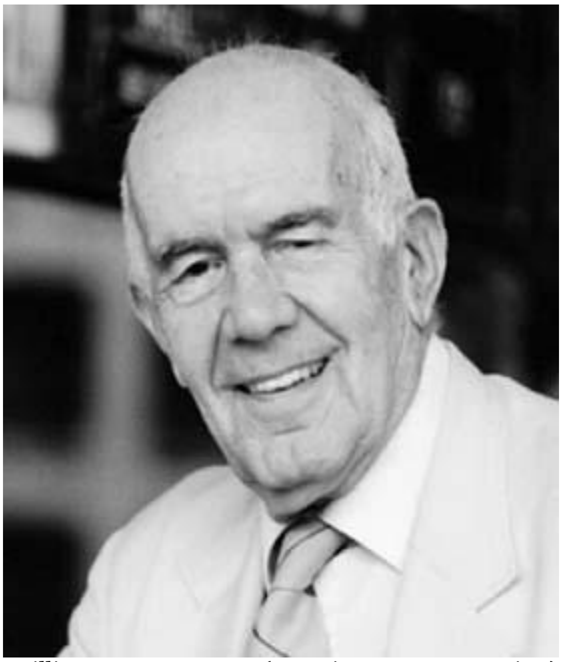
La medición de la eficiencia de las unidades de toma de decisiones

A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision making units, *European Journal of Operational Research* 2 (1978), 429-444

Un modelo de programación no lineal (no convexo) proporciona una nueva definición de eficiencia para su uso en la evaluación de actividades de entidades sin fines de lucro que participan en programas públicos. De esta manera se proporciona una medida escalar de la eficiencia de cada unidad participante, junto con métodos para determinar objetivamente los pesos con referencia a los datos de observación para los múltiples productos e insumos que caracterizan dichos programas. Se establecen equivalencias con modelos de programación lineal ordinarios para efectuar cálculos. Los duales de estos modelos de programación lineal proporcionan una nueva forma de estimar relaciones extremas a partir de datos observacionales. Son delineadas conexiones entre la ingeniería y los enfoques económicos de la eficiencia junto con nuevas interpretaciones y formas de utilizarlos en la evaluación y el control de gestión de los programas públicos.

1. Introducción

Este documento se ocupa de desarrollar medidas de "eficiencia en la toma de decisiones" con referencia especial a su posible uso en la evaluación de programas públicos. Tal como lo usaremos, el término "programa" se referirá a una colección de *unidades de toma de decisiones* (UTDs) con insumos y productos comunes. Estos productos e insumos serán generalmente de carácter múltiple y también pueden asumir una variedad de formas que admiten sólo mediciones ordinales. Por ejemplo, en un programa educativo como "*Follow Through*",[1] la eficiencia de varias escuelas, vistas como UTDs en este programa, se puede medir con referencia a los productos que involucran las categorías de educación estándar: habilidades cognitivas, afectivas y psico-motrices, mediante, respectivamente, (1) puntuaciones aritméticas, (2) pruebas psicológicas de las actitudes de los estudiantes, por ejemplo, hacia la comunidad y (3) capacidad del estudiante de entender y controlar los movimientos corporales, por ejemplo, observando su capacidad de tratar de no hundirse y de girar de adelante hacia atrás (y vice-



William Wager Cooper (Mr. Linear Programming)
1914-2012
Biografía en Mac Tutor
"Durante una larga y notablemente productiva carrera, el profesor William W. (Bill) Cooper hizo muchas contribuciones pioneras a la Investigación Operativa y a las Ciencias de la Gestión, con destacadas incursiones en las áreas de (a) programación lineal y no lineal, (b) programación de metas, (c) programación aleatoria, (d) análisis de envolvente de datos, y (e) planificación de la mano de obra, entre otras. Su legendaria colaboración con Abraham Charnes produjo resultados cuyas conexiones se remontan al siglo XVIII, con problemas planteados pero no resueltos por Laplace y Gauss." (F Glover and T Sueyoshi, Contributions of Professor William W Cooper in Operations Research and Management Science, 2009)

[1] Una discusión de este programa patrocinado por el gobierno federal que incluye un uso de las medidas de eficiencia que vamos a discutir se puede encontrar en [21]. Esto incluye el uso de varias pruebas estadísticas de significación (utilizando el llamado *estadístico de Kullback-Leibler*), que no se discutirá en el presente documento.

versa) en una piscina.[2] Todos ellos deben considerarse productos «valorados» incluso cuando no existe un mercado aparente para ellos o incluso cuando no se dispone fácilmente de otras fuentes posibles de sistemas de ponderaciones razonables. Los insumos pueden variar desde cantidades relativamente fáciles de medir y de ponderar, como "número de horas cátedra" y extenderse a las líneas más difíciles como "tiempo dedicado en las actividades del programa por líderes comunitarios y/o padres".

Nuestro uso de términos como "UTD" (unidad de toma de decisiones) y "programas" ayudará a enfatizar que nuestro interés está centrado en la toma de decisiones por parte de entidades sin fines de lucro en lugar de las "empresas" e "industrias" más habituales. También nos ayudará a enfatizar que nuestros datos (como en el ejemplo anterior) no son ponderados con precios de mercado[3] y/o con otros desiderata económicos - tales como los costos de producir la capacidad de generar ingresos de los estudiantes, con las correspondientes tasas de descuento - de acuerdo con la forma en que algunas actividades del sector público a veces se evalúan.

Naturalmente, nos interesa relacionar nuestras ideas con los desarrollos de la economía. Esto se hará por referencia a funciones de producción y conceptos relacionados como "dualidad de costos", etc. Aunque se necesitarán adaptaciones de estos conceptos, también trataremos de indicar lo que está involucrado en puntos adecuados en este artículo. (Ver más abajo sección 6, por ejemplo).

También queremos relacionar nuestras ideas con otras disciplinas, como la ingeniería, que también están relacionadas con la medición de la eficiencia. Esto se hará no sólo en aras de una mayor unidad, sino también con el interés de distinguir entre eficiencias asociadas con una "tecnología" de producción subyacente, y las debidas a la toma de decisiones gerenciales cuando la primera puede ser identificada y separada de la última como, por ejemplo, caracterizaciones ingenieriles.

Por supuesto, cuando esto no puede hacerse (el caso habitual de la economía empírica),[4] tendremos que contentarnos con el concepto algo menos satisfactorio de «eficiencia relativa». Esta última se determinará por referencia a las "clasificaciones" adecuadamente organizadas de los resultados observados de la toma de decisiones por varias UTDs en el mismo programa (por ejemplo, las diferentes escuelas en el programa *Follow Through*), al tiempo que se permite el hecho de que diferentes cantidades de insumos (en ciertos casos, legalmente estipulados) pueden estar involucradas, de manera que, por ejemplo, algunas UTDs son más similares a los miembros de un subconjunto y menos a los de otros subconjuntos, etc., en las "cantidades" de insumos y productos particulares utilizados.

El significado que se otorgará a estas caracterizaciones será aclarado en las secciones que siguen. Primero presentaremos nuestras medidas y modelos propuestos. Luego proveeremos caracterizaciones computacionales. Se delinearán las relaciones con líneas seleccionadas de

---

2 Véase la discusión en [31] para un uso de medidas como éstas en contextos de planificación-presupuesto-programa (PPBS).

3 Nos referimos aquí a los precios y los costos *reales* de mercado. Más adelante mostraremos cómo obtener estimaciones de coeficientes de producción (óptimos) y relacionarlos con costos y precios teóricos (de oportunidad).

4 Esta separación es aún más difícil en programas del sector público como la educación, la seguridad pública, etc., donde el significado de lo que es una "tecnología" es probablemente más ambiguo que en el caso de las industrias en el sector privado e incluso en muchas empresas de servicios.

investigación en curso, seguido de métodos de estimación e interpretación en términos de simples ilustraciones numéricas y caracterizaciones analíticas. Una sección final resumirá lo que se ha hecho y señalará las deficiencias pertinentes junto con otras posibles líneas de desarrollo.

2. Modelo y definición

Nuestra medida propuesta de eficiencia de cualquier UTD se obtiene como el máximo de una relación de productos ponderados a insumos ponderados sujeto a la condición de que las proporciones similares para cada UTD sean menores o iguales a la unidad. En forma más precisa,

$$(1) \qquad \max h_o = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}}$$

sujeto a

$$\frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq 1;\ j=1, 2\ldots n$$

$$u_r, v_i \geq 0,\ r=1, 2,\ldots, s;\ i=1, 2,\ldots, m.$$

En este caso, los y_{rj}, x_{ij} (todos positivos) son los productos e insumos conocidos de la UTD j-ésima, y los u_r, $v_i \geq 0$ son las variables de ponderaciones a ser determinadas por la solución del problema – por ejemplo, mediante los datos de la totalidad de las UTDs que están siendo utilizadas como conjunto de referencia. La eficiencia de un miembro de este conjunto de referencia de $j = 1,\ldots n$ UTDs debe ser clasificada en relación con las demás. Por lo tanto, está representada en el funcional, para la optimización -, así como en las restricciones - y se distingue, además, asignándole el subíndice '0' en el funcional (pero prefijando su subíndice original en las restricciones). La maximización indicada otorga entonces a esta UTD la ponderación más favorable que permiten las restricciones.

Para las UTD que nos interesa, estos valores x_{ij} y y_{rj}, que son constantes, usualmente serán observaciones de decisiones pasadas sobre insumos y los productos que resultaron de ellos. Sin embargo, podemos reemplazar algunas o todas estas observaciones por valores teóricamente determinados si lo deseamos (y podemos) realizar nuestras evaluaciones de eficiencia de esa manera.

