

## **Análisis de interpretación de los parámetros de relación de varianzas en el modelo de frontera estocástica**

RAFAELA DIOS PALOMARES.

*Grupo de Eficiencia y Productividad de la Universidad de Córdoba. Departamento de Estadística e Investigación Operativa. Universidad de Córdoba. ETSIAM*

Avda. alameda del Obispo, s/n-14080 Córdoba. Telf.: 957 21 84 79. E-mail: [ma1dipar@uco.es](mailto:ma1dipar@uco.es).

### RESUMEN

El presente trabajo recoge los resultados de un estudio sobre la interpretación de los parámetros de relación de varianzas en el modelo de frontera estocástica. Se valora el error cometido al no considerar la verdadera relación de varianzas de los dos componentes del error compuesto del modelo y que se produce al asignar a la variable truncada, la varianza de la normal de la que procede. Se definen los parámetros reales sobre los que hay que tomar conclusiones y se demuestra que los errores que se valoran, que pueden suponer una sobrevaloración de la presencia de la ineficiencia superior al 750%, dependen del coeficiente de variación y de la proporción en que intervenga la variable de ineficiencia.

*Palabras Clave:* Eficiencia, Frontera Estocástica, Error compuesto, Normal Truncada.

### ABSTRACT

This work presents the results of a study carried out to point up the correct interpretation of the variances ratio in the stochastic frontier framework. We have evaluated the error which arises when the variance of the underlying normal variable is taken instead of that which corresponds to the truncated normal in the computation of the Gamma parameter. We have defined the real parameters on which decisions must be taken and have demonstrated that the above mentioned error produces an overevaluation of more than 750%, of the inefficiency presence. We also concluded that the evaluated error depends on the  $m/s$  ratio and on the inefficiency variable incidence in the composed error of the model.

*Key words:* Efficiency, Stochastic Frontier, Composed Error, Truncated Normal.

Código UNESCO: 5302.-5306

AMS: 91B38

Artículo recibido el 26 de junio de 2001. Aceptado el 7 de marzo de 2002.

## 1. INTRODUCCIÓN

El análisis de eficiencia en producción es actualmente aplicado con mucha frecuencia en investigación económica, tanto en el ámbito de la eficiencia técnica a través de función frontera de producción, como de la eficiencia asignativa y económica, tomando como base la frontera de costes o la de beneficios.

El estudio de la eficiencia técnica para un grupo de empresas, se apoya en la comparación entre la producción real y la producción (frontera) que habría alcanzado si hubiera empleado los factores de producción de forma totalmente eficiente. La frontera de producción se encontrará en el valor máximo alcanzable por cada empresa dados unos factores de producción.

La eficiencia económica valora la capacidad con que la empresa utiliza los factores de producción de forma óptima desde el punto de vista económico. Esta situación será la del mínimo coste para producir unos productos, dados unos precios para los factores. Dicha eficiencia nos vendrá dada por la relación entre el coste mínimo (frontera) y el coste real soportado.

La eficiencia asignativa se refiere solo al uso de los factores en proporción óptima, por lo que se establece que el producto de eficiencia técnica por la asignativa es igual a la económica.

Las dos metodologías más usadas en la actualidad para estimar eficiencias mediante función frontera son la programación matemática aplicando DEA (Seiford y Thrall, 1990) y la que se denomina frontera econométrica ( Battese, 1992), que será el objeto del presente trabajo, y a la que nos referiremos a continuación.

De una forma general la función de producción se especifica de la forma

$$y = X * \mathbf{b} + v$$

siendo  $y$  la producción obtenida,  $X$  las variables que recogen los factores de producción, y  $v$  la perturbación aleatoria que convencionalmente se supone con una distribución Normal de media cero y varianza constante  $\sigma_v^2$ .

Así, una función de producción frontera se especifica por medio del modelo:

$$y = X * \mathbf{b} + \mathbf{e} \quad (1)$$

donde la variable de error  $\mathbf{e}$  recoge la diferencia entre la parte sistemática del modelo y los valores observados.

En el ámbito del estudio de la eficiencia, y en base al modelo anterior, se pueden considerar dos enfoques que son los siguientes.

### **Frontera Global o Determinística**

En este enfoque, la variable de error representa la ineficiencia del sistema. Por tanto, las diferencias ocurridas entre la  $Y$  observada y la correspondiente frontera se debieron únicamente a ineficiencia. Este planteamiento supone que la variable de error debe tomar siempre valores negativos, y la frontera estimada superará siempre a los valores observados excepto para la empresa más eficiente cuya producción se encuentra sobre la frontera.

La especificación que se plantea para este tipo de variable de error debe de ser congruente con la característica de no tomar valores positivos y suele adoptar forma asimétrica dado el fenómeno que recoge. Las variables Exponencial, Seminormal, Normal Truncada y Gamma, han sido las consideradas con más frecuencia en la literatura, tanto en este enfoque, como en el de error compuesto que analizaremos más adelante.

