

**EL ANÁLISIS DE CORRELACIÓN CANÓNICA COMO
INSTRUMENTO PARA LA EVALUACIÓN DE LA EFICIENCIA**

Moreno Sáez, Alfredo
eciop24@sis.ucm.es

Trillo del Pozo, David
trillo@poseidon.fcjs.urjc.es

1. Introducción:

En los análisis de la eficiencia técnica de las empresas o de las instituciones públicas se suele plantear, con carácter previo a la aplicación de los modelos, la necesidad de seleccionar las variables más representativas de la tecnología de producción que se pretende estudiar. En ese sentido, deberían recogerse los recursos de tipo humano, financiero o de capital que fuesen más relevantes en la producción de los servicios o de los bienes atribuibles a las unidades evaluadas. Además de este criterio económico, es preciso analizar si es justificable su introducción desde el punto de vista estadístico, mediante el análisis de las correlaciones existentes entre las variables explicativas de los niveles de producción obtenidos.

En ocasiones, la dificultad que entraña encontrar una buena base de datos podría justificar la selección de las mejores variables de entre las disponibles. Sin embargo, cuando se dispone de una base estadística suficientemente amplia es necesario prescindir de algunas variables para evitar la pérdida de grados de libertad en los modelos de eficiencia en los que se resuelven sistemas de ecuaciones. Dicha selección puede implicar una pérdida de información relacionada con las variables que no se han introducido en el estudio.

En este artículo se tratará este último supuesto, proponiendo concretamente un método para construir las variables que mejor representen al conjunto de inputs y outputs utilizados en el análisis de eficiencia. Mediante el análisis de correlación canónica se pueden obtener diferentes grupos de variables como combinación lineal de las variables de input, que maximizan la correlación con otro grupo de variables que representan al conjunto de los outputs. Sobre la base de estas variables, previamente transformadas para facilitar los análisis de eficiencia, se calcularán los índices de eficiencia relacionados con cada uno de los pares de variables canónicas a través de modelos de tipo estocástico y datos de panel.

La aplicación toma como base de análisis la producción científica perteneciente a los departamentos de la Universidad Politécnica de Cataluña en el periodo 1995-1999. Puesto que se trata de una universidad que utiliza un sistemas de

indicadores al servicio de la gestión, compararemos la concordancia existente entre los resultados obtenidos a partir de la aplicación que aquí se presenta y los derivados de la evaluación realizada por la Oficina Técnica de Programación de la universidad objeto de estudio.

1. EL ANÁLISIS DE CORRELACIÓN CANÓNICA Y SU ÁMBITO DE APLICACIÓN

El análisis de correlación canónica puede ser un instrumento adecuado para tratar los supuestos en los que se dispone de información de un conjunto numeroso de variables que pueden agruparse en dos grupos, el primero formado por las variables explicativas y el segundo por las explicadas y, además, dentro de cada grupo se observa un elevado grado de correlación entre las variables.

Formalizando el contenido del anterior párrafo, se estudiaría un fenómeno aleatorio de la realidad, del que puede diferenciarse un primer bloque de aspectos, que son aquellos sobre los que se desea explicar su comportamiento, representado por un conjunto de p variables, denotadas por la matriz Y , de orden $(n \cdot p)$; y un segundo bloque de características, que se consideran explicativas de las variables de la matriz Y , representadas por un conjunto de q variables, denotadas por la matriz X , de orden $(n \cdot q)$.

Si el conjunto de variables de la matriz X es numeroso, y existen correlaciones elevadas entre muchas de ellas, la estimación de un modelo de regresión, ya sea sobre la ecuación asociada a cada una de las variables endógenas por separado, ya sea mediante el procedimiento de mínimos cuadrados en tres etapas (MC3E) sobre el sistema de ecuaciones simultáneas generado por los p modelos de regresión, conduce a problemas de imprecisión en las estimaciones individuales de los parámetros. Por este motivo, los resultados de los contrastes de significación individuales de cada uno de ellos debería ser tomada con extrema precaución, puesto que la existencia de una elevada correlación entre un número elevado de variables explicativas provoca la excesiva sensibilidad de las

estimaciones ante cambios pequeños en la muestra obtenida, ocasionando modelos muy poco robustos y fiables.

Una posibilidad para resolver este problema es eliminar variables que estén altamente correlacionadas, seleccionando las que estén menos correlacionadas entre sí y que además mantengan una influencia sobre las variables endógenas fácilmente interpretable desde un punto de vista económico. Una segunda alternativa es la aplicación de un análisis factorial, recomendable cuando el número de variables con altas correlaciones es elevado y resulta difícil determinar cuáles de ellas constituyen las causas determinantes de la variabilidad de las variables endógenas, cuyo objetivo es reducir la dimensión de la matriz de variables explicativas. El análisis factorial se caracteriza por construir variables que son combinación lineal de las originales, de modo que se maximiza el poder explicativo de las causas comunes que ocasionan la variabilidad del conjunto de variables explicativas. Además, este análisis garantiza que las variables generadas por este procedimiento están incorrelacionadas entre sí, lo que asegura la inexistencia de multicolinealidad entre estas *variables ficticias*, pudiendo ser utilizadas como regresores en un modelo de regresión, previa interpretación de su significado económico. El inconveniente de este procedimiento es que no se garantiza que estas *variables ficticias* conserven el poder explicativo sobre las variables que se tratan de explicar¹.

Sería deseable la aplicación de un análisis multivariante que generara un conjunto reducido de variables endógenas, y otro de variables explicativas, de manera que ambos grupos estuviesen altamente correlacionadas, mientras que las causas comunes entre las variables que forman parte de cada grupo fuesen nulas. El análisis de correlación canónica proporciona variables que garantizan este resultado, y es recomendable su utilización cuando tanto el número de variables endógenas como explicativas es elevado, y existen elevadas correlaciones de las variables de cada grupo entre sí. Mediante este método de reducción de variables

¹ La realización de un análisis factorial sobre cada uno de dos conjuntos de variables (endógenas por un lado y explicativas por otro) garantiza la incorrelación intra-grupo, pero no una elevada correlación entre grupos de variables explicativas y grupos de explicadas.

se puede eliminar el ineludible problema de multicolinealidad que ocasionaría la consideración de todas las variables en un modelo de regresión.

Para el citado análisis se parte de las matrices X e Y que representan el conjunto de regresores y variables endógenas respectivamente, siendo la matriz: $Z = [X \ Y]$, de dimensión $n \times (q + p)$, que tendrá la siguiente matriz de varianzas y covarianzas:

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

Donde S_{11} es la matriz de varianzas y covarianzas de las q variables de la matriz X , S_{12} es la matriz de covarianzas entre las variables de la matriz X y la matriz Y , S_{22} es la matriz de varianzas y covarianzas de las p variables de la matriz Y , mientras que S_{21} coincide con la transpuesta de S_{12} .

El análisis de correlación canónica consiste en la construcción de pares sucesivos de variables, que son combinaciones lineales de las originales, de modo que cada par represente la mejor explicación de cada conjunto respecto al otro que no haya sido ya obtenida por los pares anteriores. En definitiva, la primera etapa del análisis de correlación canónica consiste en construir una variable CCI1 y otra CCO1², con la siguiente expresión:

$$CCI1 = \alpha_{11}X_1 + \alpha_{12}X_2 + \dots + \alpha_{1q}X_q$$

$$CCO1 = \beta_{11}Y_1 + \beta_{12}Y_2 + \dots + \beta_{1p}Y_p$$

De forma que la correlación entre estas dos variables sea la mayor de entre todas las combinaciones lineales posibles de los dos grupos de variables. Por consiguiente, el problema consiste en la obtención de los vectores de ponderaciones **a** y **b** que garanticen que la covarianza entre CCI1 y CCO1 sea la máxima posible, consiguiendo, igualmente, que la varianza de estas dos nuevas

² En la aplicación de este trabajo el primer grupo de variables denotadas como CCI corresponden a combinaciones lineales de las variables reales de input y las denotadas como CCO corresponden a las variables canónicas derivadas de variables reales de output.

variables sean igual a la unidad, para garantizar la homogeneidad en la dispersión de éstas respecto a los pares de variables sucesivos.

Como señala Novales (1990) puede demostrarse que los vectores de ponderaciones deseados son el resultado de la resolución del sistema de ecuaciones³:

$$\begin{pmatrix} -\lambda S_{11} & S_{12} \\ S_{21} & -\lambda S_{22} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{a}} \\ \hat{\mathbf{b}} \end{pmatrix} = 0$$

Donde el valor de λ que cumple ese sistema será la raíz del polinomio de grado p obtenido del cálculo del siguiente determinante:

$$\begin{vmatrix} -\lambda S_{11} & S_{12} \\ S_{21} & -\lambda S_{22} \end{vmatrix} = 0 \quad (\text{ec. 1})$$

De manera que el valor de λ que genera los vectores \mathbf{a} y \mathbf{b} que proporcionan una correlación máxima entre CCI1 y CCO1 son los que resultan del mayor de los valores de λ que cumplen (ec. 1), puesto que el valor de λ representa la correlación existente entre las nuevas variables construidas.

La siguiente etapa consiste en la construcción de un segundo par de variables CCI2 y CCO2, también de varianza unitaria, que experimentan la mayor correlación entre sí de todas las posibles combinaciones lineales que pueden construirse con los dos grupos de variables de partida, sin tener en cuenta el par de combinaciones obtenidos en la etapa anterior. Las ponderaciones \mathbf{a} y \mathbf{b} que garantizan tal circunstancia son las resultantes del segundo mayor valor de λ que cumplen (ec. 1).

Procediendo de forma recursiva de la manera descrita, se pueden determinar r pares de variables artificiales, llamadas *variables canónicas* ($r < p$), hasta conseguir que la proporción de las causas comunes entre los grupos de variables (endógenas e influyentes) explicadas por las mismas sea suficientemente grande.

Lo habitual es que un número reducido de pares de variables tenga un poder explicativo conjunto de la mayor parte de las causas comunes a los dos conjuntos de variables, de modo que éstas puedan sustituir a las originales, eliminando los problemas de multicolinealidad entre los regresores, puesto que el procedimiento descrito garantiza que las variables explicativas están todas ellas incorrelacionadas entre sí.

2. ANÁLISIS EXPLORATORIO Y SELECCIÓN DE VARIABLES A PARTIR DEL ANÁLISIS DE CORRELACIÓN CANÓNICA

2.1. Presentación de las variables objeto de estudio

En este apartado presentaremos la muestra de variables que se utilizarán para la evaluación de la eficiencia en la última parte de este trabajo y que serán objeto de exploración a través del análisis de correlación canónica. Como hemos comentado, se analizarán los datos relativos a la producción científica de los Departamentos de la Universidad Politécnica de Cataluña (UPC). Se ha seleccionado esta Universidad porque ofrece una importante experiencia en materia de gestión y planificación estratégica, lo que ha permitido el desarrollo de un sistema de indicadores para la gestión de la investigación. Este hecho hace que podamos manejar indicadores de calidad de la investigación y una serie temporal amplia, lo que en un contexto de producción pública no siempre es fácil de obtener⁴.

