

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

**Departamento de Economía Aplicada VI
(Hacienda Pública y Sistema Fiscal)**



**EFICIENCIA TÉCNICA Y REDES NEURONALES: UN
MODELO PARA EL CÁLCULO DEL VALOR AÑADIDO EN
EDUCACIÓN**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Daniel Santín González

Bajo la dirección de la doctora

Aurelia Valiño Castro

Madrid, 2003

ISBN: 84-669-2277-6

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES
Departamento de Economía Aplicada VI (Hacienda Pública y Sistema Fiscal)



**EFICIENCIA TÉCNICA Y REDES NEURONALES:
UN MODELO PARA EL CÁLCULO DEL VALOR
AÑADIDO EN EDUCACIÓN**

Tesis Doctoral realizada por Daniel Santín González,
bajo la dirección de la profesora Aurelia Valiño Castro,
Titular de Universidad de la Universidad Complutense de Madrid

Madrid, Enero de 2003

A mis padres, hermanos y a Ana, por todo

Agradecimientos

Podríamos clasificar a la gente que nos rodea en dos grupos: los que aportan y los que no aportan valor añadido a nuestras vidas. Esta tesis está dedicada a todas las personas que han aportado “valor añadido” a mi formación como persona y como investigador.

“Estudia, estudia y estudia”, éstas han sido las palabras más repetidas por mis padres desde que tengo uso de razón y por las que nunca les podré estar del todo agradecido. En todas las etapas educativas, mis padres siempre me han motivado y me han brindado con mucho esfuerzo y cariño todos los recursos que he necesitado. Además de eso os agradezco que siempre me hayáis inculcado valores para toda una vida. Fundamentalmente me habéis enseñado a saber apreciar todo lo que tengo y a que el trabajo y el esfuerzo siempre son recompensados. A vosotros y a mis hermanos, Abel y Diana, os quiero agradecer también el haberme aguantado los altibajos de ánimo por los que he pasado al elaborar esta Tesis. Sinceras gracias además para mi hermano Abel, ya que me ha proporcionado en todo momento el apoyo informático que he necesitado para elaborar este trabajo. Muy especialmente agradezco a mi novia Ana su continuo cariño y todo el valor añadido que me ha aportado en tantos años. Gracias porque siempre has estado a mi lado para alentarme. Nadie como tú sabe lo duro que ha sido realizar esta tesis doctoral, juntos lo recordaremos siempre.

En el ámbito académico muchas han sido las personas que han contribuido a mi formación como investigador. Quisiera agradecer a José Manuel González-Páramo su motivación como docente porque él fue la primera persona que despertó mi interés por la Hacienda Pública. La claridad de sus exposiciones en sus clases de licenciatura influyeron decisivamente en mi elección del Doctorado en Economía Pública. Gracias además, José Manuel, por las distintas revisiones y continua supervisión que has realizado de este trabajo.

Especiales gracias para mi directora de tesis, y amiga, Aurelia Valiño, quien desde el principio apostó por mi interés en economía de la educación. Gracias, Aurelia, por todas las horas de reuniones, discusiones, propuestas, búsquedas de material, lecturas de borradores y esfuerzo que día a día y durante estos años has dedicado a este trabajo que también es el tuyo. Sin tu guía, ayuda, planificación de tiempos y tareas, y por que no decirlo, sin el ánimo y entusiasmo que en todo momento me has transmitido, esta tesis hubiera sido prácticamente imposible de realizar. El valor añadido que has aportado a mi formación no tiene precio.

Gracias a M^a Jesús San Segundo, porque en las primeras etapas de doctorado me proporcionó todo su apoyo tanto investigador como material. El ejemplo que me has brindado al ofrecerme ayuda desinteresada, ha sido una de las lecciones que más he valorado. Desafortunadamente en este mundo la ayuda que se pide no siempre se recibe. Algún día me gustaría hacer lo mismo que tú con quien lo necesite. Gracias a Caroline M. Hoxby por todos sus comentarios y aportaciones. Ella me facilitó además el acceso a todos los recursos físicos y humanos del Departamento de Economía de la Universidad de Harvard y del National Bureau of Economic Research. Su gran valía como persona y como investigadora en materia de economía de la educación ha sido decisiva para la obtención de los resultados presentados.

Gracias a todos mis compañeros del Departamento de Economía Aplicada VI de la Universidad Complutense de Madrid que están conmigo a diario, Alfonso, Rosa, María, Raquel, Mercedes, Paco, Carmen, Miguel, Javier, Carola, Fernando y Valentín. Mención especial merecen: Emilio Albi, por sus consejos y su constante preocupación por el futuro de la generación más joven. Gracias además, Emilio, por haber revisado las últimas versiones de este trabajo aportando valiosas sugerencias y comentarios que dotan de mayor relevancia y sentido el resultado final. Quiero agradecer a Laura de Pablos la oportunidad que me brindó de poder trabajar a su lado, además de su apoyo como directora de Departamento. Gracias a Teresa López, por el doble apoyo recibido como compañera y como decana de la Facultad de CC. Económicas de la UCM. E infinitas gracias a Nuria Badenes, quién durante estos años ha sido una excelente compañera de despacho. Gracias, Nuria, por tus palabras de constante ánimo en los momentos en que más las he necesitado.

Gracias también a David Cantarero y a Anna García Altés por ser los grandes amigos que sois y por saber que siempre estaréis ahí. Gracias también a mi familia en Guatemala y a mis amigos chilenos Irma y Juan Carlos por todo vuestro cariño. Agradezco además a mis compañeros y amigos, Rafaela Dios y Francisco Delgado, tanto los comentarios recibidos como los constantes ánimos que siempre me habéis brindado.

Querría agradecer el apoyo económico del Instituto de Estudios Fiscales, especialmente a Juan José Rubio, Javier Salinas y a José Félix Sanz, que confiaron en mi proyecto. Por último, agradecer a la Universidad Complutense de Madrid, la financiación que he recibido en estos años a través de su programa de Formación de Personal Investigador, ayuda sin la cual, esta tesis nunca hubiera visto la luz.

There are few practical problems in which the economist has a more direct interest than those relating to the principles on which the expense of the education of children should be divided between the State and the parents.

(Alfred Marshall, 1890 *Principles of Economics*. Book IV, Chapter VI.)

TESIS DOCTORAL

EFICIENCIA TÉCNICA Y REDES NEURONALES: UN MODELO PARA EL CÁLCULO DEL VALOR AÑADIDO EN EDUCACIÓN

INTRODUCCIÓN GENERAL Y OBJETIVOS	1
CAPÍTULO 1: LAS REDES NEURONALES ARTIFICIALES COMO HERRAMIENTA PARA EL ANÁLISIS ECONÓMICO: CONCEPTOS Y APLICACIONES	9
1.1. Introducción.....	11
1.2. ¿Que son las Redes Neuronales Artificiales?.....	12
1.2.1 Las Redes Neuronales Biológicas	13
1.2.2. Las Redes Neuronales Artificiales	15
1.3. Un Poco de Historia.....	17
1.3.1. ETAPA 1: La Neurona Artificial Simple	18
1.3.2. ETAPA 2: El Perceptrón	22
1.3.3. ETAPA 3: El Perceptrón Multicapa	25
1.4. Características Generales de las RNAs	26
1.4.1. Arquitectura	26
1.5. El Aprendizaje del MLP: El Algoritmo <i>Backpropagation</i>.....	28
1.5.1. Características del MLP	29
1.5.2. Funcionamiento del Algoritmo <i>Backpropagation</i>	30
1.5.3. Algunas Cuestiones Prácticas: Métodos de Entrenamiento de un MLP	39
1.5.3.1. Empleo de una Submuestra Independiente o de Test.....	41
1.5.3.2. Validación Cruzada y Bootstrap.....	46
1.5.3.3. Combinación de distintas RNAs	48
1.6. Resumen: Principales Ventajas y Desventajas de las RNAs.....	51
1.7. Principales Aplicaciones de las RNAs en Economía.....	52
1.8 Consideraciones Finales	55

CAPITULO 2: LA MEDICIÓN DE LA EFICIENCIA TÉCNICA MEDIANTE REDES NEURONALES	57
2.1. Introducción.....	59
2.2. El Concepto Económico de Eficiencia	60
2.3. La Medición Empírica de la Eficiencia.....	65
2.3.1. Técnicas Paramétricas	65
2.3.2. Técnicas No Paramétricas.....	67
2.4. Las Redes Neuronales Artificiales como Herramienta para la Medición de la Eficiencia Técnica.....	69
2.4.1. Ejemplo 1: La Medición de la Eficiencia Técnica mediante RNAs.....	73
2.4.2. La Medición de la Eficiencia Técnica mediante RNAs a través de un Experimento de Monte Carlo: Ejemplo 2.....	77
2.5. Recapitulación	82
CAPÍTULO 3: LA ESTIMACIÓN DE LA FUNCIÓN DE PRODUCCIÓN EDUCATIVA: LOS MODELOS EN VALOR AÑADIDO	85
3.1. Introducción.....	87
3.2. La Función de Producción Educativa.....	88
3.3. Problemas en la Estimación de la Función de Producción Educativa.....	92
3.3.1 Características del Proceso Educativo.....	93
3.3.2. Algunas Dificultades en la Estimación Econométrica de la Función de Producción Educativa mediante el Esquema Input-Output.....	97
3.4. Debilidades Metodológicas en el Modelo Input-Output para la Medición de la Eficiencia en las Escuelas. Medidas de Resultados, Clima Escolar y Modelos en Valor Añadido.....	101
3.4.1. Inputs Individuales y Escolares, Medidas de Resultados y Clima Escolar.....	102
3.4.2. Indicadores y Modelos Educativos en Valor Añadido	108
3.5. Un Modelo en Valor Añadido para la medición de la Eficiencia Escolar.....	112
3.5.1. El Nivel de Alumno.....	112
3.5.2. La Construcción de Outputs en Valor Añadido.....	115
3.5.3. El Nivel Escolar.....	116
3.6. Recapitulación	121

