EL MÉTODO GENERALIZADO DE LOS MOMENTOS

Alfonsa Denia e Ignacio Mauleón*

WP-EC 95-06

* A. Denia: Universidad de Alicante; I. Mauleón: Instituto Valenciano de Investigaciones Económicas y Universidad de Alicante.

Editor: **Instituto Valenciano de Investigaciones Económicas, S.A.**
Primera Edición Junio 1995.
ISBN: 84-482-1008-5
Depósito Legal: V-2104-1995
Impreso por Copisteria Sanchis, S.L.,
Quart, 121-bajo, 46008-Valencia.
Impreso en España.

EL METODO GENERALIZADO DE LOS MOMENTOS

Alfonsa Denia e Ignacio Mauleón

RESUMEN

El método generalizado de los momentos es una técnica de estimación muy general, que permite interpretar la mayor parte de los métodos de estimación conocidos como casos particulares, incluido el de máxima verosimilitud bajo ciertas condiciones. Además, este procedimiento permite desarrollar métodos de estimación nuevos, especialmente adaptados a problemas planteados por la macroeconomía contemporánea. El objeto de este trabajo es presentar, de forma sistemática, el conjunto de resultados, propiedades y aplicaciones más relevantes del método mencionado.

PALABRAS CLAVE: Estimación por simulación, matriz generalizada de covarianzas, método de momentos.

ABSTRACT

The generalized method of moments is a common estimation technique which allows us to interpret most part of the known estimation methods as well as particular cases, including maximum likelihood under certain conditions. This procedure also allows us to develop new estimation methods, specially suited for the present macroeconomic problems. The aim of this paper is to systematically present the most relevant results, properties and applications of this method.

KEY WORDS: Simulation estimation, generalized covariance matrix, method of moments.

I. INTRODUCCION.

Aunque existen importantes trabajos previos, tales como los desarrollados por Sargan (1958) y Malinvaud (1970), por citar algunos de ellos, es a partir fundamentalmente del trabajo de Hansen (1982) cuando el Método Generalizado de los Momentos (MGM) adquiere un importante desarrollo teórico, y también, en especial en los últimos años, un notable incremento en los trabajos empíricos, en particular, los desarrollados en el ámbito de la macroeconomía.

La relativa proliferación de trabajos, tanto teóricos como aplicados, vinculados a este método de estimación puede estar debida, al menos en cierto sentido, a que los estimadores MGM incluyen como casos particulares a buena parte de los estimadores habituales en econometría, tales como los MCO, VI, MC2E, e incluso en determinadas ocasiones al MV. Pero también a que, como veremos en los epígrafes siguientes, este método de estimación alternativo sólo precisa de ciertas condiciones de ortogonalidad, es decir de lo que también se denomina 'condiciones de los momentos', y no de la función de densidad como tal. Como contrapartida, al no utilizar toda la información disponible en la muestra, el estimador MGM puede no resultar eficiente, hecho este no muy limitativo si el objetivo esencial es encontrar un estimador consistente.

La existencia de ciertas condiciones de ortogonalidad se emplea para encontrar el estimador MGM que minimice una determinada función objetivo. Como veremos, estas condiciones de ortogonalidad hacen referencia, en general, al hecho de que la esperanza del producto cruzado entre la perturbación aleatoria y las variables observables en un modelo econométrico es igual a cero. Estas perturbaciones aleatorias, que no son observables, pueden ser reemplazadas por una expresión equivalente, que esté expresada en función del verdadero vector de parámetros y las variables observadas.

La notable proliferación de trabajos en este tema, y a la que ya hemos aludido, probablemente dificulta la tarea de sintetizar los resultados, las aplicaciones y, en general, los muy diversos aspectos que se engloban en un título tan genérico como el aquí presentado. Sin embargo, creemos que un intento de sistematización, aun con el riesgo de obviar alguna cuestión de cierta relevancia, o de no profundizar en exceso en otras, es importante en la medida que puede ayudar a obtener una visión panorámica de este método de estimación, así como de sus aplicaciones más relevantes. Por otra parte, tampoco existen trabajos que presenten, desde nuestro punto de vista, un visión tan global como la que aquí se presenta, (al menos no en castellano). Los trabajos de Hall (1993) y de Ogaki (1993), muy pedagógicos ambos, el primero como una buena introducción al tema, y el segundo por su orientación a las principales aplicaciones de este método, no se corresponden exactamente con la idea de presentar de forma relativamente extensiva este método de estimación, tal y como los propios autores señalan. Es evidente, que también se podría haber optado por centrarnos en aspectos más específicos del mismo; sin embargo, creemos que ambas opciones son igualmente válidas, y desde luego complementarias. Es precisamente el rápido crecimiento de esta literatura lo que nos ha llevado a pensar que podría resultar de utilidad un documento de esta naturaleza.

El trabajo está desarrollado como sigue: En el epígrafe segundo se presenta una breve descripción del método de los momentos, que es un caso específico del MGM, y que sirve a su vez como introducción al MGM. A este método se dedica el apartado III, en el que se incluyen las propiedades básicas del mismo. La sección IV se centra en la estimación de la matriz de covarianzas de los momentos. El epígrafe V incluye los principales contrastes existentes, tanto para detectar una posible sobreidentificación en el modelo, como contrastes de hipótesis generales sobre θ. En la sección VI se incluyen algunas aplicaciones y ejemplos específicos del MGM. En particular, se obtiene el estimador óptimo para el caso lineal cuando las perturbaciones están correlacionadas con los regresores, son heterocedásticas y autocorrelacionadas. La VII discute la relación entre el estimador máximo verosímil y quasi máximo verosímil con el MGM. La sección VIII presenta el método de los momentos simulados. Finalmente, algunas conclusiones y campos de aplicabilidad del método expuesto (en particular, al caso de los modelos econométricos no lineales) se presenta, brevemente, en la sección IX.

II. METODO DE LOS MOMENTOS.

Supongamos que disponemos de un determinado conjunto de observaciones de una variable y_t, cuya ley de probabilidad depende de un vector de parámetros desconocidos, θ, que queremos estimar. Para ello disponemos de dos vías alternativas, i) estimar por Máxima Verosimilitud, para lo que evidentemente debemos conocer por completo la forma de la función de verosimilitud, y seleccionar el resultado que genere un valor más alto para la función de verosimilitud muestral, o bien, ii) utilizar un método alternativo que permita estimar el vector θ, sin necesidad de conocer la función de densidad de la variable aleatoria y_t, pero utilizando alternativamente algunos momentos de la variable aleatoria, así como la idea de que las condiciones de ortogonalidad (de la que se hizo una rápida referencia en la introducción y sobre la que volveremos en breve) sólo pueden ser satisfechas por los verdaderos valores de los parámetros. Para entender mejor esta segunda vía se incluyen los dos ejemplos siguientes.

Ejemplo II.1. Supongamos que Y_t procede de una χ^2 con ν grados de libertad, que disponemos de una muestra de tamaño T y que estamos interesados en estimar precisamente ν. Evidentemente, ν puede ser estimado por MV a partir de su función de densidad. Sin embargo, también sabemos que $E[\chi^2] = \nu$, y que $\hat{\mu} =(1/T)\Sigma y_t$, por lo que esto sugiere que una estimación alternativa a la MV de ν se limita a $\hat{\mu} = \nu$, o $\hat{\nu}(T) = \hat{\mu}$. Es decir, se despeja el parámetro estimado de interés, tras igualar los correspondientes momentos muestrales a los poblacionales.

En esencia esto es lo que se conoce como el **método de los momentos,** y descansa en la idea básica de que en muestras aleatorias, un estadístico muestral convergerá en probabilidad a una constante. Obviamente esta constante, es a su vez, función de parámetros desconocidos, θ (kx1), que caracterizan la función de densidad de la variable y_t, por lo que los momentos poblacionales de la distribución serán función de los mismos, tal que $E[y_t^i] = \mu_i(\theta)$. Para estimar estos k parámetros calculamos los correspondientes momentos muestrales y, tal y como se ha comentado anteriormente, se igualan los momentos muestrales con los poblacionales. Finalmente los parámetros se expresan en función de los momentos. Es decir, el método de los momentos obtiene el $\hat{\theta}$ para el cual se cumple que,

$$\mu_i(\hat{\theta}) = (1/T)\sum_{t=1}^{T} y_t^i \qquad [1]$$

Obsérvese que lo anterior implica que $E\left[y_t^i - \mu_i(\theta)\right] = 0$, que sólo puede ser satisfecha por el verdadero valor del parámetro, y que se denomina 'condición de momento' y [1] que representa la correspondiente 'condición de momento muestral'

$$(1/T) \sum_{t=1}^{T} \left[y_t^i - \mu_i(\hat{\theta})\right] = 0 \qquad [2]$$

siendo el estimador elegido justamente aquel que cumple la condición de momento muestral, y que se tratará de un estimador consistente, puesto que $\mu_i(\hat{\theta}) \xrightarrow{p} \mu_i(\theta)$ para un T suficientemente grande.

Ejemplo II.2. Si Y_t es una variable aleatoria con $E[Y_t] = \mu$, y $\sigma^2 = E[Y_t - \mu]^2$, los estimadores de momentos de μ y σ^2 serán sus correspondientes contrapartidas muestrales. Si además, Y_t es Normal, los estimadores de momentos anteriores coinciden con el Máximo Verosimil.

Bajo otra distribución el resultado general es que los estimadores MV serán más eficientes. Recuérdese que, como se comentó anteriormente y como se aprecia en estos ejemplos, el método de los momentos no precisa información completa de la función de densidad, sino que requiere únicamente ciertas especificaciones de los momentos; de hecho, de esto también se deduce una ventaja, y es que no se requiere hacer ningún supuesto sobre la distribución de la variable aleatoria.

III. METODO GENERALIZADO DE LOS MOMENTOS. DEFINICION GENERAL Y PROPIEDADES.

En los dos ejemplos previos contábamos con el mismo número de momentos poblacionales que de parámetros a estimar, sin embargo, hay casos en los cuales esta igualdad no se cumple, por lo que para utilizar toda la información disponible en la muestra es necesario utilizar una vía que reconcilie las diversas estimaciones que surgen de un sistema sobredeterminado. En estos casos, el método de los momentos no resultará eficiente, siendo necesario recurrir al denominado **Método generalizado de los momentos.**

Ejemplo III.1. Supongamos que ahora, siguiendo con el ejemplo II.2, $\sigma^2 = f(\mu)$. En este caso, dispondríamos de dos momentos y de un único parámetro a estimar. El problema es, por tanto, cómo elegir un estimador de μ, sabiendo que,

$$E\left(\sum_{t=1}^{T} y_t/T\right) = \mu \qquad [3]$$

$$E\left\{\sum_{t=1}^{T} \left(y_t - \Sigma y_t/T\right)^2/(T-1)\right\} = \sigma^2 = f(\mu)$$

Alternativamente podemos definir las funciones:

$$\sum_{t=1}^{T} y_t/T - \mu = m_1(y,\mu)$$

$$\sum_{t=1}^{T} \left(y_t - \mu\right)^2/(T-1) - \sigma^2 = m_2(s^2,\mu) \qquad [4]$$

Y tratar de elegir μ tan próximo como sea posible a las dos opciones. Como $E(m_1) = E(m_2) = 0$, una posibilidad sería minimizar una función criterio tal como, $M = m'A\, m$, donde $m' = (m_1,m_2)$, y A es una matriz simétrica, definida positiva y no singular. En realidad, A es una matriz de ponderaciones que reflejará la importancia que estemos dando a cada una de las dos posibilidades de estimar μ, y que será, en este ejemplo, de orden (2 x 2). Obviamente, si A = I la función criterio que se pretende minimizar es la suma de cuadrados, tipo MCO.

III.1. Función criterio.

Sea z_t un vector de variables aleatorias de orden (hx1) que son observadas en el momento t, y sea θ un vector (kx1) de parámetros desconocidos, y sea $m_j(z_t,\theta)$ un vector (jx1), con $k \leq J$, $E[m_j] = 0$, evaluado en el verdadero valor de θ. Las j filas de $E[m_j] = 0$ se conocen como las condiciones de ortogonalidad.

Sea Z_t la matriz de todas las observaciones de dimensión (Txh). Con carácter general, los momentos poblacionales serán función de los parámetros θ, $\gamma(\theta)$. Por otro lado, los momentos muestrales correspondientes serán distintas funciones g(.) de las observaciones. Por tanto, podemos escribir, ahora, lo siguiente,

$$\sum_t g_j(z_t)/T \; - \; \gamma_j(\theta) = (1/T)\sum_t m_j(z_t,\theta) = \bar{m}_j(Z_T,\theta) \qquad [5]$$

donde se cumplirá que,

$$\text{plim } \bar{m}_j(Z_T,\theta) = 0 \qquad [6]$$

es decir, que los momentos muestrales $\Sigma_t g_t(z_t)/T$ convergerán en probabilidad a los poblacionales $\gamma(\theta_j)$. Nótese, asímismo, que los momentos $\bar{m}_j$ serán $O(1/\sqrt{T})$.