Con el fin de seleccionar las variables más relevantes para el análisis de la eficiencia de los Departamentos de la UPC hemos utilizado una amplia base de datos en la serie 1995-1999, que dividimos en dos grupos, variables de

³ Un análisis adicional de la Teoría de la Correlación canónica puede encontrarse en Martín-Guzman, P. y Martín Pliego, J. (1993).

⁴ Un elemento de incentivo a la realización de este trabajo fue poder contar con la colaboración y los comentarios de la Oficina Técnica de Programación de la UPC, que en todo momento ha mostrado su interés en conocer los resultados de la investigación. El periodo 95-99 coincide con la fase de implantación de un nuevo sistema de indicadores de evaluación en relación con la puesta en marcha de la Planificación estratégica de la UPC.

producción y variables de input. Concretamente se recogen las siguientes variables de producción:

VARIABLES DE PRODUCCIÓN*	DESCRIPCION
PARNOT	Puntos por publicaciones en revistas notables
PARCON	Intervención en congresos de tipo notables
ART A, B, C	Artículos en revistas en función del índice de impacto
COMNAC, COMINT	Comunicaciones y ponencias nacionales e internacionales
POSTN, POSTI	Pósters nacionales e internacionales
PREMN, PREMI	Premios nacionales e internacionales
LIBRN, LIBRI	Libros nacionales e internacionales
CAPNAC, CAPINT	Capítulos de libros nacionales e internacionales
TESLEI	Tesis leídas

* Las patentes no han sido incluidas en el estudio puesto que representan un indicador poco aconsejable para el análisis temporal, en la medida en que recogen las patentes de años anteriores a la serie utilizada.

En este primer bloque se incluyen dieciséis variables de investigación, entre las que se incluyen tanto variables de calidad como de cantidad. A priori no establecemos ninguna selección basada en criterios de tipo económico, para observar si, a través del análisis de correlación canónica, existe alguna asociación estadística entre las variables de calidad o, si al contrario es difícil establecer dicha asociación. Una vez realizado este primer filtro, habría que definir qué variables definen la calidad en la producción científica. La Guía del Plan Nacional de Calidad de las Universidades sugiere orientar la evaluación hacia una triple clasificación de las publicaciones periódicas en función de indicadores bibliométricos. La Universidad Politécnica ha respondido a la citada propuesta implantando un sistema de indicadores al servicio de los departamentos evaluados.

Sobre esa base, incluimos en el análisis de este apartado los denominados puntos por publicaciones notables, artículos en revistas de gran repercusión científica ponderados por el índice de impacto del Science Citation Index⁵. También

⁵ La fuente de referencia para las evaluaciones es que las revistas estén indexadas en el Journal of Citation Report. En relación con estas variables es interesante señalar que los indicadores de calidad también adolecen de sesgos a favor de los grupos de Departamentos que publican con mayor asiduidad en el Journal of Citation Report o realizan intervenciones en congresos notables. El problema de evaluar unidades de una misma universidad es que puede existir una cierta especialización productiva por parte de determinadas unidades en labores en las que no intervienen otras. Este problema de homogeneidad de la muestra puede atenuarse mediante la inclusión de otro tipo de variables como los libros, los premios o las tesis y mediante una ponderación de esas actividades que corrija el sesgo de los indicadores de publicaciones. Una vía adicional de solución la está poniendo en marcha la UPC a partir de este año incluyendo publicaciones del Journal of Social Science en el grupo de revistas notables.

incluimos la triple clasificación de artículos clasificados por orden de repercusión científica (Revistas de tipo A, revistas de calidad aunque de menor impacto que las notables, tipo B, básicamente publicaciones nacionales y tipo C o de divulgación). La variable PREM alude a los premios científico-técnicos y de arquitecturas puede tomarse como una variable proxy de la calidad, en la medida en que los trabajos científicos han sido objeto de evaluación y selección por parte de un comité científico. Además, se han recogido otros indicadores cuantitativos que recogen otras publicaciones o actividades de investigación (comunicaciones, posters y Tesis).

Por otra parte, el cuadro siguiente recoge los catorce inputs que se emplearán en el análisis de reducción de datos:

VARIABLES DE RECURSOS	DESCRIPCION
CU, TU	Catedráticos y Titulares de Universidad tiempo completo
CEU, TEU	Catedráticos y Titulares de Escuela Universitaria t. completo
ASTC	Asociados tiempo completo
BECINV	Becarios de investigación e investigadores propios
ICTTEUR	Ingresos de transferencia de tecnología por proyectos europeos
ICTTCICYT	Ingresos de proyectos CICYT y DGCICYT
ICTTCBS	Ingresos para compra bienes y servicios
ICTTSUBV	Ingresos a través de Subvenciones
C2INV	Asignación presupuestaria capítulo dos para investigación
ESTLIC	Estudiantes de Licenciatura
ESTDOCT	Estudiantes de Doctorado
CREDOCT	Créditos impartidos de doctorado

En el terreno de los recursos económicos hay que señalar que los departamentos de la Universidad Politécnica se financian de forma muy autónoma a través del Centro de Transferencia de Tecnología, que gestiona los proyectos de investigación y los convenios con otras instituciones (los denominados ingresos por transferencia de tecnología (ICTT⁶). Por otra parte, el único recurso

⁶ En el análisis de correlación canónica se incluye todo el desglose entre ingresos obtenidos a través de la CICYT y DGCICYT –ICCTCICYT-, ingresos por proyectos europeos –ICTTEUR-, ingresos por convenios para la compra de bienes y servicios –ICTTCBS- e ingresos por subvenciones –ICTTSUBV-. Esta primera variable podría ser utilizada como variable de output, en cuanto que la obtención de ingresos a través de proyectos puede representar un indicador indirecto de la importancia de la producción investigadora de los Departamentos. Esto es así, porque generalmente los recursos obtenidos a través de la firma de convenios o de la obtención de proyectos terminan concretándose en publicaciones. Sin embargo, en este estudio hemos considerado que los ingresos pueden servir para reflejar las diferencias en la escala de producción de los Departamentos. Teóricamente los departamentos que obtuvieran más recursos para financiar sus proyectos deberían materializar un mayor nivel de investigaciones. Lo contrario para los que cuentan con escasos recursos a través de proyectos.

presupuestario descentralizado a los Departamentos corresponde a la asignación que realiza la universidad para financiar el gasto corriente por investigaciones, (el capítulo dos para investigación, que en el análisis se denomina -C2INV-)⁷.

Para finalizar hemos recogido un conjunto de variables de docencia con el fin de conocer si condicionan la producción científica. En concreto utilizamos tres indicadores de la carga docente, el número de estudiantes de licenciatura y los créditos y el número de estudiantes de doctorado.

A continuación realizamos el citado análisis de correlación canónica, donde los inputs formarían parte del primer grupo de variables y explicarían los output o segundo grupo. Como se ha señalado, cada variable canónica se obtiene como combinación lineal de las variables reales tipificadas originales y puede haber tantas como variables reales. Al igual que en el análisis factorial, se ordenan de mayor a menor las variables canónicas en función del porcentaje de la varianza de cada grupo explicada por las mismas. En nuestro análisis, los pares de variables serían CCI1, relacionado con CCO1; CCI2, relacionado con CCO2, CCI3 relacionado con CCO3, y así sucesivamente hasta la última de las componentes canónicas.

⁷ Un tercer grupo de variables podría corresponder a una valoración del capital. En nuestro caso disponíamos de indicadores de amortizaciones de bienes muebles e inmuebles y de metros cuadrados de superficie utilizada por el Departamento. Hemos rechazado ambas posibilidades de acuerdo con la sugerencia de la propia Oficina de Programación, que considera que estas variables introducen nuevos sesgos, ya que existen Departamentos que precisan de una gran cantidad de superficie para la investigación (por ejemplo, mecánica de fluidos necesita una gran superficie para inmersiones). Por otro lado, gran parte de los laboratorios están compartidos, por lo que resulta imposible imputar a cada departamento esta parte del capital. Otra variable que puede servir de aproximación es el número de ordenadores por profesor (por ejemplo Valderrama (96) lo utiliza como variable de input), aunque en el caso de la UPC no tendría demasiado sentido porque prácticamente todo el personal a tiempo completo dispone de equipo informático. Por último, el presupuesto para inversiones está centralizado y no pudimos acceder a un desglose por Departamentos, que además presentaría de nuevo el problema de las diferentes infraestructuras de producción de los departamentos. Por todos estos motivos hemos preferido no incluir ninguna de estas variables en el análisis.

2.2. Resultados del análisis e interpretación de las componentes canónicas

En el cuadro 1 recogemos las correlaciones de las componentes canónicas de input y output con las variables reales de las que proceden. El resultado más interesante para el análisis de eficiencia es que la primera componente del grupo de outputs correlaciona altamente con un grupo de variables que reflejan actividades relacionadas con la calidad de la investigación, con especial relevancia en las tesis doctorales, los puntos por actividades notables (PARNOT, PARCON) y la actividad en congresos de carácter internacional (pósters y comunicaciones). Todos estos resultados están explicados por la primera componente canónica del grupo de inputs, que representa de forma clara al presupuesto corriente en investigación y a las categorías del personal docente e investigadoras relativas a Titulares y Catedráticos de Universidad. La alta cifra de correlación con el presupuesto corriente para investigación es perfectamente lógica puesto que éste se establece sobre los parámetros de investigación de calidad (puntos PAR).

Del análisis del segundo par de variables canónicas se observa una relación directa de las variables de doctorado con las de premios y capítulos de libros. Si estudiamos las correlaciones más altas entre las variables originales y las componentes canónicas de input y de output, también sería importante la relación entre las categorías de asociados a tiempo completo con la intervención en congresos notables, lo que podría explicarse por la necesidad de incrementar el curriculum para acceder a alguna de las categorías de numerario.

El tercer par (CCI3-CCO3) relaciona de nuevo las variables de doctorado con premios. La relación de estos últimos parece ser inversa con la existencia de Catedráticos de Escuela. El cuarto indica de forma clara una correlación entre las categorías de titulares de universidad y los artículos notables y de forma negativa con los artículos con menor impacto científico.

Por otra parte, el quinto par de variables (CCI5-CCO5) liga las categorías de Catedráticos de Universidad e ingresos para la compra de bienes y servicios con la producción de libros, lo cual resulta coherente, puesto que son categorías con la carrera universitaria consolidada y, por tanto, disponen de más experiencia y posibilidades de publicación de libros.

De las correlaciones relativas al sexto par CCI6-CCO6 se pueden establecer dos variables ligadas a los estudios de licenciatura (número de estudiantes y Titulares de Escuela) relacionadas de forma inversa con las publicaciones de menor impacto y con los capítulos y los libros internacionales⁸.

Por último, el séptimo par reflejaría una alta correlación de los becarios de investigación con la producción de pósters y de forma negativa con las comunicaciones y ponencias, este último supuesto es también fácilmente interpretable puesto que son categorías de investigadores jóvenes.

