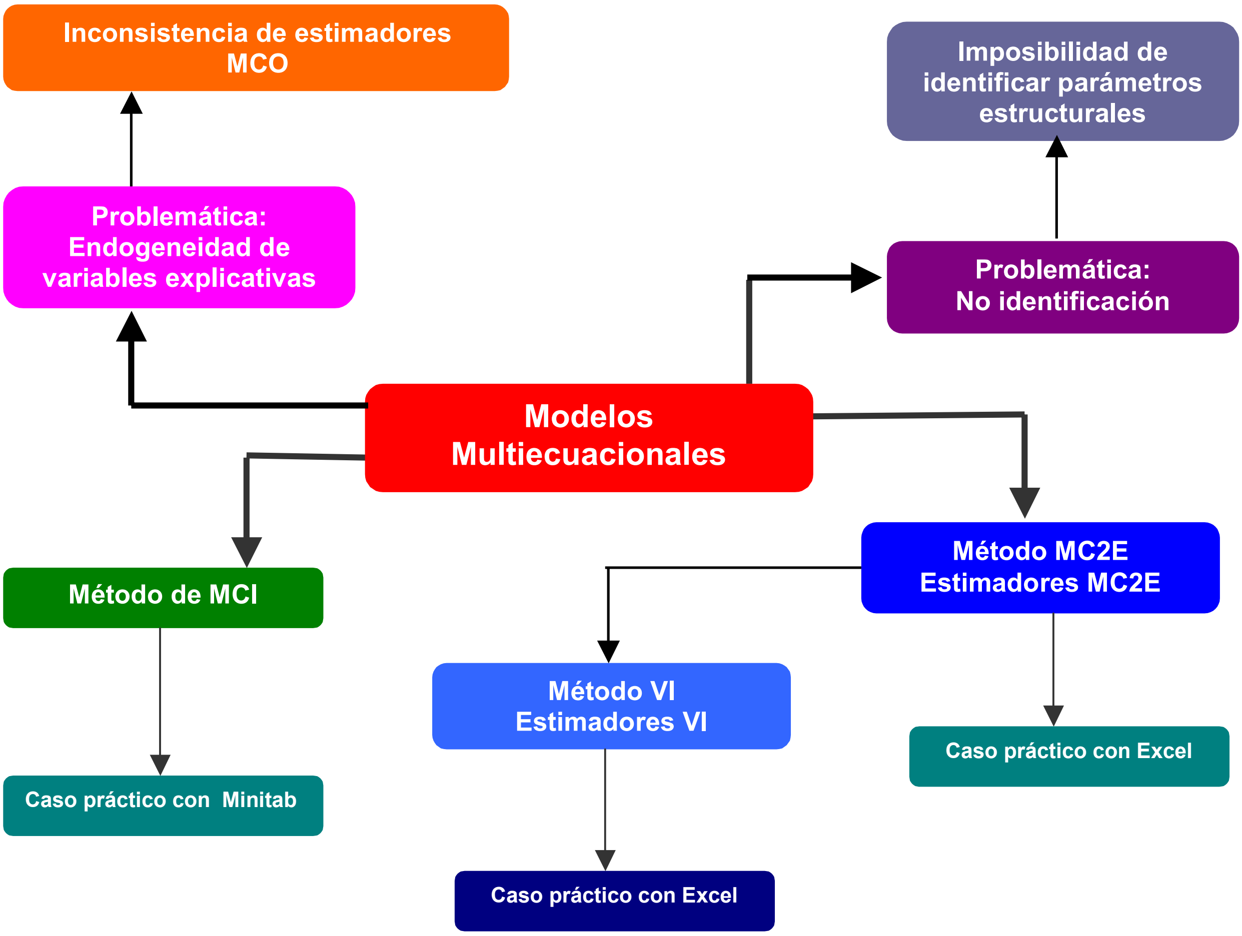
MODELOS DE ECUACIONES SIMULTÁNEAS

Autores: Ángel A. Juan (ajuanp@uoc.edu), Renatas Kizys (rkizys@uoc.edu), Luis María Manzanedo Del Hoyo (lmanzanedo@uoc.edu).

ESQUEMA DE CONTENIDOS



INTRODUCCIÓN

Como es bien sabido, el modelo de regresión lineal generalizado puede estimarse mediante el es método MCG (o de mínimos cuadrados ponderados), siendo los estimadores resultantes insesgados, consistentes y eficientes. Las características que tienen los estimadores MCG se deben al supuesto fundamental de que $\boldsymbol{u}_t$, el término de perturbación asociado a la t-ésima observación-, está incorrelacionado con $\boldsymbol{X}_t$, el vector de las variables explicativas asociado a la misma observación. Sin embargo, hoy en día nos encontramos con numerosas aplicaciones de gran importancia donde este supuesto no se cumple.

Este *math-block* trata sobre sistemas de ecuaciones en los cuales cada una de las ecuaciones representa un modelo de regresión lineal. En tales casos, es habitual encontrarnos con sistemas en los que la variable dependiente de una ecuación actúe también como variable explicativa en otra. Así pues, estaremos ante un problema de endogeneidad de los regresores –suponiendo que otros problemas, como el de la identificación, ya están solucionados. Veremos que el uso del método MCO para estimar los parámetros de las ecuaciones en forma estructural (cuando el mencionado problema persiste) da lugar a estimaciones sesgadas e inconsistentes. Consecuentemente, no será posible utilizar los métodos de estimación convencionales, teniendo que recurrir a métodos alternativos. Para obtener estimadores del modelo que verifiquen las características deseables en cualquier estimador, estudiaremos cómo pasar las ecuaciones en forma estructural a la forma reducida, y cómo determinar las condiciones de orden y de rango.

OBJETIVOS

- Saber representar un modelo multiecuacional en forma estructural y en forma reducida; conocer la relación entre los parámetros estructurales y los parámetros de las ecuaciones en forma reducida.
- Familiarizarse con el problema de la identificación y saber determinar cuando una ecuación está no identificada, cuando está sobreidentificada y cuando está exactamente identificada.
- Conocer el método de mínimos cuadrados indirectos (MCI) de las ecuaciones exactamente identificadas.
- Aprender el método de mínimos cuadrados en dos etapas (MC2E) en el caso de las ecuaciones sobreidentificadas y exactamente identificadas.
- Familiarizarse con el método de estimación de variables instrumetales (VI) como un caso particular de MC2E.

CONOCIMIENTOS PREVIOS

Aparte de estar iniciado en el uso de la hoja de cálculo Excel y del paquete estadístico Minitab, resulta muy conveniente haber leído con profundidad los siguientes *math-blocks*:

- Operaciones con matrices en Excel
- Regresión Lineal Múltiple

CONCEPTOS FUNDAMENTALES

Sistema de ecuaciones simultáneas en forma estructural y forma reducida

Consideremos un sistema de ecuaciones, cada una de las cuales representa un modelo de regresión lineal. Cuando la variable dependiente en una ecuación actúa también como variable explicativa en otra ecuación, estamos ante un **modelo de ecuaciones simultaneas** o **modelo multiecuacional**. Las variables dependientes son también llamadas **variables endógenas**. Por su parte, las variables que vienen determinadas por factores externos al modelo son llamadas **variables exógenas**.