El estimador MGM de θ será aquel que minimice una determinada función criterio. Una posibilidad es min $(\bar{m}_T'\bar{m}_T)$. Sin embargo, en un contexto más general, podemos construir una función criterio usando una matriz A, de tal forma que el objetivo será min $(\bar{m}_T'A_T\bar{m}_T)$. El estimador MGM $\hat{\theta}$ será el valor de θ que minimice

$$C_T = (\bar{m}_T'A_T\bar{m}_T), \quad \text{donde } \bar{m}' = (\bar{m}_1....\bar{m}_J). \qquad [7]$$

siendo A_T una matriz arbitraria de ponderaciones, simétrica, definida positiva, no singular, y $A_T = O(1)$ (plim $A_T = A$), que no depende de θ, pero que puede ser función de los datos.

Cuando $J = k$, el mínimo en la función objetivo C_T se conseguirá simplemente haciendo $\bar{m}_T(\hat{\theta}) = 0$. En este caso es evidente que no importa la elección de la matriz de ponderaciones, puesto que la función criterio será cero para cualquier matriz A. Si $J > k$, entonces hay más condiciones de ortogonalidad que parámetros a estimar y no se cumplirá $\bar{m}_T(\hat{\theta}) = 0$. En general, ahora la proximidad del i-ésimo elemento de $\bar{m}_T(\hat{\theta})$ a cero dependerá de la ponderación correspondiente a la i-ésima condición de ortogonalidad. Es por ello que se necesita una vía que reconcilie las diversas estimaciones que surgen de un sistema sobredeterminado.

III.2. Identificación.

Consideraremos que un modelo está asintóticamente identificado si para $\theta^1 \neq \theta^2$ existe alguna secuencia de funciones criterio C_T, tal que plim $C_T(\theta^1) \neq$ plim $C_T(\theta^2)$, donde al menos uno de los plim existe y es una constante finita. O expresado de otro modo, si plim $C_T(\theta^1) =$ plim $C_T(\theta^2)$, entonces, $\theta^1 = \theta^2$.

Para analizar la identificación supondremos que $\bar{m}$ es derivable en un entorno de θ; esto es,

$$(\partial\bar{m}/\partial\theta) = D \text{ matriz de orden } (Jxk)$$

$$(\partial\bar{m}_i/\partial\theta_j) = D_{ij} \qquad [8]$$

Además, supondremos que D es O(1), y de rango completo $\left(\text{rg(plim } D) = k\right)$ para $\theta \in \mathbb{N}_\theta$. En general, D depende de T, pero esto se omite por simplicidad de notación.

En segundo lugar supondremos que C_T evaluado en θ es dos veces diferenciable. Desarrollando por Taylor la función C(.) evaluada en un valor

$\bar{\theta}$, alrededor del verdadero valor θ, obtenemos:

$$C(\bar{\theta}) = C(\theta) + (\partial C_T/\partial\theta)'_{\theta=\bar{\theta}}(\bar{\theta} - \theta) + (1/2)(\bar{\theta} - \theta)'(\partial^2 C_T/\partial\theta\partial\theta')'_{\theta^*}(\bar{\theta} - \theta) \quad [9]$$

siendo θ^* algún punto intermedio entre θ y $\bar{\theta}$, y donde

$$(\partial C_T/\partial\theta)_\theta = 2\ D'A\bar{m}$$

$$(\partial^2 C_T/\partial\theta\partial\theta_j) = 2[D'A\ (\partial\bar{m}/\partial\theta_j) + (\partial D'/\theta_j)\ A\ \bar{m}]$$

Si $\left(\partial^2\bar{m}/\ \theta_i\ \theta_j\right)$ existe y es O(1), y dado que $\bar{m} = O\left(1/\sqrt{T}\right)$, tendremos lo siguiente,

$$(\partial C_T/\partial\theta)_\theta \longrightarrow 0.$$

$$(\partial^2 C_T/\partial\theta\partial\theta') = 2[D'AD] \quad [10]$$

Por tanto,

$$\text{plim}\ \left[C_T(\bar{\theta}) - C_T(\theta)\right] = (\bar{\theta} - \theta)'(D'AD)_{\theta^*}(\bar{\theta} - \theta) = 0$$

En esta última expresión, tanto D como A corresponden a sus plims, aunque la notación se simplifique por comodidad. Así, A y D no dependen de T, pero si de θ^*, por el momento.

Se puede asegurar que la forma cuadrática anterior es definida positiva, ya que D es de rango completo. Además, A es de rango completo en θ^*, por lo que $D(\bar{\theta} - \theta)$ es igual a cero. Ambas cosas implican que $\bar{\theta} = \theta$, por lo que θ está identificado (más concretamente, está asintóticamente identificado).

Resumen de condiciones:

1.- $\bar{m}$ es derivable dos veces con respecto a θ, cumpliéndose, además que,

1.1.- $\partial\bar{m}/\partial\theta = D_T$ es O(1),
1.2.- plim $D_T = D$; $rg(D) = k$, $\bar{\theta} \in \mathbb{N}_\theta$.
1.3.- $\partial\bar{m}/\partial\theta_i\partial\theta_j$ es O(1) en $\bar{\theta} \in N_\theta$

2.- A_T es una matriz de ponderaciones que satisface

2.1.- es simétrica y definida no negativa.
2.2.- $A_T = 0(1)$
2.3.- rg(plim A_T) es completo (es decir, es no singular).

3.- $\bar{m} = 0\left(1/\sqrt{T}\right)$

III.3. Consistencia.

Supondremos que hemos obtenido el estimador de θ por MGM ($\hat{\theta}$). Como se obtiene minimizando una función criterio que es definida no negativa para todo T, tendremos que,

$$0 \leq \bar{m}\left(Z_T,\hat{\theta}\right)' A_T \bar{m}\left(Z_T,\hat{\theta}\right) \leq \bar{m}\left(Z_T,\theta\right)' A_T \bar{m}\left(Z_T,\theta\right)$$

Por otra parte, y como $\bar{m}\left(Z_T,\theta\right) = 0\left(1/\sqrt{T}\right)$, es inmediato que,

$$0 = \text{plim } \bar{m}\left(Z_T,\hat{\theta}\right)' A_T \bar{m}\left(Z_T,\hat{\theta}\right) = \text{plim } \bar{m}\left(Z_T,\theta\right)' A_T \bar{m}\left(Z_T,\theta\right)$$

y como θ está identificado en el sentido del apartado anterior, concluimos que,

$$\text{plim } \hat{\theta} = \theta$$

es decir, $\hat{\theta}$ es un estimador consistente de θ.

III.4. Distribución asintótica.

Un estimador consistente $\hat{\theta}$ es asintóticamente normal si la secuencia de variables aleatorias $\{T^{1/2}(\hat{\theta} - \theta)\}$ tiende en distribución a una Normal multivariante, con media cero y matriz de covarianzas finita.

Dado que el estimador MGM es aquel que se obtiene al minimizar C_T, y asumiendo que existe un mínimo interior, la minimización se reduce a derivar la función criterio con respecto a θ e igualarla a cero, es decir,

$$\hat{D}'A\hat{m} = 0 \qquad [11]$$

donde ahora, $\hat{m}$ es $\bar{m}(\hat{\theta})$, y $\hat{D}$ representa la matriz de derivadas de la función $\bar{m}$ evaluadas en $\hat{\theta}$.

Linealizando la expresión anterior se obtiene,

$$\hat{D}'A\left(\bar{m} + D_*(\hat{\theta} - \theta)\right) = 0 \qquad [12]$$

donde, $m(\hat{\theta}) = \bar{m}(\theta) + D_{\theta *}(\hat{\theta} - \theta)$

y se cumplirá, lo siguiente,

i) $\theta^* \in (\hat{\theta}, \theta)$, por lo que plim $\theta^* = \theta$.

ii) $D_{\theta}{}^* = \partial\bar{m}/\partial\theta\,|_{\theta=\theta^*}$ [13]

iii) Dado que $\theta^* \xrightarrow{P} \theta$, plim $D_* =$ plim D

Despejando se obtiene la expresión,

$$\sqrt{T}(\hat{\theta} - \theta) = -\left[\hat{D}'AD_*\right]^{-1}\hat{D}'A\sqrt{T}\,\bar{m} = H_T'w,\ w = \sqrt{T}\,\bar{m} \qquad [14]$$

donde $H_T' = -\left[\hat{D}'AD_*\right]^{-1}\hat{D}'A$. Como $\hat{\theta}$ es consistente, plim $H_T = H$, donde todo está evaluado en el verdadero valor de los parámetros, θ.

El algoritmo para resolver iterativamente el estimador de θ, a partir de las ecuaciones no lineales de [11], puede obtenerse inmediatamente a partir del desarrollo de [12], y vendrá dado por,

$$\hat{\theta}_{i+1} = \hat{\theta}_i - (D'AD)_i^{-1}D_i'Am_i \qquad [15]$$

donde el subíndice 'i' indica que la correspondiente expresión está evaluada en el estimador de θ obtenido en la iteración 'i'. Como es habitual, por otra parte, puede añadirse un parámetro γ al segundo término en la expresión anterior, para asegurar que el estimador en el paso i+1, efectivamente conduce a un valor menor de la función criterio, que el estimador del paso i.

Finalmente, por el TCL, y teniendo en cuenta que $\bar{m}$ es la media muestral de un proceso con media cero, se puede asumir, bajo ciertas condiciones de regularidad (tales como la estacionariedad de z_t, la continuidad de $\bar{m}(.)$, junto con algunas restricciones para los momentos) que $\sqrt{T}\,\bar{m} \xrightarrow{D} N(0,\Phi)$. Por otra parte, dado que $\hat{\theta}$ es un estimador consistente, junto con la existencia de D, se cumple, teniendo en cuenta el Teorema de Cramer que,

$$\sqrt{T}(\hat{\theta}-\theta) \xrightarrow{D} H'\sqrt{T}\,\bar{m}$$

es decir, $\sqrt{T}\,(\hat{\theta}-\theta) \xrightarrow{D} N(0,V)$ [16]

donde $V = H'\Phi\, H$

$$H'\Phi\, H = [D'AD]^{-1}D'A\Phi AD\ [D'AD]^{-1} \qquad [17]$$

III.5. Eficiencia.

En este apartado se presenta la forma de seleccionar la matriz A_T, que de lugar a un estimador con las menores varianzas posibles.

La condición para que el estimador $\hat{\theta}$, obtenido minimizando C_T, sea eficiente es que A_T tienda a una matriz que resulte proporcional a la inversa de la matriz de covarianzas asintótica de $\bar{m}$, cuando $T \longrightarrow \infty$. Por tanto, el A óptimo es Φ^{-1}. Sustituyendo en $H'\Phi\, H$, tenemos que

$$H'\Phi\, H = \left[D'\Phi^{-1}D\right]^{-1} \qquad [18]$$

Se trata de probar que esta es la menor matriz de covarianzas asintótica, en el sentido de que la diferencia entre la matriz de covarianzas de un estimador alternativo al óptimo, $\tilde{\theta}$ y la de este, $\hat{\theta}$ es una matriz semidefinida positiva. Es decir, hay que probar alternativamente que sus inversas cumplen la condición de que $D'\Phi^{-1}D - D'AD[D'A\Phi AD]^{-1}D'AD$ es una matriz semidefinida positiva. Para probar esto, efectuaremos las operaciones siguientes,

$$\begin{aligned} &D'\Phi^{-1}D - D'AD[D'A\Phi AD]^{-1}D'AD = \\ &= D'\Phi^{-1/2}[I - \Phi^{1/2}AD[D'A\Phi AD]^{-1}D'A\Phi^{1/2}]\Phi^{-1/2}D = \\ &= \Psi'[I - P(P'P)^{-1}P']\Psi \end{aligned} \qquad [19]$$

siendo $P = \Phi^{1/2}AD$, y que es una forma cuadrática semidefinida positiva, puesto que la matriz asociada a la forma cuadrática es idempotente.