⁸ No se observa una relación negativa entre la carga docente, medida a través del número de estudiantes de licenciatura, y la producción científica, lo que en el caso de la UPC es debido a que las plantillas docentes van adecuándose a las necesidades de carga docente. De hecho, del análisis de regresión (a través de MCO) con las publicaciones notables y el número de estudiantes de licenciatura se obtiene un parámetro positivo y significativo.

Cuadro 1: Correlaciones de las variables canónicas con las variables reales

Correlac.	CCI1	CCI2	CCI3	CCI4	CCI5	CCI6	CCI7	CCO1	CCO2	CCO3	CCO4	CCO5	CCO6	CCO7
ICTTEUR	53.7%	-43.5%	24.8%	-24.7%	1.4%	-4.0%	24.2%	48.0%	-35.9%	18.6%	-15.6%	0.8%	-2.0%	11.0%
ICTCICYT	67.3%	-36.4%	19.9%	-12.3%	12.8%	0.9%	27.2%	60.2%	-30.1%	14.9%	-7.8%	7.3%	0.4%	12.4%
ICTTCBS	34.8%	-0.8%	11.5%	-18.7%	-54.5%	1.3%	24.5%	31.1%	-0.6%	8.6%	-11.8%	-31.1%	0.7%	11.2%
ICTTSUBV	7.7%	7.0%	1.8%	4.0%	-33.8%	-13.8%	-0.1%	6.9%	5.8%	1.3%	2.6%	-19.3%	-7.1%	0.0%
C2INV	90.9%	2.6%	18.2%	7.2%	-4.2%	19.6%	4.6%	81.3%	2.1%	13.6%	4.6%	-2.4%	10.1%	2.1%
CU	61.5%	-16.6%	18.5%	-9.9%	-57.7%	9.3%	1.3%	55.0%	-13.7%	13.9%	-6.2%	-33.0%	4.8%	0.6%
CEU	32.0%	10.1%	-62.8%	9.1%	-6.8%	12.9%	-7.4%	28.6%	8.3%	-47.0%	5.7%	-3.9%	6.6%	-3.4%
TU	74.8%	-8.6%	20.5%	-46.7%	-4.7%	12.5%	-30.9%	66.9%	-7.1%	15.3%	-29.6%	-2.7%	6.4%	-14.1%
TEU	12.2%	-7.8%	-20.4%	10.4%	27.3%	53.4%	10.9%	10.9%	-6.4%	-15.3%	6.6%	15.6%	27.4%	5.0%
ASTC	61.7%	-58.1%	25.7%	28.3%	8.1%	-3.6%	-23.5%	55.1%	-48.0%	19.3%	17.9%	4.6%	-1.8%	-10.7%
BECINV	68.0%	8.2%	31.1%	13.9%	21.6%	7.3%	43.7%	60.8%	6.7%	23.3%	8.8%	12.4%	3.7%	19.9%
ESTLIC	43.3%	-20.0%	37.1%	15.9%	-4.0%	44.7%	-30.8%	38.7%	-16.6%	27.8%	10.1%	-2.3%	22.9%	-14.0%
ESTDOCT	48.9%	36.8%	62.3%	-5.8%	-15.7%	-28.6%	3.2%	43.7%	30.4%	46.6%	-3.7%	-9.0%	-14.6%	1.5%
CREDOCT	46.4%	35.3%	71.8%	20.5%	-13.1%	-9.9%	-3.1%	41.5%	29.2%	53.8%	13.0%	-7.5%	-5.1%	-1.4%
PARCON	58.7%	-45.4%	23.7%	-16.4%	1.5%	-9.6%	-3.6%	65.7%	-54.9%	31.7%	-25.9%	2.6%	-18.6%	-7.8%
PARNOT	58.0%	8.5%	-24.9%	-27.2%	-3.8%	1.2%	-0.5%	64.9%	10.2%	-33.3%	-42.9%	-6.7%	2.3%	-1.2%
ARTA	50.7%	19.1%	-10.8%	9.2%	18.3%	5.6%	10.3%	56.7%	23.2%	-14.4%	14.5%	32.1%	11.0%	22.5%
ARTB	12.6%	12.4%	-3.1%	17.2%	-0.8%	-10.0%	-6.4%	14.1%	15.0%	-4.2%	27.1%	-1.4%	-19.5%	-14.0%
ARTC	13.3%	24.8%	-1.3%	24.1%	3.4%	-25.7%	-5.1%	14.9%	30.0%	-1.7%	38.1%	5.9%	-50.2%	-11.1%
COMNAC	42.5%	-30.9%	-5.3%	19.2%	-2.5%	-3.4%	-19.0%	47.6%	-37.4%	-7.0%	30.3%	-4.4%	-6.6%	-41.7%
COMINT	67.4%	-31.3%	-2.7%	2.2%	6.8%	-3.4%	-9.6%	75.4%	-37.9%	-3.6%	3.5%	11.9%	-6.5%	-20.9%
POSTN	28.1%	6.8%	-27.7%	6.6%	1.3%	4.1%	18.3%	31.4%	8.3%	-37.0%	10.4%	2.3%	8.1%	40.1%
POSTI	50.1%	16.1%	-36.8%	-11.9%	7.2%	11.3%	-2.5%	56.0%	19.5%	-49.2%	-18.7%	12.7%	22.0%	-5.5%
PREMN	8.8%	45.1%	48.5%	1.7%	11.6%	4.0%	-11.0%	9.9%	54.5%	64.7%	2.6%	20.4%	7.9%	-24.1%
PREMI	22.9%	34.9%	32.7%	-10.6%	12.8%	-3.9%	-7.5%	25.6%	42.2%	43.7%	-16.7%	22.5%	-7.6%	-16.4%
LIBRN	12.2%	23.6%	11.8%	2.9%	-38.6%	11.6%	-5.9%	13.6%	28.6%	15.8%	4.6%	-67.6%	22.5%	-12.9%
LIBRI	5.9%	32.5%	12.2%	-12.0%	-21.4%	-14.5%	-10.8%	6.7%	39.3%	16.3%	-19.0%	-37.5%	-28.4%	-23.8%
CAPNAC	20.9%	37.5%	16.1%	-5.2%	-5.9%	-4.5%	9.5%	23.4%	45.4%	21.4%	-8.2%	-10.3%	-8.9%	20.7%
CAPINT	32.4%	29.3%	0.5%	-8.0%	-5.5%	-25.8%	10.5%	36.2%	35.5%	0.6%	-12.6%	-9.5%	-50.4%	23.1%
TESLEI	76.4%	-12.1%	8.3%	6.1%	-11.8%	-1.6%	5.7%	85.4%	-14.7%	11.0%	9.6%	-20.7%	-3.1%	12.6%
%VARACUM	30.3	37.6	50.9	54.7	61.4	66.3	70.9	17.17	25.16	29.85	31.68	33.51	34.83	35.80

Cuadro 2: Resumen de relaciones

Componente Input → Output	Correlación input	Correlación output
CCI1-CCO1	RECURSOS ECON Y HUMANOS	INVESTIGACIÓN CALIDAD
CCI2-CCO2	DOCTORADO (+)	PREMIOS Y CAPÍTULOS(+)
	INGRESOS Y ASOCIADOS (-)	CONGRESOS NOTABLES (-)
CCI3-CCO3	DOCTORADO (+)	PREMIOS (+)
	CATEDR. ESCUELA (-)	POSTER (-)
CCI4-CCO4	TITULARES UNIVERSIDAD (-)	ARTÍCULOS CALIDAD (-)
		ARTÍCULOS - CALIDAD (+)
CCI5-CCO5	INGRESOS COMPRA BB.SS. (-)	LIBROS (-)
	CATEDRÁTICOS UNIVERSIDAD (-)	
CCI6-CCO6	TITULARES ESCUELA (+)	ARTÍCULOS - CALIDAD (-)
	ESTUDIANTES LICENCIATURA (+)	
CCI7-CCO7	BECARIOS INVESTIGACIÓN (+)	POSTER (+)
		COMUN. Y PONENCIAS (-)

* En el cuadro se respeta el signo de las correlaciones con las variables canónicas para mostrar si las relaciones son directas o inversas

3. TRANSFORMACIÓN DE LAS VARIABLES OBTENIDAS PARA LA EVALUACIÓN DE EFICIENCIA

Como ya se ha explicado en el epígrafe de descripción teórica del análisis de correlación canónica, las 14 componentes canónicas se obtienen a partir de las 16 variables de output y 14 variables de input, a partir de la siguiente expresión:

$$CCI_H = \sum_{i=1}^{14} \alpha_{Hi} x_i^* = \sum_{i=1}^{14} \alpha_{Hi} \left(\frac{x_i - \bar{x}_i}{\sigma_{x_i}} \right) \quad \forall H = 1, 2, \dots, 14 \quad (\text{ec. 2})$$

$$CCO_H = \sum_{j=1}^{16} \beta_{Hj} y_j^* = \sum_{j=1}^{16} \beta_{Hj} \left(\frac{y_j - \bar{y}_j}{\sigma_{y_j}} \right) \quad \forall H = 1, 2, \dots, 14 \quad (\text{ec. 3})$$

Siendo \bar{x}_i la media aritmética de la variable x_i , \bar{y}_j la media aritmética de la variable y_j , σ_{x_i} la desviación típica de x_i , mientras que σ_{y_j} es la desviación típica de y_j .

Estas componentes canónicas tienen todas ellas media nula y desviación típica uno, de modo que son adimensionales y, por consiguiente, comparables entre sí. Sin embargo, en los análisis de ratios la dimensión de las variables reales utilizadas puede influir decisivamente en la clasificación de las unidades derivadas del mismo. Algo equivalente ocurre en el análisis envolvente de datos, que es una extensión del análisis de ratios aunque establece ponderaciones para los inputs y outputs utilizados en la construcción de un único indicador de la eficiencia⁹. También influye este hecho en la estimación de frontera estocástica que presentaremos más tarde. Por ello, sería conveniente obtener unas variables que mantuviesen las propiedades deseables de las componentes canónicas¹⁰, y además, cuyos valores no fueran adimensionales, es decir, que cada una de ellas

⁹ Sin embargo, existe la posibilidad de manejar modelos envolventes aditivos, que no resultan afectados por los cambios de origen como los que planteamos en este apartado.

¹⁰ Ya hemos señalado que las variables de input son, de entre todas las combinaciones lineales posibles sobre las variables reales, las más explicativas de las de output, y además, las variables de cada grupo están incorrelacionadas entre sí.

dependiera de las unidades de aquellas variables que tuvieran una mayor ponderación en la construcción de las mismas¹¹.

