

CARACTERIZACIÓN DE EMPRESAS EFICIENTES EN EL SECTOR LECHERO ASTURIANO

Antonio Álvarez Pinilla
Eduardo González Fidalgo*
Universidad de Oviedo

En este artículo se estudia la eficiencia técnica de un grupo de explotaciones lecheras en Asturias. El objetivo es describir las características de las explotaciones más eficientes de cara a proporcionar información sobre los cambios que deben realizarse para aprovechar al máximo las posibilidades técnicas de producción. En la parte empírica se estima una función de producción utilizando un panel de datos de 82 explotaciones durante el período 1987-1991. La caracterización de las empresas eficientes se realiza en una segunda etapa utilizando dos procedimientos alternativos: análisis de regresión y análisis de la varianza. Los resultados muestran importantes diferencias entre ambos procedimientos.

Palabras clave: eficiencia, segunda etapa, explotaciones lecheras.

1. INTRODUCCIÓN

La medición del grado de eficiencia en las distintas actividades productivas constituye uno de los campos de investigación económica que mayor interés ha despertado en las últimas décadas. Desde el trabajo pionero de Farrell (1957), numerosos estudios han desarrollado técnicas de estimación cada vez más sofisticadas.

El objetivo último de los análisis de eficiencia es determinar los cambios que permitieran reducir o eliminar la ineficiencia. En efecto, la principal preocupación de una empresa no es conocer exactamente cuál es su grado de ineficiencia sino qué pasos debe seguir para ser eficiente.

Algunos estudios han tratado de dar respuesta a esta cuestión analizando las características comunes a las empresas eficientes. El procedimiento habi-

(*) Los autores agradecen las sugerencias de Carlos Arias, Rafael Álvarez, David Roibás y Luis Orea, así como los comentarios de un revisor anónimo.

tual consiste en realizar un análisis de regresión entre el índice de eficiencia y ciertas variables que se suponen relacionadas con él, como por ejemplo el tamaño de la empresa. Este procedimiento se conoce con el nombre de *análisis de segunda etapa*¹. Ejemplos de esta metodología son los trabajos de Pitt y Lee (1981) y, más recientemente, Hallam y Machado (1996).

El objetivo de este trabajo es realizar una caracterización de las explotaciones lecheras más eficientes de una muestra de 82 explotaciones lecheras asturianas observadas durante cinco años consecutivos. La caracterización se realiza mediante un análisis de varianza, cuyos resultados se comparan con los obtenidos a partir de un análisis de regresión. De este modo se pretende ilustrar las diferencias entre ambas aproximaciones.

2. ESTIMACIÓN DE ÍNDICES DE EFICIENCIA TÉCNICA CON DATOS DE PANEL

El concepto de ineficiencia técnica hace referencia a la incapacidad de algunas empresas para conseguir producir el máximo output permitido por la tecnología de producción existente, dada la dotación de inputs que utilizan (Koopmans, 1951). La frontera de producción determina la máxima cantidad de output que la empresa puede obtener, por lo que sólo las empresas situadas sobre la frontera son técnicamente eficientes. La medida habitual del grado de eficiencia técnica de una empresa viene dada por el cociente entre la cantidad de output que produce y la máxima cantidad de output que podría producir.

La estimación de la frontera requiere disponer de datos sobre la actuación de distintas empresas que utilicen una misma tecnología. De este modo, la evaluación de cada empresa se hace en relación con las mejores prácticas productivas observadas, que definen la frontera. Los índices obtenidos son, por tanto, índices de eficiencia relativa.

Las técnicas de estimación han ido evolucionando desde las primeras fronteras determinísticas hasta las fronteras estocásticas. La diferencia estriba en que las primeras atribuyen cualquier desviación de la frontera a ineficiencias, mientras que las segundas también consideran la posibilidad de que existan elementos aleatorios (no controlables por el productor) que afecten la actuación de las empresas. En palabras de Aigner, Lovell y Schmidt (1977): «... el agricultor cuya cosecha es devastada por la sequía es desafortunado con nuestra medida (estocástica), pero ineficiente con la medida habitual (determinística)».

