

# EL EFECTO DE LA DISTRIBUCIÓN A PRIORI EN MODELOS CON FRONTERA ESTOCÁSTICA. COMPARACIÓN CON MÁXIMA VEROSIMILITUD.

**FRANCISCO JAVIER ORTEGA IRIZO**

Departamento de Economía Aplicada I. Universidad de Sevilla  
Avenida Ramón y Cajal, nº1. 41018 – SEVILLA (ESPAÑA)

**JOSÉ MANUEL GAVILÁN RUIZ**

Departamento de Economía Aplicada I. Universidad de Sevilla  
Avenida Ramón y Cajal, nº1. 41018 – SEVILLA (ESPAÑA)

**JOSÉ ANTONIO CAMÚÑEZ RUIZ**

Departamento de Economía Aplicada I. Universidad de Sevilla  
Avenida Ramón y Cajal, nº1. 41018 – SEVILLA (ESPAÑA)

e-mail: [fjortega@us.es](mailto:fjortega@us.es) – [gavi@us.es](mailto:gavi@us.es) – [camunez@us.es](mailto:camunez@us.es)  
Teléfono: 954556970 – 954556970 - 954557541

## Resumen

En este trabajo se analizan y comparan las propiedades de los estimadores máximo verosímil y bayesiano en el modelo de producción con frontera estocástica a través de un estudio de tipo Monte Carlo. En el caso bayesiano, se presta especial atención al efecto que tiene la elección del hiperparámetro que determina la eficiencia mediana de la distribución a priori. Los resultados muestran un mejor comportamiento de la estimación bayesiana, salvo en el caso poco probable de que se asigne un hiperparámetro muy alejado del valor real de la población.

*Palabras clave:* Frontera Estocástica, Monte Carlo, Estimador Máximo Verosímil, Estimador Bayes.

*Área Temática:* Métodos Cuantitativos para la Economía y la Empresa.

## Abstract

In this paper, the maximum likelihood and the Bayesian methodologies are analysed and compared through a Monte Carlo study in the setting of half-normal stochastic frontier production models. In the Bayesian case, special emphasis is placed on the effect that the choice of the hyperparameter that determines the prior median efficiency has on the estimations. The results show a better behaviour of the Bayesian estimation, except in the unlikely case in which researchers assign a value to the aforementioned hyperparameter which greatly differs from its actual value.

*Key Words:* Stochastic Frontier, Monte Carlo, Maximum Likelihood, Bayesian Estimator.

*Thematic Area:* Quantitative Methods for Economics and Business Administration.

# 1. INTRODUCCIÓN

El modelo de producción con frontera estocástica (MPFE) tiene sus orígenes en los trabajos de Aigner *et al.* (1977), Battese y Corra (1977) y Meusen y van den Broeck (1977). El objetivo de este tipo de modelos es analizar la eficiencia de un proceso productivo en términos de las desviaciones observadas entre la producción real y una frontera ideal de output máximo. En términos econométricos, dichas desviaciones pueden ser identificadas con perturbaciones aleatorias en un modelo de regresión.

Estos modelos han tenido bastante difusión en la literatura científica desde su aparición, aplicados a una gran variedad de sectores productivos. Por ejemplo, han sido muy utilizados en agricultura y pesca (Battese y Broca 1997, García *et al.* 2004), en el funcionamiento de puertos y aeropuertos (Barros 2005, 2008), hospitales (O'Donnell and Nguyen 2013), sector bancario (Brissimis *et al.* 2010) o aplicados al análisis de la producción científica (Ortega y Gavilán 2013).

La formulación básica del MPFE es:

$$y_i = f(x_i, \beta) + v_i - u_i, \quad i = 1, \dots, n,$$

donde  $y_i$  es la producción de la firma  $i$ -ésima,  $x_i$  es el vector de todos sus inputs,  $\beta$  es un vector de parámetros a estimar y  $f(\bullet)$  es la función de producción.

La perturbación aleatoria  $\varepsilon_i = v_i - u_i$  se compone de dos partes (de ahí que también hablemos de modelo de error compuesto),  $v_i \in \mathbb{R}$  que recoge las fuentes de variación aleatorias y  $u_i > 0$  que recoge la ineficiencia en el proceso productivo. Comúnmente se supone  $v_i \sim N(0, \sigma_v^2)$ . Para  $u_i$  hemos de elegir una distribución de probabilidad positiva. Vamos a considerar la hipótesis más habitual, es decir, que  $u_i = |u_i^*|$ , donde  $u_i^* \sim N(0, \sigma_u^2)$ . Por definición, se dice que las perturbaciones  $u_i$  siguen una distribución Half-Normal, lo cual representamos por  $u_i \sim HN(0, \sigma_u^2)$ . Adicionalmente, se supone que todas las perturbaciones (tanto  $v_i$  como  $u_i$ ) son independientes. Es preciso señalar que se ha considerado con frecuencia otro tipo de densidad de probabilidad para las perturbaciones  $u_i$ , entre las que cabe destacar la distribución Normal truncada en un parámetro  $\mu$  no necesariamente nulo y las distribuciones Exponencial y Gamma, que han sido ampliamente utilizadas sobre todo cuando se adopta el enfoque bayesiano (Koop *et al.* 1995, Osiewalski and Steel 1998, Koop and Stell 2003).