La estimación del índice de eficiencia para cada individuo se calcula mediante el índice de Timmer (Timmer, 1970), que cuantifica de manera relativa al más eficiente, la razón entre el resultado obtenido y el máximo que podría haber conseguido (frontera) si hubiera sido totalmente eficiente.

### **Frontera Estocástica o de Error Compuesto**

En el enfoque de Error Compuesto, la variable de error no solo recoge el efecto de la ineficiencia, sino que también existe otra fuente de error incluida en la misma, que no es controlable por el individuo. Se admite por tanto, que la variable  $e$  se genera como diferencia entre una variable estocástica  $v$  (no controlable, simétrica, y definida entre  $-\infty$  e  $\infty$ ) y la variable de ineficiencia  $u$ , que en este caso será siempre positiva y asimétrica, como la definida para la frontera global. La variable  $v$  se especifica con distribución Normal con media cero y varianza  $\sigma_v^2$ , y la variable  $u$  se ha especificado con las distribuciones ya comentadas para la ineficiencia, que son la Exponencial, Seminormal, Normal Truncada y Gamma.

Para realizar el análisis de la eficiencia con este planteamiento, en base a la especificación que se establezca para  $u$ , se procede a estimar por máximoverosimilitud el modelo de función frontera, suponiendo para el error compuesto la distribución que le corresponda a  $e$  como diferencia de  $v$  y  $u$ . Como resultado inmediato de dicha estimación se obtienen los residuos que recogen dicho error compuesto, del que habrá que extraer la parte que se deba realmente a ineficiencia.

Aigner et al(1977) plantearon el modelo de función frontera con error compuesto con la especificación Normal-Seminormal resolviendo la estimación máximoverosimil de los parámetros  $\beta$ ,  $\sigma_v^2$  y  $\sigma_u^2$ . Posteriormente Stevenson (1980), generalizó la especificación a la Normal-Normal Truncada y la Normal-Gamma. La metodología desarrollada en ambos

trabajos, permiten estimar la esperanza matemática de la ineficiencia, y separar la influencia debida a cada una de las variables de error, en base a la valoración de las varianzas respectivas. Para la estimación del modelo de frontera bajo esta especificación, Aigner et al(1977), parametrizaron la función máximoverosimil definiendo el parámetro

$$I = \frac{s_u}{s_v}$$

con el fin de simplificar los desarrollos, y por la conveniencia de que de alguna forma refleja una relación entre la dispersión de ambas variables de error. Con el mismo motivo, dichos autores definen

$$s = \sqrt{s_v^2 + s_u^2}.$$

Queda por tanto la función expresada en términos de  $\lambda$  y  $\sigma$  siendo estos dos de los parámetros que se estiman, junto con  $\beta$ ,  $\sigma_v^2$ ,  $\sigma_u^2$  y  $\mu$ .

Posteriormente Battese y Corra(1977), para valorar la relevancia de la variable  $u$  en el contexto del error compuesto e definen el parámetro

$$g = \frac{s_u^2}{s_v^2 + s_u^2}.$$

Dada estas parametrizaciones, tanto el paquete econométrico LIMDEP (Greene, 1992), como el FRONTIER (Coelli,1992 y 1996) proporcionan en sus salidas de estimación de frontera estocástica los valores de los parámetros de relación de varianzas definidos anteriormente, siendo dichas estimaciones fundamentales para el establecimiento de la importancia relativa de los dos errores  $v$  y  $u$  entre sí. Sin embargo, en el caso de que  $u$  se especifique como Seminormal o Normal Truncada, la relevancia de la ineficiencia del sistema puede ser sobrevalorada debido a una mala interpretación de los parámetros  $\lambda$  y  $\gamma$ , como expondremos a continuación.

## 2. INTERPRETACIÓN DE LOS PARÁMETROS DE RELACIÓN DE VARIANZAS

En la interpretación de resultados de la estimación máximoverosimil de la función frontera, cobra un papel importante el valor que toman los parámetros  $\lambda$  y  $\gamma$  definidos anteriormente y que relacionan medidas de dispersión de las variables  $u$  y  $v$ . Estas medidas se interpretan como indicadores de la importancia relativa de cada variable en la composición del error compuesto de modo que si  $g$  toma un valor muy próximo a la unidad, se deduce que no hay efectos en el error debido a factores fuera del control de la empresa y todo se

debe a la falta de eficiencia. En este caso, la mejora de la eficiencia podrá conseguirse con mayor acierto actuando apropiadamente. Por el contrario, un valor de  $\gamma$  próximo a cero indicaría una total eficiencia en el sistema bajo estudio.

