
Indicadores bibliométricos y econométricos en la evaluación de instituciones científicas*

Ronald Rousseau¹

Resumen

Este artículo consta de 2 partes. La primera describe la evaluación de una pequeña universidad a partir de datos cuantitativos, con el objetivo principal de analizar la investigación con visibilidad internacional. La segunda parte muestra cómo un método econométrico (DEA: data enveloping analysis) puede aplicarse entre otros aspectos como medio de evaluación para determinar los recursos disponibles. Asimismo se demuestra cómo un cuerpo teórico que combina la bibliometría, la cuantimetría y la econometría puede aplicarse a problemas concretos.

Descriptores DeCS: CIENCIOMETRIA; INVESTIGACION; EVALUACION; ECONOMETRIA/métodos; FACTOR DE IMPACTO; ANALISIS DE CITAS; PRODUCCION CIENTIFICA; PUBLICACION ELECTRONICA; BIBLIOMETRIA; UNIVERSIDADES/eficiencia; UNIVERSIDADES/efectividad.

El presente artículo se ha dividido en 2 partes. La primera describe la evaluación de una pequeña universidad de Flandes, Bélgica, conocida como LUC a partir de datos cuantitativos, con el objetivo principal de analizar la investigación con visibilidad internacional.

La segunda parte muestra cómo puede utilizarse un método econométrico como instrumento de evaluación para determinar los recursos disponibles. También se demuestra cómo un cuerpo teórico que combina la bibliometría, la cuantimetría y la econometría puede aplicarse a problemas concretos,^{1,2} y que, por tanto, la bibliometría puede ser la solución adecuada para determinados problemas.³

Primera parte: Estudio cuantimétrico de las publicaciones científicas de LUC

Este estudio, autorizado por el Rector y por el Consejo de Investigaciones de LUC, investiga la posición internacional de la Facultad de Ciencia y Medicina de esa insti-

tución,⁴ y describe en particular en el período de 1981 a 1993:

- La dimensión del potencial investigativo de todos los grupos dedicados a esa actividad, medida en términos de equivalencia persona-año (person-year equivalents, PYE) consagrada a la investigación científica.
- Los resultados de la investigación, en particular, la posición internacional evidenciada por las publicaciones en revistas científicas; especialmente, la proporción en que estas publicaciones fueron citadas en la literatura profesional internacional.
- Las observaciones hechas por los científicos en relación con los métodos cuantimétricos y sus resultados. Para poder comparar los resultados de LUC con los de otras instituciones, el método empleado se aplicó también en la evaluación de otras universidades.^{5,6}

Premisas, objetivos y advertencias

La aplicación de los métodos cuantimétricos para la evaluación de investigaciones científicas^{7,8} se basa en las siguientes premisas:

* Trabajo presentado en el Seminario sobre Evaluación de la Producción Científica, realizado en São Paulo por el Proyecto SciELO, del 4 al 6 de marzo de 1998.

¹ UIA, Universiteitsplein 1, Wilrijk, Belgium.

- a) El progreso se alcanza mediante el trabajo de los científicos.
- b) En su trabajo, estos científicos se basan en la obra de colegas y de precursores en su campo.
- c) Se publican los resultados de este trabajo.
- d) Se someten a la apreciación de pares iguales.
- e) Los científicos muestran en sus publicaciones cómo se han basado en el trabajo previo de otros, al mencionar en sus textos una lista de referencias.
- f) Las revistas científicas desempeñan un papel esencial en la comunicación entre colegas científicos. Esta es la principal razón por la que la literatura científica es una representación de la actividad científica y de la red de relaciones entre subcampos.
- g) El número de publicaciones de un grupo de investigación, puede considerarse un indicador de su producción científica.
- h) El número de veces que la obra de un grupo es citada por otras publicaciones, da la medida del impacto y de la visibilidad internacional de esa obra.
- i) En el sistema global de revistas, se puede distinguir entre las "centrales", es decir, las revistas internacionales importantes, y las periféricas que son de orientación más regional.
- j) El Instituto de Información Científica (ISI) abarca por medio de su Índice de Citas Científicas (SCI), el grueso de las revistas internacionales importantes en las ciencias puras, aplicadas y médicas.

El índice de Citas de las Ciencias Sociales (SSCI) tiene la misma función para el área de las ciencias sociales.

Los indicadores cuantitativos no están destinados a sustituir a los especialistas, sino más bien a hacer que las investigaciones sean visibles y analizables, de manera que los especialistas tengan una adecuada información a su disposición y por tanto puedan formular una opinión mejor fundamentada.^{9,10} Estos indicadores son más eficaces cuanto más alto sea el nivel de agregación.¹¹

Los científicos que laboran en las universidades tienen varias responsabilidades: además de la enseñanza y la investigación, tienen una obligación para con la sociedad en su conjunto. El método cuantitativo considera sólo el aspecto investigativo, esencialmente el alcance de las contribuciones de un grupo de investigación en relación con el desarrollo de nuevos conocimientos en el frente investigativo. Esto implica que muchas de las actividades normales de los investigadores, no se toman en consideración.

Los resultados de esta investigación se distribuyeron en tres versiones. La primera contiene los resultados de cada grupo por separado. Cada grupo de investigación recibió un informe que contenía sólo sus resultados, además de una copia con la descripción de los resultados globales de LUC. La segunda versión que sólo contiene los resultados globales de LUC, se repartió a los grupos de investigadores y al Consejo de Investigaciones. Finalmente hay una versión oficial (en holandés) dada a conocer públicamente (tras ser apro-

bada por el Consejo de Investigaciones), para que los interesados en los métodos y en los resultados puedan consultarla. Esta versión oficial ha sido traducida al inglés⁴ y todos los datos en ella contenidos se han mantenido en el anonimato.

Indicadores de entrada

Se estimó la capacidad investigativa de cada grupo, expresada en equivalentes persona-año en la investigación (PYE-Res).

Para obtener esta cifra, se observaron las siguientes reglas.

Se introdujo un parámetro de investigación que expresaba una fracción de horas trabajadas en una investigación, para lo cual se emplearon los mismos porcentajes de la OECD, o sea, en cuanto a los miembros del personal pagado por la propia universidad, este parámetro se fijó en 0,4(40 %); en lo que respecta a los investigadores pagados por la Fundación Nacional de la Ciencia y otras organizaciones similares, el parámetro se estableció en 1,0 (100 %); el personal académico empleado por un período menor al 100 % se contabilizó en proporción al porcentaje de su tiempo de trabajo (los que tenían un trabajo a tiempo parcial su parámetro se valoró en 0,2). Si el porcentaje del régimen de trabajo es inferior al 20 %, las personas en este caso sólo se consideran como profesores y su parámetro investigativo se calcula en 0,0. Con el uso de estos parámetros y el número de personas activas en cada grupo de investigación, se calculó el indicador PYE-Res/año, es decir, los equivalentes promedio persona-año en labores investigativas durante el período 1981-1993.