En un reciente artículo, Ortega y Gavilán (2014) estudian el comportamiento de las estimaciones máximo verosímil (MV) y bayesiana en un MPFE a través de un análisis de tipo Monte Carlo. Para ello, se obtienen muestras simuladas para diferentes valores del espacio paramétrico y se comparan ambos estimadores en base al criterio del error cuadrático medio. Para obtener los estimadores bayesianos, los autores utilizan una única distribución a priori. La principal conclusión es que la metodología bayesiana, con la distribución a priori seleccionada, tiene un comportamiento notablemente mejor que el estimador MV en la práctica totalidad del espacio paramétrico analizado.

En este trabajo se amplía el análisis de Ortega y Gavilán (2014) considerando diferentes distribuciones a priori, con el doble objetivo de hacer un análisis de sensibilidad de la metodología bayesiana a la elección de la distribución a priori y de comprobar si el comportamiento del estimador bayesiano sigue siendo mejor que el MV independientemente de la distribución a priori elegida.

La función de producción usada es lineal, es decir,  $f(x_i, \beta) = x_i' \beta$ , donde el vector  $\beta$  contiene un término constante. Observemos que si  $x_i$  e  $y_i$  están medidos en escala logarítmica, entonces estamos ante una función de producción de Cobb-Douglas. En este caso, que es el más habitual, en lugar de interesarnos directamente en  $u_i$  como medida de ineficiencia, se suele estimar  $\exp(-u_i)$  que es una medida de eficiencia acotada entre 0 y 1. Por simplicidad, haremos las simulaciones en base a un modelo con ordenada en el origen y una única variable explicativa, pues se ha comprobado que el uso de un mayor número de variables no afecta a los resultados principales.

La estimación MV del MPFE está implementada desde hace tiempo en software estadístico diverso, por ejemplo, en FRONTIER, LIMDEP o STATA. Asimismo, en el entorno R podemos encontrar diferentes paquetes específicos con este propósito. La estimación bayesiana a través del algoritmo de Gibbs (Gelfand and Smith, 1990) es una herramienta muy adecuada en este contexto debido a la complejidad de los modelos de frontera estocástica. Uno de los problemas principales del enfoque bayesiano, sobre todo de

cara a las aplicaciones, ha sido la falta de un software fácil de utilizar y lo suficientemente versátil como para adaptarse a distintas situaciones y conjuntos de datos. En Griffin and Steell (2007), se utiliza el software libre WinBUGS para estimar diversas variantes de MPFE, comprobándose que este software es una herramienta potente y flexible en este contexto, que puede ser manejado con facilidad por parte del investigador y que consigue estimar los modelos en un tiempo razonable.

En este artículo utilizamos el software FRONTIER para obtener los estimadores MV y el software WinBUGS para el caso bayesiano. Tanto uno como otro pueden integrarse en el entorno del software R, lo que facilita la aplicación de ambas metodologías al mismo conjunto de muestras simuladas para que, de esta forma, la comparación de resultados sea más fiable.

Otros trabajos en los que se hacen comparativas a través de métodos de simulación en el MPFE son Coelli (1995), donde se compara el estimador MV con el de mínimos cuadrados corregidos y Zhang (1999) donde se compara la estimación bayesiana con la MV, aunque ésta se hace para un único punto del espacio paramétrico. En Ortega et al. (2009) se lleva a cabo una comparación entre ambas metodologías pero aplicadas sobre el modelo de producción con frontera determinista.

Como criterio de comparación usamos el error cuadrático medio (ECM), haciendo especial énfasis en el parámetro que indica qué proporción de varianza del error compuesto de debe a ineficiencia y en la estimación de las eficiencias individuales.

A partir de aquí, en la sección 2 explicamos cómo obtener los estimadores MV y bayesianos usando FRONTIER y WinBUGS, respectivamente, todo ello en el entorno del software R; en la sección 3 describimos el diseño del experimento Monte Carlo; en la sección 4 ofrecemos los resultados obtenidos, así como la interpretación de los aspectos más interesantes. Por último, en la sección 5 exponemos las principales conclusiones de nuestro trabajo.

## 2. PROGRAMACIÓN DE LOS MÉTODOS DE ESTIMACIÓN UTILIZADOS

La estimación del MPFE la vamos a hacer usando la reparametrización  $\sigma^2 = \sigma_v^2 + \sigma_u^2$  y  $\gamma = \sigma_u^2 / \sigma^2$  que es la que se considera en el paquete FRONTIER. Observemos que  $\gamma$  es un parámetro que toma valores entre 0 y 1 y es un indicador de la proporción de varianza debida a la ineficiencia. Resaltemos que  $\gamma$  no es exactamente la proporción de varianzas, ya que  $\text{var}[u] = p\sigma_u^2$ , donde  $p = (\pi - 2) / \pi$ . Concretamente, si llamamos  $\gamma^*$  a la proporción de la varianza total debida a ineficiencia (es decir,  $\gamma^* = \text{var}[u] / (\text{var}[u] + \text{var}[v])$ ) es fácil ver que se verifica  $\gamma^* = \gamma / (\gamma + (1 - \gamma)p^{-1})$ .

