

PARTE II FUNDAMENTO ECONOMÉTRICO

La econometría de series de tiempo está en constante evolución, lo que ha obligado a los economistas a replantearse los modelos y revisar la validez empírica de la teoría. En este apartado se revisan los conceptos fundamentales para que el lector pueda interpretar los resultados estadísticos del presente trabajo, se da por hecho que el lector está familiarizado con la técnica de mínimos cuadrados y con las pruebas de hipótesis. El enfoque que he decidido tomar es necesariamente teórico, en otras palabras, utilizaremos la teoría como guía en el diseño de los modelos.

PROCESOS ESTOCÁSTICOS Y ESTACIONARIEDAD

Una serie de tiempo, es el conjunto de valores que ha tomado una variable en el tiempo. Como en cada momento la variable aleatoria puede tomar un solo valor, se dice que cada valor que toma en el tiempo es una realización de la misma. Al conjunto de estas variables aleatorias ordenadas en el tiempo se le conoce como proceso estocástico y se denota como $\{y_t\}$.

Antes de realizar inferencias y de probar hipótesis, se debe probar si la serie sigue una regularidad. En otras palabras, si vamos a estudiar el comportamiento de las variables y cómo se relacionan, deseamos conocer cuál es el mecanismo que generó los datos. Para

este fin es muy útil el concepto de estacionariedad, se dice que un proceso estocástico $\{y_t\}$ es estacionario si cumple con las siguientes características:

$E(y_t) = \mu_t = \mu$ La media es constante a través del tiempo.

$Var(y_t) = \sigma_t^2 = \sigma^2$ La varianza es constante a través del tiempo.

$Cov(x_t, x_{t+j}) = \sigma_j$ La covarianza entre términos contiguos no depende del tiempo sino de la distancia entre los mismos.

Uno de los procesos estocásticos no estacionarios es el de caminata aleatoria¹⁸ y se define como:

$$y_t = y_{t-1} + u_t$$

En este caso, la variable toma el valor en el periodo anterior más una error aleatorio con media cero, varianza constante e independientemente distribuido. Este tipo de procesos es común en variables como el precio de las acciones, por lo que sería prácticamente imposible obtener ganancias extraordinarias especulando en el mercado. Por medio de iteraciones se puede demostrar que el valor de y en t es igual al valor inicial más la suma de perturbaciones o de los errores en el tiempo:

$$y_t = y_0 + \sum_{i=1}^t u_i$$

¹⁸ También conocido como proceso de raíz unitaria.

A diferencia de un proceso estacionario, el efecto de una innovación o error es permanente. Si se calcula el valor esperado o media, y la varianza de este proceso de caminata aleatoria se obtiene:

$$E(y_t) = y_0$$
$$Var(y_t) = t\sigma_u^2$$

Donde σ_u^2 es la varianza del error, se puede apreciar que aunque la media es constante en el tiempo la varianza no, por lo que la desviación estándar de un pronóstico sobre Y crecerá con la raíz cuadrada de t . Sin embargo, si diferenciamos una vez el proceso de caminata aleatoria, obtendremos un proceso estacionario:

$$y_t = y_{t-1} + u_t$$
$$y_t - y_{t-1} = \Delta y_t = u_t$$
$$E(\Delta y_t) = E(u_t) = 0$$
$$Var(\Delta y_t) = Var(u_t) = \sigma_u^2$$

Otro proceso no estacionario que frecuentemente se encuentra, es el de caminata aleatoria con deriva:

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

Como la media del error es de cero, la serie aunque no es estacionaria tendrá una tendencia ascendente si $a > 0$ y descendente si $a < 0$. Iterando hacia atrás se puede demostrar que¹⁹:

$$y_t = t\alpha + y_0 + \sum_{i=0}^t u_i$$

$$E(y_t) = t\alpha + y_0$$

$$Var(y_t) = t\sigma_u^2$$

No sólo la varianza no permanece constante sino que la media tampoco. Conforme aumenta t la media aumenta si $\alpha > 0$ y disminuye si $\alpha < 0$. Muchas veces podemos identificar qué proceso estocástico podría seguir la serie por simple inspección gráfica. A continuación se presenta la representación gráfica de los siguientes procesos:

$$y_t = y_{t-1} + u_t$$

$$y_t = 0.5 + y_{t-1} + u_t$$

Donde el valor inicial y_0 es igual que cero y la serie u_t representa números aleatorios generados de una distribución normal estándar con 500 observaciones.

¹⁹ En este caso $Cov(y_t, y_{t-k}) = \sigma_u^2(t-k)$

Gráfica 2.1: Procesos no estacionarios



Fuente: Charemza y Deadman (1997) pp.90

El caso general puede representarse como $y_t = \alpha + \delta y_{t-1} + u_t$. Si $\delta = 1$ tenemos un proceso de raíz unitaria o caminata aleatoria con deriva, mientras que si $\delta < 1$ el proceso es estacionario. Por medio de iteraciones podemos expresar el caso estacionario en la siguiente forma:

$$y_t = \alpha \sum_{i=0}^{t-1} \delta^i + \delta^t y_0 + \sum_{i=0}^{t-1} \delta^i u_{t-i} \quad \delta < 1$$

Si $t \rightarrow \infty$ el primer término converge a $\frac{\alpha}{1-\delta}$ mientras que el segundo se hace cero. La

media entonces será estacionaria, también podemos notar que el término de error no

ejerce un efecto permanente en la serie ya que conforme aumenta t su influencia disminuye²⁰. Podemos asumir también que el error sigue un proceso de tipo:

$$u_t = \rho u_{t-1} + e_t \quad \rho < 1$$

Donde $\{e_t\}$ es un proceso con media cero y varianza constante, idéntica e independientemente distribuido, también llamado ruido blanco. Aunque $\{u_t\}$ no se distribuye independientemente de sus realizaciones pasadas o rezagos, es un proceso débilmente dependiente ya que hemos establecido la condición de estacionariedad $\rho < 1$. A este tipo de procesos en la literatura se les conoce como autorregresivos de orden uno AR(1), si ρ es estadísticamente diferente de cero se dice que hay correlación serial de primer orden²¹.

