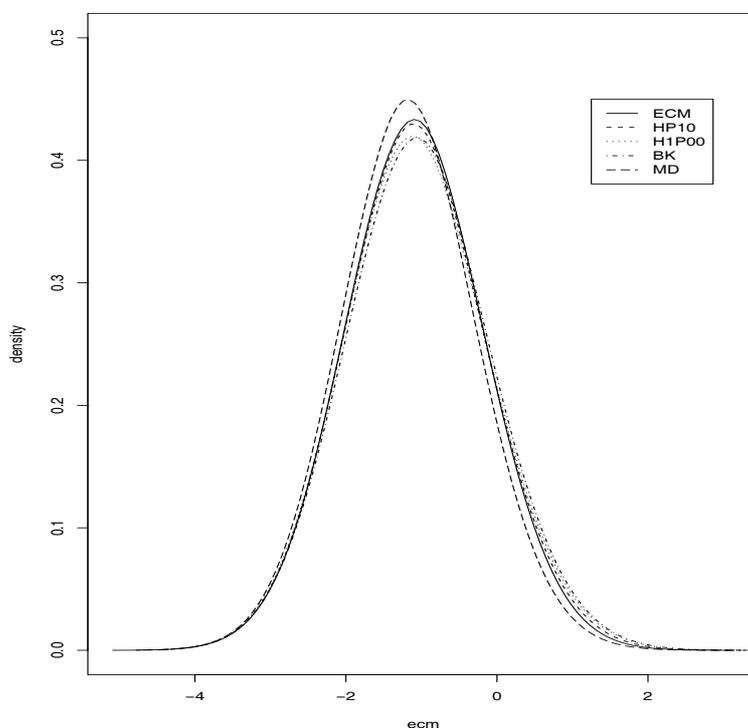


FIGURA 4.1. Estimador kernel de la densidad del estadístico del contraste MCE (Modelo 2), y MCE filtrado (Modelo 3) bajo la hipótesis nula. $T = 1000$, $a = 0$. No atípicos ($\pi = 0$)

a	T=100	T=200	T=500	T=1000
Modelo 1				
0.0	99.45	100.00	100.00	100.00
0.5	95.50	100.00	100.00	100.00
1.0	87.75	100.00	100.00	100.00
Modelo 2				
0.0	99.15	100.00	100.00	100.00
0.5	93.05	100.00	100.00	100.00
1.0	79.70	99.95	100.00	100.00
Modelo 3. HP10				
0.0	51.40	98.05	100.00	100.00
0.5	30.15	76.75	100.00	100.00
1.0	21.50	76.95	99.95	100.00
Modelo 3. HP100				
0.0	31.75	81.05	100.00	100.00
0.5	15.55	54.45	99.75	100.00
1.0	11.25	42.80	99.25	100.00
Modelo 3. BK				
0.0	74.40	99.60	100.00	100.00
0.5	46.15	92.85	100.00	100.00
1.0	21.25	76.05	100.00	100.00
Modelo 3. MD				
0.0	88.10	99.65	100.00	100.00
0.5	58.35	97.55	100.00	100.00
1.0	30.60	85.00	100.00	100.00

TABLA 4.2. Potencias de los contrastes MCE (Modelo 1 y Modelo 2) y MCE filtrado (Modelo 3). $b = -0.2$. No atípicos ($\pi = 0$).

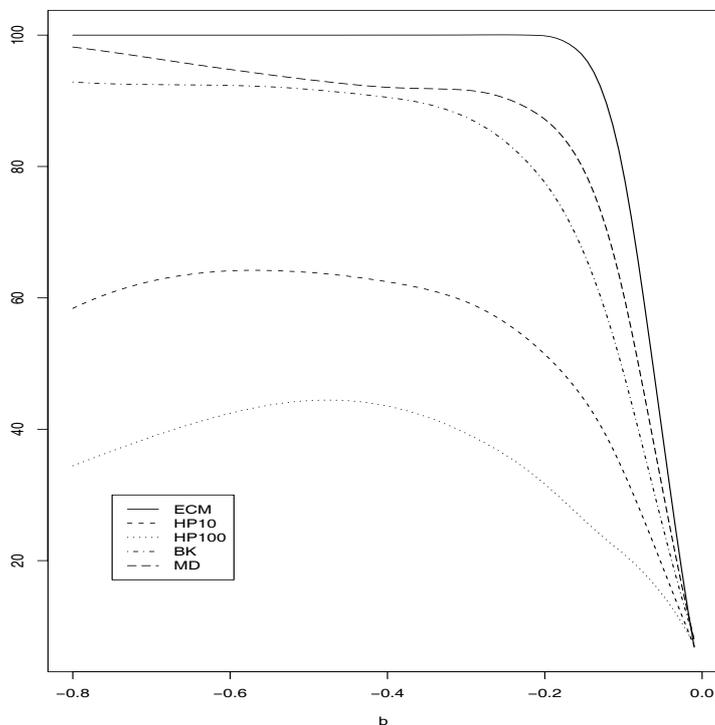


FIGURA 4.2. Comparación de potencias de los contrastes MCE (Modelos 2 y 3). $T = 100$, $a = 0$. No atípicos ($\pi = 0$)

en el que los órdenes de los polinomios $\varphi(B)$ y $\theta(B)$ se eligen mediante el criterio SBIC.

En la Figura 4.3.1 se observa que los efectos de tener 10% de contaminación de atípicos aditivos son dramáticos. Toda la distribución se desplaza a la izquierda y, por consiguiente, se detectará demasiada cointegración con los valores críticos habituales. Este resultado es congruente con la evidencia de Franses y Haldrup (1994).

Contraste MCE en el Modelo 1 con atípicos aditivos. El modelo en este caso es

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (\text{Modelo 1})$$

que sólo está correctamente especificado en el caso de *co-ruptura simultánea*. La razón es que en ese caso c_t de (5.1d) es una constante (igual a 0).

Se observa que los valores críticos (VC) son muy estables en el caso de *co-ruptura simultánea* y que sólo se ven afectados de forma marginal por el tamaño muestral (T), el parámetro a y el tamaño del salto (s).

a	θ	T=100	T=200	T=500	T=1000
0.0	-0.5	2.40	2.60	2.05	1.95
	-0.2	5.00	4.80	4.20	3.90
	0.2	6.30	6.10	5.25	5.95
	0.5	8.10	8.00	5.55	6.10
0.5	-0.5	13.10	9.60	6.45	5.10
	-0.2	8.20	7.40	6.60	5.80
	0.2	4.25	4.95	5.30	5.45
	0.5	6.95	6.90	6.30	6.05
1.0	-0.5	31.20	20.05	12.50	8.55
	-0.2	11.70	11.40	7.50	5.35
	0.2	3.75	5.15	5.30	5.00
	0.5	6.75	6.50	4.70	3.75

Tabla 4.3.1: Modelo 2

a	θ	T=100	T=200	T=500	T=1000
0.0	-0.5	2.75	3.50	2.30	2.35
	-0.2	5.25	5.85	4.35	4.25
	0.2	4.95	5.75	5.50	5.30
	0.5	6.00	7.35	6.70	6.20
0.5	-0.5	2.55	2.80	4.35	4.35
	-0.2	5.00	5.55	6.95	5.90
	0.2	4.45	3.90	4.10	4.40
	0.5	4.05	3.80	4.20	4.25
1.0	-0.5	1.25	6.55	8.05	8.40
	-0.2	4.25	8.70	6.95	6.60
	0.2	2.75	3.90	3.60	4.00
	0.5	1.90	4.40	4.75	5.00

Tabla 4.3.2: Modelo 3. HP100

a	θ	T=100	T=200	T=500	T=1000
0.0	-0.5	2.05	2.80	2.40	2.55
	-0.2	3.25	4.85	4.15	4.05
	0.2	5.30	4.70	5.20	5.50
	0.5	5.65	6.10	5.95	6.05
0.5	-0.5	3.25	1.65	3.10	2.70
	-0.2	4.00	3.40	4.60	5.05
	0.2	5.65	4.10	3.95	3.85
	0.5	4.85	3.50	3.20	3.95
1.0	-0.5	2.45	3.15	4.75	5.75
	-0.2	4.10	3.75	6.35	6.85
	0.2	4.95	3.25	2.90	4.20
	0.5	2.95	2.10	3.75	4.55

Tabla 4.3.3: Modelo 3. HP10

a	θ	T=100	T=200	T=500	T=1000
0.0	-0.5	2.70	2.80	3.05	2.90
	-0.2	3.65	4.20	4.45	4.35
	0.2	5.45	5.30	5.35	6.25
	0.5	6.15	6.90	6.85	8.05
0.5	-0.5	5.20	5.45	5.05	3.50
	-0.2	5.50	4.95	5.05	4.80
	0.2	5.60	5.05	5.75	6.00
	0.5	5.45	7.85	9.25	7.90
1.0	-0.5	10.90	6.65	3.85	3.85
	-0.2	5.80	4.40	3.75	4.20
	0.2	4.90	7.00	6.25	5.25
	0.5	7.00	9.35	7.40	6.05

Tabla 4.3.4: Modelo 3. BK Filter

a	θ	T=100	T=200	T=500	T=1000
0.0	-0.5	5.05	4.70	3.05	2.15
	-0.2	5.30	4.65	4.30	3.35
	0.2	4.95	5.25	5.65	6.20
	0.5	5.95	5.90	6.15	7.55
0.5	-0.5	5.95	5.60	4.60	2.60
	-0.2	6.60	4.45	5.10	4.05
	0.2	4.80	5.05	6.10	6.00
	0.5	4.85	6.25	7.05	7.60
1.0	-0.5	8.90	6.80	4.75	4.85
	-0.2	5.15	5.40	4.85	5.20
	0.2	4.75	4.80	5.00	5.25
	0.5	5.00	4.85	5.45	5.45

Tabla 4.3.5: Modelo 3. MD Filter

Tabla 4.3. Tamaño de los contrastes MCE (Modelo 1 y Modelo 2) y MCE filtrado (Modelo 3). Modelo simulado con perturbaciones $MA(1)$. No atípicos ($\pi = 0$).

Co-ruptura en diferencias, no en niveles En este caso $c_t = c + bc_d t - b(a - \alpha)s_z \delta_{t-1}^z$. Por tanto, en Modelo 1 hay un término omitido incluso si $a = \alpha$ (COMFAC). Los resultados de la Tabla 4.4 muestran que los VC de $t(\hat{b})$ basado en Modelo 1 están desplazados a la derecha, cambian con T y sólo marginalmente con a y s (el tamaño del salto). Es de destacar que los valores están próximos a los de la Normal, pues hay una tendencia omitida.

Co-ruptura en niveles, no en diferencias. Ahora $c_t = c + (\alpha - a)s_z \Delta \delta_t^z$, y el Modelo 1 está mal especificado. No obstante, en este caso particular, si $\alpha = a$ (COMFAC), $c_t = c$ y tenemos co-ruptura simultánea, por lo que el Modelo no estaría mal especificado.

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Co-ruptura simultánea									
100	-2.631	-2.489	-2.350	-2.779	-2.694	-2.546	-2.883	-2.858	-2.849
200	-2.519	-2.481	-2.384	-2.742	-2.731	-2.649	-2.850	-2.841	-2.840
500	-2.652	-2.591	-2.533	-2.747	-2.748	-2.669	-2.879	-2.869	-2.869
1000	-2.640	-2.627	-2.544	-2.762	-2.774	-2.745	-2.935	-2.939	-2.944
Co-ruptura en diferencias, no en niveles									
100	-2.010	-1.989	-1.997	-1.982	-1.982	-1.985	-1.970	-1.965	-1.967
200	-1.967	-1.923	-1.911	-1.936	-1.899	-1.897	-1.916	-1.902	-1.914
500	-1.813	-1.836	-1.852	-1.849	-1.831	-1.837	-1.845	-1.840	-1.839
1000	-1.763	-1.773	-1.772	-1.765	-1.760	-1.767	-1.766	-1.761	-1.758
Co-ruptura en niveles, no en diferencias									
100	-2.577	-2.739	-2.792	-2.738	-2.765	-2.799	-2.883	-2.858	-2.849
200	-2.484	-2.754	-2.818	-2.735	-2.827	-2.844	-2.850	-2.841	-2.840
500	-2.578	-2.769	-2.827	-2.750	-2.814	-2.804	-2.879	-2.869	-2.869
1000	-2.544	-2.804	-2.891	-2.765	-2.885	-2.898	-2.935	-2.939	-2.944
Shocks independientes									
100	-2.650	-4.131	-5.426	-2.964	-4.751	-5.605	-3.293	-5.281	-5.756
200	-2.582	-4.856	-7.818	-2.971	-5.855	-8.405	-3.259	-6.660	-8.744
500	-2.683	-4.837	-9.457	-2.969	-5.994	-10.591	-3.253	-6.922	-11.229
1000	-2.667	-5.202	-11.207	-2.974	-6.437	-13.026	-3.281	-7.451	-13.953
Shocks en y_t									
100	-2.624	-4.613	-7.632	-2.864	-5.467	-8.132	-3.044	-5.905	-8.437
200	-2.588	-5.108	-9.899	-2.890	-6.246	-11.059	-3.076	-6.878	-11.683
500	-2.692	-5.122	-11.485	-2.886	-6.235	-13.401	-3.081	-7.075	-14.573
1000	-2.667	-5.337	-12.553	-2.887	-6.643	-15.205	-3.089	-7.321	-16.323
Shocks en z_t									
100	-2.631	-2.489	-2.350	-2.856	-2.967	-2.540	-3.080	-3.565	-2.857
200	-2.519	-2.481	-2.384	-2.822	-3.084	-2.839	-3.024	-3.708	-3.205
500	-2.652	-2.591	-2.533	-2.838	-3.127	-2.905	-3.060	-3.790	-3.433
1000	-2.640	-2.627	-2.544	-2.882	-3.203	-3.071	-3.094	-3.929	-3.724

TABLA 4.4. Valores críticos (5%) del contraste MCE. Modelo 1, $\pi = 0.1$

De la Tabla 4.4 se deduce que los VC son más estables con el valor de T , pero pasan de -2.5 a -2.9 dependiendo de la restricción COMFAC.

Shocks independientes o shocks en sólo una de las variables. En este caso c_t nunca es constante, ver ecuaciones (2.10a)–(2.10c), por lo que el Modelo 1 siempre presenta errores de especificación importantes. Incluso para tamaños muestrales grandes (1000), los VC están entre -3.09 y -16.3 dependiendo del tamaño del salto s , ver Tabla 4.4. Esta inestabilidad disminuye cuando no hay shocks en la variable endógena y_t . La pregunta que nos hacemos es cuánto mejorará la situación al incluir retardos adicionales en MCE y pasar al Modelo 2.

Contraste MCE en el Modelo 2 con atípicos aditivos. El modelo es

$$\phi(B)\Delta y_t = c + \theta(B)\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (\text{Modelo 2})$$

Bajo *co-ruptura simultánea* o con *co-ruptura parcial en niveles* los VC son muy estables, como en el caso del Modelo 1. Análogamente con *co-ruptura parcial en diferencias* los VC se despazan a la izquierda, pero son más inestables respecto al tamaño muestral (T) y bastante estables respecto a los otros dos parámetros a y s .

T	a = 0.0			a = 0.5			a = 1.0		
	s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
Co-ruptura simultánea									
100	-2.601	-2.533	-2.407	-2.813	-2.736	-2.661	-2.934	-2.918	-2.899
200	-2.512	-2.481	-2.402	-2.772	-2.711	-2.657	-2.871	-2.874	-2.860
500	-2.650	-2.597	-2.512	-2.737	-2.757	-2.669	-2.853	-2.851	-2.840
1000	-2.664	-2.621	-2.531	-2.782	-2.801	-2.774	-2.955	-2.956	-2.954
Co-ruptura en diferencias, no en niveles									
100	-2.068	-2.072	-2.039	-2.042	-2.068	-2.029	-2.001	-1.999	-2.009
200	-1.928	-1.918	-1.903	-1.926	-1.916	-1.937	-1.897	-1.906	-1.910
500	-1.824	-1.829	-1.838	-1.819	-1.823	-1.830	-1.809	-1.826	-1.832
1000	-1.730	-1.736	-1.727	-1.729	-1.721	-1.720	-1.729	-1.707	-1.707
Co-ruptura en niveles, no en diferencias									
100	-2.587	-2.778	-2.820	-2.767	-2.800	-2.812	-2.934	-2.918	-2.899
200	-2.517	-2.786	-2.843	-2.787	-2.825	-2.843	-2.871	-2.874	-2.860
500	-2.590	-2.776	-2.821	-2.746	-2.808	-2.827	-2.853	-2.851	-2.840
1000	-2.612	-2.831	-2.885	-2.765	-2.882	-2.901	-2.955	-2.956	-2.954
Shocks independientes									
100	-2.644	-2.695	-3.111	-3.009	-4.004	-3.516	-3.303	-4.997	-5.922
200	-2.571	-2.683	-3.563	-2.945	-3.593	-4.378	-3.253	-4.589	-4.919
500	-2.645	-2.481	-2.240	-2.893	-3.387	-3.060	-3.033	-4.455	-4.046
1000	-2.674	-2.509	-2.551	-2.954	-3.567	-3.597	-2.979	-4.108	-4.540
Shocks en y_t									
100	-2.624	-3.311	-7.673	-2.866	-5.336	-8.067	-3.063	-5.921	-8.309
200	-2.555	-2.780	-4.352	-2.891	-3.631	-6.740	-3.064	-4.249	-10.992
500	-2.662	-2.635	-3.281	-2.841	-3.308	-4.634	-2.999	-3.420	-5.473
1000	-2.670	-2.686	-3.437	-2.880	-3.180	-4.664	-3.029	-3.217	-4.967
Shocks en z_t									
100	-2.601	-2.533	-2.407	-2.850	-2.982	-2.591	-3.121	-3.504	-2.887
200	-2.512	-2.481	-2.402	-2.806	-3.023	-2.793	-3.029	-3.659	-3.205
500	-2.650	-2.597	-2.512	-2.806	-3.031	-2.902	-2.968	-3.582	-3.425
1000	-2.664	-2.621	-2.531	-2.866	-3.057	-3.051	-3.020	-3.430	-3.661

TABLA 4.5. Valores críticos (5%) del contraste MCE. Modelo 2, $\pi = 0.1$

En la Tabla 4.5 se observa claramente que las mejoras del Modelo 2 sobre el Modelo 1 ocurren cuando hay *shocks independientes* o cuando ocurren *sólo en una de las variables*. Sin embargo, esto no quiere decir que sean siempre estables, pues que los shocks ocurren sólo en la y_t , los VC están entre -3 y -11 para una muestra de 200 observaciones. Por tanto, la inclusión de retardos adicionales mejora la situación, pero no es una solución satisfactoria.

La potencia del contraste depende en gran medida del tipo de co-ruptura, ver Tabla 4.6. En el caso de *co-ruptura parcial en diferencias* el contraste no tiene potencia, lo que se explica por el hecho de que estamos omitiendo una tendencia determinista con saltos. Por otra parte, la potencia del contraste es alta en los otros dos casos de co-ruptura, esto es, *co-ruptura simultánea* o *co-ruptura parcial en niveles*. En los casos en que hay *shocks independientes* o *shocks en una de las variables* la potencia depende del tipo y tamaño del shock, así como del parámetro a .

Contraste MCE en el Modelo 3 con atípicos aditivos. El modelo estimado es

$$\phi(B)\Delta y_t^g = c + \theta(B)\Delta z_t^g + b(y_{t-1}^g - z_{t-1}^g) + u_{1t} \quad (\text{Modelo 3}).$$

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Co-ruptura simultánea									
100	99.65	99.90	100.00	94.75	97.65	99.85	84.95	85.95	86.25
200	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Co-ruptura en diferencias, no en niveles									
100	5.05	5.90	7.30	0.65	0.55	0.60	0.00	0.05	0.00
200	2.85	5.55	6.20	0.40	0.30	0.15	0.00	0.00	0.00
500	0.50	4.00	5.10	0.00	0.00	0.00	0.00	0.00	0.00
1000	0.05	2.00	5.25	0.00	0.05	0.10	0.00	0.00	0.00
Co-ruptura en niveles, no en diferencias									
100	99.70	93.95	89.10	95.90	90.75	89.15	84.95	85.95	86.25
200	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks independientes									
100	99.65	90.20	44.80	94.55	52.80	14.30	84.95	48.05	3.70
200	100.00	99.95	95.60	100.00	98.85	69.00	99.85	92.80	51.15
500	100.00	100.00	100.00	100.00	100.00	98.40	100.00	100.00	89.05
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks en y_t									
100	99.60	96.00	81.40	96.00	80.75	70.45	87.15	80.05	58.50
200	100.00	100.00	99.85	100.00	100.00	99.45	100.00	99.55	98.80
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks en z_t									
100	99.60	93.10	40.10	94.65	65.95	19.45	85.10	38.05	10.25
200	100.00	99.95	85.20	100.00	98.90	48.30	99.95	89.50	23.95
500	100.00	100.00	99.80	100.00	100.00	92.50	100.00	100.00	69.70
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	97.05

TABLA 4.6. Potencia ajustada del contraste MCE. Modelo 2, $\pi = 0.1$, $b = -0.2$.

La intuición para esperar que el Modelo 3 sea una buena aproximación al modelo sin error de especificación es la siguiente: la ecuación (2.2a) se transforma en la ecuación (4.7a) basada en los componentes tendenciales, donde c_t^* es un término estocástico definido por la ecuación (4.7b). Dado que los atípicos aditicos son shocks transitorios, deben formar parte del componente tendencial al hacer la descomposición tendencia-ciclo y, así, los siguientes elementos deben estar libres de observaciones atípicas

$$(\Delta\mu_{y,t} - \Delta y_t^g), (\Delta\mu_{z,t} - \Delta z_t^g), (\mu_{y,t-1} - y_{t-1}^g), (\mu_{z,t-1} - z_{t-1}^g)$$

La pendiente estocástica c_t^* de (4.7b) debe ser una serie estacionaria que se puede aproximar mediante retardos de Δy_t^g y Δz_t^g , que es exactamente lo que hace el Modelo 3.

Las Tablas 4.7–4.10 muestran los VC obtenidos bajo las seis casos de co-ruptura y no co-ruptura analizados con anterioridad. Independientemente del filtro utilizado, los VC son estables para los distintos tamaños muestrales, el parámetro a y el tamaño del salto, s . El único caso en que los VC dependen de T es cuando hay co-ruptura parcial en diferencias. El análisis se realiza con los filtros HP10, HP100, BK y MD. Es destacar la robustez de los valores críticos a los valores de a y s incluso con un

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Co-ruptura simultánea									
100	-3.264	-3.184	-3.086	-3.725	-3.671	-3.669	-4.114	-4.048	-3.894
200	-2.725	-2.763	-2.701	-3.058	-2.958	-2.875	-3.073	-3.021	-2.959
500	-2.657	-2.634	-2.652	-2.819	-2.777	-2.788	-2.893	-2.866	-2.865
1000	-2.611	-2.589	-2.587	-2.801	-2.795	-2.791	-2.898	-2.910	-2.881
Co-ruptura en diferencias, no en niveles									
100	-2.756	-2.732	-2.633	-2.915	-2.721	-2.625	-2.878	-2.690	-2.627
200	-2.187	-2.158	-2.143	-2.234	-2.169	-2.160	-2.203	-2.185	-2.132
500	-1.920	-1.920	-1.933	-1.923	-1.909	-1.902	-1.925	-1.931	-1.911
1000	-1.830	-1.824	-1.821	-1.832	-1.830	-1.818	-1.831	-1.832	-1.815
Co-ruptura en niveles, no en diferencias									
100	-3.477	-3.783	-3.677	-3.740	-3.834	-3.688	-4.114	-4.048	-3.894
200	-2.836	-2.952	-2.931	-3.035	-2.986	-2.924	-3.073	-3.021	-2.959
500	-2.733	-2.817	-2.820	-2.790	-2.828	-2.853	-2.893	-2.866	-2.865
1000	-2.701	-2.864	-2.873	-2.806	-2.888	-2.886	-2.898	-2.910	-2.881
Shocks independientes									
100	-3.393	-3.119	-2.739	-3.725	-3.166	-2.886	-4.054	-3.526	-3.085
200	-2.782	-2.740	-2.650	-3.074	-2.981	-2.679	-3.318	-3.298	-2.952
500	-2.682	-2.707	-2.673	-2.945	-2.653	-2.675	-2.987	-2.907	-2.767
1000	-2.623	-2.607	-2.639	-2.873	-2.714	-2.740	-2.972	-2.891	-2.898
Shocks en y_t									
100	-3.403	-3.107	-3.142	-3.826	-3.248	-3.170	-4.069	-3.598	-3.317
200	-2.785	-2.746	-2.666	-3.109	-2.950	-2.779	-3.268	-3.059	-2.864
500	-2.668	-2.697	-2.673	-2.868	-2.792	-2.760	-2.951	-2.938	-2.813
1000	-2.629	-2.621	-2.624	-2.850	-2.867	-2.831	-2.971	-2.999	-2.956
Shocks en z_t									
100	-3.264	-3.184	-3.086	-3.655	-3.259	-3.179	-3.991	-3.657	-3.288
200	-2.725	-2.763	-2.701	-3.070	-2.721	-2.618	-3.166	-2.864	-2.755
500	-2.657	-2.634	-2.652	-2.887	-2.620	-2.463	-2.885	-2.776	-2.626
1000	-2.611	-2.589	-2.587	-2.838	-2.567	-2.477	-2.923	-2.790	-2.641

TABLA 4.7. Valores críticos (5%) del contraste MCE filtrado. Modelo 3, HP100, $\pi = 0.1$

10% de atípicos ($\pi = 0.1$) y no hay ninguna co-ruptura, especialmente en los casos HP100 y MD.