Consideremos, por ejemplo, la siguiente definición (citada a partir de [14]) del campo de la ingeniería de la combustión: "la eficiencia es la relación entre la cantidad real de calor liberada en un dispositivo dado y la cantidad máxima que podría liberarse dado el combustible [utilizado]”. En símbolos,

$$E_r = y_r / y_R$$

donde

y_R: calor obtenido a partir de una unidad dada de combustible por el horno que se esté evaluando;

y_r: calor máximo que se puede obtener a partir de este mismo insumo de combustible.

Aunque la definición de eficiencia varía de un campo de ingeniería a otro, la de arriba captura lo esencial, es decir, la calificación es relativa a alguna posibilidad de máxima, de modo que, siempre, $0 \leq E_r \leq 1$.

También podemos obtener el E_r definido anteriormente de (1) como sigue. Para cualquier cantidad de insumo dado x la sustitución en (1) da

$$\max h_o = uy_o / vx_o$$

$$\text{s.a:} \quad uy_R / vx_R \leq 1,$$

$$uy_r / vx_r \leq 1$$

$$u, v \geq 0.$$

En el funcional, el subíndice 0 indica que está siendo calificado el último horno.

Ahora sea u^*, v^* un par de valores óptimos. Como $y_R \geq y_r$ y $x_R = x_r = x$, ello implica que $u^* y_R = v^* x_R$ y utilizando $x_o = x$ se tiene entonces que el funcional es igual a y_r / y_R como se dijo.

Al igual que en la mayoría de las definiciones de ingeniería, aquí hemos confinado nuestro desarrollo a razones de productos y/o insumos individuales, o sumas ponderadas. Esto último se puede determinar, de nuevo por consideraciones de ingeniería (por ejemplo, combinaciones de combustible eficientes), que normalmente no están disponibles para las aplicaciones económicas que estamos considerando. Siempre que tengamos las observaciones indicadas sobre insumos y productos para UTDs individuales, sin embargo, podemos por lo menos llegar a calificaciones de "eficiencia relativa" dentro de las líneas que hemos estado sugiriendo. Esta es la forma en que se desarrollará el resto del trabajo aunque, como ya se indicó, también podemos insertar datos de ingeniería u otros datos para tales clasificaciones, si así lo deseamos, en varias combinaciones.

Obsérvese que nuestras ponderaciones, como las anteriores, están objetivamente determinadas para obtener una medida escalar (sin dimensión) de la eficiencia en cualquier caso.[5] Es decir, la elección de pesos se determina directamente a partir de datos observados sujetos solamente a las restricciones establecidas en (1). Bajo estas observaciones y restricciones ningún otro conjunto de ponderaciones comunes otorgará una calificación más favorable con respecto al conjunto de referencia. Por lo tanto, si no se alcanza una clasificación (relativa) de eficiencia del 100% bajo este conjunto de ponderaciones, entonces tampoco se alcanzará desde ningún otro conjunto.

3. Reducción a un formato de programación lineal

El modelo anterior es una formulación de programación no lineal extendida de un problema de programación fraccional ordinario. En otra parte (en [10] y [7]) hemos suministrado una teoría completa en términos de que los problemas de programación fraccionarios pueden ser reemplazados por programas equivalentes de programación lineal. Por lo tanto, proponemos utilizar esta teoría para hacer que la formulación anterior sea computacionalmente maneja-

---

[5] Las propiedades de escala e invariancia que se tratan en [12] - véase también [21] - no se discutirán en este documento.

ble para grandes números *j (n)* de observaciones, así como pequeños números de insumos *i (m)* y productos *r (s)* que puedan ser de interés al menos en aplicaciones económicas.

Lo haremos de manera de proporcionar más claridad conceptual (y flexibilidad), además de facilitar nuestra toma de contacto con desarrollos relacionados en economía. Primero consideremos el siguiente modelo que es la versión de la medida recíproca (ineficiencia) de (1):

$$(2) \qquad \min f_o = \frac{\sum_{i=1}^{m} v_i x_{io}}{\sum_{r=1}^{s} u_r y_{ro}}$$

sujeto a:

$$\frac{\sum_{i=1}^{m} v_i x_{ij}}{\sum_{r=1}^{s} u_r y_{rj}} \geq 1;\ j = 1, 2, \ldots n$$

$$v_i,\ u_r \geq 0.$$

Ahora proponemos reemplazar esta formulación no lineal no convexa por un problema de programación lineal ordinaria. Por lo tanto, consideramos primero

$$(3) \qquad \max z_o$$

Sujeto a:

$$- \sum_{j=1}^{n} y_{rj} \lambda_j + y_{ro} z_o \leq 0;\ r = 1, \ldots s.$$

$$\sum_{j=1}^{n} x_{ij} \lambda_j \leq x_{io};\ i = 1, 2 \ldots m.$$

$$\lambda_j \geq 0;\ j=1, 2 \ldots n.$$

Como (3) es un problema de programación lineal ordinario, tiene un problema de programación lineal dual que podemos escribir como sigue:

$$(4) \qquad \min g_o = \sum_{i=1}^{m} \omega_i x_{io}$$

Sujeto a:

$$- \sum_{r=1}^{s} \mu_r y_{rj} + \sum_{i=1}^{m} \omega_i x_{ij} \geq 0,$$

$$\sum_{r=1}^{s} \mu_r y_{ro} = 1,$$

$$\mu_r,\ \omega_i \geq 0.$$

Por la estructura de (4) se puede reconocer que es equivalente a un problema de programación fraccional lineal ordinario. (Ver [10] y [7].) De hecho, utilizando la teoría de la programación fraccional lineal con la transformación

$$\omega_i = t v_i;\ i = 1, \ldots, m,$$
$$\mu_r = t u_r;\ r = 1, \ldots, s,$$
$$t^{-1} = \sum_r u_r y_{ro},$$

que, con $t > 0$, da lugar a

(6) $$\min f_o = \frac{\sum_{i=1}^{m} v_i x_{io}}{\sum_{r=1}^{s} u_r y_{ro}}$$

Sujeto a

$$\sum_{i=1}^{m} v_i x_{ij} - \sum_{r=1}^{s} u_r y_{rj} \geq 0;\ j=1,\dots, n.$$

$$v_i, u_r \geq 0,$$

equivalente al problema de programación fraccional lineal (4). Por manipulaciones evidentes, sin embargo, podemos ver que (6) es lo mismo que (2). Luego podemos usar (4) para resolver (6) y por lo tanto (2) y (1) también. Q.E.D.

Ahora estamos en una posición ventajosa desde varios puntos de vista. Tenemos una definición completamente simétrica de eficiencia que generaliza definiciones de una relación de producción no sólo en economía, sino también en ingeniería y otras ciencias naturales. No es necesario resolver los problemas no lineales (y no convexos) en los que fueron formalizadas estas definiciones. Sólo necesitamos resolver el problema de programación lineal ordinaria (4) a fin de obtener tanto las f_o^* y las h_o^* óptimas y las ponderaciones $v_i^*, u_r^* \geq 0$, ya que el cambio de variables no altera el valor del funcional.

Luego,

(7.1) $f_o^* = g_o^* = z_o^*$

y por consiguiente,

(7.2) $h_o^* = 1/z_o^*$.

También se tienen los pesos relativos buscados. Por lo tanto, no se requiere más que la solución de (4) o (3) para determinar si $f_o^* > 1$ o, correspondientemente, si $h_o^* < 1$, con eficiencia prevalente si y sólo si

(7.3) $f_o^* = h_o^* = 1$.

También podemos hacer extensiones en una variedad de nuevas direcciones.[6] Aquí, sin embargo, preferimos tomar contacto con otros desarrollos y también esbozar algunas de las ideas que se describen en otros documentos como Análisis de Envolvente de Datos.[7] A tal propósito, introducimos

(8) $$P_j = \begin{bmatrix} Y_j \\ \\ X_j \end{bmatrix}$$

[6] Por ejemplo, podríamos utilizar la teoría de la dualidad asociada con la programación fraccional (como se discute en [18] y [25] - ver también [5] y [7]), diferenciada de la teoría de la dualidad de la programación lineal ordinaria y de la teoría de la dualidad en la teoría de los costos y la producción (véase más adelante).

[7] Este es un método para ajustar datos a requerimientos teóricos prescritos tales como superficies de producción óptimas, etc., antes de realizar distintas pruebas estadísticas para propósitos de análisis de políticas públicas. Véase [21].

donde el subvector Y_j contiene los valores de producto observados y_{rj}, $r = 1, \ldots, s$ en sus componentes y el subvector X_j contiene valores de insumo observados x_{ij}, $i = 1, \ldots m$.

Sea ahora la siguiente reformulación vectorial de (3):

(9) $\max z_o$

con

$$-\sum_{j=1}^{n} Y_j \lambda_j + Y_o z_o \leq 0,$$

$$\sum_{j=1}^{n} X_j \lambda_j \leq X_o,$$

$$\lambda_j \geq 0; j=1,\ldots, n.$$

Representamos su solución óptima bajo la forma de ecuaciones con variables de holgura:

(10) $z_o^*, s^{*+}, s^{*-}, \lambda_j^*$

en donde s^{*+} representa un vector de variables de holgura no negativas asociadas a las desigualdades de los productos y s^{*-} representa un vector de variables de holgura no negativas asociadas a las desigualdades de los insumos. Si $z_o^* > 1$ por (7.1)-(7.3) la frontera eficiente de la superficie de posibilidades de producción *no fue alcanzada*.

Aquí, sin embargo, podemos observar algo más. Si s^{*+} tiene componentes positivas, entonces es posible aumentar los productos asociados en las cantidades de estas variables de holgura sin alterar ninguno de los valores λ_j^* y sin violar ninguna restricción. Similarmente si s^{*-} tiene componentes positivas entonces podemos reducir los insumos de X_o a $X_o - s^{*-}$ de manera análoga. Así, en cualquier caso, la UTD que está siendo evaluada no ha logrado eficiencia (relativa) aunque $z_o^* = 1$. Es decir que, a diferencia de (1) y (2), los modelos subsiguientes para caracterizar la eficiencia no determinan necesariamente si la UTD es eficiente sólo por referencia al valor funcional óptimo.