Por otro lado, Schmidt y Sickles (1984) señalan que la estimación de fronteras estocásticas a partir de datos de corte transversal presenta tres importantes limitaciones. En primer lugar, los índices de eficiencia estimados son inconsistentes. En segundo lugar, la separación de la ineficiencia y la pertur-

(1) Bravo-Ureta y Pinheiro (1992) analizan las ventajas e inconvenientes de esta aproximación. Un artículo reciente de Battese y Coelli (1995) critica la realización de una segunda etapa debido a que los índices de eficiencia no están idénticamente distribuidos. En su lugar, proponen integrar ambas etapas en una sola.

bación aleatoria requiere formular supuestos sobre la distribución de la ineficiencia. Por último, es necesario asumir que la ineficiencia es independiente de los regresores. El desarrollo de técnicas de estimación con datos de panel ha permitido superar estos tres problemas.

La especificación logarítmica del modelo de frontera con datos de panel puede escribirse del siguiente modo:

$$(1) \quad \ln y_{it} = \alpha + \ln x_{it} \beta + \lambda_t + \varepsilon_{it} - \mu_i$$

donde y_{it} representa el output obtenido por la empresa i en el período t , α es una constante, x_{it} el vector de inputs utilizado por la empresa i y ε_{it} es un término de error que se supone simétricamente distribuido con media cero. Los parámetros λ_t son efectos temporales que controlan factores que varían de un año a otro, pero son comunes a todas las empresas, como por ejemplo un cambio general en las condiciones meteorológicas. El término μ_i es una variable aleatoria que sólo toma valores positivos y representa la ineficiencia de la unidad productiva i .

La frontera de producción representa el máximo output alcanzable dada la dotación de inputs, es decir:

$$(2) \quad \ln y_{it}^* = \alpha + \ln x_{it} \beta + \lambda_t + \varepsilon_{it}$$

El índice de eficiencia vendrá dado por el cociente entre el nivel de producción real y el potencial:

$$(3) \quad \ln ET_i = \ln \left(\frac{y_{it}}{y_{it}^*} \right) = \ln y_{it} - \ln y_{it}^* = -\mu_i \Rightarrow ET_i = e^{-\mu_i}$$

El modelo (1) puede escribirse también de la siguiente forma:

$$(4) \quad \ln y_{it} = \alpha_i + \ln x_{it} \beta + \lambda_t + \varepsilon_{it}, \quad \alpha_i = \alpha - \mu_i$$

donde los coeficientes α_i reciben el nombre de efectos individuales y recogen el efecto de la heterogeneidad inobservable, es decir, el efecto sobre la producción de características específicas de cada empresa invariantes en el tiempo que no se observan.

La elección de un método de estimación para el modelo (4) depende de si existe correlación entre los efectos individuales y las variables explicativas. En caso de existir correlación, se utiliza el estimador Intragrupos (WITHIN) que consiste en estimar el modelo tomando las variables en desviaciones sobre sus medias temporales. Este procedimiento es equivalente a estimar los parámetros α_i por Mínimos Cuadrados Ordinarios como los coeficientes de variables dummy, una para cada empresa. Por esta razón, se suele conocer a este método como modelo de efectos fijos. Este estimador es siempre consistente.

Si los efectos individuales no están correlacionados con las variables explicativas, el modelo puede estimarse por Mínimos Cuadrados Generalizados, aplicando el Estimador de Componentes de Varianza (Balestra y Nerlove, 1966). La elección entre uno u otro modelo puede hacerse basándose en el test de Hausman (1978) sobre la existencia de dicha correlación.

En el modelo de efectos fijos, los índices de eficiencia relativa pueden obtenerse a través de la comparación entre los efectos individuales de las dis-

tintas empresas de la muestra. Tener un efecto individual relativamente pequeño implica que la empresa posee una serie de características inobservables que le impiden producir en las mismas condiciones que otras empresas, cuyos efectos individuales son mayores. Teniendo en cuenta la expresión (4), el mayor efecto individual corresponde a una empresa eficiente situada sobre la frontera, para la que $u_i = 0$, de manera que $\max(\alpha_j) = \alpha_j$. Por tanto, el índice de eficiencia puede expresarse en función de los efectos individuales de la siguiente manera:

$$(5) \quad ET_i = e^{-u_i} = e^{(\alpha_i - \alpha_i^*)} = e^{(\alpha_i - \max \alpha_j)}$$

El índice de eficiencia técnica toma el valor 1 para la empresa con el mayor efecto individual. El resto obtienen índices inferiores a la unidad, que reflejan su grado de ineficiencia relativa.