La potencia del contraste se analiza en la Tablas 4.11–4.14. Es de destacar en primer lugar que en el caso de co-ruptura parcial en diferencias no hay potencia, tal y como sucedía en el Modelo 2. Es interesante que la potencia del contraste dependa de a en muchos casos. Como era de esperar, la potencia del Modelo 3 es menor que la del Modelo 2, aunque esto no sea así en todos los casos. En concreto, debemos reseñar que cuando tenemos shocks sólo en la y_t , con $a = 1$ y $s = 1, 16$, el contraste más potente es el basado en el Modelo 3 con el filtro BK para $T = 100$. En el caso de shocks en z_t , con $a = 1$ el contraste más potente para $T = 100, 200$ es el basado en el filtro MD. Otra propiedad del contraste MCE basado en el Modelo 3 con el filtro MD es que, aparte del caso de co-ruptura en diferencias que se ha comentado anteriormente, la potencia no depende del tipo de co-ruptura o no co-ruptura considerado. Además, los contrastes que utilizan el filtro MD son los que tienen potencia más alta con el Modelo 3.

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Co-ruptura simultánea									
100	-2.978	-2.921	-2.785	-3.267	-3.204	-3.090	-3.442	-3.314	-3.248
200	-2.697	-2.661	-2.591	-3.058	-3.066	-3.028	-3.175	-3.148	-3.128
500	-2.662	-2.659	-2.576	-2.863	-2.818	-2.756	-2.948	-2.912	-2.901
1000	-2.608	-2.608	-2.612	-2.874	-2.805	-2.783	-2.906	-2.907	-2.900
Co-ruptura en diferencias, no en niveles									
100	-2.343	-2.362	-2.320	-2.402	-2.308	-2.277	-2.417	-2.297	-2.263
200	-2.037	-2.027	-2.007	-2.027	-1.999	-1.957	-2.006	-1.982	-1.933
500	-1.834	-1.839	-1.852	-1.820	-1.825	-1.840	-1.814	-1.819	-1.830
1000	-1.771	-1.776	-1.781	-1.773	-1.778	-1.771	-1.774	-1.772	-1.783
Co-ruptura en niveles, no en diferencias									
100	-3.054	-3.210	-3.160	-3.299	-3.264	-3.157	-3.442	-3.314	-3.248
200	-2.737	-3.113	-3.137	-3.085	-3.142	-3.133	-3.175	-3.148	-3.128
500	-2.711	-2.900	-2.915	-2.863	-2.888	-2.882	-2.948	-2.912	-2.901
1000	-2.770	-2.852	-2.873	-2.850	-2.911	-2.913	-2.906	-2.907	-2.900
Shocks independentes									
100	-2.984	-2.836	-2.732	-3.266	-2.809	-2.946	-3.400	-2.953	-3.094
200	-2.725	-2.594	-2.661	-2.949	-2.616	-2.989	-3.123	-2.815	-3.140
500	-2.673	-2.656	-2.835	-2.807	-2.656	-3.180	-3.003	-2.941	-3.494
1000	-2.620	-2.668	-2.997	-2.859	-2.705	-3.486	-2.949	-2.970	-3.794
Shocks en y_t									
100	-2.979	-2.926	-2.851	-3.247	-3.130	-3.057	-3.412	-3.236	-3.292
200	-2.713	-2.596	-2.628	-3.037	-2.818	-2.974	-3.128	-2.879	-3.252
500	-2.674	-2.645	-2.792	-2.846	-2.692	-3.089	-3.020	-2.843	-3.420
1000	-2.621	-2.637	-2.775	-2.910	-2.782	-3.115	-2.985	-2.907	-3.324
Shocks en z_t									
100	-2.978	-2.921	-2.785	-3.294	-2.908	-2.666	-3.357	-3.251	-2.765
200	-2.697	-2.661	-2.591	-2.996	-2.675	-2.562	-3.155	-2.925	-2.673
500	-2.662	-2.659	-2.576	-2.847	-2.577	-2.474	-2.984	-2.667	-2.662
1000	-2.608	-2.608	-2.612	-2.882	-2.583	-2.541	-2.963	-2.752	-2.688

TABLA 4.8. Valores críticos (5%) del contraste MCE filtrado. Modelo 3, HP10, $\pi = 0.1$

5.4. Robustez de los valores críticos a la presencia de atípicos. La pregunta que nos hacemos es la siguiente: ¿Es seguro utilizar los valores críticos de la Tabla 4.1 para realizar el contraste MCE si hay observaciones atípicas?

Para responder a esta pregunta, se calcula el porcentaje de rechazos obtenidos al utilizar los valores críticos al 5% de la Tabla 4.1 cuando el PGD dado por las ecuaciones (5.1a)–(5.1f) presenta contaminación del peor tipo posible, es decir, shocks independientes en las variables y_t y z_t , bajo la hipótesis nula $H_0 : b = 0$. Si dichos valores críticos son robustos, dicho porcentaje de rechazos (tamaño), debería estar próximo al valor nominal del tamaño del contraste (5%).

Evidentemente, el Modelo 2 no es robusto, puesto que el porcentaje de rechazos toma valores de hasta 43.5% para muestras de tamaño $T = 100$ con $s = 16$, ver Figura 4.3.1. Al utilizar el Modelo 3 los resultados mejoran espectacularmente, ver Figuras 4.3.3–4.3.5. Con HP10 y BK, la cifra anterior se reduce a 21.8% y 26.1%, pero con HP100 y MD esos valores son 5.1% y 6.2%, respectivamente, ver Tabla 4.15. Recordemos que sólo se incluyen los resultados para el peor caso de contaminación. Para la mayoría de los parámetros estudiados todos los filtros funcionan bastante bien,

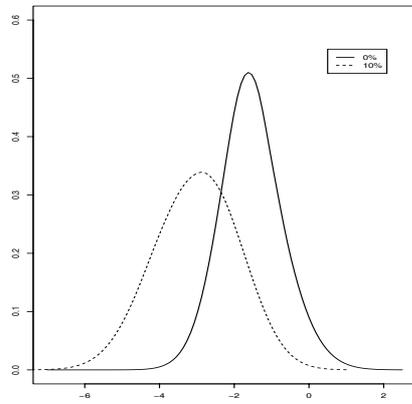


Figura 4.3.1: MCE

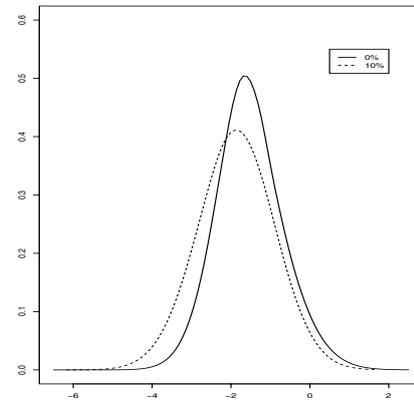


Figura 4.3.2: HP10

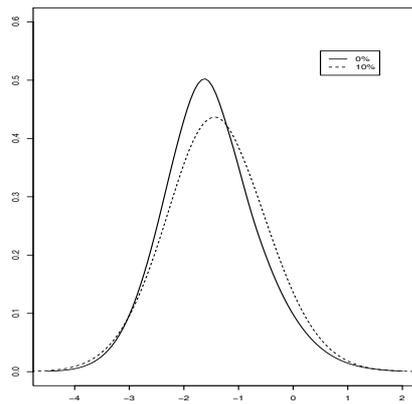


Figura 4.3.3: HP100

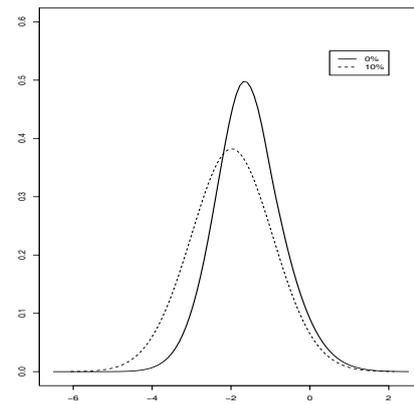


Figura 4.3.4: BK

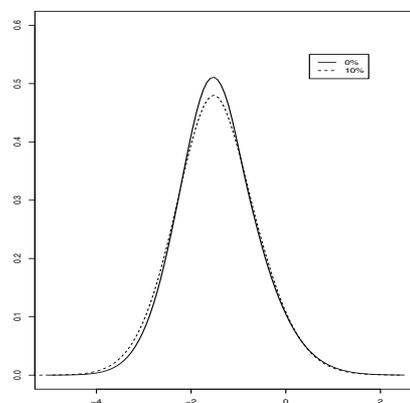


Figura 4.3.5: MD

FIGURA 4.3. Estimador Kernel de la densidad del estadístico del contraste MCE (Modelo 2) y MCE filtrado (Modelo 3) bajo H_0 . $T = 1000$, $a = 1$, sin observaciones atípicas ($\pi = 0$) y con shocks independientes en las variables y_t y z_t , $\pi = 0.1$ $s = 16$.

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Co-ruptura simultánea									
100	-2.843	-2.712	-2.680	-3.073	-2.766	-2.682	-2.889	-2.812	-2.752
200	-2.654	-2.605	-2.563	-2.665	-2.717	-2.721	-2.773	-2.793	-2.722
500	-2.652	-2.682	-2.588	-2.670	-2.792	-2.832	-2.883	-2.927	-2.928
1000	-2.589	-2.566	-2.556	-2.752	-2.882	-2.887	-2.961	-3.044	-3.050
Co-ruptura en diferencias, no en niveles									
100	-2.471	-2.450	-2.416	-2.504	-2.474	-2.465	-2.487	-2.496	-2.480
200	-2.114	-2.124	-2.114	-2.212	-2.221	-2.221	-2.246	-2.234	-2.204
500	-1.962	-1.937	-1.929	-2.015	-1.991	-1.905	-1.985	-1.949	-1.896
1000	-1.854	-1.854	-1.853	-1.880	-1.764	-1.626	-1.755	-1.654	-1.624
Co-ruptura en niveles, no en diferencias									
100	-2.890	-2.797	-2.784	-2.990	-2.796	-2.728	-2.889	-2.812	-2.752
200	-2.712	-2.737	-2.771	-2.722	-2.778	-2.799	-2.773	-2.793	-2.722
500	-2.686	-2.944	-2.932	-2.740	-2.894	-2.909	-2.883	-2.927	-2.928
1000	-2.603	-2.999	-3.028	-2.837	-3.044	-3.111	-2.961	-3.044	-3.050
Shocks independientes									
100	-2.815	-3.079	-4.165	-3.144	-3.201	-4.391	-3.242	-3.471	-4.578
200	-2.637	-2.720	-4.015	-2.868	-2.934	-4.553	-2.749	-3.033	-4.901
500	-2.689	-2.706	-2.500	-2.652	-2.740	-3.055	-2.813	-2.941	-3.578
1000	-2.587	-2.641	-2.720	-2.720	-2.863	-3.496	-2.882	-2.995	-4.089
Shocks en y_t									
100	-2.808	-2.990	-3.187	-3.175	-3.201	-3.410	-3.007	-3.362	-3.580
200	-2.629	-2.536	-2.938	-2.783	-2.770	-3.329	-2.770	-2.959	-3.539
500	-2.670	-2.615	-2.746	-2.700	-2.739	-3.105	-2.855	-2.867	-3.358
1000	-2.579	-2.598	-2.755	-2.728	-2.862	-3.286	-2.940	-2.879	-3.508
Shocks en z_t									
100	-2.843	-2.712	-2.680	-3.100	-2.762	-2.661	-3.148	-2.971	-2.757
200	-2.654	-2.605	-2.563	-2.783	-2.608	-2.532	-2.767	-2.877	-2.756
500	-2.652	-2.682	-2.588	-2.663	-2.646	-2.521	-2.867	-2.824	-2.713
1000	-2.589	-2.566	-2.556	-2.718	-2.687	-2.546	-2.894	-2.845	-2.802

TABLA 4.9. Valores críticos (5%) del contraste MCE filtrado. Modelo 3, BK, $\pi = 0.1$

pero el que mejor funciona en términos de robustez de los valores críticos y potencia es el filtro MD.

6. El Enfoque Bootstrap

Una de las áreas más productivas en los campos de la Econometría teórica y aplicada es el de los métodos intensivos en computación, tales como el bootstrap³. A pesar de que se ha aplicado tradicionalmente en situaciones *iid*, también puede ser útil en el caso de series temporales (ver por ejemplo Li y Maddala, 1996, Kiviet, 1984, Buhlmann, 1999, Berkowitz y Kilian, 2000).

Hay una serie de consideraciones previas acerca de la utilización de la metodología bootstrap para realizar inferencia. El primer punto es resaltar la importancia de los estadísticos (asintóticamente) pivote⁴ en el bootstrap. Se puede demostrar que el bootstrap proporciona aproximaciones de mayor orden de los valores críticos de estadísticos pivote ‘suaves’ (Hall, 1988). En este grupo se incluyen los estadísticos de

³Davidson y Hinkley (1997), Hall (1992), y Efron y Tibshirani (1993) son las referencias básicas sobre el tema.

⁴La propiedad fundamental de los estadísticos (asintóticamente) pivote es que sus distribuciones (asintóticas) no dependen de parámetros poblacionales desconocidos.

T	a = 0.0			a = 0.5			a = 1.0		
	s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
Co-ruptura simultánea									
100	-2.775	-2.813	-2.813	-2.810	-2.873	-2.850	-2.816	-2.729	-2.733
200	-2.609	-2.622	-2.626	-2.589	-2.622	-2.615	-2.718	-2.656	-2.638
500	-2.701	-2.658	-2.667	-2.580	-2.591	-2.594	-2.710	-2.666	-2.671
1000	-2.664	-2.663	-2.663	-2.655	-2.590	-2.598	-2.770	-2.767	-2.763
Co-ruptura en diferencias, no en niveles									
100	-2.503	-2.520	-2.527	-2.543	-2.502	-2.353	-2.481	-2.284	-2.249
200	-2.171	-2.174	-2.174	-2.187	-2.231	-2.159	-2.196	-2.111	-2.111
500	-1.987	-1.980	-1.975	-1.973	-2.043	-2.183	-2.018	-2.136	-2.198
1000	-1.848	-1.858	-1.858	-1.865	-1.900	-1.887	-1.929	-1.934	-1.902
Co-ruptura en niveles, no en diferencias									
100	-2.831	-2.727	-2.727	-2.835	-2.853	-2.850	-2.816	-2.729	-2.733
200	-2.611	-2.576	-2.576	-2.627	-2.615	-2.615	-2.718	-2.656	-2.638
500	-2.655	-2.599	-2.623	-2.574	-2.617	-2.560	-2.710	-2.666	-2.671
1000	-2.621	-2.538	-2.538	-2.651	-2.600	-2.598	-2.770	-2.767	-2.763
Shocks independientes									
100	-2.789	-2.769	-2.769	-2.845	-2.942	-2.942	-2.876	-3.078	-3.070
200	-2.611	-2.648	-2.648	-2.632	-2.737	-2.737	-2.739	-2.856	-2.831
500	-2.683	-2.641	-2.641	-2.584	-2.645	-2.666	-2.729	-2.802	-2.770
1000	-2.657	-2.634	-2.634	-2.668	-2.713	-2.713	-2.829	-2.855	-2.854
Shocks en y_t									
100	-2.801	-2.760	-2.760	-2.818	-2.875	-2.873	-2.866	-2.958	-2.971
200	-2.637	-2.652	-2.652	-2.633	-2.672	-2.678	-2.705	-2.767	-2.785
500	-2.679	-2.661	-2.661	-2.600	-2.628	-2.631	-2.728	-2.744	-2.752
1000	-2.644	-2.632	-2.632	-2.664	-2.696	-2.696	-2.808	-2.807	-2.793
Shocks en z_t									
100	-2.775	-2.813	-2.813	-2.828	-2.849	-2.854	-2.890	-2.956	-2.956
200	-2.609	-2.622	-2.626	-2.581	-2.684	-2.684	-2.741	-2.831	-2.831
500	-2.701	-2.658	-2.667	-2.558	-2.593	-2.595	-2.712	-2.762	-2.742
1000	-2.664	-2.663	-2.663	-2.668	-2.711	-2.710	-2.815	-2.830	-2.830

TABLA 4.10. Valores críticos (5%) del contraste MCE filtrado. Modelo 3, MD, $\pi = 0.1C$

contraste cuyas distribuciones asintóticas son Normales tipificadas o chi-cuadrado. La capacidad del bootstrap de proporcionar refinamientos asintóticos para dichos estadísticos constituye una razón de peso para su utilización en aplicaciones empíricas.

El bootstrap también puede aplicarse a estadísticos que no son asintóticamente pivote, como pueden ser los coeficientes de una regresión, pero en tal caso no proporciona una aproximación de mayor orden de su distribución. Los estimadores bootstrap de las distribuciones de los estadísticos no pivotaes tienen la misma precisión que las aproximaciones asintóticas de primer orden.

Se pueden obtener aproximaciones de mayor orden de las distribuciones de los estadísticos no pivotaes utilizando las técnicas de *prepivotting* o *bootstrap iterado* (Beran, 1987, 1988) o bien método de corrección del sesgo (Efron, 1987). El bootstrap iterado es computacionalmente muy costoso, lo que hace que tenga muy poco interés en los casos en que tenemos un estadístico pivote.

Así pues, una vez vista la importancia de los estadísticos (asintóticamente) pivote para la construcción de intervalos de confianza y los órdenes de aproximación de los diferentes métodos, se ve clara la necesidad de utilizar estadísticos (asintóticamente)

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Co-ruptura simultánea									
100	25.15	21.30	13.05	12.00	12.80	8.15	9.20	8.95	7.65
200	73.40	58.10	61.55	50.90	51.50	51.50	39.30	38.65	39.30
500	99.60	99.65	99.90	99.05	99.10	99.40	98.25	98.45	98.70
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Co-ruptura en diferencias, no en niveles									
100	2.20	3.90	6.00	1.40	2.15	3.25	1.70	2.10	1.55
200	1.25	2.35	4.50	0.70	1.15	1.95	0.55	0.70	0.50
500	0.05	0.25	1.55	0.05	0.00	0.25	0.05	0.05	0.05
1000	0.00	0.25	2.85	0.00	0.00	0.90	0.00	0.00	0.00
Co-ruptura en niveles, no en diferencias									
100	21.05	11.90	11.55	13.20	12.25	11.05	9.20	8.95	7.65
200	68.90	46.45	40.45	51.70	43.80	41.50	39.30	38.65	39.30
500	99.50	97.00	96.60	98.80	98.05	97.95	98.25	98.45	98.70
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks independientes									
100	23.10	16.70	9.65	12.05	14.80	5.35	8.90	9.05	2.60
200	73.05	64.45	29.40	45.10	56.10	19.75	40.45	38.15	10.15
500	99.65	98.75	94.45	99.20	98.30	71.40	98.65	86.60	39.95
1000	100.00	100.00	100.00	100.00	100.00	99.95	100.00	100.00	98.85
Shocks en y_t									
100	22.70	17.05	3.20	10.75	12.75	1.95	8.65	8.40	1.15
200	73.05	64.30	28.65	43.50	57.30	16.45	42.95	47.50	11.50
500	99.65	98.80	94.45	99.40	97.50	67.20	98.85	85.65	37.85
1000	100.00	100.00	100.00	100.00	100.00	99.90	100.00	100.00	98.75
Shocks en z_t									
100	25.50	21.80	9.30	12.80	15.55	6.25	9.75	7.10	3.65
200	73.90	70.75	40.15	48.85	50.35	20.60	41.90	28.65	9.65
500	99.75	99.25	80.55	99.25	91.90	50.05	98.50	71.40	21.45
1000	100.00	100.00	99.95	100.00	100.00	94.70	100.00	99.85	63.60

TABLA 4.11. Potencia (%) ajustada del contraste MCE filtrado. Modelo 3, HP100, $\pi = 0.1$, $b = -0.2$.

pivotales. Caso contrario, no debemos esperar ninguna mejora respecto a los resultados asintóticos.

Un problema que se plantea cuando aplicamos el bootstrap a series temporales es la dependencia de nuestras observaciones. Una de las soluciones propuestas en la Literatura es dividir los datos (o residuos) en bloques y aplicar las técnicas de remuestreo con reemplazamiento a esos bloques en lugar de utilizar los datos originales. Esta es la base del denominado *moving block bootstrap* (MBB) (Kunsch, 1989, Davidson y Hall, 1993). Podemos hacer que los bloques no se solapen (Carlstein, 1986) o bien permitir su solapamiento (Politis y Romano, 1994). Una opción es tomar la longitud de los bloques como fija o bien, tal y como proponen Politis y Romano (1994), suponer que las longitudes de los bloques son variables aleatorias que siguen una distribución geométrica. Esta técnica de bootstrap es la que la longitud de los bloques es aleatoria es lo que se conoce como *stationary bootstrap*, pues garantiza que los datos remuestreados son estacionarios supuesto que lo sean los datos originales, lo que no está garantizado en el caso de utilizar el MBB con longitudes de los bloques fijas (no aleatorias). Lahiri (1999) analiza detalladamente las propiedades de las diferentes

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Co-ruptura simultánea									
100	52.25	48.15	47.45	28.10	29.60	34.60	22.20	26.25	27.80
200	97.05	94.90	96.15	78.80	84.30	86.35	76.65	78.95	78.95
500	100.00	100.00	100.00	100.00	100.00	100.00	99.95	99.95	99.95
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Co-ruptura en diferencias, no en niveles									
100	0.95	1.70	3.75	1.05	0.85	1.25	0.85	0.50	0.30
200	0.80	1.70	4.10	0.40	0.50	1.35	0.15	0.05	0.05
500	0.15	0.80	2.90	0.00	0.00	0.55	0.00	0.00	0.00
1000	0.05	0.55	3.50	0.00	0.00	0.60	0.00	0.00	0.00
Co-ruptura en niveles, no en diferencias									
100	50.70	32.65	32.35	27.55	29.50	32.40	22.20	26.25	27.80
200	97.15	82.15	78.75	82.30	80.10	78.90	76.65	78.95	78.95
500	100.00	100.00	99.95	100.00	99.95	99.95	99.95	99.95	99.95
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks independientes									
100	51.50	51.80	29.00	30.80	26.45	9.70	20.00	12.10	4.50
200	97.20	97.40	87.25	78.75	82.60	43.90	75.30	58.45	20.25
500	100.00	100.00	100.00	100.00	100.00	99.95	99.95	99.85	93.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks en y_t									
100	52.30	63.95	60.80	30.95	42.45	41.25	22.05	29.05	26.85
200	97.20	99.25	99.70	76.80	93.25	96.90	78.55	81.85	90.75
500	100.00	100.00	100.00	100.00	100.00	100.00	99.95	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks en z_t									
100	51.90	40.75	15.60	29.55	24.60	8.70	23.65	10.80	5.05
200	97.10	95.10	62.60	77.25	73.65	28.35	75.95	45.30	13.15
500	100.00	100.00	98.60	100.00	99.90	80.20	100.00	95.10	44.65
1000	100.00	100.00	100.00	100.00	100.00	99.50	100.00	100.00	85.45

TABLA 4.12. Potencia (%) ajustada del contraste MCE filtrado. Modelo 3, HP10, $\pi = 0.1$, $b = -0.2$.

variedades de técnicas MBB. Una alternativa muy reciente es el *threshold bootstrap*, propuesto por Park y Willemain (1999).

Cualquiera que sea el tipo de bloque utilizado, se sabe que su longitud (o su longitud media) debe incrementarse con el tamaño muestral. A pesar de que se han propuesto en la Literatura distintas longitudes ‘óptimas’ de los bloques (Carlstein et al., 1998, Buhlmann y Kunsch, 1999), no está nada claro cuál es la longitud (o longitud media) óptima y, por otra parte, ésta depende de cuál sea la finalidad que se dé al bootstrap.