Como se ha indicado en la introducción, la estimación MV la hemos llevado a cabo usando el software FRONTIER y la estimación bayesiana a través de WinBUGS, ambos integrados en el entorno del software R. Concretando más, para el estimador MV usamos el paquete *frontier* en su versión 0.996-6 el cual utiliza el código fuente *Fortran* del software FRONTIER 4.1 (Coelli, 1996). Para la inferencia bayesiana se ha usado el paquete *R2WinBUGS*, el cual enlaza y abre el programa WinBUGS v1.4, le transfiere los datos y recoge los resultados hacia el entorno R. Una detallada exposición del funcionamiento del paquete *R2WinBUGS* junto con algunos ejemplos de aplicación puede encontrarse en Sturtz et al. (2005).

Destaquemos también que para la implementación del modelo en WinBUGS, se han generado tres cadenas de Markov de 1500 iteraciones cada una, de las cuales se desechan las 500 primeras, lo que conlleva un total de 3000 muestras simuladas. Como punto de partida se toman los estimadores de Mínimos Cuadrados Corregidos para los parámetros  $\beta_0$  y  $\beta_1$  (Green, 1980), mientras que para el resto de parámetros se generan valores aleatorios.

Siguiendo a Ortega y Gavilán (2014), para los parámetros  $\beta_0, \beta_1$  y  $\sigma_v^{-2}$  usamos las distribuciones no informativas  $\beta_0 \sim N(0, 10^6)$ ,  $\beta_1 \sim N(0, 10^6)$  y  $\sigma_v^{-2} \sim Ga(10^{-3}, 10^{-3})$ , donde  $N(\mu, \delta^2)$  representa una distribución Normal de media  $\mu$  y varianza  $\delta^2$ , mientras que  $Ga(a_0, a_1)$  representa una distribución Gamma con media  $a_0/a_1$  y parámetro de forma  $a_0$ . Estas son las mismas distribuciones que consideran Griffin y Steel (2007).

La distribución a priori para el parámetro  $\sigma_u^{-2}$  se suele elegir dentro de la familia Gamma, fijando los hiperparámetros de forma que la mediana a priori recoja las creencias acerca de la eficiencia mediana del

sector productivo en el que se está llevando a cabo el análisis (Van den Broeck et al. 1994, Zhang, X. 1999). Griffin and Steel (2007), en su ejemplo aplicado al mercado de la electricidad en Estados Unidos, usan la distribución a priori  $\sigma_u^{-2} \sim Ga(1, 0.0267)$ , lo que supone asumir que la eficiencia mediana a priori del sector es 0.875. Como se expondrá con mayor detalle en la siguiente sección, nosotros vamos a considerar  $\sigma_u^{-2} \sim Ga(1, \lambda_0)$ , asignando al parámetro  $\lambda_0$  diferentes valores para así poder analizar el efecto de esta elección sobre las estimaciones.

### 3. DISEÑO DEL EXPERIMENTO MONTE CARLO.

Como hemos indicado en la introducción, las simulaciones se hacen sobre un modelo con una variable explicativa más un término constante. El espacio muestral del experimento viene dado inicialmente por los distintos valores de los parámetros  $\beta_0$ ,  $\beta_1$ ,  $\sigma^2$  y  $\gamma$ , del tamaño muestral  $n$  así como de los valores del vector de observaciones de la covariable  $x$ . Centramos el análisis en el comportamiento de las estimaciones del parámetro  $\gamma$ , así como de las estimaciones de las eficiencias individuales.

Sin pérdida de generalidad, podemos fijar los valores de  $\beta_0$  y  $\beta_1$  (Zhang 1999, Coelli, 1995). Se ha considerado  $\beta_0 = \beta_1 = 1$ . Con respecto a los valores de la covariable  $x$ , se han generado de acuerdo a un modelo uniforme en el intervalo  $[0, 10]$ . Pretendemos analizar el comportamiento de ambos métodos tanto en muestras pequeñas como en muestras de gran tamaño, por lo que hemos seleccionado  $n \in \{30, 50, 100, 500\}$ . Con respecto al parámetro  $\gamma$  se consideran los valores  $\{0.1, 0.3, 0.5, 0.7, 0.9\}$ .

Para estudiar el efecto de la elección de la distribución a priori del parámetro  $\sigma_u^{-2}$ , todas las muestras simuladas proceden de una población con un valor de eficiencia mediana  $r^* = 0.5$  y en todos los casos, además del EMV, se obtienen los estimadores bayesianos con cinco distribuciones cuyas eficiencias medianas a priori son, respectivamente,  $r \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ .

Dado que la eficiencia se ha definido como  $\exp(-u_i)$ , la relación entre la mediana de la eficiencia  $r$  y el parámetro  $\sigma_U^2$  viene dada por la ecuación (Zhang, 1999)

$$\int_0^{-Lnr} \frac{2}{\sqrt{2\pi\sigma_U^2}} \text{Exp}\left\{-\frac{U^2}{2\sigma_U^2}\right\} dU = 0.5.$$

Resolviendo esta ecuación para  $r^* = 0.5$  obtenemos que  $\sigma_U^{2*} = 1.056$ . Puesto que  $\sigma_V^2 = \sigma_U^2(1-\gamma)/\gamma$ , deducimos que  $\sigma_V^2 \in \{9.504, 2.464, 1.056, 0.453, 0.117\}$ . Observemos que al fijar  $r^* = 0.5$  (y por tanto considerar un único valor  $\sigma_U^{2*}$ ), para recorrer el espacio paramétrico de  $\gamma$  es necesario considerar muestras cuya variabilidad total es muy diferente, pues  $\sigma^2 = \sigma_U^2 + \sigma_V^2$  es igual a 10.56 para  $\gamma = 0.1$  mientras que toma el valor 1.173 cuando  $\gamma = 0.9$ .

Por otra parte, tomando la distribución a priori  $\sigma_u^{-2} \sim Ga(1, \lambda_0)$ , la mediana a priori de la eficiencia es  $r = \text{Exp}\{-t_{2,0.75}\sqrt{\lambda_0}\}$ , donde  $t_{2,0.75}$  es el cuantil de orden 0.75 de una distribución t de Student con 2 grados

de libertad (Van den Broeck et al. 1994). Despejando, obtenemos  $\lambda_0 = \left(-\frac{Lnr}{t_{2,0.75}}\right)^2$ . La correspondencia entre

los valores de eficiencia a priori y los hiperparámetros de la distribución de  $\sigma_u^{-2}$  se refleja en la siguiente tabla:

|             |       |       |       |       |       |
|-------------|-------|-------|-------|-------|-------|
| $r$         | 0.1   | 0.3   | 0.5   | 0.7   | 0.9   |
| $\lambda_0$ | 7.962 | 2.177 | 0.722 | 0.191 | 0.017 |

Se simulan, pues, cinco valores de  $\gamma$  y cuatro de  $n$ , lo que conlleva un total de veinte combinaciones. En cada una de ellas, se hacen  $m=1000$  replicaciones del modelo. Para obtener los números pseudoaleatorios, utilizamos los generadores implementados por defecto en el software R. La estimación

MV y las cinco estimaciones bayesianas en cada caso, se efectúa según el procedimiento descrito en la sección anterior.

Para cada parámetro, se calcula el sesgo medio (SM) y el error cuadrático medio (ECM) observado en las  $m$  replicaciones con cada uno de los dos métodos. Para el análisis de las eficiencias, no nos limitaremos a considerar la eficiencia media, sino que calcularemos tanto el SM como el ECM para cada una de las eficiencias individualmente. Posteriormente, como indicador conjunto ofrecemos el promedio de los SM y los ECM de cada una de las firmas. Es decir, si llamamos  $E_i$  a la eficiencia de la firma  $i$ -ésima y  $\hat{E}_{ij}$  al correspondiente valor estimado (por cualquiera de los dos métodos) en la replicación  $j$ -ésima, tenemos que

$$SM_i = m^{-1} \sum_{j=1}^m (\hat{E}_{ij} - E_i) \quad \text{y} \quad ECM_i = m^{-1} \sum_{j=1}^m (\hat{E}_{ij} - E_i)^2, \quad \text{para } i=1, \dots, n.$$

Posteriormente, el SM y el ECM de las eficiencias se obtienen como  $SM = n^{-1} \sum_{i=1}^n SM_i$  y  $ECM = n^{-1} \sum_{i=1}^n ECM_i$ , respectivamente.