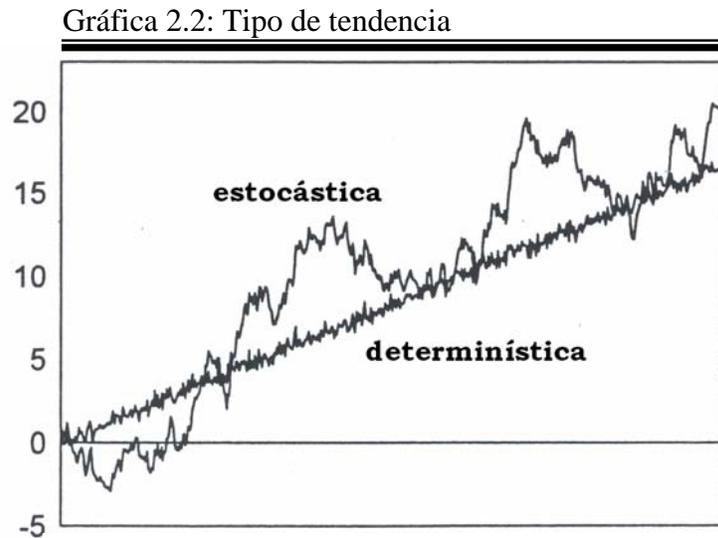
Finalmente puede ser el caso que la serie sea estacionaria, pero alrededor de una tendencia determinística, en el caso especial de que ésta sea lineal el proceso se puede representar como:

$$y_t = \alpha + \beta t + u_t$$

²⁰ En este caso $Var(y_t) = \frac{\sigma_u^2}{1 - \rho^2}$, por lo tanto es constante.

²¹ En presencia de correlación serial los estimadores por mínimos cuadrados son inseguros y consistentes, aunque ya no serán eficientes. No se pueden aplicar las pruebas convencionales de significancia estadística a los estimadores. Las pruebas estadísticas para detectar autocorrelación incluyen el estadístico Durbin-Watson (DW) y la prueba de Breusch-Godfrey entre otras.

La diferencia entre un proceso estacionario alrededor de una tendencia determinística y un proceso no estacionario o con tendencia estocástica²² puede observarse en el siguiente gráfico:



Fuente: Charemza y Deadman (1997) pp.91

En el siguiente apartado se describen brevemente las pruebas estadísticas para determinar si una serie sigue un proceso de raíz unitaria. Como se mencionó al principio, esta sección es un breve repaso de la econometría de series de tiempo, y contiene los conceptos necesarios para interpretar los principales resultados del análisis empírico en la cuarta parte. El lector interesado en profundizar puede consultar la bibliografía.

²² En este caso la serie presenta tanto tendencia estocástica como determinística, ésta última representada por la deriva.

PARTE II FUNDAMENTO ECONOMÉTRICO

La econometría de series de tiempo está en constante evolución, lo que ha obligado a los economistas a replantearse los modelos y revisar la validez empírica de la teoría. En este apartado se revisan los conceptos fundamentales para que el lector pueda interpretar los resultados estadísticos del presente trabajo, se da por hecho que el lector está familiarizado con la técnica de mínimos cuadrados y con las pruebas de hipótesis. El enfoque que he decidido tomar es necesariamente teórico, en otras palabras, utilizaremos la teoría como guía en el diseño de los modelos.

PROCESOS ESTOCÁSTICOS Y ESTACIONARIEDAD

Una serie de tiempo, es el conjunto de valores que ha tomado una variable en el tiempo. Como en cada momento la variable aleatoria puede tomar un solo valor, se dice que cada valor que toma en el tiempo es una realización de la misma. Al conjunto de estas variables aleatorias ordenadas en el tiempo se le conoce como proceso estocástico y se denota como $\{y_t\}$.

Antes de realizar inferencias y de probar hipótesis, se debe probar si la serie sigue una regularidad. En otras palabras, si vamos a estudiar el comportamiento de las variables y cómo se relacionan, deseamos conocer cuál es el mecanismo que generó los datos. Para

este fin es muy útil el concepto de estacionariedad, se dice que un proceso estocástico $\{y_t\}$ es estacionario si cumple con las siguientes características:

$E(y_t) = \mu_t = \mu$ La media es constante a través del tiempo.

$Var(y_t) = \sigma_t^2 = \sigma^2$ La varianza es constante a través del tiempo.

$Cov(x_t, x_{t+j}) = \sigma_j$ La covarianza entre términos contiguos no depende del tiempo sino de la distancia entre los mismos.

Uno de los procesos estocásticos no estacionarios es el de caminata aleatoria¹⁸ y se define como:

$$y_t = y_{t-1} + u_t$$

En este caso, la variable toma el valor en el periodo anterior más una error aleatorio con media cero, varianza constante e independientemente distribuido. Este tipo de procesos es común en variables como el precio de las acciones, por lo que sería prácticamente imposible obtener ganancias extraordinarias especulando en el mercado. Por medio de iteraciones se puede demostrar que el valor de y en t es igual al valor inicial más la suma de perturbaciones o de los errores en el tiempo:

$$y_t = y_0 + \sum_{i=1}^t u_i$$

¹⁸ También conocido como proceso de raíz unitaria.