Una forma, quizá más intuitiva, de obtener el resultado anterior, puede derivarse teniendo en cuenta que podemos escribir lo siguiente,

$$\hat{m} = m(\theta) + u \qquad [20]$$

siendo $\hat{m}$ los momentos observados, m(.) una función conocida, y 'u' un error con las propiedades E(u) = 0 y Var(u) = Φ (al menos asintóticamente). El problema de estimación de θ en este contexto puede enfocarse como un caso de mínimos cuadrados generalizados no lineales. Así, los pasos a efectuar son, (i) transformar la ecuación anterior premultiplicándola por $\Phi^{-1/2}$ y, (ii) linealizar $m(\theta)$ alrededor de algún valor de θ. El estimador que se obtiene para θ con este procedimiento conduce, precisamente, al algoritmo propuesto en el apartado anterior (véase [15]) para obtener el estimador MGM de θ.

Finalmente, aunque el óptimo de A sea Φ^{-1} no se puede decir, en general, que el estimador MGM sea globalmente eficiente. El estimador de MV será más eficiente, en principio, aunque hay casos en los que ambos coinciden.

III.6. Combinación óptima de momentos.

Una cuestión que puede surgir de forma natural, es qué momentos o combinaciones de momentos es más conveniente utilizar para estimar los parámetros. Este apartado discute los diversos aspectos relacionados con la misma.

Consideremos ahora el MGM a partir de los momentos transformados $\bar{m}^* = R'\bar{m}$, siendo R una matriz de constantes de orden (Jxr), rg(R) = r, K ≤ r ≤ J. De la simple aplicación directa de los supuestos y de las fórmulas anteriores, la varianza asintótica de $\hat{\theta}$ basada en los momentos transformados será,

$$AV_{m^*}(\hat{\theta}) = \left[D'R[R'\Phi R]^{-1}R'D\right]^{-1} \qquad [21]$$

puesto que ahora: $\partial\bar{m}^*/\partial\theta = R'D \qquad (rxk)$

$$\sqrt{T}\bar{m}^* \xrightarrow{D} N\left(0,\ (R'\Phi R)\right)$$

Procediendo de forma similar al correspondiente apartado de eficiencia, la diferencia de varianzas asintótica entre los momentos originales y los transformados es una matriz semidefinida positiva. Alternativamente consideraremos la diferencia de las inversas, esto es,

$$\begin{aligned}
&\left[AV_{\bar{m}}(\hat{\theta})\right]^{-1} - \left[AV_{\bar{m}^*}(\hat{\theta})\right]^{-1} \\
&= (D'\Phi^{-1}D) - \left[D'R[R'\Phi R]^{-1}R'D\right] \\
&= D'\Phi^{-1/2}\left[I - \Phi^{1/2}R(R'\Phi R)^{-1}R'\Phi^{1/2}\right]\Phi^{-1/2}D \qquad [22] \\
&= \Psi'\left[I - P(P'P)^{-1}P'\right]\Psi
\end{aligned}$$

donde $P = \Phi^{1/2}R$, y que es una forma cuadrática semidefinida positiva, puesto que, de nuevo, la matriz correspondiente es idempotente.

Si r = J entonces existe R^{-1}, y la diferencia anterior es cero, de lo que se deduce que cualquier transformación no singular de los momentos genera idénticamente el mismo estimador. El resultado anterior es igualmente válido si R es una función, no necesariamente lineal, de $\bar{m}$, con tal que sea derivable una vez, y el rango asintótico de la matriz de derivadas en un entorno de θ sea J (aunque en este caso la equivalencia sería asintótica, y los estimadores no serían necesariamente idénticos para un T finito).

Debe precisarse, no obstante, que toda esta discusión se refiere a momentos asintóticamente no dependientes, en el sentido de que su matriz de varianzas y covarianzas debe ser invertible. Así, consideremos, por ejemplo, el siguiente modelo dinámico,

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t$$

donde los ε_t son errores iid $(0,\sigma^2)$. La estimación MCO del vector α depende de los momentos muestrales siguientes: $\Sigma y_{t-1} y_{t-2}/T$, $\Sigma y^2_{t-1}/T$, $\Sigma y^2_{t-2}/T$, $\Sigma y_t y_{t-1}/T$, $\Sigma y_{t-2} y_t/T$. Aunque asintóticamente los momentos $\Sigma y_t y_{t-1}/T$ y $\Sigma y_{t-1} y_{t-2}/T$, por ejemplo, sean equivalentes, para T finito no lo son. Así, el método MGM no es estrictamente equivalente al MCO, para T finito, aunque asintóticamente si se cumple dicha equivalencia (esta matización puede tener importancia en el caso de los momentos simulados; véase la discusión en epígrafes posteriores, y Smith (1993)).

IV. ESTIMACION DE LA MATRIZ DE COVARIANZAS DE LOS MOMENTOS.

En epígrafes anteriores hemos denominado a Φ como la varianza asintótica de la media muestral $\bar{m}(\theta,z)$, por lo que, si $\sqrt{T}\,\bar{m} \xrightarrow{D} N(0,\Phi)$, se cumplirá que $\Phi = \lim_{T\to\infty} T\, E[\bar{m}\bar{m}']$ bajo condiciones bastante generales. Si el proceso $m(\theta,z_t)$ está serialmente incorrelado entonces, Φ podría estimarse consistentemente mediante la expresión, $\Phi^* = (1/T) \sum_{t=1}^{T} [m(\theta,z_t)]\,[m(\theta,z_t)]'$, que requiere conocer θ. Sin embargo, para un $\hat{\theta}$ consistente, y bajo la hipótesis de no correlación serial se cumplirá que $\hat{\Phi} = (1/T) \sum_{t=1}^{T} [m(\hat{\theta},z_t)]\,[m(\hat{\theta},z_t)]' \xrightarrow{P} \Phi$. Si las funciones $m(z_t\theta)$ están correlacionadas serialmente tendremos que, $TE[\bar{m}\bar{m}'] = T^{-1}E \sum_{t=1}^{T} \sum_{s=1}^{T} (f_t f_s')$, donde $f_t = m_t(z_t,\theta)$.

El problema que surge aquí es que el número de autocovarianzas crece cuando $T \longrightarrow \infty$, y, por tanto, no pueden ser estimadas consistentemente. La solución a este problema es no paramétrica, y esquemáticamente puede decirse que la vía es introducir una secuencia de ponderaciones que decrecen cuando p aumenta, siendo p el orden de la autocovarianza, y que por tanto tenderán a cero cuando p tienda a ∞. Esta es la idea básica del estimador de Φ propuesto por Newey-West (1987), y al que vamos a referirnos a continuación.

Bajo el supuesto de autocorrelación serial, y de heterocedasticidad de forma desconocida, Newey-West (1987), proponen el siguiente estimador no paramétrico de Φ:

$$\hat{\Phi} = \hat{\phi}_{0,T} + \sum_{s=1}^{p} [1 - s/(p+1)] \, [\hat{\Phi}_{s,T} + \hat{\Phi}'_{s,T}] \tag{23}$$

$$\hat{\Phi}_{s,T} = (1/T) \sum_{t=s+1}^{T} [m(\hat{\theta},z_t)] \, [m(\hat{\theta},z_{t-s})]' \tag{24}$$

siendo s el orden de la autocorrelación. Finalmente, como m(.) depende de θ, es necesario que $\hat{\theta}$ sea un estimador consistente de aquel. Si no hay autocorrelación, $\hat{\phi}_{0,T}$ es un estimador consistente de Φ en presencia de heterocedasticidad [White (1980)].

El problema que se plantea aquí es cómo determinar la longitud de s, teniendo en cuenta que necesitamos un estimador consistente de θ. En general, la vía propuesta es comenzar con un estimador consistente pero ineficiente de θ, haciendo A = I. Esto permite obtener una estimación inicial de $\hat{\Phi}^0_T$, y por tanto, una primera matriz de ponderaciones $A = (\hat{\Phi}^0_T)^{-1}$, y un nuevo estimador de θ. Sucesivamente se puede ir alargando el valor de p hasta que $\hat{\theta}^{(s)} = \hat{\theta}^{(s+1)}$.

Un ejemplo puede clarificar la fórmula anterior. Así, en el caso del modelo lineal general, sin correlación entre errores y regresores, la fórmula anterior conducirá a la siguiente expresión,

$$\hat{\Phi} = T^{-1} \sum_{t=1}^{T} \left\{ x_t x_t' \hat{u}_t^2 + \sum_{s=1}^{p} \big(1 - s/(p+1)\big)\big(x_t x'_{t-s} + x_{t-s} x_t'\big)\big(\hat{u}_t \hat{u}_{t-s}\big) \right\}$$

que sería un estimador consistente de Φ, bajo heterocedasticidad y autocorrelación de formas desconocidas.

Existen diferentes vías para estimar Φ, y desde luego alguna de ellas puede resultar preferible a la desarrollada por Newey-West en determinados contextos. Sin embargo, el método de Newey y West tiene en general buenas propiedades. Esta circunstancia, junto a su relativa simplicidad lo hace uno de los métodos de estimación más conocidos de Φ.

Recientemente, no obstante, un estimador propuesto por Andrews y Monahan (1992), está ganando popularidad. Para presentar este método comenzaremos por obrservar que la mayoría de los estimadores de Φ se basan en expresiones del tipo,

$$\Phi = (T/(T\text{-}K))\sum_{s=-T+1}^{T-1} k\left(s/S_t\right)\Phi(s) \qquad [25]$$

donde, k(.) es el Kernel, S_t es el parámetro que determina la amplitud de la banda, y que depende de los datos. El factor T/(T-K) es un factor de corrección de grados de libertad para muestras pequeñas, y $\Phi(s)$ ya ha sido definida en [24]. De hecho, el estimador de Newey y West se puede considerar un caso particular de este.

Dentro de este esquema genérico presentado se encuadra el trabajo desarrollado por Andrews y Monahan (1992). La idea central se basa en utilizar un estimador de Kernel preblanqueado, que requiere de tres etapas. En la primera fase se estima un VAR para $m(\hat{\theta},z_t)$ (para el caso MCO sería $x_t\hat{u}_t$,). Evidentemente, el VAR no tiene por qué ser el 'verdadero modelo', pero se usa para eliminar parte de la dependencia temporal de $m(\hat{\theta},z_t)$; de hecho, en situaciones específicas conocidas pueden existir modelos alternativos al VAR que den mejores resultados. En la segunda fase se obtiene una estimación de Φ (expresión [25]) basada en los residuos del VAR, a partir de la expresión del estimador de Kernel, que incluye como posibilidad precisamente, aunque no única, al estimador propuesto por Newey-West. La expresión final para $\Phi(s)$ sería,

$$\hat{\Phi}(s) = \begin{cases} (1/T)\sum_{t=s+1}^{T} \hat{V}^*_t \hat{V}^{*\prime}_{t-s} & s \geq 0 \\ (1/T)\sum_{t=-s+1}^{T} \hat{V}^*_{t+s} \hat{V}^{*\prime}_t & s < 0 \end{cases} \qquad [26]$$

donde $\hat{V}^*$ son los residuos obtenidos en el VAR. Finalmente, el estimador de Kernel preblanqueado mediante el VAR, $\hat{\Phi}^*$ tiene la expresión

$$\hat{\Phi}^* = \hat{D}^{-1}\,\hat{\Phi}\,(\hat{D}^{-1})', \qquad \hat{D} = \left[I - \sum_{r=1}^{b}\hat{A}_r\right] \qquad [27]$$

donde $\hat{A}_r$ son los coeficientes estimados en el VAR ($m_t = \sum_{r=1}^{b} A_r m_{t-r} + V_t$ sería la forma del VAR utilizado para ajustar m_t).

Entre las ventajas que presenta el método desarrollado por Andrews y Monahan destaca el hecho de que permite obtener mejores resultados cuando las perturbaciones presentan estructuras autorregresivas, y, en general, en muestras finitas su comportamiento parece ser más preciso que el estimador más general de Newey y West. Conviene subrayar, por último, que para una matriz A en [7], paramétrica y distinta, aunque próxima, a Φ^{-1}, puede que sea mejor definir el MGM con dicha A para T finito (la matriz de varianzas sería la de [17]).

V. CONTRASTES.

V.1. Contraste de sobreidentificación.

Cuando se dispone del mismo número de parámetros a estimar (k) que de ecuaciones de momentos (J), el modelo está exactamente identificado por lo que la función criterio C_T será mínima con $\hat{m}$ igual a cero, al existir un conjunto de estimaciones para el cual $\hat{m}$ se anula, siendo irrelevante, por tanto, la matriz de ponderaciones.