Para facilitar la referencia, resumimos lo implicado para estos últimos casos de la siguiente manera. Ninguna UTD puede ser calificada como eficiente a menos que las siguientes dos condiciones sean satisfechas:

(11) (i) $z_o^* = 1$

(ii) Todas las variables de holgura son nulas.

Se puede observar que estas condiciones son también condiciones para la *eficiencia de Pareto*[8] - extendida para cubrir tanto la producción como el consumo. Téngase en cuenta que esto supone que una reducción en cualquier insumo o una expansión en cualquier producto tiene algún valor, no requiere que estos valores sean estipulados o prescritos de antemano de ninguna manera. De hecho, si las medidas de eficiencia se limitan sólo a una medida escalar,

[8] También llamada eficiencia de Pareto-Koopmans. Véase el capítulo IX en [8].

entonces el cálculo objetivo de las ponderaciones de (4),[9] como ya se ha discutido, bastará para producir lo que se desea por sustitución directa en (1).

Supongamos ahora, por ejemplo, que queremos ajustar todas las observaciones con el propósito de evaluar el potencial de un programa para una UTD dada, suponiendo que este programa es administrado eficientemente por la UTD especificada. Esto puede hacerse aplicando (11) de la siguiente manera.

Primero, para una UTD seleccionada procedemos mediante (9) a obtener la solución (10). Luego formulamos un nuevo problema a partir de estos datos y su solución - a saber,

(12) $\max z^{\blacksquare}_{o}$

Con

$$-\sum_{j=1}^{n} Y_j \lambda^{\blacksquare}_j + (Y_o z_o^* + s^{*+})\, z^{\blacksquare}_o \leq 0,$$

$$\sum_{j=1}^{n} X_j \lambda^{\blacksquare}_j \qquad \leq X_o - s^{*-}$$

$$\lambda^{\blacksquare}_j \geq 0;\ j = 1, 2, \ldots, n.$$

Nos referiremos a (12) como el *problema variado* y vamos a demostrar que puede ser utilizado para eliminar todas las ineficiencias detectadas en el procedimiento de pasar de (9) a (10). Esto incluye: (a) reducir los insumos del vector de observaciones X_o original al nuevo vector de insumos (ajustado) $X_o - s^{*-}$, y también (b) aumentar el vector de producto observado originalmente Y_o a los nuevos valores (ajustados) $[Y_o z_o^* + s^{*+}]$.

Ahora mostramos que las observaciones así ajustadas satisfacen las condiciones de eficiencia en (11) como sigue. Evidentemente, debemos tener $z^{\blacksquare}_o{}^* > 1$ puesto que $z^{\blacksquare}_o{}^* = 1$ en (12) junto a (10) nos da el valor óptimo ya asegurado de la solución a (9). Ahora supongamos que tuviéramos $z^{\blacksquare}_o{}^* > 1$ en (12). Esto daría lugar a

$$-\sum_{j=1}^{n} Y_j \lambda^{\blacksquare}_j + Y_o z^{\blacksquare}_o{}^* z_o \leq \sum_{j=1}^{n} Y_j \lambda^{\blacksquare}_j + (Y_o z_o^* + s^{*+})\, z^{\blacksquare}_o{}^* \leq 0,$$

$$\sum_{j=1}^{n} X_j \lambda^{\blacksquare}_j{}^* \leq X_o - s^{*-} \leq X_o,$$

dado que tanto s^{*+} y s^{*-} son ambos no negativos. Evidentemente, las expresiones del primer miembro satisfarían el "problema no variado" (9) con $z^{\blacksquare}_o{}^*$ en lugar de z_o^*, y $\lambda^{\blacksquare}_j{}^*$ en lugar de λ_j^*. Empero, en tal caso también

$$\mathit{Max}\ z_o \geq z_o^*\, z^{\blacksquare}_o{}^* > z_o^*$$

toda vez que $z^{\blacksquare}_o{}^* > 1$. Pero, por hipótesis, $z_o^* = \max z_o$. Esta contradicción demuestra que el valor óptimo del problema variado (12) es $z^{\blacksquare}_o{}^*$.

Ahora queremos demostrar que la solución óptima, $\lambda_j^*, j = 1\ldots., n$, al problema no variado (9) es una solución óptima al problema variado (12) con cero de holgura, es decir, los vectores $s^{\blacksquare *+}$ y $s^{\blacksquare *-}$ tienen ceros en todos los componentes como se requiere para la eficiencia. Primero, vía (10),

[9] Cuando se utiliza un método de puntos extremos, tal como el método *simplex*, entonces se puede utilizar ya sea (3) o (4) ya que, como es bien sabido, estos métodos simultáneamente producen soluciones óptimas a ambos problemas. Véase, por ejemplo, [8].

$$-\sum_{j=1}^{n} Y_j \lambda_j^* + Y_o z_o^* + s^{*+} = 0,$$

$$\sum_{j=1}^{n} X_j \lambda_j^* \qquad\qquad = X_o - s^{*-},$$

Luego λ_j^* es una solución factible del problema variado con $z^{\blacksquare}_o = 1$. Es decir,

$$-\sum_{j=1}^{n} Y_j \lambda_j^* + (Y_o z_o^* + s^{*+})\, z^{\blacksquare}_o = 0,$$

$$\sum_{j=1}^{n} X_j \lambda_j^* \qquad\qquad = X_o - s^{*-},$$

con $z^{\blacksquare}_o = 1$. También es óptimo pues, como se ha mostrado, $z^{\blacksquare}_o{}^* = 1$. Adicionalmente, las holguras óptimas $s^{\blacksquare *+}$ y $s^{\blacksquare *--}$ son todas nulas. QED

En resumen, los ajustes indicados, de hecho, siempre llevan las observaciones originales al conjunto de producción eficiente relevante. No se requieren nuevos cálculos después de que se efectúen los ajustes z_o^*, s^*- para los datos Y_o, X_o originales al efecto de las comparaciones de eficiencia que posteriormente querríamos realizar.

Como veremos a continuación - en la sección (6) - podemos utilizar estos resultados para obtener una superficie que corresponda a una relación bien definida entre producción e insumos. En el caso de producción simple esta relación corresponde a una función en la que el producto es máximo para todos los insumos indicados. Por lo tanto, cumple formalmente los requisitos de una «función de producción» o, más en general, una «superficie de posibilidades de producción»[10] en el caso de múltiples productos. De esta manera obtendremos un nuevo tipo de función de producción que tiene una variedad de ventajas que indicaremos a medida que avancemos. Aquí podemos detallar algunas de estas ventajas de modo resumido como sigue. A diferencia de otros tipos de funciones de producción, éste deriva de (y es por lo tanto directamente aplicable a) observaciones empíricas. También elimina los problemas intratables de las agregaciones asociadas con otros tipos de funciones de producción.[11] Y, finalmente, se presta a la estática comparativa para propósitos tales como determinar si está ocurriendo cambio tecnológico. Estos usos estáticos "comparativos" se pueden lograr de varias maneras, como la adopción de la convención de que la misma UTD debe ser considerada como una entidad diferente en cada período de tiempo relevante.

Mediante estos conceptos de función de producción y los procedimientos que asociaremos, los requerimientos de la teoría económica podrán entonces ser traídos de nuevas maneras, modificadas para las evaluaciones de las políticas públicas y, al mismo tiempo, proporcionando una variedad de nuevas predicciones y de posibilidades de control para los administradores de programas - es decir, gerentes, legisladores, etc., que tienen responsabilidad sobre el programa. Por ejemplo, se torna posible distinguir entre «eficiencia del programa» - que se puede predecir con una gestión eficiente- y distinguirla de otras predicciones (y evaluaciones) que se podrían efectuar partiendo del supuesto de que todos los gestores conti-

---

[10] Estamos usando este término en este sentido de los vectores de actividad discutidos en Arrow y Hahn [4]. De hecho, en la Sección 5, a continuación, mostraremos explícitamente cómo nuestra caracterización de la dualidad puede utilizarse para obtener estimaciones numéricas de estos valores de coeficientes (o, más bien, sus contrapartes de la función de producción) a partir de datos observacionales.

[11] La desagregación puede ser necesaria, sin embargo, cuando los datos sobre UTD individuales no están disponibles. Ver las directrices suministradas en [12] y [21]. Para una discusión y un intento de hacer frente a las dificultades de la agregación en otros tipos de funciones de producción, véase [24].

nuarán operando sólo con los niveles de eficiencia anteriores.[12] De manera similar, es posible prever futuros cambios tecnológicos en lugar de suponer que se ha alcanzado una función de producción estática a largo plazo, como se hace actualmente en muchos de los estudios existentes sobre la orientación de la política energética en los países occidentales.