Para ilustrar los problemas de estimación de los multiecuacionales consideremos un sencillo ejercicio de la demanda de naranjas en la provincia de Alicante. Sea P_t el precio (en logaritmos) de naranjas en un determinado año y D_t, la cantidad (en logaritmos) demandada de naranjas. Además, supongamos que las variables son medidas en desviaciones de la media. Así pues, la curva de demanda tiene la siguiente forma:

$$D_t = \alpha P_t + \varepsilon^D_t, \quad (1)$$

con $\alpha < 0$; ε^D_t representa el resto de las variables que pueden influir en la demanda de naranjas. Por otro lado, el precio de naranjas influye, de forma positiva, en la oferta:

$$S_t = \beta P_t + \varepsilon^S_t, \quad (2)$$

con $\beta > 0$; ε^s_t representa el resto de las variables que pueden influir en la oferta de naranjas. Supongamos que ε^d_t y ε^s_t se distribuyen idénticamente e independientemente con varianzas σ_D^2 y σ_S^2, respectivamente, con $\sigma_{DS} = 0$.

Sea Q_t la cantidad de equilibrio en el mercado de naranjas de Alicante. En el equilibrio se cumple:

$$Q_t = S_t = D_t, \text{ o bien}$$

$$\alpha P_t + \varepsilon^d_t = \beta P_t + \varepsilon^s_t, \quad (3)$$

El sistema de ecuaciones (1) y (2) puede rescribirse en la siguiente forma estructural:

$$Q_t - \alpha P_t = \varepsilon^d_t,$$

$$Q_t - \beta P_t = \varepsilon^s_t$$

Matricialmente, tendríamos $B \cdot Y_t + \Gamma \cdot Z_t = U_t$:

$$\underbrace{\begin{pmatrix} 1 & -\alpha \\ 1 & -\beta \end{pmatrix}}_{B} \underbrace{\begin{pmatrix} Q_t \\ P_t \end{pmatrix}}_{Y_t} = \underbrace{\begin{pmatrix} \varepsilon^d_t \\ \varepsilon^s_t \end{pmatrix}}_{U_t} \quad (4)$$

siendo $\Gamma \cdot Z_t = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Como ya hemos comentado, el uso del método MCO para estimar los parámetros de las ecuaciones en forma estructural da lugar a estimaciones sesgadas e inconsistentes. Para solventar este problema, es necesario calcular antes la **forma reducida** del sistema, en la cual cada una de las variables endógenas del sistema es expresada únicamente como función lineal de las variables exógenas del modelo.

En el caso de un sistema como (1) y (2), a partir de la condición de equilibrio, podemos obtener la forma reducida a partir de la condición de equilibrio. Para ello, a partir de la ecuación (3) despejamos P_t:

$$P_t = \underbrace{(\varepsilon^S_t - \varepsilon^D_t)/(\alpha - \beta)}_{v_{1t}} \quad (5)$$

Ahora bien, la expresión del precio de equilibrio sustituimos en ecuación (2), obteniendo:

$$Q_t = \beta(\varepsilon^S - \varepsilon^D_t)/(\alpha - \beta) + \varepsilon^S_t = \alpha\varepsilon^S_t/(\alpha - \beta) - \beta\varepsilon^D_t/(\alpha - \beta);$$

$$Q_t = \underbrace{(\alpha\varepsilon^S_t - \beta\varepsilon^D_t)/(\alpha - \beta)}_{v_{2t}} \quad (6)$$

Siendo v_{1t} y v_{2t} las perturbaciones del modelo en forma reducida. Puesto que no hay variables exógenas en el modelo, el sistema de ecuaciones simultáneas viene expresado en términos aleatorios.

Al mismo resultado se puede llegar usando matrices. De forma general, la forma reducida del sistema vendrá dada por:

$$Y_t = \Pi \cdot Z_t + V_t \qquad (7)$$

donde $\Pi = -B^{-1} \cdot \Gamma$ y $V_t = -B^{-1} \cdot U_t$.

Una vez obtenida la forma reducida, es posible hallar por MCO estimadores insesgados y consistentes de los parámetros Π. No obstante, la ausencia de las variable exógenas en el modelo de demanda-oferta no nos permite estimar la matriz Π. Supongamos que nuestro objetivo consiste en estimar la elasticidad de demanda α. A tales efectos, la ecuación (1) la estimamos por MCO, obteniéndose la siguiente estimación de α a partir de una muestra de T observaciones:

$$a_T = (1/T)\Sigma P_t Q_t/((1/T)\Sigma P_t^2) \quad (8)$$

Sustituyendo las ecuaciones (5) y (6) en la ecuación (8), puede demostrarse que el estimador de la elasticidad de demanda converge en probabilidad a una combinación lineal de las elasticidades poblacionales de demanda y de oferta [5]:

$$a_T \to^P \alpha\gamma + \beta(1-\gamma), \quad (9)$$

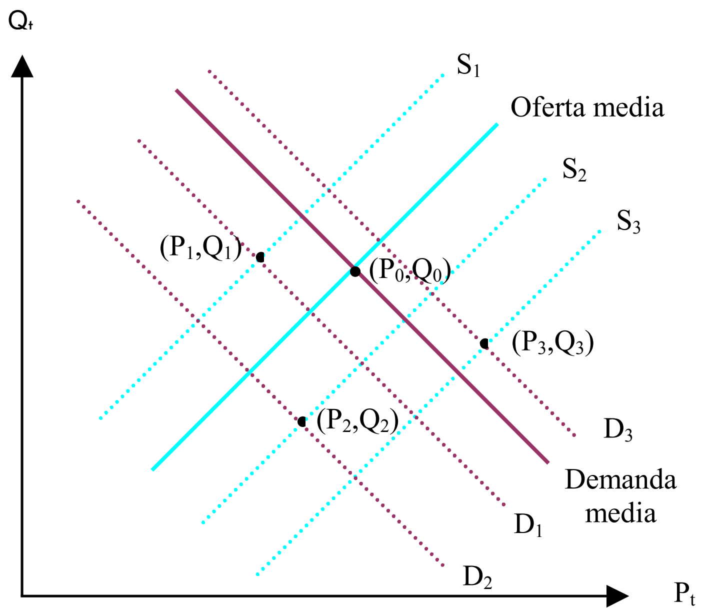
con

$$\gamma = \sigma^2_D/(\sigma^2_D + \sigma^2_S), \quad (10)$$

de modo que γ depende de las varianzas σ^2_S y σ^2_D, siendo $\gamma \in [0, 1]$. Puede verse que si $\sigma^2_S \to 0$, o bien $\sigma^2_D \to \propto$, entonces a_T será un estimador consistente de α. Por otro lado, si $\sigma^2_D \to 0$, o bien $\sigma^2_S \to \propto$, tenemos que a_T es un estimador consistente de β. Para el resto de los valores, a_T es una estimador consistente de la combinación lineal de α y de β. En general, a_T no es un estimador consistente de la elasticidad de demanda. Éste es un frecuente problema de los modelos multiecuacionales. La simultaneidad de las ecuaciones da lugar a las interpretaciones erróneas de los resultados de estimación. Así pues, en el último caso, un economista equivocadamente cree que a_T es un estimador MCO de la elasticidad de demanda, mientras que otro economista considera a_T es un estimador de la elasticidad de oferta, cuando en realidad éste es un estimador de la combinación lineal de las dos elasticidades. Las Figuras 1, 2 y 3 analizan gráficamente el problema de simultaneidad.

