

ECONOMÍA DE LA DISTRIBUCIÓN ELÉCTRICA: INTRODUCCIÓN A LA TEORÍA DE LA MEDICIÓN DE LA EFICIENCIA.

Darío Ezequiel Díaz

Doctorando en Ciencias Económicas con Mención en Economía.
Facultad de Ciencias Económicas. Universidad Nacional de Córdoba.
Becario de la Agencia Nacional de Promoción Científica y Tecnológica.
dariodiaz10@gmail.com,
dariodiaz222@gmail.com

Resumen

En los últimos años la distribución de energía eléctrica en el mundo ha sido objeto de una amplia variedad de reformas regulatorias con el fin de alcanzar mayores niveles de eficiencia económica. Medirla proporciona información acerca de la validez del argumento que afirma que este sector es un monopolio natural y contribuye, por otro lado, a la evaluación respecto a la definición del tamaño óptimo de las áreas de servicio.

La medición del desempeño económico de las empresas en términos de la eficiencia de costos o indicadores similares, tiene un rol esencial en asegurar la eficiencia productiva como así también, en la limitación del poder de mercado del monopolista.

Este paper tiene como objetivo principal, bajo el marco teórico de la economía de la energía y la microeconomía de la producción, hacer una revisión exhaustiva de la teoría de la medición de la eficiencia productiva de las distribuidoras eléctricas, y de la evaluación de los principales trabajos empíricos del área en la actualidad.

Palabras claves: regulación, distribución de energía eléctrica, medición de eficiencia.

Abstract.

During the last years, the distribution of electricity in the world has been subject to a wide variety of regulatory reforms in order to achieve higher levels of economic efficiency.

The efficiency measurement provides information about the validity of the argument that this sector is a natural monopoly and contributes, on the other hand, to the assessment against the definition of the optimal size of the service areas.

Measuring the economic performance of enterprises in terms of cost efficiency or similar indicators have a key role in ensuring productive efficiency as well as in limiting the market power of the monopolist.

Under the theoretical framework of energy economics and microeconomics of production, this paper has the main objective of a comprehensive review of the theory of measuring the productive efficiency of electricity distributors, and evaluation of major empirical work area today.

Key words: regulation, distribution of energy electric, measuring efficiency

1. Introducción

En los últimos años la distribución de energía eléctrica en el mundo ha sido objeto de una amplia variedad de reformas regulatorias con el fin de alcanzar mayores niveles de eficiencia económica. Conceptos como la eficiencia de escala, de alcance y de costos, se transformaron en insumos necesarios para la toma de decisiones en la gestión pública y privada. Los dos primeros tipos de eficiencia se relacionan directamente a las economías de escala y de alcance (vinculadas con la tecnología de la producción), y el tercer tipo, al desempeño y rendimiento económico de la empresa cuando la misma se enfrenta a las condiciones del mercado y de la tecnología.

El desempeño de una empresa que minimiza costos se puede descomponer en dos tipos de eficiencia: la técnica y la asignativa. La primera se mide como la distancia hasta la frontera óptima de producción y, consiste en el grado en que las empresas de distribución eléctrica podrían reducir los inputs para un nivel dado de outputs (orientación hacia los inputs) o incrementar los outputs, para un nivel dado de inputs (orientación hacia los outputs). La segunda se basa en la reducción de los costos a partir de la utilización de diferentes inputs en proporciones óptimas, operando en la senda de expansión del costo mínimo.

El análisis de la medición de la eficiencia en las empresas de distribución eléctrica es importante por dos motivos: la regulación y la “administración del negocio”. Medir la eficiencia proporciona información acerca de la validez del argumento que afirma que este sector es un monopolio natural y contribuye, por otro lado, a la evaluación respecto a la definición del tamaño óptimo de las áreas de servicio. También, proporciona información acerca de la importancia de las sinergias potenciales a través de la integración *horizontal*, que permite ahorrar en costos mediante la explotación de las economías de alcance. Finalmente, los indicadores de desempeño, tales como la

eficiencia de costos, son utilizados para monitorear el desempeño económico de las empresas a partir de la comparación con la(s) más eficiente(s).

Los reguladores usan cada vez más prácticas de *benchmarking*¹ en varios esquemas de regulación por incentivos². Por ejemplo, uno de los métodos más utilizados en las redes eléctricas es la regulación *price-cap*³ (RPI-X). Este método establece una tasa máxima para el incremento de los precios regulados igual a la diferencia entre la tasa de inflación (medida con el índice de precios minorista (RPI: *retail price index*)) y el crecimiento de la productividad compensado, conocido como “X-factores”. Los nuevos regímenes regulatorios, en Europa particularmente, igualan a los “X-factores” con las metas anuales de eficiencia de costos para cada empresa individualmente. Por lo tanto, el regulador puede establecer precios máximos diferenciados a partir del desempeño de la eficiencia de la empresa estimada a partir de un análisis empírico.

No obstante, el creciente uso del análisis de la eficiencia en el sector de la electricidad ha causado serias preocupaciones entre los reguladores y las empresas con respecto a la fiabilidad de las estimaciones de la eficiencia. De hecho, la evidencia empírica sugiere que las estimaciones son sensibles al enfoque de la eficiencia medida adoptada⁴. Sin embargo, existen estrategias alternativas que pueden ser utilizados para mejorar la metodología de medición de la eficiencia que se discutirán en este trabajo.

Este *paper* tiene como objetivo principal, a partir de la economía de la energía y la microeconomía de la producción, hacer una revisión exhaustiva de la teoría de la medición de la eficiencia productiva de las distribuidoras eléctricas y de la evaluación de los principales trabajos empíricos del área en la actualidad.

2. Revisión de la Teoría Microeconómica de la Producción

La teoría microeconómica de la producción ha sido desarrollada ampliamente en la literatura (por ejemplo, Shephard (1953); Chambers (1988); Jehle and Reny (1998); Hackman (2008)).