A diferencia de un proceso estacionario, el efecto de una innovación o error es permanente. Si se calcula el valor esperado o media, y la varianza de este proceso de caminata aleatoria se obtiene:

$$E(y_t) = y_0$$
$$Var(y_t) = t\sigma_u^2$$

Donde σ_u^2 es la varianza del error, se puede apreciar que aunque la media es constante en el tiempo la varianza no, por lo que la desviación estándar de un pronóstico sobre Y crecerá con la raíz cuadrada de t . Sin embargo, si diferenciamos una vez el proceso de caminata aleatoria, obtendremos un proceso estacionario:

$$y_t = y_{t-1} + u_t$$
$$y_t - y_{t-1} = \Delta y_t = u_t$$
$$E(\Delta y_t) = E(u_t) = 0$$
$$Var(\Delta y_t) = Var(u_t) = \sigma_u^2$$

Otro proceso no estacionario que frecuentemente se encuentra, es el de caminata aleatoria con deriva:

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

Como la media del error es de cero, la serie aunque no es estacionaria tendrá una tendencia ascendente si $a > 0$ y descendente si $a < 0$. Iterando hacia atrás se puede demostrar que¹⁹:

$$y_t = t\alpha + y_0 + \sum_{i=0}^t u_i$$

$$E(y_t) = t\alpha + y_0$$

$$Var(y_t) = t\sigma_u^2$$

No sólo la varianza no permanece constante sino que la media tampoco. Conforme aumenta t la media aumenta si $\alpha > 0$ y disminuye si $\alpha < 0$. Muchas veces podemos identificar qué proceso estocástico podría seguir la serie por simple inspección gráfica. A continuación se presenta la representación gráfica de los siguientes procesos:

$$y_t = y_{t-1} + u_t$$

$$y_t = 0.5 + y_{t-1} + u_t$$

Donde el valor inicial y_0 es igual que cero y la serie u_t representa números aleatorios generados de una distribución normal estándar con 500 observaciones.

¹⁹ En este caso $Cov(y_t, y_{t-k}) = \sigma_u^2(t-k)$

Gráfica 2.1: Procesos no estacionarios



Fuente: Charemza y Deadman (1997) pp.90

El caso general puede representarse como $y_t = \alpha + \delta y_{t-1} + u_t$. Si $\delta = 1$ tenemos un proceso de raíz unitaria o caminata aleatoria con deriva, mientras que si $\delta < 1$ el proceso es estacionario. Por medio de iteraciones podemos expresar el caso estacionario en la siguiente forma:

$$y_t = \alpha \sum_{i=0}^{t-1} \delta^i + \delta^t y_0 + \sum_{i=0}^{t-1} \delta^i u_{t-i} \quad \delta < 1$$

Si $t \rightarrow \infty$ el primer término converge a $\frac{\alpha}{1-\delta}$ mientras que el segundo se hace cero. La

media entonces será estacionaria, también podemos notar que el término de error no

ejerce un efecto permanente en la serie ya que conforme aumenta t su influencia disminuye²⁰. Podemos asumir también que el error sigue un proceso de tipo:

$$u_t = \rho u_{t-1} + e_t \quad \rho < 1$$

Donde $\{e_t\}$ es un proceso con media cero y varianza constante, idéntica e independientemente distribuido, también llamado ruido blanco. Aunque $\{u_t\}$ no se distribuye independientemente de sus realizaciones pasadas o rezagos, es un proceso débilmente dependiente ya que hemos establecido la condición de estacionariedad $\rho < 1$. A este tipo de procesos en la literatura se les conoce como autorregresivos de orden uno AR(1), si ρ es estadísticamente diferente de cero se dice que hay correlación serial de primer orden²¹.

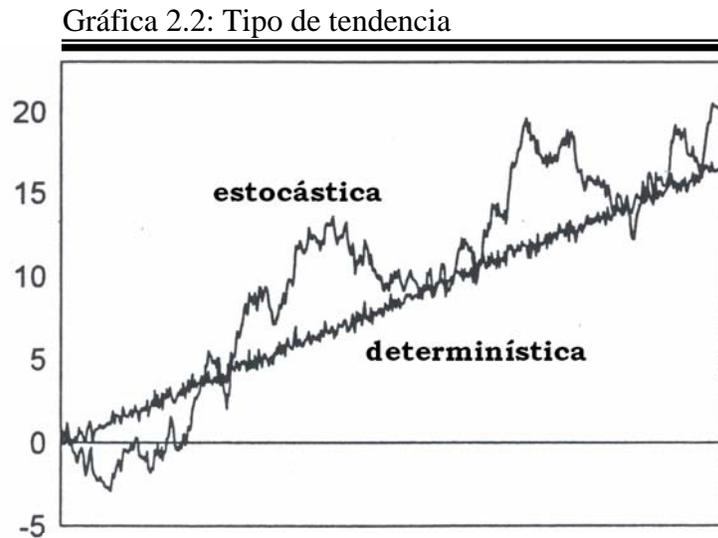
Finalmente puede ser el caso que la serie sea estacionaria, pero alrededor de una tendencia determinística, en el caso especial de que ésta sea lineal el proceso se puede representar como:

$$y_t = \alpha + \beta t + u_t$$

²⁰ En este caso $Var(y_t) = \frac{\sigma_u^2}{1 - \rho^2}$, por lo tanto es constante.

²¹ En presencia de correlación serial los estimadores por mínimos cuadrados son inseguros y consistentes, aunque ya no serán eficientes. No se pueden aplicar las pruebas convencionales de significancia estadística a los estimadores. Las pruebas estadísticas para detectar autocorrelación incluyen el estadístico Durbin-Watson (DW) y la prueba de Breusch-Godfrey entre otras.