Sin embargo, hay que tener en cuenta que el parámetro  $\sigma_u^2$  es la varianza de la distribución Normal ( $N(\mu, \sigma_u^2)$ ) cuyo truncamiento en cero da lugar a la variable  $u$ . En consecuencia,  $\sigma_u^2$  no es la varianza de  $u$ , sino que al tratarse de una truncación de la primera, su varianza será tanto menor que  $\sigma_u^2$  cuanto mayor sea la truncación, dependiendo esta a su vez del valor que tome el inverso del coeficiente de variación de la Normal completa, que viene dado por  $\mu/\sigma_u$ . Así, el error de interpretación cometido al no tener en cuenta este hecho, depende de la magnitud del truncamiento, de modo que puede llegar a ser bastante importante para valores muy pequeños de dicho coeficiente.

Consideramos por tanto, que la valoración de la incidencia de  $u$  sobre el error compuesto, no debe hacerse mediante el estudio de los parámetros  $\lambda$  y  $\gamma$ , sino analizando los que definiremos como  $\lambda_R$  y  $\gamma_R$  que calcularemos como:

$$\mathbf{I}_R = \left( \frac{V(u)}{\mathbf{s}_v^2} \right)^{1/2} \quad \text{y} \quad \mathbf{g}_R = \frac{V(u)}{\mathbf{s}_v^2 + V(u)}$$

siendo  $V(u)$  la varianza de  $u$ .

En los trabajos de investigación sobre eficiencia mediante función frontera econométrica realizados recientemente, existe una interesante línea de trabajo enfocada hacia la detección de los factores socioeconómicos que pueden haber influido en la eficiencia de cada empresa. Para ello, se han establecido dos metodologías que se conocen como el método en dos etapas (Pitt y Lee, 1981) y el método de modelo conjunto (Battese y Coelli, 1995).

Este último se apoya en la teoría de que la distribución de la variable  $u_i$  no se va manteniendo igual para todas las empresas, sino que el parámetro  $m_i$  depende de algunos factores socioeconómicos. Se especifica de esta forma el modelo de frontera como hemos planteado hasta aquí en (1), pero éste se estima conjuntamente con el modelo de ineficiencia que recoge la influencia de los factores socioeconómicos sobre el parámetro  $\mu_i$  y que se expresa del modo

$$\mathbf{m}_i = g(\mathbf{Z}_i, \mathbf{d}) \quad (2)$$

siendo  $\mathbf{Z}_i$  un vector de variables socioeconómicas asociadas con los efectos de ineficiencia;  $\mathbf{d}$  es un vector de parámetros; y  $g(\cdot)$  es una forma funcional apropiada, normalmente lineal.

Según este planteamiento, a la variable que recoge la ineficiencia de cada empresa se le supone un valor distinto de  $\mu_i$  que es la media de la Normal completa de la que procede  $u_i$ , asumiendo el mismo parámetro  $\sigma_u^2$  para todas. En base a los anteriores supuestos, queda claro que si bien los parámetros  $\lambda$  y  $\gamma$  son iguales para toda la muestra, no lo son en cambio

los valores reales de las varianzas de las  $u_i$ , por lo que cabe calcular un valor distinto de  $\lambda_{Ri}$  y  $\gamma_{Ri}$  para cada unidad muestral, debido a que el cambio en  $m_i$  supone también cambio en el coeficiente de variación, en el truncamiento y por tanto en el valor de  $V(u_i)$ .

Dado que la salida de los paquetes econométricos como el LIMDEP y el FRONTIER, aportan estimaciones de  $\lambda$  y  $\gamma$  y no de los valores  $\lambda_R$  y  $\gamma_R$ , es frecuente encontrar trabajos aplicados donde se realiza la interpretación de resultados en base a los primeros, con el consiguiente error de infravaloración. Como ejemplo, podemos hacer referencia al trabajo de Bagi (1984) y las recientes publicaciones de Tadesse, B. y Krishnamoorthy, S. (1997) y Hallan y Machado (1996), en que se comenta el valor estimado de  $\gamma$  como si fuera la  $\gamma_R$ . Igualmente, es interesante destacar que en lo que conocemos, todas las aplicaciones de la estimación del modelo conjunto se limitan a aportar e incluso a comentar el valor único que toma el parámetro  $g$ , como se puede ver en, Coelli (1996), Coelli y Battese (1996), Tian, y Wan (2000), Tveras y Battese (2001), y Fried et al (2002) entre otros.

Pasamos a continuación a realizar un análisis de valoración del error que se puede cometer en la interpretación de los parámetros que recogen la relación de varianzas de los errores  $u_i$  y  $v_i$ , cuando la especificación del error compuesto es Normal-Normal.Truncada, estudiando a su vez la Normal-Seminormal como el caso particular de la anterior en que  $\mu = 0$ .