Estas consideraciones son las que hace que se utilice en su lugar la alternativa del *sieve bootstrap* (Buhlmann, 1997, 1998). La base es la búsqueda de un modelo que nos proporcione residuos independientes de forma que podamos utilizar la metodología bootstrap habitual. Dada la naturaleza de nuestro problema, se consideran los modelos de la familia AR, que son la base de los contrastes aplicados en las secciones anteriores. Un aspecto fundamental es la necesidad de generar los residuos del modelo restringido (Li y Maddala, 1996, Datta, 1996). Así pues, el procedimiento que seguimos es el siguiente:

- (1) Estimar el modelo MCE

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Co-ruptura simultánea									
100	76.65	83.15	94.90	44.55	51.55	73.70	24.30	24.15	26.40
200	99.80	99.50	99.70	94.60	96.00	98.20	76.15	74.70	78.85
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Co-ruptura en diferencias, no en niveles									
100	1.80	3.00	5.40	0.80	0.70	1.05	0.15	0.10	0.05
200	1.35	3.50	6.15	0.20	0.15	0.80	0.00	0.00	0.00
500	0.40	1.75	5.20	0.05	0.05	1.10	0.00	0.00	0.00
1000	0.05	0.90	4.20	0.00	0.05	0.65	0.00	0.00	0.00
Co-ruptura en niveles, no en diferencias									
100	73.35	41.15	27.60	39.00	30.25	28.55	24.30	24.15	26.40
200	99.60	92.70	82.70	92.25	84.35	77.00	76.15	74.70	78.85
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks independientes									
100	77.70	77.75	27.30	43.50	47.90	9.65	28.25	19.00	3.65
200	99.80	99.45	91.85	93.70	93.75	75.85	84.65	71.65	47.75
500	100.00	100.00	100.00	100.00	100.00	99.00	100.00	99.45	80.40
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks en y_t									
100	78.35	83.40	98.15	45.45	62.95	97.10	26.45	66.00	96.25
200	99.70	99.85	99.95	93.95	98.85	99.85	82.75	97.95	99.95
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks en z_t									
100	76.90	70.05	27.15	45.95	35.70	12.65	25.95	14.40	5.75
200	99.85	99.25	79.45	93.20	85.00	40.50	80.60	50.95	13.90
500	100.00	100.00	99.95	100.00	100.00	89.05	100.00	96.90	52.45
1000	100.00	100.00	100.00	100.00	100.00	99.80	100.00	100.00	87.60

TABLA 4.13. Potencia (%) ajustada del contraste MCE filtrado. Modelo 3, BK, $\pi = 0.1$, $b = -0.2$.

$$\Delta y_t^g = c + a\Delta z_t^g + \sum_{i=1}^p \phi_i \Delta y_{t-i}^g + \sum_{i=1}^q \theta_i z_{t-i}^g + b(y_{t-1}^g - z_{t-1}^g) + u_t.$$

and obtain the t -stat $t(\hat{b})$

(2) Estimar el modelo *restringido*

$$\Delta y_t^g = c^* + a\Delta z_t^g + \sum_{i=1}^{p^*} \phi_i^* \Delta y_{t-i}^g + \sum_{i=1}^{q^*} \theta_i^* \Delta z_{t-i}^g + \nu_t.$$

(3) Remuestrear residuos ν_t^* y construir $y_t^{g^*}$ con el modelo

$$\Delta y_t^{g^*} = c^* + \hat{a}\Delta z_t^g + \sum_{i=1}^{p^*} \hat{\phi}_i^* \Delta y_{t-i}^{g^*} + \sum_{i=1}^{q^*} \hat{\theta}_i^* \Delta z_{t-i}^g + \nu_t^*.$$

(4) Estimar el modelo

$$\Delta y_t^{g^*} = c + a\Delta z_t^g + \sum_{i=1}^p \phi_i \Delta y_{t-i}^g + \sum_{i=1}^q \theta_i z_{t-i}^g + b(y_{t-1}^{g^*} - z_{t-1}^g) + \xi_t.$$

y obtener el estadístico t , $t^*(\hat{b}^*)$

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Co-ruptura simultánea									
100	88.25	85.35	57.10	58.75	53.60	33.70	33.95	35.40	35.30
200	99.70	99.75	98.10	97.80	97.65	94.65	85.80	89.10	89.75
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Co-ruptura en diferencias, no en niveles									
100	3.20	3.35	3.25	1.25	1.55	1.10	0.50	0.35	0.25
200	2.05	2.20	2.15	0.80	0.75	0.55	0.05	0.05	0.05
500	0.30	0.35	0.35	0.05	0.20	0.15	0.00	0.00	0.00
1000	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Co-ruptura en niveles, no en diferencias									
100	86.90	86.60	86.50	56.45	54.25	54.25	33.95	35.40	35.30
200	100.00	99.95	99.95	97.10	97.45	97.45	85.80	89.10	89.75
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks independientes									
100	86.70	84.90	84.75	58.85	57.25	57.35	31.90	29.75	30.25
200	99.70	99.60	99.65	97.05	94.90	94.95	85.05	80.20	81.60
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks en y_t									
100	87.05	87.25	87.25	59.60	57.60	57.80	30.70	32.00	32.15
200	99.65	99.60	99.60	96.95	96.45	96.35	85.90	86.20	86.10
500	100.00	100.00	100.00	100.00	100.00	100.00	99.95	99.95	99.95
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Shocks en z_t									
100	88.30	86.30	86.25	57.10	59.05	58.65	31.15	30.75	30.90
200	99.65	99.75	99.75	97.75	96.50	96.50	85.25	80.55	80.50
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.95	99.95
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

TABLA 4.14. Potencia (%) ajustada del contraste MCE filtrado. Modelo 3, MD, $\pi = 0.1$, $b = -0.2$.

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Modelo 2									
100	5.35	6.30	13.95	7.55	17.55	16.45	10.05	38.90	19.25
200	6.05	7.75	30.85	7.00	19.40	37.45	10.95	43.10	50.15
500	5.05	3.25	1.25	6.65	16.35	10.75	7.35	34.80	33.30
1000	5.35	3.40	3.60	6.45	16.80	20.45	5.50	21.75	43.55
Modelo 3, HP10									
100	4.75	3.45	2.50	4.95	2.00	2.00	4.15	1.65	2.05
200	5.60	4.35	4.90	3.75	1.95	4.05	4.75	2.15	4.55
500	5.00	4.90	7.20	4.70	2.95	9.35	6.10	5.00	15.40
1000	5.10	6.00	11.00	4.75	3.35	13.80	5.50	5.40	21.85
Modelo 3, HP100									
100	5.85	3.70	1.95	5.05	1.70	0.65	4.50	1.55	0.40
200	5.60	5.15	4.35	5.00	3.90	2.20	7.10	6.75	3.70
500	5.30	5.35	5.20	6.40	3.50	3.40	6.65	5.10	3.85
1000	5.35	4.95	5.35	6.05	4.15	4.45	6.15	5.00	5.15
Modelo 3, BK filter									
100	3.95	7.20	34.75	5.60	7.00	41.10	6.65	10.00	48.55
200	4.65	5.80	22.15	5.90	6.85	34.15	4.65	9.35	45.70
500	5.05	5.25	3.75	4.80	5.80	11.35	3.85	5.55	18.65
1000	4.85	5.85	7.50	4.65	5.85	16.70	4.00	5.50	26.10
Modelo 3, MD filter									
100	5.55	5.25	5.25	5.35	6.50	6.50	5.10	7.60	7.60
200	4.90	5.20	5.20	5.20	6.30	6.25	5.25	6.40	6.40
500	5.15	4.90	4.90	5.15	6.25	6.30	4.90	5.80	5.50
1000	4.85	4.35	4.35	5.25	5.90	5.90	5.40	6.15	6.25

TABLA 4.15. Robustez frente a valores atípicos. Porcentaje de rechazos cuando el modelo se simula con el PGD (5.1a)–(5.1f) con shocks independientes en las variables y_t y z_t con $\pi = 0.1$ bajo $b = 0$, usando los valores críticos de la Tabla 4.1.

Repetir pasos 3–4 NB veces, donde NB indica el número de remuestreos bootstrap. El valor crítico bootstrap (VC_b) se obtiene a partir del cuantil 5% de la distribución empírica de t^* . Si $t(\hat{b}) < VC_b$ se rechaza H_0 ($b = 0$).

Nuestro experimento de simulación de Monte Carlo se basa en el mismo PGD que en la Sección anterior, con 2000 réplicas y $NB = 100$ remuestreos bootstrap, que se puede considerar ‘seguro’ según Davidson y MacKinnon (1998, 2000).

a	T=100	T=200	T=500	T=1000
Modelo 2				
0.0	6.25	5.70	5.15	5.65
0.5	6.05	5.05	6.40	5.40
1.0	6.05	5.85	5.70	6.15
Modelo 3. HP10				
0.0	7.50	6.25	4.75	4.25
0.5	9.95	9.40	7.65	6.55
1.0	11.15	9.95	6.40	5.50
Modelo 3. HP100				
0.0	9.90	6.20	5.80	5.40
0.5	12.75	7.50	6.00	6.15
1.0	14.70	6.90	6.30	5.05
Modelo 3. BK				
0.0	8.50	5.50	5.40	4.60
0.5	9.15	6.35	5.70	6.00
1.0	7.80	6.30	6.90	6.70
Modelo 3. MD				
0.0	7.58	5.88	5.92	5.02
0.5	6.70	5.36	3.98	4.17
1.0	5.88	4.92	4.37	4.58

TABLA 4.16. Tamaño (%) del contraste bootstrap MCE (Modelo 2) y MCE filtrado (Modelo 3) sin observaciones atípicas ($\pi = 0$).

a	T=100	T=200	T=500	T=1000
Modelo 2				
0.0	99.45	100.00	100.00	100.00
0.5	94.95	100.00	100.00	100.00
1.0	87.75	100.00	100.00	100.00
Modelo 3. HP10				
0.0	52.30	95.70	100.00	100.00
0.5	41.45	86.15	100.00	100.00
1.0	42.20	85.00	100.00	100.00
Modelo 3. HP100				
0.0	30.05	70.60	99.45	100.00
0.5	23.95	55.80	98.95	100.00
1.0	27.75	45.80	98.70	100.00
Modelo 3. BK				
0.0	75.25	99.50	100.00	100.00
0.5	51.90	91.40	100.00	100.00
1.0	24.05	71.60	100.00	100.00
Modelo 3. MD				
0.0	86.70	99.45	100.00	100.00
0.5	59.80	95.60	100.00	100.00
1.0	29.70	81.35	100.00	100.00

TABLA 4.17. Potencia ajustada (%) del contraste bootstrap MCE (Modelo 2) y MCE filtrado (Modelo 3) sin observaciones atípicas ($\pi = 0$). $b = -0.2$

Tal y como se observa en la Tabla 4.16, cuando no hay observaciones atípicas necesitamos muestras con al menos 200 observaciones para conseguir un tamaño próximo

a su valor nominal (5%). Esto sucede tanto si aplicamos el contraste a las series observadas a los componentes tendenciales con el filtro MD. Al aplicar los filtros HO y BK se necesitan muestras aún más grandes. Recordemos que en el Capítulo 3 habíamos obtenido tamaños del contraste bootstrap adecuados para muestras con tan sólo 25 observaciones. La única diferencia es que ahora incluimos términos dinámicos en la regresión del contraste y los resultados se ven sesgados por el procedimiento de selección del modelo.

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Modelo 2									
100	7.25	19.40	17.35	24.65	40.20	27.30	43.90	60.75	33.95
200	8.90	6.20	21.15	29.70	29.25	36.00	49.85	55.45	52.95
500	6.45	2.80	1.65	23.15	18.75	11.05	26.95	42.85	37.75
1000	6.25	5.70	3.20	20.70	29.60	20.70	17.30	52.95	50.25
Modelo 3. HP10 Filter									
100	5.60	9.15	8.25	5.25	9.85	10.40	7.40	13.70	13.10
200	5.00	6.25	5.40	5.45	4.70	8.55	7.60	6.35	13.10
500	5.05	7.25	9.70	6.70	9.15	17.55	4.80	13.85	26.05
1000	5.50	4.85	9.70	8.20	7.60	18.45	6.90	12.65	28.40
Modelo 3. HP100 Filter									
100	6.65	10.15	4.25	8.60	10.80	4.70	10.85	13.90	6.15
200	5.70	6.55	4.05	5.25	7.55	4.70	7.05	10.85	7.75
500	5.45	4.05	5.55	5.70	4.55	5.20	6.65	6.15	5.40
1000	5.20	4.40	5.40	3.90	5.50	5.45	6.30	6.35	7.45
Modelo 3. BK Filter									
100	8.65	13.70	46.60	7.45	17.35	53.75	8.80	23.00	57.45
200	5.10	7.10	23.90	5.85	10.15	38.75	7.40	12.25	46.55
500	4.90	6.35	4.55	4.85	12.00	13.05	7.90	19.85	24.40
1000	4.40	4.25	7.90	6.65	10.90	18.90	8.05	19.30	34.25
Modelo 3. MD Filter									
100	7.55	8.40	7.70	6.15	7.35	8.25	6.35	6.25	8.55
200	5.65	6.20	7.30	4.75	4.85	5.45	5.10	4.90	6.65
500	4.90	5.65	5.60	4.55	5.05	5.50	4.85	5.30	5.50
1000	5.65	5.55	5.30	4.80	4.75	6.15	5.35	4.80	6.15

TABLA 4.18. Tamaño (%) del contraste bootstrap MCE (Modelo 2) y MCE filtrado (Modelo 3) con shocks independientes en las variables, $\pi = 0.1$.

T	$a = 0.0$			$a = 0.5$			$a = 1.0$		
	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$	$s=1$	$s=6$	$s=16$
Modelo 3. MD Filter									
100	87.85	86.70	86.95	58.55	64.15	62.00	32.35	40.05	37.75
200	99.45	99.60	99.30	95.05	94.60	94.80	80.05	80.05	81.45
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.90	99.70
1000	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

TABLA 4.19. Potencia (%) ajustada del contraste bootstrap MCE (Modelo 2) y MCE filtrado (Modelo 3) con shocks independientes en las variables, $\pi = 0.1$, $b = -0.2$.

La Tabla 4.17 muestra que la potencia del contraste bootstrap es alta en el caso de que no haya observaciones atípicas. Sin embargo, debemos recordar que el tamaño del contraste bootstrap es adecuado sólo si la muestra tiene más de 100 observaciones.

La cuestión fundamental es si nuestro enfoque de filtrar las series antes de aplicar el contraste, combinado con la aplicación de técnicas bootstrap hace que el contraste

MCE sea robusto a la presencia de observaciones atípicas en las series. Nos vamos a centrar en el caso potencialmente más dañino, aquél en que hay shocks independientes en ambas variables. La respuesta a nuestra pregunta está en la Tabla 4.18 y es negativa excepto en el caso del filtro MD. Al aplicar el contraste bootstrap a las series tras aplicar el filtro MD, el tamaño del contraste es adecuado para muestras de 100 observaciones, a no ser que los saltos sean muy grandes, y con muestras de 200 observaciones, un caso muy habitual en la práctica, el contraste tiene siempre un tamaño cercano al valor nominal. Al comparar la Tablas 4.14 y 4.19 se observa que la potencia del contraste bootstrap es muy similar a la obtenida con los valores críticos tabulados en la Tabla 4.10.

7. Ejemplo Empírico

En esta Sección se ilustra la utilidad práctica de nuestro enfoque mediante el análisis de las observaciones anuales del tipo de cambio real, basado en el IPC, entre USA y Finlandia, en el período 1980–88. Este conjunto de datos ha sido analizado con anterioridad por Perron y Vogelsang (1992), Franses y Haldrup (1994) y Franses y Lucas (1998), y sirvió de base para el ejemplo empírico del Capítulo 1. La pregunta que nos hacemos es si la serie de tipo de cambio real es estacionaria o no, es decir, si se cumple la paridad del poder adquisitivo (PPP). Para ello debemos realizar el contraste para saber si el tipo de cambio nominal y el ratio de precios están cointegrados.

En primer lugar se realiza el análisis imponiendo la restricción de factor común (COMFAC) con las siguientes variables: y_t es el logaritmo neperiano del tipo de cambio nominal y z_t es el logaritmo neperiano de ratio de IPC. La relación de cointegración es conocida e igual al tipo de cambio real, $r_t = y_t - z_t$. El número de retardos en el modelo MCE viene determinado por el criterio SBIC. Según el teorema de representación de Granger (Engle y Granger, 1987), si las variables están cointegradas, el término r_{t-1} debe ser significativo en al menos una de las dos ecuaciones del modelo MCE. Realizamos el contraste MCE examinando los estadísticos t del coeficiente correspondiente a r_{t-1} en ambas ecuaciones.

Al realizar el análisis con las variables observadas, los valores del estadístico t son -6.65 y -1.49 , cuando las variables dependientes son Δy_t y Δz_t , respectivamente. Es obvio que la causalidad a largo plazo va de z_t , el log del ratio de IPC, a y_t , el log del tipo de cambio nominal, pero no está claro si las series están cointegradas debido a los atípicos aditivos que hay en las series. A partir de las Tablas 4.1 y 4.5, los valores críticos indicarían que las variables están cointegradas sólo si los shocks

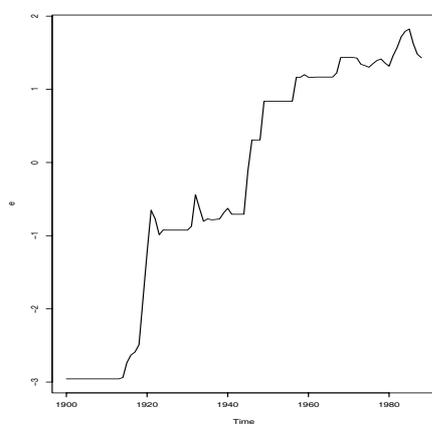


Figura 4.4.1: Tipo de cambio nominal

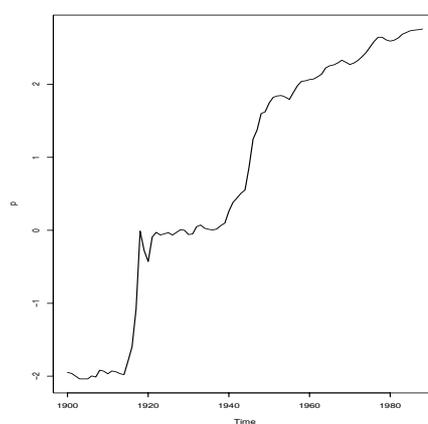


Figura 4.4.2: Ratio de precios

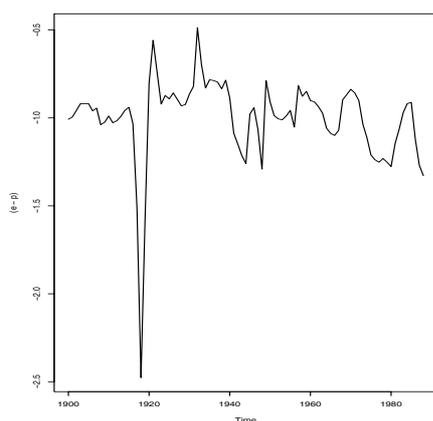
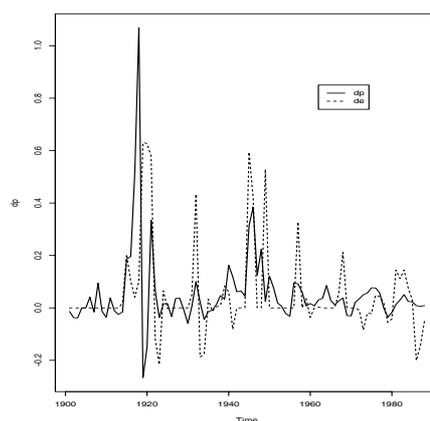
Figura 4.4.3: $e_t - p_t$ 

Figura 4.4.4: Primeras diferencias de las series observadas.

FIGURA 4.4. Tipo de cambio nominal (e) y ratio de precios (p) USA/Finlandia, entre 1900 y 1988, series observadas.

ocurren únicamente en z_t , el log del ratio de IPC. Sin embargo, no se puede asegurar que las series estén cointegradas si los shocks en ambas variables son independientes u ocurren sólo en la variable y_t , el log del tipo de cambio nominal, y éstos casos no son en absoluto descartables. Al aplicar los contrastes a los componentes tendenciales, contraste MCE filtrado, con el filtro HP100, los valores de los estadísticos t son -1.03 y -2.15 respectivamente, en tanto que con el filtro BK los estadísticos del contraste son -1.84 y -0.47 , y, por último, -2.16 y -0.37 con el filtro MD.

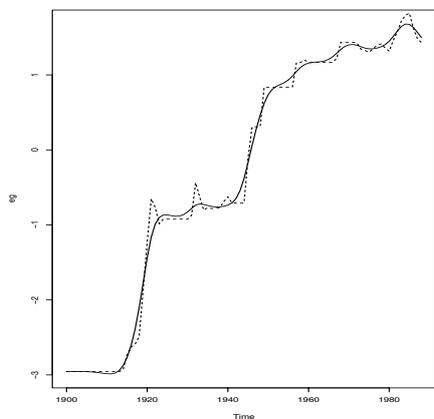
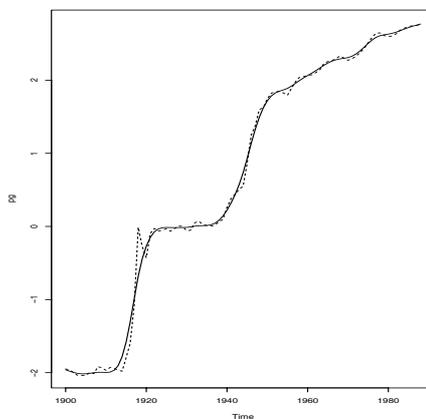
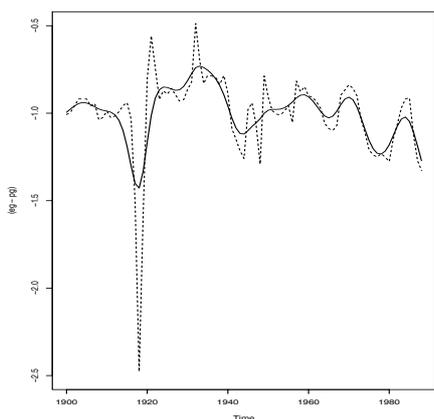
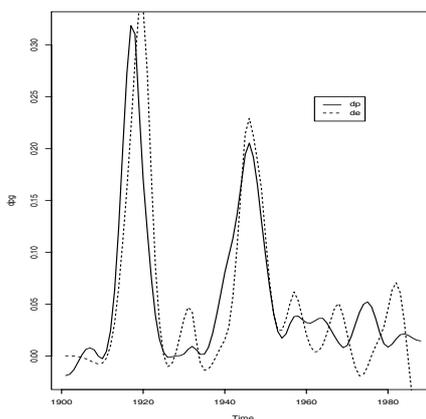
Figura 4.5.1: e_t^g Figura 4.5.2: p_t^g Figura 4.5.3: $e_t^g - p_t^g$ Figura 4.5.4: Primeras diferencias de e_t^g y p_t^g .

FIGURA 4.5. Tipo de cambio nominal (e) y ratio de precios (p) USA/Finlandia, entre 1900 y 1988, series filtradas con HP10.

Como se deduce de los valores críticos de la Tablas 4.7, 4.9, y 4.10, la conclusión es que no se puede rechazar la hipótesis de que las series no están cointegradas, lo que concuerda con los resultados de Franses y Haldrup (1994) y Franses y Lucas (1998).

Sin embargo, hemos de recordar el hecho de que al imponer la restricción COM-FAC la potencia del contraste se reduce, ver Kremers et al. (1992). Así pues, se procede a aplicar los contrastes sin imponer la restricción de factor común, tal y como se vio en la Sección 5. En el caso de filtro BK, los estadísticos del contraste MCE

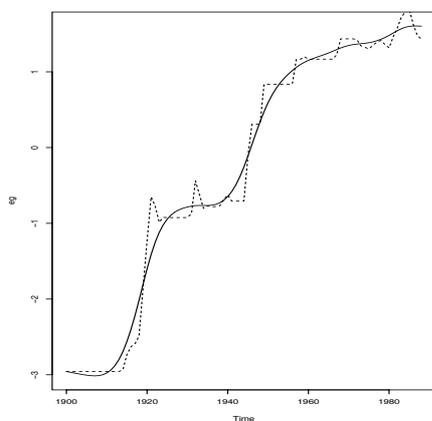
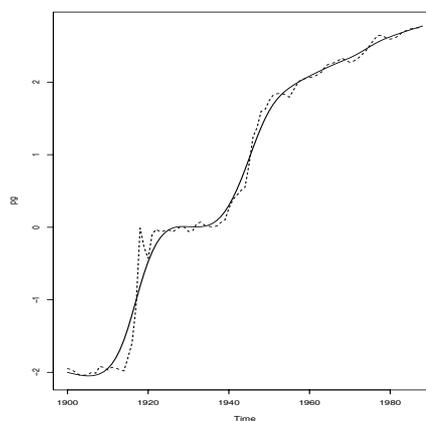
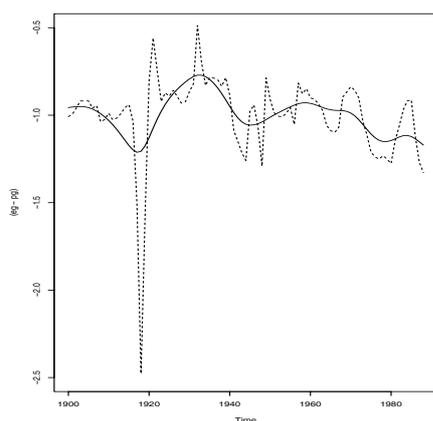
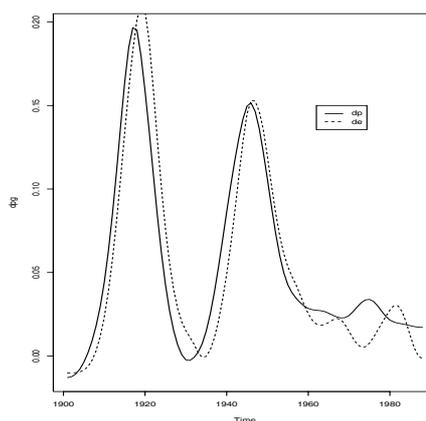
Figura 4.6.1: e_t^g Figura 4.6.2: p_t^g Figura 4.6.3: $e_t^g - p_t^g$ Figura 4.6.4: Primeras diferencias de e_t^g y p_t^g .