3. DATOS

Los datos utilizados en el análisis empírico proceden de un programa de gestión de vacuno lechero de la Consejería de Agricultura y Pesca del Principado de Asturias. Para la estimación de los índices de eficiencia técnica se dispone de un panel completo de datos técnicos y contables de 82 explotaciones lecheras durante el período 1987-1991. Todas las explotaciones son de carácter familiar, pudiendo asumirse que comparten una misma tecnología de producción.

Las variables utilizadas para estimar la frontera de producción son:

Output (Y): Producción de leche (miles de litros).

Trabajo (L): Coste total de la mano de obra (cientos de miles de ptas.).

Tierra (T): Superficie total de la explotación (hectáreas).

Vacas (V): Número de vacas.

Pienso (P): Cantidad total de piensos (toneladas). Debido a que las explotaciones presentan tasas de recría diferentes, esta variable ha sido ajustada de manera que sólo incluye los concentrados suministrados a las vacas en ordeño.

Forraje (F): Gastos totales necesarios para producir forrajes (cientos de miles de ptas.). Incluye gastos en semillas, pesticidas, fertilizantes, etc. También incluye la depreciación de la maquinaria ya que ésta se utiliza fundamentalmente para producir forrajes.

En el cuadro 1 se presentan estadísticas descriptivas de las variables empleadas en el modelo². El elevado valor que toman los coeficientes de variación sugiere la existencia de un alto grado de heterogeneidad entre los procesos productivos de las explotaciones de la muestra.

(2) Las variables Trabajo y Forraje, al estar medidas en unidades monetarias, fueron deflactadas por un índice de precios pagados por los ganaderos.

Cuadro 1
ESTADÍSTICAS DESCRIPTIVAS DE LOS DATOS DE ACTIVIDAD

Variable	Media	Coef. variación	Min.	Max.
Output	110,1	0,52	27,6	386,3
Trabajo	12,1	0,25	6,3	24,7
Tierra	12,1	0,32	5,5	25,0
Vacas	21,0	0,39	8,1	56,0
Pienso	31,2	0,67	1,1	149,3
Forraje	8,9	0,59	0,95	36,5

4. MODELO EMPÍRICO

La frontera propuesta para estimar los índices de eficiencia técnica se especifica bajo la forma funcional Translog, con variables dummy de empresa para obtener los efectos individuales³:

$$(6) \quad \ln Y_{it} = \alpha_i + \sum_{j=1}^5 \beta_j \ln x_{jit} + \frac{1}{2} \sum_{j=1}^5 \sum_{k=1}^5 \beta_{jk} \ln x_{jit} \ln x_{kit} + \varepsilon_{it}$$

donde el subíndice i indica la explotación, los subíndices j, k denotan genéricamente los cinco inputs utilizados en la estimación y el subíndice t indica el año.

No se han incluido efectos temporales debido a problemas de multicolinealidad con el resto de las variables explicativas. En el caso estudiado, los efectos temporales deberían recoger los efectos de las variaciones en las condiciones climatológicas de los distintos períodos, como por ejemplo la pluviometría o la temperatura. Sin embargo, los ganaderos responden a las variaciones en las condiciones climatológicas ajustando las cantidades de pienso y forraje utilizadas, por lo que estas variables ya controlan el efecto de tales variaciones.

La frontera especificada se estimó por Mínimos Cuadrados Ordinarios mediante el estimador de efectos fijos (WITHIN). Con objeto de obtener estimadores eficientes se aplicó el estimador de la matriz de varianzas-covarianzas de White (1980). El cuadro 2 recoge los coeficientes de la estimación del modelo (6). El coeficiente de determinación es 0,98.