CAPITULO 4: EL MODELO EN VALOR AÑADIDO PARA LA MEDICIÓN DE LA EFICIENCIA ESCOLAR: UNA APLICACIÓN AL CASO ESPAÑOL	122
4.1. Introducción.....	124
4.2. ¿Qué es el TIMSS?	125
4.3. El Nivel de Alumno	133
4.3.1 Modelo 1: Variables Individuales	133
4.3.2. Modelo 2: Variables Individuales y de Inteligencia Emocional	139
4.4. La Construcción de los Outputs Escolares en Valor Añadido.....	151
4.5. El Nivel Escolar.....	153
4.5.1. Los Inputs Significativos para Una Sola Dimensión de Output.....	153
4.5.2. Los Inputs Significativos para Varias Dimensiones de Output.....	161
4.5.3. Los Inputs Significativos para Todas las Dimensiones de Output.....	163
4.5.4. La Medición de la Eficiencia en las Escuelas.	163
4.6. La Eficiencia Regional	168
4.7. Recapitulación	171
REFLEXIONES FINALES Y LÍNEAS FUTURAS DE INVESTIGACIÓN	175
REFERENCIAS	181
ANEXOS	195
ANEXO 1: Etiquetas y Códigos de las Variables de Alumno creadas a partir del fichero original del TIMSS para los datos de España: bsgcode.cdt.....	197
ANEXO 2: Etiquetas y Códigos de las Variables del Colegio y el Profesor creadas a partir del fichero original del TIMSS para los datos de España: bcgcode.cdt y btmcode.cdt.....	216
ANEXO 3: Valores de los outputs reales y las predicciones para los centros evaluados.....	227
ANEXO 4: Eficiencia técnica de los centros evaluados para cada output.....	230
ANEXO 5: Ranking de eficiencia técnica de cada escuela para cada output.....	233

INTRODUCCIÓN GENERAL Y OBJETIVOS

INTRODUCCIÓN GENERAL Y OBJETIVOS

Las recientes reformas educativas llevadas a cabo en España en la educación no universitaria por diferentes gobiernos, vendrían a tratar de ofrecer respuestas a los objetivos que las sociedades modernas buscan en el actual sistema educativo. Desde la consolidación en los años ochenta, también en el campo de la educación, del llamado Estado del Bienestar, se generalizó en nuestro país la educación básica, universal y gratuita, como pilar imprescindible de lo que empezaba a ser una sociedad democrática moderna con ambición de convergencia hacia los países más ricos de nuestro entorno europeo. Actualmente, tanto la educación básica como parte de la educación secundaria es obligatoria hasta los 16 años.

Una vez conseguido este objetivo, que para algunos es considerado “el mínimo decente”, se han seguido detectando en el ámbito internacional comparado notables diferencias en las grandes cifras objetivas que conciernen a la educación. Así, a pesar del enorme esfuerzo realizado en las dos últimas décadas, España, sigue estando por debajo de la media de países de la OCDE en porcentaje del PIB dedicado a educación, a la cabeza en fracaso escolar y en una muy mediocre posición en cuanto a rendimiento escolar en asignaturas clave como las matemáticas y las ciencias.

Junto a este retraso histórico que arrastra España en los guarismos educativos respecto a otros países de su entorno, se han generado nuevos retos y necesidades sociales. Por un lado, España ha pasado de ser un país de emigrantes a convertirse en una economía rica, que debe hacer frente a una creciente inmigración de personas procedentes de otras realidades, como pueden ser los colectivos latinoamericanos, de la comúnmente llamada “Europa del Este” o del Magreb. A pesar de los indudables beneficios económicos y la riqueza cultural que la inmigración puede aportar y aporta a nuestra sociedad, resulta evidente que estos colectivos vienen con unas necesidades educativas distintas a los alumnos que serán sus compañeros en las escuelas. Dificultades como el idioma, el retraso educativo, o el contexto familiar, deben ser abordados desde los primeros días de su incorporación a la escuela, con el fin de lograr una plena integración del alumno en la sociedad de la que formará parte.

Pero no sólo las especiales necesidades educativas están asociadas al colectivo de población inmigrante. El interés de los padres por alcanzar el máximo nivel educativo en sus hijos, lleva a que demanden un trato especial cuando el alumno tiene problemas de aprendizaje debidos a multitud de factores, tanto psicológicos; de psicomotricidad o incluso por una inteligencia excesiva, por poner sólo algunos ejemplos. Hoy estos problemas pueden ser conocidos, sin embargo tiempo atrás cualquier síntoma de desinterés por el estudio o cualquier

problema en la escuela se resumía con la terrible afirmación de: “este niño no sirve para estudiar”.

Por otro lado, las nuevas tecnologías han pasado tan rápidamente a formar parte de nuestra realidad, que resulta difícil cualquier planteamiento que cuestione que ya son el presente y materia clave en la formación del alumno. Junto a este reto, el aprendizaje de idiomas, el conocimiento cultural y artístico, o la práctica de deporte, resultan claves para lograr una educación integral del alumno.

Así, los nuevos objetivos sociales en el mundo educativo comprenderían un amplio abanico, formado por la reducción del fracaso escolar; el aumento del rendimiento académico; la mejora en la calidad de la enseñanza (entendida como el acceso a las nuevas tecnologías; instalaciones educativas; educación más personalizada y métodos pedagógicos exitosos) y la igualdad de oportunidades educativas para todos los alumnos.

La solución a todos estos problemas y retos planteados serían para el gestor educativo una tarea relativamente fácil, y tan solo dependiente de los recursos disponibles, si dispusiéramos de lo que en economía llamamos, una función de producción educativa. Mediante esta herramienta podríamos relacionar los factores productivos, el contexto y los resultados educativos y usar esta información para asignar el monto de recursos que la sociedad decida aportar a la educación maximizando así los objetivos educativos. Desafortunadamente, después de miles de publicaciones que discuten en torno a las características de esta función productiva y que especulan con su obtención, las conclusiones de carácter general alcanzadas han sido muy escasas. Siguen siendo muchas las preguntas que quedan en el aire en torno a los recursos que más influyen a la hora de aumentar el rendimiento académico y reducir el fracaso escolar educativo.

Así, deben quedar claras las enormes carencias de conocimiento científico en economía de la educación. Cuando una sociedad aborda reformas educativas o decide dedicar un mayor esfuerzo en términos del PIB a la educación de sus ciudadanos, no dispone de ningún modelo teórico que garantice que más recursos, que en principio parecerían lógicos, tales como menos alumnos por clase, mejores salarios y más formación del profesorado, ordenadores en las aulas, etc., produzcan, *per se*, mejores resultados.

A pesar de los notables avances metodológicos, sobre todo en las tres últimas décadas, los modelos económicos tradicionales, basados en su mayoría en el esquema input-output, no han contestado a todas las cuestiones pendientes relacionadas con la educación. Es por ello que

desde algunos colectivos se plantea la inquietante pregunta; ¿importan las escuelas?. La desazón que esta realidad produce en los ciudadanos debe motivar los esfuerzos investigadores en torno al proceso educativo. El objetivo fundamental de esta Tesis es el de aportar nueva luz en torno a la función de producción y la medición de la eficiencia escolar a partir del análisis no lineal de una realidad tan compleja como la educativa. Este análisis no lineal será llevado a cabo mediante un modelo basado en redes neuronales artificiales (RNAs).

Para alcanzar este objetivo final se han desarrollado cuatro capítulos, cuyos contenidos y justificación desarrollamos a continuación.

En el capítulo 1, se pasará revista a las RNAs que será la herramienta estadística, aunque basada en la inteligencia artificial, que posteriormente se utiliza en el capítulo educativo. Así, se presentan las características principales de las RNAs y se justifican sus ventajas y desventajas sobre los métodos paramétricos en el estudio de fenómenos no lineales. Igualmente, se discuten algunas de las aplicaciones de esta técnica tanto en economía en general como en economía pública y de la educación en particular. El objetivo final de este primer capítulo es familiarizar al lector con esta novedosa técnica de *data mining*, cada vez más aplicada en distintas áreas de la investigación económica, resaltando sus bondades para analizar el problema de la educación.

El segundo capítulo podría ser considerado como una ampliación del primero, aunque su independencia en cuanto al tema que trata ha llevado a dejarlo independiente. En él se estudia el concepto económico de eficiencia farreliana, así como los principales métodos tradicionales para su estimación en distintos ámbitos. El objetivo de este capítulo es presentar a las RNAs como una herramienta alternativa a los modelos frontera tradicionales a la hora de estimar funciones de producción desconocidas. Concluiremos que las RNAs son una opción más a la hora de estimar la eficiencia de unidades productivas homogéneas en diferentes sectores económicos siguiendo la metodología input-output.

Sin embargo, el proceso productivo de la educación no universitaria no es tan simple como el de la fábrica de alfileres de Adam Smith. Es por ello que se dedican los capítulos tercero y cuarto al estudio teórico y empírico de cómo estimar la función de producción educativa y la eficiencia, en este sector de provisión mayoritariamente pública.

En el capítulo 3, se realiza una revisión histórica de los modelos de producción educativa. Desde un punto de vista crítico se analizan las debilidades de estos modelos en multitud de aspectos. Por un lado, los actuales modelos educativos consideran exclusivamente

el resultado medio de la clase, o de los alumnos de un nivel en el colegio, como el output educativo. En esta Tesis se defiende el uso del output *valor añadido escolar*, o el valor en términos de conocimientos que la escuela aporta al alumno una vez tenidas en cuenta las características de éste. Además, se construyen distintos outputs en valor añadido para distintos tipos de alumno, con el fin de investigar las políticas educativas que resultan eficaces sobre distintos alumnos. Para alcanzar este objetivo, se define el valor añadido como la diferencia entre el resultado escolar realmente obtenido y lo que se espera podría obtener el alumno en función de sus características individuales, no sólo las familiares, sino también las relativas a su contexto personal. Para ello utilizaremos RNAs que han demostrado ser útiles en el campo de la predicción no lineal. El supuesto clave que se defiende para la construcción de este tipo de output y modelo en dos etapas es que los alumnos son unidades heterogéneas que no pueden por tanto ser agregadas, ya que ello puede conducir a resultados erróneos.