Indicadores de producción, productividad e impacto

La producción de un grupo de investigación se expresa en su cantidad de publicaciones. La productividad se expresa como el número de publicaciones por equivalentes persona-año en las investigaciones.

El impacto se ha medido sobre la base de la cantidad de citas por publicación y está relacionada con un patrón internacional (un promedio mundial).

Indicadores de producción (PA, Pt, P, Pf)

Para cada grupo de investigación se ha determinado el número total de publicaciones y el número de publicaciones incluidas en el archivo LUC-JCR (Journal Citation Reports), durante el período 1981-1993. Estos indicadores se identifican con los símbolos PA y Pt. El número de publicaciones incluidas en el JCR de 1981 a 1993 se denota con la letra P. El

SCI utiliza una clasificación de las publicaciones según se trate de artículos originales, cartas, notas, artículos de revisión, resúmenes, conferencias, editoriales, reseñas de libros y correcciones. Solamente las cuatro primeras categorías mencionadas se consideran para los parámetros Pt y P.

A menudo sucede que un grupo de investigación publica artículos en cooperación con científicos de otros grupos, tanto dentro como fuera de la universidad.¹² Vale preguntar si estas publicaciones pueden o no ser completamente atribuidas a cada grupo de investigación participante. Esto constituye un problema, en especial cuando las publicaciones se relacionan con los datos de entrada, como son los equivalentes persona-año para una investigación.

De hecho, los artículos que surgen de una colaboración se basan en parte en los esfuerzos y recursos de otros grupos. Por esta razón, se han utilizado aquí dos esquemas de contabilización: el usual y directo (una publicación es una unidad) y el fraccionado (tomando en consideración la colaboración). Si, de acuerdo con las indicaciones sobre publicaciones encontradas en los archivos del SCI, un artículo es publicado por miembros de dos grupos, cada grupo recibe la mitad del crédito. Si son tres grupos, entonces cada uno recibe la tercera parte del crédito y así sucesivamente. A las publicaciones calculadas por el método fraccionario e incluidas en el archivo LUC-JCR, se les asigna el símbolo Pf. Este indicador sólo se ha utilizado para medir la productividad de un grupo de investigación o de una facultad.

Obsérvese que los indicadores P, Pt y Pf sólo se relacionan con artículos del LUC-JCR, es decir, con artículos originales, cartas, notas y artículos de revisión publicados en un número de la JCR y que tienen la dirección de LUC.

Indicadores de productividad (PA/PYE-Res por año; Pf/PYE-Res por año)

La productividad se mide utilizando las relaciones PA/PYE-Res y Pf/PYE-Res (número total de publicaciones y número fraccionado en JCR por equivalente persona-año en actividades de investigación) en el lapso de un año. Como el objetivo de este informe es determinar la posición internacional de LUC, sería conveniente comparar la productividad de los grupos de investigación y de las facultades con un patrón internacional. Esto permitiría determinar si la productividad es alta o baja desde un punto de vista global (internacional). Sin embargo, no es posible en estos momentos obtener datos fiables para hacer tal comparación.

La comparación de la productividad entre diferentes grupos de investigación no es por lo general significativa. Además del hecho de que esto sólo determinaría la posición relativa local de un grupo, existe la complicación de que los hábitos de publicación varían marcadamente entre los distintos campos y departamentos científicos. Por ejemplo, tiene poca importancia comparar la productividad de un grupo experimental con la de un grupo dedicado a la ciencia aplica-

da. La principal función de este indicador es encontrar los grupos de investigación con una productividad muy baja, es decir, grupos que publican muy poco con respecto a su capacidad investigadora.

Indicadores de impacto (Ct, Ctex, C, Cex, CpP, CexpP, CpP/FCSm, CpP/JCSm, JCSm/FCSm, PSt, PS)

Estos indicadores sólo tienen que ver con artículos del LUC-JCR. Para cada científico se ha determinado el número total de citas durante el período 1981-1993 (y esto es con respecto a los artículos publicados en este período). Esto se ha identificado con el símbolo Ct, en el que también se incluyen las autocitas. Para dar una idea de la cantidad de autocitas, se determinó también el número total de citas sin incluir autocitas, y el porcentaje de autocitas (símbolos Ctex, PSt). Estos indicadores se determinaron para un período de cuatro años tras la publicación (símbolos C, Cex, PS). Además se obtuvo *la cantidad promedio de citas por artículo con autocitas y sin estas* (símbolos CpP y CexpP).

Por regla general, un grupo publica artículos en determinadas revistas, las cuales constituyen la colección de revistas de ese grupo. Es posible obtener el factor de impacto para cada revista de esta colección, es decir, el número de veces promedio en que se cita el artículo publicado en esa revista. Obsérvese que para el cálculo del factor de impacto de una revista, se toman en consideración todos los artículos y no sólo aquéllos publicados por el grupo de investigación que se evalúa. En este informe, el factor de impacto se denomina JCS (Journal Citation Score: puntuación de las citas de la revista).

El uso de factores de impacto de revistas posibilita formular una media ponderada del factor de impacto para la colección de revistas donde publica el grupo (símbolo JCSm). Ejemplificaremos esta construcción de la siguiente manera: supongamos que un grupo publicó en un año determinado 3 artículos en la revista A y 2 en la revista B; luego supongamos que los artículos de la revista A tuvieron como promedio 6 citas (JCS = 6) y los de la B, 2 (JCS = 2). La media ponderada del factor de impacto para la colección de revistas de este grupo en ese año sería (JCSm): $(3 \times 6 + 2 \times 2) : 5 = 4.4$.

Los factores de impacto pueden calcularse de diversas formas. La utilizada en este informe difiere de la oficial, es decir de la del ISI. La noción del factor de impacto siempre se refiere a una cantidad promedio de citas de artículos publicados en cierta revista. El factor de impacto del ISI se refiere al número promedio de citas obtenido durante el segundo o tercer año posterior a la publicación. Para presentarlo de una forma más precisa: el factor de impacto del ISI para una revista X en 1998 se obtiene al dividir la suma de las citas en el año 1998 de los artículos publicados en esta revista en los años 1996 y 1997 entre la suma de la cantidad de artículos publicados en esa revista X durante esos mismos años.

Nuestro trabajo se ha limitado a un período predeterminado, llamado *un período de 4 años*,¹³ que consiste en el año de publicación (año uno) y los 3 años siguientes. Una consecuencia de esta elección del período de 4 años es que tan sólo las publicaciones del período 1981-1989 pueden analizarse en su totalidad.