Los coeficientes de una función Translog en su forma exacta no tienen interpretación económica directa. Para facilitar la interpretación de los resultados, el cuadro 3 recoge la estimación de las elasticidades de los inputs en la media geométrica de la muestra.

Como se esperaba, todas las elasticidades de los inputs son positivas y significativamente distintas de cero. Los coeficientes sugieren la especial importancia de los inputs Vacas y Pienso en la producción de leche. La eficiencia media estimada para la muestra es del 72%. Existe por tanto un alto grado de ineficiencia en las explotaciones analizadas.

(3) La especificación Cobb-Douglas fue rechazada al nivel de significación 0,01 al contrastarla frente a la Translog.

Cuadro 2
ESTIMACIÓN DE LA FRONTERA DE PRODUCCIÓN TRANSLOG

Parámetro	Coeficiente	t-ratio	Parámetro	Coeficiente	t-ratio
β_L	0,06	0,18	β_{LT}	0,01	0,10
β_T	-0,77	-1,35	β_{LV}	-0,19	-1,20
β_V	0,43	0,83	β_{LP}	0,12	2,00*
β_P	0,21	0,92	β_{LF}	0,13	2,30*
β_F	0,004	0,02	β_{TV}	0,65	2,51*
β_{LL}	-0,03	-0,19	β_{TP}	-0,17	-1,22
β_{TT}	-0,07	-0,22	β_{TF}	-0,16	-1,77
β_{VV}	-0,26	-0,80	β_{VP}	-0,06	-0,66
β_{PP}	0,16	4,23*	β_{VF}	0,05	0,56
β_{FF}	0,08	1,56	β_{PF}	-0,06	-1,75*

* Nivel de Significación 0,1.

Cuadro 3
ELASTICIDADES DE LOS INPUTS EN LA MEDIA GEOMÉTRICA

Input	Elasticidad	t-ratio
Trabajo	0,10	2,88*
Tierra	0,14	2,56*
Vacas	0,68	12,43*
Pienso	0,26	8,87*
Forraje	0,04	1,99*

* Nivel de Significación 0,1.

5. CARACTERIZACIÓN DE LAS EXPLOTACIONES EFICIENTES

En este apartado se analiza la relación existente entre algunas variables de gestión y los índices de eficiencia estimados en el apartado 4. Se utilizarán dos métodos alternativos: análisis de regresión y análisis de varianza. Posteriormente se realiza una reflexión sobre la comparación de los resultados obtenidos mediante ambos métodos.

Para realizar la caracterización de las explotaciones más eficientes, se dispone de una encuesta auxiliar efectuada únicamente en el año 1991 que recoge datos sobre ciertas características de las explotaciones de la muestra. Las variables incluidas en esta encuesta son:

Edad: Edad del cabeza de familia (años).

Pradera artificial: Proporción de pradera artificial y cultivos forrajeros sobre el total de la superficie de la explotación. El resto del terreno es pradera natural, menos productiva.

- Genética:** Esta variable mide el coste de la dosis de semen más cara empleada por la explotación (ptas.). Es una variable proxy del nivel genético del rebaño.
- Ensilado:** Cantidad total de ensilado por vaca (metros cúbicos). Se considera que este tipo de alimento aumenta la producción de leche.
- ATP:** Esta variable refleja el nivel de dedicación de los miembros de la familia a las tareas agrarias. Es una variable dummy que toma el valor 1 si los miembros de la familia obtienen rentas procedentes de otras actividades, incluidas pensiones.
- Zona:** Es una variable binaria que toma el valor 1 para las explotaciones costeras y 0 para las explotaciones del interior.

El cuadro 4 recoge una descripción estadística de las variables anteriores. Se observa una gran variación en cuanto a las características distintivas de las explotaciones.

Cuadro 4
ESTADÍSTICAS DESCRIPTIVAS DE LAS CARACTERÍSTICAS
DE LAS EXPLOTACIONES

Variable	Media	Coef. variación	Min.	Max.
Edad	46,2	0,19	25,0	63,0
Pradera artificial	0,42	0,51	0,03	1,0
Genética	6.163	0,72	1.000	15.000
Ensilado	2,57	0,97	0,02	11,4
ATP	0,46	1,08	0	1
Zona	0,41	1,19	0	1

El modelo definitivo a estimar incluye como variables independientes Edad, Genética y Ensilado, tomadas en logaritmos. Asimismo, con el fin de aproximar el tamaño de la explotación, se incluye la variable Vacas (también en logaritmos); por último, se incluyen las variables ZONA, ATP y Pradera artificial. La variable dependiente (Eficiencia) también se expresa en logaritmos.