FIGURA 4.6. Tipo de cambio nominal (e) y ratio de precios (p) USA/Finlandia, entre 1900 y 1988, series filtradas con HP100.

son -3.63 y -0.86 , y -6.71 y -1.42 en el caso del filtro MD. Dado que los valores absolutos del segundo estadístico hacen que no sean significativos, esto indica que la segunda variable es exógena en sentido débil para el parámetro de largo plazo de interés. No obstante, el primer estadístico nos lleva a la conclusión, a partir de los valores críticos de las Tablas 4.7, 4.9, y 4.10, de que las variables están cointegradas, es decir, que el tipo de cambio real sigue un proceso estacionario con observaciones

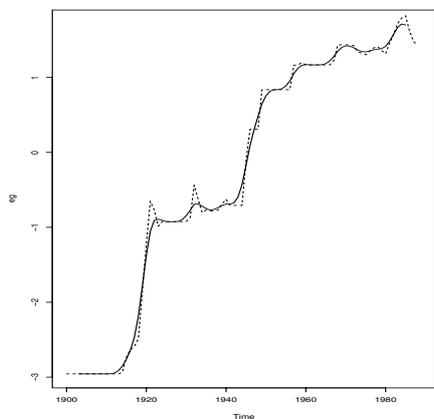
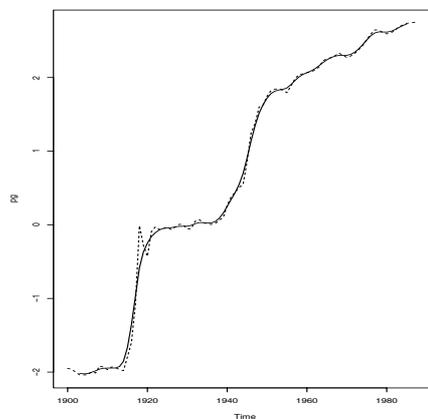
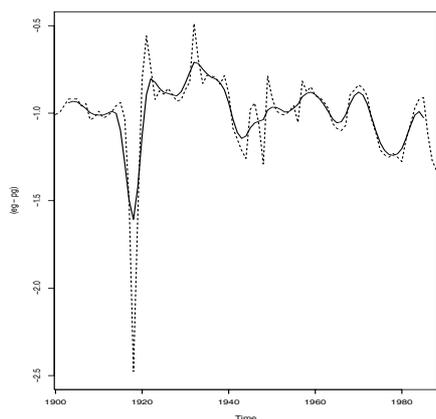
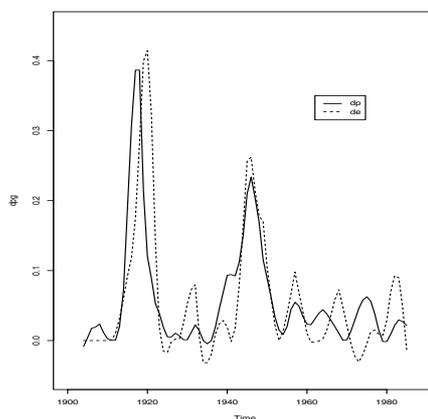
Figura 4.7.1: e_t^g Figura 4.7.2: p_t^g Figura 4.7.3: $e_t^g - p_t^g$ Figura 4.7.4: Primeras diferencias de e_t^g y p_t^g .

FIGURA 4.7. Tipo de cambio nominal (e) y ratio de precios (p) USA/Finlandia, entre 1900 y 1988, series filtradas con BK.

atípicas, por lo que se cumple la hipótesis de la PPP. Este resultado es congruente con los resultados obtenidos en Vogelsang (1999) y el Capítulo 1.

8. Conclusiones

En este Capítulo se han analizado los efectos que tienen los atípicos aditivos sobre las variables cointegradas en un contexto multivariante. Los estudios anteriores mostraban que los contrastes de no-cointegración habituales (como el contraste MCE) tienden a encontrar demasiada cointegración en tales casos. Este problema se palía

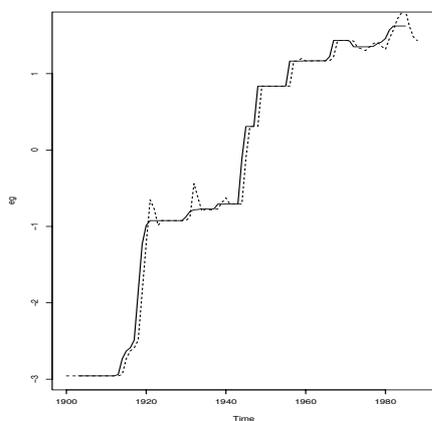
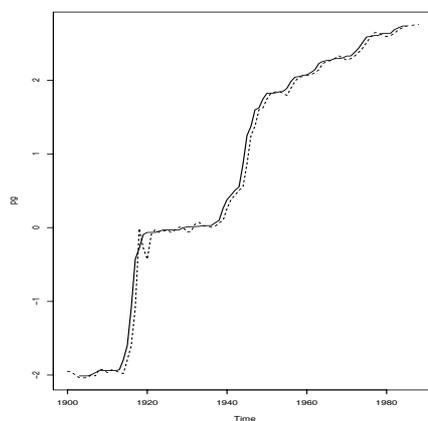
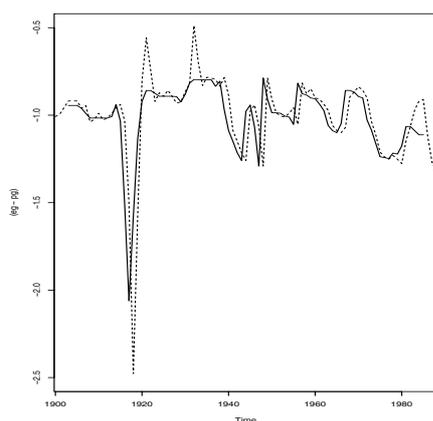
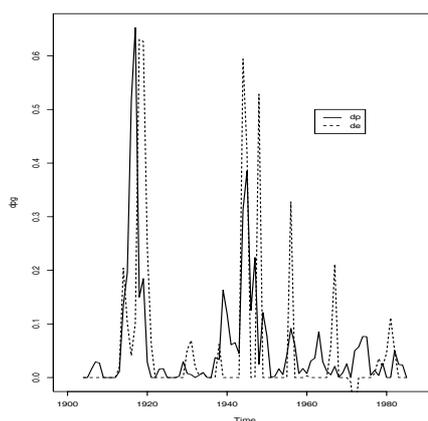
Figura 4.8.1: e_t^g Figura 4.8.2: p_t^g Figura 4.8.3: $e_t^g - p_t^g$ Figura 4.8.4: Primeras diferencias de e_t^g y p_t^g .

FIGURA 4.8. Tipo de cambio nominal (e) y ratio de precios (p) USA/Finlandia, entre 1900 y 1988, series filtradas con MD.

si utilizamos modelos sobrep parametrizados que incluyen retardos adicionales de las variables en el modelo de regresión sobre el que se realiza el contraste, pero no se soluciona totalmente.

Se han estudiado los efectos de los atípicos aditivos en un marco multivariante, incluyendo casos de co-ruptura simultánea, co-rupturas parciales y otros en los que no había ningún de tipo de co-ruptura, como es el caso de que hay shocks independientes en ambas variables o cuando los shocks ocurren sólo en una de las variables.

Evidentemente, la peor distorsión se produce cuando tenemos shocks independientes en ambas variables.

Nuestra recomendación para solucionar este problema es realizar el contraste de no-cointegración sobre el modelo MCE basado en los componente tendenciales de las series en lugar de utilizar el modelo con las series observadas. Se han analizado diferentes filtros de extracción de tendencia por medio de experimentos de Monte Carlo, tales como el filtro de Hodrick–Prescott, (HP10, HP100), Baxter y King (BK) y de la mediana (MD). A pesar de que en casi todos los casos se obtienen buenos resultados en términos de robustez de los valores críticos del contraste, los mejores resultados se obtienen con los filtros HP100 y MD. No obstante, atendiendo al criterio de maximizar potencia, los mejores filtros en este aspecto son BK y MD. Estos resultados nos hacen pensar que debemos prestar mayor atención a los filtros no lineales, tales como el de la mediana (MD). Este resultado se confirma al aplicar la metodología bootstrap al contraste. El único problema que plantea el bootstrap es que necesitamos al menos 200 observaciones para que el contraste tenga un tamaño próximo al su valor nominal.

La metodología empírica propuesta se aplica al tipo de cambio real basado en el IPC entre USA y Finlandia. Este conjunto de datos es particularmente importante, dado que en torno al mismo hay una controversia en la Literatura al obtenerse distintas conclusiones dependiendo de cuál sea el contraste empleado. Por una parte, Franses y Haldrup (1994) y Franses y Lucas (1998) no encuentran evidencia de cointegración entre el tipo de cambio nominal y el ratio de precios, en tanto que Vogelsang (1999) llega a la conclusión de que las series están cointegradas, lo que coincide con el resultado encontrado en el Capítulo 1. La conclusión a la que se llega tras utilizar la versión más potente del contraste sobre las componentes tendenciales es que las series están cointegradas, en tanto que si se impone la restricción COMFAC la conclusión es que no hay cointegración, lo que justifica que haya resultados empíricos contradictorios.

Por tanto, el procedimiento propuesto es robusto a la presencia de observaciones atípicas aditivas y no requiere la identificación previa de las mismas.

Parte 3

Estimación de relaciones de largo plazo

Un Enfoque No-paramétrico de la Cointegración de Pares de Variables $I(d)$

1. Introducción

Muchas series temporales presentan shocks permanentes que afectan tanto a la media como a la varianza de la serie. Dichas series temporales se conocen habitualmente como *integradas*, dado que sus características fundamentales pueden generarse como sumas de una número creciente de variables aleatorias débilmente dependientes. El hecho de que los shocks remotos tengan un efecto permanente sobre los niveles de dichas series es lo que se conoce como propiedad de *memoria larga* de las mismas. Las series integradas pueden expresarse como sumas de *componentes no observados*, en la que uno de dichos componentes es una *tendencia estocástica*. En algunos casos, los movimientos permanentes en los niveles de dichas series pueden presentar un alto nivel de correlación para distintas series. En el caso de que las series de memoria larga presenten un componente de *tendencia estocástica común*, decimos que las series están *cointegradas*.

El concepto de *cointegración* fue introducido originalmente por Granger (1981), y posteriormente desarrollado por Engle y Granger (1987) y Johansen (1988) en un contexto del mecanismo de corrección del error que se aplica asiduamente en Macroeconomía y Finanzas. Ejemplos de series cointegradas son los precios de bienes y acciones, renta y gasto, y tipos de cambio entre diferentes monedas. También es posible encontrar relaciones de cointegración entre variables utilizadas en Ingeniería, tales como meteorología y telecomunicaciones.

Estas tendencias estocásticas comunes que afectan a las series cointegradas nos permiten predecir su evolución relativa en el largo lazo y proporcionan importante información sobre los procesos generadores de datos. Por ejemplo, si los tipos de cambio y/o los precios de los activos financieros de diferentes países se mueven conjuntamente en el largo plazo, esto puede ser un signo de integración de los mercados financieros internacionales. Además, el hecho de que los precios de los bienes de

distintos países estén cointegrados indica que la paridad del poder adquisitivo debe cumplirse en el largo plazo.

En el campo de la Ingeniería, supongamos que estamos interesados en reconocer una serie de características con fines de identificación o clasificación, y supongamos que nuestras señales discretizadas tienen un componente de energía de baja frecuencia predominante. El problema que se plantea es que en estas condiciones es que se van a detectar correlaciones significativas incluso cuando las señales no tienen ninguna relación entre sí. Este problema es lo que se conoce en Econometría como *regresiones espurias*, ver Granger y Newbold (1974). Evidentemente, no se nos pueden escapar las consecuencias de las mismas, pues vamos a tomar decisiones a partir de las mismas que son absolutamente erróneas. Los contrastes de cointegración nos permiten solventar el problema de las regresiones espurias, y nos permiten conocer cuándo dos señales están relacionadas en el largo plazo.

Detrás del concepto de cointegración subyace la idea de un *equilibrio* a largo plazo (una relación determinista que se cumple en la media) entre dos variables integradas (de memoria larga), x_t, y_t . Existe una relación estricta (lineal) de equilibrio cuando para algún $\alpha \neq 0$, se tiene $y_t = \alpha x_t$. Esta situación determinista tan poco realista se sustituye en la práctica por la relación de cointegración, en la que error de equilibrio estocástico $z_t = y_t - \alpha x_t$ puede tomar valores distintos de cero, pero debe fluctuar alrededor de su valor medio con una frecuencia mucho mayor que las series individuales (este comportamiento por parte de los errores de equilibrio es lo que conoce como *memoria corta*).

Las definiciones habituales de cointegración y los contrastes más comunes parten de un determinado tipo de proceso generador de datos para ambas series individuales y la relación existente entre ambas, y suponen que dicho proceso es lineal, característica que también comparte la relación de cointegración. Como consecuencia, la cointegración suele interpretarse como una tendencia uniforme de las series a moverse hacia un único equilibrio de largo plazo. Esto contrasta con los equilibrios múltiples que se observan par numerosos pares de series económicas (ver, por ejemplo, Escribano (1986) y Escribano y Granger (1998)).

Los modelos de equilibrio general dinámico son generalmente no lineales por construcción. Otras relaciones no lineales se derivan, por ejemplo, de las Teorías del consumidor o de la producción, y pueden tomar la forma de asimetrías en los precios de ajuste o no convexidades en la función de producción, como en Escribano y Pfann (1998). Esta teoría sugiere que los agentes económicos realizarán ajustes de forma

continua sólo si los beneficios derivados de los ajustes son mayores que los costes de los mismos (Balke y Fomby, 1997). Por ejemplos, en los mercados financieros los costes de transacción permiten la aparición de una banda en la que los rendimientos pueden ser divergentes, introduciendo así ineficiencias y la posibilidad de arbitrajes. Las políticas intervencionistas, tales como las de control de tipo de cambio por parte de los bancos centrales, o la estabilización de los precios de algunos bienes a partir de compras o ventas de stocks por parte de las distintas administraciones, también pueden dar lugar a no-uniformidades en el ajuste de los agentes. Estas no-uniformidades se traducen en desviaciones de la hipótesis de cointegración lineal, tal y como se recoge en Escribano (1986) para un ejemplos de la demanda de dinero en Reino Unido.

Las series de memoria larga no lineales y las relaciones de cointegración no lineales también aparecen en aplicaciones de Ingeniería. Este tipo de series aparece, por ejemplo al analizar y comparar señales de salida de sensores en respuesta a algunos tipo de flujos persistentes. El problema de la estimación de las relaciones de cointegración aparece de forma natural al comparar los comportamientos de los sensores ante el mismo flujo de entrada. También aparecen problemas similares en Ingeniería de Telecomunicaciones al estudiar las señales transmitidas a través de un canal de comunicaciones distorsionante, algunas de las cuales tienen correlaciones de orden muy alto, tal y como se recoge, por ejemplo, en Mandelbrot (1965, 1967), Leland et al. (1994), Beran et al. (1991), y Willinger et al. (1995).

Es importante introducir medidas de cointegración no paramétricas (independientes del modelo) para que se pueda extender el concepto de cointegración en un contexto no lineal, tal y como se hace con las medidas basadas en teoría de la información en Aparicio y Escribano (1999). Para ello se empieza analizando una serie de conceptos generales, como son *memoria corta*, *memoria larga* y *cointegración*, de tal forma que nos lleven a una nueva caracterización de la cointegración.

La estructura del presente Capítulo es la siguiente: la Sección 2 presenta diferentes medidas de dependencia lineal (correlación) en el contexto de series temporales integradas de orden d , $I(d)$, donde d puede ser cualquier número entero o real. En la Sección 3 se introduce una nueva medida no paramétrica de cointegración lineal en el dominio temporal, encontrándose la relación existente con los conceptos habituales para series $I(1)$ y de memoria larga (cointegración fraccional). Se demuestra que tiene una interpretación natural como estimador de variables instrumentales en el que el instrumento es a algún retardo de la variable explicativa. La Sección 4 incluye un análisis de las propiedades en muestras pequeñas del estimador de variables instrumentales

basado en experimentos de simulación de Monte Carlo con variables $I(1)$ y variables con órdenes de integración fraccionales. El concepto de cointegración se analiza en la Sección 5 en el dominio de frecuencias. La relación de los conceptos de cointegración definidos en los dominios temporal y de frecuencias se estudia en la Sección 6, basándonos en resultados analíticos y alguna simulación. Por último, la Sección 7 incluye algunas observaciones finales y avanza algunas líneas de investigación futura.

2. Definiciones de memoria en series temporales

La caracterización habitual de las series integradas se realiza en función de los modelos *ARIMA*.

DEFINICIÓN 2.1. Una serie temporal x_t se dice *ARIMA*(p, d, q), donde $d \in \mathfrak{R}$, si tras haberla diferenciado d veces, tiene una representación *ARMA*(p, q) estacionaria e invertible, siendo p, q enteros no negativos.

Por tanto, si $x_t \sim \text{ARIMA}(p, d, q)$ existen los polinomios $\Phi(B)$ y $\Theta(B)$ en el operador B , de órdenes $p \geq 0$ y $q \geq 0$ respectivamente, con todas las raíces fuera del círculo unitario y sin factores en común, de forma que se puede escribir

$$\Phi(B)(1 - B)^d x_t = \Theta(B)\epsilon_t, \quad (2.1)$$

donde ϵ_t se supone habitualmente que es una sucesión de errores de media cero, idéntica e independiente distribuidos con distribución Normal. Sea Δ el operador diferencia, de forma que $\Delta^d x_t = (1 - B)^d x_t$. Según Hosking (1981), cuando d no es entero se puede expresar:

$$\begin{aligned} \Delta^d x_t &= \sum_{k=0}^{\infty} \frac{d!}{k!(d-k)!} (-B)^k x_t \\ &= x_t - dx_{t-1} - \frac{1}{2}d(1-d)x_{t-2} \\ &\quad - \frac{1}{6}d(1-d)(2-d)x_{t-3} - \dots \end{aligned}$$

Cuando el parámetro d es positivo, suele denominarse *parámetro de memoria larga*, y determina la tasa de decrecimiento de la dependencia lineal de x_t al aumentar los retardos. Es decir, si $d > 0$ el proceso x_t tiene memoria larga, en tanto que tiene memoria corta si $d = 0$. Para valores $d < 0$ x_t se suele decir que es *antipersistente*. Además, si $d < \frac{1}{2}$, entonces x_t es estacionario, en tanto que es no estacionario para valores $d \geq \frac{1}{2}$, ver Granger y Joyeux (1980). Por último, sólo en el caso en que $d < 1$ es x_t reversible a la media.

Una propiedad de los procesos x_t Gaussianos de memoria corta es que su función de autocorrelación (ACF) converge a cero a una tasa *exponencial* cuando τ crece a infinito, ver Box y Jenkins (1970), es tanto que esta tasa es *hiperbólica* si $0 < d < 1$, ver Granger y Joyeux (1980).

En el caso estacionario lineal, las medidas de dependencia serial para una serie temporal x_t con media $E(x_t) = \mu_x$ se basan en su función de autocovarianzas, $cov(x_t, x_{t-\tau})$, definida como

$$cov(x_t, x_{t-\tau}) = E[(x_t - \mu_x)(x_{t-\tau} - \mu_x)]. \quad (2.2)$$

Al tratar con series temporales integradas, estamos permitiendo que tanto la función de autocovarianzas como la varianza del proceso estocástico dependan del valor inicial finito que toman en el momento $t = 0$, tal como $x_0 = 0$. Recordemos que en este caso la varianza de un proceso no estacionario es finita sólo para valores finitos de t . Por tanto, a partir de ahora, todas las esperanzas consideradas en este Capítulo, $E(x_t) = E(x_t|x_0)$, $cov(x_t, x_{t-\tau}) = E[(x_t - \mu_x)(x_{t-\tau} - \mu_x)|x_0]$, y $cov(x_t, y_{t-\tau}) = E[(y_t - \mu_{y_t})(x_{t-\tau} - \mu_{x_{t-\tau}})|x_0, y_0]$, son todas ellas condicionales respecto a los valores iniciales en $t = 0$, pongamos $x_0 = 0$, e $y_0 = 0$.

DEFINICIÓN 2.2. Un proceso estocástico x_t se dice que tiene **memoria corta**¹ si existe un real finito y positivo b tal que $\lim_{T \rightarrow \infty} \sum_{\tau=1}^T |cov(x_t, x_{t-\tau})| = b$.

DEFINICIÓN 2.3. Un proceso estocástico x_t se dice que tiene **memoria larga** si $\lim_{T \rightarrow \infty} \sum_{\tau=1}^T |cov(x_t, x_{t-\tau})| = \infty$.

Basado en estos conceptos, es natural definir el concepto de cointegración de una forma independiente de modelos.

DEFINICIÓN 2.4. Un proceso estocástico x_t se dice que es **integrado de orden d** , ($x_t \sim I(d)$) si es de memoria larga y d es el menor real tal que $(1 - B)^d x_t$ es de memoria corta.

Obsérvese que las series temporales generadas por modelos $ARIMA(p, d, q)$, definidos en la ecuación (2.1) satisfacen esta definición de procesos $I(d)$. En la Sección siguiente se extiende este enfoque al concepto de cointegración.

3. Cointegración lineal en el dominio temporal

Sean x_t, y_t dos series temporales de media cero e integradas de órdenes $d_x, d_y \in \mathbb{R}^+$, respectivamente. Es decir, $x_t \sim I(d_x)$, $y_t \sim I(d_y)$, lo que significa que $(1 - B)^{d_x} x_t =$

¹Este concepto de memoria corta está directamente relacionado con el de incorrelación asintótica. Ver, por ejemplo, White (1984) y Escribano (1987b).

ϵ_t , $(1 - B)^d y_t = \xi_t$, donde ϵ_t, ξ_t son series de *memoria corta*. Sea $z_t = y_t - \alpha x_t$, para algún numeral real α distinto de cero.

DEFINICIÓN 3.1. (Granger, 1981) Dos series temporales x_t, y_t $I(d)$, con $d > 0$, se dice que están (linealmente)² **cointegradas** si $\exists \alpha \in \Re - \{0\}$ tal que la serie $z_t = y_t - \alpha x_t$ es $I(d_z)$ con $d_z < d$.

Observaciones:

- Cuando x_t, y_t están cointegradas, z_t tiene media cero y es $I(0)$, y_t , y x_t tienden a moverse conjuntamente en el largo plazo, a pesar de que sus movimientos en el corto plazo pueden no estar ‘alineados’.
- Desde el punto de vista económico, el caso más interesante es cuando $d = 1$, $d_z = 0$, puesto que se puede interpretar como la existencia de equilibrio (lineal) a largo plazo. Sin embargo, la definición anterior de cointegración no implica necesariamente la existencia de un equilibrio entre las dos series $I(d)$ ($d > 0$), puesto que sólo requería los residuos de cointegración z_t sean $I(d_z)$ con $d_z < \min(1, d)$. Para que se observe un equilibrio se requiere que z_t sea reversible a la media ($d_z < 1$).

Se propone una caracterización alternativa del concepto de cointegración lineal, restringiendo su aplicación al caso no trivial de que todas las series sean mutuamente dependientes. Sean x_t, y_t dos series temporales $I(d)$ con $d > 0$, con valores iniciales en $t = 0$ $x_0 = 0$ e $y_0 = 0$, respectivamente. Sea $cov(x_t, y_{t-\tau})$ la *función de covarianzas cruzadas* de x_t, y_t condicional a x_0, y_0 para $t \geq \tau$, definida como

$$cov(y_t, x_{t-\tau}) = E[(y_t - \mu_{y_t})(x_{t-\tau} - \mu_{x_{t-\tau}})] \quad (3.1)$$

donde $\mu_{x_t} = E(x_t)$ y $\mu_{y_t} = E(y_t)$ son las esperanzas de x_t e y_t condicionadas a $x_0 = 0$ e $y_0 = 0$ respectivamente.

Se propone la siguiente caracterización no paramétrica de la cointegración:

TEOREMA 5.3.1. *Un par de series temporales x_t, y_t , $I(d)$ están (linealmente) cointegradas con vector de cointegración $\beta' = (1, -\alpha)$ si, para algún $\tau < t$ fijo o para $\tau = o(t)$*

$$\lim_{\tau \rightarrow \infty} \frac{cov(y_t, x_{t-\tau})}{cov(x_t, x_{t-\tau})} = \alpha. \quad (3.2)$$

DEMOSTRACIÓN. En Sección 8. □

²En Granger (1981) no hay ninguna mención explícita del término *lineal*.