En el capítulo 4, se presentan los resultados obtenidos con el modelo en valor añadido propuesto para los datos españoles del TIMSS (*Third International Mathematics and Science Study*), llevado a cabo por una organización tan prestigiosa como la IEA (*The International Association for the Evaluation of Educational Achievement*), de la que actualmente forman parte más de 58 países. El TIMSS es, hasta la fecha, el estudio de evaluación de Matemáticas y Ciencias más extenso en recopilación de datos. En la primera oleada del TIMSS, llevada a cabo en 1995, participaron en total más de 500.000 alumnos de tres niveles no universitarios (4º, 8º y 12º nivel, que se corresponden respectivamente con las edades de 10, 13 y 17 años), de 15.000 escuelas, en más de 40 países de todo el mundo y en más de 30 idiomas, siendo así la evaluación internacional con mayor número de países llevada a cabo hasta esa fecha. Además de la evaluación en matemáticas y ciencias, se llevó a cabo una enorme recolección de información tanto personal, como de los métodos educativos referente a los alumnos, profesores y directores de las escuelas, que se han utilizado en este trabajo.

En este capítulo, se muestra como distintos inputs educativos influyen en distinta cuantía sobre diferentes alumnos a la hora de explicar el output en valor añadido de la escuela. A continuación, una vez calculado y descontado el efecto de los factores escolares, se calcula la eficiencia escolar. El objetivo es demostrar que una escuela puede ser más o menos eficiente según el tipo de alumno que está recibiendo la educación. Seguidamente, se han agregado los datos a nivel regional para calcular la eficiencia de las ocho comunidades autónomas que participaron en este estudio. Una vez descentralizado el gasto educativo en España, el poder de decisión a la hora de tomar iniciativas de política educativa queda en manos de los consejeros regionales. Se muestra así la utilidad del modelo que presentamos para gestionar y comparar los recursos y resultados obtenidos en las distintas regiones españolas.

El objetivo final que se pretende alcanzar en este capítulo es demostrar de forma empírica que la escuela SÍ importa, pero que se debe ser cauteloso a la hora de medir sus efectos debido a las especiales características de este proceso productivo.

Finalmente se desarrollan algunas de las líneas futuras que quedan abiertas en esta Tesis. Por un lado, se hará referencia al potencial y los retos de las RNAs en futuras aplicaciones. Por otro lado, se repasarán los desafíos y las cuestiones pendientes referentes al progreso de la economía de la educación como disciplina.

**CAPÍTULO 1: LAS REDES NEURONALES ARTIFICIALES COMO
HERRAMIENTA PARA EL ANÁLISIS ECONÓMICO:
CONCEPTOS Y APLICACIONES**

1.1. Introducción

Durante las dos últimas décadas, se han producido en el ámbito de la investigación económica tres acontecimientos que han contribuido a una mejora significativa, tanto cuantitativa como cualitativa, de los resultados que se obtienen en los trabajos teóricos y, fundamentalmente, empíricos que se llevan a cabo en esta disciplina.

En primer lugar, el desarrollo de equipos informáticos cada vez más potentes, rápidos y baratos, ha favorecido la difusión de programas software dedicados al análisis de bases de datos muy amplias. Este avance supone que determinados análisis científicos que eran ambiciosos hace algunos años, como la construcción de complicados modelos econométricos, la simulación de distintos escenarios financieros o el tratamiento de datos y posterior análisis estadístico sean hoy asequibles, a un coste razonable, para los investigadores económicos.

En segundo lugar, la difusión de internet ha favorecido el rápido acceso a la información, tanto de datos como de publicaciones electrónicas, imprescindible para cualquier trabajo de investigación económica. Además, ha supuesto la puesta en contacto de grupos de investigación lejanos en el espacio pero muy cercanos ahora en el *ciberespacio*, lo que ha dado lugar a fuertes economías de escala a través de la puesta en común de bases de datos, conocimientos teóricos y resultados empíricos.

En tercer lugar, y favorecido por los dos anteriores, se ha producido un fuerte incremento en el conocimiento y la aplicación de modernas técnicas econométricas y estadísticas que tiempo atrás eran consideradas como sofisticadas y sólo al alcance de unos pocos. Nos estamos refiriendo a las regresiones no lineales, el análisis estadístico multivariante y en particular a las técnicas de minería de datos o *data mining*. La aplicación de estas técnicas está contribuyendo, como ya se ha indicado, a una significativa mejora de los trabajos y los resultados que se desarrollan en el amplio campo de la investigación económica y, en muchos casos, han supuesto nuevas formas de enfrentarse a problemas económicos tradicionales.

Si bien el efecto de estos avances ha sido en general muy positivo, el “bombardeo” de todo tipo de nuevas técnicas, a cada cual más prometedora, ha generado cierto grado de confusión en torno a las expectativas de los resultados que podemos alcanzar a través de estas herramientas. En un extremo de la distribución, se situarían aquellos investigadores recelosos de las nuevas técnicas que siguen fieles a las tradicionales, cuestionando cualquier novedad. En el otro lado de la balanza tendríamos aquellos otros con un elevado grado de entusiasmo, que

creen, sin demasiados argumentos, que estas nuevas herramientas se postulan como panaceas que pueden resolver mejor cualquier problema y solucionar aquellos que no eran fácilmente solubles con los métodos tradicionales.

Aunque las técnicas modernas no son infalibles en todos los casos, las tradicionales han demostrado sus limitaciones en determinados problemas, principalmente en el análisis de problemas de naturaleza no lineal. Uno de ellos es el del cálculo del valor añadido en educación como más adelante se explicará. A fin de superar estos problemas se defiende la utilización de una de estas nuevas técnicas: las redes neuronales artificiales (en adelante RNAs). El objetivo de este primer capítulo es tratar de arrojar algo más de luz sobre esta técnica de *data mining*, que en economía ha sido rodeada de un cierto “halo misterioso”, como todo aquello relacionado con la inteligencia artificial. Para ello, se analizan sus orígenes, características y principales aplicaciones actuales como herramienta de análisis económico, señalando con detalle sus principales ventajas e inconvenientes en este ámbito y frente a otras técnicas.

Se pretende, por tanto, dar a conocer la técnica al investigador económico poco o nada familiarizado en la misma, con el fin de contribuir a la difusión y materialización de sus potenciales aplicaciones en aquellos problemas en los que se justifique su uso. Por otro lado, se pretende desmitificar el uso de esta técnica, dando respuestas a lo que potencialmente puede y no puede hacer. Una vez conocida esta herramienta estadística, a lo largo de los siguientes capítulos se mostrará cómo puede ser aplicada para dos problemas concretos: la medición de la eficiencia técnica y la predicción del rendimiento económico en educación.

1.2 ¿Que son las Redes Neuronales Artificiales?

Las RNAs son el resultado de los intentos por reproducir mediante computadoras el funcionamiento del cerebro humano (tabla 1.1). Los modelos de RNAs creados hasta ahora son extremadamente simples y lo que se busca no es ya imitar a las neuronas auténticas, sino lograr una máquina de computación en paralelo formada por la interconexión de muchos elementos simples de cálculo.

Tabla 1.1. Diferencias entre el cerebro y el ordenador.

CEREBRO	ORDENADOR
<ul style="list-style-type: none"> • Unos 100.000 millones de uds. de proceso. • Cientos de operaciones por segundo. • Precisión aritmética muy escasa. • Paralelismo masivo. • Lógica difusa. • Memoria de tipo asociativo y almacenada de forma dispersa. • Tolera fallos (muerte de neuronas). • Maneja todo tipo de información, incluso sujeta a incertidumbre, en poco tiempo pero no necesariamente con exactitud 	<ul style="list-style-type: none"> • Una unidad de proceso. • Millones de operaciones por segundo. • Precisión aritmética absoluta. • Operaciones en serie. • Lógica rígida. • La información se guarda en posiciones de memoria de acceso directo. • Los pequeños fallos (averías) son críticos. • Sistemas altamente especializados con capacidad para procesar información muy concreta, siguiendo instrucciones dadas.

Fuente: Elaboración propia.

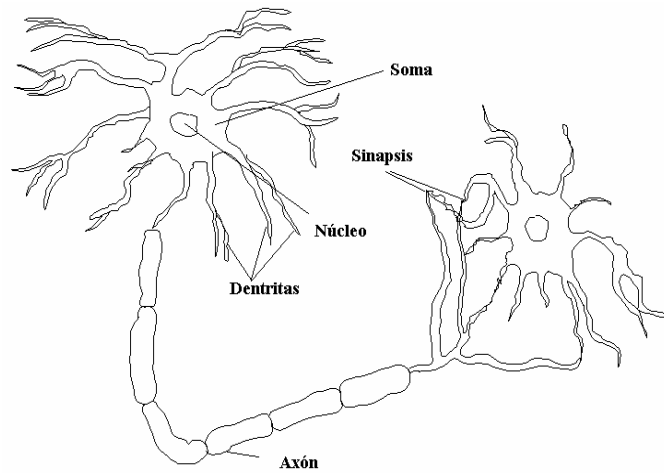
Actualmente, salvo en operaciones simples de cálculo aritmético, el cerebro humano es superior a cualquier ordenador. Las principales características del cerebro son su *robustez* y *paralelismo masivo*, ya que su funcionamiento no se ve alterado ante cambios de pequeña importancia, y su *flexibilidad*, ya que se adapta con rapidez a entornos cambiantes y, además, puede tratar con información ambigua o incompleta. Como se verá más adelante, la propiedad del cerebro humano más buscada en las RNAs es la capacidad de aprender de la experiencia y generalizar a partir de ella.

1.2.1 Las Redes Neuronales Biológicas

Se estima que el cerebro humano está compuesto por más de cien mil millones de neuronas interconectadas de forma variable y compleja. Su especial forma de trabajo le convierte en el dispositivo más versátil y eficaz para procesar todo tipo de información procedente del mundo real.

En la figura 1.1 se puede ver la representación esquemática de dos neuronas biológicas. En el cuerpo de la célula o *soma* se localiza el *núcleo* de la misma. De este *soma* se extienden fibras en forma de árbol llamadas *dentritas*, que son receptoras de impulsos procedentes de otras neuronas. También del *soma* sale una única y muy larga fibra llamada *axón*, que transporta la señal procesada por la neurona al exterior. El axón se subdivide arborizándose en más fibras denominadas *ramificaciones axonales*. Al final de éstas se encuentran las uniones sinápticas o *sinapsis*, que constituyen los puntos de conexión con las dentritas de otras neuronas. Una neurona recibe información de miles de otras neuronas y a su vez envía información a miles de neuronas más.