Se considera que  $v_i$  sigue una distribución Normal de media cero y varianza  $\sigma_v^2$ ,  $v_i \sim N(0, \sigma_v^2)$  con  $-\infty \leq v_i \leq \infty$  y  $u_i$  se distribuye como una Normal Truncada en 0, de moda  $\mu$ , es decir  $u_i \sim N(\mu, \sigma_u^2)$  con  $u_i \geq 0$ .

Aunque la distribución de  $u_i$ , tal como la hemos definido aquí, depende de los dos parámetros  $\mu$  y  $\sigma_u^2$ , podemos demostrar que la magnitud del truncamiento se puede expresar en función del coeficiente de variación. Esto supone que todas las variables normales truncadas en cero que tengan el mismo coeficiente de variación, tendrán el mismo truncamiento y por tanto la misma reducción en su varianza al sufrir dicho truncamiento.

Para analizar sus características, definimos previamente el inverso del coeficiente de variación que será:

$$icv = \frac{m}{s_u}$$

La esperanza matemática de  $u_i$  será:

$$E(u_i) = m + s_u * I(-icv), \text{ con } I(-icv) = \frac{f(-icv)}{1 - F(-icv)}$$

siendo  $\phi(\cdot)$  y  $\Phi(\cdot)$  las funciones de densidad y distribución de la Normal Tipificada.

La varianza de  $u_i$ , tomará el valor:

$$V(u_i) = s_u^2 * [1 - d(-icv)] \text{ con } d(-icv) = I(-icv) * [I(-icv) + icv]$$

Podemos por tanto establecer que la relación entre la varianza  $V(u)$  de la Normal truncada y la de la Normal entera de la que procede será:

$$\frac{V(u_i)}{s_u^2} = [1 - d(-icv)], \text{ con lo que queda claro que dicha relación solo dependerá del}$$

valor que toma el inverso del coeficiente de variación.

Si estudiamos algunos casos particulares, podemos valorar la importancia que puede llegar a tener el error de interpretación que estamos analizando. Teniendo en cuenta que estudiaremos variaciones sobre el valor de  $icv$ , nos podemos centrar en la familias de distribuciones de desviación típica igual a la unidad, con lo que pondremos nuestra atención solo en el valor de  $\mu$ , ya que coincidirá con el de  $icv$ .

Los casos de menor truncamiento corresponden a valores muy altos de  $\mu$ . Si nos centramos en truncamientos que coincidan con percentiles exactos y habituales en el análisis estadístico, estudiaremos el siguiente:

$$\mathbf{m = 1.96, s_u^2 = 1, y icv = 1.96}$$

Si truncamos esta variable en el valor cero, estaremos prescindiendo por la izquierda de un área de la función de densidad igual a 0.025, que sería el valor de truncamiento que va desde  $-\infty$  hasta cero. En este caso, aunque cabe pensar que la reducción de la varianza no sería muy importante, pasamos a valorarla a continuación.

$$f(-icv) = f(-1.96) = 0.0584$$

$$F(-icv) = F(-1.96) = 0.025$$

$$I(-icv) = I(-1.96) = \frac{f(-1.96)}{1 - F(-1.96)} = \frac{0.0584}{1 - 0.025} = 0.0599$$

La esperanza matemática de  $u_i$  será:

$$E(u) = \mathbf{m + s_u * I(-icv)} = 1.96 + 1 * 0.0599 = 2.0199$$

La relación de varianzas tomará el valor:

$$\begin{aligned} d(-icv) &= I(-icv) * [I(-icv) + icv] = I(-1.96) * [I(-1.96) + 1.96] = \\ &= 0.0599 * (0.0599 + 1.96) = 0.1209 \end{aligned}$$

$$\frac{V(u)}{s_u^2} = [1 - d(-icv)] = [1 - d(-1.96)] = (1 - 0.1209) = 0.879$$

Considerando el caso extremo al anterior, estudiamos la variable con:

$$\mathbf{m = -1.96, s_u^2 = 1, y icv = -1.96}$$

El valor del truncamiento asciende a 0.975, ya que estamos prescindiendo de la parte de densidad de la variable igual a dicha área, y que va desde  $-\infty$  hasta cero, quedando solo el 2.5% de la variable original con valores positivos, entre cero e  $\infty$ , que definen el espacio de la Normal truncada en cero.

Los valores que toman las funciones son:

$$f(-icv) = f(1.96) = 0.0584$$

$$F(-icv) = F(1.96) = 0.975$$

$$I(-icv) = I(1.96) = \frac{f(1.96)}{1 - F(1.96)} = \frac{0.0584}{1 - 0.975} = 2.336$$

La esperanza matemática de  $u_i$  será:

$$E(u) = m + s_u * I(-icv) = -1.96 + 1 * 2.336 = 0.376$$

La relación de varianzas tomará el valor:

$$d(-icv) = I(-icv) * [I(-icv) + icv] = I(1.96) * [I(1.96) - 1.96] = 2.336 * (2.336 - 1.96) = 0.8783$$

$$\frac{V(u)}{s_u^2} = [1 - d(-icv)] = [1 - d(1.96)] = (1 - 0.8783) = 0.1217$$

Vemos por tanto, que el factor de reducción de varianza puede alcanzar un valor igual al 12.17 %, e incluso menor si fuera menor el valor de  $icv$ .