COROLARIO 5.3.1. Sean y_t, x_t dos series $I(d)$ linealmente cointegradas con vector de cointegración $\beta' = (1, -\alpha)$, con d restringido al intervalo $(0, 1]$. si $\lim_{\tau \rightarrow \infty} \frac{\text{cov}(y_t, x_{t-\tau})}{\text{var}(x_{t-\tau})} = \alpha$ con $\alpha \neq 0$, entonces $x_t, y_t \sim I(1)$

DEMOSTRACIÓN. En Sección 8. □

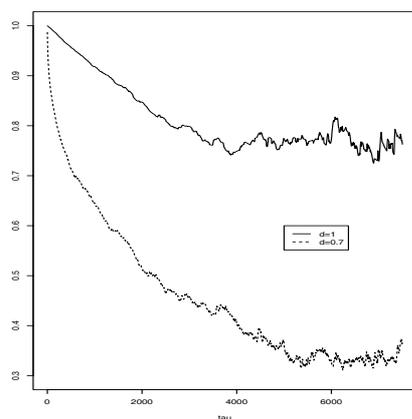


Figura 5.1.1: $\hat{\alpha}_{ols}$, $b = -0.5$

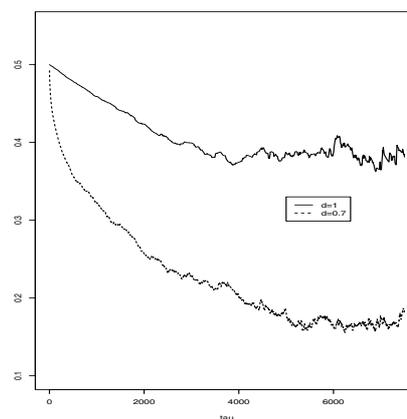


Figura 5.1.2: $\hat{\alpha}_{ols}$, $b = 0.0$

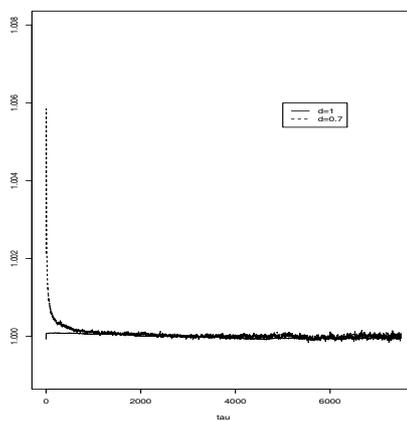


Figura 5.1.3: $\hat{\alpha}_{ive}(\tau)$, $b = -0.5$

FIGURA 5.1. Simulaciones de $\hat{\alpha}_{ols}(\tau)$, ecuación (3.5) y $\hat{\alpha}_{ive}(\tau)$, ecuación (3.4) $a = 0.5$, $s = 1$, tamaño muestral 10000, media de 500 réplicas.

Observaciones:

- Intuitivamente, el Teorema 5.3.1 indica que, si hay cointegración, el pasado remoto de x_t debe ser tan útil como el pasado remoto de y_t en la predicción a largo plazo de y_t .
- El Teorema 5.3.1 implica que las tasas de convergencia de $cov(y_t, x_{t-\tau})$ y de $cov(x_t, x_{t-\tau})$ cuando τ crece sin límite, deben ser iguales. Por ejemplo, supongamos $cov(y_t, x_{t-\tau}) \sim \alpha \tau^{-b} var(x_{t-\tau})$ y $cov(x_t, x_{t-\tau}) \sim \tau^{-a} var(x_{t-\tau})$ para τ grande. En general, esperamos $a \leq b$, pero bajo cointegración lineal debe cumplirse la igualdad. En un análisis preliminar, un gráfico de $cov(y_t, x_{t-\tau})/cov(x_t, x_{t-\tau}) \sim \alpha \tau^{a-b}$ versus τ debería ser útil para identificar la existencia de una relación de cointegración lineal ($a = b$). La Figura 5.1.3 muestra que esto es cierto incluso para valores pequeños de τ .
- Si y_t, x_t tiene media cero, el ratio $\frac{cov(y_t, x_{t-\tau})}{cov(x_t, x_{t-\tau})} = \frac{E(y_t, x_{t-\tau})}{E(x_t, x_{t-\tau})}$, y proporciona un estimador de variables instrumentales (VI) del parámetro de cointegración, en el que el instrumento es $x_{t-\tau}$. Esto es, consideremos la relación de cointegración lineal

$$y_t = \alpha x_t + z_t, \quad (3.3)$$

donde x_t, y_t son $I(d)$ con $d > 0$ y z_t es $I(0)$, entonces el estimador VI de α viene dado por

$$\hat{\alpha}_{iv} = \frac{\sum_{t=\tau+1}^N y_t x_{t-\tau}}{\sum_{t=\tau+1}^N x_t x_{t-\tau}}, \quad (3.4)$$

donde N indica el tamaño muestral disponible. La Figura 5.1.3 muestra la rápida convergencia de $\hat{\alpha}_{iv}$ incluso para valores pequeños de τ . Además, Engle y Granger (1987) y Stock (1987) para $I(1)$, y Dolado y Marmol (1998) para $I(d)$ con d fraccional, muestran que α puede estimarse consistentemente mediante MCO a partir de la regresión (3.3), donde

$$\hat{\alpha}_{ols} = \frac{\sum_{t=1}^N y_t x_t}{\sum_{t=1}^N x_t^2}. \quad (3.5)$$

- La condición límite en el Teorema 5.3.1 no puede comprobarse en la práctica a no ser que τ sea finita. Sin embargo, cuando se analiza la cointegración lineal en aplicaciones empíricas, esta condición se satisface habitualmente para valores pequeños de τ , tal y como se comprobará en la Sección 4. Un ejemplo importante ocurre cuando $y_t = \alpha x_t + z_t$, con $\alpha \neq 0$, y donde x_t, y_t son $I(1)$ y z_t es una sucesión de variables aleatorias *i.i.d.* e independientes de x_t . En este caso, $cov(y_t, x_{t-\tau})/cov(x_t, x_{t-\tau}) = \alpha$ para cualquier $\tau < t$.

- Si las series son de memoria corta pero mutuamente correladas, el ratio de la función de covarianza cruzada sobre la función de autocovarianza puede también converger eventualmente a un número no negativo, y esta es la razón por la que se impone en el Teorema 5.3.1 que las series sean de memoria larga.
- La expresión del Corolario 5.3.1 puede estimarse por MCO regresando y_t sobre $x_{t-\tau}$.

Los dos ejemplos siguientes demuestran que el Teorema 5.3.1 no lo satisfacen pares de series temporales no cointegradas (Ejemplo 3.1), en tanto que sí se verifica para variables generadas por un modelo (lineal) de *factores comunes* (Ejemplo 3.2)

EJEMPLO 3.1. Consideremos el siguiente par de series no cointegradas:

$$x_t = w_t + \xi_t \quad (3.6)$$

$$y_t = q_t + v_t \quad (3.7)$$

donde w_t, q_t son series $I(d)$ mutuamente dependientes, y ξ_t, v_t siguen procesos $ARMA(p, q)$, con órdenes AR y MA generalmente distintos, e independientes de w_t y q_t , respectivamente. Con estas condiciones es evidente que

$$\text{cov}(y_t, x_{t-\tau}) = \text{cov}(v_t, \xi_{t-\tau}). \quad (3.8)$$

La covarianza $\text{cov}(v_t, \xi_{t-\tau})$ decrecerá exponencialmente al crecer τ a infinito dado que las series v_t, ξ_t son ambas $I(0)$. Por el contrario, $\text{cov}(x_t, x_{t-\tau})$ crecerán sin límite si $d = 1$, es decir $\text{cov}(x_t, x_{t-\tau}) = (t - \tau)\sigma_\epsilon^2$, o decaen hiperbólicamente al crecer τ para $\tau = o(t)$. En ambos casos tenemos

$$\lim_{t \rightarrow \infty} \frac{\text{cov}(y_t, x_{t-\tau})}{\text{cov}(x_t, x_{t-\tau})} = 0. \quad (3.9)$$

EJEMPLO 3.2. Consideremos el siguiente modelo lineal de de factor común:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} \alpha \\ 1 \end{pmatrix} w_t + \begin{pmatrix} v_t \\ \xi_t \end{pmatrix} \quad (3.10)$$

con $\alpha \neq 0$ y donde $w_t = w_{t-1} + \epsilon_t$ y (v_t, ξ_t, ϵ_t) son sucesiones de variables aleatorias idéntica e independientemente distribuidas con distribución Normal con media cero y matriz de covarianzas conjunta igual a la matriz identidad. Sea $\beta'_\perp = (\alpha, 1)$, donde β'_\perp es la transpuesta de β_\perp . Por tanto, el complemento ortogonal de β'_\perp es $\beta' = (1, -\alpha)$.

Por tanto, la relación de cointegración se obtiene como

$$z_t = \beta' \begin{pmatrix} y_t \\ x_t \end{pmatrix} = y_t - \alpha x_t = v_t - \alpha \xi_t, \quad (3.11)$$

lo que indica que z_t es $I(0)$.

A partir de la ecuación (2.2), la autocovarianza de x_t es

$$\text{cov}(x_t, x_{t-\tau}) = (t - \tau)\sigma_\epsilon^2 \quad (3.12)$$

en tanto que la covarianza cruzada de y_t y x_t viene dada por

$$\text{cov}(y_t, x_{t-\tau}) = \alpha(t - \tau)\sigma_\epsilon^2. \quad (3.13)$$

Obviamente, su ratio es α , y puesto que $\alpha \neq 0$ por hipótesis, se deduce que las series y_t, x_t presentan cointegración lineal.

En la Sección siguiente se analiza la evidente mejora en muestras pequeñas obtenida con el análogo muestral del concepto de cointegración no paramétrico.

4. Experimento de simulación de Monte Carlo

El proceso generador de datos (PGD) sigue el diseño del experimento del Capítulo 2. Este PGD es una extensión del introducido por Engle y Granger (1987), Banerjee et al. (1986), y Gonzalo (1994) para analizar las propiedades en muestras finitas del estimador superconsistente MCO del vector de cointegración.

El PGD es un sistema cointegrado bivalente de series fraccionalmente integradas $I(d)$ con dos representaciones equivalentes. La primera se basa en la *regresión de cointegración*, con errores de equilibrio (z_t) estacionarios $I(0)$

$$y_t = \alpha x_t + z_t \quad (4.1a)$$

$$z_t = (a - \alpha)\Delta^d x_t + (1 + b)z_{t-1} + u_{1,t} \quad (4.1b)$$

$$\Delta^d x_t = u_{2,t} \quad (4.1c)$$

donde los términos de error $u_{1,t}, u_{2,t}$ son $iidN(0, \sigma_1^2)$, e $iidN(0, \sigma_2^2)$ con $\text{cov}(u_{1,t}, u_{2,t}) = 0$.

La segunda representación se basa en el siguiente *modelo de corrección del error* (MCE) de series con integración fraccional

$$\Delta y_t = \alpha \Delta x_t + (a - \alpha)\Delta^d x_t + b(y_{t-1} - \alpha x_{t-1}) + u_{1,t} \quad (4.2a)$$

$$\Delta^d x_t = u_{2,t} \quad (4.2b)$$

Obsérvese que el MCE no tiene la especificación habitual en el sentidos de que no aparecen sólo variables en primeras diferencias y en niveles. No obstante, si se satisface la restricción de factor común, ($a - \alpha = 0$), el MCE es de la forma

$$\Delta y_t = \alpha \Delta x_t + b(y_{t-1} - \alpha x_{t-1}) + u_{1,t} \quad (4.3)$$

incluso cuando las variables presentan integración fraccional, $I(d)$, como puede ser $d = 0.7$.

El Teorema 5.4.1 muestra la importancia de los parámetros a y b a la hora de hacer inferencia sobre los estimadores mínimo cuadráticos del vector de cointegración α .

TEOREMA 5.4.1. *Dado el PGD (4.1a)–(4.1c) con $d = 1$, bajo condiciones de regularidad estándar, cuando $T \rightarrow \infty$,*

$$\begin{aligned} T(\hat{\alpha}_{ols} - \alpha) \xrightarrow{d} & \left\{ \int_0^1 B_2^2(r) dr \right\}^{-1} \left\{ \frac{\sigma_1}{-b} \left[1 - (a - \alpha)^2 \frac{\sigma_2^2}{\sigma_1^2} \right] \int_0^1 B_2(r) dW_1(r) + \right. \\ & \left. + \left(\frac{a - \alpha}{-b} \right) \int_0^1 B_2(r) dB_2(r) + \left(\frac{(a - \alpha)\sigma_2}{-b} \right) \right\} \end{aligned} \quad (4.4)$$

donde $W_1(r)$ es un proceso de Wiener estándar, $B_2(r)$ es un movimiento Browniano con varianza de largo plazo Σ_{22} y $W_1(r)$ y $B_2(r)$ son independientes.

DEMOSTRACIÓN. En Sección 8 □

En las simulaciones de Monte Carlo se toma el siguiente conjunto de parámetros: $\alpha = 1$, $a = -1, -0.5, 0$, y 1 (COMFAC), $b = -0.2, -0.5, -1$, $\sigma_1^2 = 1$, $\sigma_2^2 = s^2$, donde $s = 1, 6, 16$. El tamaño de la muestra es $N = 100$, las variables son integradas $I(d)$ para $d = 1, 0.7, 1.3$ y el número de réplicas es $T = 20000$.

Se analizan los sesgos en muestras pequeñas del estimador de variables instrumentales (3.4) para diferentes valores del retardo τ de la variable instrumental ($x_{t-\tau}$). Los resultados se muestran en las Tablas 5.1 a 5.5. Obsérvese que cuando $\tau = 0$, el estimador VI coincide con el de MCO (3.5), con la distribución asintótica dada en la ecuación (4.4).

Para estudiar las propiedades en muestras finitas del estimador se toman cinco estadísticos:

- (1) Sesgo en media: $(1/T) \sum_{i=1}^T \hat{\alpha}_i - \alpha$
- (2) Sesgo en mediana: $med(\hat{\alpha}_1, \dots, \hat{\alpha}_T) - \alpha$
- (3) ECM: (sesgo en media)² + varianza de $(\hat{\alpha}_1, \dots, \hat{\alpha}_T)$
- (4) Rango intercuartílico (RIQ): $Q_3 - Q_1$

(5) Probabilidad de concentración: $\Pr(|\hat{\alpha} - \alpha| \leq 0.05)$

Las Tablas 5.1 a 5.3 de la Sección 9 muestran los resultados de las simulaciones para variables $I(1)$, en tanto que los resultados para variables con integración fraccional, $I(d)$, se recogen en las Tablas 5.4 y 5.5 para $d = 0.7, 1.3$. Se muestran los resultados obtenidos para $\tau = 1, 2, \dots, 5$, y los diferentes valores del parámetro a . A pesar de que se han estudiado diferentes valores del parámetro s , concretamente $s = 1, 6, 16$, los resultados mostrados son sólo para $s = 6$, manteniéndose las conclusiones para los demás casos. Tal y como se comentó con anterioridad, el parámetro b del término de ajuste del error en la ecuación (4.2a), es fundamental en términos del sesgo del estimador. Cuanto más próximo está b a 0, mayor es la autocorrelación de los errores de equilibrio, z_t . Por ejemplo, $b = -0.2$ se corresponde a un coeficiente $AR(1)$ de valor 0.8 en la ecuación (4.1b).

Cuando $a = 1$ se cumple la *restricción de factor común*, ($a - \alpha = 0$), y por tanto el estimador MCO de la ecuación (4.1a) debe tener un buen comportamiento en muestras pequeñas (sesgos bajos y probabilidad de concentración próxima a 1). La razón es obvia, puesto que el regresor contemporáneo, Δx_t , desaparece de la esperanza condicional de la ecuación (4.1b), y Δy_t y Δz_t no están correlados en el corto plazo. Por otra parte, el sesgo del estimador MCO debe aumentar con el valor absoluto de $(a - 1)$. En concreto, como se observa en la Tabla 5.1, para $b = -0.2$, y $(a - 1) = -0.5$, el sesgo en media es -4.7% , el sesgo en mediana -3.5% , el ECM 0.4% , el RIQ 4.5% , y la probabilidad de concentración 66% . Para $(a - 1) = -1$ el sesgo en media -9.5% , el sesgo en mediana -6.9% , el ECM es 17% , el RIQ 8.8% y la probabilidad de concentración 35% . Los peores resultados se obtienen para $(a - 1) = -2$: sesgo en media -19% , sesgo en mediana -13.9% , ECM 6.6% , RIQ 17.4% , y probabilidad de concentración 14% .

Las mejoras del nuevo estimador VI para $\tau = 3$ son evidentes. De nuevo, cuando se cumple la restricción de factor común, $(a - 1) = 0$, tanto el estimador MCO como el VI funcionan bien. Para otros parámetros, tales como $(a - 1) = -0.5$, el estimador VI tiene mejor comportamiento que el de MCO en términos de sesgos y dispersión: el sesgo en media es 0.87% , el sesgo mediano -1.1% , el ECM 0.1% , el RIQ 3% , y la probabilidad de concentración es 90% . Para $(a - 1) = -1$ la diferencia del comportamiento es aún más favorable al estimador VI: el sesgo en media es -1.7% , el sesgo en mediana -2.1% , el ECM 0.3% , el RIQ 5.4% y la probabilidad de concentración es 69% . Por último, para el valor más alto, $(a - 1) = -2$, se obtiene mejor resultado al utilizar el estimador VI que con el estimador MCO: el sesgo en media es

−3.4%, el sesgo en mediana −4%, el ECM 1.3%, el RIQ 10.3% y la probabilidad de concentración es 41%.

Cuando $b = -0.5$, el coeficiente $AR(1)$ de los errores de equilibrio (z_t) de la ecuación (4.1b) es 0.5. Por tanto, la memoria de dependencia temporal se reduce, originando sesgos menores con valores de τ más bajos, ver Tabla 5.2. Los mejores resultados se obtienen para $\tau = 1$ y $\tau = 2$. El estimador MCO tiene un comportamiento peor que el estimador VI, pero mejor que el estimador de MCO cuando teníamos $b = -0.2$.

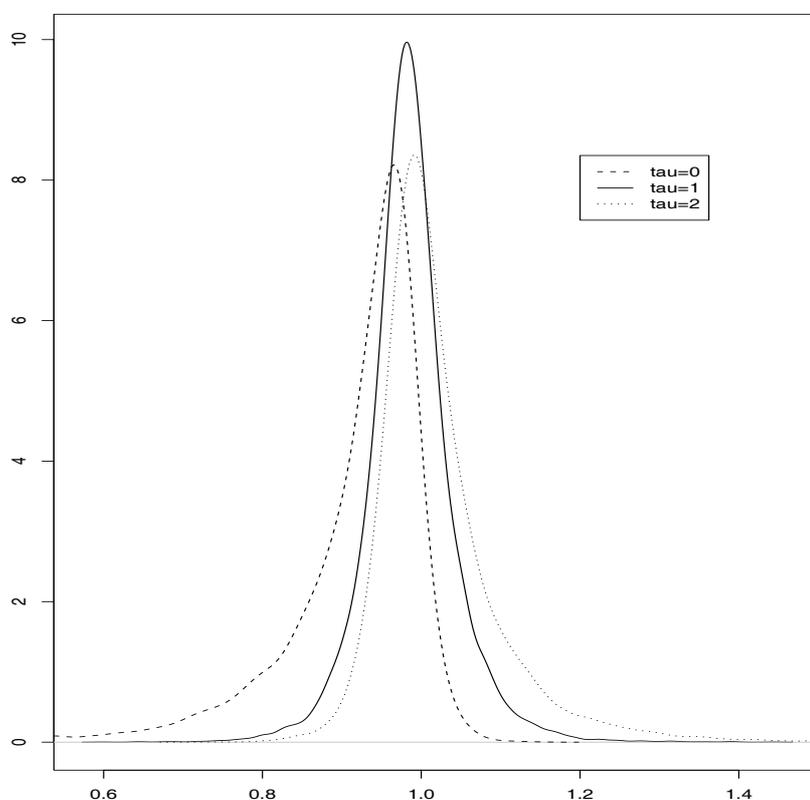


FIGURA 5.2. Densidad del estimador VI para $b = -0.5$, $a - \alpha = -2$, $N = 100$

La Figura 5.2 resulta muy reveladora acerca de los efectos de τ en la distribución del estimador VI. Se observa que cuando $\tau = 0$ (MCO) la distribución del estimador de α es asimétrica a la izquierda. Al hacer $\tau = 1$, la media y la mediana de $\hat{\alpha}_{iv}$ (las medidas centrales) están más próximas a 1.0, el verdadero valor de α , la dispersión disminuye y la distribución es simétrica. No obstante, si hacemos $\tau = 2$ el sesgo en

media aumenta, aunque la mediana se aproxima aún más a valor 1.0, y la dispersión se sitúa en un punto entre las que se obtenían con los valores $\tau = 0$ y $\tau = 1$.

Tal y como era de esperar para *variables débilmente exógenas* para el parámetro de cointegración, la mejora del estimador VI sobre el estimador MCO es marginal cuando los errores de equilibrio son ruido blanco, $b = -1$, ver Tabla 5.3. El mejor estimador VI en términos de sesgo en media, ECM y RIQ es $\tau = 1$, pero $\tau = 2$ da como resultado prácticamente el mismo RIQ y el sesgo en mediana es ligeramente inferior.

En el caso de que las variables presenten integración fraccional, $I(d)$, con $d = 0.7$, la mejora del estimador VO sobre el estimador MCO es evidente, tal y como se desprende de la Tabla 5.4. El estimador MCO sólo es mejor cuando se cumple la restricción COMFAC, $(a-1) = 0$. En los restantes casos analizados, las probabilidades de concentración del estimador MCO nunca son mayores de 11%, y en la mayoría de los casos están por debajo de la barrera de 1%, y los sesgos en media son muy grandes, alcanzando el valor 53% cuando $(a-1) = -2$. Por otra parte, al considerar el estimador VI con $\tau = 3$, la mejora es espectacular. Incluso cuando no hay COMFAC, en el peor caso, $(a-1) = -2$, la probabilidad de concentración es 28%, el sesgo en media 5.6%, el sesgo en mediana 1.54% (comparado con 46.3% con el estimador MCO), y el RIQ es 19.2% (comparado con 36.5% del estimador MCO). Por tanto el nuevo estimador VI del vector de cointegración produce importantes reducciones del sesgo respecto al estimador MCO.

En la siguiente Sección se procede a realizar una interpretación del nuevo concepto de cointegración en el dominio de las frecuencias.

5. Cointegración lineal en el dominio de frecuencias

Sean de nuevo las series $x_t \sim I(d_x)$, $y_t \sim I(d_y)$, y la serie z_t formada como $z_t = y_t - \alpha x_t$. Para ilustrar el concepto de cointegración lineal en el dominio de frecuencias se considera el espectro y el espectro cruzado para las diferentes series³. A partir de Granger (1981), es fácil deducir de la definición de z_t que para $0 < \lambda < \pi$,

$$S_z(\lambda) = S_y(\lambda) + \alpha^2 S_x(\lambda) - \alpha(S_{y,x}(\lambda) + S_{y,x}^*(\lambda)), \quad (5.1)$$

donde $S_u(\lambda)$ y $S_{y,x}(\lambda)$ representan el espectro de u_t ($u = x, y, z$) y el espectro cruzado del par x_t, y_t , respectivamente, y $S_{y,x}^*(\lambda)$ indica el complejo conjugado de $S_{y,x}(\lambda)$. Puesto que $|S_{y,x}(\lambda)|^2 \leq S_x(\lambda)S_y(\lambda)$, para todo λ , y $S_x(\lambda) \sim A_x \lambda^{-2d_x}$, $S_y(\lambda) \sim$

³Para series integradas no estacionarias siempre se pueden considerar el *pseudo-spectrum* o *pseudo-cross-spectrum* como en Harvey (1989, p. 64).