Los indicadores mencionados, son indicadores básicos que responden a preguntas como: ¿cuál es la frecuencia con que se citan los artículos dentro del conjunto de revistas del SCI? y ¿cuál es la media ponderada del factor de impacto de las revistas en las que se publican las investigaciones? Estos indicadores no responden a la pregunta de si un grupo de investigación es o no más citado que otros grupos similares y si la puntuación de las citas en la revista del grupo es alta o baja vista desde una perspectiva internacional. Una comparación directa entre los grupos o los departamentos de una universidad, basada en números absolutos de citas (C), de citas por publicación (CpP) o el factor de impacto de la colección de revistas (JCSm) no es por lo general significativa, dada la diversidad de hábitos a la hora de citar en diferentes campos de la ciencia. En campos como la inmunología, las listas de referencias son mucho más largas que en matemática o en computación. Por lo tanto, hemos optado por un método donde al menos para el análisis del impacto se ha establecido un nivel de referencia internacional. Las puntuaciones de los grupos de investigación de LUC se comparan entonces con este nivel de referencia internacional. Hemos calculado, por subcampo, la cantidad promedio de citas que recibe un artículo de ese campo, lo cual se denomina *promedio global por campo* y se denota como FCS (field citation score). Para la clasificación de revistas por subcampos, se parte de la clasificación de revistas en categorías según el ISI. Luego se comparan las puntuaciones de las citas de LUC con las existentes en el mundo y entonces puede establecerse si los artículos de un grupo de investigación son más o menos citados que el promedio global imperante en el campo donde opera el grupo de investigación. Por lo general, un grupo trabaja en más de un campo y, por consiguiente, la media ponderada que se utiliza (símbolo FCSm) es similar al cálculo de la media ponderada del factor de impacto (JCSm).¹⁴

Se elaboraron 3 importantes indicadores sobre la base de las variables CpP, JCSm y FCSm:

- a) La relación entre el impacto obtenido y el impacto de la colección de revistas: $CpP/JCSm$. Aquí se compara la puntuación del promedio de citas de un grupo de investigación (CpP) con la media ponderada del factor de impacto de las revistas en la cual publica el grupo (JCSm). Este indicador permite determinar si la obra de un grupo -dentro de un parámetro de citas en 4 años-¹ es más citada que el promedio. El poder evaluativo de este indicador es limitado por el hecho de que no se toma en cuenta el nivel de citación de la colección de revistas.

Un grupo que publica en revistas menos importantes- con un bajo factor de impacto- y que es citado con la misma frecuencia promedio que se citan los artículos de su colección de revistas, es clasificado por este indicador en el mismo nivel que un grupo que publica en las principales revistas con un alto factor de impacto y obtiene una puntuación de citas promedio igual a la de su colección de revistas.

- b) La relación entre el impacto obtenido y el impacto global promedio de los campos en los cuales trabaja el grupo: $CpP/FCSm$. Este indicador compara la puntuación de citas promedio de la obra del grupo con el promedio internacional de citas de los campos en los que se desenvuelve el grupo. Una relación mayor que 1 significa que el grupo es citado con más frecuencia que el promedio global de sus subcampos. Este indicador es el más apropiado para medir la posición internacional de un grupo de investigación.
- c) La relación entre factor de impacto promedio de la colección de revistas con el impacto global promedio de los campos en que actúa el grupo: $JCSm/FCSm$. Este indicador compara la media ponderada del factor de impacto de la colección de revistas con el promedio global de citas del campo en que participa el grupo. Si el $JCSm/FCSm$ es mayor que 1, esto significa que el grupo publica en revistas que tienen un factor de impacto alto con respecto al promedio global (en el campo).

Por su puesto que estos 3 indicadores no son independientes. El conocimiento de dos de ellos determina el tercero.

Análisis de tendencias

Una limitación de los indicadores mencionados es que ofrecen un valor promedio por encima período estudiado. Así, este cuadro se ha complementado con una serie de análisis de tendencias.

Rousseau¹⁵ aborda otro método para los análisis de tendencias. Los indicadores se estudian aquí sobre una base anual para hallar la evolución del impacto del grupo. Los gráficos que se muestran a continuación, reflejan la evolución anual y los promedios de 3 años. Considerando que la evolución de un año al otro ocurre a menudo de forma irregular, se recomienda concentrar la atención en la evolución de los promedios de 3 años.

1. Número de publicaciones por año
La figura muestra el total de trabajos publicados anualmente por un grupo de investigación o facultad. Todas las publicaciones, incluidas las categorizadas en (S), C y P, se toman en cuenta.

2. Número de artículos JCR por año
Sólo se muestran los artículos publicados en las revistas incluidas en JCR.
3. Productividad.
Esta figura representa la relación entre el número de publicaciones y el factor persona-año equivalente en la investigación (PYE-Res).
4. Productividad de JCR
Esta figura muestra lo mismo que el anterior excepto que ahora sólo se analizan las publicaciones de JCR.
5. Número total de citas por año
Aquí se muestra el número total de citas (en las revistas de JCR) que el grupo o facultad ha obtenido por año. Sólo se analizan las citas de publicaciones de 1981 a 1993.
A medida que se incrementa el número acumulativo de artículos citables durante el período, es normal que también aumente este gráfico. Si ello no ocurre así, constituye un punto negativo del grupo.
6. Número de citas por publicación-publicaciones que tienen 4 años como máximo
Esta figura muestra los mismos datos que el anterior, con la diferencia de que las publicaciones tienen a lo sumo 4 años. En consecuencia, los primeros 3 años están incompletos.
7. Número de citas de artículos JCR con 4 años como máximo
Otra figura similar a las 2 anteriores, pero que sólo considera las publicaciones JCR. De ahí que se muestran las publicaciones por artículo y los 3 primeros años están incompletos.
8. Impacto de citas con 4 años
Esta figura, por artículo del JCR, el número total de citas de artículos publicados en el primer año de un período de 4 (el año reflejado en el eje horizontal). Esto significa que los últimos 3 años (1991-1993) están incompletos. Por lo tanto, es lógico que estas figuras disminuyan al final. Antes de este período (1991-1993) se utilizaba un parámetro de 4 años.
9. Impacto en relación con la colección de revistas.

Los valores de los puntos de esta figura son los cocientes de los números de la figura anterior, divididos por el número promedio de citas obtenido por un artículo publicado en la misma revista durante 4 años. Si el grupo obtiene la misma puntuación, que su colección de revistas, este cociente es 1.

10. Impacto en relación con el promedio mundial de citas
Los valores de los puntos encontrados en esta figura son los cocientes de los números de la figura 8 y del número promedio de citas que un artículo en el mismo campo obtiene durante el mismo período de 4 años (en la misma categoría de JCR)

Cobertura del SCI

Los indicadores de impacto se basan enteramente en artículos con la dirección de LUC, publicados en revistas cubiertas por el SCI. La hipótesis central es que el SCI es un instrumento adecuado para seleccionar artículos publicados en las principales revistas internacionales.¹⁶ En esta sección se analizará con más profundidad la validez de esta afirmación. A tales efectos, se elaboraron 2 indicadores adicionales:

a) Pt/Pjtot

Este indicador brinda el porcentaje de artículos en el archivo LUC-JCR(Pt) con respecto al total de artículos de revistas (Pjtot).

