UNIVERSIDAD DEL BÍO-BÍO

FACULTAD DE CIENCIAS

Transmisión Asimétrica de Precios en el sector de la palta en Chile: Evidencia desde un modelo TVECM

Autor:
Héctor Garrido Henríquez

Profesor(es) Guía(s):
Dr. Sergio Contreras Espinoza
Dra. Monia Ben Kaabia

Tesis para optar al grado de Magíster en Matemática con mención en Estadística

Departamento de Estadística

Departamento de Matemática

4 de septiembre de 2017

Índice general

Índice de figuras

Índice de cuadros

Agradecimientos

Abstract

Capítulo 1

Introducción

1.1. introducción

1.2. justificación del trabajo

Capítulo 2

Análisis univariante de las series de precios

2.1. Fuentes de información

2.1.1. Imputación de valores perdidos

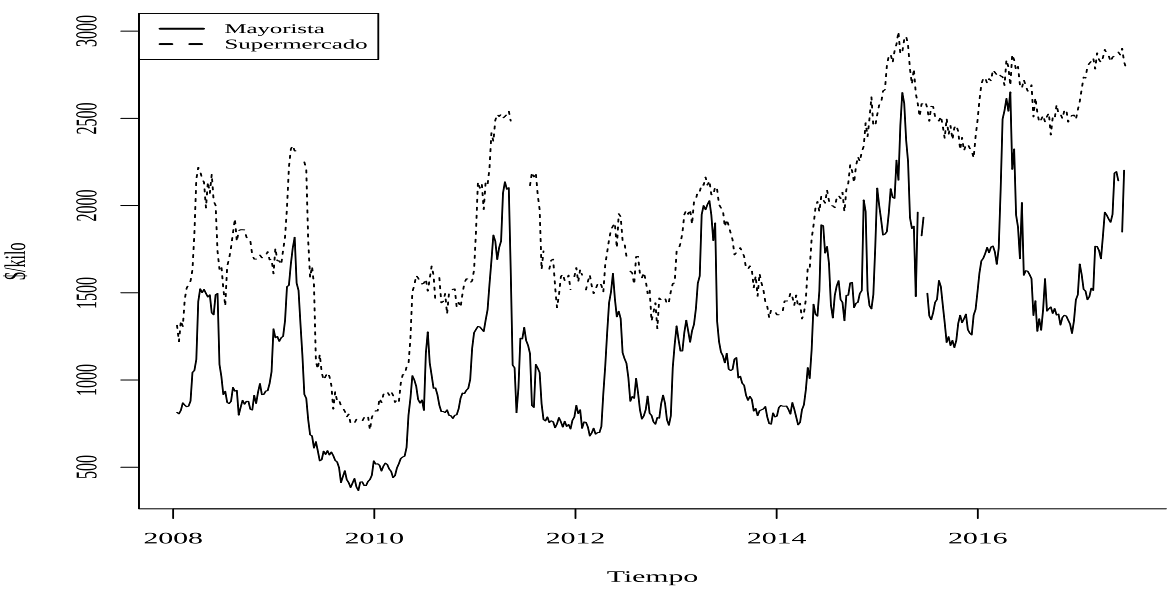
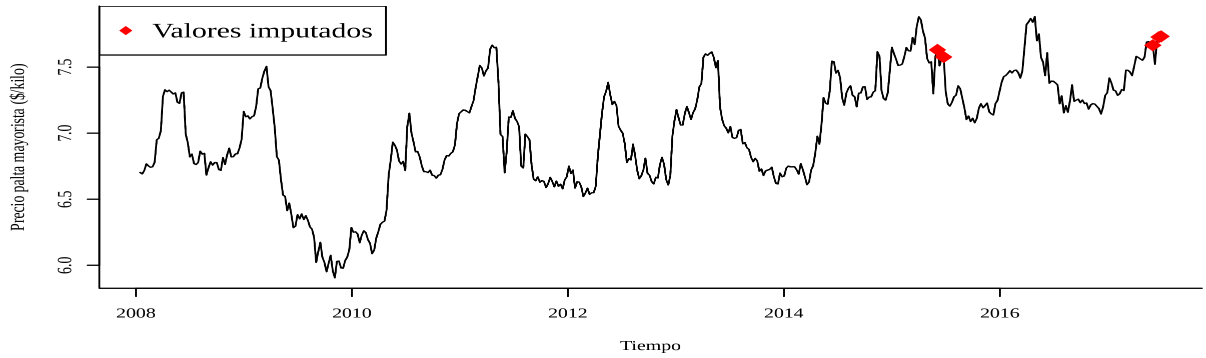


Figura 2.1: Evolución de precios del palta de larga vida de primera calidad,2008-2016