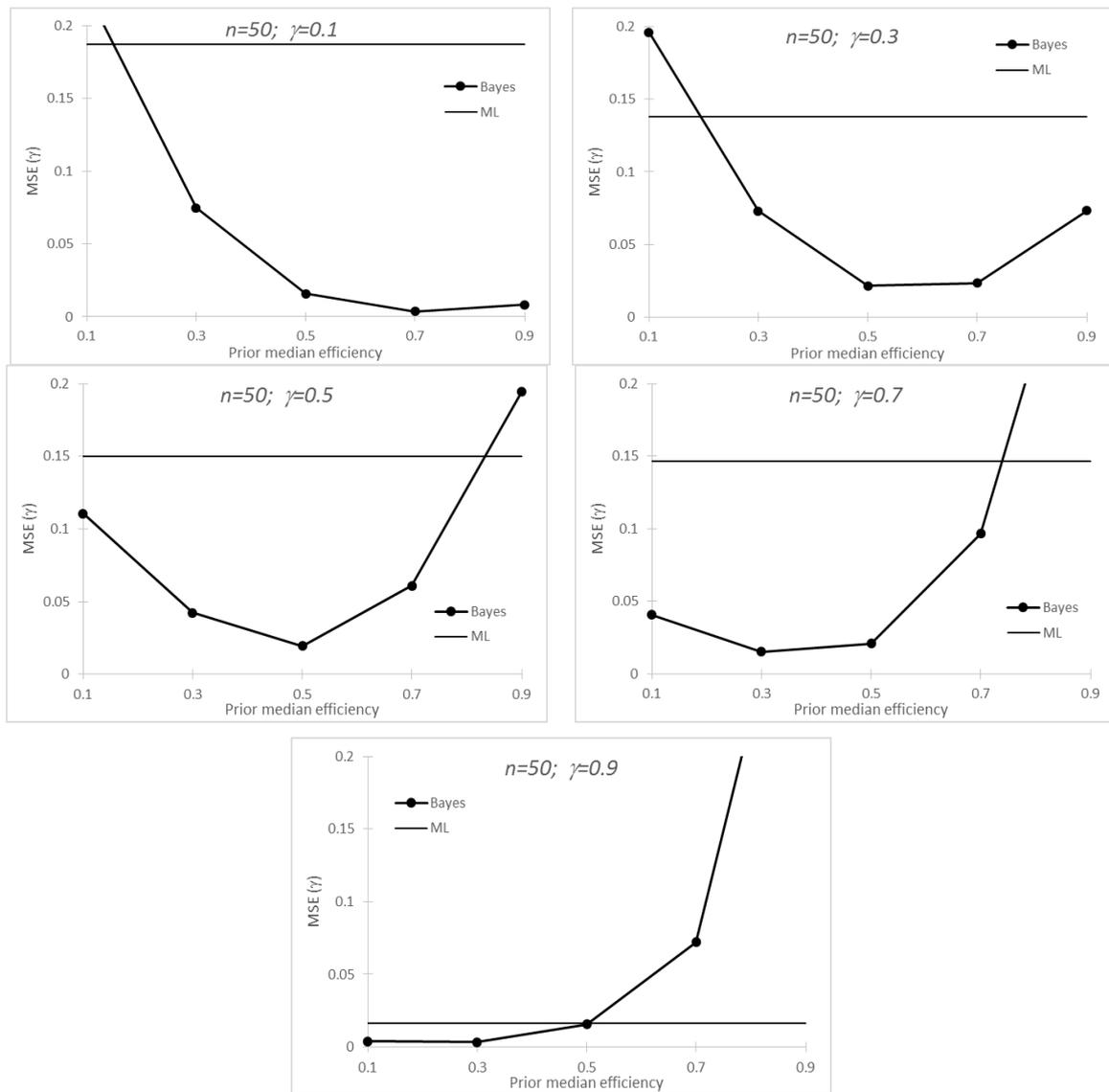
Antes de analizar los resultados, señalemos que en Ortega y Gavilán (2014), se generan muestras en las que  $\sigma^2 = 1$  y  $\gamma$  varía desde 0 hasta 1, lo que conlleva que las eficiencias medianas de las muestras se muevan desde  $r=1$  (correspondiente a  $\gamma=0$ ) hasta  $r=0.509$  (correspondiente a  $\gamma=1$ ). Además, se utiliza una única distribución  $\sigma_u^{-2} \sim Ga(1, 0.7216)$  para asumir una eficiencia mediana a priori igual a 0.5. En ese artículo se aprecia que las estimaciones bayesianas tienen mejor comportamiento a medida que el parámetro  $\gamma$  se acerca a 1 (es decir, a medida que la eficiencia mediana a priori es más cercana a la eficiencia mediana real de la población).

#### 4. RESULTADOS DEL EXPERIMENTO MONTE CARLO.

Los resultados completos del experimento Monte Carlo consisten en el sesgo y ECM de todos los parámetros y de las eficiencias individuales en las diferentes situaciones consideradas. Aquí vamos a analizar en mayor profundidad los resultados más relevantes, centrando nuestra atención en el criterio del ECM, que es al fin y al cabo el que indica qué metodología consigue reproducir mejor los verdaderos valores del modelo considerado. Nos centraremos en el comportamiento de las estimaciones del parámetro  $\gamma$  y de las eficiencias individuales. Como ya hemos indicado, el parámetro  $\gamma$  es el que recoge la estructura del error compuesto y, como es conocido, uno de los inconvenientes de este tipo de modelos es la dificultad para identificar qué proporción del error total se debe a ineficiencia y qué parte se debe a los efectos aleatorios. Por otra parte, la correcta estimación de  $\gamma$  es muy importante para establecer las eficiencias individuales de cada firma, que es uno de los objetivos fundamentales al usar este tipo de modelos.

En el Anexo ofrecemos el ECM del parámetro  $\gamma$  y de las eficiencias individuales en todas las simulaciones efectuadas. Con rellenos de fondo hemos resaltado el menor ECM obtenido para cada valor del espacio paramétrico y tamaño muestral. Hemos hecho representaciones gráficas de los resultados obtenidos para facilitar su interpretación.

Al igual que se aprecia en Ortega y Gavilán (2014), no hay grandes diferencias en los resultados generales cuando analizamos el ECM del parámetro  $\gamma$  o el de las eficiencias. Obsérvese, en particular, que los valores mínimos de ECM correspondientes al parámetro  $\gamma$  y a la eficiencia se alcanzan casi siempre en la misma fila de la tabla. Este hecho pone de manifiesto la incidencia directa que tiene una correcta estimación del parámetro  $\gamma$  sobre la determinación de las eficiencias de las firmas. Tampoco se aprecian diferencias importantes en función del tamaño de la muestra, salvo la lógica disminución del ECM a medida que aumentamos el tamaño muestral. Por ello y para simplificar, se ofrece en la figura 1 el ECM correspondiente a todos los valores del parámetro  $\gamma$  para el tamaño muestral  $n = 50$ . Hemos preferido mantener en los cinco gráficos la misma escala en el eje de ordenadas, aunque algunos valores son mayores que el máximo seleccionado (es decir, 0.2).