Sin embargo, si J > k el modelo está sobreidentificado, habiendo más ecuaciones de momentos que parámetros a estimar θ, por lo que las ecuaciones de momentos en estos casos suponen implícitamente un conjunto de (J - k) restricciones de sobreidentificación. El objetivo central de este epígrafe es presentar un contraste de sobreidentificación, es decir, de la validez de las (J-k) ecuaciones adicionales. Vamos a demostrar que este contraste es el siguiente,

$$T(\hat{m}\ A\ \hat{m}) \xrightarrow{D} \chi^2_{(J-k)} \qquad [28]$$

Este contraste sólo es válido si J > k, obviamente, y si $A = \Phi^{-1}$.

Por otro lado sabemos, por la distribución de las formas cuadráticas y si las condiciones de ortogonalidad se cumplen, que

$$\left[\sqrt{T}\ \bar{m}\right]' A \left[\sqrt{T}\ \bar{m}\right] \xrightarrow{D} \chi^2_J \qquad [29]$$

donde $\bar{m}$ está evaluada en el verdadero valor de θ. Sin embargo, la expresión anterior no se mantiene exactamente cuando se evalua en $\hat{\theta}$. La razón es que ahora J combinaciones lineales del vector $\hat{m}$ (kx1) son iguales a cero, por lo que los grados de libertad de la consiguiente forma cuadrática serán ahora (J - k). A continuación se prueba este resultado.

Teniendo en cuenta la expresión [13] $\hat{m} = \bar{m} + D_*(\hat{\theta} - \theta) = \bar{m} + D_*(-H_T\bar{m}) = \left[I - D(D'AD)^{-1}D'A\right]\bar{m}$. Obsérvese que tanto A como D depende de T, y están evaluadas para diferentes estimadores de θ. Como todos ellos son consistentes y el contraste se desarrolla en el contexto del análisis asintótico se han omitido, por simplicidad, estos detalles, sin pérdida de generalidad.

Ahora,

$$\begin{aligned} A^{1/2}\ \hat{m}\sqrt{T} &= \left[I - A^{1/2}D(D'AD)^{-1}D'A^{1/2}\right]\left[A^{1/2}\ \bar{m}\sqrt{T}\right] \\ &= \left[I - \psi(\psi'\psi)^{-1}\psi\right] w = \Psi w \end{aligned} \qquad [30]$$

$$\text{Si } A = \phi^{-1}\text{, entonces, } w = A^{1/2}\ \bar{m}\sqrt{T} \xrightarrow{D} N(0, I_J) \qquad [31]$$

Siendo Ψ una matriz simétrica e idempotente, y con $rg(\Psi) = tr(\Psi) = J - k$, es inmediato que se cumple,

$$T\hat{m}'A\hat{m} = w'\Psi w \xrightarrow{D} \chi^2_{(J-k)} \qquad [32]$$

V.2. Contrastes de restricciones generales sobre θ.

Supongamos que estamos interesados en contrastar un conjunto q de restricciones no lineales sobre θ, de la forma H_0: $g(\theta) = R(\theta) - r = 0$, siendo θ el vector de parámetros a estimar, y $g(\theta)$ una función diferenciable conocida, con $rg\ (\partial g/\partial\theta) = rg\ (G) = q$. Denominaremos a $\tilde{\theta}$ el estimador MGM restringido, y a $\hat{\theta}$ el obtenido sin imponer ninguna restricción.

TEST de WALD

$$W = T\left[g(\hat{\theta})\right]'\left\{AVARE\left(R(\hat{\theta}) - r\right)\right\}^{-1}\left[g(\hat{\theta})\right] \sim \chi^2_q \qquad [33]$$

$$\text{con } AVARE\left(R(\hat{\theta}) - r\right) = \hat{G}'\left(AVARE(\hat{\theta})\right)\hat{G}$$

$$AVARE(\hat{\theta}) = \hat{D}'\hat{\Phi}^{-1}\hat{D}$$

En general, la matriz de covarianzas asintótica AVARE se obtendría a partir de [23] y [24].

CONTRASTE DE MULTIPLICADORES DE LAGRANGE

Para el contraste LR sólo es preciso minimizar la función objetivo sujeta al conjunto de restricciones $R(\theta)$. Cómo el test LR está basado en el gradiente del logaritmo de la función de verosimilitud, evaluado en $\hat{\theta}$, y puesto que a su vez esta función es una función criterio, parece obvio que la vía será, por tanto, obtener el gradiente de la función objetivo $m'\hat{\Phi}\ m$. Así,

$$LR = T\ \left[\bar{m}(\tilde{\theta})'\ \tilde{\Phi}^{-1}\ \tilde{D}\right]\ \left[\tilde{D}'\tilde{\Phi}^{-1}\tilde{D}\right]^{-1}\ \left[\tilde{D}'\ \tilde{\Phi}^{-1}\ \bar{m}(\tilde{\theta})\right] \xrightarrow{D} \chi^2_q \qquad [34]$$

CONTRASTE DEL "TIPO RATIO DE VEROSIMILITUD"

Este contraste es, al igual que el obtenido en el contexto de los modelos estimador por MV, la diferencia entre la función criterio evaluada sin incluir las restricciones e incluyéndolas.

$$LR = T\left[\bar{m}(\tilde{\theta})'\tilde{\Phi}^{-1}\bar{m}(\tilde{\theta}) - \bar{m}(\hat{\theta})'\hat{\Phi}^{-1}\bar{m}(\hat{\theta})\right] \sim \chi^2_q \qquad [35]$$

En el caso del contraste de Wald, la distribución se obtiene por aplicación inmediata de resultados generales. En los otros dos, sin embargo, el resultado no es inmediato, puesto que aquí no se trata del contexto de estimación MV. No obstante, es automático demostrar que la distribución asintótica es la indicada. Para ello basta con observar que el criterio propuesto para obtener el $\hat{\theta}$ por MGM, dividido por dos, cumple la condición siguiente,

$$(\partial C/\partial\theta)\sqrt{T} \xrightarrow{d} N\,(0,\ \text{plim}(\partial^2 C/\partial\theta\partial\theta'))$$

cuando están todas las expresiones evaluadas en el verdadero valor, θ, y $C(\theta) = (m'\Phi^{-1}m)/2$.

Esta es la propiedad fundamental que satisface la función de verosimilitud, y que se requiere para obtener las distribuciones asintóticas analizadas. Por tanto, los contrastes aquí presentados convergen a las distribuciones indicadas.

Es importante subrayar que se requiere que A converja a Φ^{-1}, para que los resultados anteriores sean válidos (véase, también, Newey (1985) y Davidson-Mckinnon (1993)).

VI. APLICACIONES Y EJEMPLOS.

Tal y como se comentó en el epígrafe introductorio, muchos estimadores pueden verse como casos particulares del método generalizado de los momentos. Concretamente, MCO, VI, MC2E, entre otros, e incluso en algunos casos MV, lo son del MGM. En este epígrafe se desarrollaran como ejemplos el método de las Variables instrumentales y los Mínimos Cuadrados en dos Etapas, para más casos concretos se puede consultar Hamilton (1994).

VI.1. Método de las variables instrumentales.

Consideremos el modelo lineal: $Y = X\beta + u$, donde X es una matriz de variables explicativas (Txk), algunas de las cuales son endógenas, de tal forma que plim $[X'u/T] \neq 0$. Sea Z una matriz de variables predeterminadas (Txr) que están correlacionadas con X, pero no correlacionadas con u, de tal foma que,

$$\begin{aligned} &\text{plim } [Z'u/T] = 0 \\ &\text{rg plim } [Z'X/T] = k \\ &\text{rg plim } [Z'Z/T] = r \end{aligned} \qquad [36]$$

Las r condiciones de ortogonalidad son ahora,

$$E\left[z_t(y_t - x_t'\beta)\right] = E\left[z_t u_t\right] = 0 \qquad [37]$$

o en forma matricial $E[Z'u] = 0$, con $\bar{m} = (Z'u/T)$. La función criterio en este ejemplo será, por tanto,

$$C_T(\beta) = \left(u'Z/T\right) A \left(Z'u/T\right) \qquad [38]$$

Con respecto a la identificación,

$$\bar{u} = Y - X\bar{\beta} = u - X(\bar{\beta} - \beta)$$

$$Z'\bar{u} = Z'u - Z'X\ (\bar{\beta} - \beta)$$

$$\begin{aligned} \left(\bar{u}'Z/T\right) A \left(Z'\bar{u}/T\right) = \left(u'Z/T\right) A \left(Z'u/T\right) - 2\left(u'Z/T\right) A \left(Z'u/T\right)(\bar{\beta} - \beta) \\ + (\bar{\beta} - \beta)'\left(X'Z/T\right)A\left(Z'X/T\right)(\bar{\beta} - \beta) \end{aligned}$$

Supongamos que,

$$\text{plim } \left(\bar{u}'Z/T\right) A \left(Z'\bar{u}/T\right) = \text{plim}\left(u'Z/T\right) A \left(Z'u/T\right) \qquad [39]$$

Entonces, dado que plim $[Z'u/T] = 0$, tendremos lo siguiente,

$$0 = (\bar{\beta} - \beta)' \left[\text{plim}\left(X'Z/T\right) A \ \text{plim}\left(Z'X/T\right)\right](\bar{\beta} - \beta) \qquad [40]$$

Dado que el rango de la forma cuadrática es k, la única solución posible es que $\bar{\beta} = \beta$.

Por lo que respecta a la consistencia, sabemos que,

$Z'\hat{u} = Z'u - Z'X\ (\hat{\beta} - \beta)$, por lo que sustituyendo y aplicando un argumento similar al anterior,

$$\text{plim}\left(\hat{u}'Z/T\right)A\left(Z'\hat{u}/T\right) = \text{plim}(\hat{\beta} - \beta)'\left(X'Z/T\right) A \left(Z'X/T\right)](\hat{\beta} - \beta) \qquad [41]$$

y, consecuentemente, plim $\hat{\beta} = \beta$

La solución se obtiene derivando C_T con respecto a $\hat{\beta}$ e igualando a cero, de donde,

$$\hat{\beta} = [X'ZAZ'X]^{-1}X'ZAZ'Y \qquad [42]$$

Según se elijan Z y A, contamos con un amplio conjunto de estimadores conocidos que son casos particulares del VI. Así, por ejemplo, si Z = X y $\hat{A} = \left(X'X/T\right)^{-1}$ estamos ante el estimador MCO. Si Z = X, pero hay heterocedasticidad, el estimador MGM es el MCO, pero la matriz de covarianzas asintótica es la dada por White, mientras que corresponderá a la expresión dada por Newey-West si hay autocorrelación. Por contra, si hay autocorrelación entre X y u, Z representa el conjunto de k variables instrumentales. Si J > k, pero se mantienen los supuestos acerca de u, el estimador resultante es el MC2E.

Ejemplo VI.1. Supongamos que $y_t = x_t'\beta + u_t$, y que también ahora se cumple que las J condiciones de ortogonalidad son, $E\left[z_t(y_t - x_t'\beta_0)\right] = 0$, siendo β_0 el verdadero valor del vector de parámetros. Pero se sabe que J > k. Es decir, el modelo está sobredeterminado. Ahora no se puede hacer $\bar{m}(\hat{\theta}) = 0$.

Obsérvese que si el modelo está exactamente identificado entonces es suficiente con hacer $\bar{m}(\hat{\beta}) = 0$, por lo que en este caso,

$$\bar{m}(\hat{\beta}) = Z'(Y - X\hat{\beta}) = 0, \text{ ahora } \hat{\beta} = (Z'X)^{-1}Z'Y$$

puesto que en este caso A es irrelevante.

Pero como J > k, ahora, debemos minimizar C_T, lo que, como vimos en la expresión [9], supone igualar a cero la expresión,

$$(\partial C_T/\partial\theta)_\theta = 2\ D'A\bar{m}$$

evaluada en $\hat{\theta}$, y que en este ejemplo concreto será igual a,

$$0 = \left\{-(1/T)\sum_{t=}^{T} x_t z_t'\right\} \hat{A} \left\{(1/T)\sum_{t=1}^{T} z_t\left(y_t - x_t'\hat{\beta}\right)\right\}$$

En ausencia de autocorrelación y de heterocedasticiad, una estimación natural de A sería, en este caso, $\left(\hat{\sigma}^2(1/T)\Sigma z_t z_t'\right)^{-1}$, con $\hat{\sigma}^2 = (1/T)\Sigma(y_t - x_t'\hat{\beta})^2$. Sustituyendo en la expresión anterior,

$$0 = \hat{\gamma}\left\{(1/T)\sum_{t=1}^{T} z_t\left(y_t - x_t'\hat{\beta}\right)\right\}$$

con $\hat{\gamma} = \left\{\sum_{t=1}^{T} x_t z_t'\right\}\left\{\sum_{t=1}^{T} z_t z_t'\right\}^{-1}$, que evidentemente, se trata del vector de coeficientes estimados al regresar x_t sobre z_t.