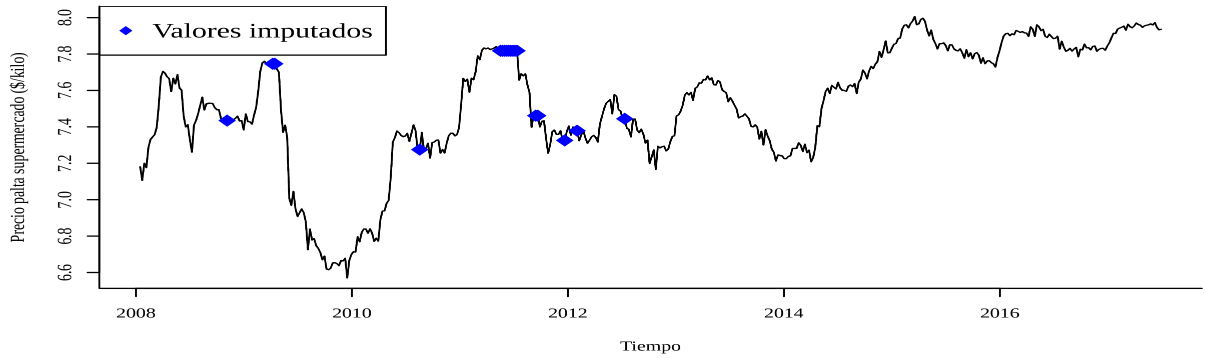


Figura 2.2: Imputación de valores perdidos a través del filtro de Kalman

2.2. Análisis del orden de integración de las series

2.2.1. Análisis gráfico

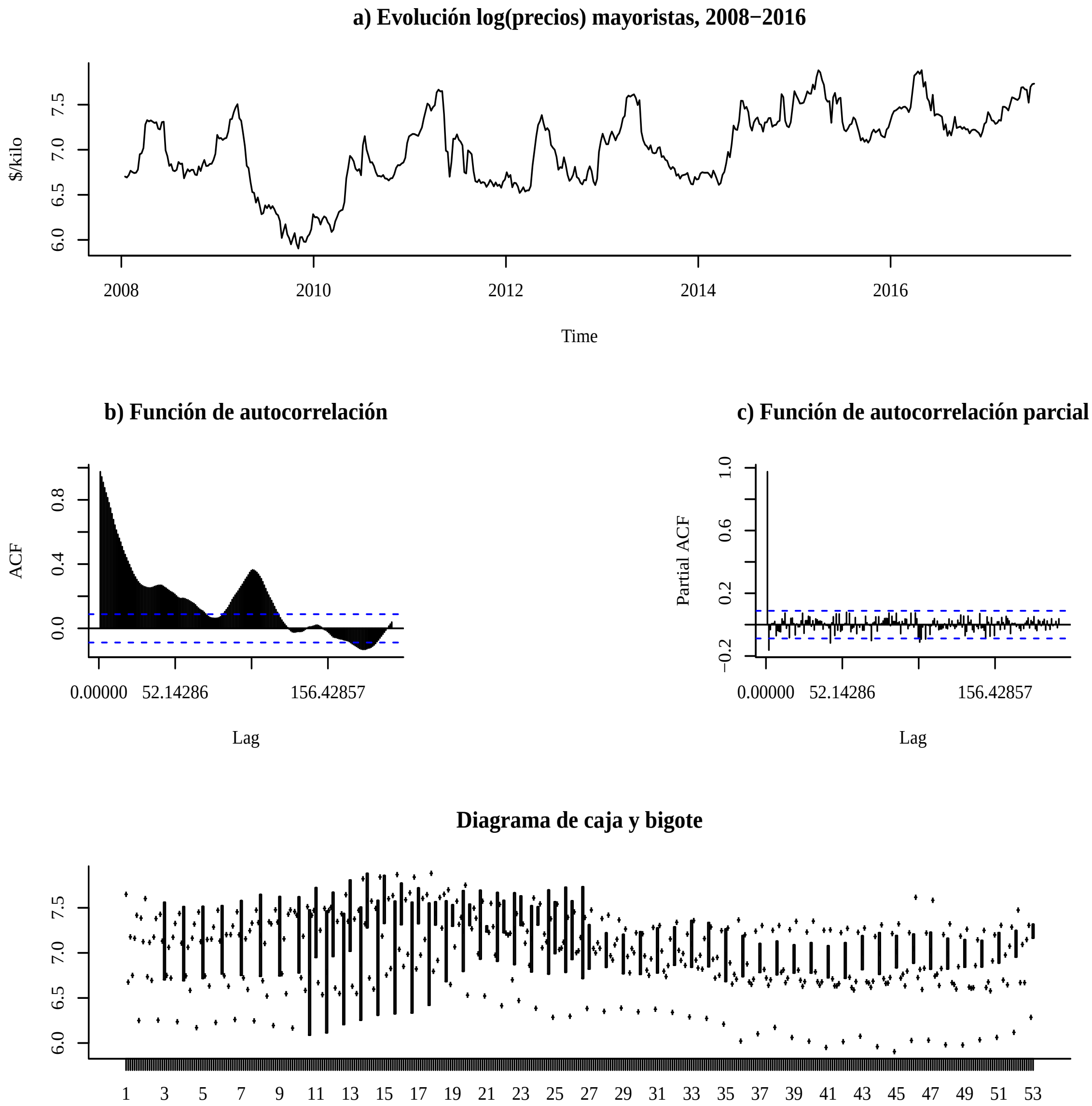


Figura 2.3: Evolución del logaritmo del precio mayorista de la palta ,2008-2016

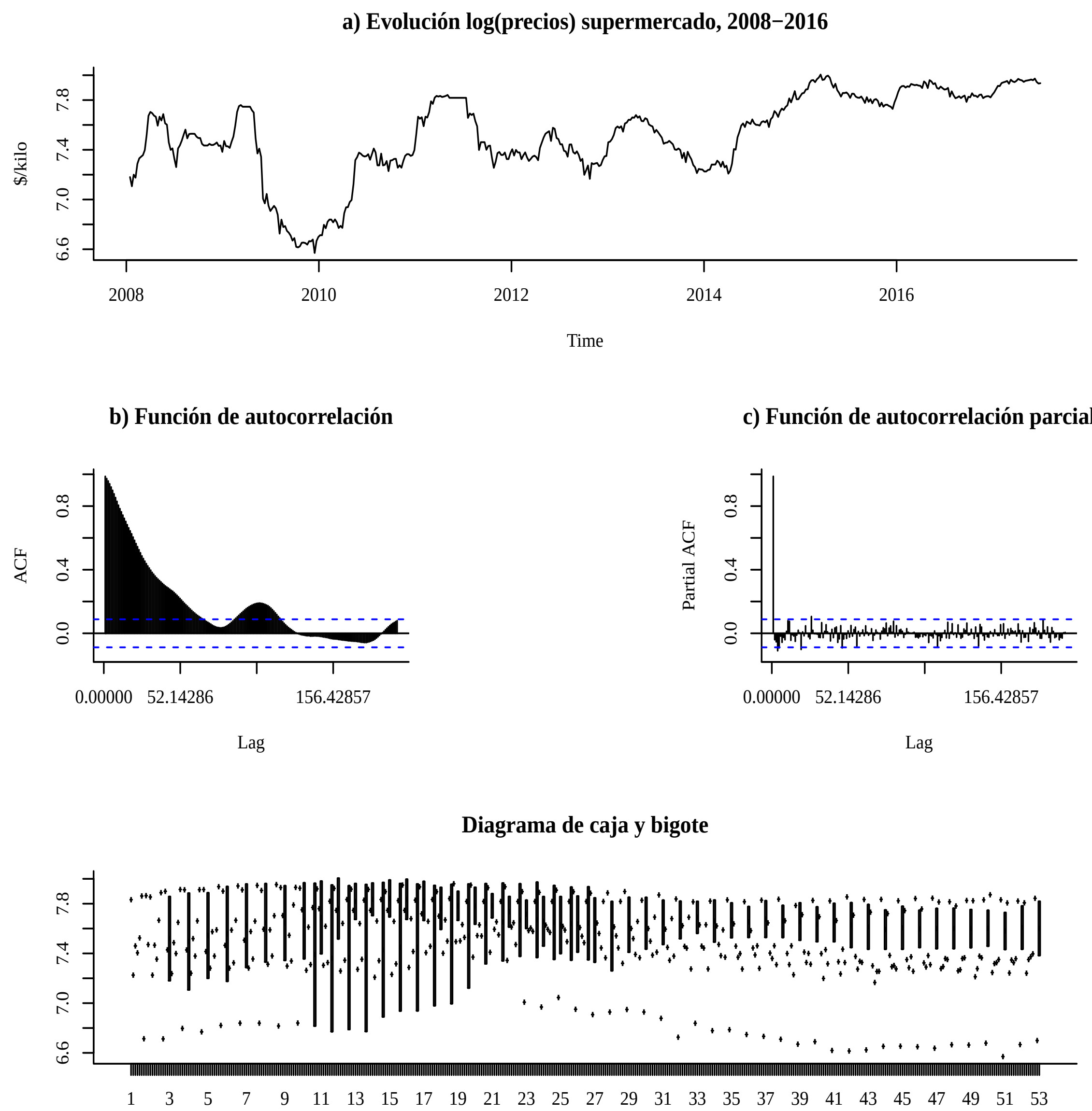


Figura 2.4: Evolución del logaritmo del precio en supermercado de la palta ,2008-2016

2.3. Contrastes de raíz unitaria

2.3.1. Contraste de Dickey-Fuller Aumentado

El contraste más utilizado en la investigación aplicada, dada su simplicidad, es el contraste propuesto por [**fuller1976**] y [**dickey1981**]. Para aplicar este contraste existen dos posibles modelos

Si y_t satisface la siguiente ecuación

$$y_t = \alpha + \rho y_{t-1} + \epsilon_t \qquad (t = 1, ..., n) \tag{2.3.1}$$

Donde $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$.

Si y_t satisface la siguiente ecuación

Como puede observarse en el cuadro 2.1, existen 3 estadísticos, Φ_1, Φ_2 y Φ_3, y sus respectivas hipótesis que pueden ser utilizados. Mientras Φ_1

$$y_t = \alpha + \beta \left(t - 1 - \frac{1}{2}n \right) + \rho y_{t-1} + \epsilon_t \qquad (t = 1, ..., n) \tag{2.3.2}$$

Donde $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$.

Cuadro 2.1: Hipótesis del contraste de Dickey-Fuller

Estadístico	$\mathcal{H}_0$	$\mathcal{H}_a$
τ	$\rho = 1$	$\rho = 0$
Φ_1	$(\alpha, \rho) = (0, 1)$	$(\alpha, \rho) \neq (0, 1)$
Φ_2	$(\alpha, \beta, \rho) = (0, 0, 1)$	$(\alpha, \beta, \rho) \neq (0, 0, 1)$
Φ_3	$(\alpha, \beta, \rho) = (\alpha, 0, 1)$	$(\alpha, \beta, \rho) \neq (\alpha, 0, 1)$

Fuente: Elaboración propia basado en Dickey y Fuller (1981)

2.3.2. Contraste de Phillips Perron (1992)

De manera similar al contraste anterior [**phillips1988**] proponen un contraste no paramétrico para la hipótesis nula de raíz unitaria. A diferencia del contraste de Dickey Fuller este contraste resiste dependencia débil y heteroscedasticidad del término de error. El contraste está construido sobre la base de las siguientes formas funcionales:

$$y_t = \mu + \alpha y_{t-1} + \varepsilon_t, \tag{2.3.3}$$

$$y_t = \mu + \beta \left(t - \frac{1}{2}T \right) + \alpha y_{t-1} + \varepsilon_t \tag{2.3.4}$$

Luego de esto definen los siguientes estadísticos de prueba:

$$Z(\hat{\alpha}) = T(\hat{\alpha} - 1) - \hat{\lambda}/\bar{m}_{yy}, \tag{2.3.5}$$

$$Z(\tau_{\hat{\alpha}}) = (\hat{s}/\hat{\sigma}T_l)t_{\hat{\alpha}} - \hat{\lambda}'\hat{\sigma}T_l/\bar{m}_{yy}^{1/2}, \tag{2.3.6}$$

$$Z(\tau_{\hat{\mu}}) = (\hat{s}/\hat{\sigma}_{Tl})t_{\hat{\mu}} + \hat{\lambda}'\hat{\sigma}_{Tl}m_y/\bar{m}_{yy}^{1/2}m_{yy}^{1/2} \tag{2.3.7}$$