Para ello, lo que haremos es operar sobre (ec. 2), de modo que obtendremos:

$$CCI_H = \sum_{i=1}^{14} \alpha_{Hi} X_i^* = \sum_{i=1}^{14} \alpha_{Hi} \frac{X_i}{\sigma_{X_i}} - \sum_{i=1}^{14} \alpha_{Hi} \frac{\bar{X}_i}{\sigma_{X_i}} \quad \forall H = 1, 2, \dots, 14$$

Despejando de esta ecuación, podemos calcular una nueva variable, a la cual le asignamos el nombre de CCIBH:

$$CCI_H + \sum_{i=1}^{14} \alpha_{Hi} \frac{\bar{X}_i}{\sigma_{X_i}} = \sum_{i=1}^{14} \alpha_{Hi} \frac{X_i}{\sigma_{X_i}} = CCIBH \quad \forall H = 1, 2, \dots, 14$$

Donde CCIBH sigue teniendo desviación típica uno, debido a que simplemente se ha sometido a cada componente canónica a un cambio de origen¹². Pero la media será, en general, distinta de cero, y su valor dependerá del ratio media-

¹¹ Un cambio de escala supondrá una alteración de la dispersión absoluta de la variable transformada. De esta forma, si el factor que multiplica a la variable es mayor que uno (es el caso de las transformaciones realizadas), la dispersión se ve incrementada, asemejándose la variabilidad de las variables transformadas a la dispersión de las variables originales más representadas por la componente canónica correspondiente. Al ser los valores de éstas más ajustadas a la realidad, la estimación de los parámetros asociados a las mismas también representarán de forma más fiel el grado de sensibilidad de la variable endógena ante un incremento de las explicativas. De igual modo, las perturbaciones del modelo de regresión con las variables sometidas al cambio de escala, representarán de forma más ajustada la variabilidad no explicada con el modelo de regresión, y por consiguiente, el cociente entre la varianza del término de ineficiencia y la varianza total (parámetro gamma de la especificación econométrica de Batesse y Coelli) también representará mejor la divergencia en los comportamientos individuales de los Departamentos en lo que se refiere a la variabilidad de sus perturbaciones. Por consiguiente, las eficiencias técnicas estimadas permitirán discriminar mejor aprovechamiento óptimo de los recursos de los diferentes departamentos.

¹² Hay que destacar que un cambio de origen, simplemente traslada hacia la derecha o la izquierda los valores la variable sobre la que se aplica, con lo que no supone una alteración de su dispersión absoluta. Por consiguiente, la estimación de los parámetros asociados a cada una de las variables explicativas no se ve alterada, cambiando únicamente de valor la estimación del término constante del modelo de regresión. No obstante, un cambio de origen sí va a alterar el cociente entre el valor máximo y el valor mínimo de la variable (coeficiente de apertura), con lo que también va a alterar la estimación del cociente entre la varianza del término de ineficiencia (recogido en el término de error del modelo) y la varianza total de las perturbaciones, denotado por gamma en la especificación econométrica de Batesse y Coelli. Por consiguiente, un cambio de origen que suponga un incremento del coeficiente de apertura, supondrá un incremento de gamma, y por tanto, un aumento de la variabilidad de las eficiencias técnicas estimadas, que permitirá discriminar mejor entre los diferentes departamento en lo que se refiere a la evaluación de su eficiencia en la gestión de sus recursos disponibles.

desviación típica de cada una de las variables ponderadas por su correspondiente coeficiente canónico.

De igual modo, se obtendrán las variables de Output CCOBH

$$CCOBH = CCOH + \sum_{j=1}^{16} \beta_{Hj} \frac{\bar{y}_j}{\sigma_{y_j}} \quad \forall H = 1, 2, \dots, 14$$

Cuya media dependerá de las variables originales ponderadas por su correspondientes coeficientes canónicos, y seguirá teniendo desviación típica igual a uno, puesto que simplemente se ha definido un cambio de origen para cada una de ellas.

Un paso más encaminado a la obtención de componentes canónicas ajustadas a la realidad de las características que están representando, es asignarle una variabilidad acorde a las variables reales a partir de las cuales han sido construidas. De esta manera, si denotamos a r_{Hi} como el coeficiente de correlación entre CCIHB¹³ y X_i , se le puede asignar a cada componente el promedio de las 14 varianzas de las variables a partir de las cuales ha sido obtenida, ponderadas por el coeficiente de correlación al cuadrado entre cada una de las variables reales y la componente canónica considerada. Por consiguiente, la varianza asignada a la componente CCIHB será:

$$\sigma_{iH}^2 = \sum_{i=1}^{14} \frac{r_{Hi}^2}{\sum_{k=1}^{14} r_{Hk}^2} \sigma_i^2 \quad \forall H = 1, 2, \dots, 14$$

Mientras que la varianza asignada a la componente de output CCOHB será:

¹³ El coeficiente de correlación entre X_i y RCCIHB tomará el mismo valor que con CCIH, puesto que RCCIHB se ha obtenido con un simple cambio de origen respecto a la componente canónica resultante del análisis de correlación canónica original, con la que ni la covarianza con otras variables, ni los momentos respecto a la media resultan alterados.

$$\sigma_{OH}^2 = \sum_{j=1}^{16} \frac{r_{Hj}^2}{\sum_{l=1}^{14} r_{Hl}^2} \sigma_j^2 \quad \forall H = 1, 2, \dots, 14$$

De este modo, se puede obtener una transformación de las variables canónicas, cuya desviación típica sea dimensional, es decir, mayor cuanto mayores sean los valores de las variables que tienen elevados grados de correlación con ellas. Para ello, haremos la siguiente transformación:

$$RCCIHC = RCCIHB \cdot \sqrt{\frac{\sum_{i=1}^{14} \frac{r_{Hi}^2}{\sum_{k=1}^{14} r_{Hk}^2} \sigma_i^2}{\sum_{k=1}^{14} r_{Hk}^2}} \quad \forall H = 1, 2, \dots, 14$$

$$RCCOHC = RCCOHB \cdot \sqrt{\frac{\sum_{j=1}^{16} \frac{r_{Hj}^2}{\sum_{l=1}^{14} r_{Hl}^2} \sigma_j^2}{\sum_{l=1}^{14} r_{Hl}^2}} \quad \forall H = 1, 2, \dots, 14$$

Las componentes obtenidas tras esta segunda transformación tienen medias y desviaciones típicas diferentes, cuyo valor dependerá de la media y la desviación típica de las variables que tienen una mayor ponderación en su construcción, pero tienen el inconveniente de que algunas de ellas tienen medias negativas, valores que no pueden ser utilizados en los análisis de eficiencia, que suponen niveles positivos para inputs y outputs. Además, dado que las variables tienen una dispersión heterogénea, para la estimación de modelos de regresión es recomendable eliminar la heterocedasticidad tomando logaritmo neperiano sobre ellas, para lo cual, es necesario que todos sus valores sean positivos. Por ello, se ha considerado conveniente eliminar los valores negativos de las componentes canónicas, mediante la siguiente transformación:

$$RCCIHD = RCCIHC + \left(\begin{matrix} \text{MIN} & \text{RCCIHC} \\ h & \\ h=1,2,\dots,38 \end{matrix} \right) \cdot (-2) \quad \forall H = 1, 2, \dots, 14$$

$$RCCOHD = RCCOHC + \left(\underset{h=1,2,\dots,38}{\text{MIN}}_h \text{ RCCOHC} \right) \cdot (-2) \quad \forall H = 1, 2, \dots, 14$$

Que nos proporciona variables cuyo valor mínimo coincide con el de la variable de partida (RCCIHC ó RCCOHC), pero con signo positivo, manteniéndose la variabilidad de cada componente inalterada, puesto que simplemente se ha realizado un cambio de origen.

Sin embargo, tras realizar las estimaciones del modelo econométrico de medición de la eficiencia con las anteriores variables, se obtuvieron resultados con una dispersión de eficiencia entre los departamentos muy reducida, y se detectó que ésta era motivada a la escasa dispersión entre los valores de las variables RCCOHD. Concretamente, se comprobó que el coeficiente de apertura de las variables originales era sensiblemente superior al de las componentes canónicas, considerando que la escasa diferencia entre los valores de las variables RCCOHD era la causa de la reducida dispersión de las eficiencias estimadas entre los 38 Departamentos utilizados para las estimaciones. Por ello, se decidió alterar el cambio de origen aplicado sobre las variables RCCOHC y RCCIHC, de modo que las variables resultantes tuvieran una misma dispersión, mediada mediante su coeficiente de apertura¹⁴.

Para realizar esta transformación, se calculó previamente el coeficiente de apertura de la media de las dieciséis variables de Output, eliminada su dimensión dividiendo previamente su valor por su desviación típica, de modo que todas ellas fuesen homogéneas, obteniendo un resultado de 108,422662. Asimismo, se calculó la media de las catorce variables de Input, eliminada igualmente su dimensión de la misma forma, obteniendo un resultado de 11,559654. De esta forma, el cambio de origen que se ha realizado sobre las variables RCCIHC y RCCOHE, respectivamente, ha sido el siguiente:

¹⁴ Se ha utilizado el coeficiente de apertura en lugar de la desviación típica, debido a que se necesita una medida de dispersión conjunta de las dieciséis variables de Output por un lado, y de las catorce variables de Input por otro. Por tanto, es imprescindible utilizar una medida de dispersión que sea adimensional, y hemos seleccionado el coeficiente de apertura, por ser la más sencilla de calcular.

$$RCCOHE = RCCOHC - \left(\min_{h=1,2,\dots,38}^h RCCOHC \right) + \frac{\left(\max_{h=1,2,\dots,38}^h RCCOHC \right) - \left(\min_{h=1,2,\dots,38}^h RCCOHC \right)}{\frac{\sum_{j=1}^{16} \left(\frac{y_j}{\sigma_y} \right)}{16} - 1}$$

$$RCCIHE = RCCIHC - \left(\min_{h=1,2,\dots,38}^h RCCIHC \right) + \frac{\left(\max_{h=1,2,\dots,38}^h RCCIHC \right) - \left(\min_{h=1,2,\dots,38}^h RCCIHC \right)}{\frac{\sum_{i=1}^{14} \left(\frac{x_i}{\sigma_x} \right)}{14} - 1}$$

$$\forall H = 1, 2, \dots, 14$$

$$\text{Donde tenemos que } \frac{\sum_{j=1}^{16} \left(\frac{y_j}{\sigma_y} \right)}{16} = 108,422662, \text{ y } \frac{\sum_{i=1}^{14} \left(\frac{x_i}{\sigma_x} \right)}{14} = 11,559654.$$

Con esta transformación se consiguen variables en las que se elimina la divergencia de dispersión entre las diferentes componentes canónicas, puesto que todas ellas tienen un mismo coeficiente de apertura, lo que garantiza una dispersión equivalente, y permitirá comparar los resultados de las eficiencias técnicas obtenidas entre las diferentes especificaciones econométricas utilizadas para cada una de las componentes canónicas analizadas.