Definiendo $\hat{x}_t = \hat{\gamma} z_t$ se llega a que

$$\hat{\beta} = \left\{\sum_{t=1}^{T} \hat{x}_t x_t'\right\}^{-1} \left\{\sum_{t=1}^{T} \hat{x}_t y_t\right\}$$

que corresponde al estimador MC2E.

Volviendo al desarrollo del tema, y por lo que respecta a la distribución asintótica, sustituyendo en [41] $Y = X\beta + u$, se llega a la expresión

$$(\hat{\beta} - \beta)\sqrt{T} = \left[(X'Z/T)\ A\ (Z'X/T)\right]^{-1}(X'Z/T)\ A(Z'u/\sqrt{T}) = H_T'w \qquad [43]$$

donde, $w = Z'u/\sqrt{T}$ se distribuye asintóticamente como una $N(0, \text{plim}(Z'\Omega Z/T))$ y, consecuentemente,

$$(\hat{\beta} - \beta)\sqrt{T} \xrightarrow{A} N\left(0, \text{plim } T(X'Z\ A\ Z'X)^{-1}X'ZA(Z'\Omega Z)AZ'X(X'ZAZ'X)^{-1}\right) \quad [44]$$

Nótese que no se está suponiendo que $\text{Var}(u) = \sigma^2 I$; de este modo la discusión es aplicable a los casos en que existe heterocedasticidad y autocorrelación serial en los residuos. Tal situación suele darse, generalmente, en los modelos con expectativas racionales sobre variables futuras (véase, también, en páginas posteriores).

La matriz A eficiente será $(Z'\Omega Z/T)^{-1}$, en cuyo caso obtenemos lo siguiente,

$$\hat{\beta} = [X'Z(Z'\Omega Z)^{-1}Z'X]^{-1}X'Z(Z'\Omega Z)^{-1}Z'Y \quad [45]$$

y su distribución asintótica vendrá dada por,

$$(\hat{\beta} - \beta)\sqrt{T} \xrightarrow{A} N\left(0, \text{plim } T(X'Z(Z'\Omega Z)^{-1}Z'X)\right)^{-1} \quad [46]$$

VI.2. Definición alternativa del MGM.

De igual forma que en el apartado anterior hemos definido el MGM por analogía al MCO, podemos construir los momentos por similitud con los MCG, es decir,

$$\bar{m} = Z'\Omega^{-1}u/T$$

$$\sqrt{T}\ \bar{m} \xrightarrow{D} N\left(0, \text{plim } (Z'\Omega^{-1}Z/T)\right)$$

La función objetivo será:

$$C_T = \left[u'\Omega^{-1}Z(Z'\Omega^{-1}Z)^{-1}Z'\Omega^{-1}u/T\right]$$

y la matriz de covarianzas asintótica del estimador,

$$AVAR(\hat{\theta}) = \text{plim } T\left(X'\Omega^{-1}Z(Z'\Omega^{-1}Z)^{-1}Z'\Omega^{-1}X\right)^{-1}$$

En lo que sigue vamos a comprobar que esta vía alternativa de definir los momentos genera un estimador que no domina ni es dominado por el estimador anterior. Para ver esta cuestión basta con analizar la diferencia entre las varianzas asintóticas de ambos estimadores alternativos, o, equivalentemente, entre sus inversas. Es decir,

$$X'Z(Z'\Omega Z)^{-1}Z'X - X'\Omega^{-1}Z(Z'\Omega^{-1}Z)^{-1}Z'\Omega^{-1}X = X'\Psi X$$

donde $\Psi = \left[Z(Z'\Omega Z)^{-1}Z' - \Omega^{-1}Z(Z'\Omega^{-1}Z)^{-1}Z'\Omega^{-1}\right]$

El paso siguiente es analizar Ψ. Para ello consideraremos como se comportan $Z'\Psi Z$ y $Z'\Omega\Psi\Omega Z$. Así, y en primer lugar,

$$\begin{aligned} Z'\Psi Z &= Z'Z(Z'\Omega^{-1}Z)^{-1}Z'Z - Z'\Omega^{-1}Z = \\ &= Z'\Omega^{-1/2}\left[\Omega^{1/2}Z(Z'\Omega Z)^{-1}Z'\Omega^{1/2} - I\right]\Omega^{-1/2}Z = \\ &= Z'\Omega^{-1/2}\left[G(G'G)^{-1}G' - I\right]\Omega^{-1/2}Z \end{aligned}$$

y $\left[G(G'G)^{-1}G' - I\right]$ es semidefinida negativa.

En segundo lugar,

$$\begin{aligned} Z'\Omega\Psi\Omega Z &= Z'\Omega Z - Z'Z(Z'\Omega^{-1}Z)^{-1}Z'Z = \\ &= Z'\Omega^{1/2}\left[I - \Omega^{-1/2}Z(Z'\Omega^{-1}Z)^{-1}Z'\Omega^{-1/2}\right]\Omega^{1/2}Z = \\ &= Z'\Omega^{1/2}\left[I - F(F'F)^{-1}F'\right]\Omega^{1/2}Z \end{aligned}$$

y $\left[I - F(F'F)^{-1}F'\right]$ es semidefinida positiva.

Por tanto, Ψ no está definida y ninguno de los dos estimadores domina al otro. La vía para compararlos requiere la estimación previa de ambas varianzas, lo que no es más que una aproximación, puesto que lo que estamos comparando son las varianzas estimadas y no las asintóticas.

El problema en discusión posee, no obstante, una solución óptima, aunque no siempre sea aplicable (resultados sobre eficiencia puede encontrarse en Hansen (1985); el análisis que se presenta aquí es una aplicación de estos resultados más generales).

Definamos ahora, $X = \bar{x} + v$, con plim $\bar{x}'v/T = 0$. Si $Z = \bar{x}$, entonces, $Z'X/T = \bar{x}'\bar{x}/T + \bar{x}'v/T \longrightarrow \bar{x}'\bar{x}/T$, $Z'\Omega^{-1}X/T \longrightarrow \bar{x}'\Omega^{-1}\bar{x}/T$, y la varianza asintótica del estimador MGM análogo al MCG será, ahora, plim $T\left(\bar{x}'\Omega^{-1}\bar{x}\right)^{-1}$.

Consideremos, alternativamente, otro estimador de momentos (θ), basado en instrumentos arbitrarios 'M' , tal que cumplen,

$$M'u/T \longrightarrow 0,\ M'X/T \longrightarrow M'\bar{x},$$
$$\left(M'u/\sqrt{T}\right) \longrightarrow N\left(0,\ (M'\Omega M)\right)$$

con $AVAR(\theta) = \text{plim } T\left(\bar{x}'M\ (M'\Omega M)^{-1}M'\bar{x}\right)^{-1}$, por lo que, ahora la diferencia entre las covarianzas asintóticas de ambos es semidefinida positiva, esto es, $\bar{x}'\Omega^{-1}\bar{x} - \bar{x}'M\ (M'\Omega M)^{-1}M'\bar{x}$ es una matriz semidefinida positiva, por lo que el estimador de los momentos basados en $\left(\bar{x}\Omega^{-1}u/T\right)$ es óptimo dentro de la clase de estimadores MGM. Un estimador equivalente asintóticamente se obtiene resolviendo $\bar{x}'\Omega^{-1}\hat{u} = 0$. En todo caso, esta argumentación proporciona el estimador óptimo, dentro de la clase MGM. El problema básico aquí, de todas formas, es que en algunos casos se desconoce E[X].

La discusión anterior es aplicable a modelos que incumplen dos de los supuestos clásicos fundamentales, es decir, plim $X'u/T \neq 0$, $Var(u) \neq \sigma^2 I$. Está claro que la solución requiere una corrección doble, que consistirá en aplicar variables instrumentales y mínimos cuadrados generalizados. El problema es en qué orden, y el método que se acaba de exponer, responde precisamente a esa pregunta. El doble problema mencionado surge naturalmente, por ejemplo, en los modelos de expectativas racionales. Así,

$$y_t = a_1\ E[x_t \,|\, I_{t-1}] + a_2 E[x_{t+1} \,|\, I_{t-1}] = a_1 x_t + a_2 x_{t+1} + u_t$$

donde 'u_t', por construcción, estará correlacionado con x_t y x_{t+1}, además, de presentar autocorrelación serial de algún tipo, que dependerá del modelo específico que siga x_t.

Otro criterio de elección se basa en el conocimiento o desconocimiento de Ω. Cuando nos encontramos ante un problema de heterocedasticidad, y desconocemos la forma de Ω podemos estimar no paramétricamente $Z'\Omega Z$.

Como,

$$Z'\Omega Z/T = 1/T \sum_{t=1}^{T} [(Z_t Z_t')\sigma_t^2]$$

es un estimador consistente de $E[(Z_t Z_t')u_t^2]$, y se cumple que,

$$\text{plim } 1/T \sum_{t=1}^{T} [(Z_t Z_t')u_t^2] = \text{plim } (Z'\Omega Z/T)$$

White (1980), si desconocemos u_t, utilizaremos en su lugar un estimador consistente $\hat{u}_t$. Así, es posible aplicar el MGM análogo al MCO.

Por otra parte,

$$1/T \sum_{t=1}^{T} z_t u_t (1/u_t^2) = (Z'\hat{\Omega}^{-1}u/T)$$

que se trata de un estimador consistente de $E[(z_t u_t)(1/u_t^2)]$, pero que evidentemente no es cero. Esto indica que el estimador MGM análogo al MCG no es consistente.

En el caso en el que $\text{plim}(X'u/T) = 0$, la comparación la debemos plantear de otra forma. Si Ω es conocida o puede estimarse, el estimador óptimo es el MCG, es decir, el basado en $(X'\Omega^{-1}u/T)$. Por eso, en presencia de heterocedasticidad conocida el MGM análogo al MCG es superior al MGM análogo al MCO.

Si Ω no puede estimarse, el estimador MGM análogo al MCG es inconsistente, como acaba de verse. Por otra parte, el estimador MGM análogo al MCO, que se basaría en minimizar el criterio $u'X(X'\Omega X)^{-1}X'u$, conduce a la solución MCO básica. Una alternativa es utilizar un estimador MGM análogo al MCO, pero con un conjunto de instrumentos más amplio que X, por ejemplo $Z = (X,X^2)$, donde X^2 incluye todos los productos cruzados de las variables en X. En este caso, la comparación de varianzas conduce a lo siguiente,

$$X'Z(Z'\Omega Z)^{-1}Z'X - X'X(X'\Omega X)^{-1}X'X$$
$$= X'\Omega^{-1/2}[\Omega^{1/2}Z(Z'\Omega Z)^{-1}Z'\Omega^{1/2} - \Omega^{1/2}X(X'\Omega X)^{-1}X'\Omega^{1/2}]\Omega^{-1/2}X$$
$$= X'\Omega^{-1/2}[\psi(\psi'\psi)^{-1}\psi' - \psi S(S'\psi'\psi S)^{-1}S\psi]\Omega^{-1/2}X$$

donde $X = ZS$ y $\psi = \Omega^{1/2}Z$ y S es una matriz de selección. La matriz de la forma cuadrática es $(P_\psi - P_{\psi S})$ y como $P_\psi P_{\psi S} = P_{\psi S}$ es idempotente y semidefinida positiva. Así, finalmente, este estimador proporciona una varianza menor, y es superior al MCO básico (Cragg(1983)). Otra alternativa posiblemente mejor en muestras finitas, consiste en utilizar una matriz Σ en lugar de Ω para aplicar la fórmula del MGM. Si Σ es, por ejemplo, una matriz paramétrica sencilla, y próxima a Ω, para T finito el estimador de los parámetros puede ser mejor. En este caso, la matriz de varianzas y covarianzas será la equivalente a la dada en [17], que, como siempre, puede estimarse no paramétricamente sin dificultad.

VI.3. Ejemplos de contrastes.