¹La idea básica es que el regulador fija la tarifa de cada empresa en función de los costos de las otras, disminuyendo con esto, la relación entre los niveles de precios y los correspondientes niveles de costos de una empresa particular.

²Es la teoría regulatoria que pretende crear mecanismos de incentivos para optimizar los efectos de la regulación, dadas las restricciones de ineficiencias productivas y de información asimétrica. Ejemplos son las llamadas *price cap*, *revenue cap*, *benchmarking*, entre otros.

³La idea general de la regulación con precio máximo es fijar las tarifas por un determinado período, usualmente 4 o 5 años, sin realizar una revisión tarifaria intermedia.

⁴Ver Jamasb and Pollitt (2003), Estache et al. (2004) and Farsi and Filippini (2004) como ejemplos.

Generalmente, los estudios empíricos modelan la producción de las empresas en una industria que utilizan g inputs $\mathbf{x} = (x_1, x_2, \dots, x_g)$ para producir m outputs $\mathbf{y} = (y_1, y_2, \dots, y_m)$. Una manera razonable de representar la tecnología de la empresa de transformar inputs en outputs en el largo plazo es a partir de una función de transformación, $T(x_1, x_2, \dots, x_g, y_1, y_2, \dots, y_m) = 0$ en el caso multiproducto, o bien como una función de producción, $y = f(x_1, x_2, \dots, x_g)$, en el caso de un único output.

Si la empresa se enfrenta a un mercado de inputs competitivos y elige una cesta de inputs para minimizar costos en el largo plazo, entonces el proceso para la minimización del costo puede representarse como:

$$\text{Min}_x C = \sum_{j=1}^g w_j x_j, \text{ S.T. } f(\mathbf{x}) \geq \mathbf{y} \quad (1)$$

Donde C representa el costo total a largo plazo, w_j el precio del input x_j , y f es la función de producción que relaciona el vector de insumos \mathbf{x} con el vector de output \mathbf{y} . La solución a (1) es de la forma $C(\mathbf{y}, \mathbf{w})$, donde $\mathbf{y} = (y_1, y_2, \dots, y_m)$ y $\mathbf{w} = (w_1, w_2, \dots, w_g)$. Siempre que la función de transformación $T(x_1, \dots, x_g, y_1, \dots, y_m) = 0$ limite con un conjunto estrictamente convexo, McFadden (1978) ha mostrado que la función de costos, $C(\mathbf{y}, \mathbf{w})$ tiene las siguientes propiedades (condiciones de regularidad):

1. $C(\mathbf{y}, \mathbf{w}) > 0$ para $\mathbf{w} > 0$ and $\mathbf{y} > 0$ (no negatividad);
2. Si $\mathbf{w}' > \mathbf{w}$, entonces $C(\mathbf{y}, \mathbf{w}') \geq C(\mathbf{y}, \mathbf{w})$ (no decreciente en \mathbf{w});
3. $C(\mathbf{y}, \mathbf{w})$ es cóncava y continua en \mathbf{w} ;
4. $C(\mathbf{y}, \mathbf{w})$ es homogénea de grado uno en los precios de los inputs: $C(\mathbf{y}, t\mathbf{w}) = t C(\mathbf{y}, \mathbf{w})$ para $t > 0$;
5. Si $\mathbf{y} > \mathbf{y}'$, entonces $C(\mathbf{y}, \mathbf{w}) \geq C(\mathbf{y}', \mathbf{w})$ (no decreciente en \mathbf{y}); y
6. $C(0, \mathbf{w}) = 0$ (sin costos fijos)

Además, de acuerdo con el lema de Shephard, las funciones de demanda de insumos que minimizan el costo se pueden derivar de las funciones de costo. Si la función de costos es diferenciable en los precios de los inputs en el punto $(\mathbf{y}^*, \mathbf{w}^*)$, entonces se cumple la siguiente propiedad:

7. El lema de Shephard: $\mathbf{x}(\mathbf{y}, \mathbf{w}) = \nabla_{\mathbf{w}} C(\mathbf{y}^*, \mathbf{w}^*)$, donde \mathbf{x} denota el vector de minimización de costos de inputs requeridos para producir el vector \mathbf{y}^* de outputs, dados los precios de los insumos \mathbf{w}^* .

Existe una relación dual entre la función de costos y la función de transformación presentada arriba. Esto implica que, si se define una función de producción y se deriva su función de costos, podemos tomar aquella función de costos para generar una función de producción. Además, cualquier función de costos que satisfaga las propiedades 1-7, implica alguna tecnología para lo cual representa los costos mínimos.

Por lo tanto, de acuerdo al teorema de la dualidad de Shephard (1953), todas las características estructurales de las posibilidades de producción son contenidas en la especificación funcional de la función de costos que satisface las propiedades 1-7.

En las aplicaciones econométricas, trabajar con una función de costos tiene muchas desventajas respecto a su transformación dual o función de producción. Por ejemplo, las funciones de costos en comparación con las funciones de producción y transformación, se centran más en la relación económica que en la relación tecnológica. Sin embargo, en las funciones de producción existe una correlación muy alta entre las variables inputs, generando un problema de multicolinealidad, que imposibilita la estimación del modelo o conduce a una interpretación imprecisa de los coeficientes estimados. No obstante, se requieren dos supuestos para que la función de costos sea una representación dual válida de la tecnología de la producción. Primero, las empresas deben enfrentar de manera exógena los precios de los inputs y outputs. Segundo, las empresas deben ajustar los niveles de insumos para minimizar costos.