La especificación **Normal-Seminormal** corresponde al caso particular en que

$$m = 0 \text{ y } icv = 0, \text{ para cualquier valor de } s_u^2.$$

Evidentemente, en este caso el valor del truncamiento es de 0.5, al quedar la variable reducida solo a la parte positiva que es exactamente la mitad.

Si particularizamos los cálculos para  $icv = 0$ , tenemos:

$$f(-icv) = f(0) = 0.3989$$

$$F(-icv) = F(0) = 0.5$$

$$I(-icv) = I(0) = \frac{f(0)}{1 - F(0)} = \frac{0.3989}{1 - 0.5} = 0.7978$$

La esperanza matemática de  $u_i$ , sí depende del valor de  $\sigma_u$ , de modo que para  $\sigma_u = 1$  será:

$$E(u) = \mathbf{m} + \mathbf{s}_u * \mathbf{I}(-icv) = 0 + 1 * 0.7978 = 0.7978$$

La relación de varianzas tomará el valor:

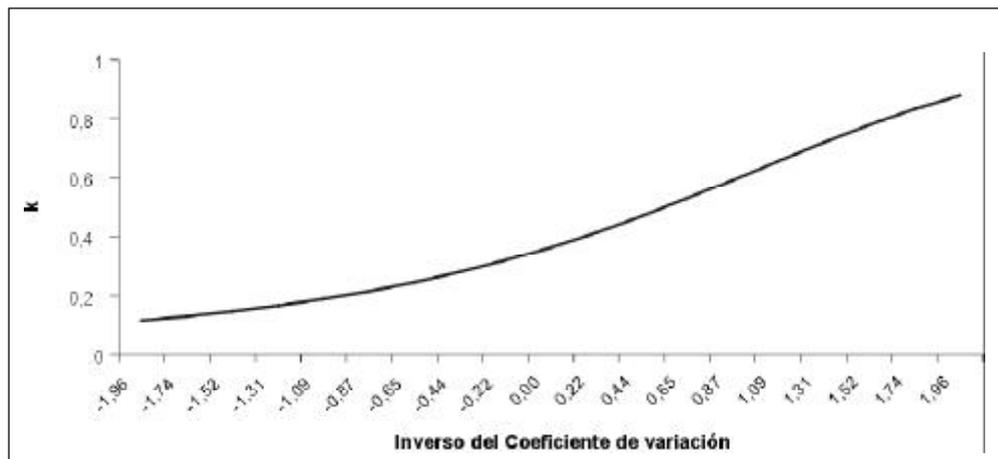
$$\mathbf{d}(-icv) = \mathbf{I}(-icv) * [\mathbf{I}(-icv) + icv] = \mathbf{I}(0) * [\mathbf{I}(0) + 0] = 0.7978 * (0.7978 + 0) = 0.6364$$

$$\frac{V(u)}{\mathbf{s}_u^2} = [1 - \mathbf{d}(-icv)] = [1 - \mathbf{d}(0)] = (1 - 0.6364) = 0.3635$$

En consecuencia, la reducción de varianza que tiene lugar al truncar una distribución Normal a Seminormal es del 36.35%, para cualquier valor  $\sigma_u^2$ .

En general, vemos que la reducción de varianzas que venimos comentando, varía según el valor que adopte el icv y además la expresión que la determina incluye las funciones de densidad y distribución de la Normal Tipificada particularizada para dicho ratio, cambiado de signo. Por este motivo, habría que calcularla para cada valor concreto de icv que nos encontremos en la práctica. Para obviar este problema, presentamos en el gráfico nº1 una representación de la relación  $k = \frac{V(u)}{\mathbf{s}_u^2}$  en función de icv, en un contexto casi continuo para valores de icv entre -1.96 y 1.96.

**Gráfico 1. Representación del factor de reducción en función del inverso del coeficiente de variación**



Fuente: Elaboración propia

Realizaremos a continuación un estudio de la repercusión que tiene dicha relación en los parámetros  $\lambda$  y  $\gamma$ , para lo cual denominaremos  $k$  al valor que toma  $\frac{V(u)}{s_u^2}$ .