Debemos notar que esto puede ser logrado sin interferir con la posterior prueba y evaluación estadística. Como en el enfoque de DEA que discutimos en otra parte, estas pruebas estadísticas, que se aplican después de los ajustes indicados, a menudo se simplifican mucho con relación a otras alternativas. Proceder sin efectuar primero tales ajustes en los datos no sólo no utilizaría la teoría subyacente, sino que también contradiría los requisitos de esa teoría a menos que (a) se suponga que todas las UTDs están operando eficientemente o (b) se pueda proponer algún método alternativo de permitir observaciones que no están en la frontera de posibilidades de producción eficiente.[13]

4. <u>Análisis de isocuantas y eficiencia en sentido de Farrell</u>

Pasamos a otras implicaciones de nuestras definiciones de eficiencia y sus implicancias de producción - economía - gestión. Para ello, procedemos a la forma más familiar de análisis de isocuantas, y conceptos relacionados con la función de producción. Esto también nos permitirá entrar en contacto con el importante trabajo iniciado por M.J. Farrell [15].[14]

En primer término realizaremos la aclaración deseada (y proporcionaremos el contacto) a través de un análisis de isocuantas que corresponde al utilizado por Farrell. A continuación, suministraremos un modelo para generar la superficie eficiente relacionada de la función de producción. También relacionaremos esta última con una función de costo asociada, y una extensión del lema de Shephard. Entonces mostraremos más relaciones entre ésta y la eficiencia de Farrell, y entre tanto se verá cómo los duales (véase [8], [12] y [13]) de nuestros modelos pueden ser usados para proporcionar una nueva forma de estimar los coeficientes de una función de producción, ya sea de por sí o asociados con técnicas de estimación estadística (por ejemplo, a través del enfoque DEA descrito al final de la última sección).

Para iniciar el tratamiento observamos primero que ahora nos referimos al caso de un solo producto, que resulta ser el mismo para cada UTD. En este caso (3) se especializa en

(13) $\max z_o$

Sujeto a:

$$-\sum_{j=1}^{n} y_j \lambda_j + y_o z_o \leq 0,$$

$$\sum_{j=1}^{n} x_{ij} \lambda_j \leq x_{io}; \qquad i=1 \ldots m$$

$$\lambda_j \geq 0; \qquad j=1,\ldots n.$$

donde, como estamos interesados en un solo producto, se suprime el subíndice adicional de y_j para cada una de las $j = 1 \ldots, n$ UTDs. Haciendo el cambio de escala de las variables

[12] Véase [21] para una discusión más detallada y una aplicación detallada que distingue entre "eficiencia de programa" y "eficiencia de gestión" en el Programa "Follow Through" de la Oficina de Educación de los Estados Unidos.
[13] Para una discusión de los problemas involucrados en el tratamiento de tales problemas de estimación, incluso en el caso de un solo producto y muy pocos insumos ver [1] y [2].
[14] Los ejemplos que implican la continuación de este trabajo pueden encontrarse en [2] y [16].

$\lambda_j=\lambda_j'/y_j$, $z_o= z'_o/y_o$, eliminando las primas y el multiplicador constante $1/y_o$ en el funcional, obtenemos la forma equivalente,

(14) $\quad max\ z_o= \sum_{j=1}^{n} \lambda_j$

sujeto a:

$$\sum_{j=1}^{n} x'_{ij}\lambda_j \leq x'_{io};\ i=1,2,\ldots m$$

$$\lambda_j \geq 0;\ j= 1,2,\ldots n$$

donde los valores $x'_{ij}= x_{ij}/y_j;\ j= 1,\ldots n$ y $x'_{io}= x_{io}/y_o$ se obtienen a partir de la operación de cambio de variables recién descrita.

Hemos hecho la transformación de (13) a (14) de la manera indicada para facilitar el contacto con la obra de Farrell.[15] También representaremos (14) de modo equivalente como

(14.1) $\quad max\ z_o = \sum_{j=1}^{n} \lambda_j$

con

$$\sum_j P_j \lambda_j \leq P_o,$$

$$\lambda_j \geq 0; \qquad j= 1, 2, \ldots n,$$

o en forma equivalente:

(14.2) $\quad max\ z_o = \sum_{j=1}^{n} \lambda_j$

con

$$\sum_{j=1}^{n} P_j \lambda_j + \sum_{i=1}^{m} e_i s_i = P_o, \text{ y}$$

$$\lambda_j, s_i \geq 0,$$

donde los P_j representan:

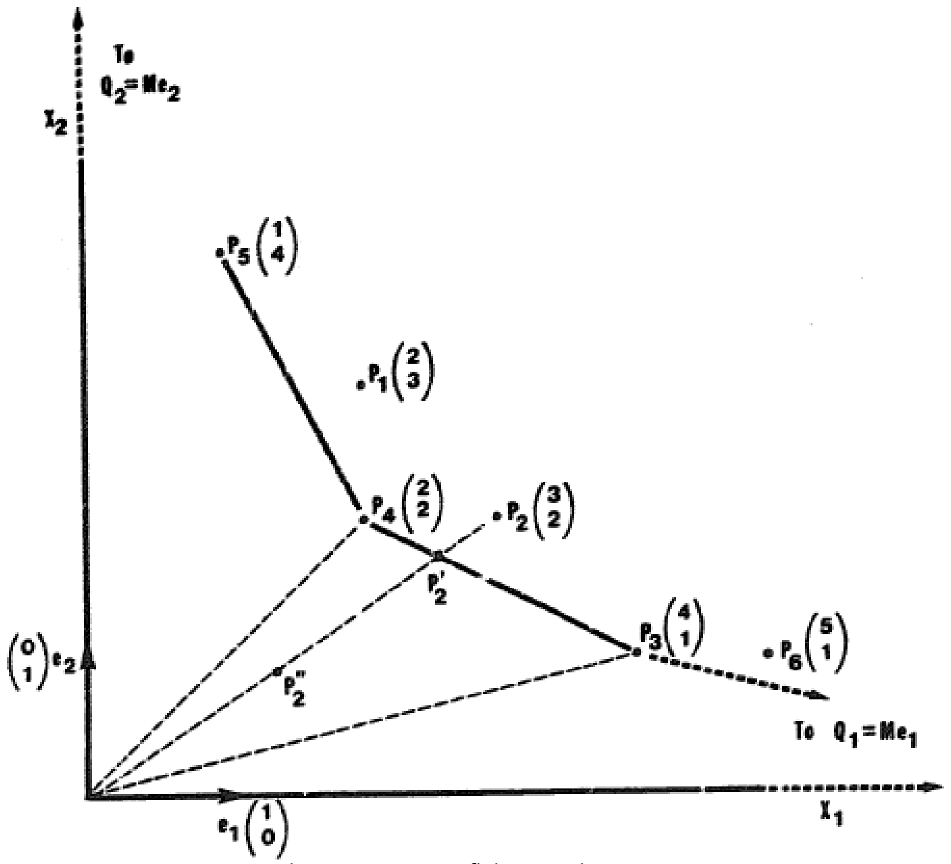
(a) una versión normalizada de (8), en la que los componentes de Xj son divididos por sus cantidades correspondientes de producto, y

(b) la producción es entonces eliminada del vector y (en versión normalizada) asociada en su lugar con los coeficientes unitarios de las variables λ_j en el funcional. En suma, $P_j = X'_j$ con sólo entradas normalizadas x'_j de sus componentes. Los e_i son vectores unitarios con la unidad en la fila $i = 1,\ldots.\ m$ y ceros en otras filas y las s_i son variables de "holgura".

Las condiciones (11) continúan aplicándose, por supuesto, pero ahora el referente para la eficiencia es sólo a los insumos.[16] Para ver lo que esto significa referirse a la Fig. 1 en la que se representa una situación para seis UTDs usando dos insumos normalizados por sus respectivos valores de producto. Aquí, las UTD están todas representadas para el caso de dos insu-

[15] Para establecer pleno contacto con el trabajo de Farrell también es necesario tratar (y eliminar) algunos de sus conceptos incómodos tales como "puntos en el infinito", etc. Esto se hace en [12] pero a efectos de la brevedad no repetiremos estos desarrollos aquí.
[16] Nos referimos sólo a los insumos para acortar la discusión que sigue, aunque: por supuesto, las manipulaciones también pueden extenderse a posibles aumentos de producción.

mos normalizados por sus producciones respectivas. Por ejemplo, P_5 tiene $x'_{15} = 1$, $x'_{25} = 4$ mientras que P_2 tiene $x'_{13} = 3$, $x'_{23} = 2$, y así sucesivamente.



<u>Figura 1.</u> Puntos eficientes e isocuanta

Supongamos que ahora se decide probar la eficiencia de P_2. Esto puede hacerse insertando $P_2 \equiv P_0$ en (14.2) y utilizando cualquier método de punto extremo adyacente, como los métodos simplex o dual,[17] para obtener:

$$(5/6)\ P_4 + (1/3)\ P_3 = P_0 = P_2.$$

Es decir, estos métodos expresan P_2 en términos de una base óptima con $\lambda_4^*= (5/6)$, $\lambda_3^* = (1/3)$ y todos los demás $\lambda_j^* = 0$. Luego tenemos $z_0^*\ (P_2) = (7/6)$ con lo cual, con referencia a (11), P_2 no es eficiente.

Ahora podemos otorgar a esta medida de eficiencia significado operativo <u>*convexificando*</u> la expresión anterior, lo que hacemos dividiendo por 7/6, para obtener:

$$(5/7)\ P_4 + (2/7)\ P_3 = (6/7)\ P_2 = P'_2.$$

Como aparece en la Figura 1, esto expresa P'_2 como una combinación convexa de los vectores de la base óptima, P_4 y P_3, y por lo tanto traslada el punto P'_2 resultante sobre el segmento de línea que los conecta. En términos de eficiencia, esto significa que si P_2 hubiera estado produciendo con la misma eficiencia que P'_2 - o, de forma equivalente, tan eficientemente como

[17] Ver la discusión de estos métodos y un análisis de sus propiedades especiales en el capítulo XIII de [8].

la combinación convexa indicada de P_4 y P_3 - entonces debería haber sido capaz de asegurar su unidad de producción con sólo (6/7) de la cantidad de cada uno de los insumos x_1 y x_2 que se observó utilizando.