Por último, para utilizar las variables en modelos econométricos que expliquen la variabilidad de cada una de las variables de Output en función de cada una de las variables de Input, es conveniente tomar logaritmo neperiano, tanto sobre la variable endógena como sobre la explicativa, para amortiguar el problema de incremento de variabilidad de la varianza de los residuos procedentes de la estimación, al aumentar el valor de la variable explicativa. Es decir, se propone la siguiente transformación:

$$RCCIHF = \text{LN}(RCCIHE) \quad \forall H = 1, 2, \dots, 14$$

$$RCCOHF = \text{LN}(RCCOHE) \quad \forall H = 1, 2, \dots, 14$$

Que son las variables que podrán utilizarse para la estimación de los distintos modelos de regresión, donde cada componente de Input explique el valor de su correspondiente componente de Output.

4. ESTIMACIONES DE FRONTERA ESTOCÁSTICA E INDICES DE EFICIENCIA

Una vez transformadas las variables, estimamos la frontera de referencia a través de modelos estocásticos adaptados al supuesto de datos de panel¹⁵. El modelo de Batesse y Coelli (1992) propone la siguiente estructura en la función de producción:

$$Y_{it} = f(X_{it}; \beta) + E_{it}$$

$$E_{it} = V_{it} - \eta(t) \cdot U_{it}, i=1, 2, \dots, n$$

Donde $f(X; \beta)$ representa la función que queremos ajustar, dependiente de un vector de inputs X y de una tecnología de producción representada por el vector de parámetros β . El subíndice "t" denota el periodo al que pertenecen los datos de input y output. La estructura de los errores es doble; el término V recogería la presencia de errores aleatorios, y sigue una distribución normal $(0, \sigma^2)$. El término U_{it} recogería la ineficiencia técnica de la empresa "i" y sobre el mismo se realizan supuestos de distribuciones diferentes a la normal¹⁶. Para incluir modelos no lineales y este supuesto de división de los errores se acude a la estimación de máxima verosimilitud.

¹⁵ Con las transformaciones planteadas cabría igualmente la aplicación de modelos no paramétricos no sensibles a los cambios de origen en las variables.

¹⁶ Se está considerando que la distribución de probabilidad correspondiente al término ineficiencias (U) se ajusta a una normal truncada en el valor cero en la especificación de Lovell y otros (1982), para evitar que unidades departamentales experimenten ineficiencias negativas. En cambio, Batesse y Coelli (1992) generalizan la anterior especificación, permitiendo que la distribución de las ineficiencias esté truncado en cualquier otro valor positivo (μ).

Los modelos estocásticos para datos de panel modifican el término relativo a los errores¹⁷, haciendo depender el término de ineficiencias U de un parámetro $\eta(t)$ que representa la evolución de las ineficiencias en el tiempo. Dicho parámetro recoge una función concreta¹⁸ que también es objeto de estudio. Una vez estimada la frontera, se obtienen los índices de eficiencia mediante el cálculo de la esperanza del error atribuible a las ineficiencias (U_{it}) condicionada al conjunto de los errores (E_{it})¹⁹.

5. ANÁLISIS DEA SOBRE UN CONJUNTO REDUCIDO DE VARIABLES

Una vez analizada la idoneidad de las variables contrastaremos los resultados con los derivados de aplicar, sobre un conjunto reducido de variables²⁰, el modelo envolvente de datos bajo el supuesto de rendimientos variables de escala²¹. El modelo envolvente supone resolver el siguiente programa matemático:

$$\begin{aligned} \max \quad & \eta + \varepsilon \left[\sum_{i=1}^m s_i^+ + \sum_{r=1}^s s_r^- \right] \\ \text{s.a.} \quad & \sum_{j=1}^n \lambda_j X_{ij} + s_i^+ = X_0 \\ & \sum_{j=1}^n \lambda_j Y_j - s_r^- = \eta Y_0 \\ & \lambda_j \geq 0; \quad j = 1, \dots, n; \quad s_i^+ \geq 0; \quad s_r^- \geq 0 \end{aligned}$$

¹⁷ El modelo original de Aigner, Lovell y Schmidt (1977), Batesse y Corra (1977) y Meeusen y van de Broeck (1977) no permitía el cálculo de un índice que midiese los residuos de eficiencia para cada empresa. Jondrow, Lovell, Materov y Schmidt (1982) solucionaron este problema y los trabajos de Kumbhakar (1990) y Batesse y Coelli (1992, 1995), entre otros, supusieron una extensión para el supuesto de datos de panel y efectos fijos o variables en el tiempo.

¹⁸ En el Working Paper 2000-28 de la Facultad de Ciencias Económicas y Empresariales de la Universidad Complutense de Madrid contrastamos diferentes hipótesis sobre la evolución de las eficiencias.

¹⁹ Una aplicación de modelos de frontera estocástica en el caso de la educación superior puede encontrarse en Cohn, E.; Rhine, S.L.W. y Santos, M.C. (1989) y Dundar, H. y Lewis, D. R. (1995), si bien utilizan una formulación diferente a la aquí presentada.

²⁰ En concreto se consideran las variables e input: EDP (profesores a tiempo completo), ICTT (ingresos por transferencia de tecnología). Como variables de producción científica se toman: PARNOT (Publicaciones periódicas notables), PARCON (intervención en congresos notables), TESLEI (tesis leídas), y PREM (premios científico-técnicos y de arquitecturas).

²¹ Así lo apuntaban los parámetros (suma superior a uno) de las regresiones múltiples efectuadas con cada variable de output como variable dependiente e idénticos inputs como regresores.

donde “ η ” es el índice de eficiencia de una unidad evaluada (\bar{X}_0, \bar{Y}_0) , las lambdas representan los números positivos por los que hay que multiplicar los inputs y outputs de otras unidades para obtener una combinación lineal o frontera de referencia. Las variables de holgura s_i^+ y s_r^- deshacen las desigualdades propuestas en el programa original de Charnes, Cooper y Rodhes (1978). En la relativa a inputs, se establece que la unidad de referencia tiene que utilizar un nivel $\sum_{j=1}^n \lambda_j X_{ij}$ menor o igual de recursos que la unidad evaluada, por lo que la variable de holgura s_i^+ se suma a la izquierda de la desigualdad. En la relativa a los outputs la unidad de referencia $\sum_{j=1}^n \lambda_j Y_j$ tiene que producir un nivel mayor o igual que la unidad evaluada, por lo que la variable de holgura s_r^- se resta para obtener la restricción de igualdad. El infinitésimo ϵ asegura que no exista ninguna unidad con un 100% de eficiencia si sus variables de holgura no son cero.

La versión dual del modelo nos permite analizar la estructura de las ponderaciones del modelo DEA, denotadas con “U” (que actúan como precios virtuales de los outputs) y “V” (precios virtuales de los inputs) respecto a la unidad evaluada “0” y el resto de las “i” unidades de la muestra:

$$\begin{aligned} \min \quad & \sum_{i=1}^m v_i X_{i0} + v_0^* \\ \text{s.a:} \quad & \sum_{r=1}^s U_r Y_{rj} - \sum_{i=1}^m V_i X_{ij} + v_0^* \leq 0 \\ & \sum_{r=1}^s U_r Y_{r0} = 1 \\ & U_r \geq \epsilon; \quad V_i \geq \epsilon; \quad v_0^* \geq \epsilon; \quad j = 1, \dots, \end{aligned}$$

El análisis temporal por medio de ventanas o “windows análisis” fue planteado por Charnes, Clark, Cooper y Golany (1985) y es un modelo idóneo para el estudio de las variaciones de la eficiencia a lo largo del tiempo a partir de un panel de observaciones. Básicamente consiste en realizar las estimaciones del modelo DEA con subpaneles de tiempo, en nuestro caso dos años, y considerar

que los datos de una unidad en el primer periodo corresponden a la unidad-año1 que se unen para el análisis con los de la unidad-año2, es decir los datos de esa misma unidad en el periodo dos. De ese modo se obtienen dos índices de eficiencia. La siguiente frontera hará lo mismo pero renunciando a los datos del primer año e incluyendo los datos del tercer año y así hasta el periodo n analizado (último año del panel) .

El modelo “window analysis” permite introducir más variables en el estudio al multiplicar al número de unidades evaluadas por el número de años que se analizan en cada subpanel o ventana. En ese sentido, da solución al problema de la pérdida de grados de libertad que se produce cuando queremos incorporar variables al estudio y el conjunto de inputs y outputs es relativamente grande con relación al tamaño de la muestra. Otra ventaja de la técnica es el de mostrar las tendencias y la estabilidad de los resultados de eficiencia. También podría servir para identificar unidades atípicas u “outliers”, cuando se produzcan resultados muy variables en el tiempo y oscilen entre índices de eficiencia completamente dispares.

PRESENTACIÓN DE RESULTADOS:

En el cuadro 3 se recogen los índices obtenidos a partir de la aplicación del análisis de eficiencia con el primer par de variables canónicas. Como observamos en el apartado correspondiente a la descripción de los resultados derivados del análisis de correlación canónica, las componentes CCI1 y CCO1 eran las que mejor representaban la producción de calidad en la investigación. Para ello, se ha utilizado una de la función de producción tipo Cobb-Douglas y se han supuesto efectos fijos para las eficiencias²².

²² La función ajustada es $\ln CCO1 = \alpha + \beta \cdot \ln CCI1$. Los parámetros son claramente significativos en concordancia con la naturaleza de las variables canónicas. La estimación se ha realizado en el programa FRONTIER (<http://www.une.edu.au>). El hecho de recoger efectos fijos pone de manifiesto que las ordenaciones de eficiencia no varían en el tiempo y el parámetro η será igual a 1. En un escenario de corto plazo parece razonable pensar en que los gestores no van a poder modificar sus posiciones en la clasificación, por ejemplo, compitiendo con el comportamiento de otras unidades. Los resultados de la estimación son:

Para estudiar la concordancia de los resultados obtenidos con otros procesos de evaluación hemos cruzado los resultados de este análisis con otros dos tipos de evaluación. La primera es la realizada en Trillo (2000), con la muestra reducida de variables de las utilizadas, que, como hemos dicho en el apartado anterior, procede de un análisis envolvente de datos con rendimientos variables²³. Los resultados corresponden a la media de los índices de eficiencia obtenidos en todos los subpaneles. Las clasificaciones que se obtienen con este procedimiento pueden estar sesgadas por el hecho de que los resultados extremos (0% o 100%) afectan mucho a la media, aunque en este caso sólo utilizamos este análisis para contrastar los resultados. Además, el hecho de que no sean exactamente las mismas variables hace que haya que tomar con cautelas sus posibles correlaciones²⁴.

La segunda clasificación corresponde al ratio PAR/EDP (publicaciones ponderadas por los puntos asignados por la UPC en relación al número de profesores a tiempo completo). Dicha clasificación es utilizada por la Universidad en los procesos de asignación de presupuesto para investigación, lo que puede servir para establecer comparaciones con procesos reales de evaluación.

	coefficient	standard-error	t-ratio
beta 0	-13.28	1.17	-11.32
beta 1	1.07	0.07	15.84
sigma-squared	0.45	0.19	2.40
gamma	0.77	0.12	6.39
mu	-1.17	0.91	-1.29

* eta is restricted to be zero
log likelihood function = -0.73835919E+02

LR test of the one-sided error = 0.19651343E+02
with number of restrictions = 2
[note that this statistic has a mixed chi-square distribution]

²³ El análisis se realiza a partir de Banker y otros (1984).