### Estudio del parámetro $l$

Para valorar el error de interpretación cometido al analizar  $\lambda$  en lugar de  $\lambda_R$  hay que tener en cuenta la relación:

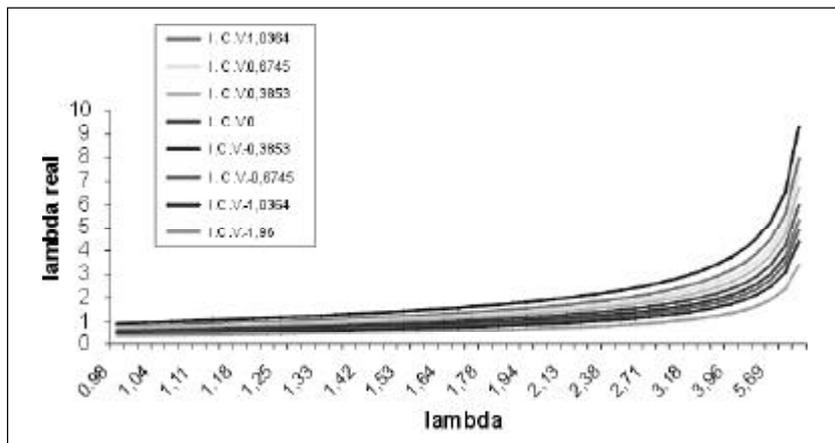
$$I_R = I * \left( \frac{V(u)}{s_u^2} \right)^{1/2} = I * k^{1/2}$$

De la expresión anterior se deduce que la reducción en el parámetro  $\lambda$  coincide exactamente con la raíz cuadrada de la de las varianzas, siendo para el caso de la especificación Normal-Seminormal la siguiente:

$$I_R = I * (0.3635)^{1/2} = I * 0.6029$$

Teniendo esto en cuenta, dicha relación depende del coeficiente de variación. Hemos considerado interesante mostrar en el gráfico n° 2 la representación de  $\lambda_R$  en función de  $\lambda$ , para cada valor de icv.

**Gráfico 2. Representación de valores de lambda real en función de lambda**



### Estudio del parámetro $g$

Para realizar el análisis del error de interpretación del parámetro  $g$ , nos centraremos en la relación que liga los valores de  $\gamma$  y  $\gamma_R$  que será:

$$g_R = \frac{V(u)}{s_v^2 + V(u)} = \frac{k * s_u^2}{s_v^2 + k * s_u^2} = \frac{k * g}{1 + (k - 1) * g}$$

Teniendo en cuenta que para cada valor de icv tendremos el correspondiente de  $\kappa$ , está claro que habrá también una relación distinta entre los parámetros  $\gamma$  y  $\gamma_R$ . Por tanto, el error de interpretación también será distinto para cada  $\gamma$  y para cada icv.

Particularizando para algunos valores concretos de icv, estudiaremos el siguiente:

$$m = 1.96, s_u^2 = 1, \text{ y icv} = 1.96,$$

Dado que  $\kappa$  es igual a 0.879, la función que liga los parámetros de interés será:

$$g_R = \frac{0.879 * g}{1 - 0.121 * g}, \text{ de modo que para } \gamma = 0.1, \gamma_R = 0.0889, \text{ con lo que la relación entre los}$$

parámetros será  $\frac{g_R}{g} = 0.889$  y  $\frac{g}{g_R} = 1.125$ , lo que supone una sobrevaloración de la inci-

dencia de la ineficiencia en el error compuesto de un 12.5%. En el caso opuesto de  $\gamma = 0.9$ ,

$\gamma_R = 0.8877$ , con lo que  $\frac{g_R}{g} = 0.9863$  y  $\frac{g}{g_R} = 1.014$ , siendo en este caso insignificante el

efecto estudiado.

Analizamos a continuación el caso extremo:

$$m = -1.96, s_u^2 = 1, \text{ y icv} = -1.96.$$

Para estas características de la variable, el truncamiento es del 97.5%, y el valor de  $\kappa$  es igual a 0.1217. Por tanto, la relación que debemos de aplicar será:

$$g_R = \frac{0.1217 * g}{1 - 0.8783 * g}.$$

Cuando  $\gamma = 0.1$ , resulta  $\gamma_R = 0.013$ , tomando la relación de parámetros los siguientes

valores:  $\frac{g_R}{g} = 0.13$  y  $\frac{g}{g_R} = 7.69$ . Este resultado deja patente que se puede producir una

sobrevaloración de 769 % de la incidencia de la ineficiencia al compararla con el error que queda fuera de control de la empresa. Estaríamos en este caso analizando un sistema totalmente eficiente y le asignaríamos por error un 10% de presencia de la variable que representa la ineficiencia en el error del modelo econométrico.

Para  $\gamma = 0.9$ , el parámetro  $\gamma_R$  es igual a 0.5227, y por tanto  $\frac{g_R}{g} = 0.58$  y  $\frac{g}{g_R} = 1.55$ .