La diferencia entre un proceso estacionario alrededor de una tendencia determinística y un proceso no estacionario o con tendencia estocástica²² puede observarse en el siguiente gráfico:



Fuente: Charemza y Deadman (1997) pp.91

En el siguiente apartado se describen brevemente las pruebas estadísticas para determinar si una serie sigue un proceso de raíz unitaria. Como se mencionó al principio, esta sección es un breve repaso de la econometría de series de tiempo, y contiene los conceptos necesarios para interpretar los principales resultados del análisis empírico en la cuarta parte. El lector interesado en profundizar puede consultar la bibliografía.

²² En este caso la serie presenta tanto tendencia estocástica como determinística, ésta última representada por la deriva.

PRUEBAS DE RAÍZ UNITARIA

A pesar de que la inspección gráfica puede ayudar bastante, no es suficiente. Se necesita una prueba estadística para sostener que efectivamente un proceso estocástico es estacionario. Las pruebas más usadas en la literatura son la desarrollada por Dickey y Fuller (1979), y más recientemente la prueba de Phillips y Perron (1988).

PRUEBA DE DICKEY FULLER

En las pruebas de raíz unitaria la hipótesis nula es que la serie no es estacionaria.

Podemos expresar el proceso estocástico en la siguiente manera:

$$\begin{aligned}y_t &= \delta y_{t-1} + e_t \\y_t - y_{t-1} &= \delta y_{t-1} - y_{t-1} + e_t \\ \Delta y_t &= (\delta - 1)y_{t-1} + e_t = \theta y_{t-1} + e_t\end{aligned}$$

Donde bajo la hipótesis nula de raíz unitaria $\theta = (\delta - 1) = 0$, por lo que se trata de una prueba sobre la negatividad de θ . Si estadísticamente el estimador de dicho parámetro es menor que cero, podemos rechazar la hipótesis nula de no estacionariedad habiendo seleccionado el nivel de significancia apropiado. En otras palabras, si el estadístico es más negativo que el valor crítico en la tabla correspondiente al nivel de significancia elegido, se rechaza la hipótesis nula de no estacionariedad.

PRUEBAS DE RAÍZ UNITARIA

A pesar de que la inspección gráfica puede ayudar bastante, no es suficiente. Se necesita una prueba estadística para sostener que efectivamente un proceso estocástico es estacionario. Las pruebas más usadas en la literatura son la desarrollada por Dickey y Fuller (1979), y más recientemente la prueba de Phillips y Perron (1988).

PRUEBA DE DICKEY FULLER

En las pruebas de raíz unitaria la hipótesis nula es que la serie no es estacionaria.

Podemos expresar el proceso estocástico en la siguiente manera:

$$\begin{aligned}y_t &= \delta y_{t-1} + e_t \\y_t - y_{t-1} &= \delta y_{t-1} - y_{t-1} + e_t \\ \Delta y_t &= (\delta - 1)y_{t-1} + e_t = \theta y_{t-1} + e_t\end{aligned}$$

Donde bajo la hipótesis nula de raíz unitaria $\theta = (\delta - 1) = 0$, por lo que se trata de una prueba sobre la negatividad de θ . Si estadísticamente el estimador de dicho parámetro es menor que cero, podemos rechazar la hipótesis nula de no estacionariedad habiendo seleccionado el nivel de significancia apropiado. En otras palabras, si el estadístico es más negativo que el valor crítico en la tabla correspondiente al nivel de significancia elegido, se rechaza la hipótesis nula de no estacionariedad.

A pesar que se usa el estadístico t de Student para realizar la prueba, éste no se distribuye como t^{23} . Dickey y Fuller (1979) calcularon las tablas apropiadas utilizando experimentos de Monte Carlo.²⁴ También se puede incluir una constante o una constante y tendencia determinística. El caso más general es:

$$y_t = \alpha + \beta t + \delta y_{t-1} + e_t$$

$$\Delta y_t = \alpha + \beta t + \theta y_{t-1}$$

Donde se puede probar simultáneamente si no hay tendencia estocástica $\theta < 0$ y si presenta tendencia determinística $\beta \neq 0$. En este caso ya que se prueba más de un parámetro, se calcula el estadístico de multiplicador de Lagrange (LM) y se utilizan las tablas apropiadas. En caso de que el error presente correlación serial, tendremos que optar por la prueba de Dickey Fuller Aumentada (ADF), en ésta se incluyen términos autorregresivos para eliminar la correlación serial. En el caso de una prueba sin constante y tendencia, la ecuación queda como:

$$\Delta y_t = \theta y_{t-1} + \sum_{i=1}^k \theta_i \Delta y_{t-i} + e_t$$

Se utiliza la misma prueba sobre θ , sólo que los valores críticos para rechazar la hipótesis nula son distintos. Éstos fueron calculados por Dickey y Fuller (1981). Mientras algunos sugieren que es criterio del econometrista el número de rezagos incluidos, otros sugieren utilizar un criterio de información como el de Akaike o el de Schwarz. Estos

²³ La razón es porque si y es un proceso estacionario una vez que diferenciamos, estamos regresando un $I(0)$ sobre un $I(1)$.

²⁴ Experimentos de simulación por computadora o muestrales.

criterios imponen una penalización mayor que la \bar{R}^2 al aumentar regresores, el modelo con el menor AIC²⁵ ó SIC²⁶ es preferido. En el caso del criterio de Akaike se tiene:

$$\ln AIC = \left(\frac{2k}{n}\right) + \ln\left(\frac{RSS}{n}\right)$$

Donde k es el número regresores, n el tamaño de la muestra, y RSS la suma del cuadrado de los residuos. Como se puede observar, al incluir regresores RSS disminuye porque se tiene un mejor ajuste pero k aumenta con cada regresor.