En el modelo lineal de regresión convencional, donde se cumplen todos los supuestos clásicos, excepto el de no correlación entre regresores y errores, consideraremos el contraste de sobreidentificación, en un contexto de estimación mediante variables instrumentales. Así,

$$\begin{aligned}\hat{u} &= Y - X\hat{\beta} = u - X(\hat{\beta} - \beta) = u - X(X'P_zX)^{-1}X'P_zu \\ &= \left(I - X(X'P_zX)^{-1}X'P_z\right)u\end{aligned}$$

donde $P_z = Z(Z'Z)^{-1}Z'$. Consideremos ahora,

$$\begin{aligned}&(Z'Z)^{-1/2}Z'\hat{u}/\sigma = \\ &\left[I - (Z'Z)^{-1/2}Z'X\left(X'Z(Z'Z)^{-1}Z'X\right)^{-1}X'Z(Z'Z)^{-1/2}\right](Z'Z)^{-1/2}Z'u/\sigma = \Psi w\end{aligned}$$

donde $w = (Z'Z)^{-1/2}Z'u/\sigma \xrightarrow{d} N(0,I_J)$. Por otra parte Ψ es una matriz simétrica, idempotente y de rango J - k, de modo que, finalmente,

$$\begin{aligned}&T\left(\hat{u}'Z/T\right)\left(Z'Z/T\right)^{-1}\left(Z'\hat{u}/T\right)/\sigma^2 \\ &= \left(\hat{u}'Z/\sigma\right)\left(Z'Z\right)^{-1/2}\left(Z'Z\right)^{-1/2}\left(Z'\hat{u}/\sigma\right) = w'\Psi w \xrightarrow{d} \chi^2_{J-k}\end{aligned}$$

que es el resultado obtenido de forma general anteriormente.

En el caso del contraste análogo al de cociente de MV, la expresión sería,

$$(\tilde{u}'P_Z\tilde{u} - \hat{u}'P_Z\hat{u})/\sigma \xrightarrow{d} \chi^2_{J-k}$$

siendo $\tilde{u}$ y $\hat{u}$, respectivamente, los residuos obtenidos con el estimador de variables instrumentales restringido y sin restringir.

Finalmente, en el caso del contraste basado en los multiplicadores de Lagrange, la expresión concreta sería,

$$\left(\tilde{u}'P_ZX(X'P_ZX)^{-1}X'P_Z\tilde{u}\right)/\sigma \xrightarrow{d} \chi^2_{J-k}$$

VII. MAXIMA VEROSIMILITUD, MGM Y OTROS PRINCIPIOS DE ESTIMACION.

Supongamos que disponemos de un conjunto de observaciones muestrales de tamaño T de una variable aleatoria y_t, con función de densidad condicionada para la t-ésima observación $f\left(y_T|y_{t-1},....\theta\right)$, y $L = \Pi f\left(y_T|y_{t-1},....\theta\right)$ que asumimos cumple las condiciones habituales de regularidad. Por tanto se cumplirá $\int\partial L/\partial\theta \; dy_t = 0$, puesto que hemos asumido que el dominio de y es independiente de θ, y que $\int\partial L/\partial\theta$ existe. Por otro lado sabemos que $\int\partial L/\partial\theta \; dy_t = \int\left(\partial \ln L/\partial\theta\right) L \; dy = 0$.

Llamaremos a $\partial \ln L/\partial\theta = m$, y como $E\left[\partial \ln L/\partial\theta\right] = \int\left(\partial \ln L/\partial\theta\right)L dy$ tenemos que $E[m] = 0$, ya se trate de la esperanza condicionada o de la no condicionada. Con $J = k$, $E[m] = 0$ representa el conjunto de las condiciones de ortogonalidad que se requieren para estimar el vector θ. Como, además, estamos asumiendo el mismo número de condiciones de ortogonalidad que de parámetros a estimar, y basándonos en los criterios que establece el MGM, parece natural pensar que la solución para θ sería $\bar{m} = 0$, es decir;

$$(1/T)\sum_{i=1}^{T} m(\theta) = 0 \qquad [47]$$

que coincide con las condiciones necesarias de primer orden para un máximo en la función de verosimilitud $\partial \text{Ln } L(\theta)/\partial\theta = 0$, suponiendo que esta función es diferenciable y que su máximo no ocurre en un extremo de su campo de definición.

Por lo que respecta a la obtención de la matriz de covarianzas de θ, según la expresión [17] $V = [D'\Phi^{-1}D]$. Cómo $D = \left(\partial\bar{m}/\partial\theta\right)$, la matriz de covarianzas del estimador MV puede aproximarse mediante la expresión:

$$AV(\hat{\theta}) = (1/T)\left\{\hat{D}'\hat{\phi}^{-1}\hat{D}\right\}^{-1} \qquad [48]$$

con,

$$\hat{D} = (\partial\bar{m}/\partial\theta)\Big|_{\hat{\theta}} = (1/T)\sum \frac{\partial m(\theta)}{\partial\theta}\Big|_{\hat{\theta}} = (1/T)\sum \frac{\partial^2 \text{Ln } f(\theta, y_t \ldots)}{\partial\theta\ \partial\theta'}\Big|_{\hat{\theta}}$$

$$\hat{\phi} = (1/T)\sum_{t=1}^{T}\left[m(\hat{\theta})\right]\left[m(\hat{\theta})\right]' \text{ en ausencia de correlación serial.}$$

Finalmente, si el modelo está correctamente especificado podemos esperar que $\hat{\phi}$ converja en probabilidad a la matriz de información, I, y que $\hat{D}$, asímismo, converja a I, por lo que a partir de [46], y dando por supuesto todas las especificaciones previamente señaladas, tendremos que,

$$\sqrt{T}\left(\hat{\theta} - \theta\right) \xrightarrow{D} N(0, I^{-1}) \qquad [49]$$

En definitiva, el estimador MV es un caso particular del estimador MGM, exactamente identificado. En general, no obstante, las funciones utilizadas para definir los momentos muestrales en la aplicación del MGM no coinciden con las del MV, y así, serán menos eficientes. En ciertos casos, de todos modos, ambos pueden coincidir: un ejemplo de esto es el caso en el que la función de verosimilitud pueda escribirse como una función monótona del criterio MGM.

En ciertos casos se escoge una función de densidad incorrecta para las observaciones, o bien se desconoce si la elección ha sido adecuada. La función de densidad en estos casos se escoge por razones de conveniencia, y porque se considera razonable en algún sentido. El estimador obtenido es entonces el de pseudo máxima verosimilitud (PMV). Gourieroux, et al (1984), analizan la consistencia y la distribución asintótica de estos estimadores con bastante generalidad. Como los estimadores serán generados por alguna expresión del tipo [47], estos estimadores pueden considerarse, asímismo, como un caso particular del método de los momentos. No obstante, la matriz asintótica de varianzas y covarianzas no vendrá dada por [49] generalmente, sino obviamente por [48]. Como en el caso de la comparación de los métodos MGM y MV, las funciones utilizadas para definir los momentos no tienen por qué coincidir y, por consiguiente, no es posible realizar una comparación general sobre la eficiencia relativa de ambos estimadores. Se puede afirmar, no obstante, que los estimadores MGM serán más robustos que los PMV, en general. En este sentido, no es ilógico considerar el estimador MGM como preferible al PMV, ya que es tanto o más robusto, y el PMV no tiene por qué ser más eficiente, en general.

Los procedimientos de estimación indirecta propuestos por Gourieroux et. al. (1993) por un lado, y Gallant y Tauchen (1994), pueden considerarse como una prolongación del procedimiento PMV, en cierto sentido. Estos autores proponen maximizar un criterio determinado, por ejemplo una función de verosimilitud, seleccionada por algún motivo intuitivo de proximidad con la verdadera verosimilitud, pero que no coincidirá con ella (hasta aquí el planteamiento sería similar al caso PMV). Esta función se optimiza con respecto a un conjunto de parámetros auxiliares, y si estos parámetros son funciones bien definidas de los parámetros de interés, es posible generar estimadores consistentes y asintóticamente Normales de estos últimos, bajo condiciones muy generales. Todas las operaciones a las que se ha hecho referencia se llevan a cabo mediante simulaciones (optimización e identificación de la relación funcional entre parámetros auxiliares y de interés). La aplicabilidad de este método es, en principio, tan amplia como la del método de los momentos simulados. No obstante, la elección del criterio a optimizar puede ser arbitraria en muchos casos, mientras que la elección de los momentos, como pieza esencial de información a ajustar, sigue siendo una elección natural en ausencia de otro tipo de información.

Otros principios de estimación relacionados con el MGM son el de distancia mínima (Malinvaud, (1970)) y el analógico (Manski, (1994)). El principio de distancia mínima sugiere obtener estimadores de modo que se minimice la distancia (geométrica) entre un vector de características poblacionales, funciones a su vez de los parámetros de interés, y funciones equivalentes de los datos muestrales. Es inmediato ver que, en este sentido, el estimador de distancia MGM es un caso particular del estimador de distancia mínima. No obstante, si tenemos en cuenta que el estimador MGM puede definirse equivalentemente mediante una transformación funcional no singular de los momentos, la diferencia entre los dos principios de estimación es mínima. Por último, el principio analógico propone obtener estimadores de alguna característica poblacional (no necesariamente un parámetro) reproduciendo en la muestra las operaciones análogas a las que definen en la población la característica de interés. Un ejemplo es la estimación de la mediana y, en general, de los quantiles, sin efectuar ningún supuesto previo acerca de la forma de la función de densidad subyacente. Enunciado de esta manera, el principio analógico es más amplio que el MGM y que el MV, y posiblemente sea el principio de estimación más general que existe. No obstante, generalidad y utilidad no son dos propiedades que vayan necesariamente unidas. De hecho, las propiedades generales del principio analógico no están bien desarrolladas, y es dudoso que puedan llegar a estarlo. Tampoco es del todo claro, además, que este principio cuente con un excesivo interés. De hecho, la mayor parte de los problemas de estimación conocidos pueden tratarse con los procedimientos más concretos ya discutidos, y cuyas propiedades, por otra parte, son suficientemente bien conocidas. Además, incluso estimadores de características como la mediana o los quantiles pueden obtenerse a partir de un estimador de momentos: para ello se puede suponer una función de densidad suficientemente general, por ejemplo, un producto de la Normal por una suma de polinomios hermitianos ponderados, a su vez, por ciertas funciones de los momentos hasta un orden preespecificado (a veces denominada función de densidad de Sargan); a partir de aquí, obtener la mediana y los quantiles es una simple cuestión numérica.

En definitiva, podemos concluir este apartado afirmando que, posiblemente, los dos principios de estimación más productivos para la econometría, con sus múltiples variantes, obviamente, son el MV y el MGM. Aunque el MV pueda considerarse formalmente un caso particular del MGM, en muchos casos de interés práctico no será así. El MV tiende a ser más eficiente, pero puede ser menos robusto que el MGM. El MV puede adaptarse para que genere estimadores consistentes en cualquier caso (estimación indirecta), pero el MGM sigue ofreciendo un criterio de estimación 'más natural'.

Otros principios, como el de distancia mínima o el analógico, pueden ser entendidos, en opinión de los autores, más como una característica de algunos estimadores, que como auténticos métodos o principios de estimación independientes y operativos.

VIII. ESTIMACION MEDIANTE MOMENTOS SIMULADOS.

Ejemplo VIII.1. Sea $y_i = m^* + e_i$, $i = 1,...,I$, donde y_i son observaciones, y la distribución de 'e' es conocida, con E[e] = 0, y Var[e] = σ^2, y son independientes.

Supongamos que se generan J observaciones x_j, de la siguiente forma, $x_j = m + u_j$, donde u_j se ha generado a partir de la distribución de 'e'. Consideremos, ahora, el estimador m^*, obtenido al minimizar la siguiente expresión con respecto a m,

$$\left[\sum_{i=1}^{I} y_i/I - \sum_{j=1}^{J} x_j/J\right]^2 = C(m) \qquad [50]$$

Para analizar las propiedades del estimador, es útil escribir C(m) del modo siguiente,

$$C(m) = \left[m^* - m + \bar{e} - \bar{u}\right]^2 \qquad [51]$$

donde $\bar{e} = \Sigma e_i/I$, $\bar{u} = \Sigma u_j/J$.

Derivando e igualando a cero, obtenemos el estimador de m, es decir,

$$\partial C/\partial m = -2\,[m^* - \hat{m} + \bar{e} - \bar{u}] = 0$$

de donde $\hat{m} = m^* + \bar{e} - \bar{u} = \bar{y} - \bar{u}$

Es inmediato, ahora, que $\hat{m} + \bar{u} = \bar{x} = \bar{y}$ (donde $\bar{y} = \Sigma y_i/I$ y $\bar{x} = \Sigma x_j/J$), y por tanto, $C(\hat{m}) = 0$.

Así, el estimador es único y alcanza el mínimo valor posible del criterio propuesto. Por otra parte, el resultado es intuitivamente aceptable: se simulan valores para y_i con sucesivas medias, hasta que la media muestral de ambas series, observada y simulada (que aquí hemos denominado x_j) coinciden.