$A_y \lambda^{-2d_y}$ cuando $\lambda \rightarrow 0$, es obvio que el término $\lambda^{-2\max(d_x, d_y)}$ dominará en las frecuencias bajas, y, por tanto, $z_t \sim I(\max[d_x, d_y])$ a no ser que se cumpla una restricción de cointegración. En efecto, bajo cointegración (lineal), el álgebra anterior no es cierta, y existe un real distinto de cero, α , tal que $z_t = y_t - \alpha x_t \sim I(d_z), d_z < d$. Esto es lo mismo que decir que existe un número real finito positivo c tal que $\lim_{\lambda \rightarrow 0} (S_y(\lambda)/S_x(\lambda)) = c$, y, por otro lado, que $\lim_{\lambda \rightarrow 0} (S_z(\lambda)/S_x(\lambda)) = 0$. Por tanto, $\lim_{\lambda \rightarrow 0} [(\alpha^2 + c)S_x(\lambda) - \alpha S_{y,x}(\lambda)] = 0$, y puesto que α y c son distintos de cero, el límite $\lim_{\lambda \rightarrow 0} S_{y,x}(\lambda)/S_x(\lambda)$ debe ser finito y distinto de cero. Intuitivamente, esto quiere decir que x_t, y_t tienen el componente de onda larga en común, lo que equivale al enunciado del Teorema 5.3.1. En efecto, definamos

$$\sigma_x^2(t) = \text{var}(x_t) \quad (5.2)$$

$$S_x(\lambda, t) = \sigma_x^2(t) \left(1 + 2 \sum_{\tau=1}^t \text{cov}(x_t, x_{t-\tau}) \exp(-j\lambda\tau) \right) \quad (5.3)$$

$$S_{y,x}(\lambda, t) = \sigma_x^2(t) \left(1 + 2 \sum_{\tau=1}^t \text{cov}(y_t, x_{t-\tau}) \exp(-j\lambda\tau) \right), \quad (5.4)$$

donde $j^2 = -1$. Se observa que al emplear esta notación, los (pseudo) espectros $S_x(\lambda)$ e $S_{y,x}(\lambda)$ se pueden obtener como el límite de t , es decir,

$$S_x(\lambda) = \lim_{t \rightarrow \infty} S_x(\lambda, t) \quad (5.5)$$

$$S_{y,x}(\lambda) = \lim_{t \rightarrow \infty} S_{y,x}(\lambda, t). \quad (5.6)$$

La ventaja de introducir esta notación es que tanto $S_x(\lambda, t)$ como $S_{y,x}(\lambda, t)$ existen para t finito. Por tanto, para valores de t finitos, se tiene

$$\lim_{\lambda \rightarrow 0} \frac{S_{y,x}(\lambda, t)}{S_x(\lambda, t)} = \frac{1}{1 + 2s_t^{(x,x)}} + \frac{2s_t^{(y,x)}}{1 + 2s_t^{(x,x)}}. \quad (5.7)$$

Pero, puesto que $s_t^{(x,x)}$ es divergente para $t \rightarrow \infty$, para todo T tal que $\tau < T < t$ se tiene

$$\lim_{t \rightarrow \infty, \lambda \rightarrow 0} \frac{S_{y,x}(\lambda, t)}{S_x(\lambda, t)} = \lim_{t \rightarrow \infty} \frac{\sum_{\tau=1}^T \text{cov}(y_t, x_{t-\tau}) + \sum_{\tau>T}^t \text{cov}(y_t, x_{t-\tau})}{\sum_{\tau=1}^T \text{cov}(x_t, x_{t-\tau}) + \sum_{\tau>T}^t \text{cov}(x_t, x_{t-\tau})}, \quad (5.8)$$

y puesto que el primer término del numerador y el denominador del lado derecho de la fórmula anterior está acotado para T finito, se pueden eliminar para valores de t grandes, por lo que

$$\lim_{t \rightarrow \infty, \lambda \rightarrow 0} \frac{S_{y,x}(\lambda, t)}{S_x(\lambda, t)} = \lim_{t \rightarrow \infty} \frac{\sum_{\tau>T}^t \text{cov}(y_t, x_{t-\tau})}{\sum_{\tau>T}^t \text{cov}(x_t, x_{t-\tau})}. \quad (5.9)$$

La expresión anterior es válida para todo T finito. Si permitimos que T crezca sin límite al tender t a infinito, se tiene que, para $\tau = o(t)$

$$\lim_{t \rightarrow \infty, \lambda \rightarrow 0} \frac{S_{y,x}(\lambda, t)}{S_x(\lambda, t)} = \lim_{t \rightarrow \infty} \frac{\text{cov}(y_t, x_{t-\tau})}{\text{cov}(x_t, x_{t-\tau})} = \alpha. \quad (5.10)$$

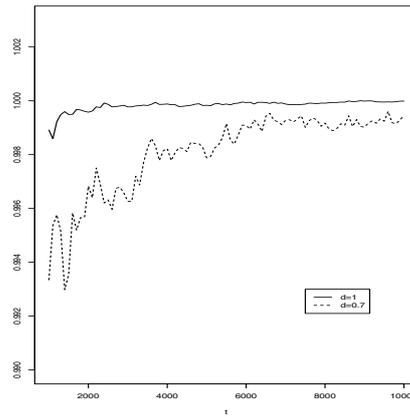


Figura 5.3.1: $b = -0.5$,

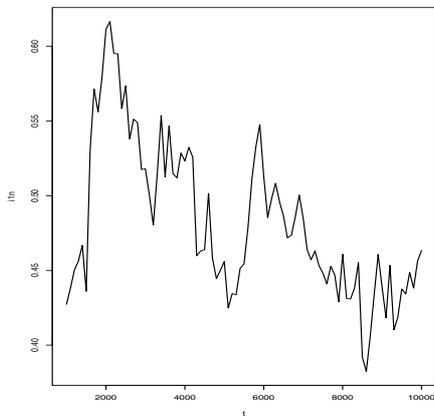


Figura 5.3.2: $b = 0.0, I(1)$

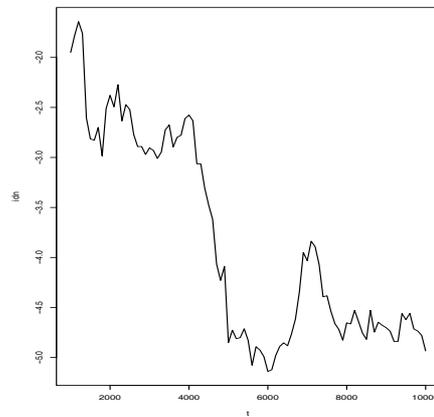


Figura 5.3.3: $b = 0.0, I(0.7)$

$$\text{FIGURA 5.3.} \quad \left[\sum_{\tau > T}^t \widehat{\text{cov}}(y_t, x_{t-\tau}) \right] / \left[\sum_{\tau > T}^t \widehat{\text{cov}}(x_t, x_{t-\tau}) \right]$$

Con el fin de comprobar la condición (5.9) se procede a simular 500 pares de series x_t, y_t de tamaño muestral 10000 con el PGD de las ecuaciones (4.2a)–(4.2b) para $d = 1, 0.7$, $a = 0.5$, con $\alpha = 1$. Como se observa en la Figura 5.3.1, la media de los ratios $\frac{\sum_{\tau > T}^t \text{cov}(y_t, x_{t-\tau})}{\sum_{\tau > T}^t \text{cov}(x_t, x_{t-\tau})}$ se encuentra alrededor del verdadero valor de α si las

variables están cointegradas, en tanto que no es convergente si las series no presentan cointegración, ver Figuras 5.3.2 y 5.3.3.

6. Relación entre cointegración en el dominio temporal y en el dominio de frecuencias

Es de destacar que tanto el Teorema 5.3.1 como la ecuación (5.10) apuntan al mismo resultado espectral, esto es, que bajo cointegración lineal $\lim_{\lambda \rightarrow 0} (S_{y,x}(\lambda)/S_x(\lambda)) = \alpha$, donde α es una constante finita distinta de cero. Si suponemos que las series x_t, y_t tienen los espectros acotados con conta distinta de cero en $\lambda = 0$, este resultado nos sugiere la posibilidad de utilizar el estimador del dominio del tiempo de $\frac{cov(y_t, x_{t-\tau})}{cov(x_t, x_{t-\tau})}$ (estimador VI), cuando t tiende a infinito, o, alternativamente, cuando λ tiende a 0, en lugar del estimador espectral de $\frac{S_{y,x}(\lambda)}{S_x(\lambda)}$ del parámetro de cointegración lineal.

LEMA 6.1. Sean $\sum p_n$ y $\sum q_n$ dos series divergentes de términos positivos. Si $q_n/p_n \rightarrow 0(\infty)$ entonces $\sum q_n$ diverge con menor (mayor) rapidez que $\sum p_n$.

DEMOSTRACIÓN. Ver Knopp (1990, p. 280) □

TEOREMA 5.6.1. Sea $s_n^{(y,x)} = \sum_{\tau=1}^n cov(y_t, x_{t-\tau})/var(x_{t-\tau})$, donde $n = o(t)$. Las series y_t, x_t tienen **memoria larga** y presentan **cointegración lineal** si cuando $n \rightarrow \infty$, se tiene que:

- (1) la sucesión de sumas parciales $s_n^{(y,x)}$ es divergente y,
- (2) el ratio de las sucesiones $s_n^{(y,x)}$ y $s_n^{(x,x)}$ converge a un número finito distinto de cero, α (el parámetro de cointegración).

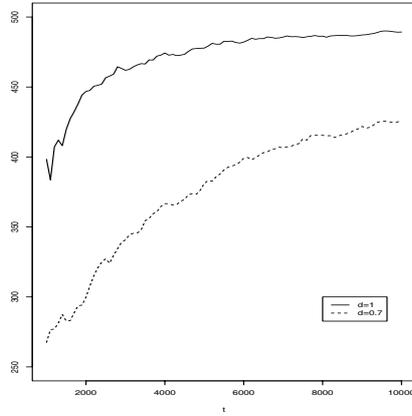
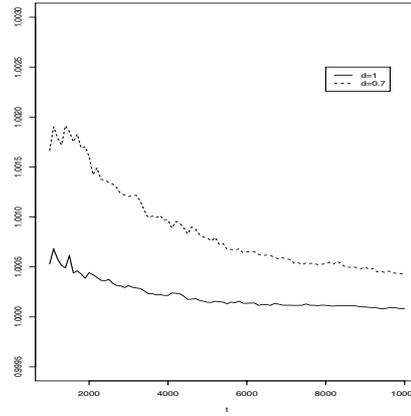
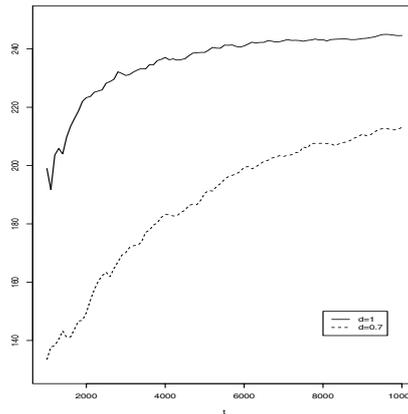
DEMOSTRACIÓN. En Sección 8. □

COROLARIO 5.6.1. Si y_t, x_t son $I(d_y), I(d_x)$, respectivamente, con d_x, d_y en el intervalo $[0, 1]$, entonces una condición necsaria y suficiente para que las series presenten **cointegración fraccional** o que estén linealmente cointegradas con $0 < d_x = d_y = d < 1$ es que, cuando $n \rightarrow \infty$, se tenga:

- (1) La sucesión $s_n^{(y,x)}$ es divergente.
- (2) El ratio de sucesiones $s_n^{(y,x)}$ y $s_n^{(x,x)}$ converge a un valor finito distinto de cero, α .
- (3) Para un $\tau < t$ fijo o para $\tau = o(t)$, $\lim_{t \rightarrow \infty} cov(y_t, x_{t-\tau})/var(x_{t-\tau}) = 0$.

DEMOSTRACIÓN. En Sección 8. □

De nuevo, la Figura 5.4 nos muestra la intuición el Teorema 5.6.1 y del Corolario 5.6.1. Se han simulado 500 pares de series x_t e y_t de tamaño t con el PGD de las

Figura 5.4.1: $s_n^{(y,x)}$, $b = -0.5$ Figura 5.4.2: $s_n^{(y,x)} / s_n^{(x,x)}$, $b = -0.5$ Figura 5.4.3: $s_n^{(y,x)}$, $b = 0.0$ FIGURA 5.4. Simulación del Teorema 5.6.1 y Corolario 5.6.1, con $d = 1$, y $d = 0.7$, $\alpha = 1$, $a = 0.5$, $n = 500$

ecuaciones (4.2a)–(4.2b) para $d = 1$, y 0.7 , $a = 0.5$, con $\alpha = 1$, y se toman las medias de los análogos muestrales de los estadísticos de interés. Los resultados son los esperados en el caso de que las variables estén cointegradas: las sumas parciales divergen y el ratio converge incluso para valores pequeños de τ . En el caso en que las variables no estén cointegradas, las sumas parciales divergen y el ratio es exactamente el valor del parámetro a .

Observaciones:

- Nótese que no hay pérdida de generalidad por restringir d_x, d_y al intervalo unitario, dado que se pueden diferenciar apropiadamente las series para determinar los enteros más próximos a d_x y d_y .
- una vez conocido que $0 < d < 1$, podemos preguntarnos si $d < 1/2$ (en cuyo caso las series son estacionarias), o no. Para ello se puede comprobar si la varianza de x_t es divergente o no.

7. Conclusiones y extensiones

En este Capítulo se ha propuesto una caracterización no paramétrica de la memoria larga y la cointegración lineal, para variables $I(d)$, con órdenes de integración enteros o fraccionales, basada en estadísticos construidos a partir de los ratios de las funciones de covarianzas cruzadas y covarianzas de las series. Se demuestra la relación entre esta definición y el concepto habitual de cointegración.

Este concepto no paramétrico de cointegración tiene una interpretación natural como estimador de variables instrumentales en el que el instrumento es algún retardo de la variable explicativa, $(x_{t-\tau})$. Se realiza el análisis tanto en el dominio temporal como en el de frecuencias. A partir de la comparación de los resultados, se recomienda un estimador alternativo del parámetro de cointegración, α , basado en el ratio de las sumas parciales de las covarianzas cruzadas y las sumas parciales de las autocovarianzas.

Por último, se realizan experimentos de simulación de Monte Carlo y se comprueba el excelente comportamiento en muestras pequeñas de este nuevo estimador respecto de los estimadores habituales basados en el principio de mínimos cuadrados. La derivación formal de las propiedades asintóticas de la distribución de este estimador están fuera del objeto de estudio, aunque se está trabajando en la misma.

8. Demostraciones

Demostración del Teorema 5.3.1

Por simplicidad, vamos a suponer que x_t e y_t tienen media cero. Nos vamos a referir a la función de covarianzas de un par de procesos y_t, x_t , condicional a las condiciones iniciales, tales como $y_0 = x_0 = 0$, mediante $cov(x_t, y_{t-\tau}) = \gamma_{x,y}(\tau, t)$. Obsérvese que si tenemos un proceso no estacionario, la covarianza será una función de t . Análogamente, denotaremos la función de autocovarianzas de un proceso x_t , mediante $cov(x_t, x_{t-\tau}) = \gamma_x(\tau, t)$, condicional a $x_0 = 0$.

Usando la Definición 3.1 y haciendo $\tau = o(t)$, bajo cointegración lineal, $y_t = \alpha x_t + z_t$, existe un número α finito y distinto de cero tal que

$$\lim_{\tau \rightarrow \infty} cov(y_t, x_{t-\tau}) = \alpha \lim_{\tau \rightarrow \infty} cov(x_t, x_{t-\tau}) + \lim_{\tau \rightarrow \infty} cov(z_t, x_{t-\tau}) \quad (8.1)$$

Por tanto

$$\lim_{\tau \rightarrow \infty} \frac{cov(y_t, x_{t-\tau})}{cov(x_t, x_{t-\tau})} = \alpha + \lim_{\tau \rightarrow \infty} \frac{\gamma_{z,x}(\tau, t)}{\gamma_x(\tau, t)}. \quad (8.2)$$

Pero si hay cointegración z_t debe ser una serie $I(d_z)$, con $d_z < d$, en tanto que x_t es $I(d)$, es decir

$$\Delta^d x_t = u_t \quad (8.3)$$

$$\Delta^{d_z} z_t = v_t, \quad (8.4)$$

donde u_t, v_t representan series $I(0)$ con media cero. la inversión del operador retardo nos lleva a una expansión de media móvil infinita. Al truncar dicha expansión para tener en cuenta que la serie sólo está disponible desde $t = 0$ dadas las condiciones iniciales $x_0 = y_0 = z_0 = 0$, se obtiene:

$$z_t = \sum_{k=0}^t \theta_k v_{t-k} \quad (8.5)$$

$$x_{t-\tau} = \sum_{l=0}^t \phi_l u_{t-\tau-l}. \quad (8.6)$$

Así pues:

$$\gamma_{z,x}(\tau, t) = \sum_{k=0}^t \sum_{l=0}^t \phi_l \theta_k \gamma_{u,v}(\tau + l - k). \quad (8.7)$$

Recordando que la función de covarianzas de dos procesos estacionarios u_t y v_t , $\gamma_{u,v}(\tau)$, está relacionada con su *cross-spectrum*, $S_{u,v}(\lambda)$, mediante

$$\gamma_{u,v}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{u,v}(\lambda) \exp(j\lambda\tau) d\lambda, \quad (8.8)$$

con $j^2 = -1$, (ver Granger y Hatanaka (1964) y Priestley (1981)), con un poco de álgebra se obtiene:

$$\gamma_{z,x}(\tau, t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_t^*(\lambda) \Theta_t(\lambda) S_{u,v}(\lambda) \exp(j\lambda\tau) d\lambda, \quad (8.9)$$

donde $\Phi_t(\lambda)$ y $\Theta_t(\lambda)$ se definen como

$$\Phi_t(\lambda) = \sum_{k=0}^t \phi_k \exp(-j\lambda k) \quad (8.10)$$

$$\Theta_t(\lambda) = \sum_{l=0}^t \theta_l \exp(-j\lambda l), \quad (8.11)$$

y $\Phi_t^*(\lambda)$ denota el complejo conjugado de $\Phi_t(\lambda)$. Obsérvese que a medida que t tiende a infinito, estos operadores tienden a

$$\Phi(\lambda) = \lim_{t \rightarrow \infty} \Phi_t(\lambda) = [1 - \exp(-j\lambda)]^{-d} \quad (8.12)$$

$$\Theta(\lambda) = \lim_{t \rightarrow \infty} \Theta_t(\lambda) = [1 - \exp(-j\lambda)]^{-d_z}. \quad (8.13)$$

Además, no es difícil demostrar que cuando λ tiende a cero tienden a

$$\Phi(\lambda) \sim \lambda^{-d} \quad (8.14)$$

$$\Theta(\lambda) \sim \lambda^{-d_z}. \quad (8.15)$$

Consecuentemente, al multiplicar $S_{u,v}(\lambda)$ en la ecuación (8.9), ambos operadores $\Phi(\lambda)$ y $\Theta(\lambda)$ actúan como *filtros pasa-baja*, es decir, enfatizan los valores de $S_{u,v}(\lambda)$ para λ muy próximos a cero, en tanto que difuminan los correspondientes a las frecuencias altas. Puesto que u_t, v_t son procesos $I(0)$, su espectro cruzado, $S_{u,v}(\lambda)$, tendrá una cota finita distinta de cero en $\lambda = 0$, para valores muy grandes de t $\gamma_{z,x}(\tau, t)$ vendrá dominado por el término $\int_{-\pi}^{\pi} \lambda^{-d-d_z} \exp(j\lambda\tau) d\lambda$. Por último, hay que resaltar que la transformada inversa de Fourier de la función $f(\lambda) = \lambda^{-d-d_z}$ es τ^{d+d_z-1} , por lo que tenemos

$$\frac{\gamma_{z,x}(\tau, t)}{\gamma_x(\tau, t)} \sim \tau^{d_z-d}. \quad (8.16)$$

La prueba del Teorema es inmediata puesto que $d_z < d$.

Demostración del Corolario 5.3.1

Puesto que las series y_t y x_t están cointegradas, del Teorema 5.3.1 se deduce que

$$\lim_{\tau \rightarrow \infty} \frac{\text{cov}(y_t, x_{t-\tau})}{\text{cov}(x_t, x_{t-\tau})} = \alpha \quad \text{for } \tau < t \text{ or } \tau = o(t)$$

Lo mismo ocurre si multiplicamos y dividimos por $var(x_{t-\tau})$

$$\lim_{\tau \rightarrow \infty} \frac{\frac{cov(y_t, x_{t-\tau})}{var(x_{t-\tau})}}{\frac{cov(x_t, x_{t-\tau})}{var(x_{t-\tau})}} = \alpha$$

Si $\lim_{\tau \rightarrow \infty} \frac{cov(y_t, x_{t-\tau})}{var(x_{t-\tau})} = \alpha$ se debe cumplir $\lim_{\tau \rightarrow \infty} \frac{cov(x_t, x_{t-\tau})}{var(x_{t-\tau})} = 1$. Dado que $d \in (0, 1]$, ésto sólo es cierto para $d = 1$, puesto que para cualquier otro valor de d , x_t es reversible a la media.

Demostración del Teorema 5.4.1

Gonzalo (1994) demostró que cuando el PGD es $y_t = \alpha x_t + z_t$, donde $z_t = \varphi z_{t-1} + \eta_{1,t}$, y $\Delta x_t = u_{2,t}$, con

$$\begin{pmatrix} \eta_{1,t} \\ u_{2,t} \end{pmatrix} \sim NID \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \theta \sigma_1 \sigma_2 \\ \theta \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix} \right]$$

se cumple

$$\begin{aligned} T(\hat{\alpha}_{ols} - \alpha) &\xrightarrow{d} \left\{ \int_0^1 B_2^2(r) dr \right\}^{-1} \left\{ \left(\frac{\sigma_1}{1-\varphi} \right) [1-\theta^2]^{1/2} \int_0^1 B_2(r) dW_1(r) + \right. \\ &\quad \left. + \left(\frac{1}{1-\varphi} \right) \int_0^1 B_2(r) dB_2(r) + \left(\frac{1}{1-\varphi} \right) \theta \sigma_1 \sigma_2 \right\} \end{aligned}$$

El PGD utilizado en Gonzalo (1994) se puede escribir como (4.1a)–(4.1c) con $d = 1$,

$$y_t = \alpha x_t + z_t$$

$$z_t = (1+b)z_{t-1} + (a-\alpha)\Delta x_t + u_{1,t}$$

$$\Delta x_t = u_{2,t}$$

where

$$\begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} \sim NID \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \right]$$

El resultado se obtiene al observar que $\varphi = (1+b)$ y que $\theta = (a-\alpha)\sigma_2/\sigma_1$.

Demostración del Teorema 5.6.1

En primer lugar, demostremos que las condiciones del teorema son necesarias para que las series tengan memoria larga y estén linealmente cointegradas. Bajo esta última condición, del Teorema 5.3.1 se deduce que existe un número finito y distinto de cero α tal que

$$\text{cov}(y_t, x_{t-\tau}) = \alpha \text{cov}(x_t, x_{t-\tau}) + o(\text{cov}(x_t, x_{t-\tau})), \quad (8.17)$$

Por otra parte, puesto que x_t es de memoria larga, $s_n^{(x,x)}$ es divergente. Si tomamos sumas parciales en la expresión anterior, es obvio que $s_n^{(y,x)}$ es también divergente, en tanto que el ratio $(s_n^{(y,x)})/s_n^{(x,x)}$ converge a α cuando n tiende a infinito.

Demostración del Corolario 5.6.1

Los procesos con integración fraccional con parámetro de memoria larga $d < 1$ satisfacen la condición 1, en tanto que los Teoremas 5.3.1 y 5.4.1 implican las condiciones 2 y 3 para procesos linealmente cointegrados.

9. Tablas

TABLA 5.1. Estimadores MCO y VI de α en la ecuación $y_t = \alpha x_t + z_t$, donde el PGD es $\Delta y_t = a\Delta x_t - 0.2(y_{t-1} - \alpha x_{t-1}) + u_{1t}$, $\Delta x_t = u_{2t}$, $\alpha = 1$, $u_{1t} \sim N(0, 1)$, $u_{2t} \sim N(0, s^2)$. $N = 100$, $T = 20000$ réplicas.

$\tau = 0$ (MCO)					
$a - \alpha$	Sesgo en media	Sesgo en mediana	ECM	RIQ	$\Pr(\hat{\alpha} - \alpha \leq 0.05)$
-2.0	-0.190	-0.139	0.664E-01	0.174	0.139
-1.0	-0.948E-01	-0.691E-01	0.168E-01	0.882E-01	0.351
-0.5	-0.474E-01	-0.346E-01	0.433E-02	0.455E-01	0.659
0.0	-0.920E-04	-0.184E-03	0.190E-03	0.134E-01	0.994
$\tau = 1$					
-2.0	-0.129	-0.102	0.322E-01	0.128	0.216
-1.0	-0.647E-01	-0.512E-01	0.820E-02	0.656E-01	0.479
-0.5	-0.324E-01	-0.257E-01	0.221E-02	0.350E-01	0.774
0.0	-0.106E-03	-0.185E-03	0.209E-03	0.136E-01	0.991
$\tau = 2$					
-2.0	-0.782E-01	-0.706E-01	0.158E-01	0.102	0.320
-1.0	-0.392E-01	-0.353E-01	0.413E-02	0.528E-01	0.615
-0.5	-0.196E-01	-0.177E-01	0.120E-02	0.298E-01	0.872
0.0	-0.109E-03	-0.159E-03	0.232E-03	0.138E-01	0.988
$\tau = 3$					
-2.0	-0.346E-01	-0.398E-01	0.129E-01	0.103	0.407
-1.0	-0.173E-01	-0.208E-01	0.340E-02	0.538E-01	0.692
-0.5	-0.872E-02	-0.109E-01	0.104E-02	0.301E-01	0.904
0.0	-0.107E-03	-0.198E-03	0.261E-03	0.140E-01	0.984
$\tau = 4$					
-2.0	0.408E-02	-0.184E-01	0.243E-01	0.124	0.393
-1.0	0.200E-02	-0.993E-02	0.632E-02	0.630E-01	0.675
-0.5	0.963E-03	-0.581E-02	0.182E-02	0.342E-01	0.885
0.0	-0.777E-04	-0.185E-03	0.308E-03	0.143E-01	0.981
$\tau = 5$					
-2.0	0.436E-01	-0.437E-02	0.374	0.147	0.369
-1.0	0.218E-01	-0.228E-02	0.992E-01	0.745E-01	0.643
-0.5	0.110E-01	-0.163E-02	0.279E-01	0.389E-01	0.854
0.0	0.110E-03	-0.195E-03	0.746E-03	0.146E-01	0.976

TABLA 5.2. Estimadores MCO y VI de α en la ecuación $y_t = \alpha x_t + z_t$, donde el PGD es $\Delta y_t = a\Delta x_t - 0.5(y_{t-1} - \alpha x_{t-1}) + u_{1t}$, $\Delta x_t = u_{2t}$, $\alpha = 1$, $u_{1t} \sim N(0, 1)$, $u_{2t} \sim N(0, s^2)$. $N = 100$, $T = 20000$ réplicas.