3. Eficiencia de escala, de alcance y de costos

Una gran parte de la literatura de estimación de costos se focaliza en la estimación de los rendimientos a escala. Chambers (1988) los define como el cambio proporcional en el output cuando una cesta de inputs es modificada por un escalar. Si un incremento proporcional en todos los inputs produce un aumento mayor (menor) del output, entonces la función de producción tiene rendimientos crecientes a escala (rendimientos decrecientes a escala). Si un incremento proporcional en todos los inputs resulta en un incremento en la misma proporción del output, entonces la función de producción exhibe retornos de escala constantes. Ya sea una tecnología de la producción que muestre cualquiera de los tres tipos de rendimientos de escala (constante, decreciente o creciente), ésta tiene implicaciones para la determinación de la estructura más eficiente de la industria. En particular, si una industria estuviese caracterizada por una tecnología de la producción para un único output con rendimientos crecientes de escala para un rango de output dado, la estructura eficiente de la industria implicaría un monopolio natural dentro de ese rango.

Los rendimientos de escala se definen usualmente en términos del incremento relativo del output a partir de un incremento proporcional de todos los inputs. En general, los rendimientos de escala necesitan definirse a lo largo de una recta específica de inputs combinados que se elige arbitrariamente. Considerando este problema, Hanoch (1975)

observó que es más relevante medir rendimientos a escala mediante la relación entre el costo total y el output a lo largo del sendero de expansión.

Para trasladar la definición de rendimientos constantes de escala en términos de la función de costos, Chambers (1988) propuso el concepto de “costo de flexibilidad” que puede medirse con el ratio del costo marginal y el costo medio, o bien como la elasticidad del costo del tamaño. El grado del costo de flexibilidad e_{cy} puede medirse utilizando la siguiente expresión:

$$e_{cy} = \sum_{k=1}^m \frac{\partial \ln C}{\partial \ln y_k} = \sum_{k=1}^m \left(\frac{\frac{\partial C}{\partial y_k}}{C/y_k} \right) \quad (2)$$

Con m diferentes outputs y_k .

El recíproco de la elasticidad del costo del tamaño se define como las economías de escala (ES), es decir:

$$ES = \frac{1}{\left(\sum_{k=1}^m \frac{\partial \ln C}{\partial \ln y_k} \right)} \quad (3)$$

Hay economías y deseconomías de escala si ES es mayor o menor que 1 respectivamente. Chambers expone que la definición de economías de escala es equivalente a la definición derivada de la tecnología de la producción si y solo si, esta última, es homotética. Generalmente, los estudios empíricos de la estimación de las funciones de costos de las empresas de distribución eléctrica no imponen este tipo de supuesto y, por lo tanto, utilizan el concepto de economías de escala.

En la teoría neoclásica de la producción, todas las unidades de producción son homogéneas. Por lo que, la producción de cualquier unidad de output no se diferencia del próximo, y puede intercambiarse sin ningún impacto en los costos. En la distribución de energía, sin embargo, una industria se caracteriza no solo por el output total producido, sino también por la estructura de la red servida. Es decir, los costos en las industrias de redes son influenciados no solo por el output total producido, sino también, por la estructura de red y el tamaño del área.

Por estas razones, la teoría económica aplicada ha sugerido incluir varias características del output en la especificación del modelo de costos para las industrias de redes. Estas características deberían capturar la heterogeneidad de los outputs. Por ejemplo, en la siguiente función de costos, se incluyen algunas variables características del output,

$$C = f(\mathbf{y}, \mathbf{w}, \mathbf{q}_{h+k}), \quad (4)$$

Donde C es el costo total, \mathbf{y} es un vector de outputs, \mathbf{w} es un vector de precios de los inputs, y \mathbf{q}_{h+k} es un vector de variables que reflejan las características del output (como

por ejemplo el factor de carga, el número de clientes y el tamaño del área de servicio de la empresa eléctrica).

En una función de costos la inclusión del número de clientes y/o del tamaño del área de servicio permite la distinción de economías de densidad del output, economías de la densidad del cliente y economías de escala.

Siguiendo a Roberts (1986), para empresas de distribución de gas y electricidad es posible definir economías de densidad de output (*EOD*) como el incremento proporcional en los costos totales provocados por un incremento proporcional del output (*y*), manteniendo todos los precios de inputs y las variables características del output tales como, el número de clientes (*CU*) y el tamaño del territorio del servicio (*AS*) constantes. Esto es equivalente a la inversa del costo total de la elasticidad con respecto al output:

$$EOD = 1 / \left(\frac{\partial \ln C}{\partial \ln y} \right) \quad (5)$$

Las economías (deseconomías) de la densidad del output existe si *EOD* es mayor (menor) que 1. Las economías de la densidad del output existen si los costos medios disminuyen a medida que aumenta el volumen de electricidad vendido a un número fijo de clientes en un área de servicio dado. Esta medida es relevante para decidir si la competencia o el monopolio local es la forma más eficiente en la industria de la distribución eléctrica.

Las economías de la densidad del cliente (*ECD*) se definen como el incremento proporcional en los costos totales provocados por un incremento proporcional del output y del número de clientes, manteniendo todos los precios de los inputs, y las otras características del output constantes. *ECD* puede definirse como:

$$ECD = 1 / \left(\frac{\partial \ln C}{\partial \ln y} + \frac{\partial \ln C}{\partial \ln CU} \right) \quad (6)$$

Las deseconomías (economías) de escala se presentan si *ECD* es menor (mayor) a 1. Esta medida es relevante para analizar el costo de distribuir más electricidad para un área de servicio fijo.

Finalmente, las economías de escala (*ES*) se definen como el incremento proporcional en los costos totales provocados por un incremento proporcional del output, el número de clientes y el tamaño del área de servicio, manteniendo todos los precios de los inputs constantes. *ES* puede definirse como:

$$ES = 1 / \left(\frac{\partial \ln C}{\partial \ln y} + \frac{\partial \ln C}{\partial \ln CU} + \frac{\partial \ln C}{\partial \ln AS} \right) \quad (7)$$

La presencia de las economías (deseconomías) de escala corresponde a valores ES mayores (menores) que 1. Esta medida es relevante para analizar el impacto en los costos de la fusión de dos empresas de distribución eléctricas adyacentes.