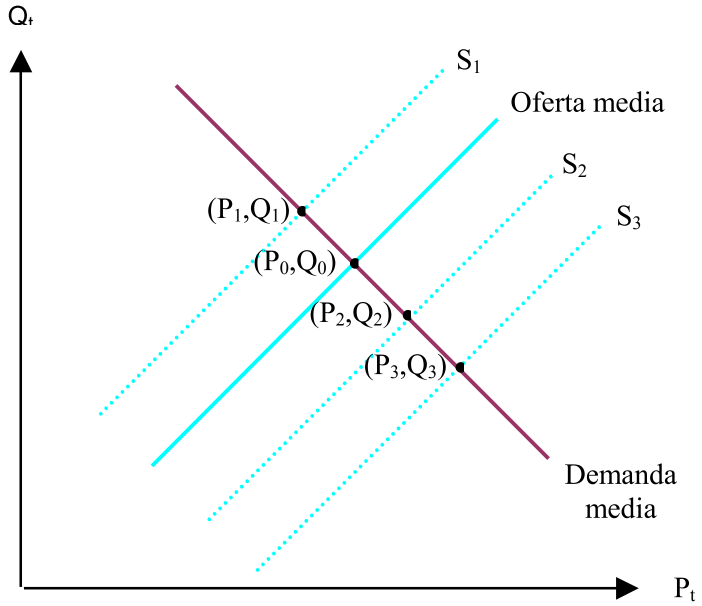
En el gráfico, un equilibrio viene dado por la intersección de la curva de demanda y la curva de oferta. Sea **(P_0,Q_0)** un equilibrio inicial. Supongamos que en t = 1 se produce un shock pequeño y negativo de demanda, dando lugar a D_1, combinando con un shock elevado de signo positivo de oferta que, a su vez, genera S_1. El efecto conjunto sitúa al punto de intersección en **(P_1,Q_1)**. Por otro lado, un shock grande negativo de demanda y un shock grande negativo de oferta se traducen en **(P_2,Q_2)**. Finalmente, el punto **(P_3,Q_3)** representa un efecto conjunto de una innovación moderada positiva de demanda y una innovación grande negativa a oferta. Así pues, las combinaciones de los shocks forman una nube sobre la cual puede cruzarse la recta de regresión. En este caso, a_T es un estimador consistente de "la mixtura" de elasticidades de oferta y demanda, $\alpha\gamma + \beta(1-\gamma)$.

Figura 1. Comportamiento del mercado de naranjas en respuesta a las perturbaciones de demanda y de oferta en conjunto



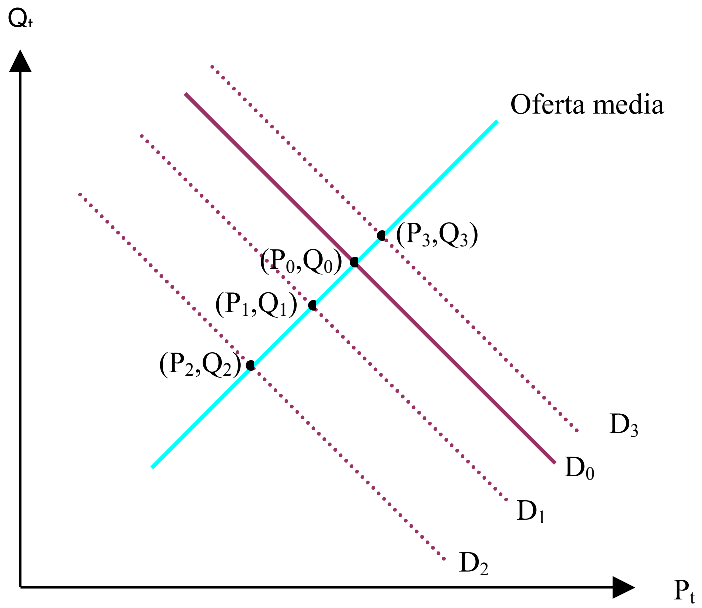
En la Figura 2, el equilibrio viene dado por la intersección de la curva de demanda y la curva de oferta. Sea **(P_0,Q_0)** un equilibrio de partida. Supongamos que en t =1 se produce una innovación positiva de oferta, desplazando a la curva de oferta a S_1, y, en consecuencia, produciendo un nuevo punto de equilibrio en **(P_1,Q_1)**. Por otro lado, un shock negativo de oferta en t = 2 se traduce en el punto **(P_2,Q_2)**. Finalmente, el punto **(P_3,Q_3)** representa un efecto de una innovación elevada negativa a oferta. Así pues, las innovaciones a oferta permiten trazar la curva de demanda. En este caso, a_T es un estimador consistente de la elasticidad de demanda, **α**.

Figura 2. Comportamiento del mercado de naranjas en respuesta a las innovaciones a oferta.



En la Figura 3, el equilibrio viene dado por la intersección de la curva de demanda y la curva de oferta. Sea **(P_0,Q_0)** un punto de equilibrio en t = 0. Ante una perturbación moderada negativa a demanda en t = 1, la curva de demanda se desplaza a D_1, generando el punto de equilibrio en **(P_1,Q_1)**. Por otro lado, un shock elevado negativo de demanda sitúa al equilibrio en el punto **(P_2,Q_2)**. Finalmente, el punto **(P_3,Q_3)** representa un efecto conjunto de una innovación moderada positiva a demanda. Así, a través de todos los puntos de equilibrio, obtenemos la curva de oferta Concluimos, por tanto, que a_T es un estimador consistente de la elasticidad de oferta **β**.

Figura 3. Comportamiento del mercado de naranjas en respuesta a las innovaciones a demanda.



Identificación

El **problema de la identificación** hace referencia a la posibilidad o no de calcular los parámetros estructurales de un modelo de ecuaciones simultaneas (elementos de las matrices $\boldsymbol{B}$ y $\boldsymbol{\Gamma}$) a partir de los parámetros de la forma reducida asociada (elementos de la matriz $\boldsymbol{\Pi}$), los cuales sí se podían estimar mediante MCO.

- Diremos que una **ecuación** está **no identificada** cuando no tengamos suficiente información para estimar los parámetros de la forma estructural de la ecuación.
- Diremos que una **ecuación** está **sobreidentificada** cuando haya más de una combinación posible de valores estimados para los parámetros de la forma estructural.
- Finalmente, una **ecuación** estará **exactamente identificada** cuando sólo sea posible obtener una única estimación de los parámetros estructurales.
- Dado un modelo multiecuacional en forma estructural, diremos que es un **sistema exactamente identificado** cuando todas sus ecuaciones lo sean. En tal caso, será posible obtener, de forma unívoca, estimaciones para los elementos de las matrices $\boldsymbol{B}$ y $\boldsymbol{\Gamma}$ a partir de los elementos de la matriz $\boldsymbol{\Pi}$.

Condición de orden:

Sean:

N1 = nº de variables exógenas del sistema no incluidas en una determinada ecuación
N2 = nº de variables endógenas de dicha ecuación

Dada una ecuación identificada,

- Si N1 = N2 – 1, entonces la ecuación está **exactamente identificada**.
- Si N1 > N2 – 1, entonces la ecuación está **sobreidentificada**.

Observación: Aunque la condición de orden es sólo una condición necesaria y no suficiente (i.e.: el hecho de que se cumpla N1 ≥ N2 – 1 no implica de por sí que la ecuación esté identificada), en la gran mayoría de los casos proporciona la respuesta correcta al problema de la identificación sin necesidad de recurrir a la condición de rango [1].