Donde $\bar{m}_{yy} = T^{-2}\sum(y_t - \bar{y})^2$, $m_{yy} = T^{-2}\sum y_t^2$, $m_y = T^{-3/2}\sum y_t$ y $\hat{\lambda} = \frac{1}{2}(\hat{\sigma}_{Tl}^2 - \hat{s}^2)$, donde $\hat{\sigma}^2$ es la varianza muestral de los residuos, $\hat{\lambda}' = \hat{\lambda}/\hat{\sigma}_{Tl}^2$. Luego, la varianza de largo plazo es estimada de la siguiente forma:

$$\hat{\sigma}_{Tl}^2 = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t^2 + 2T^{-1}\sum_{s=1}^{l}w_{sl}\sum_{t=s+1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_{t-s} \tag{2.3.8}$$

Donde $w_{sl} = 1 - s/(l+1)$

De manera similar, el contraste permite la inclusión de una tendencia determinista, modificando los estadísticos de prueba de la siguiente manera:

$$Z(\tilde{\alpha}) = T(\hat{\alpha} - 1) - \hat{\lambda}/M, \tag{2.3.9}$$

$$Z(t_{\tilde{\alpha}}) = (\tilde{s}/\tilde{\sigma}_{Tl})t_{\tilde{\alpha}} - \tilde{\lambda}'\tilde{\sigma}_{Tl}/M^{1/2}, \tag{2.3.10}$$

$$Z(t_{\tilde{\mu}}) = (\tilde{s}/\tilde{\sigma}_{Tl})t_{\tilde{\mu}} - \tilde{\lambda}'\tilde{\sigma}_{Tl}m_y/M^{1/2}(M + m_y^2)^{1/2}, \tag{2.3.11}$$

$$Z(t_{\tilde{\beta}}) = (\tilde{s}/\tilde{\sigma}_{Tl})t_{\tilde{\beta}} - \tilde{\lambda}'\tilde{\sigma}_{Tl}\left(\frac{1}{2}m_y - m_{ty}\right)/(M/12)^{1/2}\bar{m}_{yy}^{1/2} \tag{2.3.12}$$

Donde m_y, $\bar{m}_{yy}$, $\tilde{\lambda}$, $\tilde{\lambda}'$ y $\tilde{\sigma}_{Tl}$ son definidos al igual que en las ecuaciones 2.3.5 a 2.3.7 y $m_{ty} = T^{5/2}\sum t_{yt}$, $t_{\tilde{\mu}}$, $t_{\tilde{\beta}}$ y $t_{\tilde{\alpha}}$ son los estadísticos t de $\tilde{\mu}$, $\tilde{\alpha}$ y $\tilde{\beta}$, respectivamente. Por último la constante $M = (1 - T^{-2})m_{yy} - 12m_{ty}^2 + 12(1 + T^{-1})m_{ty}m_y - (4 + 6T^{-1} + 2T^{-2})m_y^2$

2.3.3. Contraste Kiatkowsky, Pesaran, Schmidt & Shin (1992)

2.3.4. Contraste de Elliot, Rothenberg & Stock (1993)

2.3.5. Contraste de Canova & Hansen (1995) para estacionalidad estable

2.4. Resultados de los contrastes de raíz unitaria

A continuación se aplicará el contraste de Dickey-Fuller aumentado bajo el supuesto de que el proceso subyacente tiene drift y tendencia y bajo el supuesto de que sólo tiene drift.

Cuadro 2.2: Contraste de Dickey-Fuller aumentado (con drift)

	Estadístico		Valores críticos		
$\mathcal{H}_0$	Mayorista[a]	Supermercado[b]	90 %	95 %	99 %
τ_2	-2.5725	-1.6393	-2.57	-2.87	-3.44
ϕ_1	3.4095	1.6354	3.79	4.61	6.47

[a]: Con un rezago, de acuerdo a criterio BIC.
[b]: Con un rezago, de acuerdo a criterio BIC.

El cuadro 2.2 muestra que la evidencia estadística provista por la realización de la serie es apenas suficiente para rechazar la hipótesis $\mathcal{H}_0 : \rho = 1$ a un nivel de significancia de 10 %. Por otro lado, la hipótesis $\mathcal{H}_0 : (\alpha, \rho) = (0, 1)$ no puede ser rechazada.

Con respecto a la serie de precios de supermercados la evidencia va en la misma dirección.

De acuerdo a lo anterior, la serie no contendría una raíz unitaria. De todas maneras, es necesario formular otras representaciones del proceso, como bien podría ser incluir una tendencia determinista (a continuación) o utilizar otros contrastes.

Cuadro 2.3: Contraste de Dickey-Fuller aumentado (con tendencia determinista)

	Estadístico		Valores críticos		
$\mathcal{H}_0$	Mayorista[a]	Supermercado[b]	90 %	95 %	99 %
τ_3	-3.4073	-2.1143	-3.13	-3.42	-3.98
ϕ_2	3.9655	1.6848	4.05	4.71	6.15
ϕ_3	5.84	2.2351	5.36	6.30	8.34

[a]: Con un rezago, de acuerdo a criterio BIC.
[b]: Con un rezago, de acuerdo a criterio BIC.

El cuadro 2.3 muestra que para el caso de la serie de precios mayoristas, la hipótesis $\mathcal{H}_0 : \rho = 1$ se rechaza a un nivel de significancia de 10 %. Mientras que la hipótesis $\mathcal{H}_0 : (\alpha, \beta, \rho) = (0, 0, 1)$ no muestra evidencia estadística suficiente para ser rechaza a los niveles de significancia propuestos. Por último, la hipótesis $\mathcal{H}_0 : (\alpha, \beta, \rho) = (\alpha, 0, 1)$ puede ser rechaza sólo a un 10 % de significancia.

Para el caso de la serie de precios de supermercado, la hipótesis $\mathcal{H}_0 : \rho = 1$ no puede ser rechazada a ninguno de los niveles de significancia propuestos. Mientras que el contraste tampoco provee información suficiente para rechazar $\mathcal{H}_0 : (\alpha, \beta, \rho) = (0, 0, 1)$. Por último, la hipótesis $\mathcal{H}_0 : (\alpha, \beta, \rho) = (\alpha, 0, 1)$ puede ser rechaza sólo a un 10 % de significancia.

El resultado de este contraste es más bien contradictorio, pues no permite concluir que ambas series contengan una raíz unitaria, bajo los diferentes escenarios que sus hipótesis configuran. Queda la posibilidad entonces de que las series tengan un comportamiento estacionario en tendencia, situación que será abordada con el contraste KPSS.

Cuadro 2.4: Contraste Phillips & Perron[a] (con tendencia determinista)

	$Z(t_{\hat{\alpha}})$	$Z(t_{\hat{\mu}})$	$Z(t_{\hat{\beta}})$
Mayorista	-22.1367	1.5224	2.2482
Supermercado	-11.958	1.0702	1.6795

[a]: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

Si se observa el cuadro 2.4 puede observarse que el contraste claramente rechaza la hipótesis de raíz unitaria, aunque no permite identificar con claridad la estructura de los términos deterministas.

Cuadro 2.5: Contraste Phillips & Perron[a] (con tendencia determinista)

	$Z(t_{\hat{\alpha}})$	$Z(t_{\hat{\mu}})$
Mayorista	-13.3185	2.5034
Supermercado	-6.8785	1.846

[a]: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

Al observar el cuadro 2.5 queda de manifiesto el rechazo de la hipótesis de raíz unitaria, aunque por cierto no se puede afirmar con exactitud el comportamiento de los términos deterministas.

Cuadro 2.6: Contraste de Elliot, Rothenberg & Stock (con tendencia determinista)

Estadístico		Valores críticos		
Mayorista[a]	Supermercado[b]	90 %	95 %	99 %
-3.1558	-2.0504	-2.57	-2.89	-3.48

[a]: Con un rezago, de acuerdo a criterio BIC.
[b]: Con un rezago, de acuerdo a criterio BIC.

Cuadro 2.7: Contraste de Elliot, Rothenberg & Stock (con constante)

Estadístico		Valores críticos		
Mayorista[a]	Supermercado[b]	90 %	95 %	99 %
-2.3519	-0.8837	-1.62	-1.94	-2.57

[a]: Con un rezago, de acuerdo a criterio BIC.
[b]: Con un rezago, de acuerdo a criterio BIC.