$\tau = 0$ (MCO)					
$a - \alpha$	Sesgo en media	Sesgo en mediana	ECM	RIQ	$\Pr(\hat{\alpha} - \alpha \leq 0.05)$
-2.0	-0.819E-01	-0.561E-01	0.139E-01	0.754E-01	0.449
-1.0	-0.410E-01	-0.279E-01	0.350E-02	0.381E-01	0.729
-0.5	-0.205E-01	-0.141E-01	0.902E-03	0.197E-01	0.911
0.0	-0.397E-04	-0.679E-04	0.352E-04	0.559E-02	1.00
$\tau = 1$					
-2.0	-0.149E-01	-0.149E-01	0.193E-02	0.418E-01	0.799
-1.0	-0.746E-02	-0.767E-02	0.510E-03	0.217E-01	0.961
-0.5	-0.376E-02	-0.408E-02	0.156E-03	0.122E-01	0.997
0.0	-0.521E-04	-0.566E-04	0.390E-04	0.569E-02	1.000
$\tau = 2$					
-2.0	0.211E-01	0.461E-02	0.575E-02	0.645E-01	0.711
-1.0	0.105E-01	0.214E-02	0.147E-02	0.324E-01	0.889
-0.5	0.523E-02	0.900E-03	0.398E-03	0.169E-01	0.969
0.0	-0.504E-04	-0.681E-04	0.439E-04	0.575E-02	1.000
$\tau = 3$					
-2.0	0.408E-01	0.130E-01	0.122E-01	0.794E-01	0.654
-1.0	0.204E-01	0.638E-02	0.309E-02	0.398E-01	0.830
-0.5	0.102E-01	0.311E-02	0.809E-03	0.205E-01	0.936
0.0	-0.418E-04	-0.695E-04	0.507E-04	0.585E-02	0.999
$\tau = 4$					
-2.0	0.530E-01	0.172E-01	0.220E-01	0.882E-01	0.628
-1.0	0.265E-01	0.853E-02	0.556E-02	0.442E-01	0.803
-0.5	0.132E-01	0.418E-02	0.144E-02	0.228E-01	0.915
0.0	-0.140E-04	-0.973E-04	0.615E-04	0.596E-02	0.998
$\tau = 5$					
-2.0	0.647E-01	0.193E-01	0.234	0.932E-01	0.616
-1.0	0.324E-01	0.957E-02	0.609E-01	0.468E-01	0.788
-0.5	0.163E-01	0.464E-02	0.165E-01	0.240E-01	0.904
0.0	0.105E-03	-0.789E-04	0.183E-03	0.603E-02	0.997

TABLA 5.3. Estimadores MCO y VI de α en la ecuación $y_t = \alpha x_t + z_t$, donde el PGD es $\Delta y_t = a\Delta x_t - (y_{t-1} - \alpha x_{t-1}) + u_{1t}$, $\Delta x_t = u_{2t}$, $\alpha = 1$, $u_{1t} \sim N(0, 1)$, $u_{2t} \sim N(0, s^2)$. $N = 100$, $T = 20000$ réplicas.

$\tau = 0$ (MCO)					
$a - \alpha$	Sesgo en media	Sesgo en mediana	ECM	RIQ	$\Pr(\hat{\alpha} - \alpha \leq 0.05)$
-2.0	-0.422E-01	-0.282E-01	0.391E-02	0.390E-01	0.727
-1.0	-0.211E-01	-0.141E-01	0.984E-03	0.196E-01	0.906
-0.5	-0.106E-01	-0.705E-02	0.253E-03	0.100E-01	0.984
0.0	-0.195E-04	-0.294E-04	0.941E-05	0.284E-02	1.00
$\tau = 1$					
-2.0	0.277E-01	0.112E-01	0.387E-02	0.461E-01	0.789
-1.0	0.138E-01	0.551E-02	0.973E-03	0.230E-01	0.914
-0.5	0.691E-02	0.270E-02	0.250E-03	0.117E-01	0.981
0.0	-0.322E-04	-0.394E-04	0.106E-04	0.289E-02	1.00
$\tau = 2$					
-2.0	0.283E-01	0.109E-01	0.430E-02	0.463E-01	0.786
-1.0	0.142E-01	0.540E-02	0.108E-02	0.233E-01	0.911
-0.5	0.707E-02	0.264E-02	0.279E-03	0.119E-01	0.979
0.0	-0.226E-04	-0.347E-04	0.120E-04	0.292E-02	1.00
$\tau = 3$					
-2.0	0.290E-01	0.107E-01	0.483E-02	0.470E-01	0.785
-1.0	0.145E-01	0.535E-02	0.122E-02	0.234E-01	0.908
-0.5	0.723E-02	0.261E-02	0.315E-03	0.120E-01	0.976
0.0	-0.146E-04	-0.431E-04	0.141E-04	0.300E-02	1.000
$\tau = 4$					
-2.0	0.303E-01	0.105E-01	0.656E-02	0.473E-01	0.784
-1.0	0.151E-01	0.524E-02	0.166E-02	0.238E-01	0.905
-0.5	0.757E-02	0.255E-02	0.434E-03	0.122E-01	0.973
0.0	0.649E-05	-0.482E-04	0.171E-04	0.302E-02	1.000
$\tau = 5$					
-2.0	0.330E-01	0.103E-01	0.513E-01	0.480E-01	0.781
-1.0	0.166E-01	0.504E-02	0.133E-01	0.241E-01	0.902
-0.5	0.831E-02	0.248E-02	0.362E-02	0.124E-01	0.969
0.0	0.726E-04	-0.323E-04	0.473E-04	0.306E-02	1.000

TABLA 5.4. Estimadores MCO y VI de α en la ecuación $y_t = \alpha x_t + z_t$, donde el PGD es $\Delta y_t = \alpha \Delta x_t + (a-1)\Delta^d x_t - 0.5(y_{t-1} - \alpha x_{t-1}) + u_{1t}$, $\Delta x_t = u_{2t}$, $\alpha = 1, d = 0.7$, $u_{1t} \sim N(0, 1)$, $u_{2t} \sim N(0, s^2)$. $N = 100$, $T = 20000$ réplicas.

$\tau = 0$ (MCO)					
$a - \alpha$	Sesgo en media	Sesgo en mediana	ECM	RIQ	$\Pr(\hat{\alpha} - \alpha \leq 0.05)$
-2.0	-0.526	-0.463	0.365	0.430	0.001
-1.0	-0.263	-0.232	0.914E-01	0.215	0.009
-0.5	-0.131	-0.116	0.230E-01	0.108	0.116
0.0	0.480E-05	-0.134E-04	0.182E-03	0.157E-01	0.998
$\tau = 1$					
-2.0	-0.250	-0.227	0.829E-01	0.195	0.038
-1.0	-0.125	-0.113	0.209E-01	0.980E-01	0.128
-0.5	-0.625E-01	-0.560E-01	0.544E-02	0.504E-01	0.434
0.0	-0.609E-04	-0.907E-04	0.272E-03	0.174E-01	0.989
$\tau = 2$					
-2.0	-0.648E-01	-0.852E-01	0.263E-01	0.132	0.220
-1.0	-0.324E-01	-0.427E-01	0.678E-02	0.678E-01	0.468
-0.5	-0.163E-01	-0.214E-01	0.194E-02	0.375E-01	0.809
0.0	-0.118E-03	0.306E-04	0.373E-03	0.186E-01	0.975
$\tau = 3$					
-2.0	0.558E-01	-0.154E-01	0.103	0.192	0.282
-1.0	0.278E-01	-0.764E-02	0.259E-01	0.971E-01	0.528
-0.5	0.138E-01	-0.413E-02	0.669E-02	0.518E-01	0.791
0.0	-0.177E-03	-0.406E-05	0.541E-03	0.195E-01	0.962
$\tau = 4$					
-2.0	0.184	0.202E-01	45.2	0.247	0.268
-1.0	0.921E-01	0.947E-02	11.7	0.125	0.493
-0.5	0.464E-01	0.475E-02	3.12	0.650E-01	0.728
0.0	0.726E-03	0.880E-04	0.141E-01	0.206E-01	0.948
$\tau = 5$					
-2.0	0.253	0.387E-01	38.2	0.285	0.253
-1.0	0.125	0.192E-01	8.48	0.144	0.466
-0.5	0.613E-01	0.962E-02	1.64	0.744E-01	0.691
0.0	-0.255E-02	0.107E-04	0.141	0.216E-01	0.931

TABLA 5.5. Estimadores MCO y VI de α en la ecuación $y_t = \alpha x_t + z_t$, donde el PGD es $\Delta y_t = \alpha \Delta x_t + (a-1)\Delta^d x_t - 0.5(y_{t-1} - \alpha x_{t-1}) + u_{1t}$, $\Delta x_t = u_{2t}$, $\alpha = 1, d = 1.3$, $u_{1t} \sim N(0, 1)$, $u_{2t} \sim N(0, s^2)$. $N = 100$, $T = 20000$ réplicas.

$\tau = 0$ (MCO)					
$a - \alpha$	Sesgo en media	Sesgo en mediana	ECM	RIQ	$\Pr(\hat{\alpha} - \alpha \leq 0.05)$
-2.0	-0.120E-02	-0.146E-02	0.283E-03	0.156E-01	0.983
-1.0	-0.605E-03	-0.797E-03	0.753E-04	0.788E-02	0.997
-0.5	-0.308E-03	-0.421E-03	0.227E-04	0.414E-02	1.000
0.0	-0.119E-04	0.394E-05	0.434E-05	0.151E-02	1.000
$\tau = 1$					
-2.0	0.791E-02	0.216E-02	0.620E-03	0.207E-01	0.947
-1.0	0.395E-02	0.111E-02	0.157E-03	0.104E-01	0.991
-0.5	0.197E-02	0.532E-03	0.421E-04	0.525E-02	0.999
0.0	-0.158E-04	-0.268E-05	0.446E-05	0.151E-02	1.000
$\tau = 2$					
-2.0	0.124E-01	0.365E-02	0.111E-02	0.237E-01	0.916
-1.0	0.619E-02	0.179E-02	0.278E-03	0.118E-01	0.977
-0.5	0.309E-02	0.887E-03	0.719E-04	0.604E-02	0.997
0.0	-0.147E-04	-0.459E-05	0.461E-05	0.152E-02	1.000
$\tau = 3$					
-2.0	0.146E-01	0.422E-02	0.144E-02	0.250E-01	0.904
-1.0	0.727E-02	0.208E-02	0.361E-03	0.124E-01	0.970
-0.5	0.363E-02	0.103E-02	0.925E-04	0.642E-02	0.994
0.0	-0.170E-04	-0.492E-05	0.480E-05	0.150E-02	1.000
$\tau = 4$					
-2.0	0.157E-01	0.434E-02	0.166E-02	0.258E-01	0.896
-1.0	0.782E-02	0.217E-02	0.415E-03	0.127E-01	0.968
-0.5	0.390E-02	0.108E-02	0.106E-03	0.656E-02	0.993
0.0	-0.180E-04	0.381E-05	0.521E-05	0.151E-02	1.000
$\tau = 5$					
-2.0	0.161E-01	0.437E-02	0.177E-02	0.262E-01	0.893
-1.0	0.805E-02	0.219E-02	0.445E-03	0.130E-01	0.965
-0.5	0.402E-02	0.107E-02	0.115E-03	0.662E-02	0.991
0.0	-0.118E-04	0.745E-05	0.559E-05	0.153E-02	1.000

Bibliografía

- Andrés, J., A. Escribano, C. Molinas y D. Taguas (1990), *La inversión en España: Econometría con restricciones de equilibrio*, A. Bosch.
- Andrews, D. (1993), 'Tests of parameter instability and structural change with unknown break point', *Econometrica* **61**, 821–856.
- Andrews, D., W. Lee y W. Ploberger (1996), 'Optimal changepoints tests for normal linear regression', *Journal of Econometrics* **70**, 9–38.
- Aparicio, F.M. y A. Escribano (1999), 'Information-theoretic analysis of serial dependence and cointegration', *Studies in Nonlinear Dynamics and Econometrics* **3**(3).
- Arranz, M. y A. Escribano (1998), Detrending procedures and cointegration testing: ECM tests under structural breaks, Working Paper 98–15, Universidad Carlos III de Madrid.
- Arranz, M. y A. Escribano (2000), 'Cointegration testing under structural breaks: A robust extended error correction model', *Oxford Bulletin of Economics and Statistics* **61**, 23–52.
- Bai, J. (1997), 'Estimating multiple breaks one at a time', *Econometric Theory* **13**, 315–352.
- Bai, J. y P. Perron (1998), 'Estimating and testing linear models with multiple structural changes', *Econometrica* **66**, 47–78.
- Balke, N. S. y T. B. Fomby (1997), 'Threshold cointegration', *International economic review* **38**(3), 627–645.
- Banerjee, A., J. Dolado, D. Hendry y G. Smith (1986), 'Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence', *Oxford Bulletin of Economics and Statistics* **48**, 853–877.
- Banerjee, A., J. Dolado y F. Mestre (1998), 'ECM test for cointegration in a single equation framework', *Journal of Time Series Analysis* **19**, 267–283.
- Banerjee, A., J. Dolado, J.W. Galbraith y D.F. Hendry (1993), *Cointegration, error-correction, and the econometric analysis of non-stationary data*, Oxford University Press.
- Banerjee, A., R. Lumsdaine y J. Stock (1992), 'Recursive and sequential tests of unit root and trend break hypothesis: Theory and international evidence', *Journal of Business and Economic Statistics* **10**, 271–287.
- Banerjee, A., S. Lazarova y G. Urga (1998), Bootstrapping sequential tests for multiple structural breaks, Discussion Paper 17–98, Center for Economic Forecasting, London Business School.

- Baxter, M. y R. King (1995), Measuring business cycles: Approximate band-pass filters for economic time series, Working Paper 5022, NBER.
- Baxter, M. y R. King (1999), 'Measuring business cycles: Approximate band-pass filters for economic time series', *Review of Economics and Statistics* **81**, 575–593.
- Beran, J., R. Sherman, M.S. Taqqu y W. Willinger (1991), Variable-bit-rate video traffic and long range dependence. Working paper (Dept. of Mathematics of Boston University).
- Beran, R. (1987), 'Prepivoting to reduce level error of confidence sets', *Biometrika* **74**, 457–468.
- Beran, R. (1988), 'Prepivoting test statistics: a bootstrap view of asymptotic refinements', *Journal of the American Statistical Association* **83**, 681–697.
- Berkowitz, J. y L. Kilian (2000), 'Recent developments in bootstrapping time series', *Econometric Reviews* **19**, 1–48.
- Box, G.E.P. y G.M. Jenkins (1970), *Time series analysis: forecasting and control*, Holden-Day, San Francisco.
- Buhlmann, P. (1997), 'Sieve bootstrap for time series', *Bernoulli* **3**, 123–148.
- Buhlmann, P. (1998), 'Sieve bootstrap for smoothing in non-stationary time series', *The Annals of Statistics* **48**, 48–83.
- Buhlmann, P. (1999), Bootstrap for time series, Research report, ETH, Zurich.
- Buhlmann, P. y H. Kunsch (1999), 'Block length selection in the bootstrap for time series', *Computational Statistics and Data Analysis* **31**, 295–310.
- Burns, A. y W. Mitchell (1946), *Measuring Business Cycles*, NBER.
- Campos, J., N. Ericsson y D. Hendry (1996), 'Cointegration tests in the presence of structural breaks', *Journal of Econometrics* **70**, 187–220.
- Canova, F. (1998), 'Detrending and business cycle facts', *Journal of Monetary Economics* **41**, 475–512.
- Carlstein, E. (1986), 'The use of subseries methods for estimating the variance of general statistics from a stationary time series', *The Annals of Statistics* **14**, 1171–1179.
- Carlstein, E., K. Do, P. Hall, T. Hesterberg y H. Kunsch (1998), 'Matched-block bootstrap for dependent data', *Bernoulli* **4**, 305–328.
- Chang, I., G. Tsiao y C. Chen (1988), 'Estimation of time series parameters in the presence of outliers', *Technometrics* **30**, 193–204.
- Chen, C. y L. Liu (1993), 'Joint estimation of model parameters and outlier effects in time series', *Journal of the American Statistical Association* **88**, 284–297.
- Christiano, L. (1992), 'Searching for a break in GNP', *Journal of Business and Economic Statistics* **10**, 237–250.
- Christiano, L. y T. Fitzgerald (1999), The band pass filter, Working Paper 7257, NBER.
- Clements, M. y D. Hendry (1999), *Forecasting Non-stationary Economic Time Series*, MIT Press, chapter 9.

- Cogley, F. y J. Nason (1995), 'Effects of the Hodrick-Prescott filter on trend and difference stationary time series: Implications for business cycle research', *Journal of Economic Dynamics and Control* **19**, 253–278.
- Datta, S. (1996), 'On asymptotic properties of bootstrap for AR(1) processes', *Journal of Statistical Planning and Inference* **53**, 361–374.
- Davidson, A. y B. Hinkley (1997), *Bootstrap Methods and Their Applications*, Cambridge University Press.
- Davidson, A. y P. Hall (1993), 'On studentizing and blocking methods for implementing the bootstrap with dependent data', *Australian Journal of Statistics* **35**, 212–224.
- Davidson, R. y J. MacKinnon (1998), 'Graphical methods for investigating the size and power of hypothesis tests', *The Manchester School* **66**, 1–26.
- Davidson, R. y J. MacKinnon (2000), 'Bootstrap tests: How many bootstraps?', *Econometric Reviews* **19**, 55–68.
- Dickey, D. y W. Fuller (1979), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* **74**, 427–431.
- Dolado, J. y F. Marmol (1998), FM-OLS estimation of cointegrating relationships among nonstationary fractionally integrated processes, Working Paper 98–41, Dpt. of Statistics and Econometrics, Universidad Carlos III de Madrid.
- Dolado, J. y H. Lutkepohl (1996), 'Making Wald tests work for cointegrated VAR systems', *Econometric Reviews* **15**, 369–386.
- Efron, B. (1987), 'Better bootstrap confidence intervals', *Journal of the American Statistical Association* **82**, 171–185.
- Efron, B. y R. Tibshirani (1993), *Introduction to the Bootstrap*, Chapman and Hall.
- Ehglén, J. (1998), 'Distortionary effects of the optimal Hodrick–Prescott filter', *Economics Letters* **61**, 345–349.
- Engle, R. F. y C. W. J. Granger (1987), 'Cointegration and error-correction: Representation, estimation, and testing', *Econometrica* **55**, 251–276.
- Escribano, A. (1986), Nonlinear error-correction. The case of money demand in the U.K. (1878-1970). Chapter IV, PhD thesis, Dept. of Economics of the University of California at San Diego, La Jolla, USA.
- Escribano, A. (1987a), Co-integration, time co-trends and error-correction systems: an alternative approach, Discussion Paper 8715, CORE.
- Escribano, A. (1987b), Error-correcting systems: nonlinear adjustment to linear long-run relationships. CORE Discussion paper 8730, C.O.R.E.
- Escribano, A. y C. W. J. Granger (1998), 'Investigating the relationships between gold and silver prices', *Journal of forecasting* **17**, 81–107.
- Escribano, A. y G. A. Pfann (1998), 'Nonlinear error correction, asymmetric adjustment and cointegration', *Economic modelling* **15**, 197–216.

- Franses, P. y A. Lucas (1997a), Outlier robust cointegration analysis of Dutch interest rates, *in* 'Proceedings of the 1996 Meeting of the American Statistical Association', ASA, pp. 106–109.
- Franses, P. y A. Lucas (1997b), Outlier robust cointegration tests, VU Research Memorandum 97–45, Vrije Universiteit Amsterdam.
- Franses, P. y A. Lucas (1998), 'Outlier detection in cointegration analysis', *Journal of Business and Economic Statistics* **16**, 459–468.
- Franses, P. y N. Haldrup (1994), 'The effects of additive outliers on unit roots and cointegration', *Journal of Business and Economic Statistics* **12**, 471–478.
- Ghysels, E. y P. Perron (1993), 'The effect of seasonal adjustment filters on tests for a unit root', *Journal of Econometrics* **55**, 57–98.
- Giersbergen, N. y J. Kiviet (1993), A Monte Carlo comparison of asymptotic and various nonparametric inference procedures in first-order dynamic models, Discussion Paper TI 93–187, Tinbergen Institute.
- Giersbergen, N. y J. Kiviet (1994), How to implement bootstrap hypothesis testing in static and dynamic regression models, Discussion Paper TI 94–130, Tinbergen Institute.
- Giersbergen, N. y J. Kiviet (1996), 'Bootstrapping a stable AD model: Weak vs. strong exogeneity', *Oxford Bulletin of Economics and Statistics* **48**, 631–656.
- Gómez, V. y A. Maravall (1996), Programs TRAMO and SEATS, Working Paper 9628, Bank of Spain.
- Gonzalo, J. (1994), 'Five alternative methods of estimating long-run relationships', *Journal of Econometrics* **60**, 203–233.
- Granger, C. (1981), 'Some properties of time series and their use in econometric model specification', *Journal of Econometrics* **16**, 121–130.
- Granger, C. W. J. y M. Hatanaka (1964), *Spectral analysis of economic time series*, Princeton University Press, New Jersey.
- Granger, C. W. J. y P. Newbold (1974), 'Spurious regressions in econometrics', *Journal of Econometrics* **2**, 111–120.
- Granger, C. W. J. y R. Joyeux (1980), 'An introduction to long-memory time series models and fractional differencing', *Journal of Time Series Analysis* **1**(1), 15–29.
- Guay, A. y P. St-Amant (1997), Do the Hodrick-Prescott and Baxter-King filters provide a good approximation of business cycles?, Cahiers de Recherche 53, CREFE.
- Hall, P. (1988), 'Theoretical comparison of bootstrap confidence intervals', *The Annals of Statistics* **16**, 927–953.
- Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*, Springer-Verlag.
- Hall, P. (1994), Methodology and theory for the bootstrap, *in* R. Engle y D. McFadden, eds, 'Handbook of Econometrics', Vol. 4, Elsevier Science, chapter 39.

- Hall, P. y S. Wilson (1991), 'Two guidelines for bootstrap hypothesis testing', *Biometrika* **47**, 757–762.
- Hansen, B. (1995), 'Rethinking the univariate approach to unit root testing: Using covariates to increase power', *Econometric Theory* **11**, 1148–1172.
- Harvey, A. C. y A. Jaeger (1993), 'Detrending, business cycles facts and the business cycle', *Journal of Applied Econometrics* **8**, 231–47.
- Harvey, A.C. (1989), *Forecasting structural time series models and the Kalman filter*, Cambridge University Press, Cambridge.
- Hendry, D. y A. Neale (1990), The impact of structural breaks on unit root tests, in P.Hackl y A. Westlung, eds, 'Economic Structural Change: Analysis and Forecasting', Berlin: IIASA and Springer-Verlag, pp. 95–119.
- Hodrick, R. y E. Prescott (1980), Post-war U.S. business cycle: An empirical investigation, Technical report, Carnegie-Mellon University.
- Hodrick, R. y E. Prescott (1997), 'Post-war U.S. business cycle: An empirical investigation', *Journal of Money, Credit and Banking* **29**, 1–16.
- Horowitz, J. (2000), The bootstrap, in R.Engle y D.McFadden, eds, 'Handbook of Econometrics', Vol. 5, North-Holland. To appear.
- Hosking, J.R.M. (1981), 'Fractional differencing', *Biometrika* **68**, 165–176.
- Johansen, S. (1988), 'Statistical analysis of cointegration vectors', *Journal of economic dynamics and control* **12**, 231–254.
- Kaiser, R. (1998), Detection and estimation of structural changes and outliers, Working Paper 98–23, Universidad Carlos III de Madrid.
- Kaiser, R. y A. Maravall (1999), 'Estimation of the business cycle: A modified Hodrick-Prescott filter', *Spanish Economic Review* **1**, 175–206.
- King, R. y S. Rebelo (1993), 'Low frequency filtering and real business cycles', *Journal of Economic Dynamics and Control* **17**, 207–31.
- Kiviet, J. (1984), Bootstrap inference in lagged-dependent variable models, Working paper, University of Amsterdam.
- Knopp, K. (1990), *Theory and application of infinite series*, Dover, N.Y.
- Kremers, J., N. Ericsson y J. Dolado (1992), 'The power of cointegration tests', *Oxford Bulletin of Economics and Statistics* **54**, 349–67.
- Kunsch, P. (1989), 'The jackknife and the bootstrap for general stationary observations', *The Annals of Statistics* **17**, 1217–1241.
- Kydland, F. y E. C. Prescott (1990), 'Business cycles: Real facts and a monetary myth', *Federal Reserve Bank of Minneapolis Quarterly Review* **14**, 3–18.
- Lahiri, S. (1999), 'Theoretical comparison of block bootstrap methods', *The Annals of Statistics* **37**, 386–404.