PRUEBA DE PHILLIPS-PERRON

Phillips y Perron (1988) desarrollaron un método no paramétrico para controlar por correlación serial al probar la hipótesis nula de raíz unitaria. Se estima la ecuación de la prueba de Dickey Fuller no aumentada y se modifica el estadístico t de los coeficientes para que la correlación serial no afecte la distribución asintótica del mismo. Los paquetes estadísticos calculan esta prueba, mostrando los valores críticos con los que se rechazaría la hipótesis nula.

²⁵ *Akaike Information Criterion*

²⁶ *Schwarz Information Criterion*

ORDEN DE INTEGRACIÓN

El orden de integración es el número de veces que una serie se debe diferenciar para obtener una serie estacionaria, aunque no todas las series al diferenciarse se pueden convertir en estacionarias. Por ejemplo si una serie la diferenciamos d veces y aplicando las pruebas de raíz unitaria rechazamos la hipótesis nula de no estacionariedad, entonces se dice que la serie es integrada de orden d , $I(d)$. En macroeconomía encontramos series que a lo mucho son $I(2)$.

REGRESIONES ESPURIAS

No tomar en cuenta que las series siguen un proceso con tendencia estocástica o que son integradas de cierto orden es peligroso, ya que podríamos obtener coeficientes significativos a pesar de que éstas hayan sido generadas independientemente. Newbold y Davies (1978) en un ejercicio de simulación generaron dos caminatas aleatorias independientes. Como hemos visto, este tipo de series son $I(1)$ ya que al diferenciarlas una vez obtendríamos series estacionarias ó $I(0)$. Estas series son:

$$\begin{aligned}y_t &= y_{t-1} + e_{1t} \\x_t &= x_{t-1} + e_{2t}\end{aligned}$$

e_{1t} y e_{2t} son variables aleatorias independientes que se distribuyen normalmente con media cero y varianza de uno. Fueron generadas 50,000 muestras de 50 observaciones

cada una para e_{1t} y e_{2t} , con estos valores se obtuvieron los procesos de caminata aleatoria utilizando valores iniciales de cero. Luego se regresó x_t sobre y_t (sin constante). Los resultados muestran que tenemos una probabilidad de $2/3$ de que el estimador en la regresión sea significativo, aunque se sabe que en realidad no hay relación entre las variables porque fueron generadas independientemente.²⁷

Convencionalmente se consideraba que una regresión con variables no estacionarias no era confiable porque no serían válidas las pruebas de significancia estadística, y además al no contar con estacionariedad casi nada podríamos decir de la relación entre las series ya que ésta cambiaría con el tiempo. Sin embargo Engle y Granger (1987) mostraron que no siempre sería el caso, las variables podrían relacionarse describiendo una relación de largo plazo no espuria, si la combinación lineal de las mismas era estacionaria. A continuación se expone esta reciente contribución a la econometría de series de tiempo.

COINTEGRACIÓN Y MODELO DE CORRECCIÓN DE ERRORES

Una limitación de realizar el análisis con variables diferenciadas, es que perderemos información sobre la relación de largo plazo que guardan las variables. De acuerdo con Engle y Granger (1987), se dice que las variables en el vector columna \mathbf{x}_t son cointegradas de orden d, b donde $d \geq b \geq 0$ si cada variable es integrada de orden d y existe un vector de cointegración λ tal que $\mathbf{x}'_t \lambda \sim I(d - b)$. Formalmente:

²⁷ Una señal de alerta en los resultados es que se obtuvo un estadístico DW (Durbin Watson) muy bajo, lo que nos indica que hay una fuerte autocorrelación de primer orden.

$$\mathbf{x}'_t \lambda \sim CI(d,b)$$

Por ejemplo si dos series x_t, y_t están cointegradas, y cada una es $I(1)$, existe una combinación lineal de las variables $\beta_1 y_t + \beta_2 x_t$ que es $I(0)$. En este caso $[\beta_1, \beta_2]$ es el vector de cointegración. Matemáticamente: $\mathbf{x}'_t \beta \sim CI(1,1)$.

La interpretación económica de la cointegración de variables es muy agradable pues se puede decir que éstas guardan una relación de equilibrio en el largo plazo. Por lo tanto, la combinación lineal de las mismas no cambia sistemáticamente en el tiempo, es estacionaria. Así, es útil realizar la prueba de cointegración para evitar una regresión espuria.

MÉTODO DE ENGLE-GRANGER

La prueba de cointegración consiste en estimar el vector de cointegración por mínimos cuadrados provisto que se trate de un modelo uniecuacional²⁸, de esta forma si los residuos de la regresión son estacionarios hallamos evidencia de que las variables en efecto están cointegradas. En el ejemplo anterior la regresión de cointegración sería:

$$x_t = \alpha + \delta y_t + e_t$$

²⁸ A diferencia del caso de regresión espuria MCO provee estimadores consistentes de los parámetros de cointegración. Para métodos de sistema aplíquese el método de Johansen.

Podemos probar si los residuos \hat{e}_t son estacionarios aplicando una prueba ADF. Sin embargo, los valores críticos de la prueba no son los mismos que los calculados por Dickey y Fuller debido a que los residuos están basados en el parámetro de cointegración estimado²⁹ aunque algunos autores no toman en cuenta esta diferencia.