²⁴ Por ejemplo, en el análisis de correlación canónica, al introducirse más variables, ningún departamento presenta niveles extremadamente bajos, mientras que en el análisis de eficiencia realizado con el modelo DEA sí cabía esa posibilidad.

Cuadro 3: Indicadores de eficiencia y productividad científica

DPTO	EFDEA	EFCC1	PAREDP
1	75.77%	87.69%	11.28
2	72.14%	90.69%	23.01
3	66.69%	89.87%	19.43
4	38.77%	83.28%	22.53
5	33.02%	36.31%	1.55
6	71.62%	92.84%	28.82
7	93.60%	88.01%	18.54
8	70.94%	88.40%	26.55
9	42.79%	90.23%	5.62
10	88.22%	93.65%	12.33
11	80.81%	88.13%	17.59
12	28.59%	80.92%	7.00
13	72.38%	91.40%	17.98
14	21.57%	85.88%	13.29
15	26.96%	88.76%	12.54
16	82.07%	76.33%	8.24
17	14.14%	57.97%	1.57
18	68.62%	67.65%	3.05
19	0.00%	55.54%	0.17
20	51.60%	90.32%	15.64
21	64.59%	87.29%	12.79
22	97.71%	90.83%	29.80
23	63.80%	88.27%	9.68
24	48.07%	93.17%	12.19
25	62.78%	72.23%	10.23
26	84.39%	80.40%	8.65
27	74.20%	92.29%	10.45
28	81.45%	88.89%	8.51
29	90.43%	93.50%	10.13
30	42.39%	84.17%	9.03
31	41.09%	85.49%	7.76
32	71.44%	86.51%	19.44
33	47.12%	88.52%	27.00
34	72.91%	86.97%	21.08
35	71.10%	86.39%	12.33
36	46.69%	70.03%	12.10
37	39.88%	78.51%	19.37
38	30.92%	56.42%	4.44

Correlaciones

		EFCC1G	EFDEA	PAREDP
EFCC1G	Correlación de Pearson	1,000	,544**	,595**
	Sig. (bilateral)	,	,000	,000
	N	38	38	38
EFDEA	Correlación de Pearson	,544**	1,000	,429**
	Sig. (bilateral)	,000	,	,007
	N	38	38	38
PAREDP	Correlación de Pearson	,595**	,429**	1,000
	Sig. (bilateral)	,000	,007	,
	N	38	38	38

** . La correlación es significativa al nivel 0,01 (bilateral).

CONCLUSIONES PRELIMINARES Y REFLEXIONES FINALES:

Mediante el presente trabajo se ha intentado resaltar las posibilidades que ofrece el análisis de correlación canónica para el estudio y selección de las variables más relevantes en el proceso de producción de instituciones privadas o públicas, así como proponer un método para la utilización de las variables canónicas en los análisis de eficiencia.

En cuanto al primer aspecto, hemos observado que el estudio de las correlaciones de las variables originales de input y output con las diferentes variables canónicas puede aportar una importante información sobre las relaciones entre las variables y los agrupaciones entre variables dentro de cada subconjunto de inputs o outputs.

En concreto, para el supuesto de la Universidad Politécnica de Cataluña, se puede comprobar a través de la primera componente canónica la relación entre las publicaciones de calidad y los inputs conducentes a las mismas. Ello puede ser útil en dos sentidos, primero para conocer el grado de asociación natural o estadística en un análisis previo al de regresión. En segundo lugar, desde el punto de vista del planificador o evaluador puede ser interesante conocer si los sistemas de seguimiento y control de resultados tienen consecuencias sobre las agrupaciones entre variables. Esto podría observarse si se realizase el análisis de correlación canónica al principio y al final de un determinado periodo de evaluación. En el caso de la UPC, la implantación de la Planificación estratégica y la obtención de financiación condicionada a la obtención de resultados por parte de los departamentos, podrían explicar por qué las variables de calidad están representados por el par de variables que explican un porcentaje mayor de la varianza total acumulada.

Otra posible utilidad de este análisis sería como método para establecer o revisar las ponderaciones asignadas a cada variable de producción en los mecanismos de evaluación. Piénsese por ejemplo en el indicador de publicaciones ponderadas en relación al número de investigadores. Igualmente podría servir para imponer, sobre una base objetiva, restricciones a las ponderaciones en el análisis envolvente de datos cuando los resultados

derivados de la asignación libre de los inputs y outputs virtuales no concuerden con el objetivo o plan de evaluación deseado. Un ejercicio similar se realiza en Friedman, L. y Sinuany-Stern, Z. (1997) para obtener una ordenación completa de todas las unidades evaluadas.

La aplicación de modelos de eficiencia a partir de las variables resultantes del análisis de correlación canónica plantea dificultades que hemos intentado solventar mediante diferentes transformaciones sobre las variables. La concordancia entre los resultados obtenidos y los derivados del análisis de indicadores de productividad y de un modelo DEA realizado a partir de una muestra reducida de las variables reales originales, apunta a que pueden realizarse análisis de eficiencia con variables canónicas, con todas las cautelas derivadas de un método poco experimentado; y a que dicho método podría introducir nuevos elementos de reflexión, además de solucionar parte de los problemas de ajuste que tradicionalmente se plantean en las estimaciones de los modelos de frontera. El hecho de reducir al mínimo las variables del modelo, aunque se encuentren representadas todas las variables originales, disminuye las posibilidades de que las estimaciones de los parámetros sean ineficientes y simplifica enormemente la estructura de producción, lo que hace innecesario la utilización de funciones de producción más flexibles en las que se establecen términos donde se multiplican los inputs entre sí.

Seguimos cuestionando si las propuestas de transformación de las variables son las que mejor responden al objetivo de la obtención de componentes que, resumiendo toda la información procedente de las variables reales originales, son las que mejor reflejan el comportamiento de eficiencia de las instituciones objeto de evaluación. Estudiando la correlación con otros análisis parecidos realizados con parte de la muestra de indicadores utilizados únicamente obtenemos indicios de que los resultados son coherentes. Sería preciso establecer la misma base de comparación, realizando modelos exactamente con las mismas variables, si bien esto implicaría reducir el tamaño de la base de datos. Sirva este primer análisis de muestra para plantear el problema y mejorar el análisis con los comentarios que podamos incorporar en el futuro.

BIBLIOGRAFÍA

Aigner, D.; Lovell, C. K. y Schmidt, P. J. (1977): Formulation and estimation of stochastic frontier production models. *Journal of Econometrics* nº 6, pp. 21-37.

Banker, R. D., Charnes, A. y Cooper, w. W. (1984): Models for estimation of technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30. PP. 1078-92

Batesse, G. E. y Coelli, T. J. (1992): Frontier Production Functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of Productivity Analysis*, nº 3, pp. 153-169.

Battese, G. E. y Coelli, T. J. (1995): A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* nº 20, pp. 325-332.

Charnes, A., Cooper, D. B. y Rhodes, E. (1978): Measuring Efficiency of Decision Making Units. *European Journal of Operational Research* nº 1. pp. 429-444.

Cohn, E.; Rhine, S.L.W. y Santos, M.C. (1989): Institutions of higher education as multi-product firms: economies of scale and scope. *Review of Economics and Statistics* nº 71, pp. 284-290.

Dundar, H. y Lewis, D. R. (1995): Departmental productivity in American universities Economies of scale and scope. *Economics of Education Review* vol. 14 nº 2, pp. 119-144.

Friedman, L. y Sinuany-Stern, Z. (1997): Scaling units via the canonical correlation analysis in the DEA context. *European Journal of Operational Research* nº 100, pp. 629-637.

Jondrow, J.; Lovell, C. A. K.; Materov, S. y Schmidt, P. (1977): On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics* 19, pp. 233-238.

Kumbhakar, S. C. (1990): Production frontiers, panel data, and time-varying technical inefficiency. *Journal of Econometrics*, 46, pp. 201-211

Martín-Guzman, P. y Martín Pliego, J. (1993): "Curso básico de estadística económica". AC. Madrid.

Meeusen, W. y Van Den Broeck, J. (1977): Efficiency estimation from Cobb-Douglas productions functions with composed error. *International Economic Review*, vol. 18, nº 2 june, pp. 435-444.

Novalés, A. (1990): *Econometría*. McGraw-Hill. Madrid.

Trillo, D. (2000): Modelos dinámicos de medición de la eficiencia: una comparación de métodos en educación superior. VIII Encuentro de Economía Pública. Universidad de Extremadura. Febrero de 2001

Universitat Politècnica de Catalunya (1994-2000): Dades estadístiques i de gestió. Oficina Tècnica de Programació.

ANEXO ESTADÍSTICO:

1. EL ANÁLISIS DE CORRELACIÓN CANÓNICA

Canonical Correlations

Num.	Eigenvalue	Canonical	Wilks	Chi-Square	D.F.	Correlation	Lambda	P-Value
1	0.798913	0.893819	0.00447315	938.577	224	0.0000		
2	0.682908	0.826382	0.0222448	660.28	195	0.0000		
3	0.561019	0.749012	0.0701525	461.004	168	0.0000		
4	0.401052	0.633287	0.159808	318.162	143	0.0000		
5	0.326461	0.571368	0.266814	229.229	120	0.0000		
6	0.262829	0.512669	0.396137	160.66	99	0.0001		
7	0.208281	0.456378	0.537376	107.754	80	0.0210		
8	0.136063	0.368867	0.678745	67.233	63	0.3343		
9	0.104758	0.323663	0.785641	41.8578	48	0.7214		
10	0.060681	0.246335	0.877574	22.6581	35	0.9467		
11	0.0344942	0.185726	0.934266	11.797	24	0.9821		
12	0.0215635	0.146845	0.967644	5.70659	15	0.9843		
13	0.00840809	0.0916956	0.98897	1.9244	8	0.9832		
14	0.00264448	0.0514245	0.997356	0.459425	3	0.9277		