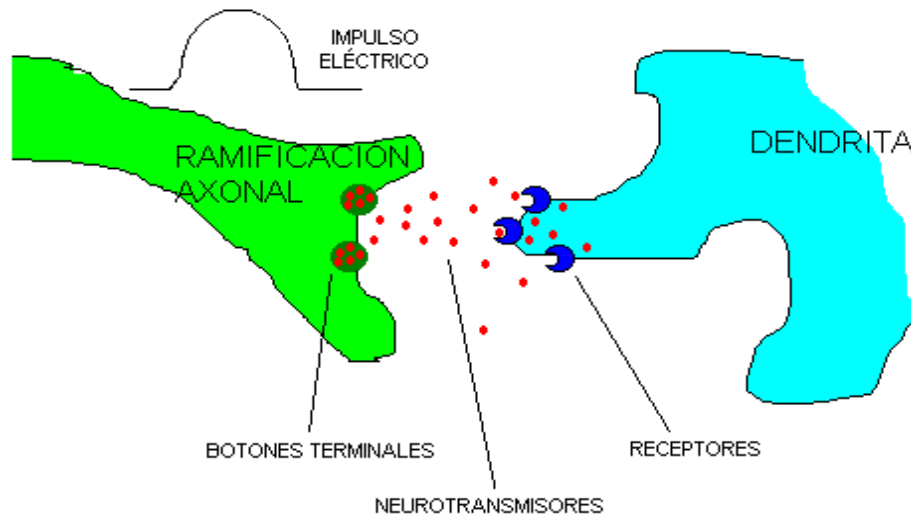
Figura 1.1. Características básicas de la neurona biológica.



La transmisión de una señal de una neurona a otra en la sinapsis es un proceso químico altamente complejo (figura 1.2). El proceso comienza cuando la neurona transmisora libera neurotransmisores, especialmente iones de sodio y potasio, en la sinapsis. Su misión es incrementar, (si la neurona es excitadora), o decrementar (si es inhibidora), el potencial eléctrico del soma de la neurona receptora y su efecto puede ser más o menos intenso en función de la cantidad de neurotransmisores liberados. Las señales excitadoras e inhibidoras recibidas por una neurona se combinan en el soma y, en función de la estimulación total recibida, la neurona toma un cierto nivel de activación. Si esta activación supera un determinado nivel umbral, se traduce enviando breves pulsos, o potencial de acción, de una determinada fuerza y duración a lo largo del axón. Se dice entonces que la neurona se ha activado.

Estos pulsos viajan por el axón hasta alcanzar las sinapsis con otras neuronas desde donde el proceso se repite otra vez. No todas las sinapsis son iguales y por tanto, no todas tienen el mismo efecto sobre la neurona receptora ya que algunas señales son más fuertes que otras. Otro factor que importa a la hora de activar una neurona es el número de conexiones que recibe.

Figura 1.2: Detalle de una neurona. Transmisión de una señal en la sinapsis.



Aunque todavía se desconoce mucho sobre la forma en que el cerebro aprende a procesar la información, las teorías actuales sobre el aprendizaje nos dicen que su efecto en el cerebro resulta de la modificación de las conexiones o sinapsis entre neuronas, lo que unido al procesamiento de información de la neurona determinan el mecanismo básico de la memoria, [Arbib, 1995]. Este mecanismo de modificación de los pesos sinápticos se produce en el tiempo por medio de la experiencia del sujeto y se realiza aumentando o disminuyendo el número de neurotransmisores liberados en determinadas sinapsis ante un mismo estímulo. Así, una neurona cuyo potencial antes no lograba alcanzar el umbral, ahora sí lo hace y se activa, y viceversa, otra que antes se activaba ahora puede que no lo haga y así la respuesta que el individuo da ante una misma situación real sea bien distinta, decimos entonces que hemos aprendido. Normalmente el aprendizaje requiere además de un maestro que indique como realizar bien una tarea. Tras un proceso de ensayo-error el individuo es capaz de realizar correctamente una tarea que antes no era capaz de resolver.

1.2.2. Las Redes Neuronales Artificiales

Bajo esta descripción, las RNAs tratan de simular este complicado proceso biológico, que se acaba de describir, en un equipo computacional basándose en el siguiente modelo:

- La información se procesa en elementos simples llamados **neuronas** (a veces también llamados nodos en la literatura sobre RNAs).
- Las señales o **inputs**, x_i , $i = 1, 2, \dots, n$; se transmiten entre las neuronas a través de **conexiones**.

- Cada conexión tiene asociado un **peso** w_{ij} que representa la intensidad de la señal recibida correspondiente al input i en la neurona j . La mayor o menor intensidad de cada input representa la influencia del mismo en el resultado final.
- Cada neurona j suma todas las señales recibidas, $\sum w_{ij} x_i$ procedentes de otras neuronas y las transforma mediante una **función de activación** f , generalmente no lineal, para determinar la señal de salida **output**, que transmite a otras neuronas, donde el proceso se repite. Finalmente, tras la interacción de las neuronas que componen la red, la señal o respuesta de la red se transmite al exterior.
- El aprendizaje se realiza mediante el ajuste de los pesos que ponderan las conexiones entre las neuronas que componen la red. El **algoritmo de aprendizaje** es el método por el que se obtienen los ajustes óptimos evitando tener que supervisar todo el proceso. Este método consta de una serie de pasos sencillos e iterativos.
- Para procesar la información las neuronas se organizan en **capas**: la capa de entrada transmite las variables input utilizadas, la capa de salida presenta el output final y entre medias pueden existir una o más capas ocultas que procesan la información.

En definitiva, se puede definir una RNA como un conjunto de procesadores sencillos, organizados en capas y altamente interconectados, que son capaces de aprender y generalizar la relación desconocida entre varias variables a partir de ejemplos reales. El objetivo de la red es, por tanto, aprender a asociar un vector input, X , a un vector output, Y , a partir de la interacción entre las neuronas, W . Esto es, dada la función de activación f , la red es entrenada para buscar una matriz de pesos W tal que:

$$\begin{aligned}
 & f \\
 & \mathbb{R}^n \rightarrow \mathbb{R}^m \\
 & f(W, X) = Y \\
 & y_k = f_0 \left[\sum_{i \rightarrow k} w_{ik} x_i + \sum_{j \rightarrow k} w_{jk} f_h \left(\sum_{i \rightarrow j} w_{ij} x_i \right) \right]
 \end{aligned}$$

A partir de una muestra $\{X(p), Y(p)\}$, $p = 1, 2, \dots, H$ donde $X(p) \in \mathbb{R}^n$, $Y(p) \in \mathbb{R}^m$. Donde i denota la información en capa de entrada, j la capa oculta, k la capa de salida, h la función de activación en las neuronas de la capa oculta y 0 en las neuronas de la capa de salida.

1.3. Un Poco de Historia

La inteligencia artificial (IA), entendida como el modelado y la simulación de las actividades cognitivas complejas; percepción, memoria, solución de problemas, etc. que caracterizan a los organismos avanzados, y en particular a los seres humanos se separó, tal y como señalan Waltz et al. (1988), casi desde su inicio en dos ramas bien diferenciadas:

Por un lado, se trató de modelar la actividad racional compleja mediante sistemas formales de reglas y manipulación simbólica. Esta es la rama más conocida de la IA, que se puede denominar simbólico-deductiva, en donde se postulan una serie de reglas y el sistema resuelve los problemas realizando deducciones sobre las reglas programadas. Las reglas *Si-Entonces* le dicen al sistema cómo se debe reaccionar ante distintos estados. De esta forma, un sistema experto es un método de representación y procesamiento del conocimiento, mucho más rico y poderoso que un programa clásico de ordenador. Sin embargo, con respecto al cerebro humano las limitaciones son múltiples. Los problemas planteados en términos difusos o ambiguos son difíciles de analizar con sistemas de procesamiento simbólico.

Por otro lado, se desarrollaron modelos computacionales conexionistas inspirados en las redes neuronales biológicas. Su principal característica es su inductividad, ya que extraen la información necesaria para resolver un problema a partir de un conjunto de ejemplos, sin necesidad de indicarle las reglas necesarias para resolver dicho problema.

Si bien es mucho más conocida la aproximación simbólico-deductiva y su principal aplicación: los sistemas expertos, (en general, sistemas basados en el conocimiento), existe un considerable y renacido interés por los modelos conexionistas. Los avances tecnológicos ofrecen, como ya ha sido apuntado, recursos cada vez mayores para representar estructuras muy complejas y realizar cálculos a gran velocidad y en paralelo, apoyando y fomentando la investigación en este campo y sus aplicaciones a distintas áreas del conocimiento. En la tabla 1.2 se puede observar un resumen de las principales diferencias entre ambos enfoques.

Tabla 1.2 Inteligencia Artificial: Comparación entre redes neuronales artificiales y sistemas expertos.

REDES NEURONALES	SISTEMAS EXPERTOS
Paralelo	Secuencial
Construidas a partir de ejemplos	Construidas a partir de reglas
Encuentran reglas	Necesitan reglas
Programadas internamente	Programadas externamente
Fáciles de mantener	Difíciles de mantener
Tolerantes a fallos	No tolerantes a fallos
Necesitan (sólo) una base de datos	Necesitan un experto
Lógica difusa	Lógica rígida
Sistemas adaptativos	Requieren reprogramación

Fuente: Samdani, G. (1990) y elaboración propia.

Se puede dividir el desarrollo histórico de las RNAs en tres grandes etapas:

1.3.1. ETAPA 1: La Neurona Artificial Simple

Con el nacimiento de la electrónica moderna, una de las vías de investigación en IA fue tratar de modelar el comportamiento del cerebro humano. Los orígenes de las RNAs suelen situarse a partir del modelo de neurona básico debido al neurofisiólogo Warren McCulloch y al matemático Warren Pitts. Estos autores escribieron un estudio en el que modelaban el funcionamiento de una neurona simple con circuitos eléctricos. En este trabajo proponían, basados en los estudios del funcionamiento de la neurona de Ramón y Cajal (1906), así como en los estudios de sinapsis excitadoras e inhibitoras de Sherrington (1906), una teoría general del procesamiento de la información. Esta teoría estaba basada en redes de elementos binarios, capaces de realizar cálculos similares a los ordenadores digitales, pero con una ejecución paralela y no serial, en donde los pesos hacían el papel que el programa juega en un ordenador convencional.