Por otra parte, es inmediato que, $E[\hat{m}] = m^* + E[\bar{e}] - E[\bar{u}] = m^*$, es decir, es un estimador insesgado. La varianza se obtiene fácilmente puesto que 'e_i' y 'u_j' son independientes. Así,

$$V(\hat{m}) = \sigma^2/I + \sigma^2/J = \left(\sigma^2/I\right)\left(1 + I/J\right) \qquad [52]$$

Finalmente, la distribución de $\hat{m}$ puede obtenerse con facilidad, si se conoce la de e_i, y por tanto, las de $\bar{e}$ y $\bar{u}$. Por ejemplo, si e_i es normal,

$$\hat{m} - N\left(m^*,\left(\sigma^2/I\right)\left(1 + I/J\right)\right)$$

Debe subrayarse que los valores simulados deben permanecer constantes para todos los valores de 'm', es decir, deben simularse sólamente una vez: de otro modo la distribución del estimador no sería la indicada (computacionalmente el ahorro es pequeño, no obstante).

Si 'e_i' no está distribuido normalmente, el resultado es válido asintóticamente, con tal de que I/J no tienda a infinito. Si J tiende a infinito más rápidamente que I, obtenemos,

$$\left(\hat{m} - m\right)\sqrt{T} \xrightarrow{d} N(0,\sigma^2)$$

que es un resultado standard, para el estimador usual de la media.

En definitiva, la idea básica de este procedimiento es sustituir en el criterio de estimación C, la media m por una estimación simulada de dicha media $\Sigma x_j/J$. Así, para J alto, es evidente que, como esta expresión será un estimador consistente de m^*, el estimador final de la media será el convencional. Además, la varianza adicional del estimador (recogida en el factor de corrección (1 + I/J)) refleja el hecho de que se está utilizando una media simulada.

El ejemplo anterior ha sido simplificado deliberadamente. De hecho, en ese caso carece de sentido aplicar el estimador propuesto, puesto que la varianza será, en general, mayor que la del estimador convencional.

Sin embargo, en el análisis económico actual, este tipo de estimador puede ser útil (especialmente en relación a modelos de optimización intertemporal, y de expectativas racionales). Su utilidad estriba en que no es preciso conocer la forma analítica de los momentos en los que se basan los estimadores: es suficiente con simularlos consistentemente para un número de observaciones elevado. En este caso el método tiene sentido. Debe subrayarse, no obstante, que el estimador de momentos analíticos es siempre superior, por lo que es el que se debería elegir, siempre que sea posible.

A continuación presentamos un ejemplo de aplicabilidad.

Ejemplo VIII.2. Sea el modelo dinámico no lineal (y estacionario), $y_t = h(y_{t-1},\beta,e_t)$, donde β es un conjunto de parámetros y 'e_t' un error iid con distribución conocida. En ocasiones es posible resolver para 'e_t' la función anterior, con lo que se obtendría $f(y_t,y_{t-1},\beta) = e_t$.

En este caso es posible aplicar cualquiera de los métodos convencionales (MC no lineales o MV). Pero si esta función no existe, lo que puede ocurrir con facilidad, la alternativa viable es el MGM con momentos simulados. Es decir, pueden generarse u_t observaciones (t = 1,...,nT) con igual distribución que e_t, y a partir de ahí las observaciones $\hat{y}_t = h(\hat{y}_{t-1},\beta,u_t)$, y los momentos,

$$\sum_{t=1}^{nT} \hat{y}_t^2/nT$$

$$\sum_{t=s}^{nT} \hat{y}_t\hat{y}_{t-s}/nJ \qquad [53]$$

de donde se puede estimar el vector de parámetros β.

VIII.1. El método generalizado de momentos simulados.

El método se basa, como ya se ha indicado, en valores simulados de la variable dependiente y_t, a partir de los cuales se forman los correspondientes momentos, tal y como se indicó en el epígrafe III; es decir,

$$\sum_{i=1}^{n}\sum_{t=1}^{T} f_j\,(\hat{y}_{ti}, x_t, \bar{\theta})/(nT) = m(\bar{\theta}) \qquad [54]$$

Partiendo de esta nueva definición, todos los resultados de los apartados anteriores son aplicables con alguna salvedad, que se comenta a continuación (no obstante, las pruebas formales pueden encontrarse en Lee e Ingram (1991) y Duffie y Singleton (1993)). La mayor parte de los resultados son extensiones más o menos automáticas de las derivaciones expuestas en los epígrafes anteriores.

En primer lugar, la varianza de los estimadores debe ser multiplicada por el factor (1 + n), así como todos los criterios en los que se basan los contrastes de sobreidentificación y restricciones sobre parámetros.

La única prueba que no es obvia se refiere a la estacionariedad de las simulaciones, es decir, en este caso, a su independencia del valor inicial y_0, dado un valor concreto del vector de parámetros β. En el caso lineal estacionario, esta independencia se obtiene fácilmente. En otros casos no es así, por lo que es necesario imponer condiciones especiales (más concretamente, lo que Duffie y Singleton denominan 'ergodicidad geométrica'). En la práctica, no obstante, puede ser difícil comprobar el cumplimiento de esta condición.

También es preciso observar, que si la distribución de los errores afecta a la forma de la dependencia funcional de los momentos respecto de los parámetros, el estimador de momentos simulados será sesgado, a no ser que se conozca exactamenta dicha distribución. Así, en este contexto es especialmente importante contrastar la distribución utilizada para simular.

Es interesante señalar, por último, que una de las aplicaciones de este método es el contraste estadístico de los modelos 'calibrados' y el enfoque 'calibracionista', en general. Más allá de las discusiones metodológicas acerca de los fundamentos filosóficos del método econométrico, lo cierto es que si se acepta la utilidad del paradigma estadístico clásico aplicado a las ciencias sociales, y en este caso a la economía, el método de los momentos simulados proporciona la solución estadística más natural a la calibración de los modelos (otros aspectos como los criterios de estimación indirecta, y la incorporación del enfoque Bayesiano, se discuten en otros apartados).

IX. EXTENSIONES Y CONCLUSIONES.

El método generalizado de los momentos engloba a un conjunto importante de estimadores conocidos; sin embargo, no siempre se trata de un estimador eficiente. De hecho, la eficiencia de este estimador está vinculada a la clase de estimadores VI definida por las condiciones de ortogonalidad. Así, por ejemplo, en presencia de autocorrelación o de heterocedasticidad MCO no es tan eficiente como MCG (dado que el MCO es un forma del MGM, incluso en presencia de autocorrelación o heterocedasticidad, será en general no eficiente, aunque si consistente).

Por otro lado, aunque los ejemplos presentados aquí se han desarrollado en el contexto de modelos lineales, la extensión, desde el punto de vista teórico, a modelos no lineales es inmediata. En la práctica, se debe recurrir a algún algoritmo de optimización, que minimice numéricamente la función criterio para obtener los estimadores. La única cuestión relevante en este tipo de modelos, no discutida en este trabajo, es la selección del conjunto de variables instrumentales. En general es posible demostrar que los instrumentos óptimos son $E[\partial u_t/\partial\theta] = Z_t$, siendo u_t el error de la ecuación no lineal a estimar. Como Z_t es, evidentemente, desconocida, existen una variedad de métodos para estimarla, sin menoscabo de la eficiencia asintótica del estimador de θ (o con la menos pérdida posible; véase, por ejemplo, Bowden-Turkington (1984)).

Desde el punto de vista práctico, también es conveniente señalar que todas las propiedades presentadas se desarrollan en el ámbito de la teoría asintótica, pero no se puede decir mucho acerca del comportamiento de este estimador para tamaños de la muestra relativamente pequeños. En dos puntos hay que ser especialmente cuidadoso: i) no tomar demasiadas VI, pues en el caso de que J = T, el estimador correspondiente de VI es idéntico al MCO y, ii) los métodos no paramétricos en la estimación de la matriz de covarianzas de los estimadores, que, aunque asintóticamente consistentes, suelen ser muy ineficientes (siempre debe preferirse un estimador paramétrico si está disponible).

Por otra parte, en todas las páginas previas no se ha hecho ninguna referencia directa al supuesto de estacionariedad. Para los casos en los que W_t no es estacionario (incluye por ejemplo una tendencia determinista) puede consultarse el trabajo realizado por Ogaki (1993). Alternativamente, puede utilizarse un criterio de estimación indirecta basado en una transformación estacionaria de los datos, definiendo a continuación un criterio de máxima verosimilitud (Laroque y Salanié, (1993)), o bien un criterio de momentos simulados. Este aspecto está abierto a la investigación futura, pero no es dificultad insuperable, por las razones que se acaban de aducir.

También es posible contrastar la existencia de un cambio estructural utilizando el test de Wald para contrastar la H_0: $\theta_1 = \theta_2$, basándose en el supuesto de que ambos estimadores son asintóticamente independientes (veáse Andrews-Fair, 1988). Similatmente pueden definirse contraste de restricciones lineales, aplicando los tres criterios discutidos en secciones anteriores. Contrastes de hipótesis no anidadas también pueden llevarse a cabo en el marco del estimador MGM (Smith, (1992)).

En cuanto a la incorporación del método Bayesiano al método MGM, diremos que, similarmente al caso de la regresión lineal, la información a priori puede resumirse en un conjunto de restricciones estocásticas, lineales o no. Estas restricciones pueden combinarse con las condiciones de primer orden para generar unos estimadores 'mejorados'; es decir, insesgados asintóticamente y con menor varianza. La combinación, por analogía al caso lineal, puede realizarse siguiendo el principio de mínimos cuadrados generalizados. Este procedimiento, aunque todavía no discutido en la literatura, es una extensión inmediata de la sugerencia originaria de Goldberger para incorporar la información a priori, y permite responder fácilmente a las críticas al MGM desde el enfoque Bayesiano (ver, a este respecto Cánova, (1993)).

Para finalizar diremos que una de las aplicaciones más relevantes del MGM es a datos de panel. En este caso el estudio de las propiedades asintóticas se encuadra en el comportamiento de las mismas cuando el número de individuos N, y no el número de observaciones temporales T, tiende a ∞. Por tanto, también $E[m_j] = 0$ hace referencia a la información sobre los individuos. No obstante, discutir las peculiaridades propias del MGM, y en especial de la estimación de los momentos simulados a los datos de panel excede del propósito de este documento. Existen, no obstante, un importante conjunto de trabajos centrados en este tema, entre los que podemos reseñar los realizados por Arellano y Bond (1991), Keane y Runkle (1992) y Keane (1993).

REFERENCIAS BIBLIOGRAFICAS.

Andrews, D.W.K. - Fair, C. (1988): "Inference in Nonlinear Econometric Models with Structural Change". *Review of Economic Studies*, 55, pp 615 - 640.

Andrews, D.W.K. - Monahan, J.K. (1992): "An Improved Heteroskedasticity and Autocorrelation Consistent Matrix Estimator". *Econometrica*, 60. pp 953-966.

Arellano, M. - Bond, S. (1991): "Some test of specification for panel data: Monte Carlo evidence and an application to employment equations". *Rev. Econom. Stud*, 58. pp 277-297.

Bowden, R.J. - Turkington, D. (1984): Instrumental Variables. Econometric Society Monographs in Quantitative Economics. CUP.

Cragg, J.G. (1983): "More efficient estimation in the presence of heteroskedasticity of unknown form". *Econometrica*, 51. pp 751-763.

Davidson, R. - Mackinnon, J. (1993): *Estimation and Inference in Econometrics.* Oxford University Press. Caps 5 y 17.

Duffie, D. - Singleton, K.J. (1993): "Simulated Moments Estimation of Markov Models". *Econometrica.*

Gallant, R. - Tauchen, G. (1994): "Which moments to match". University of North Carolina (mimeo)

Gourieroux, C. - Monfort, A. - Trognon, A. (1984): "Pseudo maximum likelihood methods: Theory". Econometrica, 52, pp. 681-700.

Gourieroux, C. - Monfort, A. - Renault, E. (1993): "Indirect inference". *Journal of Applied Econometrics*, 8. pp. 85-118.

Green, N. (1993): Econometric Analysis. MacMillan. caps 4 y 13. (2ª Ed).

Hamilton, J.D. (1994): Time Series Analysis. Priceton. Cap 14.

Hansen, L.P.(1982): "Large Sample Properties of Generalized Method of Moments Estimators". *Econometrica*, 50. pp. 1029-1054.

Hansen, L.P. (1985): "A method for calculating bounds on the asymptotic covariance matrices of generalizad method of moments estimators". *Journal of Econometrics*, 30. pp 203-238.

Keane, M.P. - Runkle, D.E. (1992): "On the Estimation of Panel Data with Serial Correlation when Instrument are not strictly exogenus". *J.Bussiness Economic Statistic*, 10. pp 1-9.