Si este porcentaje se aproxima al 100 %, ello significa que el grupo de investigación publica sus artículos principalmente en las revistas cubiertas por JCR. En ese caso, el SCI tiene una buena cobertura de la producción de este grupo.

b) Pjtot/AA

Este indicador brinda el porcentaje de artículos de revistas con respecto al número total de publicaciones (PA). Si este porcentaje es cercano al 100 %, ello significa que el grupo publica mayormente artículos y que es menor la importancia de otros tipos de publicaciones.

¿Cómo deben interpretarse estos indicadores? Hablando en términos comparativos, si un grupo de investigación publica una gran cantidad de artículos en revistas no cubiertas por JCR, ello puede interpretarse de dos formas. Si se asume que la cobertura del SCI es adecuada y realmente abarca todas las principales revistas, entonces se ha de llegar a la conclusión de que obviamente este grupo no publica en revistas im-

Fig. 1. Número de publicaciones por año

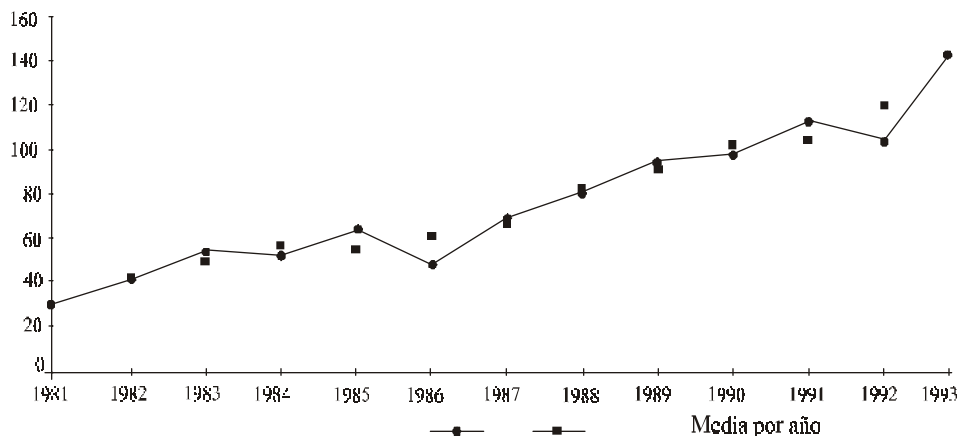


Fig. 2. Número de artículos JCR por año

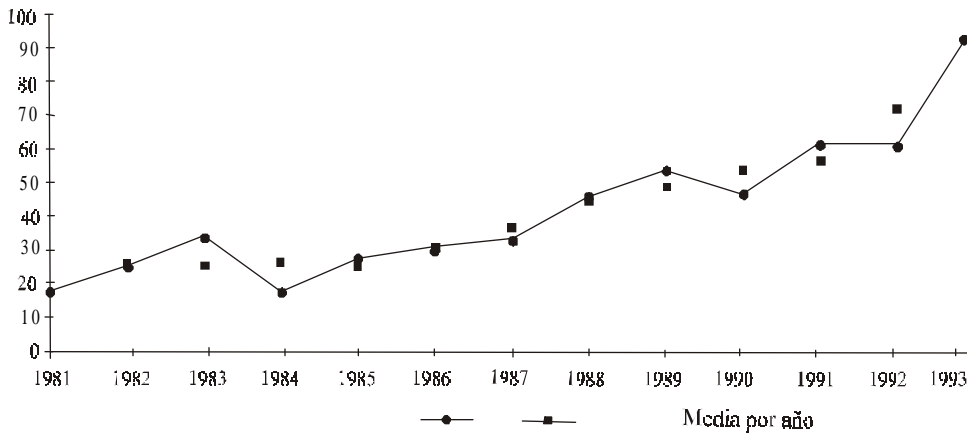


Fig. 3. Número total de citas por año

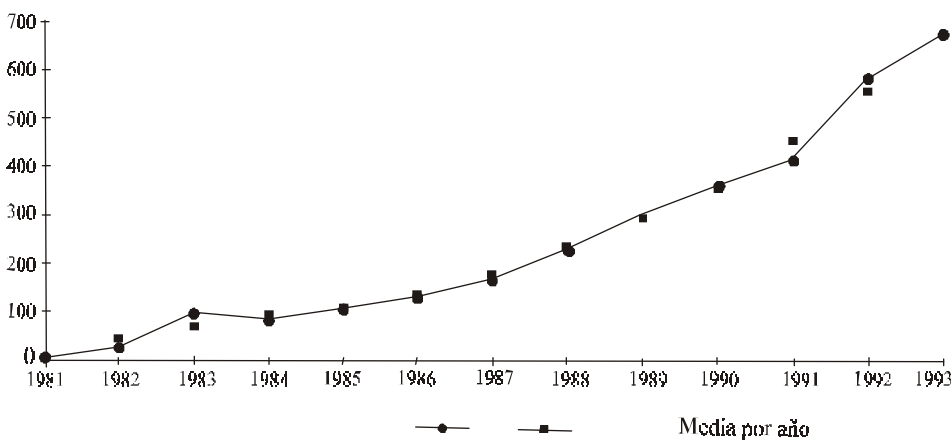
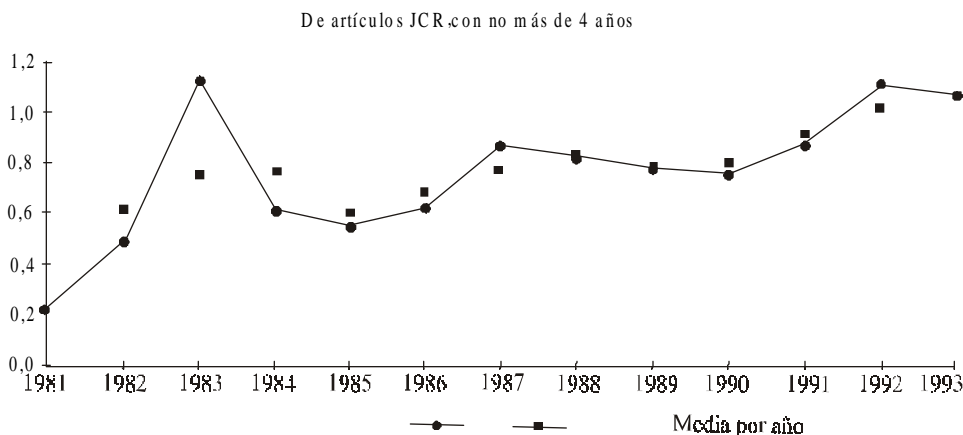


Fig. 4. Número de citas por publicación



portantes. En ese caso, la relación Pt/P_{jtot} es un indicador de la posición internacional del grupo como lo muestra su comportamiento en cuanto a la publicación de trabajos, es decir, de la elección de revistas para publicar artículos (o puede ser que los artículos del grupo sólo se acepten en revistas de menos importancia). Por eso una relación pequeña significa una posición internacional débil.