La eficiencia de la sinapsis en la transmisión de señales al cuerpo de la neurona pudo ser modelada en la neurona de Mc Culloch y Pitts (1943), asociando un factor multiplicativo a cada línea de entrada a la neurona. Si la neurona tiene n entradas procedentes de otras neuronas, el modelo queda representado por la ecuación 1.1:

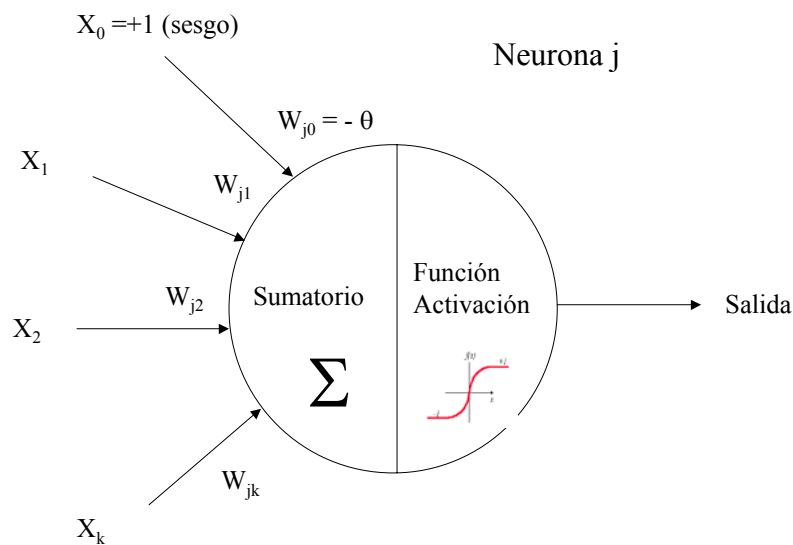
$$\text{Input total} = \text{peso línea 1} \times \text{input 1} + \dots + \text{peso línea } n \times \text{input } n = \sum_{i=1}^n \omega_i x_i \quad (\text{ec. 1.1})$$

Si esta suma supera el valor umbral, la señal de salida es 1 y, si no 0. Utilizando la notación $X = (x_1, \dots, x_n) \in \mathbb{R}^n$ para el vector de entradas, $W = (\omega_1, \dots, \omega_n) \in \mathbb{R}^n$ para el vector de pesos sinápticos y θ para el valor del umbral, la activación (salida) de la neurona para la entrada X pudo modelarse por la ecuación 1.2:

$$g(X) = \begin{cases} 1, & \text{si } \sum_{i=1}^n \omega_i x_i \geq \theta \equiv W^T \cdot X - \theta \geq 0 \\ 0, & \text{si } \sum_{i=1}^n \omega_i x_i < \theta \equiv W^T \cdot X - \theta < 0 \end{cases} \quad (\text{ec. 1.2})$$

La existencia del umbral se modela introduciendo entre los inputs un sesgo. Esto equivale a considerar otro input adicional cuyo valor siempre es 1 y cuyo peso asociado representa el umbral (figura 1.3).

Figura 1. 3. Estructura de la neurona artificial básica y modelado del umbral.



El proceso de la neurona simple puede interpretarse geoméricamente de la siguiente manera: dado un vector de entrada X , se calcula el producto escalar $W^T X$ y se le resta el umbral θ . Si el resultado es positivo, la neurona produce una respuesta y si es negativo otra. Así, una neurona con función de activación o función *escalón* (Heaviside), funciona exactamente igual que un clasificador lineal.

La Función de Activación

El modelo fue posteriormente extendido y en la actualidad la función de activación de la neurona, en el caso anterior la función de Heaviside (también conocida como función signo o escalón), se ha cambiado por una función no lineal. Esta extensión se debe a que casi todos los avances recientes en conexionismo se atribuyen, como veremos más adelante, a arquitecturas multicapa que utilizan funciones de activación no lineales como la función exponencial, la función umbral o muy comúnmente funciones de tipo sigmoideal [Quinlan, 1991]. El problema de trabajar con modelos no lineales radica en que son difíciles de describir en términos lógicos o matemáticos convencionales, [Jordan, 1986].

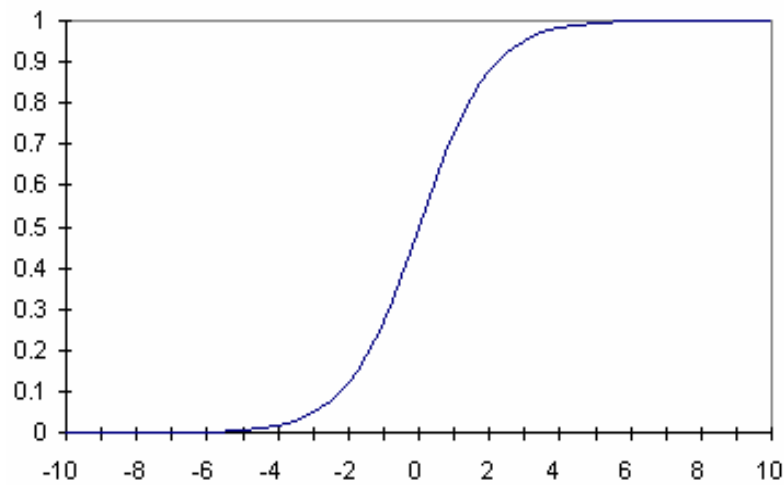
Además, en un principio se pensaba que las neuronas biológicas usaban una función de umbral, es decir, que permanecían inactivas y se activaban sólo si la estimulación total superaba cierto valor límite. Después se comprobó que las neuronas emitían impulsos de actividad eléctrica con una frecuencia variable, dependiendo de la intensidad de la estimulación recibida, y que tienen cierta actividad hasta en reposo, con estimulación nula.

Estos descubrimientos, llevaron a modelar este comportamiento usando funciones no lineales con esas características, como la función sigmoideal, con un perfil parecido al escalón de una función de umbral, pero continua. La función de activación logística o sigmoideal es probablemente la función de activación más empleada en la actualidad, (ecuación 1.3).

$$f(x) = \frac{1}{(1 + e^{-ax})} \quad (\text{ec. 1.3})$$

Se trata de una función continua, no lineal, con bastante plausibilidad fisiológica. La función sigmoide posee un rango comprendido entre 0 y 1. Esto, aplicado a las unidades de proceso de una RNA significa que, sea cual sea la entrada, la salida estará comprendida entre 0 y 1, (figura 1.4).

Figura 1.4: Función sigmoidal.



La salida de una unidad vale 0.5 cuando la entrada es nula. Esto significa que la unidad tiene cierta actividad aún en ausencia de estimulación. Al aumentar la estimulación, la unidad aumenta su activación y la disminuye si la estimulación es inhibitoria, de forma parecida a las neuronas reales. Sin embargo, la función sigmoidal tiene un proceso de transición más o menos suave aunque en un intervalo corto. Esto hace que esta función se considere a veces como una “función de Heaviside suavizada”.

Presenta además las siguientes características deseables, [Wassermann, 1989].

1. Acomodación de señales muy intensas sin producir saturación.
2. Admite señales débiles sin excesiva atenuación
3. Fácilmente derivable

$$f'(x) = f(x) \cdot (1 - f(x))$$

4. Es una función no lineal

La principal limitación de esta función es que no sirve para expresar polaridades, da siempre valores positivos. Una función alternativa, con cualidades parecidas pero con un rango entre -1 y 1, es la función tangente hiperbólica, aunque una solución más sencilla consiste en

reescribir de nuevo la ecuación sigmoide para trabajar en el intervalo $[+0.5, -0.5]$ (ecuación medio-sigmoidal) o en el intervalo $[-1, +1]$ ¹, (ecuación 1.4).

$$f(x) = \frac{2}{(1 + e^{-\alpha x})} - 1 \quad (\text{ec. 1.4})$$

1.3.2. ETAPA 2: El Perceptrón

El primer modelo de RNA capaz de asociar patrones fue el Perceptrón. Rosenblatt (1958) propuso el perceptrón simple con tan solo dos capas y un mecanismo de aprendizaje, basado en las ideas de Hebb (1949)², como un sistema de clasificación capaz de interpretar patrones tanto abstractos como geométricos. Esta RNA era capaz de asociar patrones de entrada formados por variables continuas, con una variable de salida discreta binaria del tipo todo-nada, 0/1, dado que el perceptrón aplicaba una función de umbral. Si la respuesta de la red es la correcta las conexiones existentes no se modifican, cosa que sí ocurre en caso contrario.

El aprendizaje de la neurona consiste en la modificación del vector de pesos W , de modo que la salida sea la adecuada para la tarea considerada. El algoritmo de aprendizaje del perceptrón funciona secuencialmente por pasos discretos. Se utiliza la siguiente notación:

$X(t) = (-1, x_1(t), \dots, x_n(t))$ = vector de entradas en el paso t

$W(t) = (\theta(t), w_1(t), \dots, w_n(t))^T$ = vector de pesos en el paso t

$\theta(t)$ = valor del umbral en el paso t

$y(t)$ = respuesta de la neurona en el paso t

$d(t)$ = respuesta deseada de la neurona en el paso t

α = valor del parámetro de aprendizaje: constante real entre 0,0 y 1,0

El esquema del algoritmo es el siguiente:

¹ La función $\tanh: \mathfrak{R} \Rightarrow [-1, 1]$ $G(a) = \frac{e^{\alpha a} - e^{-\alpha a}}{e^{\alpha a} + e^{-\alpha a}}$ es, como acabamos de señalar, muy utilizada, ya

que permite expresar esta polaridad.