**Figura 1.** ECM obtenido para  $\gamma \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$  en función de la eficiencia mediana a priori asumida y para  $n=50$

Podemos observar en la tabla del Anexo como el mínimo de ECM se alcanza en la mayoría de los casos con la estimación bayesiana cuando se asume una eficiencia mediana a priori  $r=0.5$ , que es la mediana real de la población, sobre todo en los valores centrales del parámetro  $\gamma$ . En algunos casos dicho mínimo se obtiene asumiendo a priori  $r=0.3$  o  $r=0.7$  si bien debemos resaltar que en todo el espacio paramétrico y para todos los tamaños muestrales las diferencias de ECM en la metodología bayesiana para los valores de  $r \in \{0.3, 0.5, 0.7\}$  no son, en general, importantes.

Otra conclusión interesante que podemos extraer es que el uso de las distribuciones a priori con  $r \in \{0.1, 0.9\}$  (es decir, asumir a priori un valor de eficiencia mediana muy alejado del valor real de la población que es  $r^* = 0.5$ ) empeora sensiblemente los resultados, superando con creces en muchos casos al ECM obtenido con el estimador MV. Se observa también que este efecto depende de la zona del espacio paramétrico de  $\gamma$  en el que nos encontremos. En efecto, una eficiencia mediana a priori muy inferior a la real conlleva un aumento importante del ECM para valores bajos del parámetro  $\gamma$  mientras que si la mediana a priori es muy superior a la real el aumento de ECM se produce en valores altos de  $\gamma$ .

Según se desprende del análisis de simulación efectuado y tratando de dar respuesta al doble objetivo planteado en la introducción de este trabajo, podemos decir que la metodología bayesiana es sensible a la elección de la distribución a priori sólo cuando se asigna una mediana a priori muy alejada de la real, en cuyo caso el comportamiento en cuanto al ECM es peor que el obtenido con la EMV. No obstante, y como conclusión más importante, si usamos una distribución correspondiente a una mediana a priori que esté en

un entorno relativamente amplio del valor real, la estimación bayesiana no es muy sensible a tal elección y presenta mejor comportamiento que el estimador MV.

En muchas situaciones prácticas, el investigador tiene información previa para poder asignar razonablemente un valor de eficiencia mediana a priori no muy alejado del valor real. En todo caso, el uso de una distribución a priori “moderada” parece garantizar un mejor comportamiento del estimador bayesiano frente al MV, por lo que recomendamos el uso de la metodología bayesiana para estimar el MPFE.

## 5. CONCLUSIONES.

La principal conclusión es que la estimación bayesiana en general es preferible a la estimación MV en el modelo de producción con frontera estocástica analizado. Esta conclusión se ha obtenido en base a que el ECM del parámetro  $\gamma$  y de las eficiencias estimadas es inferior en la inferencia bayesiana salvo en el caso poco probable de que el investigador asigne a priori un valor de eficiencia mediana muy alejado del valor real de la población.

Podemos añadir que, exceptuando el caso mencionado con anterioridad, la estimación bayesiana en el MPFE no es excesivamente sensible a la elección de la distribución a priori para el parámetro  $\sigma_u^{-2}$ .

Queremos destacar también que la estimación bayesiana se ha llevado a cabo usando el software bayesiano de propósito general WinBUGS, usado desde el entorno R, lo que hace que su uso sea viable y atractivo para investigadores aplicados con unos conocimientos básicos de estadística bayesiana.

Para finalizar, resaltemos que nuestro análisis se ha centrado en un modelo con datos de sección cruzada y en el que se ha supuesto que las perturbaciones  $u_i$  siguen un modelo Half-Normal. Ampliaciones naturales de este trabajo consisten en suponer otro tipo de modelos para las perturbaciones  $u_i$  y en la consideración de datos de panel.