La línea que conecta P_4 y P_3 representa un segmento de lo que Farrell denomina la *isocuanta unitaria*. También es eficiente. Es decir, no puede haber un punto tal como P''_2 entre esta isocuanta unitaria y el origen ya que esto implicaría que P_4 y P_3 no eran una base óptima.[18]

El análisis anterior ha empleado dos hipótesis, a las que nos referiremos como hipótesis de "isocuanta" e hipótesis "radial", respectivamente. Esta última, es decir, el supuesto radial, que corresponde a asumir retornos constantes a escala,[19] puede ser relajada pero a un costo en complicaciones y explicaciones resultantes que no vamos a realizar aquí. La primera, es decir, el supuesto de representación de la isocuanta es crítica y no puede ser relajada en el sentido de que proporciona el punto de comparación entre P_2 y P'_2.[20]

Acabamos de describir la importancia operativa que deseamos otorgar a nuestra medida de eficiencia para el caso de P_2. Consideremos P_1 para el cual obtenemos

$$(5/6)\ P_4 + (2/6)\ P_5 = P_1$$

cuando empleamos métodos de puntos extremos adyacentes. Es decir, la base óptima consiste en P_4 y P_3 (que se caracterizan ambas por ser eficientes) con $z_0^*(P_1) = \lambda_4^* + \lambda_5^* = (7/6)$ de modo que, en virtud de (11), P_1 *tampoco es eficiente*.

Como se aprecia, $z_0^*(P_1) = z_0^*(P_2) = (7/6)$. Sin embargo, P_1 y P_2 están expresadas en términos de diferentes bases y por lo tanto tienen referentes diferentes. El proceso de convexificación utilizado para P_2 también se puede emplear para P_1 obteniéndose:

$$(5/7)\ P_4 + (2/7)\ P_5 = (6/7)\ P_1 = P'_1$$

De modo que también se necesita la misma relación de contracción para todos los recursos para llevar P_1 a la superficie eficiente. Sin embargo, se encuentra en el cono a través del origen formado por P_4 y P_5, mientras que P_2 está en el cono formado a partir de P_4 y P_3. Esta condición (que surge de la no negatividad impuesta sobre los valores λ_j admisibles) tiene una ventaja para los tipos de aplicaciones de programas públicos que estamos considerando. Como se observa en la sección inicial, muchas UTDs, tales como diferentes distritos escolares, etc., trabajan bajo diversas restricciones con respecto a los insumos (así como los productos) para el mismo programa realizado en diferentes lugares o diferentes partes del país. Por lo tanto, es bueno tener los referentes utilizados para calificar la eficiencia de cada UTD tan iguales a ella como sea posible, al menos en algún sentido lato, sin interferir "demasiado" con las calificaciones de eficiencia deseadas.[21]

---

[18] Las pruebas de estas proposiciones, que son bastante transparentes, no serán dadas en el presente trabajo. Pueden encontrarse en [12]. Esta última fuente también puede ser consultada sobre las formas en que las formulaciones aquí utilizadas difieren de la sugerida a Farrell por A. Hoffman en la discusión asociada con [15] y posteriormente empleadas por Farrell y Fieldhouse en [16].

[19] Obsérvese, sin embargo, que esta constante será diferente, en general, para cada P_j.

[20] No obstante, debe entenderse que nuestros análisis son también aplicables cuando se utilizan diversas transformaciones para llevar otras funciones tales como, por ejemplo, las funciones de Cobb-Douglas, a representaciones apropiadas lineales por tramos.

[21] Obsérvese que tanto P_1 como P_2 están también en el cono formado de P_5 y P_3 en la Fig. 1 pero el segmento de línea que conecta P_3 y P_5 no es eficiente - y por lo tanto P_3 y P_5 no serán una base óptima

Concluimos ahora esta sección con el caso de P_6 en la Fig. 1, para la cual están presentes los óptimos alternativos ya que

$$1\ P_6 = P_6 \text{ y } 1\ P_3 + 1\ e_1 = P_6$$

con $z_0^*(P_6) = 1$ en ambos casos. Obsérvese, sin embargo, que la segunda de estas dos soluciones tiene $s_1^* = 1$, es decir que la eficiencia no se alcanza para P_6 sino hasta que se reste la holgura en la cantidad de $s_1^* = 1$ de la primera componente en P_6 - después de lo cual es coincidente con P_3 en la Fig. 1; véase (11).

Para evitar la ambigüedad y mantener el significado operativo deseado para la eficiencia, es necesario maximizar los valores de holgura, pero de manera de otorgar a ésta una menor prioridad que maximizar los valores de λ_j. Esto se hace reemplazando (14.2) por

(14.3) $\max z_0 = \sum_{j=1}^{n} \lambda_j + (1/M) \sum_{i=1}^{m} s_i$

Sujeto a:

$$P_0 = \sum_{j=1}^{n} P_j \lambda_j + \sum_{i=1}^{m} e_i s_i,$$

$$\lambda_j, s_i \geq 0;\ j = 1,\ldots n;\ i = 1,\ldots m$$

en donde M es la "cantidad" usual amplia (no arquimediana) (cf, por ejemplo, [22]) que asegura que $1/M > 0$ es siempre menor que cualquier real positivo que pueda ser asumido por cualquier λ_j.

5. Relaciones de dualidad para la estimación de coeficientes

Farrell distinguió entre la ineficiencia antes mencionada,[22] a la que denominó «ineficiencia técnica», y otros tipos de ineficiencia a las que se refirió como «ineficiencia precio» o «asignativa», e «ineficiencia global», caracterizada tanto por «ineficiencias técnica y asignativa». Aquí podemos notar que Farrell restringió sus estudios principalmente a la ineficiencia técnica - y por las razones expuestas en nuestra introducción haremos lo mismo. En lo que respecta a la ineficiencia precio (en el sentido de «precios reales de mercado»), Farrell se contentó, en su mayor parte, con apuntar las enormes dificultades que entraña evaluar incluso la ineficiencia de precios *relativos*, por ejemplo, por los diversos motivos de los compradores y vendedores. Sin embargo, se puede llegar a este problema desde el punto de vista de "precios en principio" y/o "costos de oportunidad", como los que se obtienen a través de teoremas económicos estándar. En cualquier caso éste es el camino que seguiremos para mostrar cómo las formulaciones duales a las anteriores de programación lineal pueden usarse para obtener los valores de los coeficientes de la pendiente de las isocuantas eficientes.

Por supuesto, generalmente se tratará con representaciones m-dimensionales en las que segmentos de línea tales como los representados en la Fig. 1 son reemplazados por "facetas" eficientes. Sin embargo, ahora disponemos de un modelo y un método para generar estos

---

a menos que se quiera imponer más restricciones sobre las opciones de base. Ver la discusión en la Sección 2, más arriba.

[22] La representación de más arriba, dada al final de la sección anterior, retiene implícitamente conceptos como "puntos en el infinito", por ejemplo, como se representa por Q_1 y Q_2 en la Fig. 1 - pero dejamos a un lado el desarrollo necesario para eliminarlos. En cualquier caso, podemos recurrir a (1) para la medida escalar de eficiencia deseada sin referencia a estas consideraciones. Es decir, $h_o^* = 1$ entonces es suficiente para determinar si el P_o correspondiente es o no eficiente.

factores para cualquier número finito de insumos,[23] ya que estas facetas corresponden a todas las combinaciones convexas de puntos que pueden ser generados a partir de bases óptimas. La especificación de estas bases óptimas constituye así una manera de representar estas facetas.

Otra forma está disponible por referencia a las *relaciones de dualidad de la programación lineal* como ahora demostraremos. Por lo tanto, escribimos el dual de (14.1) como

(15) $\min g_o = \omega^T P_o$

Con

$\omega^T P_j \geq 1;\ j=1,\ldots, n,$

$\omega^T \geq 0,$

donde el superíndice T representa trasposición, como de costumbre, por lo que ω^T representa la traspuesta del vector ω con componentes $\omega_1, \omega_2, \ldots \omega_m$ y ω^{*T} denota un vector óptimo de estas variables en el problema mencionado.

Ahora observamos que $\omega^{*T} P_i = 1$ para cada P_i en una base óptima. Para obtener nuestra representación alternativa de esta faceta en términos de las pendientes de la superficie eficiente de la isocuanta, por lo tanto, sólo necesitamos mostrar que ω^{*T} es ortogonal (ver [20] y/o [8]) a la faceta eficiente generada por estos P_i. Para ello basta con demostrar que ω^* es ortogonal a cualquier dirección situada en la faceta, por ejemplo, a cualquier vector que sea la diferencia, $\bar{P} - \bar{\bar{P}}$, de dos vectores en la faceta. Puesto que, por supuesto, $\bar{P}$ y $\bar{\bar{P}}$ están en la faceta, podemos expresarlos en términos de estos mismos P_i mediante

(16.1)

$$\bar{P} = \sum_i P_i \bar{v}_i, \qquad \bar{\bar{P}} = \sum_i P_i \bar{\bar{v}}_i,$$

$$0 \leq \bar{v}_i, \bar{\bar{v}}_i \leq 1, \qquad \sum_i \bar{v}_i = \sum_i \bar{\bar{v}}_i = 1$$

donde la suma es realizada sobre todos los índices de estos P_i. Pero en tal caso,

(16.2)

$$\omega^{*T}(\bar{P} - \bar{\bar{P}}) = \sum_i (\omega^{*T} P_i \bar{v}_i - \omega^{*T} P_i \bar{\bar{v}}_i)$$

$$= \sum_i (1\, \bar{v}_i - 1\, \bar{\bar{v}}_i)$$

$$= \sum_i \bar{v}_i - \sum_i \bar{\bar{v}}_i = 0,$$

dado que $\Sigma_i \bar{v}_i = \Sigma_i \bar{\bar{v}}_i = 1$. QED. Por lo tanto, el ω^* correspondiente a esta faceta eficiente determinada por esta base óptima es ortogonal (o normal) a ella. Así, ω^* es normal al hiperplano que contiene esta faceta. La ecuación de este hiperplano es

[23] Cualquier método de punto extremo adyacente lo hará.