Cuadro 2.8: Contraste KPSS (con tendencia determinista)

	Estadístico		Valores críticos		
$\mathcal{H}_0$	Mayorista[a]	Supermercado[b]	90 %	95 %	99 %
τ_3	0.2657	0.3246	0.119	0.146	0.216

[a]: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.
[b]: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

Cuadro 2.9: Contraste KPSS (sin tendencia determinista)

	Estadístico		Valores críticos		
$\mathcal{H}_0$	Mayorista[a]	Supermercado[b]	90 %	95 %	99 %
τ_3	2.5592	2.9497	0.347	0.463	0.739

[a]: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.
[b]: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

```
## Error in x[, (1 + cs[i]):cs[i + 1]] <- xx: número de items para para
sustituir no es un múltiplo de la longitud del reemplazo
```

Cuadro 2.10: Contraste de Canova & Hansen[a]

	Estadístico	valor-p
$2\pi/52$,14	0.227	0.6999
$4\pi/52$,14	0.2603	0.6133
$6\pi/52$,14	0.32	0.4755
$8\pi/52$,14	0.2532	0.6312
$10\pi/52$,14	0.3436	0.4287
$12\pi/52$,14	0.2876	0.547
$14\pi/52$,14	0.2109	0.7432
$16\pi/52$,14	0.2976	0.5239
$18\pi/52$,14	0.2948	0.5304
$20\pi/52$,14	0.383	0.3595
$22\pi/52$,14	0.306	0.5052
$24\pi/52$,14	0.3091	0.4985
$26\pi/52$,14	0.2195	0.7199
$28\pi/52$,14	0.4032	0.3282
$30\pi/52$,14	0.3891	0.3498
$32\pi/52$,14	0.1906	0.7979
$34\pi/52$,14	0.2511	0.6367
$36\pi/52$,14	0.2742	0.5789
$38\pi/52$,14	0.6599	0.0982
$40\pi/52$,14	0.2673	0.5959
$42\pi/52$,14	0.2401	0.6652
$44\pi/52$,14	0.1919	0.7945
$46\pi/52$,14	0.5093	0.2012
$48\pi/52$,14	0.303	0.5119
$50\pi/52$,14	0.2531	0.6314
$52\pi/52$,14	0.476	0.235
joint	5.1191	1

[a]: El contraste utiliza términos trigonométricos

Capítulo 3

Análisis de cointegración

3.1. Cointegración y análisis de las relaciones de largo plazo

Apuntes de Farías (2017a)

Dutoit et al. (2010, p. 15) define la transmi- sión en este contexto como “(...) la relación entre los precios de dos mercados relacio- nados; por ejemplo, entre el precio interna- cional de un producto y su precio doméstico

En econometría, la estimación de las re- laciones entre variables que presentan tendencia reviste complejidad, porque su estructura puede provocar que se consideren significativas relaciones comple- tamente espurias (Granger y Newbold, 1974)

Estas técnicas (cointegración) corresponden a los modelos de cointegración, que en la actualidad son utilizados con frecuencia en los estudios de transmisión de precios (aná- lisis horizontal) y de transferencia de costos (análisis vertical)

En este caso, $\mathbf{\Pi}$ pue- de ser factorizado en una matriz $\boldsymbol{\alpha\beta'}$, donde $\boldsymbol{\alpha}$ es una matriz de dimensión $n \times r$ que re- presenta la velocidad de ajuste al equilibrio, mientras que $\boldsymbol{\beta}$ es una matriz de dimensión $n \times r$ que representa los coeficientes de largo plazo

Apuntes de Juselius

- The time series describing cumulated trend-adjusted shocks is usually called a stochas- tic trend. It is a cumulation of random shocks with zero mean and constant variance. If

- with a linear deterministic trend component. Thus, the difference between a stochastic and deterministic trend is that the increments of a stochastic trend change randomly, whereas those of a deterministic trend are constant over time.

- It is easy to see that if inflation rate is I(1) with a non-zero mean, then prices will contain a integrated twice cumulated of order stochastic two, or in trend, t s=1 notation i=1 s i . pt We say I(2). that trend-adjusted prices are

- We shall argue below that, unless a unit root is given a structural interpretation, the choice of one representation or the other is as such not very important, as long as there is consistency between the economic analysis and the choice. However, from an econometric point of view the choice between the two representations is usually crucial for the whole empirical analysis and should therefore be carefully considered.

- variable. Because a cointegrating relation does not necessarily correspond to an interpretable economic relation, we make a further distinction between the statistical concept of a 'cointegration relation' and the economic concept of a 'long-run equilibrium relation'.

- say second that stochastic the distinction trend, between u2i , as a long-run a long-run and structural medium-run trend stochastic or not. trend Thus,in one this might case is between an I(1) stochastic trend with no linear trend and a near I(1) stochastic trend with a linear trend.

Definición 3.1.1. *Sea* $\{\mathbf{x}_t\}$ *un proceso estocástico para* $t = ..., -1, 0, 1, 2, ...$ *Si*

$$\mathbb{E}[\mathbf{x}_t] = -\infty < \mu < \infty \tag{3.1.1}$$

$$\mathbb{E}[\mathbf{x}_t - \mu]^2 = \mathbf{\Sigma}_0 < \infty \qquad \forall t \tag{3.1.2}$$

$$\mathbb{E}[(\mathbf{x}_t - \mu)(\mathbf{x}_{t+h} - \mu)] = \mathbf{\Sigma}_h \qquad \forall t \text{ y } h \tag{3.1.3}$$

Entonces $\mathbf{x}_t$ *es debilmente estacionario. La estacionariedad estricta requiere que la distribución de* $(x_{t1}, ..., x_{tk})$ *es la misma que* $(x_{t1+h}, ..., x_{tk+h})$ *para* $h = ..., -1, 0, 1, 2, ...$

for time t based on the available information at time t 1. For example, a VAR model with autocorrelated and or heteroscedastic residuals would describe agents that do not use all information in the data as efficiently as possible. This is because by including the

For example, a VAR model with autocorrelated and or heteroscedastic residuals would describe agents that do not use all information in the data as efficiently as possible.

Simulation studies have shown that valid statistical inference is sensitive to violation of some of the assumptions, such as parameter non-constancy, autocorrelated residu- als (the higher, the worse) and skewed residuals, while quite robust to others, such as excess kurtosis and residual heteroscedasticity. This will be discussed in more detail in

• the use of intervention dummies to account for significant political or institutional events during the sample; • conditioning on weakly or strongly exogenous variables; • checking the measurements of the chosen variables; • changing the sample period to avoid fundamental regime shift or splitting the sample into more homogenous periods.

and the model has been extended to contain Dt , a vector of deterministic components, such as a constant, seasonal dummies and intervention dummies. The autoregressive for- mulation is useful for expressing hypotheses on economic behaviour, whereas the moving average representation is useful when examining the properties of the proces.

Si asumimos un modelo $VAR(2)$ bi-dimensional

$$(\mathbf{I}-\mathbf{\Pi}_1 L-\mathbf{\Pi}_2 L^2)\mathbf{x}_t=\mathbf{\Phi}\mathbf{D}_t+\varepsilon_t \tag{3.1.4}$$

La función características es entonces

$$\mathbf{\Pi}(z)=\mathbf{I}-\begin{bmatrix}\pi_{1,11} & \pi_{1,12}\\ \pi_{1,21} & \pi_{1,22}\end{bmatrix}z-\begin{bmatrix}\pi_{2,11} & \pi_{2,12}\\ \pi_{2,21} & \pi_{2,22}\end{bmatrix}z^2 \tag{3.1.5}$$

$$=\mathbf{I}-\begin{bmatrix}\pi_{1,11}z & \pi_{1,12}z\\ \pi_{1,21}z & \pi_{1,22}z\end{bmatrix}-\begin{bmatrix}\pi_{2,11}z^2 & \pi_{2,12}z^2\\ \pi_{2,21}z^2 & \pi_{2,22}z^2\end{bmatrix} \tag{3.1.6}$$

$$=\begin{bmatrix}(1-\pi_{1,11}z-\pi_{2,11}z^2) & (-\pi_{1,12}z-\pi_{2,12}z^2)\\ (-\pi_{1,21}z-\pi_{2,21}z^2) & (1-\pi_{1,22}z-\pi_{2,22}z^2)\end{bmatrix} \tag{3.1.7}$$

y

$$|\mathbf{\Pi}(z)|=(1-\pi_{1,11}z-\pi_{2,11}z^2)(1-\pi_{1,22}z-\pi_{2,22}z^2)-(\pi_{1,12}z+\pi_{2,12}z^2)(\pi_{1,21}z+\pi_{2,21}z^2) \tag{3.1.8}$$

$$=1-a_1z-a_2z^2-a_3z^3-a_4z^4 \tag{3.1.9}$$

$$= (1-\rho_1 z)(1-\rho_2 z)(1-\rho_3 z)(1-\rho_4 z) \tag{3.1.10}$$

El determinante entrega información valiosa sobre el comportamiento dinámico del proceso.