## REFERENCIAS

- AIGNER, D.J.; LOVELL, C.A.; SCHMIDT, P. (1977), “Formulation and estimation of stochastic frontier production function models”, *Journal of Econometrics*, 6, pp. 21-37.
- BARROS, C.P. (2005), “Decomposing growth in Portuguese Seaports: A frontier cost approach”, *Maritime Economics&Logistic*, 7, pp. 297-315.
- BARROS, C.P. (2008), “Technical efficiency of UK airports”, *Journal of Air Transport Management*, 14, pp. 175-178.
- BATTESE, G.E.; CORRA, G.S. (1977), “Estimation of a production frontier model: With application to the Pastoral Zone of Eastern Australia”, *Australian Journal of Agricultural Economics*, 21, pp.169-179.
- BATTESE, G.E.; BROCA, S.S. (1997), “Functional forms of stochastic frontier production functions and models for technical inefficiency effects: A comparative study for wheat farmers in Pakistan”, *Journal of Productivity Analysis*, 8, pp.395-414.
- BRISSIMIS, S.N.; DELIS, M.D.; TSIONAS, E. (2010), “Technical and allocative efficiency in European banking”, *European Journal of Operational Research*, 204, pp. 153-163.
- COELLI, T. (1995), “Estimators and hypothesis test for a stochastic frontier function: A Monte Carlo analysis”, *The Journal of Productivity Analysis*, 6, pp. 247-268.
- COELLI, T. (1996), A guide to FRONTIER version 4.1: a computer program for frontier production function estimation. CEPA Working Paper 96/07, Department of Econometrics, University of New England, Armidale, Australia. <http://www.uq.edu.au/economics/cepa/software/FRONT41-xp1.zip>
- GARCÍA, J.J.; CASTILLA, D.; JIMÉNEZ, R. (2004): “Determination of technical efficiency of fisheries by stochastic frontier models: a case on the Gulf of Cádiz (Spain)”, *ICES Journal of Marine Science*, 61, pp.416-421.
- GELFAND, A.E.; SMITH, A.F. (1990): “Sampling-based approaches to calculating marginal densities”, *Journal of the American Statistical Association*, 85, pp. 398-409
- GREEN, W.H. (1980): “Maximum Likelihood Estimations of econometric frontier functions”, *Journal of Econometrics*, 13, pp. 27-56.

- GRIFFIN, J.E.; STEEL, M.F.J. (2007), "Bayesian stochastic frontier analysis using WinBUGS", *Journal of Productivity Analysis*, 27, pp. 163-176.
- KOOP, G.; STEEL, F.J. (2003), Koop, G.; Steel, M.F. (2003), "Bayesian analysis of stochastic frontier models". En: Baltagi, B.H. (editor), *A companion to theoretical econometrics*, Blackwell.
- KOOP, G.; STEEL, M.F.J.; OSIEWALSKI, J. (1995), "Posterior analysis of stochastic frontier models using Gibbs sampling", *Computational Statistics*, 10, pp. 353-373.
- MEEUSEN, W.; VAN DEN BROECK, J. (1977), "Efficiency estimation from Cobb-Douglas production functions with composed error", *International Economic Review*, 18, pp. 435-444.
- O'DONNELL, C.J.; NGUYEN, K. (2013), "An econometric approach to estimating support prices and measures of productivity change in public hospitals", *Journal of Productivity Analysis*, 40, pp. 323-335.
- ORTEGA, F.J.; BASULTO, J.; CAMÚÑEZ, J.A. (2009), "Comparing Bayesian and corrected least-squares estimators in frontier production models", *Boletín de Estadística e Investigación Operativa*, 25 (2), pp. 86-96.
- ORTEGA, F.J.; GAVILAN, J.M. (2013), "The measurement of production efficiency in scientific journals through stochastic frontier analysis models: Application to quantitative economics journals", *Journal of Informetrics*, 7, pp. 959-965.
- ORTEGA, F.J.; GAVILAN, J.M. (2014), "A comparison between maximum likelihood and Bayesian estimation of stochastic frontier production models", *Communications in Statistics-Simulation and Computation*, 43 (7), pp. 1714-1725.
- OSIEWALSKI, J.; STEEL, M.F.J. (1998), "Numerical tools for the Bayesian analysis of stochastic frontier models", *Journal of Productivity Analysis*, 10, pp. 133-117.
- STURZ, S.; LIGGES, U.; GELMAN, A. (2005), "R2WinBUGS: A package for running WinBUGS from R", *Journal of Statistical Software*, 12 (3), pp. 1-16.
- VAN DEN BROEK, J.; KOOP, G., OSIEWALSKI, J.; STEEL, M.F.J., (1994), "Stochastic frontier models. A Bayesian perspective", *Journal of Econometrics*, 61, pp. 273-303.
- ZHANG, X. (1999), "A Monte Carlo study on the finite sample properties of the Gibbs sampling method for a stochastic frontier model", *Journal of Productivity Analysis*, 14, pp. 71-83.

## ANEXO

En este anexo se ofrecen los valores de ECM del parámetro  $\gamma$  y de las eficiencias individuales en todas las simulaciones efectuadas. Con rellenos de fondo se ha resaltado el menor ECM obtenido para cada valor del espacio paramétrico y tamaño muestral.

| n=30  | $\gamma=0.1$    |        | $\gamma=0.3$    |        | $\gamma=0.5$    |        | $\gamma=0.7$    |        | $\gamma=0.9$    |        |        |
|-------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|--------|
|       | ECM( $\gamma$ ) | ECM(E) |        |
| MV    | 0.2034          | 0.2266 | 0.1723          | 0.1948 | 0.1699          | 0.1591 | 0.1979          | 0.1334 | 0.1072          | 0.0577 |        |
| BAYES | r=0.1           | 0.2676 | 0.1436          | 0.2553 | 0.1171          | 0.1471 | 0.0929          | 0.0541 | 0.0654          | 0.0053 | 0.0307 |
|       | r=0.3           | 0.1080 | 0.0947          | 0.1170 | 0.0823          | 0.0681 | 0.0672          | 0.0251 | 0.0491          | 0.0033 | 0.0246 |
|       | r=0.5           | 0.0274 | 0.0664          | 0.0404 | 0.0645          | 0.0266 | 0.0559          | 0.0182 | 0.0448          | 0.0131 | 0.0248 |
|       | r=0.7           | 0.0053 | 0.0821          | 0.0237 | 0.0776          | 0.0547 | 0.0715          | 0.0875 | 0.0614          | 0.0841 | 0.0399 |
|       | r=0.9           | 0.0087 | 0.1841          | 0.0724 | 0.1800          | 0.1942 | 0.1750          | 0.3483 | 0.1632          | 0.4509 | 0.1321 |