(17) $\omega^{*T}x = 1;$

es decir, cualquier x que satisfaga esta ecuación es un punto en el espacio lineal generado por la totalidad de los P_j y e_i. La porción del hiperplano que consiste en la faceta eficiente es el conjunto de todas las x que son combinaciones convexas de los P_i que forman la base óptima.

Por supuesto, necesitaremos obtener generalmente diferentes ω^{*T} para las "pendientes" asociadas con las diferentes facetas de las superficies eficientes y los rangos para los que son válidas. Estos valores de ω^{*T}, sin embargo, se pueden obtener dentro del proceso computacional para las λ^*. Como bien se sabe (véase, por ejemplo, [8]), los métodos de puntos extremos adyacentes (tales como los métodos simplex y dual) resuelven simultáneamente tanto (14.1) como (15). Es decir, los valores óptimos λ^* y ω^{*T} se obtienen a partir de la misma tabla. Por lo tanto, la solución a uno de estos problemas también proporciona la solución al otro sin esfuerzo computacional adicional.

Por ejemplo, la tabla que proporcionó la solución y, por lo tanto, la base óptima para caracterizar la eficiencia de P_1 con $\lambda_4^* = 5/6$; $\lambda_5^* = 2/6$ - ver sección precedente - también proporciona las variables duales asociadas $\omega_1^* = 1/3$, $\omega_2^* = 1/6$. Pero, como se ha señaló anteriormente, podemos expresar cualquier $x^T = (x_1, x_2)$ en esta faceta por medio de los vectores de la base, P_4, P_5, como

$$x = P_4 v_4 + P_5 v_5$$

con $v_4, v_5 \geq 0$, y $v_4 + v_5 = 1$. Luego también

$$\omega^{*T}x = \omega_1^* x_1 + \omega_2^* x_2 \ .$$

Para este segmento, se tiene $\omega_1^* = \frac{1}{3}, \omega_2^* = \frac{1}{6}$ y por tanto,

$$S[P_5, P_4] \equiv$$
$$\{(x_1, x_2): \tfrac{1}{3}x_1 + \tfrac{1}{6}x_2 = 1 \ ; 1 \leqslant x_1 \leqslant 2 \ , 2 \leqslant x_2 \leqslant 4\} \ ,$$

donde los corchetes significan que los puntos P_5, P_4 indican esta faceta (= segmento) con la relación prescrita para los rangos indicados entre las llaves.

En forma similar, para $S\,[P_4, P_3]$ se tiene

$$S[P_4, P_3] \equiv$$
$$\{(x_1, x_2): \tfrac{1}{6}x_1 + \tfrac{1}{3}x_2 = 1 \ ; 2 \leqslant x_1 \leqslant 4 \ , 1 \leqslant x_2 \leqslant 2\} .$$

En tanto, el segmento de isocuanta desde P_3 a Q_1 puede ser representado por

$$S[P_3, Q_1) \equiv$$
$$\{(x_1, x_2): \ 1 = \frac{1}{M}x_1 + x_2 ; \ 4 \leqslant x_1, 1 \leqslant x_2\}$$

donde el corchete indica que P_3 debe ser incluido mientras que el paréntesis de la derecha indica que $Q_1 = Me_1$, es excluido del conjunto y similares características son aplicables a los puntos como Q_2, etc.

Ahora reemplazamos (17) con

(17.1) $S\,[P_i\text{: para todo } i \in I] =$

$$= \{x\text{: } \omega^{*T} = 1;\ x = \textstyle\sum_{i \in I} P_i v_i \text{ para todo } v_i \geq 0,\ \sum_{i \in I} v_i = 1\}$$

a efectos de extender el análisis precedente a mayores dimensiones. Aquí las superficies eficientes se componen de facetas, y ω^{*T} es la normal a tal faceta (un simplex) (ver [8], p.242) que es el conjunto convexo generado por la base óptima. La teoría que se ha desarrollado sobre los métodos usuales de puntos adyacentes extremos, incluyendo sus extensiones no lineales, nos permite obtener toda la información necesaria para la interpretación, así como la prolongación de una faceta eficiente a otra por medio de las estrategias computacionales que estos métodos permiten en conjunción con las mismas.[24]

Para concluir esta sección podemos obtener una manera directa de interpretar los componentes del vector normal, ω^{*T}, como productividades marginales de los insumos indicados. Hacemos esto volviendo al caso bidimensional, tal como el representado en la Fig. 1. La condición para permanecer en la misma isocuanta al efectuar sustituciones es

$$\omega_1^* x_1 + \omega_2^* x_2 = \hat{\omega}_1^* \hat{x}_1 + \hat{\omega}_2^* \hat{x}_2\ ,$$

donde $x \neq \hat{x}$ son dos puntos en la isocuanta indicada, pero en segmentos posiblemente diferentes con vectores de pendiente ω^* y $\hat{\omega}^*$. Sumando $\omega_1^* \hat{x}_1 - \omega_1^* \hat{x}_1 = 0$ a la izquierda y la expresión $\omega_2^* \hat{x}_2 - \omega_2^* \hat{x}_2 = 0$ a la derecha, haciendo algunas manipulaciones algebraicas se obtiene[25]

$$\frac{\Delta x_2}{\Delta x_1} = -\frac{\omega_1^*}{\omega_2^*} - \frac{\Delta\omega_1}{\Delta x_1}\frac{\hat{x}_1}{\omega_2^*} - \frac{\Delta\omega_2}{\Delta x_1}\frac{\hat{x}_2}{\omega_2^*}$$

en cuya expresión:

$$\Delta x_1 = x_1 - \hat{x}_1\ ; \qquad \Delta x_2 = x_2 - \hat{x}_2$$
$$\Delta\omega_1 = \omega_1^* - \hat{\omega}_1^*\ ; \qquad \Delta\omega_2 = \omega_2^* - \hat{\omega}_2^*\ .$$

Cuando x y $\hat{x}$ están ambos en el mismo segmento de una isocuanta, se tendrá $\Delta\omega_1 = \Delta\omega_2 = 0$. De hecho, en este caso tendremos una derivada bien definida y por lo tanto podemos reemplazar la formulación de diferencias finitas anterior con

$$\frac{dx_2}{dx_1} = -\frac{\omega_1^*}{\omega_2^*}\ ,$$

que es la expresión usual que relaciona la tasa marginal de sustitución entre dos insumos, caracterizada por la derivada a la izquierda, y la recíproca negativa de la relación de sus productividades marginales, a la derecha, con igualdad requerida entre los dos lados para permanecer en la misma isocuanta. En resumen, $\omega_1{}^*$ es la productividad marginal del i-ésimo factor, $i = 1, 2$. Puesto que, por los supuestos de la economía, estas productividades margina-

[24] Por ejemplo, el método simplex de G.B. Dantzig y el método dual de C.E. Lemke pueden ser considerados para estos fines. Ver [12] y [21] para discusión adicional así como ilustraciones numéricas. Para una discusión sobre la riqueza de información relevante para las interpretaciones económicas y de gestión (políticas) disponibles en los cuadros resultantes, véase el cap. VI en [8].

[25] Si se desea, esto puede ser reformulado en términos de elasticidades de sustitución.

les nunca son negativas, se deduce que la tasa marginal de sustitución es no positiva y la isocuanta también se supone convexa (y continua), etc. si queremos mantener contacto con los teoremas usuales en estas partes de la teoría económica.

6. Funciones de producción y relación costo-precio

Ahora estamos listos para construir la función de producción y las relaciones de costo a partir de los valores derivados de ω^*. Antes de proporcionar el modelo para hacerlo, sin embargo, podemos refinar algunos de los puntos que ya hemos hecho haciendo una pausa para ver lo que hemos logrado y cómo se relaciona con otros tipos de modelos y enfoques en economía.

Nuestra función se deriva evidentemente de observaciones empíricas. Aunque estas observaciones están todas al nivel de las "firmas" individuales,[26] evidentemente tenemos algo que difiere de otros estudios y estimaciones de función de producción a nivel individual de la empresa (véase [19]) ya que, por hipótesis, estamos considerando todas las UTDs pertinentes.

Enfoques que incluyen todos estos datos se han llevado a cabo habitualmente sólo a niveles agregados, con todas las dificultades (y supuestos) necesarios para asegurar que las funciones así estimadas tengan las propiedades extremales que debe poseer una "función de producción".[27] Nuestras funciones difieren evidentemente de estas últimas en que (a) los datos no son agregados antes de la estimación y (b) las estimaciones resultantes son óptimas - es decir, tienen las propiedades extremales requeridas - en la medida en que los datos lo permitan.[28] Se deduce que tenemos derecho a utilizar estas propiedades de optimalidad para deducir los teoremas y relaciones adicionales que reseñaremos a continuación,[29] ya que también se refieren a estos mismos datos o bien a estimaciones derivadas de los mismos a través de métodos y principios de optimización rigurosamente establecidos.

En cierto modo, estas funciones de producción recuerdan las que podrían estar asociadas con el concepto de Alfred Marshall de "empresa representativa". (Discusión en [28], también [9]). Aquí, sin embargo, el referente es más bien a "empresas representativas eficientes". Téngase en cuenta que el plural es necesario en la medida en que hay más de una sola faceta eficiente. El continuo dentro de cada faceta es entonces "representativo" de la eficiencia para la cual las firmas eficientes originalmente observadas sirven como referentes.