- Leland, W.E., M.S. Taquq, W. Willinger y D. Wilson (1994), 'On the self-similar nature of ethernet traffic', *IEEE/ACM Transactions on Networking* **2**, 1–15.
- Leybourne, S., T. Mills y P. Newbold (1998), 'Spurious rejections by Dickey–Fuller tests in the presence of a break under the null', *Journal of Econometrics* **87**, 191–203.
- Li, H. y G. Maddala (1996), 'Bootstrapping time series models', *Econometric Reviews* **15**, 115–158.
- Li, H. y G. S. Maddala (1997), 'Bootstrapping cointegrating regressions', *Journal of Econometrics* **80**, 297–318.
- Lucas, A. (1995a), 'An outlier robust unit root tests with an application to the extended Nelson–Plosser data', *Journal of Econometrics* **66**, 153–173.
- Lucas, A. (1995b), 'Unit roots tests based on M estimators', *Econometric Theory* **11**, 331–346.
- Maddala, G. y Y. Yin (1996), Outliers, unit roots and robust estimation of non–stationary time series, in G.Maddala y C.Rao, eds, 'Handbook of Statistics', Vol. 15, North–Holland.
- Mandelbrot, B. (1965), 'Self-similar error clusters in communication systems and the concept of conditional stationarity', *IEEE Transactions on communication technology* **13**, 71–72.
- Mandelbrot, B. (1967), 'Some noises with 1/f spectrum: a bridge between direct current and white noise', *IEEE Transactions on information theory* **13**, 289–298.
- Mantalos, P. y G. Shukur (1998), 'Size and power of the error correction model cointegration test. A bootstrap approach', *Oxford Bulletin of Economics and Statistics* **60**, 249–255.
- Martin, R. y V. Yohai (1986), 'Influence functionals for time series', *The Annals of Statistics* **14**, 781–818.
- Ng, S. y P. Perron (1998), Lag length selection and the construction of unit root tests with good size and power, Working Paper 319, Dpt. of Economics, Boston College.
- Park, D. y T. Willemain (1999), 'The threshold bootstrap and the threshold jackknife', *Computational Statistics and Data Analysis* **31**, 187–202.
- Park, J. (1992), 'Canonical cointegrating regression', *Econometrica* **60**, 119–43.
- Park, J. y P. Phillips (1988), 'Statistical inference in regressions with integrated processes: Part I', *Econometric Theory* **4**, 458–497.
- Park, J. y P. Phillips (1989), 'Statistical inference in regressions with integrated processes: Part II', *Econometric Theory* **5**, 95–131.
- Perron, P. (1989), 'The great crash, the oil price shock and the unit root hypothesis', *Econometrica* **57**, 1361–1401.
- Perron, P. y S. Ng (1996), 'Useful modifications to some unit root tests with dependent errors and their local asymptotic properties', *Review of Economic Studies* **63**, 435–465.
- Perron, P. y T. Vogelsang (1992), 'Nonstationarity and level shifts with an application to purchasing power parity', *Journal of Business and Economic Statistics* **10**, 301–320.
- Phillips, P. (1987a), 'Time series regression with a unit root', *Econometrica* **58**, 277–301.

- Phillips, P. (1987*b*), 'Towards a unified asymptotic theory for autoregression', *Biometrika* **74**, 535–47.
- Phillips, P. y B. Hansen (1990), 'Statistical inference in instrumental variables regression with I(1) processes', *Review of Economic Studies* **57**, 99–125.
- Politis, D. y J. Romano (1994), 'The stationary bootstrap', *Journal of the American Statistical Association* **89**, 1303–1313.
- Priestley, M.B. (1981), *Spectral Analysis and Time Series*, Academic Press.
- Rappoport, P. y L. Reichlin (1989), 'Segmented trends and nonstationary time series', *Economic Journal* **99**, 168–177.
- Shao, J. y D. Tu (1995), *The Jackknife and the Bootstrap*, Springer–Verlag.
- Shin, D., S. Sarkar y J. Lee (1996), 'Unit root tests for time series with outliers', *Stochastic and Probability Letters* **30**, 189–197.
- Stock, J. (1987), 'Asymptotic properties of least squares estimators of cointegrating regressions', *Econometrica* **51**, 153–174.
- Stock, J. (1994), Unit roots, structural breaks and trends, in R. F.Engle y D. L.McFadden, eds, 'Handbook of Econometrics', Vol. 4, North–Holland, chapter 46, pp. 2270–2381.
- Toda, H. y T. Yamamoto (1995), 'Statistical inference in vector autorregression with possibly integrated processes', *Journal of Econometrics* **66**, 225–250.
- Tsay, R. (1986), 'Time series model specification in the presence of outliers', *Journal of the American Statistical Association* **81**, 132–141.
- Vogelsang, T. (1997), 'Wald–type tests for detecting breaks in the trend of a dynamic time series', *Econometric Theory* **13**, 818–849.
- Vogelsang, T. (1999), 'Two simple procedures for testing for a unit root when there are additive outliers', *Journal of Time Series Analysis* **20**, 173–192.
- Wen, Y. y B. Zeng (1999), 'A simple nonlinear filter for economic time series analysis', *Economics Letters* **64**, 151–160.
- White, H. (1984), *Asymptotic theory for econometricians*, Academic Press, San Diego.
- Willinger, W., M.S. Taqqu, W.E. Leland y D.V. Wilson (1995), 'Self-similarity in high-speed packet traffic: analysis and modelling of ethernet traffic measurements', *Statistical Science* **10**(1), 67–85.
- Yin, Y. y G. Maddala (1997), The effects of different outliers on unit root tests, in T.Fornby y R.Hill, eds, 'Advances in Econometrics', Vol. 13, JAI Press.
- Zivot, E. (2000), 'The power of single equation tests for cointegration when the cointegration vector is prespecified', *Econometric Theory* **16**, 407–439.
- Zivot, E. y D. Andrews (1992), 'Further evidence on the great crash, the oil price shock and the unit root hypothesis', *Journal of Business and Economic Statistics* **10**, 251–270.

Bibliografía

- Andrés, J., A. Escribano, C. Molinas y D. Taguas (1990), *La inversión en España: Econometría con restricciones de equilibrio*, A. Bosch.
- Andrews, D. (1993), 'Tests of parameter instability and structural change with unknown break point', *Econometrica* **61**, 821–856.
- Andrews, D., W. Lee y W. Ploberger (1996), 'Optimal changepoints tests for normal linear regression', *Journal of Econometrics* **70**, 9–38.
- Aparicio, F.M. y A. Escribano (1999), 'Information-theoretic analysis of serial dependence and cointegration', *Studies in Nonlinear Dynamics and Econometrics* **3**(3).
- Arranz, M. y A. Escribano (1998), Detrending procedures and cointegration testing: ECM tests under structural breaks, Working Paper 98–15, Universidad Carlos III de Madrid.
- Arranz, M. y A. Escribano (2000), 'Cointegration testing under structural breaks: A robust extended error correction model', *Oxford Bulletin of Economics and Statistics* **61**, 23–52.
- Bai, J. (1997), 'Estimating multiple breaks one at a time', *Econometric Theory* **13**, 315–352.
- Bai, J. y P. Perron (1998), 'Estimating and testing linear models with multiple structural changes', *Econometrica* **66**, 47–78.
- Balke, N. S. y T. B. Fomby (1997), 'Threshold cointegration', *International economic review* **38**(3), 627–645.
- Banerjee, A., J. Dolado, D. Hendry y G. Smith (1986), 'Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence', *Oxford Bulletin of Economics and Statistics* **48**, 853–877.
- Banerjee, A., J. Dolado y F. Mestre (1998), 'ECM test for cointegration in a single equation framework', *Journal of Time Series Analysis* **19**, 267–283.
- Banerjee, A., J. Dolado, J.W. Galbraith y D.F. Hendry (1993), *Cointegration, error-correction, and the econometric analysis of non-stationary data*, Oxford University Press.
- Banerjee, A., R. Lumsdaine y J. Stock (1992), 'Recursive and sequential tests of unit root and trend break hypothesis: Theory and international evidence', *Journal of Business and Economic Statistics* **10**, 271–287.
- Banerjee, A., S. Lazarova y G. Urga (1998), Bootstrapping sequential tests for multiple structural breaks, Discussion Paper 17–98, Center for Economic Forecasting, London Business School.

- Baxter, M. y R. King (1995), Measuring business cycles: Approximate band-pass filters for economic time series, Working Paper 5022, NBER.
- Baxter, M. y R. King (1999), 'Measuring business cycles: Approximate band-pass filters for economic time series', *Review of Economics and Statistics* **81**, 575–593.
- Beran, J., R. Sherman, M.S. Taqqu y W. Willinger (1991), Variable-bit-rate video traffic and long range dependence. Working paper (Dept. of Mathematics of Boston University).
- Beran, R. (1987), 'Prepivoting to reduce level error of confidence sets', *Biometrika* **74**, 457–468.
- Beran, R. (1988), 'Prepivoting test statistics: a bootstrap view of asymptotic refinements', *Journal of the American Statistical Association* **83**, 681–697.
- Berkowitz, J. y L. Kilian (2000), 'Recent developments in bootstrapping time series', *Econometric Reviews* **19**, 1–48.
- Box, G.E.P. y G.M. Jenkins (1970), *Time series analysis: forecasting and control*, Holden-Day, San Francisco.
- Buhlmann, P. (1997), 'Sieve bootstrap for time series', *Bernoulli* **3**, 123–148.
- Buhlmann, P. (1998), 'Sieve bootstrap for smoothing in non-stationary time series', *The Annals of Statistics* **48**, 48–83.
- Buhlmann, P. (1999), Bootstrap for time series, Research report, ETH, Zurich.
- Buhlmann, P. y H. Kunsch (1999), 'Block length selection in the bootstrap for time series', *Computational Statistics and Data Analysis* **31**, 295–310.
- Burns, A. y W. Mitchell (1946), *Measuring Business Cycles*, NBER.
- Campos, J., N. Ericsson y D. Hendry (1996), 'Cointegration tests in the presence of structural breaks', *Journal of Econometrics* **70**, 187–220.
- Canova, F. (1998), 'Detrending and business cycle facts', *Journal of Monetary Economics* **41**, 475–512.
- Carlstein, E. (1986), 'The use of subseries methods for estimating the variance of general statistics from a stationary time series', *The Annals of Statistics* **14**, 1171–1179.
- Carlstein, E., K. Do, P. Hall, T. Hesterberg y H. Kunsch (1998), 'Matched-block bootstrap for dependent data', *Bernoulli* **4**, 305–328.
- Chang, I., G. Tsiao y C. Chen (1988), 'Estimation of time series parameters in the presence of outliers', *Technometrics* **30**, 193–204.
- Chen, C. y L. Liu (1993), 'Joint estimation of model parameters and outlier effects in time series', *Journal of the American Statistical Association* **88**, 284–297.
- Christiano, L. (1992), 'Searching for a break in GNP', *Journal of Business and Economic Statistics* **10**, 237–250.
- Christiano, L. y T. Fitzgerald (1999), The band pass filter, Working Paper 7257, NBER.
- Clements, M. y D. Hendry (1999), *Forecasting Non-stationary Economic Time Series*, MIT Press, chapter 9.

- Cogley, F. y J. Nason (1995), 'Effects of the Hodrick-Prescott filter on trend and difference stationary time series: Implications for business cycle research', *Journal of Economic Dynamics and Control* **19**, 253–278.
- Datta, S. (1996), 'On asymptotic properties of bootstrap for AR(1) processes', *Journal of Statistical Planning and Inference* **53**, 361–374.
- Davidson, A. y B. Hinkley (1997), *Bootstrap Methods and Their Applications*, Cambridge University Press.
- Davidson, A. y P. Hall (1993), 'On studentizing and blocking methods for implementing the bootstrap with dependent data', *Australian Journal of Statistics* **35**, 212–224.
- Davidson, R. y J. MacKinnon (1998), 'Graphical methods for investigating the size and power of hypothesis tests', *The Manchester School* **66**, 1–26.
- Davidson, R. y J. MacKinnon (2000), 'Bootstrap tests: How many bootstraps?', *Econometric Reviews* **19**, 55–68.
- Dickey, D. y W. Fuller (1979), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* **74**, 427–431.
- Dolado, J. y F. Marmol (1998), FM-OLS estimation of cointegrating relationships among nonstationary fractionally integrated processes, Working Paper 98–41, Dpt. of Statistics and Econometrics, Universidad Carlos III de Madrid.
- Dolado, J. y H. Lutkepohl (1996), 'Making Wald tests work for cointegrated VAR systems', *Econometric Reviews* **15**, 369–386.
- Efron, B. (1987), 'Better bootstrap confidence intervals', *Journal of the American Statistical Association* **82**, 171–185.
- Efron, B. y R. Tibshirani (1993), *Introduction to the Bootstrap*, Chapman and Hall.
- Ehglén, J. (1998), 'Distortionary effects of the optimal Hodrick–Prescott filter', *Economics Letters* **61**, 345–349.
- Engle, R. F. y C. W. J. Granger (1987), 'Cointegration and error-correction: Representation, estimation, and testing', *Econometrica* **55**, 251–276.
- Escribano, A. (1986), Nonlinear error-correction. The case of money demand in the U.K. (1878-1970). Chapter IV, PhD thesis, Dept. of Economics of the University of California at San Diego, La Jolla, USA.
- Escribano, A. (1987a), Co-integration, time co-trends and error-correction systems: an alternative approach, Discussion Paper 8715, CORE.
- Escribano, A. (1987b), Error-correcting systems: nonlinear adjustment to linear long-run relationships. CORE Discussion paper 8730, C.O.R.E.
- Escribano, A. y C. W. J. Granger (1998), 'Investigating the relationships between gold and silver prices', *Journal of forecasting* **17**, 81–107.
- Escribano, A. y G. A. Pfann (1998), 'Nonlinear error correction, asymmetric adjustment and cointegration', *Economic modelling* **15**, 197–216.

- Franses, P. y A. Lucas (1997a), Outlier robust cointegration analysis of Dutch interest rates, *in* 'Proceedings of the 1996 Meeting of the American Statistical Association', ASA, pp. 106–109.
- Franses, P. y A. Lucas (1997b), Outlier robust cointegration tests, VU Research Memorandum 97–45, Vrije Universiteit Amsterdam.
- Franses, P. y A. Lucas (1998), 'Outlier detection in cointegration analysis', *Journal of Business and Economic Statistics* **16**, 459–468.
- Franses, P. y N. Haldrup (1994), 'The effects of additive outliers on unit roots and cointegration', *Journal of Business and Economic Statistics* **12**, 471–478.
- Ghysels, E. y P. Perron (1993), 'The effect of seasonal adjustment filters on tests for a unit root', *Journal of Econometrics* **55**, 57–98.
- Giersbergen, N. y J. Kiviet (1993), A Monte Carlo comparison of asymptotic and various nonparametric inference procedures in first-order dynamic models, Discussion Paper TI 93–187, Tinbergen Institute.
- Giersbergen, N. y J. Kiviet (1994), How to implement bootstrap hypothesis testing in static and dynamic regression models, Discussion Paper TI 94–130, Tinbergen Institute.
- Giersbergen, N. y J. Kiviet (1996), 'Bootstrapping a stable AD model: Weak vs. strong exogeneity', *Oxford Bulletin of Economics and Statistics* **48**, 631–656.
- Gómez, V. y A. Maravall (1996), Programs TRAMO and SEATS, Working Paper 9628, Bank of Spain.
- Gonzalo, J. (1994), 'Five alternative methods of estimating long-run relationships', *Journal of Econometrics* **60**, 203–233.
- Granger, C. (1981), 'Some properties of time series and their use in econometric model specification', *Journal of Econometrics* **16**, 121–130.
- Granger, C. W. J. y M. Hatanaka (1964), *Spectral analysis of economic time series*, Princeton University Press, New Jersey.
- Granger, C. W. J. y P. Newbold (1974), 'Spurious regressions in econometrics', *Journal of Econometrics* **2**, 111–120.
- Granger, C. W. J. y R. Joyeux (1980), 'An introduction to long-memory time series models and fractional differencing', *Journal of Time Series Analysis* **1**(1), 15–29.
- Guay, A. y P. St-Amant (1997), Do the Hodrick-Prescott and Baxter-King filters provide a good approximation of business cycles?, Cahiers de Recherche 53, CREFE.
- Hall, P. (1988), 'Theoretical comparison of bootstrap confidence intervals', *The Annals of Statistics* **16**, 927–953.
- Hall, P. (1992), *The Bootstrap and Edgeworth Expansion*, Springer-Verlag.
- Hall, P. (1994), Methodology and theory for the bootstrap, *in* R. Engle y D. McFadden, eds, 'Handbook of Econometrics', Vol. 4, Elsevier Science, chapter 39.

- Hall, P. y S. Wilson (1991), 'Two guidelines for bootstrap hypothesis testing', *Biometrika* **47**, 757–762.
- Hansen, B. (1995), 'Rethinking the univariate approach to unit root testing: Using covariates to increase power', *Econometric Theory* **11**, 1148–1172.
- Harvey, A. C. y A. Jaeger (1993), 'Detrending, business cycles facts and the business cycle', *Journal of Applied Econometrics* **8**, 231–47.
- Harvey, A.C. (1989), *Forecasting structural time series models and the Kalman filter*, Cambridge University Press, Cambridge.
- Hendry, D. y A. Neale (1990), The impact of structural breaks on unit root tests, in P. Hackl y A. Westlung, eds, 'Economic Structural Change: Analysis and Forecasting', Berlin: IIASA and Springer-Verlag, pp. 95–119.
- Hodrick, R. y E. Prescott (1980), Post-war U.S. business cycle: An empirical investigation, Technical report, Carnegie-Mellon University.
- Hodrick, R. y E. Prescott (1997), 'Post-war U.S. business cycle: An empirical investigation', *Journal of Money, Credit and Banking* **29**, 1–16.
- Horowitz, J. (2000), The bootstrap, in R. Engle y D. McFadden, eds, 'Handbook of Econometrics', Vol. 5, North-Holland. To appear.
- Hosking, J.R.M. (1981), 'Fractional differencing', *Biometrika* **68**, 165–176.
- Johansen, S. (1988), 'Statistical analysis of cointegration vectors', *Journal of economic dynamics and control* **12**, 231–254.
- Kaiser, R. (1998), Detection and estimation of structural changes and outliers, Working Paper 98–23, Universidad Carlos III de Madrid.
- Kaiser, R. y A. Maravall (1999), 'Estimation of the business cycle: A modified Hodrick-Prescott filter', *Spanish Economic Review* **1**, 175–206.
- King, R. y S. Rebelo (1993), 'Low frequency filtering and real business cycles', *Journal of Economic Dynamics and Control* **17**, 207–31.
- Kiviet, J. (1984), Bootstrap inference in lagged-dependent variable models, Working paper, University of Amsterdam.
- Knopp, K. (1990), *Theory and application of infinite series*, Dover, N.Y.
- Kremers, J., N. Ericsson y J. Dolado (1992), 'The power of cointegration tests', *Oxford Bulletin of Economics and Statistics* **54**, 349–67.
- Kunsch, P. (1989), 'The jackknife and the bootstrap for general stationary observations', *The Annals of Statistics* **17**, 1217–1241.
- Kydland, F. y E. C. Prescott (1990), 'Business cycles: Real facts and a monetary myth', *Federal Reserve Bank of Minneapolis Quarterly Review* **14**, 3–18.
- Lahiri, S. (1999), 'Theoretical comparison of block bootstrap methods', *The Annals of Statistics* **37**, 386–404.

- Leland, W.E., M.S. Taquq, W. Willinger y D. Wilson (1994), 'On the self-similar nature of ethernet traffic', *IEEE/ACM Transactions on Networking* **2**, 1–15.
- Leybourne, S., T. Mills y P. Newbold (1998), 'Spurious rejections by Dickey–Fuller tests in the presence of a break under the null', *Journal of Econometrics* **87**, 191–203.
- Li, H. y G. Maddala (1996), 'Bootstrapping time series models', *Econometric Reviews* **15**, 115–158.
- Li, H. y G. S. Maddala (1997), 'Bootstrapping cointegrating regressions', *Journal of Econometrics* **80**, 297–318.
- Lucas, A. (1995a), 'An outlier robust unit root tests with an application to the extended Nelson–Plosser data', *Journal of Econometrics* **66**, 153–173.
- Lucas, A. (1995b), 'Unit roots tests based on M estimators', *Econometric Theory* **11**, 331–346.
- Maddala, G. y Y. Yin (1996), Outliers, unit roots and robust estimation of non–stationary time series, in G.Maddala y C.Rao, eds, 'Handbook of Statistics', Vol. 15, North–Holland.
- Mandelbrot, B. (1965), 'Self-similar error clusters in communication systems and the concept of conditional stationarity', *IEEE Transactions on communication technology* **13**, 71–72.
- Mandelbrot, B. (1967), 'Some noises with 1/f spectrum: a bridge between direct current and white noise', *IEEE Transactions on information theory* **13**, 289–298.
- Mantalos, P. y G. Shukur (1998), 'Size and power of the error correction model cointegration test. A bootstrap approach', *Oxford Bulletin of Economics and Statistics* **60**, 249–255.
- Martin, R. y V. Yohai (1986), 'Influence functionals for time series', *The Annals of Statistics* **14**, 781–818.
- Ng, S. y P. Perron (1998), Lag length selection and the construction of unit root tests with good size and power, Working Paper 319, Dpt. of Economics, Boston College.
- Park, D. y T. Willemain (1999), 'The threshold bootstrap and the threshold jackknife', *Computational Statistics and Data Analysis* **31**, 187–202.
- Park, J. (1992), 'Canonical cointegrating regression', *Econometrica* **60**, 119–43.
- Park, J. y P. Phillips (1988), 'Statistical inference in regressions with integrated processes: Part I', *Econometric Theory* **4**, 458–497.
- Park, J. y P. Phillips (1989), 'Statistical inference in regressions with integrated processes: Part II', *Econometric Theory* **5**, 95–131.
- Perron, P. (1989), 'The great crash, the oil price shock and the unit root hypothesis', *Econometrica* **57**, 1361–1401.
- Perron, P. y S. Ng (1996), 'Useful modifications to some unit root tests with dependent errors and their local asymptotic properties', *Review of Economic Studies* **63**, 435–465.
- Perron, P. y T. Vogelsang (1992), 'Nonstationarity and level shifts with an application to purchasing power parity', *Journal of Business and Economic Statistics* **10**, 301–320.
- Phillips, P. (1987a), 'Time series regression with a unit root', *Econometrica* **58**, 277–301.

- Phillips, P. (1987*b*), 'Towards a unified asymptotic theory for autoregression', *Biometrika* **74**, 535–47.
- Phillips, P. y B. Hansen (1990), 'Statistical inference in instrumental variables regression with I(1) processes', *Review of Economic Studies* **57**, 99–125.
- Politis, D. y J. Romano (1994), 'The stationary bootstrap', *Journal of the American Statistical Association* **89**, 1303–1313.
- Priestley, M.B. (1981), *Spectral Analysis and Time Series*, Academic Press.
- Rappoport, P. y L. Reichlin (1989), 'Segmented trends and nonstationary time series', *Economic Journal* **99**, 168–177.
- Shao, J. y D. Tu (1995), *The Jackknife and the Bootstrap*, Springer–Verlag.
- Shin, D., S. Sarkar y J. Lee (1996), 'Unit root tests for time series with outliers', *Stochastic and Probability Letters* **30**, 189–197.
- Stock, J. (1987), 'Asymptotic properties of least squares estimators of cointegrating regressions', *Econometrica* **51**, 153–174.
- Stock, J. (1994), Unit roots, structural breaks and trends, in R. F.Engle y D. L.McFadden, eds, 'Handbook of Econometrics', Vol. 4, North–Holland, chapter 46, pp. 2270–2381.
- Toda, H. y T. Yamamoto (1995), 'Statistical inference in vector autorregression with possibly integrated processes', *Journal of Econometrics* **66**, 225–250.
- Tsay, R. (1986), 'Time series model specification in the presence of outliers', *Journal of the American Statistical Association* **81**, 132–141.
- Vogelsang, T. (1997), 'Wald–type tests for detecting breaks in the trend of a dynamic time series', *Econometric Theory* **13**, 818–849.
- Vogelsang, T. (1999), 'Two simple procedures for testing for a unit root when there are additive outliers', *Journal of Time Series Analysis* **20**, 173–192.
- Wen, Y. y B. Zeng (1999), 'A simple nonlinear filter for economic time series analysis', *Economics Letters* **64**, 151–160.
- White, H. (1984), *Asymptotic theory for econometricians*, Academic Press, San Diego.
- Willinger, W., M.S. Taqqu, W.E. Leland y D.V. Wilson (1995), 'Self-similarity in high-speed packet traffic: analysis and modelling of ethernet traffic measurements', *Statistical Science* **10**(1), 67–85.
- Yin, Y. y G. Maddala (1997), The effects of different outliers on unit root tests, in T.Fornby y R.Hill, eds, 'Advances in Econometrics', Vol. 13, JAI Press.
- Zivot, E. (2000), 'The power of single equation tests for cointegration when the cointegration vector is prespecified', *Econometric Theory* **16**, 407–439.
- Zivot, E. y D. Andrews (1992), 'Further evidence on the great crash, the oil price shock and the unit root hypothesis', *Journal of Business and Economic Statistics* **10**, 251–270.