Coefficientes de las Correlaciones Canónicas del primer conjunto

PARCON	0.164034	-0.629579	0.669275	-0.600797	0.0523427	-0.133126	
0.348709	-0.0630052	-0.125789	0.257458	-0.887981	-0.065		
PARNOT	0.105682	0.198431	-0.525863	-0.634579	-0.175838	-0.0489568	
-0.465244	0.326293	0.624872	0.132527	0.437695	-0.523		
ARTA	0.193975	0.0657907	0.0724398	0.366717	0.541249	0.465769	
0.197949	-0.616553	0.533765	0.196843	-0.418187	-0.108		
ARTB	0.126115	0.165764	-0.26198	0.17866	-0.0899752	-0.186522	-
0.148789	0.079138	-0.130314	0.444009	0.339424	-0.381		
ARTC	0.0240243	0.137155	-0.0942979	0.276033	0.0324643	-0.623428	-
0.19647	0.372178	0.146185	-0.205887	-0.411773	-0.090		
COMNAC	0.0162625	-0.158596	-0.149187	0.433513	-0.0474308	0.16373	
-0.486046	-0.382829	-0.0904375	0.344012	-0.0691647	-0.176		
COMINT	0.129084	-0.0890029	-0.10614	0.161979	0.256514	-0.279945	
-0.501694	-0.161745	-0.220569	-0.511314	0.638305	0.4686		
POSTN	-0.074814	-0.0785061	0.0555459	0.223831	-0.27659	-0.148838	
0.810246	0.217366	-0.364137	0.506474	-0.195688	0.6247		
POSTI	0.276689	0.22296	-0.295433	-0.305381	0.236726	0.369099	-
0.237949	0.0854171	-0.508115	-0.333717	-0.483186	0.1774		
PREMN	0.150622	0.222532	0.64861	0.185553	0.211939	0.433955	-
0.301449	0.392308	0.256829	-0.428133	0.0321782	0.2939		
PREMI	-0.0548803	0.281973	-0.00387285	-0.240588	0.220518	-0.316564	
0.0951868	-0.0761568	-0.455271	0.619604	0.0680365	-0.213		
LIBRN	0.0673554	0.0589616	0.206961	0.150982	-0.640974	0.361347	-
0.0522123	-0.173049	-0.104641	-0.107756	-0.38538	-0.248		
LIBRI	-0.0847001	0.121024	-0.0252277	-0.174066	-0.346991	-0.268707	-
0.263282	-0.167361	0.166828	0.335806	-0.129594	0.8094		
CAPNAC	0.0306358	0.217097	0.0604236	-0.222338	-0.00926539	0.152862	
0.177793	-0.415432	-0.380728	0.107009	0.293795	-0.275		
CAPINT	0.0783872	0.0809721	0.0451086	-0.00354375	0.0656918	-0.598327	
0.30622	-0.227085	-0.123487	-0.463533	0.034649	-0.088		
TESLEI	0.486697	0.0651443	0.106217	0.458933	-0.497046	0.15847	
0.441026	0.580106	0.243397	-0.152189	0.460053	0.0565		

Coefficientes de las Correlaciones Canónicas del segundo conjunto

ICTTEUR	-0.209066	-0.473475	0.21881	-0.534548	0.0722523	
0.0000510925	0.18824	0.0167167	0.542784	-0.359023	-0.413379	0.3610
ICTCICYT	0.10704	-0.179236	0.164775	-0.20002	0.195106	-0.139449
0.310017	0.648149	-0.770923	0.950832	-0.0594823	0.3042	
ICTTCBS	0.00526497	-0.25458	0.0294974	0.0743074	-0.553062	0.0953008
0.296912	0.449642	0.482819	-0.397495	0.588782	0.0891	
ICTTSUBV	0.0706622	0.00519979	-0.0272944	0.0879845	-0.114681	-
0.0814042	-0.0636281	0.607603	-0.64687	0.518942	0.0626022	-0.360
C2INV	0.568828	0.659296	-0.420456	0.604569	0.37081	0.502172
0.556021	-0.183295	-0.862643	-0.836973	0.890012	0.6283	-
CEU	0.286225	0.182049	-0.659619	0.055617	-0.130395	-0.44525
0.157386	0.113304	0.540247	0.455308	-0.388747	0.2388	-
CU	-0.0175133	-0.148245	-0.053273	0.0108526	-1.02833	0.208686
0.503193	-0.704101	-0.104011	0.0422534	-0.733845	-0.382	
TEU	-0.0149668	-0.174552	0.056068	-0.0400833	-0.211819	0.527641
0.717148	0.352915	-0.113009	-0.773894	-0.717139	-0.243	
TU	0.323282	0.149109	-0.0071002	-0.859948	0.33731	0.026406
0.666997	0.368086	0.307347	-0.0423774	-0.0633774	-0.687	-
ASTC	0.315186	-0.806788	-0.162029	0.818514	-0.060812	-0.840518
0.219603	0.024668	0.268461	-0.593731	-0.0104337	-0.316	-
EDPBEC	0.180654	0.153488	-0.0348196	0.0941915	0.416052	0.019898
0.725624	-0.547991	0.514394	0.5811	0.210323	-0.713	
ESTLIC	-0.350234	-0.0641428	0.473768	-0.25988	0.0729289	0.871867
0.397392	-0.434279	-0.0829855	1.09255	0.475578	0.6103	-
ESTDOCT	0.101319	0.223563	0.0142328	-0.738027	-0.0505261	-0.834875
0.52136	-1.04969	-0.508006	0.0356848	-0.0437472	0.7850	
CREDOCT	-0.043902	0.254006	0.645468	0.793212	-0.239029	0.163844
-0.411571	1.41319	0.727783	-0.330887	-0.792993	-0.513	

NOTA:

Las combinaciones lineales de los dos conjuntos de variables que tienen una correlación más alta entre sí. Pueden formarse catorce grupos. El primero se formaría utilizando la siguiente expresión (que incluye las ponderaciones de cada variable real original):

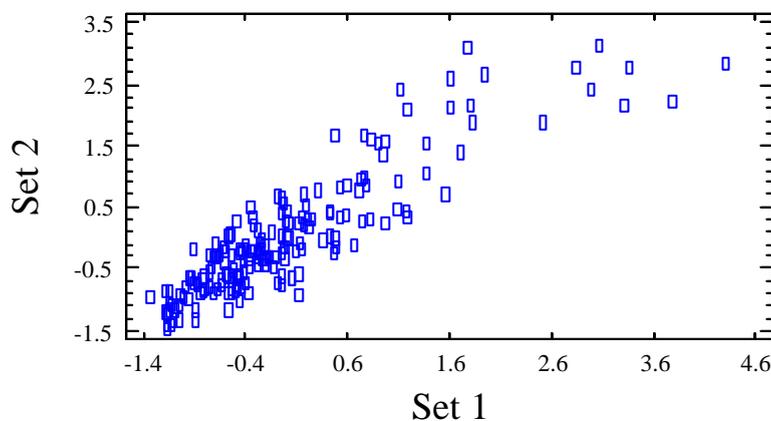
$$0.164034*PARCON + 0.105682*PARNOT + 0.193975*ARTA + 0.126115*ARTB + \\ 0.0240243*ARTC + 0.0162625*COMNAC + 0.129084*COMINT - 0.074814*POSTN \\ +0.276689*POSTI + 0.150622*PREMN - 0.0548803*PREMI + 0.0673554*LIBRN - \\ 0.0847001*LIBRI + 0.0306358*CAPNAC + 0.0783872*CAPINT + 0.486697*TESLEI$$

y

$$-0.209066*ICTTEUR + 0.10704*ICTCICYT + 0.00526497*ICTTCBS + \\ 0.0706622*ICTTSUBV + 0.568828*C2INV + 0.286225*CEU - 0.0175133*CU - \\ 0.0149668*TEU + 0.323282*TU + 0.315186*ASTC + 0.180654*EDPBEC - \\ 0.350234*ESTLIC + 0.101319*ESTDOCT - 0.043902*CREDOCT$$

donde las variables se han tipificado por su media y dividiendo por su desviación típica. Siete conjuntos tienen un P-valor menor que un 0.05, siendo sus correlaciones estadísticamente significativos al 95%.

Plot of Canonical Variables #1



2. EL ANÁLISIS DEA:

DMU	1/w	EDP (I)(V)	ICTT (I)(V)	CREDOCT(I)(V)	PARNOT (O)(V)	PARCON (O)(V)	TESLEI(O)(V)	PREM (O)(V)
D195	40.14%	0.84	0.16	0	0	1	0	0
D295	57.74%	0.3	0.1	0.59	0	0.3	0.14	0.56
D395	48.01%	0.52	0.48	0	0	0	0.85	0.15
D495	44.75%	0.55	0	0.45	0	0	0.49	0.51
D595	0.00%	1	0	0	0	0.14	0.19	0.67
D695	38.68%	0.64	0.36	0	0	0.34	0.58	0.09
D795	71.02%	0.31	0.1	0.59	0	0.54	0	0.46
D895	70.18%	0.96	0.04	0	0	0.29	0.52	0.19
D995	100.00%	0.37	0.63	0	0	0.97	0.03	0
D1095	32.33%	0	0.5	0.5	0	0.65	0.35	0
D1195	80.62%	1	0	0	0	0.91	0	0.09
D1295	33.70%	0.44	0.51	0.05	0	0	1	0
D1395	75.16%	0.32	0	0.68	0	0	0.72	0.28
D1495	48.59%	0.57	0.03	0.4	0	0	0.29	0.71
D1595	48.84%	0.41	0.59	0	0	0	1	0
D1695	49.18%	0.61	0.08	0.31	0	0	1	0
D1795	0.00%	0.57	0.27	0.16	0	0.04	0	0.96
D1895	28.90%	0.45	0.14	0.42	0	0	0	1
D1995	100.00%	0.94	0.06	0	0	0.68	0.04	0.28
D2095	27.38%	0.42	0.58	0	0.56	0.03	0.41	0
D2195	39.20%	0	1	0	1	0	0	0
D2295	100.00%	1	0	0	0	0.58	0.42	0
D2395	42.30%	0	1	0	0	0	1	0
D2495	36.74%	0.43	0	0.57	0.06	0	0.22	0.72
D2595	100.00%	0.66	0.34	0	0.94	0.03	0.03	0
D2695	100.00%	0.98	0.01	0.01	0	0.02	0	0.98
D2795	100.00%	0.17	0.83	0	0	1	0	0
D2895	42.57%	0	0.81	0.19	0	0	0.86	0.14
D2995	100.00%	0.44	0.27	0.29	0	0	0.9	0.09
D3095	5.85%	0.33	0.27	0.39	1	0	0	0
D3195	48.74%	1	0	0	0	0	0.95	0.05
D3295	100.00%	0.23	0.77	0	0	0	0	1
D3395	0.00%	0.34	0	0.66	0	0	0.44	0.56
D3495	35.87%	0.35	0.28	0.37	0	0.58	0.18	0.24
D3595	61.05%	1	0	0	0	0	0.91	0.09
D3695	100.00%	0.45	0.13	0.43	0	0	0	1
D3795	100.00%	0.47	0.2	0.33	0.05	0.46	0	0.5
D3895	100.00%	0.83	0.17	0	0	0.31	0.5	0.19
D196	67.12%	1	0	0	0	0.52	0.43	0.06
D296	64.26%	0.31	0	0.69	0.06	0.17	0.28	0.5
D396	36.01%	0.44	0.56	0	0	0	1	0
D496	52.31%	1	0	0	0	0	0.8	0.2
D596	100.00%	0.8	0.2	0	0	0	0	1
D696	42.17%	0.29	0	0.71	0.07	0.09	0.33	0.5
D796	74.38%	0.33	0.13	0.54	0	0.62	0	0.38
D896	54.36%	1	0	0	0	0.33	0.58	0.09
D996	49.84%	0.47	0.2	0.34	0	0.55	0.2	0.26
D1096	52.90%	0	0.15	0.85	0	0.64	0	0.36
D1196	80.93%	1	0	0	0	0.47	0.53	0
D1296	11.89%	0.34	0	0.66	0.08	0	0	0.92
D1396	54.27%	0.83	0	0.17	0.22	0	0.61	0.18
D1496	22.89%	0.4	0.03	0.57	0	0	0.45	0.55
D1596	28.00%	0.33	0	0.67	0	0	0.54	0.46
D1696	100.00%	0.43	0.28	0.29	0	1	0	0
D1796	0.00%	0.3	0.42	0.28	0	0	0.63	0.37
D1896	23.02%	0	1	0	0	0	0.86	0.14
D1996	0.00%	0.93	0.07	0	0	0.26	0	0.74
D2096	32.43%	0.35	0.49	0.15	0.56	0.06	0.18	0.2
D2196	32.99%	0	0.8	0.2	0.85	0.01	0	0.14
D2296	100.00%	0.38	0	0.62	0	0.05	0.47	0.47
D2396	25.46%	0	1	0	0	0.8	0.2	0
D2496	52.28%	0.3	0	0.7	0	0.08	0.4	0.52
D2596	15.57%	0.95	0.05	0	1	0	0	0