Una medida tradicional de la ineficiencia técnica (TIN) de la empresa que produce el output, y , con inputs, x , está dado por:

$$TIN = \frac{y}{y^*} \quad (8)$$

Donde y^* es la frontera del output asociado con el nivel de input x . Esta medida no incluye la ineficiencia asignativa, es decir, los ahorros potenciales por la reasignación de los inputs. La ineficiencia de costos es una medida orientada hacia los inputs de la ineficiencia total, que se define como la distancia desde la frontera de costos⁵. Una medida de la ineficiencia (CIN) de la empresa que produce el output y con el costo C está dado por:

$$CIN = \frac{C}{C^*} \quad (9)$$

Donde C^* es el costo de la frontera asociada con el nivel de output y .

Una ventaja importante del TIN es que no requiere datos de costos y precios. Estos datos son usualmente difíciles de obtener pero requeridos para estimar el CIN . Sin embargo, el TIN no proporciona información alguna del proceso de minimización de costos, que es más interesante desde el punto de vista económico.

La medida de ineficiencia de costos incluye las ineficiencias técnicas y asignativas pero no proporciona un modo fácil para separar estos dos componentes⁶. Otro rasgo clave del enfoque de la eficiencia de costos es que considera al output como dado. Esto es un supuesto realista en la mayoría de las industrias reguladas donde el nivel de output es establecido por el regulador o determinado por los factores de demanda.

4. Enfoques econométricos de la medida de eficiencia

Hay dos corrientes en la literatura empírica. La primera corriente consiste en estudios que estiman una función de costos utilizando, por ejemplo, los mínimos cuadrados ordinarios (OLS: *ordinary least squares*) o métodos econométricos más elaborados sin un componente estocástico para la ineficiencia, donde todas las empresas se asumen

⁵ La frontera de producción representa el output máximo producido por un conjunto dado de inputs, donde la "frontera de costos" define los costos mínimos de producir un nivel de output dados los precios de los inputs.

⁶La única manera de desenmarañar las ineficiencias técnicas y asignativas en el marco de la frontera de costos es a través de las ecuaciones de demanda de los inputs. Debido a la complejidad de la estructura del error resultante, una solución econométrica satisfactoria todavía no ha sido desarrollada (Greene, 1997; Kumbhakar and Lovell, 2000).

que operan en la frontera de costos (por lo que no se observa ineficiencia de costos). De la estimación de esta función de costos es posible calcular el nivel de ineficiencia de escala y de alcance para las empresas. La segunda corriente consiste en estudios que adoptan el enfoque de frontera asumiendo que la eficiencia del costo total se limita a aquellas empresas que se identifican como los mejores productores. Todas las otras compañías se asumen que operan por arriba de la frontera de costos, por lo que la ineficiencia no toma el valor cero. En este caso, la estimación econométrica de las mejores prácticas en la frontera de costos permite el cálculo del nivel de ineficiencia de los costos de las empresas además de su eficiencia de escala y de alcance. Nos referimos a los primeros enfoques como **métodos de no frontera**, en el cual la función de costos estimada está basada en el desempeño medio, como oposición al **análisis de frontera** basado en las mejores prácticas observadas.

La estimación de la función de costos bajo el enfoque de **no frontera** está basada en la estimación OLS de una función de costos paramétrica, usualmente expresada en logaritmos:

$$\ln C_i = f(\mathbf{y}_i, \mathbf{w}_i) + \varepsilon_i, \quad (10)$$

Donde C es el costo total incurrido por la empresa i , $f(\cdot)$ es la función de costos, \mathbf{y} es un vector de outputs, \mathbf{w} es un vector de precios de los inputs, y ε_i es el término del error estocástico. La eficiencia de los parámetros estimados puede mejorarse a partir de la combinación de la función de costos con las ecuaciones de participación de los factores que implica el lema de Shephard. De acuerdo a la teoría de la minimización de costos, los parámetros verdaderos de la función de costos son idénticos a los parámetros verdaderos en las ecuaciones de participación. Se obtienen grados adicionales de libertad sin la necesidad de estimar algún parámetro adicional. Las ecuaciones de participación de los inputs toman la siguiente forma:

$$S_{ji} = g(\mathbf{y}_i, \mathbf{w}_i) + \varepsilon_{ji} \quad (11)$$

Donde S_{ji} es la participación del costo incurrido por la empresa i para el input j , $g(\cdot)$ es la función de participación de los inputs, \mathbf{y} es un vector de outputs, \mathbf{w} es un vector de precios de inputs, y ε_{ji} es el término de error estocástico. El sistema de costos es generalmente estimado utilizando la técnica iterativa de Zellner (Zellner, 1962) para regresiones que aparentemente no están relacionadas (*SUR: Seemingly Unrelated Regressions*).

El enfoque del análisis de frontera está desarrollado con una amplia variedad de métodos paramétricos. Generalmente, la estimación de la frontera de producción o de costos puede ser realizada utilizando enfoques no paramétricos. Estos enfoques, tales

como **DEA** (*data envelopment analysis*) y “*free disposal hull*”, utilizan programación lineal para determinar la frontera de eficiencia de una empresa. En estos enfoques, la frontera de costos se considera como una función determinista de las variables observadas. Estos métodos son no paramétricos en cuanto a que no impone algún supuesto sobre la forma funcional o de distribución. Las ventajas y desventajas han sido extensivamente discutidas en la literatura.

En este paper, la atención se centra en los métodos econométricos con una especificación paramétrica. Una clasificación general de los enfoques de medición basados en métodos paramétricos sería la siguiente. En primer lugar, el análisis determinista y el análisis estocástico. Y este último, a partir de sus extensiones para datos de panel, se divide en: modelos con efectos fijos, mínimos cuadrados generalizados (GLS), estimaciones por máxima verosimilitud (MLE), y efectos aleatorios.