Por otra parte, se puede asumir que la cobertura del SCI es inadecuada al no considerar muchas revistas importantes. Si se asume asimismo que el grupo publica en revistas importantes, entonces la relación Pt/P_{jtot} se convierte en un indicador de cobertura del SCI con el grupo de investigación (departamento, facultad) como norma.

Puede que surja un problema con la adecuación del SCI en aquellos campos donde la publicación de alto nivel no está limitada a un grupo relativamente pequeño de las mejores revistas internacionales en idioma inglés, pero en él existe un amplio espectro de revistas de orientación local (es decir, dirigidas a una nación o a un grupo idiomático) que publican artículos de gran calidad. En general, el problema no está en las ciencias. Sin embargo, este aspecto ha de considerarse en la interpretación de los resultados.

Se pueden hacer observaciones análogas con respecto a la relación P_{jtot}/PA . Si un grupo tiene una relación P_{jtot}/PA baja, ello puede significar que el grupo no quiere o no es capaz de publicar en revistas. Sin embargo, también puede significar que en ese campo de investigación, las revistas tienen un menor papel en la divulgación de los conocimientos científicos.

Discusión de estos resultados

La Facultad de Ciencia y Medicina publicó 1 038 trabajos durante el período 1981-1993 que significa un promedio de 80 trabajos al año; el 82 % de estas publicaciones son artículos de revistas.

Por consiguiente, el 18 % se refiere a otros tipos de publicaciones, lo que demuestra que en la Facultad de Ciencia y Medicina son también

importantes los anales de congresos y las contribuciones en libros editados. De estos artículos de revistas, el 72 % pertenece a las categorías de la JCR: artículos originales, cartas, artículos de las categorías de revisión y notas. En general, no hay problema de cobertura para el SCI porque una gran parte de las demás publicaciones son resúmenes de revistas de JCR.

Fig. 5. Impacto

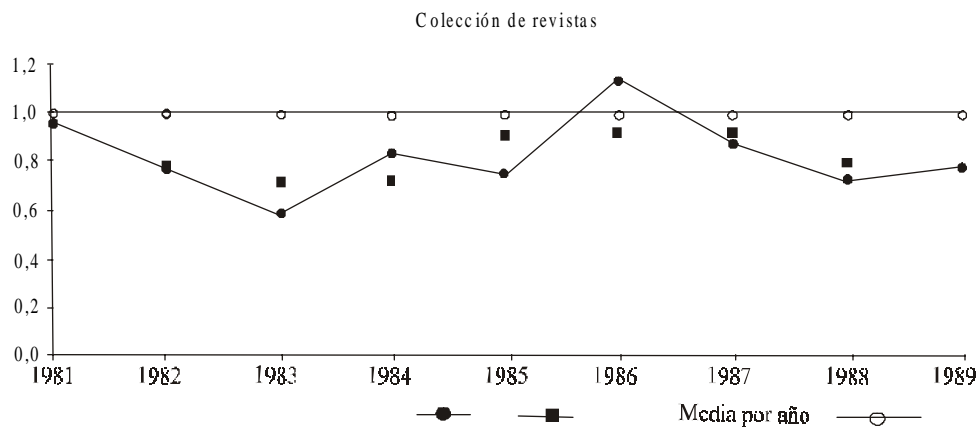
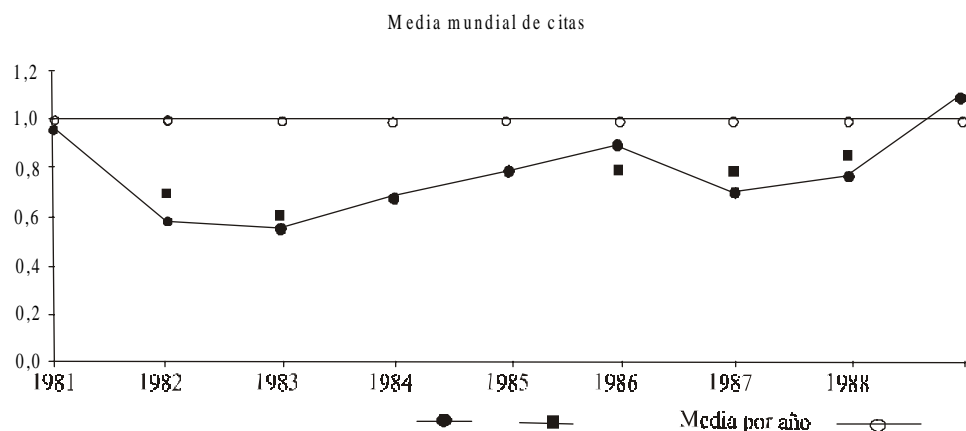


Fig. 6. Impacto



Los 586 artículos de la Facultad en JCR se citaron 2 915 veces. El porcentaje de autocitas fue igual al 38 %, que es bastante alto. En relación con el impacto a corto plazo, se observa que estos artículos (publicados en el período 1981-1989) se citaron como promedio 3,38 veces durante los primeros 4 años tras la publicación (39 % son autocitas).

El factor de impacto promedio de la colección de revistas de la facultad fue 4,77 y la puntuación promedio mundial de citas fue 4,83, lo que demostró que la Facultad en su conjunto publica dentro del nivel promedio. El impacto promedio de la Facultad se sitúa justo por debajo del promedio mundial ($3,88/4,83=0,8$) donde se toman en cuenta los márgenes de errores. El impacto de la colección de revistas fue de 0,81.

Los análisis de las tendencias demuestran una trayectoria irregular, aunque terminan de forma ascendente. Se puede decir con toda seguridad que en 1989 la Facultad de Ciencias y Medicina alcanzó el promedio mundial.

Discusiones con los científicos

Consultamos a los investigadores durante dos rondas de validación, donde se les dio la posibilidad de corregir los datos numéricos.

Adicionalmente se compararon los resultados de los análisis cuantitativos con la visión que los propios científicos tenían de su Facultad y de su grupo de investigación. Durante las entrevistas personales, se recopiló la información que tuviera alguna importancia concreta para dar una interpretación correcta a los resultados cuantitativos. Estas interpretaciones, puestas a la consideración del Consejo de Investigaciones para su procesamiento, hicieron de este informe un instrumento válido a la política de investigaciones en la Universidad. Los investigadores formularon una serie de observaciones importantes, a veces críticas, pero nunca totalmente negativas.