Ejemplo 1: Consideremos el sistema de ecuaciones simultáneas consistiendo de (1) y (2). Podríamos decir que ambas ecuaciones están no identificadas (pues no excluyen ninguna variable exógena, e incluyen dos endógenas, $\boldsymbol{Q}$ y $\boldsymbol{P}$). Siguiendo el análisis anterior, para poder recuperar la elasticidad de demanda a partir de la estimación MCO, necesitamos información adicional acerca la oferta. En concreto, hemos de encontrar una variable que pueda desplazar a la curva de oferta y no a la de demanda. Sea $\boldsymbol{W_t}$ el número de días cuando las temperaturas caen por debajo de 0ºC. Técnicamente, la nueva variable $\boldsymbol{W_t}$ forma parte de $\boldsymbol{\varepsilon^S_t}$, pues ésta última incorpora los factores que influyen en la cantidad ofertada de naranjas diferentes de precio. Por consiguiente, $\boldsymbol{\varepsilon^S}_t$ puede descomponerse en dos términos, el término sistemático y el término idiosincrático:

$$\boldsymbol{\varepsilon^S_t = \eta w_t + u^S_t}, \quad (11)$$

siendo $\boldsymbol{\eta}$, el coeficiente de proyección lineal de $\boldsymbol{\varepsilon^S_t}$ sobre $\boldsymbol{w}_t$. Por tanto, $\boldsymbol{u^S_t}$ está incorrelado con $\boldsymbol{W_t}$. Aunque el indicador de temperatura pueda influir directamente en la cosecha de naranjas y, consecuentemente, en la cantidad ofertada de los cítricos, pero es natural suponer que el tiempo afectará la demanda de naranjas a través del precio de la mismas. Cambios en el precio que pueden ser asociados a los cambios en tiempo, representan desplazamientos de la curva de demanda y no de la curva de oferta. De esta manera, además de lograr identificar la demanda, el comportamiento de la misma puede sistematizarse con una mayor precisión. La elasticidad de demanda ahora puede estimarse por MCO consistentemente. Para ello, sustituimos (11) en (5) obteniéndose

$$P_t = (u^S_t + \eta W_t - \varepsilon^D_t)/(\alpha - \beta) \quad (12)$$

Sea P^*_t la proyección lineal de P_t sobre W_t:

$$P^*_t = \eta W_t/(\alpha - \beta) \quad (13)$$

La expresión de P_t queda:

$$P_t = P^*_t + (u^S_t - \varepsilon^D_t)/(\alpha - \beta) \quad (14)$$

Ahora bien, sustituyendo la ecuación (14) en (1) y agrupando los términos, tenemos:

$$Q_t = \alpha[P^*_t + (u^S_t - \varepsilon^D_t)/(\alpha - \beta)] + \varepsilon^D_t = \alpha P^*_t + \xi_t, \quad (15)$$

Siendo

$$\xi_t = \alpha u^S_t/(\alpha - \beta) - \beta\varepsilon^D_t/(\alpha - \beta) \quad (16)$$

Puesto que u^S_t y ε^D_t están incorrelacionadas con W_t, tenemos que ξ_t está incorrelado con P^*_t. En consecuencia, la estimación de MCO de la ecuación de regresión (15) nos proporcionará un estimador consistente de α:

$$a^*_T = (1/T)\Sigma P^*_t Q_t/((1/T)\Sigma P^{*2}_t) \quad (17)$$

Es fácil ver que $a^*_T \rightarrow^P \alpha$ [5].

Ecuaciones exactamente identificadas: estimación por MCI:

El método de **mínimos cuadrados indirectos** (MCI) permite estimar los parámetros estructurales (elementos de las matrices ***B*** y ***Γ***) en el caso de ecuaciones exactamente identificadas. El método MCI consiste en:

- Aplicar MCO para estimar los parámetros del sistema en forma reducida (elementos de la matriz ***Π***), y
- Usar dichas estimaciones para calcular los parámetros estructurales.

Ejemplo 2: La tabla siguiente, referida a los datos obtenidos a partir de una simulación., muestra los valores para ***Q***, ***P*** y ***W*** en términos logarítmicos (***W*** viene definida por el logaritmo de una función lineal de las condiciones climatológicas) durante el período 1981 – 2000.

Observación	Año	Q6	P	W
1	1981	-0,085181	-1,04982	-1,20397
2	1982	-0,159996	-0,65393	-0,10536
3	1983	-0,087284	-0,79851	-0,69315
4	1984	0,005580	-0,69315	-1,20397
5	1985	-0,097901	-0,43078	-0,10536
6	1986	-0,004575	-0,59784	-0,91629
7	1987	-0,035900	-0,43078	-0,35667
8	1988	-0,068519	-0,10536	0,33647
9	1989	0,164610	-0,28768	-1,60944
10	1990	0,078521	0,13976	-0,35667
11	1991	0,053681	0,30010	0,09531
12	1992	0,241651	0,22314	-1,20397
13	1993	0,115055	0,37156	-0,10536
14	1994	0,143971	0,40547	-0,35667
15	1995	0,211917	0,43825	-0,69315
16	1996	0,153309	0,53063	-0,10536
17	1997	0,171513	0,55962	-0,35667
18	1998	0,294913	0,50078	-1,20397
19	1999	0,238418	0,66783	-0,51083
20	2000	0,241904	0,85442	-0,22314

Primeramente, especificamos el sistema de ecuaciones simultáneas en forma estructural:

$$\begin{aligned} Q_t - \alpha P_t &= u^D{}_t \\ Q_t - \beta P_t - \eta W_t &= u^S{}_t \end{aligned} \qquad (18)$$

Reescribiendo el sistema en forma matricial, tenemos $B \cdot Y_t + \Gamma \cdot Z_t = U_t$:

$$\begin{pmatrix} 1 & -\alpha \\ 1 & -\beta \end{pmatrix} \begin{pmatrix} Q_t \\ P_t \end{pmatrix} + \begin{pmatrix} 0 \\ -\eta \end{pmatrix} W_t = \begin{pmatrix} u^D{}_t \\ u^S{}_t \end{pmatrix} \qquad (19)$$

Sabemos que la forma reducida del sistema viene dada dada por:

$$\boldsymbol{Y_t = \Pi \cdot Z_t + V_t}$$

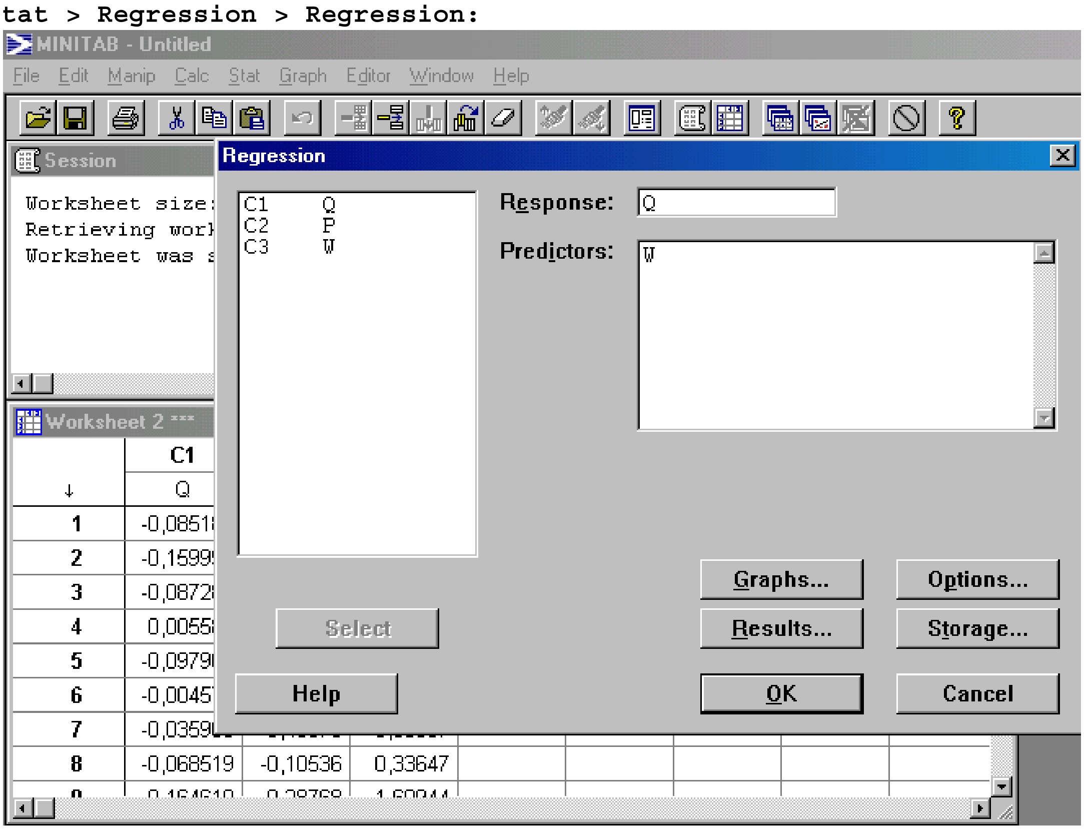
siendo $\boldsymbol{\Pi} = -\boldsymbol{B}^{-1}\cdot\boldsymbol{\Gamma}$ y $\boldsymbol{V_t} = -\boldsymbol{B}^{-1}\cdot\boldsymbol{U_t}$.

Finalmente, el sistema en forma reducida viene dado por las siguientes ecuaciones:

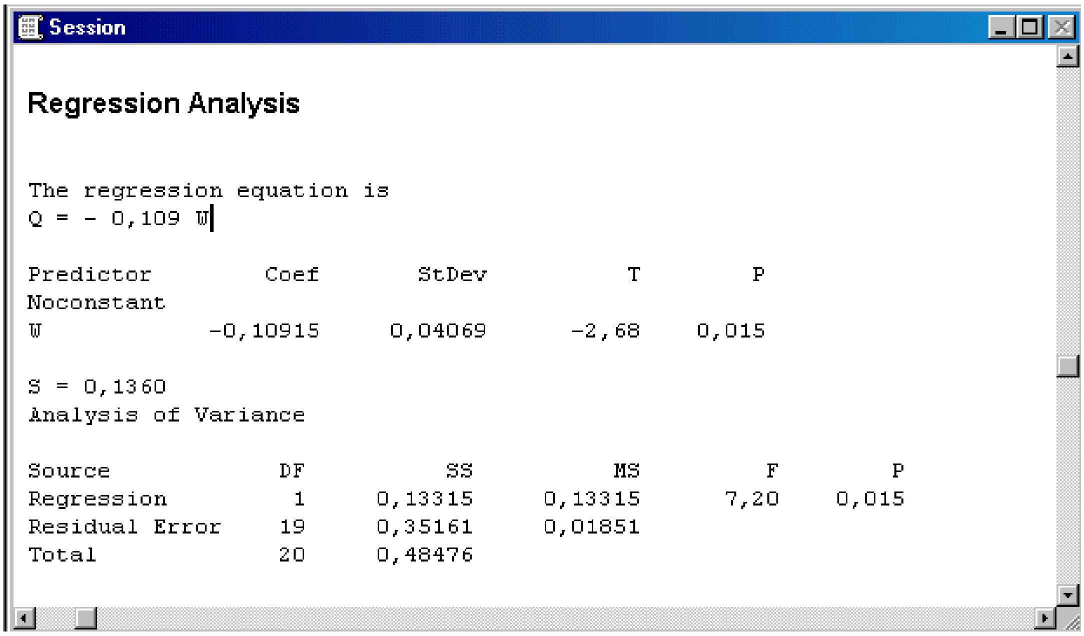
$$\begin{aligned} \boldsymbol{Q_t} &= \boldsymbol{\pi_1 W_t + v_{1t}} \\ \boldsymbol{P_t} &= \boldsymbol{\pi_2 W_t + v_{2t}} \end{aligned} \qquad (20)$$

Además, sabemos que la primera ecuación esta exactamente identificada. Por tanto, aplicaremos MCO sobre las ecuaciones de (21). Para ello, en el entorno del programa Minitab, seleccionamos:

`Stat > Regression > Regression:`



El programa nos proporciona los siguientes resultados la estimación MCO de la primera ecuación :



Regression Analysis

```
The regression equation is
Q = - 0,109 W

Predictor        Coef       StDev          T        P
Noconstant
W            -0,10915     0,04069      -2,68    0,015

S = 0,1360
Analysis of Variance

Source             DF         SS          MS          F         P
Regression          1    0,13315     0,13315       7,20     0,015
Residual Error     19    0,35161     0,01851
Total              20    0,48476
```

A partir de los resultados de estimación, el parámetro $\Pi_1 = -0{,}109$ resulta de nuestro interés.. A continuación, volvemos a emplear el mismo procedimiento a fin de estimar la segunda ecuación. Los resultados de estimación aparecen en el cuadro siguiente:

Regression Analysis

```
The regression equation is
P = 0,155 W

Predictor        Coef       StDev          T        P
Noconstant
W              0,1546      0,1664       0,93    0,365

S = 0,5562
Analysis of Variance

Source             DF         SS          MS          F         P
Regression          1     0,2671      0,2671       0,86     0,365
Residual Error     19     5,8786      0,3094
Total              20     6,1457
```

De la salida de estimación, se deduce que $\Pi_2 = 0{,}155$.

Ahora bien, una vez conocidos los estimadores de π_1 y π_2, procedemos a calcular la elasticidad de demanda estimada a_T:

Puesto que $\Pi = -B^{-1}\cdot\Gamma$, tenemos que $B\cdot\Pi = -\Gamma$, o bien, en forma más detallada,

$$\begin{pmatrix} 1 & -a_T \\ 1 & -b_T \end{pmatrix} \begin{pmatrix} \Pi_1 \\ \Pi_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \eta \end{pmatrix} \qquad (21)$$

O bien,

$$\Pi_1 - a_T \Pi_2 = 0$$
$$\Pi_1 - b_T \Pi_2 = \eta_T \qquad (22)$$

Consecuentemente, $a_T = \Pi_1 / \Pi_2 = -0{,}109/0{,}155 = -0{,}703$. Concluimos, por tanto, que entre la cantidad demandada de naranjas y el precio de equilibrio existe relación inversa. El método de mínimos cuadrados indirectos nos permite estimar sólo la elasticidad de demanda. Puesto que la segunda ecuación de (19) está no identificada, no podemos recuperar, a partir de la forma reducida, el resto los parámetros estructurales b_T y η_T. No obstante, veremos más adelante que, bajo unas condiciones específicas, dichos parámetros pueden estimarse por el método de variables instrumentales.