Keane, M.P. (1993): "Simulation estimation for panel data models with limited dependent variables". En Maddala,G.S - Rao, C.R - Vinod, H.D (ed). *Handbook of Statistics*, vol xi. Amsterdam, North Holland.

Laroque, G. - Salante, B. (1993): "Simulation based estimation of models with lagged latent variables". *Journal of Applied Econometrics.*

Lee, B.S. - Ingram, B.F. (1991): "Simulation estimation of time-series models". *Journal of Econometrics*, 47. pp 197-205.

Malinvaud, E. (1970): *Statistical Method of Econometrics.* Amsterdam, North Holland.

Manski, Ch. (1994): "Analog Estimation of Econometric Models", en Handbook of Econometrics vol IV. North-Hollan. Amsterdam.

Newey, W.K. (1985): "Generalized Method of Moments Specificaction Testing". *Journal of Econometrics*, 29. pp 229-256.

Newey, W.K. - West, K.D. (1987): "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix". *Econometrica*, 55. pp 703-708.

Ogaki, M. (1993): "Generalized Method of Moments: Econometric Applications". En Maddala,G.S - Rao, C.R - Vinod, H.D (ed). *Handbook of Statistics*, vol xi. Amsterdam, North Holland

Sargan, J.D. (1958): "The Estimation of Economic Relationship Using Intrumental Variables". *Econometrica*, 26. pp 393-415.

Smith, R. (1992): "Non-nested test for computing models estimated by generalized method of moments", *Econometrica*, 60.pp 973-980.

Smith, A. (1993): "Estimating Nonlinear Time Series Models using simulated vector autorregresion". *Journal of Applied Econometrics*, 8. pp 63-85.

White, H. (1980): "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity". *Econometrica*, 48, pp 817-838.

DOCUMENTOS PUBLICADOS

WP-EC 90-01	"Los Determinantes de la Evolución de la Productividad en España" M. Mas, F. Pérez. Diciembre 1990.
WP-EC 90-02	"Mecanización y Sustitución de Factores Productivos en la Agricultura Valenciana" A. Picazo, E. Reig. Diciembre 1990.
WP-EC 90-03	"Productivity in the Service Sector" H. Fest. Diciembre 1990.
WP-EC 90-04	"Aplicación de los Modelos de Elección Discreta al Análisis de la Adopción de Innovaciones Tecnológicas. El Caso del Sector Azulejero" E.J. Miravete. Diciembre 1990.
WP-EC 90-05	"Rentabilidad y Eficiencia del Mercado de Acciones Español" A. Peiró. Diciembre 1990.
WP-EC 90-06	"La Coordinación de Políticas Fiscales en el Marco de una Unión Económica y Monetaria" J.E. Boscá, V. Orts. Diciembre 1990.
WP-EC 91-01	"Medición de la Segregación Ocupacional en España: 1964-1988" M. Sánchez. Mayo 1991.
WP-EC 91-02	"Capital Adequacy in the New Europe" E.P.M. Gardener. Mayo 1991.
WP-EC 91-03	"Determinantes de la Renta de los Hogares de la Comunidad Valenciana. Una Aproximación Empírica." M.L. Molto, C. Peraita, M. Sánchez, E. Uriel. Mayo 1991.
WP-EC 91-04	"Un Modelo para la Determinación de Centros Comerciales en España". A. Peiró, E. Uriel. Septiembre 1991.
WP-EC 91-05	"Exchange Rate Dynamics. Cointegration and Error Correction Mechanism". M.A. Camarero. Septiembre 1991.
WP-EC 91-06	"Aplicación de una Versión Generalizada del Lema de Shephard con Datos de Panel al Sistema Bancario Español". R. Doménech. Septiembre 1991.
WP-EC 91-07	"Necesidades, Dotaciones y Deficits en las Comunidades Autónomas" B. Cabrer, M. Mas, A. Sancho. Diciembre 1991.
WP-EC 91-08	"Un Análisis del Racionamiento de Crédito de Equilibrio" J. Quesada. Diciembre 1991.
WP-EC 91-09	"Cooperación entre Gobiernos para la Recaudación de Impuestos Compartidos" G. Olcina, F. Pérez. Diciembre 1991.
WP-EC 91-10	"El impacto del Cambio Tecnológico en el Sistema Bancario: El Cajero Automático" J. Maudos. Diciembre 1991.

WP-EC 91-11 "El Reparto del Fondo de Compensación Interterritorial entre las Comunidades Autónomas"
C. Herrero, A. Villar. Diciembre 1991.

WP-EC 91-12 "Sobre la Distribución Justa de un Pastel y su Aplicación al Problema de la Financiación de las Comunidades Autónomas"
C. Herrero, A. Villar. Diciembre 1991.

WP-EC 92-01 "Asignaciones Igualitarias y Eficientes en Presencia de Externalidades"
C. Herrero, A. Villar. Abril 1992.

WP-EC 92-02 "Estructura del Consumo Alimentario y Desarrollo Economico"
E. Reig. Abril 1992.

WP-EC 92-03 "Preferencias de Gasto Reveladas por las CC.AA."
M. Mas, F. Pérez. Mayo 1992.

WP-EC 92-04 "Valoración de Títulos con Riesgo: Hacia un Enfoque Alternativo"
R.J. Sirvent, J. Tomás. Junio 1992.

WP-EC 92-05 "Infraestructura y Crecimiento Económico: El Caso de las Comunidades Autónomas"
A. Cutanda, J. Paricio. Junio 1992.

WP-EC 92-06 "Evolución y Estrategia: Teoría de Juegos con Agentes Limitados y un Contexto Cambiante"
F. Vega Redondo. Junio 1992.

WP-EC 92-07 "La Medición del Bienestar mediante Indicadores de 'Renta Real': Caracterización de un Indice de Bienestar Tipo Theil"
J.M. Tomás, A. Villar. Julio 1992.

WP-EC 92-08 "Corresponsabilización Fiscal de Dos Niveles de Gobierno: Relaciones Principal-Agente"
G. Olcina, F. Pérez. Julio 1992.

WP-EC 92-09 "Labour Market and International Migration Flows: The Case of Spain"
P. Antolín. Julio 1992.

WP-EC 92-10 "Un Análisis Microeconométrico de la Demanda de Turismo en España"
J.M. Pérez, A. Sancho. Julio 1992.

WP-EC 92-11 "Solución de Pérdidas Proporcionales para el Problema de Negociación Bipersonal"
M.C. Marco. Noviembre 1992.

WP-EC 92-12 "La Volatilidad del Mercado de Acciones Español"
A. Peiró. Noviembre 1992.

WP-EC 92-13 "Evidencias Empíricas del CAPM en el Mercado Español de Capitales"
A. Gallego, J.C. Gómez, J. Marhuenda. Diciembre 1992.

WP-EC 92-14 "Economic Integration and Monetary Union in Europe or the Importance of Being Earnest: A Target-Zone Approach"
E. Alberola. Diciembre 1992.

WP-EC 92-15 "Utilidad Expandida y Algunas Modalidades de Seguro"
R. Sirvent, J. Tomás. Diciembre 1992.

WP-EC 93-01 "Efectos de la Innovación Financiera sobre la Inversión: El Caso del Leasing Financiero"
M.A. Díaz. Junio 1993.

WP-EC 93-02 "El problema de la Planificación Hidrológica: Una Aplicación al Caso Español"
A. González, S.J. Rubio. Junio 1993.

WP-EC 93-03 "La Estructura de Dependencia del Precio de las Acciones en la Identificación de Grupos Estratégicos: Aplicación al Sector Bancario Español"
J.C. Gómez Sala, J. Marhuenda, F. Más. Noviembre 1993.

WP-EC 93-04 "Dotaciones del Capital Público y su Distribución Regional en España"
M. Mas, F. Pérez, E. Uriel. Noviembre 1993.

WP-EC 93-05 "Disparidades Regionales y Convergencia en las CC.AA. Españolas"
M. Mas, J. Maudos, F. Pérez, E. Uriel. Noviembre 1993.

WP-EC 93-06 "Bank Regulation and Capital Augmentations in Spain"
S. Carbó. Diciembre 1993.

WP-EC 93-07 "Transmission of Information Between Stock Markets"
A. Peiró, J. Quesada, E. Uriel. Diciembre 1993.

WP-EC 93-08 "Capital Público y Productividad de la Economía Española"
M. Mas, J. Maudos, F. Pérez, E. Uriel. Diciembre 1993.

WP-EC 93-09 "La Productividad del Sistema Bancario Español (1986-1992)"
J.M. Pastor, F. Pérez. Diciembre 1993.

WP-EC 93-10 "Movimientos Estacionales en el Mercado de Acciones Español"
A. Peiró. Diciembre 1993.

WP-EC 93-11 "Thresholds Effects, Public Capital and the Growth of the United States"
J. García Montalvo. Diciembre 1993.

WP-EC 94-01 "International Migration Flows: The Case of Spain"
P. Antolín. Febrero 1994.

WP-EC 94-02 "Interest Rate, Expectations and the Credibility of the Bank of Spain"
F.J. Goerlich, J. Maudos, J. Quesada. Marzo 1994.

WP-EC 94-03 "Macromagnitudes Básicas a Nivel Sectorial de la Industria Española: Series Históricas"
F.J. Goerlich, V. Orts, S. García. Mayo 1994.

WP-EC 94-04 "Job Search Behaviour"
P. Antolín. Mayo 1994.

WP-EC 94-05 "Unemployment Flows and Vacancies in Spain"
P. Antolín. Mayo 1994.

WP-EC 94-06 "Paro y Formación Profesional: Un Análisis de los Datos de la Encuesta de Población Activa"
C. García Serrano, L. Toharia. Mayo 1994.

WP-EC 94-07 "Determinantes de la Dinámica de la Productividad de los Bancos y Cajas de Ahorro Españolas"
J.M. Pastor. Junio 1994.

WP-EC 94-08 "Estimación Regionalizada del Stock de Capital Privado (1964-1989)"
F.J. Escribá, V. Calabuig, J. de Castro, J.R. Ruiz. Junio 1994.

WP-EC 94-09 "Capital Público y Eficiencia Productiva Regional (1964-1989)"
M. Mas, J. Maudos, F. Pérez, E. Uriel. Julio 1994.

WP-EC 94-10 "Can the Previous Year Unemployment Rate Affect Productivity? A DPD Contrast"
R. Sánchez. Septiembre 1994.

WP-EC 94-11 "Comparing Cointegration Regression Estimators: Some Additional Monte Carlo Results"
J. García Montalvo. Septiembre 1994.

WP-EC 94-12 "Factores Determinantes de la Innovación en las Empresas de la Comunidad Valenciana"
M. Gumbau. Septiembre 1994.

WP-EC 94-13 "Competencia Imperfecta y Discriminación de Precios en los Mercados de Exportación. El Caso del Sector de Pavimentos Cerámicos"
J. Balaguer. Noviembre 1994.

WP-EC 94-14 "Utilidad Expandida Estado Dependiente: Algunas Aplicaciones"
R.J. Sirvent, J. Tomás. Noviembre 1994.

WP-EC 94-15 "El Efecto de las Nuevas Tecnologías de Transacción en la Demanda de Dinero en España"
J. Maudos. Noviembre 1994.

WP-EC 94-16 "Desajustes en los Tipos de Cambio e 'Hysteresis' en los Flujos Comerciales: Las Exportaciones Españolas a EE.UU."
J. de Castro, V. Orts, J.J. Sempere. Diciembre 1994.

WP-EC 94-17 "Stock Prices and Macroeconomic Factors: Evidence from European Countries"
A. Peiró. Diciembre 1994.

WP-EC 95-01 "Margen Precio-Coste Marginal y Economías de Escala en la Industria Española: 1964-1989"
F.J. Goerlich, V. Orts. Abril 1995.

WP-EC 95-02 "Temporal Links Between Price Indices of Stock Markets with Overlapping Business Hours"
A. Peiró, J. Quesada, E. Uriel. Abril 1995.

WP-EC 95-03 "Competitive and Predatory Multi-Plant Location Decisions"
A. García Gallego, N. Georgantzis. Abril 1995.

WP-EC 95-04 "Multiproduct Activity and Competition Policy: The Tetra Pack Case"
A. García Gallego, N. Georgantzis. Junio 1995.

WP-EC 95-05 "Estudio Empírico de la Solvencia Empresarial en Comunidad Valenciana"
J. L. Gandía, J. López. R. Molina. Junio 1995.

WP-EC 95-06 "El Método Generalizado de los Momentos"
A. Denia, I. Mauleón. Junio 1995.