Los debates que sostuvo con varios científicos fueron, a mi modo de ver, muy instructivos para ambas partes. Como se ha mencionado anteriormente, son sólo los científicos quienes pueden valorar correctamente los resultados obtenidos. Un aspecto común de todas las discusiones fue que los científicos se sintieron un tanto disgustados con los resultados promedio (promedios a lo largo de todo el período), pero satisfechos con los análisis de las tendencias que muestran el avance por ellos logrando. A menudo señalaron que los programas de investigaciones más extensos requerían de un tiempo de incubación para poder ponerse en marcha. Asimismo enfatizaron que, además de las citas "objetivas" los contactos personales y las solicitudes de orientación eran también importantes para sentirse apreciados por sus iguales.

Como respuestas a preguntas más específicas, se registraron los comentarios siguientes:

- Los métodos cuantitativos son útiles siempre que se tomen en cuenta las diferencias de recursos materiales y humanos existentes y entre los grupos de investigación.
- La visibilidad es uno de los aspectos importantes de la investigación científica. Un estudio de las citas brinda una interesante retroalimentación con respecto a esto.
- Un estudio cuantitativo fiable informa a los miembros de la Universidad de manera que puedan tratar de mejorar su visibilidad.¹⁷
- Muchas personas subrayaron la naturaleza objetiva de los datos sobre citas, pero también apuntaron su unilateralidad, ya que ser miembro de una universidad exige mucho más que la investigación científica tal como se publica en las revistas cubiertas por JCR.

- Un grupo de investigación afirma que una publicación debe ser el resultado natural de la investigación y no a la inversa. Algo parecido en este sentido señala otro grupo con su declaración de que publicar más no debe ser motivo de una política publicitaria, sino inspirarse en la necesidad de divulgar los resultados de la actividad científica sostenida.
- Hay quienes consideran el estudio econométrico como un incentivo para que el Consejo de Investigaciones evalúe el trabajo de grupos dedicados a esta actividad.
- Se detectaron problemas de coberturas en las áreas de la computación y de la estadística. El problema en esta última radica en que algunas revistas no están cubiertas por el ISI; en lo que respecta a la computación, el problema se relaciona más con el hecho de que en este campo los científicos tienden a publicar con mayor frecuencia en las actas de las conferencias, las que presentan por lo general criterios de selección muy restringidos.
- Algunos científicos quisieran saber cómo se evaluarían en el futuro inmediato las revistas electrónicas y las citas en ellas contenidas.¹⁸
- Un grupo de investigación no visualiza el propósito de un estudio econométrico, en particular en una pequeña universidad como LUC, donde estos grupos son también pequeños. Para este grupo, es imposible el uso a ciegas de los resultados de este análisis con fines administrativos.
- Al parecer es más difícil obtener buenas puntuaciones cuando se publica en revistas con un alto nivel de impacto.

Comentario personal: Las revistas de un alto impacto pertenecen a menudo a campos de las ciencias conocidos como "gran ciencia". Particularmente dentro de estos campos, la pequeñez de LUC representa hoy día un aspecto adverso, lo que explica al menos en parte los resultados más bien desalentadores de algunos grupos de investigación.

Conclusión de la primera parte

Estamos convencidos que para aquellos que utilizan de forma responsable los estudios de evaluación de la ciencia a los efectos de controlar la política investigativa, nuestro informe les resultará una herramienta útil. En cuanto a los científicos de LUC, el informe les servirá como una forma de retroalimentación en relación con una de sus actividades principales.

Segunda parte: Data envelopment analysis como herramienta para elaborar indicadores econométricos

La finalidad de esta parte es mostrar la factibilidad del data envelopment analysis (DEA) en las aplicaciones

cienciométricas e introducir al lector en esta técnica. Como ejemplo se tomará en consideración la comparación de las puntuaciones del desempeño de los departamentos universitarios.

Exposición del problema y su solución: DEA¹⁹

Los principales problemas en la evaluación de la eficiencia de instituciones públicas como las universidades, radican en la falta de una norma absoluta y de un buen estimado de la función productiva en el sentido económico de la palabra. Una función de producción es la relación funcional entre insumos y productos, que define cuáles insumos y cuántos insumos se necesitan para generar un determinado producto. Otro problema es la necesidad de considerar los insumos múltiples y los productos múltiples. En este contexto, el término unidades de toma de decisiones (UTD) se utilizará para instituciones, gobiernos, etc. Una primera y bastante informal definición de eficiencia podría ser el uso que se hace de los recursos para lograr los objetivos, teniendo en cuenta los factores sociales y económicos.

La primera incursión en este terreno apareció con la investigación realizada por Charnes, Cooper y Rhodes en 1978,²⁰ primer trabajo que utilizó la técnica de DEA, cuya definición formal de la eficiencia es:

"Una unidad alcanza el 100 % de eficiencia sólo cuando:

- a) Ninguno de sus productos puede aumentar sin que aumente uno o más de sus insumos o sin que disminuyan algunos de los otros productos.
- b) Ninguno de sus insumos puede disminuir sin que disminuyan algunos de sus productos o aumenten algunos de los demás insumos".

Al no existir forma alguna de establecer una norma absoluta de eficiencia, Charnes y Cooper introdujeron la siguiente definición de eficiencia relativa: "Cualquier unidad puede alcanzar el 100 % de eficiencia relativa, sólo cuando las comparaciones con otras unidades pertinentes no ofrezca ningún indicio de ineficiencia en el uso de un insumo o producto."

El uso del DEA tiene 2 ventajas:

1. No es un método paramétrico, por lo que no es necesario conocer la relación funcional entre insumos y productos.
2. Conduce a las puntuaciones del desempeño relativo. No es necesario definir una norma de desempeño absoluto externa. Los miembros de los grupos se comparan sólo con otros miembros del mismo grupo (las unidades de toma de decisiones).

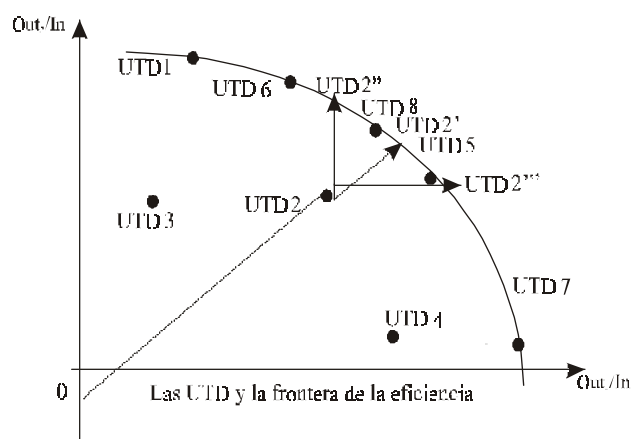
El DEA se ha utilizado en estudios de administración de hospitales, en la organización del Comando de Reclutamiento de la Marina de los Estados Unidos, en los distritos escolares, en análisis de la rentabilidad base de diferentes sectores de una misma empresa en países, universidades y departamentos universitarios.²¹⁻²³

Para mayor claridad, se explica en la figura 1 la técnica del DEA. En esta figura se representa gráficamente la eficiencia relativa de varias unidades UTD imaginarias, las cuales utilizan un insumo para generar 2 productos.