Al margen de la existencia de excepciones, todos los métodos paramétricos tienen un elemento estocástico en su función de frontera. Por lo que, este grupo de métodos se denominan **análisis de fronteras estocásticas** (SFA: *stochastic frontier analysis*). La excepción principal con una frontera determinista es el método COLS (*corrected OLS*). En este enfoque las ineficiencias son definidas a través de un movimiento constante de los residuos de OLS (ver Greene, 1980). Como todo el término estocástico se considera ineficiente, la frontera sigue siendo determinista.

El enfoque COLS está basado en la estimación OLS de una función de costos paramétrica generalmente expresada en logaritmos:

$$\ln C_i = f(\mathbf{y}_i, \mathbf{w}_i) + \varepsilon_i \quad (12)$$

Donde los parámetros se han definido más arriba. Después de corregir el término del error estocástico ε_i , desplazando la intersección de tal manera que todos los residuos sean positivos, el modelo COLS puede escribirse como:

$$\ln C_i = f(\mathbf{y}_i, \mathbf{w}_i) + \min(\varepsilon_i) + u_i, \quad \text{con } u_i = \varepsilon_i - \min(\varepsilon_i) \geq 0, \quad (13)$$

Donde u_i es un término no negativo que representa la ineficiencia de la firma. La eficiencia de costos de la firma i está dada por: $Eff_i = \exp(-u_i)$

El principal inconveniente de este método es que confunde la ineficiencia con el ruido estadístico: el residuo total es clasificado como ineficiencia.

En el modelo de frontera estocástica el término de error está dividido en dos partes incorrelacionadas: u_i es una perturbación no negativa asimétrica que refleja los efectos de ineficiencia, y v_i es una perturbación simétrica que captura el ruido aleatorio. Generalmente el ruido estadístico se asume normalmente distribuido, mientras que el

término de ineficiencia u_i se asume que sigue una distribución half-normal⁷. Un modelo básico SFA puede escribirse como:

$$\ln C_i = f(\mathbf{y}_i, \mathbf{w}_i) + u_i + v_i \quad (14)$$

Este modelo con un término de error compuesto **normal-half-normal** puede estimarse utilizando el método de máximo verosimilitud (MLE: *maximum likelihood estimates*). De manera similar, la eficiencia de costos de la empresa i está dada por: $Eff_i = \exp(u_i)$.

Los modelos SFA permiten una heterogeneidad aleatoria no observada entre empresas diferentes (representadas por v_i) pero requieren una forma funcional específica para la función de producción o de costos. La principal ventaja de tales métodos respecto a los enfoques deterministas es la separación del efecto de la ineficiencia del ruido estadístico debido a errores de datos y variables omitidas.

Muchos estudios han mostrado discrepancias en las estimaciones de eficiencia entre enfoques y especificaciones del modelo diferentes. Generalmente las discrepancias se deben particularmente a metodologías sensibles en la estimación de los resultados de eficiencia individual y no son limitados a una industria de red específica⁸. Las empresas reguladas operan en diferentes regiones con características ambientales y de redes variadas y que solo son observadas parcialmente. Puesto que los factores inobservados se consideran de manera diferente en cada método⁹, los estimadores resultantes pueden variar a través de los métodos. La magnitud de la variación depende de la importancia de los factores inobservados, que cambiarían de un caso a otro.

En la mayoría de los casos, no existe un criterio claro para la elección del modelo y del enfoque. Entonces, se asume que los resultados son válidos si son independientemente obtenidos de diversos modelos.

La literatura de la frontera estocástica ha proporcionado una variedad de modelos de datos de panel que pueden ser utilizados para superar algunos de los inconvenientes y proporcionar instrumentos más atractivos para usar en la regulación¹⁰. En particular,

⁷Otros desarrollos del modelo básico de frontera han considerado distribuciones exponenciales y normales truncadas para el término de la ineficiencia. Por ejemplo, Battese y Coelli (1992).

⁸Horrace y Schmidt (1996), Jensen (2000) y Street (2003) expusieron errores y problemas de consistencia en la estimación de resultados de eficiencia individual en los datos transversales. El primer paper muestra que tales problemas persisten en un conjunto de datos de panel con seis períodos.

⁹Mientras el enfoque econométrico utiliza términos estocásticos aditivos, el método de programación lineal permite la heterogeneidad en la producción a partir de la flexibilización de la restricción impuesta por una forma funcional específica.

¹⁰Algunas de esos modelos alternativos han sido utilizados en recientes estudios (Alvarez et al, 2004; Greene, 2005a; Farsi y Filippini, 2004; Farsi et al., 2006a)

esos modelos pueden controlar mejor la heterogeneidad no observada de las empresas o redes específicas.

El modelo de frontera de la ecuación (14) puede reescribirse para datos de panel usando subíndices i y t , respectivamente, para representar la empresa y el año de operación:

$$\ln C_{it} = f(y_{it}, w_{it}) + u_{it} + v_{it} \quad (15)$$

Generalmente, se asume que el término de heterogeneidad v_{it} está normalmente distribuido y que el término de ineficiencia u_{it} tiene una distribución half-normal que es, una distribución normal truncada en cero:

$$u_{it} \sim |N(0, \sigma_u^2)|, v_{it} \sim N(0, \sigma_v^2) \quad (16)$$

Este modelo está basado en el modelo original de la frontera de costos propuesta por Aigner et al (1977). La ineficiencia de la empresa es estimada utilizando la media condicional del término de ineficiencia como lo propuso Jondrow et al (1982): $E[u_{it} | \varepsilon_{it}]$, donde $\varepsilon_{it} = u_{it} + v_{it}$.