Ecuaciones identificadas: estimación por MC2E:

El método de **mínimos cuadrados en dos etapas** (MC2E) permite obtener estimadores consistentes para los parámetros estructurales (elementos de las matrices B y Γ) en el caso de ecuaciones sobreidentificadas o exactamente identificadas. El método MC2E consiste en:

- Para cada variable endógena explicativa de la ecuación, hallar la ecuación de regresión de ésta sobre todas las variables exógenas del sistema.

- Con las ecuaciones de regresión obtenidas, hallar los valores estimados para cada variable endógena, y realizar la regresión de la variable endógena dependiente sobre las variables explicativas usando dichos valores estimados (en lugar de los valores observados).

Ejemplo 3: Volvemos a considerar el sistema de ecuaciones simultáneas, definida en (18).

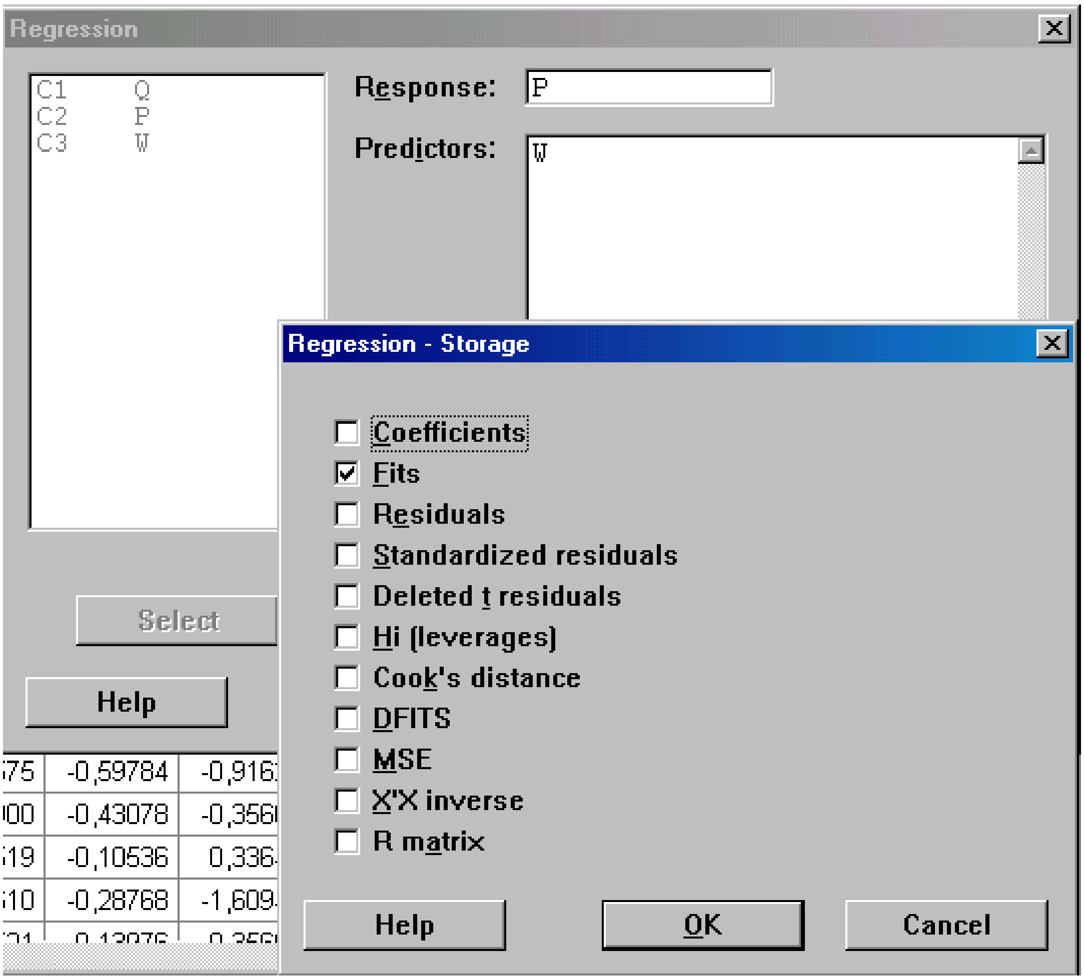
$$Q_t - \alpha P_t = u^D_t$$
$$Q_t - \beta P_t - \eta W_t = u^S_t.$$

Recordemos que la primera ecuación está identificada, mientras que la segunda está no identificada. Por tanto, podemos utilizar el método de MC2E para estimar consistentemente la elasticidad de demanda α.

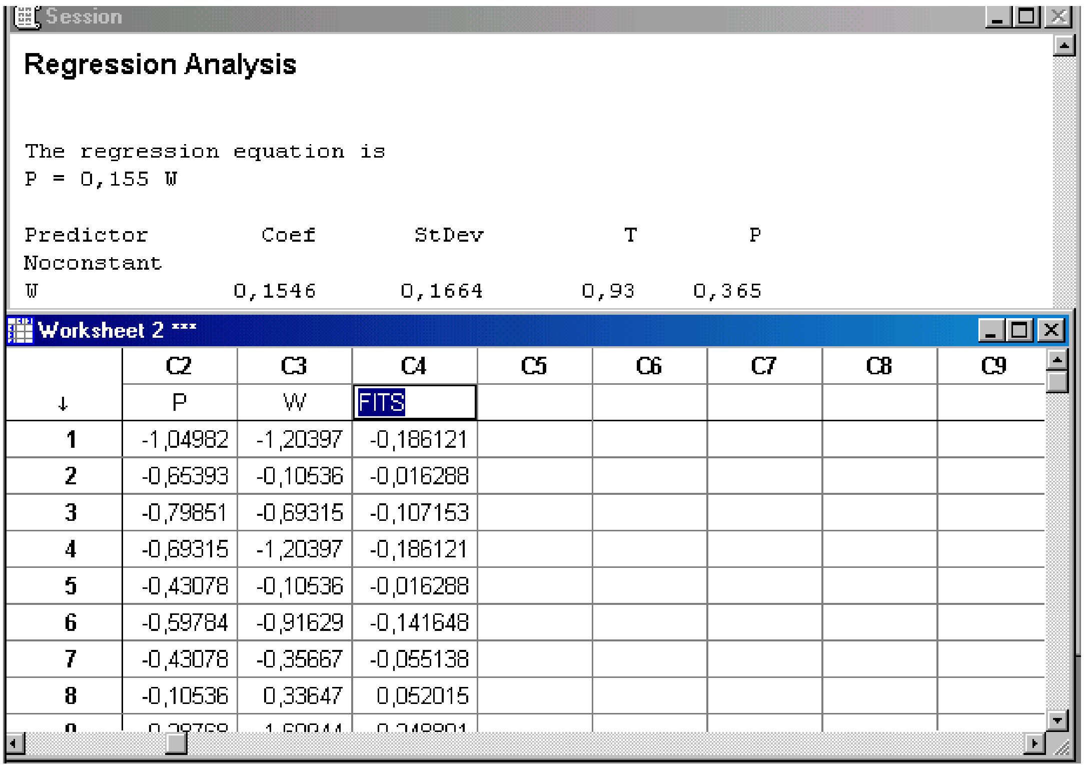
En el sistema, hay sólo una variable endógena explicativa (P_t) y una variable exógena (W_t). De modo que en una primera etapa regresamos el precio de naranjas sobre el indicador de condiciones climatológicas, en la siguiente ecuación de regresión:

Primera etapa: $P_t = \delta W_t + v_t$

Desplegamos `Stat > Regression > Regression > Storage`; una vez dentro marcamos la casilla "Fits" que permite generar, en la hoja de cálculo, los valores ajustados de P_t:



Obteniéndose los siguientes resultados:



En segunda etapa, partimos de la estimación consistente especificada en (17):

$$a^*_T = (1/T)\Sigma P^*_t Q_t/((1/T)\Sigma P^{*2}_t)$$

Esta expresión sólo sirve los propósitos de teoría, ya que en práctica desconocemos los valores de los parámetros β y η necesarios para construir P^*_t. Sin embargo, este término lo podemos sustituir por el estimador consistente de P_t determinado en la primera etapa. Así pues, con el programa Minitab, regresamos cantidad sobre el valor estimado de P_t (llamándolo "PHAT"):

Segunda etapa:

Regression Analysis

```
The regression equation is
Q = - 0,706 PHAT

Predictor         Coef        StDev          T        P
Noconstant
PHAT           -0,7061       0,2632      -2,68    0,015

S = 0,1360
Analysis of Variance

Source            DF          SS          MS          F        P
Regression         1     0,13315     0,13315       7,20    0,015
Residual Error    19     0,35161     0,01851
Total             20     0,48476
```

Así pues, el estimador de MC2E queda $a_{MC2E} = -0,706$.

Ecuaciones identificadas: estimación por VI

El método de **variables instrumentales** (VI) permite obtener estimadores consistentes para los parámetros estructurales (elementos de las matrices **B** y **Γ**) en el caso de ecuaciones sobreidentificadas o exactamente identificadas. Es un caso particular del método de MC2E y se utiliza cuando el número de instrumentos es igual al número de variables endógenas explicativas. Un instrumento resulta válido cuando está correlacionado con una de las variables endógenas y no con el término perturbación. El método VI consiste en:

- Seleccionar tantas variables instrumentales cuantas hay variables endógenas explicativas (correlacionadas con término de perturbación) y variables predeterminadas (incorrelacionadas con término de perturbación) .

- Estimar el modelo en el cual regresamos la variable dependiente sobre las variables explicativas (endógenas y predeterminadas), utilizando las variables instrumentales para corregir la endogeindad de los regresores.

En general, el estimador de variables instrumentales puede determinarse de la siguiente manera:

$$B_{VI} = [\Sigma X_t Z_t']^{-1}[\Sigma X_t Y_t] \quad (23)$$

Donde: Y_t es la variable dependiente, Z_t es el vector de las variables explicativas (endógenas y predeterminadas) y X_t es el vector de las variables instrumentales. La condición imprescindible es que el número de instrumentos, r, sea igual al número de variables explicativas, k.

Ejemplo 4: Volvemos a considerar el sistema de ecuaciones simultáneas, definida en (18).

$$Q_t - \alpha P_t = u^D_t$$

$$Q_t - \beta P_t - \eta W_t = u^S_t.$$

Puesto que la primera ecuación está identificada, podemos utilizar el método de VI para estimar consistentemente la elasticidad de demanda α. A fin de realizar dicha estimación, como instrumento utilizaremos W_t que puede estar correlacionado con P_t pero no con el término de perturbación u^D_t. Así pues, hay una variable explicativa y un instrumento. El estimador de la elasticidad viene dado por la siguiente expresión:

$$a_{VI} = [\Sigma W_t P_t]^{-1}[\Sigma W_t Q_t] \quad (24)$$

La estimación no puede realizarse por vía del Minitab, por lo que emplearemos la hoja de cálculo Excel.

	A	B	C	D	E	F
1	ESTIMACIÓN POR MÉTODO DE VARIABLES INSTRUMENTALES					
2						
3						
4	Año	Qt	Pt	Wt	WtQt	WtPt
5	1981	-0,085181	-1,04982	-1,20397	0,1025554	1,2639518
6	1982	-0,159996	-0,65393	-0,10536	0,0168572	0,0688981
7	1983	-0,087284	-0,79851	-0,69315	0,0605009	0,5534872
8	1984	0,00558	-0,69315	-1,20397	-0,006718	0,8345318
9	1985	-0,097901	-0,43078	-0,10536	0,0103148	0,045387
10	1986	-0,004575	-0,59784	-0,91629	0,004192	0,5477948
11	1987	-0,0359	-0,43078	-0,35667	0,0128045	0,1536463
12	1988	-0,068519	-0,10536	0,33647	-0,023055	-0,03545
13	1989	0,16461	-0,28768	-1,60944	-0,26493	0,4630037
14	1990	0,078521	0,13976	-0,35667	-0,028006	-0,049848
15	1991	0,053681	0,3001	0,09531	0,0051163	0,0286025
16	1992	0,241651	0,22314	-1,20397	-0,290941	-0,268654
17	1993	0,115055	0,37156	-0,10536	-0,012122	-0,039148
18	1994	0,143971	0,40547	-0,35667	-0,05135	-0,144619
19	1995	0,211917	0,43825	-0,69315	-0,14689	-0,303773
20	1996	0,153309	0,53063	-0,10536	-0,016153	-0,055907
21	1997	0,171513	0,55962	-0,35667	-0,061174	-0,1996
22	1998	0,294913	0,50078	-1,20397	-0,355066	-0,602924
23	1999	0,238418	0,66783	-0,51083	-0,121791	-0,341148
24	2000	0,241904	0,85442	-0,22314	-0,053978	-0,190655
25						
26	W'Q=	-1,219833		ALFA_VI =	W'Q/W'P=	-0,706095
27	W'P=	1,7275773				

Estimador de VI de la elasticidad de demanda

Hoja1 Hoja2 Hoja3

En definitiva, el estimador de VI de la elasticidad de demanda es a_{VI} = **-0,706**.

Consideramos ahora estimación de la elasticidad de oferta, β y del efecto de las condiciones climatológicas, η:

$$Q_t = \beta P_t + \eta W_t + u^S_t$$

Ahora bien, suponiendo conocida la elasticidad de demanda, α, podemos generar el término de perturbación de demanda:

$$u^D_t = Q - \alpha P_t$$

Dado que el error está correlacionado con P_t, la variable explicativa endógena, pero no correlacionado con u^S_t, entonces cumple los requisitos de una variable instrumental. Por otra parte, W_t está obviamente correlacionada con W_t, pero no correlacionada con con u^S_t. Así pues, habremos agregado un vector de variables instrumentales, $X_t = (u^D_t; W_t)$ para poder estimar consistentemente los parámetros de la ecuación de oferta:

$$\begin{pmatrix} b_T \\ \eta_T \end{pmatrix} = \begin{pmatrix} U^D{'}P & U^D{'}W \\ W'P & W'W \end{pmatrix}^{-1} \begin{pmatrix} U^D{'}Q \\ W'Q \end{pmatrix}$$

Aunque en práctica desconocemos el valor de la elasticidad de demanda, ya sabemos que este parámetro puede estimarse por el método de variables instrumentales, produciendo el resultado inmediatamente anterior. Precisamente este valor puede utilizarse para generar el residuo de demanda y utilizarlo posteriormente a fin de realizar la estimación de VI de β y η:

$$\begin{pmatrix} b_{VI} \\ \eta_{VI} \end{pmatrix} = \begin{pmatrix} \hat{U}^D{'}P & \hat{U}^D{'}W \\ W'P & W'W \end{pmatrix}^{-1} \begin{pmatrix} \hat{U}^D{'}Q \\ W'Q \end{pmatrix}$$

Puede demostrarse que la estimación de dichos parámetros es consistente [5].