Los ejes horizontal y vertical representan la relación de cada producto (OUT₁ y OUT₂) con respecto a la cantidad de insumos empleados (IN). En el DEA, el desempeño de cada UTD se compara con los de otras unidades consideradas en ese caso. Las unidades con mejor desempeño utilizan sus insumos de una forma más óptima que las demás. Otra vez la relatividad de este enfoque tiene que subrayarse. Las unidades más óptimas forman una frontera llamada la "frontera de la eficiencia". Las unidades con menor desempeño precisan de más insumos para generar la misma cantidad de productos y por lo tanto están situadas a cierta distancia de la frontera. La ineficiencia de las mismas se representa con la distancia radial. Es obvio que todas las unidades situadas en la frontera de la eficiencia tendrán un índice de desempeño relativo igual a 1 (son totalmente eficientes).

La frontera de la eficiencia incluye a todas las unidades de toma de decisiones, como se puede ver en la figura 7.

Fig. 7. Técnica del DEA



Es por eso que la técnica se denomina DEA. A fin de incrementar el desempeño de una UTD no óptima como la UTD 2, una nueva estrategia tiene que estar dirigida hacia la evolución en pos de la frontera de la eficiencia. La técnica DEA sugiere una evolución radial a partir de la fuente hacia el punto de meta en la frontera que representa lo que esa unidad pudiera lograr. De esa forma, se logra llegar a la unidad de desempeño óptimo UTD 2: Se observa que la UTD 2 está situada entre las 2 unidades óptimas UTD 8 y UTD 5,

que forman un grupo de referencia para UTD 2. El grupo de referencia permite calcular las reducciones de insumos y ocasionalmente, las reducciones de productos necesarios para obtener una unidad con desempeño eficiente, además de mostrar dónde y cómo actuar. Hay otras vías no radiales para convertir en eficientes a ciertas unidades. Por ejemplo, una evolución puramente vertical lleva a una unidad eficiente UTD 2".

Ahora, para resolver un problema específico, hay que calcular el desempeño de cada UTD K (k=1...n) comparado con las unidades de mejor desempeño del grupo. Por lo tanto, es necesario maximizar la relación de productos ponderados (O_{rk}; r=1...s; k=1...n) con respecto a los insumos ponderados (I_{ik}; i=1...m, k=1...n), sujetos a la condición de que las relaciones similares para cada UTD sean inferiores o iguales a la unidad, es decir, a 1. El siguiente modelo matemático brinda una mejor precisión.

$$\text{Max } h_k = \frac{\sum_{r=1}^s v_{rk} * O_{rk}}{\sum_{i=1}^m u_{ik} * I_{ik}} \tag{1}$$

sujeto a

$$\frac{\sum_{r=1}^s v_{rk} * O_{rj}}{\sum_{i=1}^m u_{ik} * I_{ij}} \leq 1 \quad \text{para } j=1, \dots, n \tag{2}$$

$$v_{rk} \geq \epsilon \quad r=1 \dots s \tag{3}$$

$$u_{ik} \geq \epsilon \quad i=1 \dots m \tag{4}$$

Las variables de decisiones de este modelo son coeficientes de ponderación (V_{rk} e U_{ik}): un conjunto para cada unidad. Están constituidas de esa forma para que arrojen el resultado más beneficioso para la unidad en cuestión. Esto significa que los valores altos se atribuyen a los mejores resultados, y que los valores bajos se atribuyen a los insumos utilizados con más frecuencia. El resultado más elevado posible, es por supuesto el 1. Cuando ello ocurre, la unidad se denomina entonces eficiente. Las ecuaciones (3) y (4) definen por razones técnicas un límite por debajo del cual no pueden aparecer ponderaciones. Hemos utilizado e=0.000001.

Este modelo DEA no es lineal ni convexo. Como no existe una forma fácil de resolver este tipo de problema, es procedente transformar el modelo en un modelo de programación lineal; ello se hace otorgando al denominador un valor igual a 1. El nuevo modelo permite una fácil solución

con los paquetes de software existentes como LINGO. Así se obtiene el siguiente modelo lineal y convexo.

$$\begin{aligned} \text{Max } h_k &= \sum_{r=1}^s V_{rk} * O_{rk} \\ \text{sujeto a} \\ \sum_{r=1}^s V_{rk} * O_{rj} - \sum_{i=1}^m U_{ik} * I_{ij} &\leq 0 \\ &\text{com } j=1\dots n \\ \sum_{i=1}^m U_{ik} * I_{ik} &= 1 \\ V_{rk} &\geq \epsilon \\ U_{ik} &\geq \epsilon \end{aligned}$$

Puede demostrarse que el DEA nunca sobrestimaré el potencial de mejoramiento con respecto al identificado por cualquier otro método de ponderación. Obsérvese que las comparaciones se hacen siempre con respecto al desempeño logrado y nunca con respecto a un objetivo hipotético que puede no alcanzarse.

Departamentos universitarios²³

- Selección de resultados y recursos.

Los resultados responden la pregunta: ¿cuáles son los objetivos de un departamento universitario?

Los recursos dan respuesta a la pregunta: ¿quién o qué permite a un departamento alcanzar sus objetivos?

Resultados

Los esfuerzos en la educación, medidos por el número de estudiantes obtenidos por el departamento. Los estudiantes se clasifican según su nivel (pre-grado o posgrado) y si necesitan o no equipos de laboratorio (costosos) (una o dos variables).

Investigación básica y aplicada.

Los recursos obtenidos para hacer esta investigación²³ (una o dos variables).

Clasificaciones de los departamentos por parte de una comisión externa.²²

Recursos

El personal (FTE) se paga con los recursos de la Universidad que brinda el gobierno o con los fondos salariales.

Otros recursos de la Universidad subvencionados por el gobierno (incluye el financiamiento de equipos y de los fondos para la biblioteca).

Resultados de la Universidad Católica de Leuven: 36 departamentos de los cuales 13 tienen una puntuación de eficiencia igual a 1.

Son 100 % eficientes: derecho eclesiástico, filosofía, criminología, economía aplicada, historia, psicología y antropología, física, química, electrónica, arquitectura, ciencias biológicas aplicadas, medicina y kinestesia. Sólo para que se conozcan, se mencionan los menos eficientes: construcción (46 %), geología y geografía (42 %), estomatología (27 %), preparación física (gimnástica y deportiva: 23 %).