El primer uso de los modelos de datos de panel en los modelos de frontera estocástica fue de Pitt y Lee (1981), que asumieron que el término de ineficiencia u_{it} era constante a lo largo del tiempo, es decir $u_i \sim |N(0, \sigma_u^2)|$. El modelo de Pitt y Lee es diferente respecto al modelo convencional de efectos aleatorios donde los efectos individuales específicos se asumen que siguen una distribución half-normal. Schmidt y Sickles (1984) presentaron importantes variaciones de este modelo, quienes relajaron el supuesto de distribución y utilizaron el estimador de mínimos cuadrados generalizados (GLS: *generalized least squares*). También Battese y Coelli (1988) hicieron su aporte asumiendo una distribución normal truncada. En los casos donde los efectos de la empresa individual (u_i) están correlacionados con las variables explicativas, los parámetros estimados pueden ser sesgados. Schmidt y Sickles propusieron un enfoque de efectos fijos para evitar tales sesgos.

La restricción principal de esos modelos radica en que los factores no observados son aleatorios a lo largo del tiempo y entre las firmas, implicando que las redes y las características ambientales no observadas (que son generalmente invariantes temporalmente) no se consideren como heterogéneas. Además, con algunas pocas excepciones (como en el trabajo de Sickles (2005)), la variación de la eficiencia a lo largo del tiempo es determinista y sigue la misma forma funcional de todas las firmas. Dada de que las fuentes de la ineficiencia dependen de los shocks tecnológicos y de otras variaciones en el mercado de inputs como también de las habilidades

gerenciales, se puede argumentar que las ineficiencias son aleatorias a lo largo del tiempo¹¹.

Los modelos “verdaderos” de panel extienden el modelo de frontera estocástica a los datos de panel a partir del agregado de un efecto individual invariante en el tiempo. Estos modelos (Polachek y Yoon, 1996; y recientemente Greene, 2005a) incluyen dos términos estocásticos para la heterogeneidad no observada: uno para los factores que varían en el tiempo y otro para las características constantes específicas de la empresa. Asumiendo que los factores ambientales y de redes, y que sus efectos sobre la producción no varían considerablemente a lo largo del tiempo y, además, que la ineficiencia varía a lo largo del tiempo, esos modelos ayudan a separar esos efectos no observados de los estimadores de eficiencia.

Algunos de estos modelos han sido exitosamente utilizados en las redes de distribución eléctrica (Farsi et al., 2006a). Los modelos pueden escribirse sumando un término estocástico de la empresa específica (α_i) del lado derecho de la ecuación (15):

$$\ln C_{it} = f(\mathbf{y}_{it}, \mathbf{w}_{it}) + \alpha_i + u_{it} + v_{it} \quad (17)$$

El término α_i es un componente aleatorio independiente e idénticamente distribuido (iid) en un marco de efectos aleatorios (RE: *random-effects*), o un parámetro constante en un enfoque de efectos fijos (FE: *fixed-effects*). El término de ineficiencia se asume como una variable aleatoria iid con distribución half-normal. Esto implica que la ineficiencia no es persistente y que cada período trae consigo nuevos elementos idiosincráticos, es decir, nuevas fuentes de ineficiencia. Tales modelos tienen una ventaja importante ya que permiten que la ineficiencia varíe en el tiempo mientras se “*controla*” la heterogeneidad no observada a nivel empresa a través de efectos fijos o aleatorios. Un problema surge cuando los efectos específicos de la firma están correlacionados con las variables explicativas. En tales casos, los estimadores RE son afectados por el sesgo de heterogeneidad. Además, hay un importante problema práctico con el modelo FE: requiere la estimación de un gran número de parámetros. Cuando los datos de panel están disponibles, los reguladores tienen la posibilidad de utilizar métodos paramétricos con datos de panel para la predicción en intervalos de los costos de la empresa. Los intervalos predichos pueden utilizarse en la aplicación de una competencia por comparación (*yardstick competition*) o en la evaluación de los costos informados por las empresas, bajo una regulación *price-cap*.

¹¹ Alvarez et al. (2004) muestra que incluso en casos donde la ineficiencia se debe a factores invariantes temporales tales como la capacidad de los gerentes, las ineficiencias resultantes pueden variar a lo largo del tiempo porque depende de una serie de factores variantes en el tiempo que tienen un efecto de interacción con las habilidades del manager.

La elección de la forma funcional $f(\cdot)$ es otro aspecto importante de la estimación paramétrica de la frontera de costos. Una variedad de formas funcionales han sido empleadas para estimar las funciones de costos en la literatura¹². Lo que resta decir es que con respecto a las empresas eléctricas, el estudio de Nerlove (1963) fue el primero en analizar la estructura de costos de las empresas eléctricas de Estados Unidos empleando la especificación tradicional de Cobb-Douglas. El paper de Christensen et al. (1971) utilizó la forma funcional translog generando un impacto significativo en la literatura empírica del análisis de la producción aplicada.

5. Estudios empíricos de eficiencia de escala y de costos.

Filippini (1998) es el primer paper con una función de costo total sólo para las empresas de distribución.¹³ El costo total para distribuir electricidad de una empresa ha sido representado por la función de costos translogarítmica:

$$TC = C(y, P_c, P_p, P_l, LF, CU, ST, T),$$

Donde TC representa el costo total, y es el output representado por la cantidad total de kWh (kilovatios hora) entregados, y P_c , P_p y P_l son los precios del capital, del input kWh y el trabajo, respectivamente. LF es el factor de carga, ST el tamaño del territorio del servicio de la empresa de distribución medida en kilómetros cuadrados y CU el número de clientes. Estas variables últimas son introducidas en el modelo como características del output. El factor de carga ha sido incluido con el objetivo de capturar el impacto en los costos de la intensidad en el uso de la planta. T es una variable temporal que captura el cambio en la tecnología, es decir, representa el cambio en la eficiencia tecnológica. El modelo de costos ha sido estimado para muestras de datos de corte transversal de empresas públicas de distribución eléctrica que operan en las ciudades de Suiza. El conjunto de datos se compone de una muestra de 39 ciudades de empresas de distribución eléctrica observadas en 4 años, de 1988 a 1991. Los resultados indican la existencia de economías de producción y de densidad de cliente como así también economías de escala para la mayoría de niveles de output.