MODELO DE CORRECCIÓN DE ERRORES

Una vez que se ha detectado la relación que describe el equilibrio en el largo plazo, se obtienen los residuos de la regresión de cointegración. Cualquier valor distinto de cero de los mismos, significa que las variables no están en equilibrio. Como ejemplo, supongamos que la regresión de cointegración que se obtuvo es:

$$x_t = \hat{\alpha} + \hat{\delta}y_t + \hat{e}_t$$

Podemos obtener los residuos restando los valores ajustados a la variable dependiente:

$$\hat{e}_t = x_t - \hat{\alpha} - \hat{\delta}y_t$$

Llamemos a éstos término de corrección de error (*CE*). El próximo paso es construir el modelo de corto plazo. Ya que estamos trabajando con variables $I(1)$, al diferenciarlas obtendremos variables estacionarias que describen los cambios en las variables originales. Introducimos como variables independientes el término de corrección de error

²⁹ Los valores críticos apropiados pueden consultarse en Charemza y Deadman (1997), pp.281-295.

rezagado un periodo, y las variables explicativas diferenciadas con sus respectivos rezagos. También se pueden incluir algunas variables exógenas y estacionales, el modelo puede representarse como:

$$\Delta x_t = \delta + \lambda CE_{t-1} + \sum_{i=0}^k \alpha_i \Delta y_{t-i} + \sum_{i=1}^l \theta_i w_i + u_t$$

Donde Δy_{t-i} es la diferencia de la variable explicativa y sus rezagos, w_i son variables exógenas y u_t es el término de error. El coeficiente λ es la velocidad de ajuste del modelo ante desequilibrios en la relación de largo plazo, el signo de éste debe llevar el modelo en la dirección correcta. En otras palabras, el signo ante desequilibrios en la relación de largo plazo $CE_{t-1} \neq 0$, *ceteris paribus*, debe llevarnos de regreso al equilibrio. Por ejemplo, si la relación entre x_t y y_t en el largo plazo es positiva, un residuo negativo en la regresión de cointegración indicaría que hay un desequilibrio ya que x_t estaría por debajo de su nivel de largo plazo, de modo que si $\Delta y_t = \Delta y_{t-1} = \dots = \Delta y_{t-k} = 0$ tendría que aumentar x_t , o sea $\lambda < 0$.

VECTORES DE COINTEGRACIÓN (VEC)

Para simplificar la exposición, utilizaré un modelo VAR³⁰ con tres variables³¹. En este tipo de modelos simplemente se regresa cada variable sobre los rezagos de ella misma y las demás, de modo que tenemos:

³⁰ Modelo de vectores autorregresivos.

$$\mathbf{Z}_t = \mathbf{A}_1 \mathbf{Z}_{t-1} + \mathbf{A}_2 \mathbf{Z}_{t-2} + \dots + \mathbf{A}_k \mathbf{Z}_{t-k} + \mathbf{E}_t$$

$$\begin{bmatrix} w_t \\ x_t \\ y_t \end{bmatrix} = \mathbf{A}_1 \begin{bmatrix} w_{t-1} \\ x_{t-1} \\ y_{t-1} \end{bmatrix} + \mathbf{A}_2 \begin{bmatrix} w_{t-2} \\ x_{t-2} \\ y_{t-2} \end{bmatrix} + \dots + \mathbf{A}_k \begin{bmatrix} w_{t-k} \\ x_{t-k} \\ y_{t-k} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{bmatrix}$$

Donde como es usual \mathbf{E}_t es un vector de errores o innovaciones con media cero y varianza constante. Podemos transformar el sistema restando el primer rezago en cada lado de la igualdad:

$$\begin{bmatrix} \Delta w_t \\ \Delta x_t \\ \Delta y_t \end{bmatrix} = \Pi \begin{bmatrix} w_{t-1} \\ x_{t-1} \\ y_{t-1} \end{bmatrix} + \Gamma_1^* \begin{bmatrix} \Delta w_{t-1} \\ \Delta x_{t-1} \\ \Delta y_{t-1} \end{bmatrix} + \Gamma_2^* \begin{bmatrix} \Delta w_{t-2} \\ \Delta x_{t-2} \\ \Delta y_{t-2} \end{bmatrix} + \dots + \Gamma_{k-1}^* \begin{bmatrix} \Delta w_{t-k+1} \\ \Delta x_{t-k+1} \\ \Delta y_{t-k+1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{bmatrix}$$

Donde Π es una matriz de 3×3 que puede expresarse como un producto de dos matrices del mismo orden:

$$\Pi = \alpha\beta'$$

Debido a las dimensiones de Π , sabemos que su rango $r(\Pi)$ ³² máximo es de tres y su rango mínimo de cero. Revisemos cada caso en particular:

³¹ Exposición basada en Charemza y Deadman (1997).

³² Número máximo de columnas linealmente independientes. El rango del producto de dos matrices es igual al de la matriz con el rango menor.

$$\text{Si } r(\Pi) = 0 \quad \begin{bmatrix} \Delta w_t \\ \Delta x_t \\ \Delta y_t \end{bmatrix} = \Gamma_1^* \begin{bmatrix} \Delta w_{t-1} \\ \Delta x_{t-1} \\ \Delta y_{t-1} \end{bmatrix} + \Gamma_2^* \begin{bmatrix} \Delta w_{t-2} \\ \Delta x_{t-2} \\ \Delta y_{t-2} \end{bmatrix} + \dots + \Gamma_{k-1}^* \begin{bmatrix} \Delta w_{t-k+1} \\ \Delta x_{t-k+1} \\ \Delta y_{t-k+1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{bmatrix}$$

Como las variables $I(1)$ están diferenciadas una vez, las pruebas de significancia estadística serán válidas.

$$\text{Si } r(\Pi) = 1 \quad \Pi \mathbf{Z}_{t-1} = \alpha \beta' \mathbf{Z}_{t-1} = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{bmatrix} [\beta_{11} \ \beta_{12} \ \beta_{13}] \begin{bmatrix} w_{t-1} \\ x_{t-1} \\ y_{t-1} \end{bmatrix}$$