| n=50  | $\gamma=0.1$    |        | $\gamma=0.3$    |        | $\gamma=0.5$    |        | $\gamma=0.7$    |        | $\gamma=0.9$    |        |        |
|-------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|--------|
|       | ECM( $\gamma$ ) | ECM(E) |        |
| MV    | 0.1870          | 0.2090 | 0.1378          | 0.1821 | 0.1502          | 0.1544 | 0.1463          | 0.1043 | 0.0582          | 0.0406 |        |
| BAYES | r=0.1           | 0.2233 | 0.1336          | 0.1958 | 0.1022          | 0.1107 | 0.0787          | 0.0406 | 0.0555          | 0.0040 | 0.0261 |
|       | r=0.3           | 0.0747 | 0.0865          | 0.0731 | 0.0720          | 0.0422 | 0.0593          | 0.0152 | 0.0436          | 0.0034 | 0.0224 |
|       | r=0.5           | 0.0159 | 0.0645          | 0.0217 | 0.0605          | 0.0195 | 0.0536          | 0.0210 | 0.0440          | 0.0156 | 0.0248 |
|       | r=0.7           | 0.0037 | 0.0827          | 0.0235 | 0.0780          | 0.0608 | 0.0736          | 0.0967 | 0.0621          | 0.0722 | 0.0366 |
|       | r=0.9           | 0.0084 | 0.1836          | 0.0734 | 0.1806          | 0.1947 | 0.1736          | 0.3529 | 0.1639          | 0.3798 | 0.1140 |

| n=100 | $\gamma=0.1$    |        | $\gamma=0.3$    |        | $\gamma=0.5$    |        | $\gamma=0.7$    |        | $\gamma=0.9$    |        |        |
|-------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|--------|
|       | ECM( $\gamma$ ) | ECM(E) |        |
| MV    | 0.1309          | 0.1992 | 0.0978          | 0.1705 | 0.1151          | 0.1311 | 0.0953          | 0.0780 | 0.0162          | 0.0247 |        |
| BAYES | r=0.1           | 0.1634 | 0.1214          | 0.1398 | 0.0884          | 0.0753 | 0.0671          | 0.0247 | 0.0465          | 0.0024 | 0.0220 |
|       | r=0.3           | 0.0469 | 0.0804          | 0.0511 | 0.0673          | 0.0233 | 0.0540          | 0.0096 | 0.0407          | 0.0033 | 0.0210 |
|       | r=0.5           | 0.0089 | 0.0635          | 0.0122 | 0.0595          | 0.0177 | 0.0536          | 0.0224 | 0.0429          | 0.0103 | 0.0225 |
|       | r=0.7           | 0.0030 | 0.0878          | 0.0261 | 0.0831          | 0.0680 | 0.0765          | 0.1006 | 0.0629          | 0.0430 | 0.0297 |
|       | r=0.9           | 0.0085 | 0.1887          | 0.0761 | 0.1870          | 0.2022 | 0.1806          | 0.3463 | 0.1618          | 0.2470 | 0.0814 |

| n=500 | $\gamma=0.1$    |        | $\gamma=0.3$    |        | $\gamma=0.5$    |        | $\gamma=0.7$    |        | $\gamma=0.9$    |        |        |
|-------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|-----------------|--------|--------|
|       | ECM( $\gamma$ ) | ECM(E) |        |
| MV    | 0.0593          | 0.1744 | 0.0546          | 0.1344 | 0.0507          | 0.0808 | 0.0137          | 0.0411 | 0.0009          | 0.0190 |        |
| BAYES | r=0.1           | 0.1001 | 0.1083          | 0.0683 | 0.0725          | 0.0269 | 0.0542          | 0.0065 | 0.0391          | 0.0006 | 0.0190 |
|       | r=0.3           | 0.0302 | 0.0782          | 0.0226 | 0.0614          | 0.0089 | 0.0501          | 0.0050 | 0.0384          | 0.0009 | 0.0190 |
|       | r=0.5           | 0.0038 | 0.0634          | 0.0073 | 0.0588          | 0.0137 | 0.0520          | 0.0122 | 0.0398          | 0.0012 | 0.0191 |
|       | r=0.7           | 0.0030 | 0.0852          | 0.0247 | 0.0793          | 0.0559 | 0.0696          | 0.0382 | 0.0462          | 0.0016 | 0.0192 |
|       | r=0.9           | 0.0074 | 0.1568          | 0.0628 | 0.1500          | 0.1508 | 0.1350          | 0.1330 | 0.0794          | 0.0050 | 0.0201 |