¹²Griffin et al. (1987) y Chambers (1988) proporcionan una interesante crítica de varias formas funcionales en el análisis de la producción aplicada. Este tema no será tratado por la extensión del mismo.

¹³Hubieron estudios anteriores, como ser el de Roberts (1986) quien fue el primero en estimar una función de costos para la producción y distribución eléctrica conjunta utilizando la teoría de la dualidad y una forma funcional flexible.

La eficiencia de costos de una muestra de empresas distribuidoras eléctricas suizas es estimada por Farsi et al. (2006a) utilizando tres modelos de frontera alternativa para estimar la función de costo total de la empresa:

$$C = C(Y, P_K, P_L, P_P, LF, CU, AS, HGRID, DOT, DW, T),$$

Donde C representa el costo total; Y es el output en kWh; P_K , P_L y P_P son los precios del capital, del trabajo y del insumo eléctrico, respectivamente; y T es una variable temporal que representa la tendencia línea del progreso tecnológico.

Se incluyen también las seis características del output (y de la red):

- 1) LF : "factor de carga" definida como el ratio de la carga media de la empresa y su carga pico, es decir, la relación existente de la potencia promedio y la potencia máxima. El factor de carga indica el grado con que se mantiene el valor de la potencia máxima.
- 2) AS : tamaño del área de servicio de la empresa distribuidora;
- 3) CU : cantidad de clientes
- 4) $HGRID$: indicador binario para distinguir las empresas que operan en una red de transmisión de alto voltaje, además de su red de distribución
- 5) DOT : variable dummy que representa a las empresas de servicios públicos, cuya participación en los ingresos auxiliares es más del 25 por ciento de los ingresos totales
- 6) DW : indicador para los casos en que más del 40 por ciento del área de servicio está cubierta por bosques con el fin de capturar los costos relativamente más altos por el mantenimiento y el riesgo de daños en las líneas eléctricas.

Se adoptó la forma funcional de Cobb-Douglas. Los autores excluyeron la forma translog para evitar el potencial riesgo de multicolinealidad entre los términos de segundo orden debido a la fuerte correlación entre output y el gran número de variables características. Por último cabe decir que en el paper se muestra que utilizando el modelo GLS (Schmidt y Sickles, 1984) y el MLE (Pitt y Lee, 1981), los mismos otorgan resultados similares. Sin embargo, el modelo TRE (*True random effects*) (Greene, 2005a), predice una tasa de eficiencia promedio más alta. El modelo TRE asume un término de ineficiencia invariante respecto al tiempo y un término estocástico aislado de la heterogeneidad no observada de la empresa específica sugiriendo que los otros modelos sobrestimarían la ineficiencia. Esta conclusión es válida en la medida que la ineficiencia no permanezca constante a lo largo del tiempo

y los efectos de red no observados continúen siendo constantes¹⁴. Desafortunadamente estos nuevos modelos pueden solo brindar una solución parcial para los problemas de sensibilidad.

6. Conclusión

Se ha presentado una visión general de los diferentes métodos de medición de eficiencia en el sector de la distribución eléctrica. Afirmando que la eficiencia productiva debería ser considerada en el diseño y organización del sector como así también en la regulación de incentivos de las empresas individuales, se ha expuesto el concepto de eficiencia de acuerdo a tres aspectos: escala, alcance y costos. Las eficiencias de escala y de alcance están principalmente relacionadas con la estructura de mercado, particularmente con el nivel de producción de la empresa y la combinación de los inputs; mientras que la eficiencia de costos se debe principalmente al desempeño económico particular de la empresa frente a las condiciones de mercado. Una reforma regulatoria efectiva consideraría todas las dimensiones de la ineficiencia productiva.

La estimación de ineficiencias relacionadas a la escala y al alcance de las empresas distribuidoras es relativamente sencilla. De hecho, las características tecnológicas tales como las economías de escala y de alcance pueden estimarse utilizando los coeficientes de una función de costos. Estas estimaciones ayudan a determinar el tamaño óptimo y de alcance, y el grado del monopolio natural de la distribución eléctrica.

Las reformas regulatorias deberían apuntar a una organización óptima de las redes de distribución que permita a las empresas mayores posibilidades de utilizar variadas sinergias para reducir costos.

Además, los estudios empíricos confirman las características de monopolio natural de las redes de distribución eléctrica favoreciendo al monopolio local. En tales situaciones, un sistema regulatorio de incentivos efectivo es importante en ausencia de competencia directa entre las compañías. Por lo que, la medición del desempeño económico de las empresas en términos de la eficiencia de costos o indicadores similares, tiene un rol esencial en asegurar la eficiencia productiva como así también en la limitación del poder de mercado del monopolista.

¹⁴Si por el contrario, se supone que las ineficiencias son persistentes y no cambian considerablemente con el tiempo, entonces los resultados obtenidos a partir de los modelos de datos de panel convencional proporcionan mejores estimaciones de las ineficiencias

La medición de la eficiencia del costo se convertirá en un tema polémico en la medida en que el desempeño de las empresas individuales sea de interés en la política pública regulatoria. Un sistema efectivo de regulación requiere modelos apropiados para la medición de la eficiencia de costos. Tales modelos necesitan tomar en cuenta la heterogeneidad no observada entre empresas, que puede ser proporcionada por ciertos modelos de datos de panel. Además, dado que los diferentes modelos postulan varios supuestos respecto a la variación de la eficiencia entre compañías y a lo largo del tiempo, es clave que dichos supuestos sean claramente especificados y que las estimaciones de la eficiencia sean interpretadas en consecuencia. En muchos casos, es importante ver la eficiencia desde diversos ángulos, requiriendo la aplicación de distintos modelos con supuestos diferentes.