Las ecuaciones del sistema quedarían como:

$$\begin{aligned} \Delta w_t &= \alpha_{11} (\beta_{11} w_{t-1} + \beta_{12} x_{t-1} + \beta_{13} y_{t-1}) + \sum_{i=1}^{k-1} a_{1i} \Delta w_{t-i} + \sum_{i=1}^{k-1} b_{1i} \Delta x_{t-i} + \sum_{i=1}^{k-1} c_{1i} \Delta y_{t-i} + e_{1t} \\ \Delta x_t &= \alpha_{21} (\beta_{11} w_{t-1} + \beta_{12} x_{t-1} + \beta_{13} y_{t-1}) + \sum_{i=1}^{k-1} a_{2i} \Delta w_{t-i} + \sum_{i=1}^{k-1} b_{2i} \Delta x_{t-i} + \sum_{i=1}^{k-1} c_{2i} \Delta y_{t-i} + e_{2t} \\ \Delta y_t &= \alpha_{31} (\beta_{11} w_{t-1} + \beta_{12} x_{t-1} + \beta_{13} y_{t-1}) + \sum_{i=1}^{k-1} a_{3i} \Delta w_{t-i} + \sum_{i=1}^{k-1} b_{3i} \Delta x_{t-i} + \sum_{i=1}^{k-1} c_{3i} \Delta y_{t-i} + e_{3t} \end{aligned}$$

Por ejemplo en el caso de la ecuación para Δw_t , sabemos que los rezagos de las diferencias son estacionarios. Esto quiere decir que si la combinación lineal $(\beta_{11} w_{t-1} + \beta_{12} x_{t-1} + \beta_{13} y_{t-1})$ es estacionaria, toda la ecuación lo será. Formalmente:

$$(w_t, x_t, y_t) \sim CI(1,1)$$

Hay un solo vector de cointegración $[\beta_{11} \ \beta_{12} \ \beta_{13}]^{33}$ que entra en cada ecuación del VAR por medio de los coeficientes $\alpha_{11}, \alpha_{12}, \alpha_{13}$, que determinan la velocidad del ajuste.

$$\text{Si } r(\Pi) = 2 \quad \Pi \mathbf{Z}_{t-1} = \alpha \beta' \mathbf{Z}_{t-1} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \end{bmatrix} \begin{bmatrix} w_{t-1} \\ x_{t-1} \\ y_{t-1} \end{bmatrix} \quad \text{En este caso:}$$

$$\begin{aligned} \Delta w_t &= \alpha_{11}(\beta_{11}w_{t-1} + \beta_{12}x_{t-1} + \beta_{13}y_{t-1}) + \alpha_{12}(\beta_{21}w_{t-1} + \beta_{22}x_{t-1} + \beta_{23}y_{t-1}) + \dots + e_{1t} \\ \Delta x_t &= \alpha_{21}(\beta_{11}w_{t-1} + \beta_{12}x_{t-1} + \beta_{13}y_{t-1}) + \alpha_{22}(\beta_{21}w_{t-1} + \beta_{22}x_{t-1} + \beta_{23}y_{t-1}) + \dots + e_{2t} \\ \Delta y_t &= \alpha_{31}(\beta_{11}w_{t-1} + \beta_{12}x_{t-1} + \beta_{13}y_{t-1}) + \alpha_{32}(\beta_{21}w_{t-1} + \beta_{22}x_{t-1} + \beta_{23}y_{t-1}) + \dots + e_{3t} \end{aligned}$$

Si $r(\Pi) = 3$ Hay tres vectores de cointegración, y como cada relación de cointegración es estacionaria podemos resolver para las variables, sabemos que cada variable está en función los tres términos de error que vienen de las regresiones de cointegración. En este caso todas las variables en el VAR deben ser $I(0)$.

Teorema de representación de Granger (1987):

1. Si el rango de la matriz Π es igual que n , el proceso \mathbf{Z}_t es estacionario
2. Si el rango de la matriz Π es menor que n , existe una representación de Π tal que $\Pi = \alpha\beta'$ donde α y β son matrices $n \times r$

Si $\mathbf{Z}_t \sim I(1)$ y $\beta' \mathbf{Z}_t \sim I(0)$ entonces β es la matriz de cointegración y hay un máximo de

³³ El vector se puede normalizar, dividiendo por β_{11} obtenemos $[1 \ \frac{\beta_{12}}{\beta_{11}} \ \frac{\beta_{13}}{\beta_{11}}]$, o sea $[1 - \phi_{12} - \phi_{13}]$.

Donde $\phi_{12} = -\frac{\beta_{12}}{\beta_{11}}$ y $\phi_{13} = -\frac{\beta_{13}}{\beta_{11}}$

$r = n - 1$ ³⁴ vectores de cointegración.

MÉTODO DE JOHANSEN

Este método de máxima verosimilitud basado en un VAR³⁵, se considera como superior al método de una sola ecuación de Engle-Granger (EG) debido a sus propiedades estadísticas. En general, el poder de la prueba de cointegración es mayor³⁶, además se puede utilizar ante modelos uniecuacionales para validar la división entre variables endógenas y exógenas en el método EG³⁷. A continuación se expone la interpretación estadística de los estadísticos obtenidos por el paquete de regresión, para una exposición teórica del método consulte W. Charemza y D. Deadman (1997) además de W. Enders (1995).

Para identificar el número de vectores de cointegración y construir la matriz respectiva, la prueba presenta los eigenvalores³⁸ de modo que $\mu_1 > \mu_2 > \dots > \mu_n$ donde μ_i es el eigenvalor "i". También se presentan los eigenvectores asociados a estos eigenvalores \mathbf{v}_i .³⁹ Si la matriz es de rango $r < n$, los primeros r eigenvectores corresponden a los vectores de cointegración que podemos usar como columnas de la matriz β .

³⁴ n es el número de variables.

³⁵ Vector Autorregresivo.

³⁶ La probabilidad de rechazar una hipótesis siendo falsa.