Luego

$$\mathbf{x}_t = \frac{\mathbf{\Pi}^a(L)(\mathbf{\Phi D}_t + \varepsilon_t)}{(1-\rho_1 z)(1-\rho_2 z)(1-\rho_3 z)(1-\rho_4 z)} + \tilde{\mathbf{X}}^0, \qquad t = 1, ..., T \tag{3.1.11}$$

$$= \left(\frac{\mathbf{\Pi}_1^a L + \mathbf{\Pi}_2^a L^2}{(1-\rho_2 z)(1-\rho_3 z)(1-\rho_4 z)} \right) \left(\frac{\varepsilon_t + \mathbf{\Phi} D_t}{(1-\rho_1 L)} \right) + \mathbf{X}^0, \qquad t = 1, ..., T \tag{3.1.12}$$

3.2. Estimación basada en la verosimilitud para el modelo VAR irrestricto

Cuando el modelo no tiene restricciones sobre sus parámetros (como las que pueden surgir debido a la presencia de raíces unitarias) el modelo puede estimarse por MCO, caso que coincide con el estimador de *Full information maximum likelihood*

Si escribimos el modelo en su versión apilada

$$\mathbf{x}_t = \mathbf{B}'\mathbf{Z}_t + \varepsilon_t, \qquad t = 1, .., T \tag{3.2.1}$$

$$\varepsilon_t \sim IN_p(\mathbf{0}, \mathbf{\Omega}) \tag{3.2.2}$$

Donde:

- $\mathbf{B}' = [\boldsymbol{\mu}_0, \mathbf{\Pi}_1, \mathbf{\Pi}_2, ..., \mathbf{\Pi}_k]$
- $\mathbf{Z'}_t = \left[\mathbf{1}, \mathbf{x'_{t-1}}, \mathbf{x'_{t-2}}, ..., \mathbf{x'_{t-k}}\right]$
- $\mathbf{X}^0 = \left[\mathbf{x'_0}, \mathbf{x'_{-1}}, ..., \mathbf{x'_{-k+1}}\right]$

La función de verosimilitud será la siguiente:

$$\log L(\boldsymbol{B}, \boldsymbol{\Omega}, \boldsymbol{X}) = -T\frac{p}{2}\log(2\pi) - T\frac{1}{2}\log|\mathbf{\Omega}| - \frac{1}{2}\sum_{t=1}^{T}(\mathbf{x_t} - \mathbf{B}'\mathbf{Z_t})'\mathbf{\Omega}^{-1}(\mathbf{x_t} - \mathbf{B}'\mathbf{Z_t}) \tag{3.2.3}$$

Si calculamos $\frac{\partial \log L}{\partial \mathbf{B}}$, tendremos

$$\sum_{t=1}^{T} \mathbf{x_t Z_t'} = \tilde{\mathbf{B}}' \sum_{t=1}^{T} \mathbf{Z_t Z_t'}$$

Entonces, el estimador de máxima verosimilitud es

$$\tilde{\mathbf{B}}' = \sum_{t=1}^{T}(\mathbf{x_t}\mathbf{Z_t'})\left(\sum_{t=1}^{T}\mathbf{Z_t}\mathbf{Z_t'}\right)^{-1} = \mathbf{M}_{xZ}\mathbf{M}_{ZZ}^{-1} \tag{3.2.4}$$

Luego calculando $\frac{\partial \log L}{\partial \mathbf{\Omega}} = \mathbf{0}$

$$\hat{\mathbf{\Omega}} = T^{-1}\sum_{t=1}^{T}(\mathbf{x_t} - \hat{\mathbf{B}}'\mathbf{Z_t})(\mathbf{x_t} - \hat{\mathbf{B}}'\mathbf{Z_t})' = T^{-1}\sum_{t=1}^{T}\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t' \tag{3.2.5}$$

El valor máximo de la función de Verosimilitud, será el siguiente:

$$\log L_{\text{máx}} = -\frac{P}{2}T\log(2\pi) - \frac{1}{2}T\log|\hat{\mathbf{\Omega}}| - \frac{1}{2}\sum_{t=1}^{T}(\mathbf{x_t} - \hat{\mathbf{B}}'\mathbf{Z_t})'\hat{\mathbf{\Omega}}^{-1}(\mathbf{x_t} - \hat{\mathbf{B}}'\mathbf{Z_t}) \tag{3.2.6}$$

Mostraremos que $\log L_{\text{máx}} = -\frac{1}{2}T\log|\hat{\mathbf{\Omega}}| + K, \qquad K \in \mathbb{R}$

$$\begin{aligned}
(\boldsymbol{x_t} - \hat{\boldsymbol{B}}'\boldsymbol{Z_t})\hat{\mathbf{\Omega}}^{-1}(\mathbf{x_t} - \hat{\mathbf{B}}'\mathbf{Z_t}) &= \hat{\boldsymbol{\varepsilon}}_t'\hat{\mathbf{\Omega}}^{-\mathbf{1}}\hat{\boldsymbol{\varepsilon}}_t \\
&= \sum_{ij}\hat{\varepsilon}_{t,i}(\hat{\mathbf{\Omega}}^{-1})_{ij}\hat{\varepsilon}_{t,j} && (3.2.7)\\
&= \sum_{ij}(\hat{\mathbf{\Omega}}^{-1})_{ij}\hat{\varepsilon}_{t,i}\hat{\varepsilon}_{t,j} \\
&= \text{traza}\{\hat{\mathbf{\Omega}}^{-1}\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t'\} && (3.2.8)
\end{aligned}$$

Luego, se tiene que

$$\begin{aligned}
\sum_{t=1}^{T}(\mathbf{x_t} - \hat{\mathbf{B}}'\mathbf{Z_t})\hat{\mathbf{\Omega}}^{-1}(\mathbf{x_t} - \hat{\mathbf{B}}'\mathbf{Z_t})' &= \sum_{t=1}^{T}\text{traza}\{\hat{\mathbf{\Omega}}^{-1}\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t'\} && (3.2.9)\\
&= T\sum_{t=1}^{T}\text{traza}\{\hat{\mathbf{\Omega}}^{-1}\hat{\boldsymbol{\varepsilon}}_t\hat{\boldsymbol{\varepsilon}}_t'/T\} && (3.2.10)\\
&= T\text{traza}\{\hat{\mathbf{\Omega}}^{-1}\hat{\Omega}\} && (3.2.11)\\
&= T\text{traza}\{\mathbf{I}_p\} = Tp && (3.2.12)
\end{aligned}$$

De donde se desprende que

$$\log L_{\text{máx}} = -T\frac{1}{2}\log|\hat{\mathbf{\Omega}}| \underbrace{-T\frac{p}{2} - T\frac{p}{2}\log(2\pi)}_{+K} \tag{3.2.13}$$

NOTA PARA RECORDAR: si las variables del modelo están formuladas en logaritmo la desviación estándar de cada una de estas puede ser interpretada como un porcentaje de error

3.2.1. Contraste de razón de verosimilitud

$$-2\log Q(\mathcal{H}_k/\mathcal{H}_{k+1}) = T(\log|\hat{\mathbf{\Omega}}_k| - \log|\hat{\mathbf{\Omega}}_{k+1}|) \sim \chi^2_{p^2} \tag{3.2.14}$$

Criterios de selección

$$\text{AIC} = \log|\hat{\mathbf{\Omega}}| + (p^2k)\frac{2}{T} \tag{3.2.15}$$

$$\text{SC} = \log|\hat{\mathbf{\Omega}}| + (p^2k)\frac{\log T}{T} \tag{3.2.16}$$

$$\text{Hannah-Quinn} = \log|\hat{\mathbf{\Omega}}| + (p^2k)\frac{2\log\log T}{T} \tag{3.2.17}$$

Todos los criterios en común están basados en el máximo valor que alcanza la función de verosimilitud del modelo, más un factor que penaliza por el número de parámetros estimados.