Operando en el entorno de Excel, calculamos los componentes de las matrices ***X'Z*** y de ***X'Y***:

	E	F	G	H	I	J	K	L	M
4	WtQt	WtPt	Udt	UdtPt	UdtWt	UdtQt	Wt^2		
5	0,102555369	1,2639518	-0,826354	0,8675229	0,9949053	0,0703897	1,4495438		
6	0,016857179	0,0688981	-0,621671	0,406529	0,0654992	0,0994648	0,0111007		
7	0,060500905	0,5534872	-0,651032	0,5198556	0,4512629	0,0568247	0,4804569		
8	-0,006718153	0,8345318	-0,483784	0,3353348	0,5824613	-0,0027	1,4495438		
9	0,010314849	0,045387	-0,402032	0,1731872	0,0423581	0,0393593	0,0111007		
10	0,004192027	0,5477948	-0,42665	0,2550685	0,3909352	0,0019519	0,8395874		
11	0,012804453	0,1536463	-0,340031	0,1464784	0,1212787	0,0122071	0,1272135		
12	-0,023054588	-0,03545	-0,142903	0,0150563	-0,048083	0,0097916	0,1132121		
13	-0,264929918	0,4630037	-0,038492	0,0110734	0,0619507	-0,006336	2,5902971		
14	-0,028006085	-0,049848	0,1771916	0,0247643	-0,063199	0,0139133	0,1272135		
15	0,005116336	0,0286025	0,2655516	0,079692	0,0253097	0,0142551	0,009084		
16	-0,290940554	-0,268654	0,3991878	0,0890748	-0,48061	0,0964641	1,4495438		
17	-0,012122195	-0,039148	0,3773764	0,140218	-0,03976	0,043419	0,0111007		
18	-0,051350137	-0,144619	0,4302328	0,1744465	-0,153451	0,061941	0,1272135		
19	-0,146890269	-0,303773	0,5213215	0,2284691	-0,361354	0,1104769	0,4804569		
20	-0,016152636	-0,055907	0,5279338	0,2801375	-0,055623	0,080937	0,0111007		
21	-0,061173542	-0,1996	0,5666047	0,3170833	-0,202091	0,0971801	0,1272135		
22	-0,355066405	-0,602924	0,6484637	0,3247376	-0,780731	0,1912404	1,4495438		
23	-0,121791067	-0,341148	0,709906	0,4740965	-0,362641	0,1692544	0,2609473		
24	-0,053978459	-0,190655	0,8451245	0,7220913	-0,188581	0,204439	0,0497915		
25									
26	W'Q	W'P		Ud'P	Ud'W	Ud'Q	W'W		
27	-1,2198329	1,72758		5,58492	-0,0002	1,36447	11,1753		
28									
29									
30									
31									

Hoja1 / Hoja2 / Hoja3

A continuación, calcularemos la matriz inversa de **X'Z** (*InvX_Z*) y posteriormente la multiplicamos por ***X'Y*** (*X_Y*), obteniendo el vector de los parámetros estimados b_{VI} y η_{VI}. Es importante recordar que para validar fórmulas matriciales con Excel es necesario usar la combinación de teclas **[Control]+[Shift]+[Enter]** (ver *math-block* sobre operaciones con matrices en Excel):

	A	B	C	D	E	F	G	H
25								
26	W'Q=	-1,2198		ALFA_VI =	W'Q/W'P=	-0,7061		
27	W'P=	1,72758						
28	W'W=	11,1753						
29	Ud'P=	5,58492						
30	Ud'W=	-0,0002						
31	Ud'Q=	1,36447						
32								
33								
34	INPUTS:		X_Z=	5,58491709	-0,000163319		X_Y=	1,3644736
35				1,72757729	11,17526505			-1,219833
36								
37			InvX_Z=	0,17905288	2,61674E-06	Estimador de VI de		
38				-0,0276797	0,089482932	Beta (elasticidad		
39						de oferta) y de Eta		
40				BETA_VI=	0,244309736			
41	OUTPUTS:		Est_VI=InvX_Z*X_Y=	ETA_VI=	-0,146922407			
42								
43								
44								

La estimación VI nos proporciona los siguientes resultados:

$\boldsymbol{a_{VI} = -0{,}706}$; $\boldsymbol{b_{VI} = 0{,}244}$ y $\boldsymbol{\eta_{VI} = -0{,}147}$. Concluimos por tanto, que el mercado de naranjas funciona bajo las de leyes de oferta y demanda, la elasticidad de demanda siendo negativa y la elasticidad de oferta, positiva. Por otro lado, el las adversas condiciones meteorológicas producen efecto negativo sobre la oferta de naranjas.

En este math-block hemos analizado los modelos multicecuacionales, problemas de identificación y endogeneidad, y hemos conocido los principales métodos de estimación de estos modelos que dan lugar a los estimadores consistentes, insesgados e eficientes.

BIBLIOGRAFÍA

[1] Artís, M.; Suriñach, J.; et al (2002): "Econometría". Ed. Fundació per a la Universitat Oberta de Catalunya. Barcelona.

[2] Carter, R.; Griffiths, W.; Judge, G. (2000): "Using Excel for Undergraduate Econometrics". ISBN: 0-471-41237-6

[3] Doran, H. (1989): "Applied Regression Analysis in Econometrics". Ed. Marcel Dekker, Inc. ISBN: 0-8247-8049-3

[4] Gujarati, D. (1997): "Econometría básica". McGraw-Hill. ISBN 958-600-585-2

[5] Hamilton, J. D. (1994): "Time Series Analysis". Princeton University Press. Princton, N. J. 08540. ISBN: 0-691-04289-6

[6] Johnston, J. (2001): "Métodos de econometría". Ed. Vicens Vives. Barcelona. ISBN 84-316-6116-X

[7] Kennedy, P. (1998): "A Guide to Econometrics". Ed. MIT Press. ISBN: 0262611406

[8] Novales, A. (1993): "Econometría". McGraw-Hill. ISBN 84-481-0128-6

[9] Pulido, A. (2001): "Modelos econométricos". Ed. Pirámide. Madrid. ISBN 84-368-1534-3

[10] Uriel, E. (1990): "Econometría: el modelo lineal". Ed. AC. Madrid. ISBN 84-7288-150-4

[11] Wooldridge, J. (2001): "Introducción a la Econometría: un enfoque moderno". Ed. Thomson Learning. ISBN: 970-686-054-1

ENLACES

- http://www.feweb.vu.nl/econometriclinks/index.html
  The Econometrics Journal On-Line
- http://www.elsevier.com/hes/books/02/menu02.htm
  Libro on-line: Handbook of Econometrics Vols. 1-5
- http://elsa.berkeley.edu/users/mcfadden/discrete.html
  Libro on-line: Structural Analysis of Discrete Data and Econometric Applications
- http://www.oswego.edu/~kane/econometrics/stud_resources.htm
  Online Resources for Econometric Students
- http://www.econ.uiuc.edu/~morillo/links.html
  Econometric Sources: a collection of links in econometrics and computing. University of Illinois
- http://www.econometrics.net/
  Econometrics, Statistics, Mathematics, and Forecasting
- http://ideas.uqam.ca/EDIRC/ectrix.html
  Economics Departments, Institutes and Research Centers in the World: Econometrics, Mathematical Economics