Referencias Bibliográficas

- Aigner, D., C.A.K. Lovell and P. Schmidt (1977), Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, **6**, 21–37.
- Alvarez, A., C. Arias and W.H. Greene (2004), Accounting for unobservables in production models: management and inefficiency, Working Paper E2004/72, Centro de Estudios Andaluces, Seville, Spain.
- Battese, G.E. and T.J. Coelli (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data, *Journal of Econometrics*, **38**, 387–99.
- Battese, G.E. and T. Coelli (1992), Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India, *Journal of Productivity Analysis*, **3**, 153–69.
- Baumol, W.J., J.C. Panzar and R.D. Willig (1982). *Contestable Markets and the Theory of Industry Structure*, New York: Harcourt Brace Jovanovich.
- Chambers, R.G. (1988). *Applied Production Analysis: A Dual Approach*, Cambridge: Cambridge University Press.
- Christensen, L.R., D. Jorgenson and L.J. Lau (1971), Transcendental logarithmic production frontiers, *Review of Economics and Statistics*, **55**, 28–45.
- Coelli, T., A. Estache, S. Perelman and L. Trujillo (2003). *A Primer on Efficiency Measurement for Utilities and Transport Regulators*, World Bank Institute of Development Studies, Washington, DC: World Bank.
- Estache, A., M.A. Rossi and C.A. Ruzzier (2004). The case for international coordination of electricity regulation: evidence from the measurement of efficiency in South America, *Journal of Regulatory Economics*, **25** (3), 271–95.

- Farsi, M. and M. Filippini (2004), Regulation and measuring cost efficiency with panel data models: application to electricity distribution utilities, *Review of Industrial Organization*, **25** (1), 1–19.
- Farsi, M. and M. Filippini (2004). Regulation and measuring cost efficiency with panel data models: application to electricity distribution utilities, *Review of Industrial Organization*, **25** (1), 1–19.
- Farsi, M., M. Filippini and W.H. Greene (2006a), Application of panel data models in benchmarking analysis of the electricity distribution sector, *Annals of Public and Cooperative Economics*, **77** (3), 271–90.
- Filippini, M. (1996). Economies of scale and utilization in the Swiss electric power distribution industry, *Applied Economics*, **28**, 543–50.
- Filippini, M. (1998). Are municipal electricity distribution utilities natural monopolies?, *Annals of Public and Cooperative Economics*, **2**, 157–74.
- Greene, W.H. (1980). Maximum likelihood estimation of econometric frontier functions, *Journal of Econometrics*, **13**, 27–56.
- Greene, W.H. (1997). Frontier production function, in M.H. Pesaran and P. Schmidt (eds), *Handbook of Applied Econometrics, Vol. II: Microeconomics*, Oxford: Blackwell, ch. 3, pp. 81–166.
- Greene, W.H. (2005a). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model, *Journal of Econometrics*, **126** (2), 269–303.
- Griffin, R.C., J.M. Montgomery and M.E. Rister (1987), Selecting functional form in production function analysis, *Western Journal of Agricultural Economics*, December, 216–27.
- Hackman, S.T (2008). *Production Economics: Integrating the Microeconomic and Engineering Perspectives*, Springer-Verlag Berlin Heidelberg.
- Hanoch, G. (1975). The elasticity of scale and the shape of average costs. *American Economic Review*, **65**, 956–65.
- Horrace, W.C. and P. Schmidt (1996). Confidence statements for efficiency estimates from stochastic frontier models, *Journal of Productivity Analysis*, **7**, 257–82
- Jamasb, T. and M. Pollitt (2003). International benchmarking and regulation: an application to European electricity distribution utilities, *Energy Policy*, **31**, 1609–22.
- Jehle, G.A. and Ph.J. Reny (1998). *Advanced Microeconomic Theory*, Englewood Cliffs, NJ: Prentice-Hall International.

- Jensen, U. (2000). Is it efficient to analyse efficiency ranking. *Empirical Economics*, **25**, 189–208
- Jondrow, J., I. Materov, K. Lovell and P. Schmidt (1982), On the estimation of technical inefficiency in the stochastic frontier production function model, *Journal of Econometrics*, **19** (2/3), August, 233–8.
- Kumbhakar, S.C. and C.A.K. Lovell (2000). *Stochastic Frontier Analysis*. Cambridge: Cambridge University Press.
- McFadden, D. (1978), Cost, revenue, and profit functions, in Melvyn Fuss and Daniel L. McFadden (eds), *Production Economics: A Dual Approach to Theory and Applications Volume I: The Theory of Production*, Amsterdam: North-Holland, pp. 3–109.
- Murillo-Zamorano, L.R. (2004). Economic efficiency and frontier technique, *Journal of Economic Surveys*, **18** (1), 33–77.
- Nerlove, M. (1963), 'Return to scale in electricity supply', in C.F. Christ (ed.), *Measurement in Economic Studies in Honor of Yehuda Grunfeld*, Stanford, CA: Stanford University Press, pp. 167–98.
- Pitt, M. and L. Lee (1981). The measurement and sources of technical inefficiency in the Indonesian weaving industry, *Journal of Development Economics*, **9**, 43–64.
- Polachek, S. and B. Yoon (1996). Panel estimates of a two-tiered earnings frontier, *Journal of Applied Econometrics*, **11**, 169–78.
- Roberts, M.J. (1986). Economies of density and size in the transmission and distribution of electric power, *Land Economics*, **62**, 378–87.
- Schmidt, P. and R.E. Sickles (1984). Production frontiers and panel data. *Journal of Business and Economic Statistics*, **2**, 367–74.
- Shephard, R.W. (1953). *Cost and Production Functions*, Princeton, NJ: Princeton University Press.
- Sickles, R. (2005), Panel estimators and the identification of firm-specific efficiency levels in parametric, semiparametric and nonparametric settings, *Journal of Econometrics*, **126** (2), 305–34
- Street, A. (2003). How much confidence should we place in efficiency estimates? *Health Economics*, **12** (11), 895–907.
- Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and test for aggregation bias, *Journal of the American Statistical Association*, **58**, 348-68.