Al momento de la determinación del número de rezagos, volver a revisar tabla 4.5 de la página 92

$$\text{Trace correlation} = 1 - \text{traza}(\hat{\mathbf{\Omega}}[\text{Cov}(\mathbf{\Delta x_t})]^{-1})/p \tag{3.2.18}$$

El contraste de Ljung-Box

$$\text{Ljung-Box} = T(T+2)\sum_{h=1}^{T/4}(T-h)^{-1}\text{traza}(\hat{\mathbf{\Omega}}'_{\boldsymbol{h}}\hat{\mathbf{\Omega}}^{\mathbf{-1}}\hat{\mathbf{\Omega}}'_{\boldsymbol{h}}\hat{\mathbf{\Omega}}^{\mathbf{-1}}) \tag{3.2.19}$$

Donde $\hat{\mathbf{\Omega}}_h = T^{-1}\sum_{t=1}^{T}\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}\hat{\boldsymbol{\varepsilon}}'_{\boldsymbol{t-h}}$. El estadístico se considera distribuido aproximadamente según una χ^2 con $p^2(T/4-k+1)-p^2$ grados de libertad.

También puede utilizarse un contraste propuesto por Godfrey(1988), el cual consiste en regresar los residuos del modelo VAR estimado, $\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}$, sobre k variables rezagadas, $\mathbf{x_{t-1}}, \mathbf{x_{t-2}}, ..., \mathbf{x_{t-k}}$ y el j-ésimo residuo rezagado

$$\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{t}} = \mathbf{A_1x_{t-1}} + \mathbf{A_2x_{t-2}} + ... + \mathbf{A_kx_{t-k}} + \mathbf{A}_\varepsilon\hat{\boldsymbol{\varepsilon}} \tag{3.2.20}$$

3.2.2. Modelo Vectorial de Corrección del Error (VECM)

Suponga que cada componente de una serie de tiempo K-dimensional y_t es $I(1)$. Entonces, la ecuación (VAR) no será una formulación adecuada de este modelo debido a que los términos $y_t, y_{t-1}, ..., y_{t-p}$ son todos no estacionarios. De todas formas, sustituyendo

$$\mathbf{A}_1 = \mathbf{I}_k + \mathbf{\Gamma}_1 \tag{3.2.21}$$

$$\mathbf{A}_i = \mathbf{\Gamma_i} - \mathbf{\Gamma_{i-1}} \qquad i = 1, ..., p-1 \tag{3.2.22}$$

$$\mathbf{A}_p = -\mathbf{\Gamma}_{p-1} \tag{3.2.23}$$

En la ecuación (2.8), reagrupando términos y utilizando que $\Delta \mathbf{y}_i = \mathbf{y}_i - \mathbf{y}_{i-1} \quad \forall i$, podemos reescribir esta ecuación como

$$\Delta \mathbf{y}_t = \mu + \mathbf{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \mathbf{\Gamma}_2 \Delta \mathbf{y}_{t-2} + ... + \mathbf{\Gamma}_{p-1} \Delta \mathbf{y}_{t-p+1} + \mathbf{u}_t \tag{3.2.24}$$

Naturalmente, ambas ecuaciones describen el mismo modelo, pero preferimos usar la ecuación (3.2.24) cuando $\mathbf{y}_t$ es $I(1)$, debido a que cada término es estacionario en este caso. Entonces, cuando $\mathbf{y}_t$ es $I(1)$, podemos encontrar un modelo apropiado para y_t diferenciando cada componente de $\mathbf{y}_t$ una vez, y llevando a cabo la regresión basada en la ecuación (3.2.24). De todas formas, entonces no podremos tomar en cuenta que podría haber dependencias entre algunos de los componentes de $\mathbf{y}_t$. Por ejemplo, dos de los componentes podrían tener una tendencia en común, o podría existir una combinación lineal de los componente de y_t la cual fuera estacionaria. Este problema suele resolverse utilizando incluyendo un **término de corrección del error $\mathbf{\Pi y}_{t-1}$** en la ecuación (3.2.24), donde $\mathbf{\Pi}$ es una matriz $K \times K$ de cuyo rango $rank(\mathbf{\Pi}) < K$, debido a que si $\mathbf{\Pi}$ tuviera rango completo, entonces $\mathbf{\Pi}$ es invertible, de manera que la variable no estacionaria $\mathbf{y}_{t-1}$ puede ser escrita como la suma de términos estacionarios, lo que es una contradicción. Entonces, $rank(\mathbf{\Pi}) = r < K$ lo cual implica que existen $(K \times r) - matrices$ $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ de rango r tales que $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$. Entonces, cada una de las r filas de $\boldsymbol{\beta}'\mathbf{y_{t-1}}$ es una combinación lineal estacionaria de los componentes de y_t y es llamada una **relación de cointegración**. El número r, el cual es igual al número de relación de cointegración es llamado el **rango de cointegración**. Como la matriz $\boldsymbol{\beta}$ contiene todos los coeficiente de las relaciones de cointegración, es llamada **la matriz de cointegración**. La matriz $\boldsymbol{\alpha}$, la cual es la matriz de coeficientes de los términos estacionarios $\boldsymbol{\beta}'\mathbf{y_{t-1}}$ en la ecuación (3.2.25), es llamada la matriz de carga.

Definición 3.2.1. *Un modelo* ***VECM*** *de orden p se define como*

$$\Delta \mathbf{y}_t = \boldsymbol{\mu} + \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \mathbf{y}_{t-1} + ... + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{y}_{t-p+1} + u_t \quad t = 1, ..., T \tag{3.2.25}$$

Donde $\mathbf{y}_t = [y_{1t}, ..., y_{Kt}]'$ *es un vector aleatorio de* $K \times 1$, $\boldsymbol{\mu}$ *es un vector constante de* $(K \times 1)$, $\boldsymbol{\alpha}$ *y* $\boldsymbol{\beta}$ *son matrices* $(K \times r)$ *tales que* $rank(\boldsymbol{\alpha}) = rank(\boldsymbol{\beta}) < K$,

3.2.3. El contraste de cointegración de Johansen

Considere el siguiente modelo

$$\Delta \mathbf{y}_t = \Pi \mathbf{y}_{t-1} + \Gamma_1 \Delta \mathbf{y}_{t-1} + ... + \Gamma_{p-1} \Delta \mathbf{y}_{t-p+1} + u_t \tag{3.2.26}$$

Donde $\mathbf{y}_t$ es un proceso K-dimensional y $rk(\Pi) = r$ con $0 \leq r \leq K$.

$$\mathcal{H}_0 : rk(\mathbf{\Pi}) = r_0 \qquad versus \qquad \mathcal{H}_1 : r_0 < rk(\mathbf{\Pi}) \leq r_1 \tag{3.2.27}$$

Lütkepohl(2005, pp. 294) Lütkepohl(2005, pp.340) Tso (1981) Para regresión de rango reducido

Proposición 3.2.1. *Sea* $M := I_T - \Delta\mathbf{X}'(\Delta\mathbf{X}\Delta\mathbf{X}')^{-1}\Delta\mathbf{X}$, $R_0 := \Delta\mathbf{Y}M$ *y* $R_1 = \mathbf{Y}_{-1}M$. *Además*

$$S_{ij} := R_i R_j'/T, \qquad i = 0, 1, \tag{3.2.28}$$

$\lambda_1 \geq ... \geq \lambda_K$ *los autovalores de* $S_{11}^{-1/2} S_{10} S_{00}^{-1} S_{01} S_{11}^{-1/2}$, *y* $\mathbf{v}_1, ..., \mathbf{v}_K$, *los correspondientes autovectores ortonormales*

$$\begin{aligned} \log l = &-\frac{KT}{2}\log 2\pi - \frac{T}{2}\log|\boldsymbol{\Sigma}_u| \\ &-\frac{1}{2}tr[(\Delta\mathbf{Y} - \boldsymbol{\alpha}\boldsymbol{\beta}'Y_{-1} - \boldsymbol{\Gamma}\Delta\mathbf{X})\boldsymbol{\Sigma}_{\mathbf{u}}(\Delta\mathbf{Y} - \boldsymbol{\alpha}\boldsymbol{\beta}'Y_{-1} - \boldsymbol{\Gamma}\Delta\mathbf{X})] \end{aligned} \tag{3.2.29}$$

Desde el resultado anterior puede plantearse el siguiente estadístico de razón de verosimilitud (LRT) para contrastar (3.2.27)

$$
\begin{aligned}
\lambda_{LR}(r_0, r_1) &= 2[\log l(r_1) - \log l(r_0)] \\
&= T\left[-\sum_{i=1}^{r_1}\log(1-\lambda_i) + \sum_{i=1}^{r_0}\log(1-\lambda_i)\right] \\
&= -T\sum_{i=r_0+1}^{r_1}\log(1-\lambda_i)
\end{aligned}
\tag{3.2.30}
$$

Cabe destacar que bajo la hipótesis nula, el estadístico no sigue una distribución estándar por lo que lo que sus valores críticos deben obtenerse mediante simulación. En particular, depende del número de relaciones de cointegracón y del tipo de hipótesis alterna a utilizar. Dos especificaciones son las más utilizadas en la literatura:

$$
\mathcal{H}_0 : rk(\mathbf{\Pi}) = r_0 \qquad \textit{versus} \qquad \mathcal{H}_1 : r_0 < rk(\mathbf{\Pi}) \leq K \tag{3.2.31}
$$

y

$$
\mathcal{H}_0 : rk(\mathbf{\Pi}) = r_0 \qquad \textit{versus} \qquad \mathcal{H}_1 : rk(\mathbf{\Pi}) = r_0 + 1 \tag{3.2.32}
$$

El estadístico $\lambda_{LR}(r_0, K)$ para contrastar (3.2.31) se denomina comunmente como el estadístico *de la traza* para testear el rango de cointegración, mientras que $\lambda_{LR}(r_0, r_0+1)$ es llamado estadístico de *máximo autovalor*.