² Hebb apuntó que la funcionalidad de una red neuronal era determinada por la fuerza de las conexiones entre las neuronas. Así, la regla de aprendizaje de Hebb apunta que si la red responde de forma deseable ante un input, entonces los pesos deben ser ajustados para incrementar la probabilidad de una respuesta similar ante un input similar en el futuro. Por el contrario, si la red responde de forma no deseable, los pesos deben ser ajustados para que decrezca la probabilidad de obtener la misma respuesta en el futuro.

1. Inicializar aleatoriamente los pesos para $W(0)$ y elegir un coeficiente de aprendizaje α tal que $0 < \alpha \leq 1$, que module la cuantía en la que los pesos serán modificados.
2. En el instante t , activar la neurona presentando el vector de entrada $X(t)$ y la respuesta deseada $d(t)$.
3. Calcular la respuesta actual

$$y(t) = U[W^T(t) \cdot X(t)] \quad U(n) = \begin{cases} +1 & \text{si } n \geq 0 \\ -1 & \text{si } n < 0 \end{cases}$$

4. Modificar los pesos de las sinapsis de la neurona.

$$W(t+1) = W(t) + \alpha[d(t) - y(t)]X(t)$$

5. Repetir los pasos del 2 al 4 con todos los patrones hasta que el perceptrón calcule correctamente la salida.

El algoritmo de aprendizaje es iterativo, es decir, no calcula los pesos de una vez, sino que hay que aplicarlo repetidamente hasta que se obtiene un conjunto de pesos que soluciona el problema. El mecanismo es muy análogo al que se da en el aprendizaje humano. Una persona debe repetir una tarea e ir corrigiendo y aprendiendo de sus errores hasta que es capaz de ejecutarla correctamente. La diferencia $d(t) - y(t)$ desempeña el papel de la señal de error.

Este algoritmo tiene su justificación matemática en un teorema llamado, «*Teorema de la Convergencia del Perceptrón*», cuyo enunciado, propuesto por Rosenblatt (1958) es el siguiente:

Teorema de la Convergencia del Perceptrón. Sean X_1 y X_2 dos conjuntos de patrones linealmente separables. El algoritmo del perceptrón aplicado a $X = X_1 \cup X_2$ converge tras un número finito de iteraciones.³

Este teorema demuestra que el algoritmo siempre obtiene una solución, es decir, un conjunto de pesos en un número finito de iteraciones, si el problema a resolver por la red de neuronas es linealmente separable y todos los ejemplos son presentados con la misma frecuencia.

³ La demostración del teorema va más allá de los objetivos de esta tesis doctoral. Existen además distintas versiones de la demostración del mismo. Para ver una sencilla demostración del teorema puede acudirse por ejemplo a Hertz et al. (1991).

Tras este brillante inicio, dos profesores del M.I.T (Marvin Minsky y Seymour Papert) pusieron en claro el alcance y las limitaciones de estos novedosos modelos de proceso en el campo de la IA. El problema detectado por Minsky y Papert (1969) era que este tipo de redes neuronales tan solo serían capaces de hacer clasificaciones en conjuntos de datos que fueran linealmente separables, (figura 1.5), ya que la célula de decisión que utiliza es el hiperplano. ¿Pero qué ocurre si el problema a tratar no se puede separar con un hiperplano?.

Resulta obvio, que muchos problemas reales no son linealmente separables, lo que llevó al abandono masivo del estudio de las RNAs. Un ejemplo clásico de este tipo de problema, es el XOR lógico⁴. En este caso, tal y como muestra la figura 1.6, aunque claramente existen dos patrones, no existe una recta o hiperplano que pueda dejar los círculos blancos a un lado y los negros a otro. Este es un problema no lineal, pero ¿Hasta qué punto abundan y por tanto nos puede interesar resolver problemas no lineales? Podemos tratar de responder a esta pregunta a través de algunas cifras. Con dos variables binarias se pueden implementar un total de 16 funciones booleanas, de las cuales 14 son linealmente separables y sólo 2, el XOR y su negación, son no lineales.

Figura 1.5. Ejemplo de regiones de decisión linealmente separables.

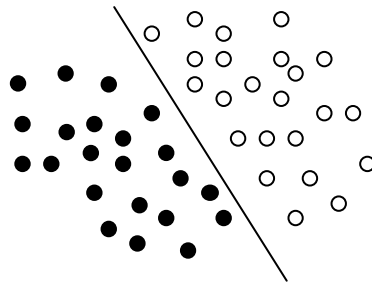
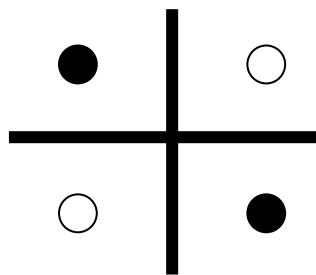


Figura 1.6. Ejemplo de patrones no separables linealmente. El XOR lógico.



⁴ El XOR lógico clasifica dos entradas iguales en un conjunto y dos distintas en otro.

Parece, por tanto, que con dos variables binarias los problemas no lineales no son muy abundantes pero si aumentamos su número, por ejemplo a 6, el número de posibles funciones a implementar es de $1,8 \times 10^{19}$, de las cuales sólo 5.028.134 son linealmente separables, es decir, aproximadamente el 10 por 100. A medida que aumenta el número de variables envueltas en el problema y nos adentrarnos en el mundo de los números reales, esta desproporción a favor de los problemas no lineales es todavía más fuerte. Por tanto, parece que sí interesa resolver este tipo de problemas que sin duda aparecen frecuentemente en el campo de la economía.

Con la imposibilidad de abordar los problemas no lineales, las posibilidades de este tipo de técnica se reducen, y a pesar de que un grupo de investigadores continuara trabajando en RNAs, el campo deja de tener atractivo durante la década de los 70. En este período Werbos (1974) desarrolló y empleó el algoritmo *backpropagation*. Sin embargo, para solucionar este problema iba a ser necesario convertir el perceptrón de un solo nivel en un perceptrón multicapa, mediante la introducción de capas ocultas.

1.3.3. ETAPA 3: El Perceptrón Multicapa

Habría que esperar al trabajo seminal del *Parallel Distributed Processing Research Group* de la Universidad de California para que la investigación en torno a las RNAs alcanzara su mayor difusión histórica. Así, Rumelhart y Mc Clelland (1986) popularizaron una red de neuronas, el Perceptrón Multicapa⁵ (MLP), que junto con un método de aprendizaje especial, el algoritmo de retropropagación supervisada o *Backpropagation*, es capaz de resolver aquellos problemas que el perceptrón no podía. El campo se convierte en multidisciplinar, y toda la comunidad científica: ingenieros, médicos, físicos, psicólogos, biólogos, matemáticos y también economistas, entre otros, comienzan a aportar y sugerir nuevas ideas hacia múltiples aplicaciones reales.

⁵ Utilizaremos en adelante las siglas inglesas de Multilayer Perceptron (MLP) por la que es más conocido.

1.4. Características Generales de las RNAs

1.4.1. Arquitectura

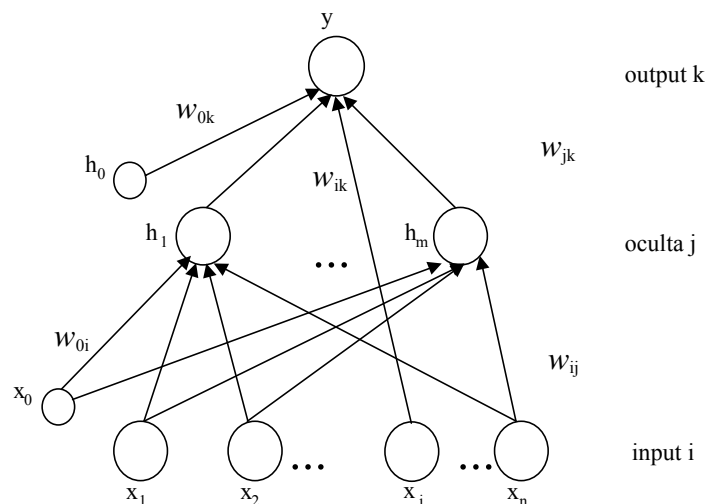
Como hemos apuntado anteriormente, dentro de una RNA, las neuronas se encuentran organizadas en distintas capas (figura 1.7):

Capa de entrada: Recibe los inputs del exterior. En esta capa normalmente no se procesan las señales recibidas sino que éstas son enviadas a la capa siguiente. El número de inputs que debe utilizarse depende del problema específico que estemos considerando. Desde el punto de vista de la econometría, las variables independientes serían los inputs del problema que presentaríamos a la capa de entrada.

Capa(s) oculta(s): Las neuronas de esta(s) capa(s) transforman la señal recibida y la envían a la capa de salida. Para la mayoría de los problemas es suficiente con una sola capa oculta. De la modificación de los pesos de las neuronas de esta capa depende el aprendizaje de la red.

Capa de salida: Una vez la red ha transformado la señal, esta capa envía la respuesta al exterior. Las neuronas de esta capa realizan una nueva transformación de la señal recibida por la capa(s) oculta(s).

Figura 1.7: Organización de las capas de una red neuronal.



1.4.2. Una Clasificación de las RNAs

Si bien en economía la gran mayoría de trabajos han aplicado el MLP, que será la red que comentaremos con mayor profundidad a lo largo de este capítulo, hay que indicar que existen numerosos modelos de RNAs que podemos clasificar atendiendo a diversos criterios y que brevemente pasamos a comentar.

Según *la forma en que las neuronas están conectadas* podemos clasificar las RNAs en dos grupos básicos:

- **Redes de propagación hacia delante (*feed-forward*).** Es la forma más básica de red. Las neuronas de cada nivel sólo están conectadas con las neuronas de los niveles posteriores por lo que la información se propaga hacia delante. Algunas de las redes de este tipo más conocidas son: Perceptron, Adaline (ADAPTative LINear Elements), Madaline⁶ (Multiple Adaline) y el MLP. Todas ellas suelen aplicarse a la aproximación de funciones o al reconocimiento o clasificación de patrones.
- **Redes de propagación hacia atrás (*feedback*).** En estas redes, las neuronas pueden estar conectadas indistintamente con neuronas de niveles previos, posteriores, de su mismo nivel (conexiones laterales) o incluso con ellas mismas (autorrecurrentes). Algunos de los modelos más conocidos, [Fausett, 1994] son: La Bidirectional Associative Memory (BAM) y la máquina de Boltzman.