En primer lugar se realiza un análisis de regresión del índice de eficiencia sobre las variables citadas en el párrafo anterior. El cuadro 5 muestra los resultados de la estimación por Mínimos Cuadrados Ordinarios⁴. El coeficiente de determinación es de 0,2.

El análisis de regresión muestra la existencia de relaciones significativas entre algunas de las variables independientes y el índice de eficiencia técnica.

(4) Al igual que en la estimación de la función de producción, se aplicó el estimador de la matriz de varianzas-covarianzas de White (1980).

Cuadro 5
ESTIMACIÓN DE LOS DETERMINANTES DE LOS ÍNDICES DE EFICIENCIA

Variable	Coeficiente	t-ratio
Constante	-1,14	-2,95*
Log Edad	0,13	1,84*
Log Genética	0,07	2,29*
Log Ensilado	0,02	1,90*
Log Vacas	-0,09	-1,80*
Pradera artificial	0,03	0,36
Zona	0,005	0,14
ATP	0,03	1,00

* Nivel de Significación 0,1.

ca. Antes de analizar los resultados de la regresión, se presentan los resultados del análisis de varianza sobre el mismo conjunto de variables explicativas, a fin de poder comparar los resultados obtenidos con ambos procedimientos. Para realizar el análisis de varianza se dividió la muestra según el índice de eficiencia en tres grupos con igual número de observaciones cada uno, denominados respectivamente baja, media y alta. Seguidamente se contrastó la hipótesis de igualdad de medias para cada una de las variables entre los tres grupos de eficiencia. El cuadro 6 muestra los resultados del análisis.

Cuadro 6
ANÁLISIS DE VARIANZA SOBRE LOS POSIBLES DETERMINANTES DE LA EFICIENCIA

	Eficiencia técnica			Test F
	Baja	Media	Alta	
Edad	44,2	50,1	44,5	4,06*
Genética	4.752	5.880	7.796	3,58*
Ensilado	2,19	1,74	3,75	5,40*
Vacas	22,6	22,5	21,8	0,05
Pradera artificial	0,41	0,43	0,42	0,06

Para contrastar la relación entre la eficiencia y las variables dummy se realizó un contraste t de igualdad de medias entre las explotaciones con valor 1 y las explotaciones con valor 0. En cuanto a la variable Zona, la eficiencia media de las explotaciones costeras es del 73,6% mientras que en el interior alcanza tan sólo el 71,3%. No obstante, el valor del estadístico t (1,01) no permite rechazar la hipótesis de igualdad de medias. En cuanto a la variable ATP, las explotaciones cuyos miembros obtienen rentas procedentes de fuentes distintas a la actividad agraria obtienen un índice de eficiencia medio del

72,5%, mientras que el resto presenta un índice medio del 72,0%. Nuevamente, la diferencia no es estadísticamente significativa como muestra el valor t (0,05).

El análisis comparativo de los resultados obtenidos según las dos aproximaciones alternativas pone de manifiesto las importantes diferencias en las conclusiones que se pueden extraer a partir de una misma base de datos.

En primer lugar, el análisis de regresión revela la existencia de una relación positiva entre la Edad y el grado de eficiencia en la explotación. Por tanto, sería de esperar que las explotaciones gestionadas por los individuos de mayor edad presentasen índices de eficiencia también mayores. Sin embargo, el resultado del análisis de varianza sobre la variable Edad no muestra la existencia de esta relación. De hecho, se observa que los ganaderos más jóvenes se encuentran en los intervalos opuestos de mayor y menor eficiencia, mientras que la población de edad más avanzada se encuentra en un nivel intermedio de eficiencia. Por tanto, el análisis de regresión muestra un resultado que es distinto al del análisis de varianza, que no impone la cláusula *ceteris paribus*.