Supone la relación anterior una sobrevaloración del 155% sobre el valor real de  $\gamma$ .

Encontramos por tanto, que el error de la sobrevaloración aumenta lógicamente con el tamaño del truncamiento y disminuye a medida que la variable  $u_1$  cobra más importancia en la composición del error.

Estudiaremos ahora la variable con:

$$m = 0, \text{ y } icv = 0.$$

Se trata de la distribución **Normal- Seminormal** siendo en este caso  $\kappa=0.3635$ , y la relación entre  $\gamma$  y  $\gamma_R$ , la siguiente:

$$g_R = \frac{0.3635 * g}{1 - 0.6365 * g}.$$

Cuando  $\gamma = 0.1$ , el valor de  $\gamma_R = 0.0388$ , y las relaciones estudiadas serán:

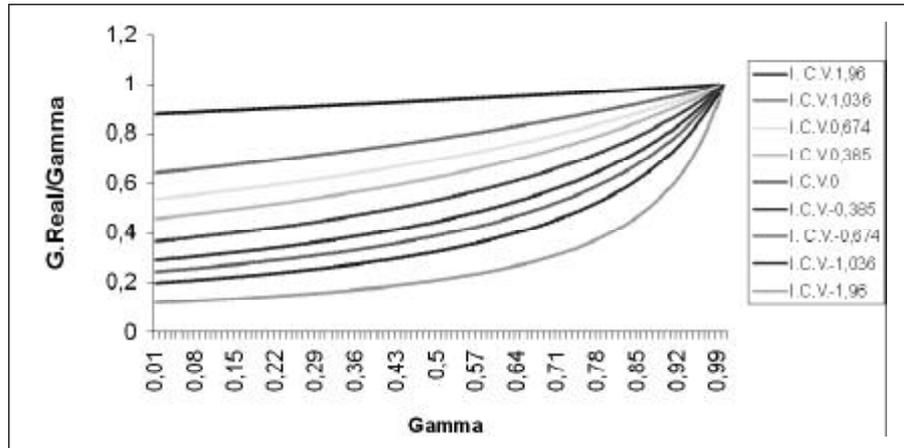
$$\frac{g_R}{g} = 0.38 \text{ y } \frac{g}{g_R} = 2.63.$$

Cuando  $\gamma = 0.9$ , resulta  $\gamma_R = 0.7659$ . En este caso los valores que toman las relaciones son:  $\frac{g_R}{g} = 0.851$ . y  $\frac{g}{g_R} = 1.175$ .

Por último, con el fin de analizar situaciones intermedias, seguimos estudiando el caso **Normal-Seminormal** que recoge una situación de truncamiento medio. Si consideramos el caso de composición de  $e$  en que  $\gamma = 0.5$ , que corresponde a una situación en que la ineficiencia tiene la misma incidencia que el error simétrico, el valor de  $\gamma_R$  será igual a 0.266. Esto supone que las relaciones toman el siguiente valor

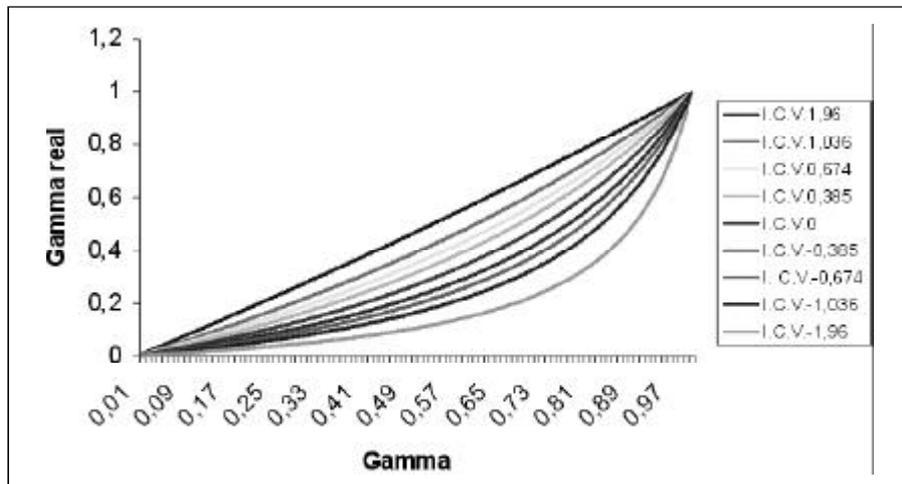
$$\frac{g_R}{g} = 0.5331 \text{ y } \frac{g}{g_R} = 1.875.$$

En el gráfico nº 3 se ha representado  $\frac{g}{g_R}$  en función de  $\gamma$ , para distintos valores de  $icv$ , con lo que se ilustra la magnitud del error de interpretación en cada caso particular.

**Gráfico 3. Representación del error de interpretación en función de Gamma**

Fuente: Elaboración propia.