Podemos también clasificar las RNAs *a partir de las reglas básicas de entrenamiento-aprendizaje*, señalando tres tipos fundamentales: supervisado, no supervisado y reforzado.

- **Aprendizaje supervisado:** Los datos para el entrenamiento están constituidos por un grupo de pares de patrones entrada-salida deseada. Existe un maestro (algoritmo) encargado de comprobar que la salida de la red coincida con la respuesta deseada, si finalmente no es así se deben modificar los pesos de las conexiones para intentar que la salida obtenida se aproxime lo más posible a la respuesta deseada. Este tipo de redes han sido las más desarrolladas, desde el inicio de las RNAs, para la aproximación de funciones y la clasificación de patrones. El MLP utiliza, como ya hemos señalado, este tipo de aprendizaje.

⁶ Los modelos de Widrow (1959), Adaline y Madaline (Multiple ADAPTive LINear Elements), permitieron usar, por primera vez, una red neuronal en un problema del mundo real: filtros adaptativos que eliminan ecos en las líneas telefónicas. Esta red neuronal aún está en uso.

- **Aprendizaje no supervisado:** En esta ocasión se suministran los inputs de entrada, pero la red no recibe ninguna información que le indique cuál es la salida. La red modifica los pesos para poder asociar a la misma unidad de salida los vectores de entrada más similares. Suele decirse que estas redes son capaces de auto-organizarse. Normalmente, se habla de dos tipos de entrenamiento no supervisado: el hebbiano y el competitivo/cooperativo. El primero generalmente más orientado a detectar las similitudes de las entradas y el segundo más orientado hacia la clusterización. El tipo de red más difundido en aplicaciones reales son las Redes Auto-Organizadas⁷ o de Kohonen [Kohonen, 1984].
- **Aprendizaje Reforzado:** A medio camino entre los anteriores. En este caso, el supervisor se limita a indicar si la salida ofrecida por la red es correcta o incorrecta, pero ni se le indica la cuantía de su error ni la respuesta que debería dar. El tipo de red más extendida de este tipo son las Redes de Hopfield [Hopfield, 1982].

1.5. El Aprendizaje del MLP: El Algoritmo *Backpropagation*

En este apartado se entra en el estudio detallado del MLP. El MLP y su algoritmo de aprendizaje *backpropagation* es el tipo de red neuronal que con más frecuencia ha sido aplicado en los estudios de economía en general, y que en particular aplicaremos posteriormente en el análisis empírico de los capítulos siguientes.

Como se verá más adelante, existen tres fases en la aplicación del MLP: la fase de aprendizaje o entrenamiento, la fase de validación y la fase de test.

- La *fase de entrenamiento*, consiste en hacer que la red sea capaz de extraer, a partir de ejemplos reales, reglas generales que la permitan en el futuro responder adecuadamente a patrones nunca vistos anteriormente. Este proceso consiste, por tanto, en la búsqueda de la matriz de pesos óptima.
- Una vez entrenados varios modelos, es la *fase de validación* o de prueba, la que nos ayuda a seleccionar cual de las redes entrenadas responde mejor a los objetivos planteados.
- Por último, una vez elegida la red definitiva que procesará los inputs, ésta resultará operativa, y es entonces cuando la arquitectura, número de neuronas, conexiones y

⁷ Conocidas también como SOM procedente del término inglés Self-Organizing Map.

pesos quedan fijos y la red está lista para funcionar. Es en este momento cuando se realiza la *fase de test*, donde se presenta a la red *un patrón de entradas nunca visto antes por ella* con el fin de obtener una predicción o clasificación insesgada de los datos. Se evalúa así, una vez comparada la respuesta real y la estimada, la precisión del MLP para la resolución del problema planteado.

1.5.1. Características del MLP

En el MLP las neuronas se agrupan por capas conectándose todas las neuronas de una capa con las de la siguiente capa y así sucesivamente. Se trata de una red *feed-forward* o en cascada, en la cual las señales se propagan únicamente hacia delante. Para empezar, veamos como procesa una red multicapa la información que le es suministrada.

Existe un vector de entrada $x_p = (x_{p1}, x_{p2}, x_{p3}, \dots, x_{pN})$ que se aplica a las neuronas de la capa de entrada de la red, y ésta distribuirá dichos valores a las neuronas de la capa oculta. Las entradas netas hacia estas neuronas se calculan mediante la siguiente fórmula:

$$net_{pj}^h = \sum_{i=1}^N w_{ji}^h \cdot x_{pi} + \theta_j^h$$

en donde,

- El superíndice h se refiere a magnitudes de la capa oculta.
- El subíndice p se refiere al p -ésimo vector de entrenamiento.
- El subíndice j se refiere a la j -ésima neurona oculta.
- El término θ representa el sesgo que puede considerarse como una entrada más cuyo término asociado es la unidad.
- w_{ji}^h es la ponderación de la señal procedente de la neurona de la capa de entrada i que llega a la j -ésima neurona de la capa oculta h .

Donde net_{pj}^h es la suma lineal de los inputs recibidos ponderados y N denota el número de inputs que estamos considerando. El resultado se hace pasar a continuación a través de una función de transferencia, generalmente sigmoideal.

$$i_{pj} = f_j^h(\text{net}_{pj}^h)$$

Asimismo, para las neuronas de la capa de salida, la entrada neta y su correspondiente salida se calculan mediante las siguientes expresiones;

$$\text{neta}_{pk}^o = \sum_{j=1}^L w_{kj}^o \cdot i_{pj} + \theta_k^o$$

Siendo L el número de neuronas en la capa oculta. Este resultado es de nuevo transformado en la neurona de la capa de salida cuyo resultado es el output definitivo.

$$y_{pk} = f_k^o(\text{neta}_{pk}^o)$$

En donde el superíndice *o* se refiere a magnitudes de la capa de salida. Si denotamos por A y B las matrices que recogen los valores de los pesos de la capa de entrada a la oculta y de la oculta a la de salida respectivamente; por Y, al vector salida de la red; X al vector input y como F, a la función de transferencia, entonces el MLP quedará descrito como:

$$Y = F [F (X \cdot A) \cdot B]$$

Una vez calculada la salida, el método a seguir para ajustar los pesos es el siguiente: los pesos se actualizarán de forma proporcional al error, que se calcula por diferencia entre el valor deseado y el obtenido.

Dado que el MLP no sólo tiene una capa de salida, sino que además puede tener una o varias capas ocultas, y todas ellas están formadas por múltiples neuronas, hay que definir el error de cada una de las neuronas que componen tanto la capa de salida como las capas ocultas. Una vez calculados estos errores, deberemos actualizar a continuación los pesos en ambas capas.

1.5.2. Funcionamiento del Algoritmo *Backpropagation*

El método de entrenamiento supervisado más extendido en la literatura sobre el MLP es el algoritmo de retropropagación supervisada (*Backpropagation*), también conocido como regla delta generalizada, [Rumelhart y Mc Clelland, 1986]. El objetivo del proceso de aprendizaje es

ajustar los pesos de la red para que los patrones de entrada introducidos permitan obtener las salidas deseadas. El aprendizaje de tipo supervisado supone la necesidad de disponer de pares de entrenamiento *entrada-salida deseada*. Este proceso se lleva a cabo a través de la minimización de una función de actuación, que generalmente se corresponde con la suma de errores al cuadrado. Veamos uno de los posibles desarrollos matemáticos de este algoritmo.

El algoritmo backpropagation constará de dos fases:

1ª Fase: Propagación Hacia Delante.

1. *Inicialización*: Se asigna aleatoriamente a la red un vector de pesos uniformemente distribuidos cuyo rango depende del usuario pero que generalmente se establece en el intervalo $[-0.5, +0.5]$.
2. *Propagación*: Se presenta a la red un vector de entrada. Esta información se va propagando a través de todas las capas (de la forma que vimos en el apartado anterior) hasta obtener el vector salida global de la red.
3. *Comprobación del error cometido*: A continuación la red compara la salida obtenida por ella con la deseada.

Actualización de los pesos de la capa de salida

Dado el p-ésimo vector de entrenamiento, el error cometido por la unidad k-ésima perteneciente a la capa de salida es el siguiente:

$$\varepsilon_{pk} = d_{pk} - y_{pk}$$

Siendo:

d_{pk} el valor de salida deseado

y_{pk} el valor de salida obtenido

La regla delta generalizada trata de minimizar la suma de los cuadrados de los errores de todas las unidades de salida,

$$\varepsilon_p = \frac{1}{2} \sum_{k=1}^M \varepsilon_{pk}^2$$

Donde multiplicar por $\frac{1}{2}$ sólo se hace por razones de conveniencia a la hora de simplificar en el cálculo de las derivadas.

Una vez calculado el error, se tiene que determinar el sentido en que los pesos van a modificarse. Para ello se calcula el gradiente de la función de error, y de este modo los cambios en los pesos se deben producir en la dirección del gradiente negativo, que es la dirección del descenso más pronunciado a lo largo de la superficie de error que se haya definida en el espacio de pesos. Los cambios de los pesos deben hacerse de forma iterativa, hasta que ε_p alcance el punto mínimo de la superficie de error.

Por tanto, la regla general para la modificación de los pesos en la última capa se realiza mediante la siguiente expresión:

$$\Delta w_{kj}^o = -\alpha \frac{\partial \varepsilon_p}{\partial w_{kj}^o}$$

Donde α es el coeficiente de aprendizaje encargado de regular la cuantía en la que los pesos serán modificados.

Y para resolver la derivada parcial se aplica la regla de la cadena:

$$\frac{\partial \varepsilon_p}{\partial w_{kj}^o} = \frac{\partial \varepsilon_p}{\partial net_{pk}^o} \cdot \frac{\partial net_{pk}^o}{\partial w_{kj}^o}$$

2ª Fase: Hacia atrás:

4. *Adaptación del vector de pesos:* Los cambios en los pesos se producen en la dirección en la que el error caiga lo más rápidamente posible (gradiente negativo descendente). Estas señales de error se propagan a los nodos de las capas, empezando por la de salida y siguiendo con las capas sucesivas. De esta forma, los nodos de cada capa sólo reciben una fracción del error global en función de su aproximada contribución relativa a la obtención de la salida. Los pesos se modificarán en base a la señal de error recibida, de manera que se reduzca el error actual de la red y la salida obtenida se vaya aproximando a la deseada.