Hemos comentado anteriormente la utilidad de estas facetas "representativas" (y los conos generados a partir de ellas)[30] para evaluar la eficiencia de las UTDs en varios programas públicos. Proporcionan un referente igualmente apropiado para hacer estimaciones de lo que

---

[26] Utilizamos el término "empresas" de forma intercambiable con las UTDs en esta sección a fin de evitar circunlocuciones al relacionar estos desarrollos con las partes pertinentes de las economías tradicionales. Véase Johnston [19].
[27] Para una discusión de las onerosas (y poco realistas) condiciones necesarias para obtener estas propiedades a nivel agregado, véase Sato [24], pp. 3-8 ff.
[28] Incluyendo las asignaciones *ex cathedra*, cuando estén disponibles, como se indica en la Sección 2.
[29] Para una crítica adicional de la utilización de estas propiedades de óptima en asociación con tales funciones agregadas (por ejemplo, en algunos estudios actuales de política energética), junto con simples contra-ejemplos a los supuestos utilizados en tales estudios ver [13].
[30] Véase la descripción de los «conjuntos convexos representativos» y los conos que «representan» en la p. 236 en [8].

las distintas UTDs deben ser capaces de producir como productos[31] dada la cantidad de factores y / o las relaciones entre los diversos insumos que pueden ser prescritos para ellos. Esto quiere decir que queremos referir cada UTD a su faceta (o cono) "representativo" en camino para realizar las traducciones y transformaciones prescritas en (12) a fin de trasladarlas a la superficie de la función de producción relevante. Teniendo en cuenta las últimas proyecciones, estamos en condiciones de estimar qué productos podemos esperar con una producción eficiente de varias asignaciones de recursos a cada UTD en los programas que se están considerando. Estos resultados pueden agregarse en una variedad de maneras para evaluar o controlar las actividades que serán generadas por estas UTDs.[32]

Para permitir todas las posibles variaciones en tales insumos y productos, por supuesto, necesitamos tener una forma adecuada de obtener la superficie de la función de producción. Por lo tanto, ahora proporcionamos una manera de hacerlo que relaciona nuestros resultados con otro tipo de "dualidad" en economía.[33]

Esta última, introducida por Samuelson [23] y Shephard [26, 27], procede introduciendo una función de costo $C\,(y, p)$ determinada por medio de

(18) $C\,(y, p) = min\; p^T x$ para $x \epsilon L\,(y)$

donde

$L\,(y) \equiv \{x:$ al menos el vector de productos y puede ser producido$\}$.

En otras palabras, $L\,(y)$ es la asignación de punto a conjunto $y \rightarrow L\,(y)$. Por ejemplo, en el caso de producción simple, como en la Fig. 1, si se obtiene a través de la isocuanta asociada a y = 1, $L\,(y)$ designaría el conjunto de todos los puntos $x = (x_1, x_2)$ interpretados como combinaciones de insumos en o al noreste de esta isocuanta. A partir del conocimiento de estas relaciones, la función de costo $C\,(y, p)$ será obtenida a través de la minimización indicada donde p es un vector 'precio' con la componente p_i que representa el 'precio' por unidad x_i, la cantidad del i-ésimo insumo factorial.

En nuestro caso deseamos que nuestra función de producción esté empíricamente basada. Es decir, queremos que nuestra función de producción se base en datos de insumos observados o estimaciones derivadas de ellos de tal manera que ninguna empresa del conjunto de observación tenga una producción más elevada para cualquier insumo que se pueda especificar. Tampoco una combinación no negativa de estas empresas puede tener una mayor producción cuando se deben realizar extrapolaciones o interpolaciones de las observaciones originales para los valores especificados de los insumos.

A estos fines ahora procedemos a obtener la función de producción deseada de la siguiente manera. Sea y una producción prescrita[34] y que a_s^T, la fila s de la matriz A, represente el conjunto de coeficientes asociados con la faceta s eficiente estimada a partir de los datos, como

[31] Incluyendo productos múltiples como se discutió en la Sección 2, ya que los principios que estamos discutiendo se extienden en formas bastante obvias al caso más general de superficies de posibilidades de producción.
[32] Los principios a utilizar para efectuar estas agregaciones son expuestos en [12].
[33] En otra parte hemos sugerido que estas relaciones podrían asociarse mejor con las matemáticas de la "teoría de las transformaciones". Véase [13].
[34] Podemos, por supuesto, extender esto al caso de producción múltiple muy fácilmente, como en (18).

se describe en la sección anterior. Es decir, $a_{si} = \omega_{si}^*$. Véase (17) o (17.1). Entonces sea también P una matriz con sus vectores columna P_j que representan los datos observacionales para cada una de las UTDs originales $j = 1,\ldots,$ n.[35]

Luego se escribe

(18.1) $\quad min\ p^T x$

sujeto a

$$Ax \geq ye,$$

$$-P\lambda + Ix = 0, \quad \lambda \geq 0,$$

donde e es un vector columna con todos los elementos iguales a la unidad mientras que I es la matriz de identidad de manera que (i) $Ix = P\lambda,\ \lambda > 0$ nos asegura que derivaremos nuestra función de producción a partir de observaciones empíricas y (ii) la forma en que los vectores a_s fueron derivados en $a_s^T x \geq y$, junto con el objetivo de minimización, nos aseguran que estaremos siempre en una frontera eficiente cuando se obtenga una solución óptima a (18.1) para cualquier y prescrito.

Los valores x^*, y resultantes de las soluciones a (18.1) para diferentes p son todos puntos de la superficie de la función de producción. Cuando esta superficie está disponible, puede evidentemente procederse también de manera inversa, si se desea, para estimar el valor de producto que se puede asegurar de cualquier x prescrito.

Además, está disponible más, ya que el modelo anterior se puede utilizar de diversas maneras. Aquí, sin embargo, queremos principalmente relacionarlo con las relaciones de dualidad de la función de costo-de producción de la teoría económica. Por tanto, escribimos el dual de programación matemática a (18.1) como

(18.2) $\quad max\ y\eta^T e$

Sujeto a

$$\eta^T A + u^T I = p^T,$$

$$-u^T P \leq 0, \qquad \eta^T \geq 0.$$

Por medio del teorema de dualidad de la programación lineal, se tendrá:

(19.1) $\quad p^T x \geq y\eta^T e$

para todos los x, λ, y η, u que satisfagan las restricciones y

(19.2) $\quad p^T x^* = y\, \eta^{*T} e,$

en el óptimo. En otras palabras,

(19.3) $\quad C\,(y, p) = y\, \eta^{*T} e = p^T x^*$

[35] Se puede omitir los Pi que no son eficientes o bien se puede ajustarlos y llevarlos al conjunto eficiente de la manera indicada en (12).

es la función de costo (minimizante) requerida, la cual varía con cada elección de y y p.

Aquí hemos pasado de la función de producción a la función de costo, pero, por supuesto, también podríamos haber avanzado por el camino inverso. Sin embargo, esta última es sólo una declaración "en principio", ya que, como se señaló en la introducción, muchos de los insumos y productos de las aplicaciones del sector público no son fáciles de tasar ni calcularse sin recurrir a procedimientos y supuestos arbitrarios y ex cathedra. Además, el uso de datos de precios también presenta dificultades como las que se discutieron en relación con la eficiencia de Farrell (en la sección anterior), a las que ahora cabría agregar otras dificultades como determinar si los datos de cada UTD se refieren a precios de lista o reales y si se han hecho provisiones por consideraciones de financiamiento, mezcladas con consideraciones de producción, tales como descuentos en efectivo, descuentos por cantidad, etc.

Aunque preferimos continuar en el lado de la producción por razones como las que acabamos de indicar, puede haber casos en donde los precios también (o alternativamente) se pueden utilizar con ventaja. En otra parte [13] hemos extendido el "lema de Shephard" dejando de lado el supuesto de diferenciabilidad, por lo que ahora adaptamos ese resultado para su uso también en el presente caso.

A tal fin volvemos a expresar (19.2) de la siguiente forma:

(20.1) $p^T x^* = y\, \eta^{*T} e = y \sum_s \eta_s^* = y\, c$

donde

(20.2) $c \equiv \sum_s \eta_s^*$,

por lo que c es el costo unitario total de producir y unidades de producto. Nótese que generalmente podremos variar los componentes de p^T dentro de algún rango sin alterar los valores de x^*.[36] Así, si seleccionamos un componente, digamos p_i para tal variación, podemos obtener

(21.2) $\Delta p_i\, x_i^* = y \Delta c$

o bien,

(21.2) $x_i^* = y \dfrac{\Delta c}{\Delta p_i} = \dfrac{\Delta C(y, p)}{\Delta p_i}$

que no es otra cosa que el lema de Shephard en forma discreta.

La definición de eficiencia en (1) nos ha permitido, por tanto, establecer contacto con esta línea de trabajo. También podemos hacer más. Podemos abrir nuevas posibilidades para emplear esta última, por ejemplo, al efectuar estimaciones de relaciones extremales a partir de datos originales, tal como lo hemos hecho para la eficiencia de Farrell.

Para mostrar brevemente tales posibilidades bastará con relacionar la medida de distancia de Shephard con la eficiencia de Farrell en el contexto de los modelos de la Sección 2. Esto nos permitirá tratar esta medida en el contexto de los datos originales. También nos permitirá cerrar esta sección volviendo al caso de producción múltiple.

---

[36] Condiciones necesarias y suficientes aplicables a cada componente de p^T se dan en [81], cap. IX.

En la notación de Shephard reemplazamos la primera porción de (18) con:

(22.1) $$C(y, p) = \min_x \{p^T x: \Psi(y, x) \geqslant 1\}$$

donde

$$\Psi(y, x) \equiv \frac{1}{v^*(y, x)},$$

con

$$v^*(y, x) = \min v,\ vx \in L(y),\ v \geqslant 0,$$

y *L (y)* se define como en la segunda parte de (18). Esta $\Psi(y, x)$ es la que Shephard llama una 'función distancia'.[37]

El desarrollo anterior supone que las relaciones de producción han sido determinadas y por lo tanto están disponibles, como en (18.1), para la determinación de *x*. Aquí, sin embargo, queremos considerar a estos últimos como datos. Por lo tanto, volvemos a (8) y escribimos

(22.2) $\min v$

con

$$vX_o - \sum_{j=1}^{n} X_j \lambda_j \geq 0,$$

$$\sum_{j=1}^{n} Y_j \lambda_j \geq Y_o,$$

$$\lambda_j \geq 0;\ j=1,\ldots, n,$$

donde ahora $Y_o \equiv y$ es un vector de productos.