Además el gráfico nº 4 nos permite conocer el valor de  $\gamma_R$ , dado un coeficiente de variación y un  $\gamma$  estimado.

**Gráfico 4. Representación de los valores de Gamma real en función de los de Gamma**

Fuente: Elaboración propia.

### 3. CONCLUSIONES

En el presente trabajo se ha valorado el error de interpretación que se puede producir cuando se toman conclusiones con respecto a la incidencia de la ineficiencia en el modelo de frontera estocástica, tomando como base los valores de  $\lambda$  y  $\gamma$ .

Los desarrollos expuestos dejan patente que los parámetros que hay que analizar son los que se han definido como  $\lambda_r$  y  $\gamma_r$ , que miden la proporción de varianza real de la variable  $u_i$  en el error compuesto  $\varepsilon$ .

Los errores de interpretación se producen debido a la disminución de varianza ocasionada en la variable Normal a ser truncada en el valor cero, siendo esta reducción dependiente solo del inverso del coeficiente de variación definido como .

Tras la valoración efectuada en el estudio, concluimos que el error en que se incurre debido a una interpretación inadecuada puede llegar a provocar una sobrevaloración superior al 750% en los casos más desfavorables.

### 4. BIBLIOGRAFÍA

- AIGNER, D.J., LOVELL, C.A.K. y SCHMIDT, P. (1977). Formulation and estimation of stochastic frontier production, function models. *J. Econometrics*, nº 6, pp 21-37.
- BAGI, F. S. (1984). Stochastic frontier production function and farm-level technical efficiency of full-time and part-time farms in west Tennessee. *North Central J. Agr. Econ.* Nº1, pp 48-55.
- BATTESE, G.E., (1992), "*Frontier Production Functions and Technical Efficiency: A Survey of Empirical Applications in Agricultural Economics*", *Agricultural Economics*, 7, 185-208
- BATTESE, G.E., COELLI, T.J., (1995), "*A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data*", *Empirical Economics*, 20, 325-332
- BATTESE, G.E. y CORRA, G.S. (1977). Estimation of a production frontier model: with application to the pastoral zone of eastern Australia. *Aust. J. Agric. Econ.*, nº 21, pp 169-179.
- COELLI, T.J., (1992), "*A Computer Program for Frontier Production Function Estimation: FRONTIER, Version 2.0*", *Economics Letters*, 39, 29-32
- COELLI, T.J., (1996), "*A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation*", CEPA Working Papers, No. 7/96, Department of Econometrics, University of New England, Armidale, Australia
- COELLI, T.J., (1996), "*Measurement and sources of technical inefficiency in australian coal-fired electricity generation*" mimeo, Department of Econometrics, University of New England, Armidale, Australia
- COELLI, T.J., BATTESE, G.E., (1996), "*Identification of factors which influence the technical inefficiency of indian farmers*". *Australian Journal of Agricultural Economics*, Vol. 40, 2, 103- 128.

- FRIED, H.O., LOVELL, C.A.K., SCHMIDT, S.S., YAISAWARNG, S., (2002), " *Accounting for Environmental Effects and Statistical Noise in Data Envelopment Analysis*" Journal of Productivity Analysis, 17, 157-174.
- GREENE, W.H. (1992). *Limdep. Versión 6.0 User's Manual and Reference Guide*. De. Econometric Software, Inc. Bellport, New York.
- HALLAN, D., MACHADO, F. (1996), " *Efficiency analysis with panel data: a study of Portuguese dairy farms*" ,European Review of Agricultural Economics 23 (1) :79-93
- PITT, M.M., LEE, L-F., (1981), " *Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry*" , Journal of Development Economics, 9, 43-64
- SEIFORD, L.M., THRALL, R.M., (1990), " *Recent Developments in DEA: The Mathematical Approach to Frontier Analysis*" , Journal of Econometrics, 46, 7-38
- STEVENSON, R.E. (1980). Likelihood function for generalized stochastic frontier estimation. J. Econometrics, nº 13, pp 57-66.
- TADESSE, B. y KRISHNAMOORTHY, S. (1997). Technical efficiency in paddy of Tamil Nadu: an analysis based on farm size and ecological zone. Agricultural Economics, nº16, pp 185-192
- TVERAS, R. y BATTESE, G.E. (2001). " *The influence of Regional Agglomeration Externalities on Efficiency in Norwegian Salmon Aquaculture*". CEPA Working Papers, No.1/2001. Department of Econometrics, University of New England, Armidale, Australia
- TIAN, W.M. y WAN, G.H. (2000). " *Technical efficiency and its determinants in China's grain production*" Journal of Productivity Analysis. Volume 13, number 2.
- TIMMER, C. P. (1970). On measuring technical efficiency. Food Research institute Studies, nº 9, pp. 99-171.