DMU	1/w	EDP (I)(V)	ICTT (I)(V)	CREDOCT(I)(V)	PARNOT (O)(V)	PARCON (O)(V)	TESLEI (O)(V)	PREM (O)(V)
D2696	37.54%	0.24	0	0.76	0.28	0	0.41	0.31
D2796	58.49%	0.45	0.55	0	0.03	0.27	0.67	0.03
D2896	45.78%	0	0.74	0.26	0	0	0.77	0.23
D2996	100.00%	0.47	0.3	0.23	0	0.02	0.98	0
D3096	62.26%	0.47	0	0.53	0	0	0	1
D3196	6.54%	1	0	0	0	0	1	0
D3296	66.79%	0	0	1	0	0	0.04	0.96
D3396	24.79%	0.16	0.02	0.82	0.01	0.13	0	0.86
D3496	63.38%	0.46	0.39	0.15	0.07	0.06	0.87	0
D3596	44.85%	0	0	1	0	0.37	0.37	0.26
D3696	47.62%	0.35	0.13	0.52	0	0	0	1
D3796	3.70%	0.88	0.12	0	1	0	0	0
D3896	0.00%	0.56	0.1	0.34	0	0.05	0.56	0.39
D197	60.33%	0.8	0.09	0.11	0	0.64	0.27	0.1
D297	100.00%	0.42	0.5	0.08	0	0	0.91	0.08
D397	53.97%	0.49	0.51	0	0.03	0	0.97	0
D497	11.71%	0.42	0.58	0	0	0	1	0
D597	0.00%	1	0	0	0	0.14	0.19	0.67
D697	70.17%	1	0	0	0	0	0.78	0.22
D797	100.00%	0.71	0.29	0	0	0.75	0.21	0.04
D897	81.51%	1	0	0	0	0.29	0.6	0.11
D997	31.95%	0.43	0.2	0.37	0	0.42	0.31	0.27
D1097	100.00%	0	0.22	0.78	0	1	0	0
D1197	87.28%	1	0	0	0	0	1	0
D1297	1.86%	0.16	0.28	0.56	1	0	0	0
D1397	66.41%	0.41	0	0.59	0.1	0	0.76	0.13
D1497	3.22%	0.77	0.23	0	1	0	0	0
D1597	20.28%	0.69	0	0.31	0	0	0.76	0.24
D1697	37.44%	0.35	0.42	0.23	0.06	0.94	0	0
D1797	0.00%	0.3	0.36	0.33	0	0	0.63	0.37
D1897	63.32%	0.55	0.16	0.28	0	0	0	1
D1997	0.00%	0.93	0.07	0	0	0.26	0	0.74
D2097	48.83%	0.59	0.41	0	0.62	0.32	0.06	0
D2197	77.24%	0	1	0	0.5	0	0.5	0
D2297	100.00%	0.48	0.52	0	0	1	0	0
D2397	59.28%	0	1	0	0	0	1	0
D2497	26.35%	0.41	0.52	0.07	0.09	0	0.91	0
D2597	100.00%	0.84	0.16	0	0.54	0.05	0.41	0.01
D2697	100.00%	0	1	0	0.63	0	0.36	0.01
D2797	86.39%	0.62	0.38	0	0	0	1	0
D2897	85.03%	0	0	1	0	0	0.71	0.29
D2997	100.00%	0.44	0	0.56	0	0	0	1
D3097	40.44%	0.62	0	0.38	0.13	0	0	0.87
D3197	40.82%	1	0	0	0	0	0.94	0.06
D3297	95.64%	0	0	1	0.02	0	0	0.98
D3397	39.58%	0.24	0	0.76	0	0	0.44	0.56
D3497	61.72%	0.44	0.39	0.17	0.09	0.41	0.42	0.08
D3597	79.24%	1	0	0	0	0.41	0.52	0.06
D3697	50.66%	0.42	0.2	0.38	0	0	0.37	0.63
D3797	78.37%	0.85	0.15	0	0.42	0.04	0	0.54
D3897	35.98%	0.52	0.24	0.24	0	0	0.66	0.34
D198	75.63%	0.69	0.14	0.17	0	0.78	0.18	0.04
D298	24.32%	0.31	0	0.69	0.34	0	0.28	0.38
D398	91.15%	0.66	0.34	0	0	0	0.88	0.12
D498	36.88%	0.5	0.5	0	0	0	0.89	0.11
D598	0.00%	1	0	0	0	0.14	0.19	0.67
D698	74.15%	0.38	0	0.62	0.05	0.05	0.63	0.27
D798	100.00%	0.63	0.37	0	0	0.27	0.71	0.01
D898	62.28%	1	0	0	0	0	0.95	0.05
D998	42.55%	0.6	0.4	0	0	0	1	0
D1098	100.00%	0	0.72	0.28	0	0.57	0.29	0.14
D1198	55.03%	0.29	0	0.71	0	0	1	0
D1298	42.35%	0.51	0	0.49	0	0	0.44	0.56

DMU	1/w	EDP (I)(V)	ICTT (I)(V)	CREDOCT(I)(V)	PARNOT (O)(V)	PARCON (O)(V)	TESLEI (O)(V)	PREM (O)(V)
D1398	68.84%	0.33	0	0.67	0.2	0	0.53	0.27
D1498	31.91%	0.49	0.32	0.19	0	0	1	0
D1598	21.46%	0.84	0.16	0	0	0.38	0.62	0
D1698	90.86%	0.52	0.28	0.2	0.12	0.88	0	0
D1798	56.58%	0.94	0.06	0	0	0	0	1
D1898	88.15%	0	0.89	0.11	0	0.33	0.46	0.21
D1998	0.00%	0.92	0.08	0	0	0.26	0	0.74
D2098	32.30%	0.51	0.49	0	0.36	0	0.64	0
D2198	48.13%	0	1	0	0.49	0	0.49	0.03
D2298	90.86%	0.98	0	0.02	0	0	1	0
D2398	70.47%	0	1	0	0	0	1	0
D2498	45.97%	0.37	0.4	0.23	0	0	1	0
D2598	35.54%	1	0	0	0.47	0	0.53	0
D2698	100.00%	0.32	0.68	0	0	1	0	0
D2798	51.93%	0.08	0.01	0.91	0.1	0	0.9	0
D2898	95.01%	0	0.13	0.87	0	0.77	0.17	0.05
D2998	73.82%	0.42	0.46	0.12	0	0.06	0.81	0.12
D3098	32.24%	0.7	0.27	0.03	0	0	1	0
D3198	53.93%	1	0	0	0	0	1	0
D3298	37.57%	0.49	0	0.51	0.04	0	0.12	0.84
D3398	88.73%	0.6	0.4	0	0	0.08	0.8	0.12
D3498	66.55%	0.4	0.23	0.37	0.09	0.29	0.2	0.42
D3598	60.34%	0	0	1	0	0.28	0.37	0.34
D3698	43.99%	0.66	0.34	0	0	0	0.84	0.16
D3798	20.58%	0.53	0.44	0.04	0	0	0.81	0.19
D3898	39.32%	0.5	0.33	0.16	0	0	1	0
D199	100.00%	0.92	0.08	0	0	0.73	0.27	0
D299	100.00%	1	0	0	0.77	0.17	0	0.06
D399	85.65%	0.61	0.39	0	0	0	0.94	0.06
D499	54.18%	0.58	0	0.42	0.11	0	0.32	0.57
D599	32.06%	1	0	0	1	0	0	0
D699	100.00%	0.28	0.11	0.61	0.16	0	0.56	0.28
D799	100.00%	0.33	0.31	0.35	0.11	0.89	0	0
D899	85.62%	1	0	0	0.5	0	0.45	0.05
D999	46.82%	0.57	0.43	0	0.17	0	0.83	0
D1099	100.00%	0	0	1	0.11	0	0.89	0
D1199	100.00%	0.71	0	0.29	0.14	0	0.86	0
D1299	58.26%	0.51	0	0.49	0.16	0	0.23	0.61
D1399	100.00%	0	1	0	0.96	0	0	0.04
D1499	28.27%	0.44	0.45	0.11	0.22	0	0.78	0
D1599	38.10%	1	0	0	0.29	0.57	0.15	0
D1699	100.00%	0.55	0.15	0.3	0.15	0	0	0.85
D1799	0.00%	1	0	0	0	0.14	0.19	0.67
D1899	100.00%	0.32	0.44	0.24	0.04	0.54	0	0.42
D1999	0.00%	0.88	0.12	0	0	0.18	0	0.82
D2099	92.82%	0.15	0.85	0	0.87	0.07	0.01	0.05
D2199	100.00%	0	1	0	0.77	0	0.23	0
D2299	100.00%	1	0	0	1	0	0	0
D2399	100.00%	0	1	0	0.43	0.57	0	0
D2499	67.65%	0.22	0	0.78	0.16	0	0.84	0
D2599	100.00%	0.92	0.08	0	1	0	0	0
D2699	100.00%	0.54	0.27	0.19	0.3	0.55	0	0.16
D2799	100.00%	1	0	0	0.33	0.42	0.26	0
D2899	100.00%	0	0.05	0.95	0.3	0.08	0.14	0.48
D2999	87.91%	0.39	0.5	0.1	0	0	1	0
D3099	34.62%	0.53	0.14	0.33	0.57	0	0	0.43
D3199	63.07%	1	0	0	0	0	0.87	0.13
D3299	85.79%	0.4	0	0.6	0.13	0	0	0.87
D3399	35.39%	0.15	0	0.85	0.15	0	0	0.85
D3499	100.00%	0.59	0.3	0.11	0.86	0	0.14	0
D3599	100.00%	0	0	1	0.17	0.67	0	0.17
D3699	44.50%	0.54	0.46	0	0	0	0.75	0.25
D3799	56.88%	0.88	0.12	0	0.8	0	0	0.2
D3899	48.39%	0.55	0.3	0.16	0.08	0	0.77	0.15