Si consideramos $z_o^* \geq 1$ como una extensión de la eficiencia de Farrell al caso de producción múltiple de (9), también podemos considerar $\min v = v^*$ en la formulación anterior como una extensión de la función de distancia de Shephard (y conceptos relacionados) para su uso en la evaluación de la eficiencia de las UTDs. Como una comparación con (9) deja claro, estos valores z_o^* y v^* son complementarios entre sí. (De hecho, son recíprocos, véase [13]). En otras palabras, la función de distancia de Shephard se convierte en una medida de eficiencia con $v^* < 1$ y $v^* = 1$ sólo cuando la UTD en cuestión es eficiente en relación con las otras observaciones y/o condiciones teóricamente prescritas (por ejemplo, a través de condiciones de ingeniería ex cathedra).

De esta manera, hemos empleado nuestra definición de eficiencia no sólo para unir dos vertientes previamente separadas de investigación económica, sino también para abrir nuevas posibilidades para cada una de ellas, que incluyen nuevas formas de estimar las relaciones extremales a partir de datos empíricos.

[37] Lau en [17], p. 179 - véase también Jacobsen en [17], p. 172 - se refiere a esta como una "función de medida" (su denotación clásica en la teoría de conjuntos convexos) ya que no tiene las propiedades habituales de una función de distancia. Véase, por ejemplo, el Apéndice A en [81].

7. Resumen y conclusión

En este documento hemos proporcionado una variedad de formas de evaluar la eficiencia de las unidades dde toma de decisiones (UTDs) en los programas públicos a fin de mejorar la planificación y el control de estas actividades. Esto se inició con una nueva definición de eficiencia, en la Sección 2, que hemos relacionado con conceptos ingenieriles y económicos. También suministramos formas operativas[38] e interpretaciones para aplicaciones reales. Además, se introdujo un nuevo tipo de función de producción y nuevos métodos de obtención de estimaciones a partir de datos empíricos de manera que nos han permitido presentar una variedad de conceptos económicos de manera nueva y potencialmente útil. Esto se hizo con nuevos modelos e interpretaciones, sin duda, pero sin alterar ninguno de los principios de optimización que constituyen la base esencial de estos conceptos.

Podemos (y extenderemos) estos desarrollos en varias formas adicionales.[39] Aquí podemos concluir mejor con algunos comentarios sobre las posibles limitaciones y alternativas a estos enfoques.

Puede surgir una limitación por la carencia de datos disponibles en los niveles individuales de UTD. Esto es probable que no sea tanto un problema de las aplicaciones en el sector público, como en el sector privado. Además, el término UTD puede tener una considerable flexibilidad en su interpretación, por ejemplo cuando la indisponibilidad de los datos puede hacer que sea deseable pasar al distrito escolar en lugar del nivel individual de la escuela y, de hecho, incluso resulte necesario pasar a los niveles estatales de oficina de educación cuando los datos no están disponibles a nivel de distrito. Después de todo, el objetivo es medir la eficiencia de utilización de los recursos en las combinaciones que estén presentes (holgadas o ajustadas) en las organizaciones así como las tecnologías utilizadas.[40] Esto también sugiere una estrategia de investigación en la que dados los resultados obtenidos en un nivel, puede obtenerse una base sistemática para pasar a otros niveles, planteando preguntas pertinentes y exigiendo una justificación adicional de los funcionarios responsables por las ineficiencias que se descubran.

En cuanto a las aplicaciones del sector privado, el argumento de nuestra medida de eficiencia propuesta comienza a debilitarse en la medida en que haya presente competencia. En particular, comienza a debilitarse tan pronto como está presente la libertad para asignar recursos de una "industria" a otra (tal vez en una región alejada). La evaluación de tales posibilidades implicaría introducir precios, u otros esquemas de ponderación, para evaluar alternativas que de otro modo no serían comparables.

Aunque nuestras medidas no están diseñadas para este tipo de aplicación, están diseñadas para programas del sector público en donde los gerentes de muchas UTDs no tienen libertad para desviar recursos a otros programas simplemente porque sean más rentables o más atractivos. Nuestra medida tiene como objetivo evaluar los logros, o las posibilidades de con-

---

[38] Es decir, hemos proporcionado modelos y procedimientos computacionales.

[39] Por ejemplo, como ya se ha observado, los desarrollos anteriores pueden extenderse a funciones que son Cobb-Douglas a trozos.

[40] El hecho de que en general se incluyan tales consideraciones organizativas como parte de la función de producción en las investigaciones empíricas entra en una tradición económica que se remonta al menos hasta Sune Carlson. Véase [6]. (En futuros trabajos, sin embargo, mostraremos cómo los conceptos que hemos introducido aquí se pueden extender para usar en la separación de estas características gerenciales en datos observados de otros aspectos de una tecnología de producción).

servación de recursos, para cada UTD con recursos asignados. En terminología golfística es, por así decirlo, una medida de "distancia" en lugar de "dirección" con respecto a lo que se ha logrado (y podría lograrse).[41]

Referencias

[1] J D.J. Aigner, T. Ameiya, P.J. Poirier, On the estimation of production frontiers: maximum likelihood estimation of the parameters of a discontinuous density function, Internat. Econ. Rev. XVII, (2) (1976) 377-396.

[2] D.J. Aigner and S.F. Chu, On estimating the industry production function, American Economic Review LVII| (1968) 826-839.

[3] R.F. Alioto and J.A. Jungherr, Operational PPBS for Education (Harper and Row, New York, 1971).

[4] K.J. Arrow and F.H. Hahn, General Competitive Analysis (Holden-Day, San Francisco, 1971).

[5] C.R. Bector, Duality in Linear Fractional Programming, Utilitas Mathematica 4 (1973) 155-168.

[6l S. Carlson, A Study on the Pure Theory of Production, (Kelley and Millman, Inc., New York, 1956).

[7] A. Charnes and W.W. Cooper, An explicit general solution in linear fractional programming, Naval Research Logistics Quarterly 20 (3) (1973).

[8] A. Charnes and W.W. Cooper, Management Models and Industrial Applications of Linear Programming (Wiley, New York, 1961).

[9] A. Charnes and W.W. Cooper, Managerial economics - past, present, future, Journal of Enterprise Management (forthcoming).

[10] A. Charnes and W.W. Cooper, Programming with linear fractional functionals, Naval Res. Logist. Quart. 9 (3, 4) (1962) 181-185.

[11] A. Charnes, W.W. Cooper and E. Rhodes, Expositions, interpretations, and extensions of Farrell efficiency measures, Management Sciences Research Group Report, Pittsburgh: Carnegie-Mellon University School of Urban and public Affairs (1975).

[12] A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision making units with some new production functions and estimation methods, Center for Cybernetic Studies Research Report CCS 276, Austin, TX, University of Texas Center for Cybernetic Studies (1977).

[41] El término dirección debe entenderse en plural para el caso de productos múltiples. (Aquí nos interesa elegir la dirección que se podría llevar a cabo a nivel legislativo, incluyendo la cuestión de si el programa debería llevarse a cabo en absoluto.)

[13] A. Charnes, W.W. Cooper and A. Schinnar, Transforms and approximations in cost and production function relations, Research Report CCS 284, Austin, TX, University of Texas Center for Cybernetic Studies (1977).

[14] Encyclopedia Americana (Encyclopedia Americana Corporation, New York, 1966) 352-353.

[15] M.J. Farrell, The measurement of productive efficiency. J. Roy. Statist. Soc. Set. A, III (1957) 253-290.

[16] M.J. Farrell and M. Fieldhouse, Estimating efficient production functions under increasing returns to scale, J. Roy. Stat. Soc. Ser. A, II (1962) 252-267.

[17] M.D. Intriligator and D.A. Kendrick, eds., Frontiers of Quantitative Economics, IL North-Holland, Amsterdam (1974).

[18] R. R. Jagannathan, Duality for nonlinear fractional programs, Z. Operations Res., 17, Physica-Verlag, Wurzburg (1971) 1-3.

[19] J. Johnston, Statistical Cost Analysis, McGraw-Hill, New York (1960).

[20] K. Lancaster, Consumer Demand, Columbia University Press, New York (1971), 40.

[21] E. Rhodes, Data Envelopment Analysis and Related Approaches for Measuring the Efficiency of Decision Making Units with an Application to Program Follow Through in U.S. Education, Ph.D. thesis, Carnegie-Mellon University School of Urban and Public Affairs, Pittsburgh (1978).

[22] A. Robinson, Non-Standard Analysis (North-Holland, Amsterdam, 1966).

[23] P.A. Samuelson, Prices of Factors and Goods in General Equilibrium, Rev. Econom. Stud. 21 (1953-1954) 1-20.

[24] K. Sato, Production Functions and Aggregation (North-Holland, Amsterdam, 1975).

[25] S. Schaible, Parameter-free Convex Equivalent and Dual Programs of Fractional Programming Problems, Z. Operations Res., 18 (1974) 187-196.

[26] R.W. Shephard, Cost and Production Functions (Princeton University Press, Princeton, 1953).

[27] R.W. Shephard, The Theory of Cost and Production Functions (Princeton University Press, Princeton, 1970).

[28] G.J. Stigler, The Xistence of X-Efficiency, Am. Econorn. Rev. (March 1976) 213-216.