³⁷ En el método de EG la división se asume.

³⁸ Los eigenvalores son las raíces de una ecuación polinomial obtenida de un determinante.

³⁹ Los eigenvectores se normalizan.

Se calcula el estadístico de razón de verosímil LR^{40} para cada eigenvalor progresivamente, la hipótesis nula es que hay como máximo r vectores de cointegración. Se inicia probando $r = 0$, si no se puede rechazar la hipótesis nula de que hay cero vectores de cointegración termina el proceso y se concluye que no hay evidencia que haya algún vector de cointegración. Si se rechaza la hipótesis nula se continúa probando progresivamente. La mayor dificultad de la prueba es determinar el número de rezagos de las variables a utilizar, el número de rezagos debe ser común para todas las variables del sistema.

PRUEBAS DE CAUSALIDAD DE GRANGER

La definición de causalidad en el sentido de Granger establece que una variable causa a otra, si esta última se puede predecir significativamente mejor haciendo uso de los valores pasados de la primera⁴¹. Sin embargo, cuando se trabaja con series no estacionarias dentro del marco de la cointegración, la prueba convencional de causalidad de Granger debe ser modificada. Simplemente se agrega un término de corrección de error y se utilizan las series en diferencias para excluir cualquier fuente de no estacionariedad en la prueba. El no incluir el término de corrección de error provocaría un sesgo por omisión de variable.

Por ejemplo, si se quiere investigar si x causa a la Granger a y , y ambas son integradas de orden uno, se corre primero la regresión de cointegración de y sobre x . Después, se

⁴⁰ *Likelihood Ratio* con una distribución asintótica aproximada por Johansen (1988).

⁴¹ No se debe confundir con causalidad estricta.

construye el modelo con la primera diferencia de y como variable dependiente y como variables independientes el término de corrección de error, la constante, y los rezagos de las dos variables en diferencias (el número de rezagos debe ser el mismo para cada variable).

Se realiza una prueba F de significancia conjunta para los rezagos de la variable (en primera diferencia) que hipotéticamente está causando en el sentido de Granger a la variable dependiente. Si el estadístico resulta significativo, se encuentra que la información proveniente de los rezagos de la primera diferencia de x ayuda a predecir la variable causada. Por tanto se concluye que x causa a la Granger a y . Formalmente:

$$x \rightarrow y$$

Para determinar si y granger causa a x , se lleva a cabo el mismo procedimiento. En este caso se usa como variable dependiente la primera diferencia de x y el término de corrección de error proviene de la regresión de x sobre y . Si el estadístico F que prueba la significancia conjunta de los rezagos de la primera diferencia de y es mayor que el valor crítico, se concluye que y causa a la Granger a x . Formalmente:

$$y \rightarrow x$$

Si las variables en cuestión están cointegradas, al menos un término de corrección de error debe ser significativo, si en las dos pruebas el estadístico F antes mencionado es significativo, se concluye doble causalidad en el sentido de Granger. Formalmente:

$$y \leftrightarrow x$$

CAMBIO ESTRUCTURAL

Sobre todo si se está trabajando con series de tiempo no estacionarias donde las perturbaciones tienen efectos permanentes, algún choque importante en la economía puede cambiar la relación que guardan las variables en cuestión. Las crisis de balanza de pagos, apertura comercial, y choques de oferta se encuentran entre las principales razones para sospechar que haya ocurrido algún cambio estructural.

PRUEBAS PARA LA DETECCIÓN DE CAMBIO ESTRUCTURAL EN EL MODELO.

Una de las pruebas más antiguas es la de Chow (1960). Sin embargo, esta prueba sólo nos dice si una regresión con datos antes del cambio estructural y otra realizada con datos a partir de éste, son diferentes. La limitación consiste en que no se sabe si el cambio es de pendiente o de constante, y además antes de aplicar la prueba tendremos que identificar cuándo ocurrió el cambio.

Una técnica muy sencilla para detectar el cambio estructural que se utiliza en este trabajo de tesis, es la de Mínimos Cuadrados Recursivos. Partiendo de la primera observación en el tiempo, se estima una primera regresión con una muestra limitada ($n > k$) y se

obtienen los estimados de los k parámetros además de los residuales. Posteriormente se añade la siguiente observación y se vuelven a obtener los estimados, el procedimiento continúa hasta que se cubre toda la muestra.

Para identificar el cambio estructural en el modelo, se hace uso de los residuales recursivos. Con los residuales disponibles después de cada estimación, se calcula el error estándar de la regresión.⁴² Una vez que se llega al periodo en que ocurre el cambio estructural, el residual se comportará de manera considerablemente diferente a los anteriores, evidenciando así el pobre ajuste del modelo inicial que no toma en cuenta el cambio estructural.⁴³

Para resolver el problema de determinar si el cambio en el modelo es de constante, de pendiente o ambas, podemos utilizar variables categóricas. Si éstas son significativas encontraremos evidencia a favor del cambio estructural. Se pueden confirmar los resultados obtenidos de incluir el cambio estructural en el modelo observando el aspecto gráfico de los parámetros recursivos en el modelo inicial.

Si hay estabilidad en el modelo, los cambios en los estimados no serán amplios y además básicamente aleatorios. Sin embargo, si encontramos algún cambio abrupto favoreceremos la evidencia de cambio estructural. La primera parte del gráfico siempre

$$^{42} \hat{\sigma}_t = \left(\frac{\sum \hat{u}^2}{t-k} \right)^{1/2}$$

⁴³ Los estimados recursivos de los coeficientes serán también relativamente estables hasta que el punto de quiebre o de cambio estructural en el modelo se alcanza.

se comportará de manera atípica porque las estimaciones iniciales incluyen casi tantas observaciones como parámetros.