Eficiencia y efectividad

Para obtener un alto grado de eficiencia, una UTD está en plena libertad de utilizar sus recursos como lo desee.

Los departamentos pueden emplear todos sus insumos para obtener un solo producto o resultado (ej: enseñanza) y aun así lograr una puntuación de 1. Sin embargo, si una unidad tiene varias funciones, no se le puede considerar realmente eficiente si no se desempeña bien en cada una de ellas. Ello conlleva la introducción de la noción de efectividad. Dentro del marco del DEA se ha tenido en cuenta la efectividad a través de la imposición de límites a los factores de ponderación.²⁴

Se imponen restricciones a los instrumentos de ponderación y se añaden las condiciones siguientes:

$$V_{rk} - C_1 \sum_{l=1}^s V_{rl} \leq 0 \quad (8) \quad r = 1\dots s$$

y

$$U_{ik} - C_2 \sum_{j=1}^m U_{ij} \leq 0 \quad (9) \quad i=1\dots m$$

donde c_1 y c_2 son constantes. Los límites de 40 y 60 % para cada insumo y de 25 a 40 % para cada producto, conduce a una disminución general del grado de eficiencia. La criminología desciende al 50 % y la física al 46 %. Sin embargo, el derecho escolástico, la kinestasia y la historia permanecen el 100 %.

Por otra parte, la estomatología continúa en el 26 % y la gimnástica y el deporte se mantienen en el 22 %, por lo que siguen siendo las más deficientes.

Conclusión de la segunda parte

Se ha demostrado que la eficiencia y la efectividad pueden tenerse en cuenta al comparar el desempeño de una universidad. El DEA es un método adecuado para alcanzar este propósito. Una forma de considerar la efectividad es imponer límites a las ponderaciones de insumos y productos. Este estudio (como otros similares) ha mos-

trado claramente la potencialidad del DEA en los estudios de desempeño.²⁵⁻²⁷

Referencias bibliográficas

1. Van Raan AFJ, ed. Handbook of quantitative studies of science and technology. Amsterdam: Elsevier, 1988
2. Egghe L, Rousseau R. Introduction to infometrics: quantitative methods in library, documentation and information science. Amsterdam: Elsevier, 1990.
3. Wallace D P. A solution in search of a problem: bibliometrics & libraries. *Libr J* 1987; (112): 43-7.
4. Rousseau R. A scientometric study of the scientific publications of LUC Period 1981-1993: report [Flanders], 1995.
5. De Bruin R, Kint A, Luwel M. A study of research evaluation and planning: the University of Ghent Res Evaluat, 1993; (3):35 -41.
6. De Bruin R, Moed H F, Spruyt E. Antwerpse analyses. Antwerpen: CWTS. Leiden en UIA, 1993.
7. Moed H F, Burger W J M, Frankfort J G. The use of bibliometric data for the measurement of university research performance. *Res Policy* 1985;(4):131-49.
8. Oppenheim C. The correlation between citation counts and the 1992 research assessment exercise ratings for British library and information science university departments. *J Document* 1995; (51):18-27.
9. Abelson P. Mechanisms for evaluating scientific information and the role of peer review *J Am Soci Inform Sci* 1990; (41):216-22.
10. Wallmark JT, Sedig K G. Quality of research measured by citation method and by peer review a comparison. *IEEE Trans Eng Manag* 1986; (33):218-22.
11. Garfield E. How to use citation analysis for faculty evaluations and when is it relevant. Part I *Curr Contents* 1983;(44):
12. Arunachalam S. Science in the European Community the extent of collaboration. *Curr Sci* 1992; (63):56-8.
13. Rousseau R. Citation distribution of pure mathematics journals. En: Egghe L, Rousseau R, eds. *Infometrics* 87/88 249-62.
14. Van Hooydonk G, Gevaert R, Milisproost G. A biblioeconomic analysis of the impact factors of scientific disciplines. *Scientometrics* 1994; (30):65-81.
15. Rousseau R. Evolution d' importantes revues pharmacologiques. *Rev Fr Bibliométrie*, n. 5, p. 102-17, 1989.
16. Rousseau R, Spinak, E. Do a field list of internationally visible journals and their journal impact factors depend on the initial set of journals? A research proposal. *J Document* 1996; (52):449-56.
17. Cole S, Cole J R. Scientific output and recognition: a study in the operation of the reward system in science. *Am Sociol Rev* 1967; (32):377-90.
18. Collins MP, Berge ZL. *IPCT Journal: a case study of an electronic journal on the internet* *J Am Soci Inform Sci* 1994; (45):771-6.
19. Norman M, Stoker B. *Data envelopment analysis: the assessment of performance*. Chicester: Wiley 1991.
20. Charnes A, Cooper W W, Rhodes E. Measuring the efficiency of decision making units. *Eur J Operat Res* 1978; 2(6):429-44.
21. Arcelus F J, Coleman D F. An efficiency review of university departments. *Int J Syst Sci* 1997; 28(7):721-9.
22. Beasley J E. Determining teaching and research efficiencies. *J Operat Res Soc* n. 46. p 1995; (45):441-52.
23. Degraeve Z, Lambrechts M, Van Puymbroeck V. E en vergelijkende prestatie studie van de departementen van de Katholieke Universiteit Leuven. *Tijdsch Econ Manag* 1996; (41):165-93.
24. Pedraja-Chaparro F, Salinas Jimenez J, Smith P. On the role of weight restrictions in data envelopment analysis. *J Product Anal* 1997; (8):215-30.
25. Johnes G, Johnes J. Apples and oranges: the aggregation problem in publications analysis *Scientometrics* 1992; (25):353-65.
26. Rousseau S, Rousseau R. Data envelopment analysis as a tool for constructin scientometric indicators. *Scientometrics* 1997; (40):45-56.
27. ———. Measuring the scientific wealth of European nations: taking effectiveness into account *Scientometrics* 1998; (42):75-87.

Ronald Rousseau

ronald.rousseau @kh.khbo.be

Abstract

This article consists of two parts: the first part describes the evaluation of a small university it is based on scientometric data and aims mainly at the evaluation of internationally visible research. The second part shows how an econometric method (DEA: data enveloping analysis) can be used to include teaching and other aspects, e.g. fund raising, into the evaluation. Both approaches show how a body of bibliometric, scientometric and econometric theory can be applied to real world problems.

Subject headings: SCIENTOMETRICS; RESEARCH; EVALUATION; ECONOMETRICS/methods; IMPACT FACTOR; CITATION ANALYSIS; SCIENTIFIC PRODUCTION; ELECTRONIC PUBLISHING; BIBLIOMETRICA UNIVERSITIES/efficiency; UNIVERSITIES/effectiveness