Por otro lado, el análisis de regresión indica la existencia de una relación negativa y estadísticamente significativa entre la eficiencia y el tamaño de la explotación, aproximado por el número de vacas. Por contra, el análisis de varianza muestra que las diferencias en el tamaño medio de las explotaciones de los tres grupos de eficiencia no son estadísticamente significativas.

Para el resto de variables explicativas, los resultados obtenidos a partir de ambas aproximaciones son similares. Los dos análisis muestran la existencia de una relación positiva de la eficiencia con el nivel genético de la explotación, así como con la intensidad en el uso de ensilado. Finalmente, no se aprecian relaciones significativas entre el grado de eficiencia y las variables de Zona, ATP y Pradera artificial.

6. CONCLUSIONES

El análisis de la eficiencia técnica en la producción tiene como fin último ofrecer una orientación sobre la manera de mejorar la gestión productiva de la empresa. En este sentido, la caracterización de las empresas más eficientes cobra una importancia fundamental. Sin embargo, ha de tenerse especial cuidado a la hora de escoger la metodología más apropiada para realizar dicha caracterización.

Los resultados de este trabajo muestran como dos aproximaciones alternativas para el análisis de los factores determinantes de la eficiencia pueden conducir a resultados diferentes. En la muestra de explotaciones lecheras estudiada, el análisis de regresión muestra la existencia de relaciones significativas entre la eficiencia y la edad del cabeza de familia, así como con el tamaño de la explotación. En cambio, los resultados del análisis de varianza no apoyan la existencia de tales relaciones. La diferencia fundamental entre ambas aproximaciones es que la segunda no impone la cláusula *ceteris paribus*, con lo que puede considerarse más adecuada para caracterizar las explotaciones más eficientes.

REFERENCIAS BIBLIOGRÁFICAS

- Aigner, D. J.; Lovell, C. A. K. y Schmidt, P. J. (1977): «Formulation and Estimation of Stochastic Frontier Production Function Models», *Journal of Econometrics*, 6, págs. 21-37.
- Balestra, P. y Nerlove, M. (1966): «Pooling cross-section and time series data in the estimation of a dynamic model: the demand for natural gas», *Econometrica*, 34, págs. 387-399.
- Battese, G. E. y Coelli, T. J. (1995): «A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data», *Empirical Economics*, 20, págs. 325-332.
- Bravo-Ureta, B. E. y Pinheiro, A. E. (1993): «Efficiency Analysis of Developing Country Agriculture: A Review of the Frontier Function Literature», *Agricultural and Resource Economics Review*, 22, págs. 88-101.
- Farrell, M. J. (1957): «The measurement of production efficiency», *Journal of the Royal Statistical Society, Series A* (120), págs. 253-281.
- Hallam, D. y Machado, F. (1996): «Efficiency Analysis with Panel Data: A Study of Portuguese Dairy Farms», *European Review of Agricultural Economics*, 23, págs. 79-93.
- Hausman, J. A. (1978): «Specification test in econometrics», *Econometrica*, 46, págs. 69-85.
- Koopmans, T. C. (1951): «Analysis of production as an efficient combination of activities», en Koopmans, T. C.: *Activity Analysis of Production and Allocation*, Wiley, New York.
- Pitt, M. N. y Lee, L. F. (1981): «The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry», *Journal of Development Economics*, 9, págs. 43-64.
- Schmidt, P. y Sickles, R. C. (1984): «Production Frontiers and Panel Data», *Journal of Business and Economic Statistics*, 2, págs. 367-374.
- White, H. (1980): «A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity», *Econometrica*, 48, págs. 817-838.

ABSTRACT

This article performs a characterization of the most efficient dairy farms in Asturias. The objective is to provide a better understanding of the changes that must be made in order to reach the maximum production possibilities, that efficiency measures alone cannot give. A panel of milk production and input use for 82 dairy farms during the period 1986-1995 is employed to estimate the production function. In order to characterize the most efficient farms, a second step analysis is performed on the indexes of technical efficiency using two alternative procedures: regression analysis and analysis of variance. The results show important differences among both procedures.

Key words: efficiency, panel data, second step, dairy farms.