3.3. Resultados

3.3.1. Determinación del rango de cointegración

```
## $L01
## [1] 0.187144
##
## $df
## [1] 4
##
## $`p-value`
## [1] 0.9958859
```

```
##
## ######################
## # Johansen-Procedure #
```

```
## ######################
##
## Estimation and testing under linear restrictions on beta
##
## The VECM has been estimated subject to:
## beta=H*phi and/or alpha=A*psi
##
##       [,1]
## [1,]  1.00
## [2,] -1.06
## [3,]  0.98
##
## Eigenvalues of restricted VAR (lambda):
## [1] 0.0351
##
## The value of the likelihood ratio test statistic:
## 1.09 distributed as chi square with 2 df.
## The p-value of the test statistic is: 0.58
##
## Eigenvectors, normalised to first column
## of the restricted VAR:
##
##       [,1]
## [1,]  1.00
## [2,] -1.06
## [3,]  0.98
##
## Weights W of the restricted VAR:
##
##                   [,1]
## mayorista.d    -0.1540
## supermercado.d  0.0265
## #############
## ###Model VECM
## #############
## Full sample size: 494  End sample size: 491
## Number of variables: 2  Number of estimated slope parameters 10
```

```
## AIC -5462.592  BIC -5416.431  SSR 4.514969
## Cointegrating vector (estimated by ML):
##    mayorista supermercado    const
## r1         1    -1.168669 1.778087
##
##
##                         ECT                 mayorista -1
## Equation mayorista    -0.1155(0.0249)*** 0.1992(0.0469)***
## Equation supermercado 0.0533(0.0134)***  0.0585(0.0252)*
##                         supermercado -1     mayorista -2
## Equation mayorista    0.2295(0.0824)**    0.0547(0.0476)
## Equation supermercado -0.0270(0.0443)     0.0703(0.0256)**
##                         supermercado -2
## Equation mayorista    0.1593(0.0816).
## Equation supermercado 0.0182(0.0439)
```

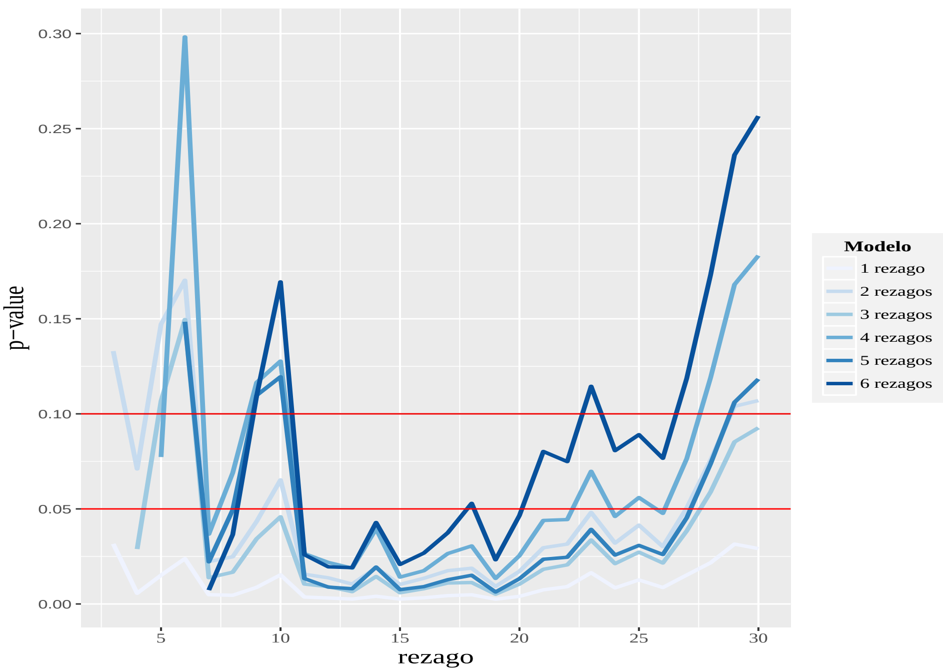


Figura 3.1: Número de Rezagos para el contraste de Independencia

```
## #############
## ###Model VECM
## #############
## Full sample size: 494  End sample size: 486
```

```
## Number of variables: 2  Number of estimated slope parameters 30
## AIC -5399.275  BIC -5269.503  SSR 4.338197
## Cointegrating vector (estimated by ML):
##    mayorista supermercado    const
## r1         1    -1.090485 1.185426
##
##
##                         ECT                mayorista -1
## Equation mayorista    -0.1371(0.0297)*** 0.2053(0.0495)***
## Equation supermercado 0.0241(0.0158)     0.0734(0.0263)**
##                         supermercado -1     mayorista -2
## Equation mayorista    0.2011(0.0890)*     0.0628(0.0503)
## Equation supermercado -0.0728(0.0473)     0.0873(0.0267)**
##                         supermercado -2     mayorista -3
## Equation mayorista    0.1627(0.0886).     0.0333(0.0500)
## Equation supermercado -0.0183(0.0471)     0.0405(0.0266)
##                         supermercado -3    mayorista -4
## Equation mayorista    0.0460(0.0888)     -0.0487(0.0493)
## Equation supermercado 0.0356(0.0472)     0.0227(0.0262)
##                         supermercado -4    mayorista -5
## Equation mayorista    0.0771(0.0872)     -0.0381(0.0490)
## Equation supermercado 0.0823(0.0463).    -0.0309(0.0260)
##                         supermercado -5     mayorista -6
## Equation mayorista    0.0163(0.0866)      0.0742(0.0487)
## Equation supermercado -0.0314(0.0460)     0.0445(0.0259).
##                         supermercado -6    mayorista -7
## Equation mayorista    0.0195(0.0844)     0.0337(0.0488)
## Equation supermercado 0.0131(0.0448)     0.0743(0.0259)**
##                         supermercado -7
## Equation mayorista    0.1745(0.0828)*
## Equation supermercado -0.0192(0.0440)
```

	Estadístico	Valores críticos		
$\mathcal{H}_0$	$p = 2$	90 %	95 %	99 %
$r \leq 1$	2.66	6.50	8.18	11.65
$r = 0$	49.54	15.66	17.95	23.52

Cuadro 3.1: Contraste de la *la traza* de cointegración de Johansen

	Estadístico	Valores críticos		
$\mathcal{H}_0$	$p = 2$	90 %	95 %	99 %
$r \leq 1$	2.66	6.50	8.18	11.65
$r = 0$	46.88	12.91	14.90	19.19

Cuadro 3.2: Contraste del *máximo autovalor* de cointegración de Johansen

Capítulo 4

Procesos de transmisión asimétricos

4.1. Modelo de vectores de corrección del error por umbrales (TVECM)

Definición 4.1.1. *Una serie de tiempo K-dimensional $\mathbf{y}_t$ se dice que sigue un modelo TVECM de k-regímenes de orden p si satisface*

$$\Delta y_t = c_j + \mathbf{\Pi}_j \mathbf{y}_{t-1} + \mathbf{\Gamma}_{1j} \Delta \mathbf{y}_{t-1} + ... + \mathbf{\Gamma}_{(p-1),j} \Delta \mathbf{y}_{t-p+1} + u_{tj}, \quad si \quad \gamma_{j-1} \leq y_{t-d-1} \leq \gamma_j \tag{4.1.1}$$

4.1.1. El contraste de Hansen & Seo (2002)

Cuando estimamos un modelo TVECM resulta vital discernir si este modelo no lineal tiene una performance superior a la que tendría un modelo lineal VECM. Hansen & Seo (2002) propusieron un contraste

4.1.2. Contrastes de linealidad

4.1.3. Estimación del modelo

4.1.4. Relaciones dinámicas a corto plazo

4.2. Resultados de la aplicación

4.2.1. Análisis de las relaciones asimétricas

4.2.2. Especificación del sistema

4.2.3. Estimación del modelo TVECM

4.2.4. Funciones de impulso respuesta

4.2.5. Diagnósticos del modelo

```
## 790 (1.6%) points of the grid lead to regimes with percentage of observations <
## There were  2  thresholds/cointegrating combinations ( -0.04397886460139 / 0.938
## which minimize the SSR in the first search, the first one  -0.04397886   0.93880
```

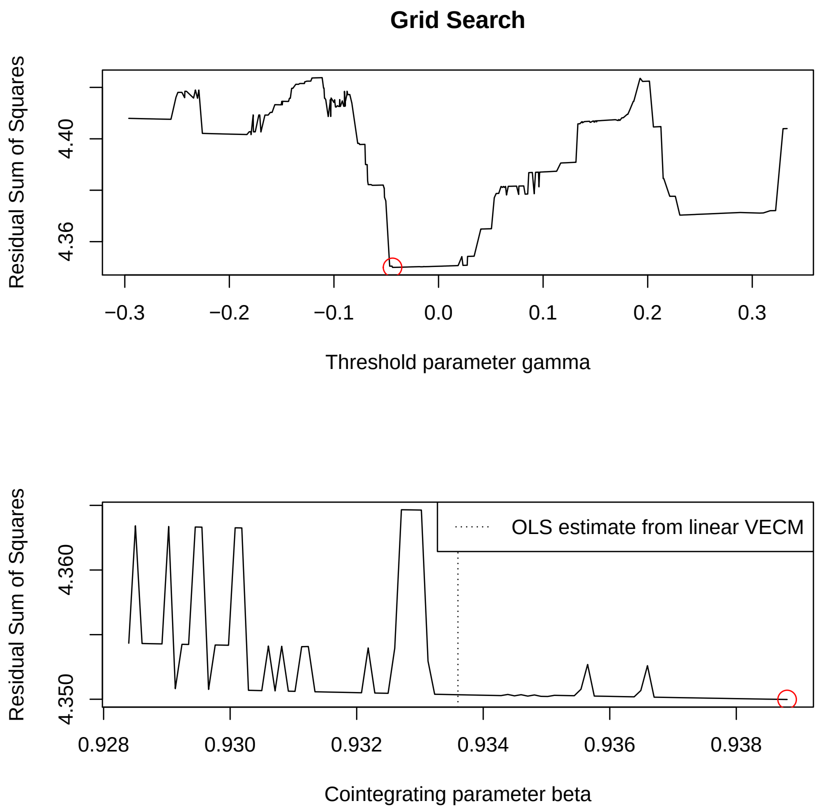
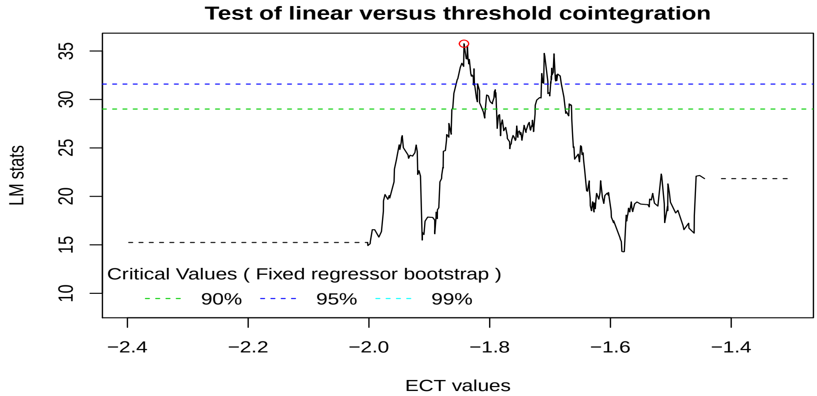
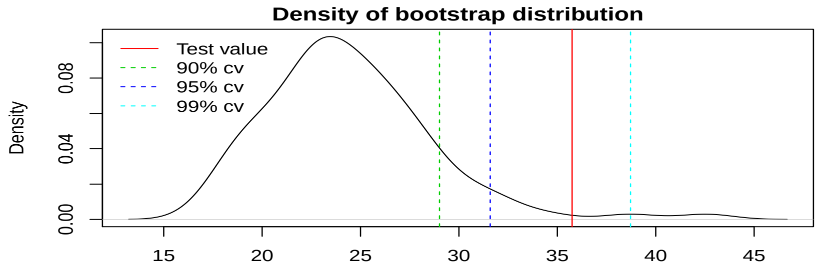
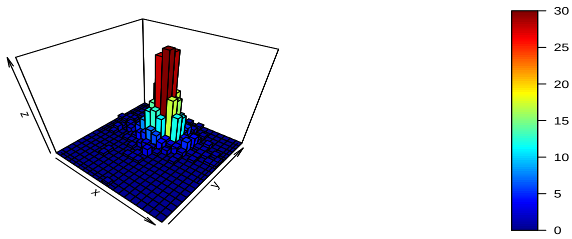


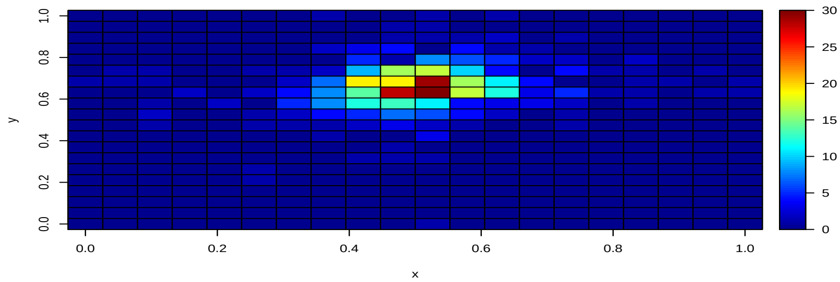
Figura 4.1: Modelo de corrección del error por umbrales

$$\begin{pmatrix}\Delta X_t^1\\ \Delta X_t^2\end{pmatrix} = \begin{cases} \begin{pmatrix}-0{,}1563\\ 0{,}1116\end{pmatrix} ECT_{-1} + \begin{pmatrix}-0{,}0193\\ 0{,}0138\end{pmatrix} + \begin{pmatrix}0{,}0979 & 0{,}2550\\ 0{,}0250 & -0{,}0168\end{pmatrix}\begin{pmatrix}\Delta X_{t-1}^1\\ \Delta X_{t-1}^2\end{pmatrix} + \begin{pmatrix}0{,}0759 & 0{,}2788\\ 0{,}0428 & -0{,}0330\end{pmatrix}\begin{pmatrix}\Delta X_{t-2}^1\\ \Delta X_{t-2}^2\end{pmatrix} & \text{if Th} \\ \begin{pmatrix}-0{,}2379\\ -0{,}0612\end{pmatrix} ECT_{-1} + \begin{pmatrix}0{,}0208\\ 0{,}0154\end{pmatrix} + \begin{pmatrix}0{,}2802 & 0{,}1576\\ 0{,}1009 & -0{,}1198\end{pmatrix}\begin{pmatrix}\Delta X_{t-1}^1\\ \Delta X_{t-1}^2\end{pmatrix} + \begin{pmatrix}0{,}0600 & -0{,}0507\\ 0{,}1142 & 0{,}0542\end{pmatrix}\begin{pmatrix}\Delta X_{t-2}^1\\ \Delta X_{t-2}^2\end{pmatrix} & \text{if Th} \end{cases} \tag{4.2.1}$$









```
## $JB
##
##  JB-Test (multivariate)
##
## data:  Residuals of VAR object mono
## Chi-squared = 758.79, df = 4, p-value < 2.2e-16
##
##
```

```
## $Skewness
##
##  Skewness only (multivariate)
##
## data:  Residuals of VAR object mono
## Chi-squared = 33.819, df = 2, p-value = 4.531e-08
##
##
## $Kurtosis
##
##  Kurtosis only (multivariate)
##
## data:  Residuals of VAR object mono
## Chi-squared = 724.97, df = 2, p-value < 2.2e-16
## $multi
##          E df   P(Chi > E)
## 1 261.0274  4 2.738934e-55
##
## $univ
##                     E df   P(Chi > E)
## mayorista     87.69408  2 9.067181e-20
## supermercado 173.33336  2 2.296880e-38
```

Capítulo 5

Conclusiones

Capítulo 6

Bibliografía