4a. *Actualización de los pesos de la capa de salida.*

- La primera derivada, de la expresión anterior, es el error negativo cometido por la unidad de salida k-ésima.

$$\frac{\partial \varepsilon_p}{\partial net_{pk}^o} = \frac{1}{2} [2 \cdot (d_{pk} - y_{pk}) \cdot (-1)] = -(d_{pk} - y_{pk}) = -\varepsilon_{pk}$$

para calcularlo volvemos a aplicar la regla de la cadena,

$$\varepsilon_{pk} = -\frac{\partial \varepsilon_p}{\partial net_{kj}^o} = -\left(\frac{\partial \varepsilon_p}{\partial y_{pk}} \cdot \frac{\partial y_{pk}}{\partial net_{pk}^o} \right)$$

- a) Al tratarse de una neurona de la capa de salida, la primera de estas derivadas parciales resulta:

$$\frac{\partial \varepsilon_p}{\partial y_{pk}} = \frac{\partial \frac{1}{2} (d_{pk} - y_{pk})}{y_{pk}} = -(d_{pk} - y_{pk})$$

- b) Al ser $y_{pk} = f_k^o(net_{pk}^o)$ la segunda de las derivadas parciales queda recogida por la siguiente expresión:

$$\frac{\partial y_{pk}}{\partial net_{pk}^o} = \frac{\partial f_k^o(net_{pk}^o)}{\partial net_{pk}^o} = f_k'^o(net_{pk}^o)$$

Uniendo lo obtenido en el apartado a y b, se llega al siguiente resultado que nos va a permitir calcular el gradiente:

$$\frac{\partial \varepsilon_p}{\partial net_{kj}^o} = (d_{pk} - y_{pk}) \cdot f_k'^o(net_{pk}^o)$$

- Desarrollando la segunda derivada,

$$\frac{\partial net_{pk}^o}{\partial w_{kj}^o} = \frac{\partial \sum_{j=1}^L w_{kj}^o \cdot i_{pj}}{\partial w_{kj}^o} = i_{pj}$$

Uniendo ambas expresiones se obtiene la variación del peso de una conexión que va de una capa oculta de la red a la capa externa,

$$\Delta w_{kj}^o = \alpha \cdot (d_{pk} - y_{pk}) \cdot f_k' (net_{pk}^o) \cdot i_{pj}$$

de esta manera, los pesos de la capa de salida se actualizan según la expresión:

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \Delta w_{kj}^o(t)$$

4b. Actualización de los pesos de las capas ocultas.

Las salidas obtenidas por las neuronas de las capas ocultas son conocidas, pero no se sabe si éstas son o no las correctas. Ahora bien, el error global de alguna manera tiene que estar relacionado con estos valores de salida. El procedimiento a seguir para calcular como varían los pesos de dichas capas es idéntico al anterior (pesos de la capa de salida), excepto por una pequeña modificación debido a que tenemos que adaptarlo a una capa oculta.

Recordemos que las entradas netas de los nodos que componen la capa oculta y su correspondiente salida vienen dadas por las siguientes expresiones:

$$net_{pj}^h = \sum_{i=1}^N w_{ji}^h \cdot x_{pi} + \theta_j^h$$

$$i_{pj} = f_j^h (net_{pj}^h)$$

Siguiendo el mismo procedimiento que en la capa de salida, la modificación de los pesos de la capa oculta se realiza mediante la siguiente expresión,

$$\Delta w_{ji}^h = -\alpha \frac{\partial \varepsilon_p}{\partial w_{ji}^h}$$

Aplicando la regla de la cadena a la derivada parcial,

$$\frac{\partial \varepsilon_p}{\partial w_{ji}^h} = \frac{\partial \varepsilon_p}{\partial net_{pj}^h} \cdot \frac{\partial net_{pj}^h}{\partial w_{ji}^h}$$

A la primera derivada se le vuelve a aplicar la regla de la cadena,

$$\frac{\partial \varepsilon_p}{\partial net_{pj}^h} = \frac{\partial \varepsilon_p}{\partial i_{pk}} \cdot \frac{\partial i_{pk}}{\partial net_{pj}^h}$$

- a) Al tratarse de una neurona de la capa oculta, a la primera de estas derivadas parciales hay que aplicarle de nuevo la regla de la cadena,

$$\frac{\partial \varepsilon_p}{\partial i_{pk}} = \sum_j^N \frac{\partial \varepsilon_p}{\partial net_{pk}^o} \cdot \frac{\partial net_{pk}^o}{\partial i_{pj}} = \sum_j^N \frac{\partial \varepsilon_p}{\partial net_{pk}^o} \cdot \frac{\partial \sum_i^N w_{kj}^o i_{pi}}{\partial i_{pj}}$$

Resolviendo queda la siguiente expresión,

$$\frac{\partial \varepsilon_p}{\partial i_{pj}} = \sum_k^N \frac{\partial \varepsilon_p}{\partial net_{pk}^o} \cdot w_{ki}$$

De la que se deduce,

$$\frac{\partial \varepsilon_p}{\partial i_{pj}} = \sum_k^N (d_{pk} - y_{pk}) \cdot f_k^{\prime o}(net_{pk}^o) \cdot w_{kj}^o$$

- b) La segunda de las derivadas parciales, al igual que en la capa de salida, queda,

$$\frac{\partial i_{pj}}{\partial net_{pj}^h} = \frac{\partial f_j^h(net_{pj}^h)}{\partial net_{pj}^h} = f_j^{\prime h}(net_{pj}^h)$$

Uniendo lo obtenido en los apartados a y b, llegamos a los resultados de la primera parte de la regla de la cadena inicial,

$$\frac{\partial \varepsilon_p}{\partial net_{pj}^h} = f_j'^h(net_{pj}^h) \cdot \sum_k^N (d_{pk} - y_{pk}) \cdot f_k'^o(net_{pk}^o) \cdot w_{kj}^o$$

- La segunda derivada parcial de la regla de la cadena inicial queda,

$$\frac{\partial net_{pk}^h}{\partial w_{ji}^h} = \frac{\partial \sum_i^N w_{ji}^h x_{pi}}{\partial w_{ji}^h} = x_{pi}$$

De nuevo si juntamos los resultados obtenidos en ambas derivadas parciales, se obtiene la expresión que permite calcular la variación que deben experimentar los pesos de las capas ocultas,

$$\Delta w_{ji}^h = \alpha \cdot x_{pi} \cdot f_j'^h(net_{pj}^h) \cdot \sum_k^N (d_{pk} - y_{pk}) \cdot f_k'^o(net_{pk}^o) \cdot w_{kj}^o$$

Como ya se ha demostrado anteriormente en la actualización de los pesos de la capa de salida,

$$\varepsilon_{pk} = (d_{pk} - y_{pk}) \cdot f_k'^o(net_{pk}^o)$$

Sustituyendo en la expresión de la variación de pesos resulta,

$$\Delta w_{ji}^h(t) = \alpha \cdot x_{pi} \cdot f_j'^h(net_{pj}^h) \cdot \sum_k^N \varepsilon_{pk} w_{kj}^o$$

Por tanto, los pesos de las capas ocultas se actualizan mediante la siguiente expresión:

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \Delta w_{ji}^h(t)$$

Resumiendo, podemos definir el error de las neuronas de salida y el error de las neuronas de la capa(s) oculta(s) expresando la actualización de los pesos de ambas capas en función de ellos:

Capa de salida

$$\text{Error: } \varepsilon_{pk}^o = (d_{pk} - y_{pk}) \cdot f_k'^o(neta_{pk}^o)$$

$$\text{Actualización del peso: } w_{kj}^o(t+1) = w_{kj}^o(t) + \alpha \cdot \varepsilon_{pk}^o \cdot i_{pj}$$

Capas ocultas

$$\text{Error: } \varepsilon_{pj}^h = f_j'^h(neta_{pj}^h) \cdot \sum_k^N \varepsilon_{pk}^o w_{kj}^o$$

$$\text{Actualización del peso: } w_{ji}^h(t+1) = w_{ji}^h(t) + \alpha \cdot \varepsilon_{pj}^h \cdot x_{pi}$$

Sólo queda por definir la función de activación que por lo general, en este tipo de redes (redes multicapa), suelen ser de tipo sigmoideal. Por tanto,

$$f_k^o = \frac{1}{1 + e^{-neta_{pk}^o}}$$

$$f_k'^o = f_k^o \cdot (1 - f_k^o) = y_{pk} \cdot (1 - y_{pk})$$

$$f_j'^h = i_{pj} \cdot (1 - i_{pj})$$

Así pues, los términos de error pueden expresarse en términos de la función de activación de tipo sigmoideal,

Error de las unidades de salida:

$$\varepsilon_{pk}^o = (d_{pk} - y_{pk}) \cdot f_k'^o(neta_{pk}^o) = (d_{pk} - y_{pk}) \cdot y_{pk} (1 - y_{pk})$$

Error de las unidades ocultas:

$$\varepsilon_{pj}^h = f_j'^h(\text{net}_{pj}^h) \cdot \sum_k^N \varepsilon_{pk}^o w_{kj}^o = i_{pj} \cdot (1 - i_{pj}) \cdot \sum_k^N \varepsilon_{pk}^o w_{kj}^o$$

5. *Actualización de los pesos:*

- Actualización de los pesos de las unidades de salida,

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \Delta w_{kj}^o(t)$$

Siendo,

$$\Delta w_{kj}^o(t) = \alpha \cdot \varepsilon_{pk}^o \cdot i_{pj}$$

- Actualización de los pesos de las unidades ocultas,

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \Delta w_{ji}^h(t)$$

Siendo,

$$\Delta w_{ji}^h(t) = \alpha \cdot \varepsilon_{pj}^h \cdot x_{pi}$$

Dentro de cada capa individual, el orden de actualización de los pesos de dicha capa no tienen porque seguir un orden establecido.

6. *Vuelta al paso 2:* Con los nuevos pesos, se introduce un nuevo par vector input-respuesta deseada y el proceso se repite hasta que el algoritmo converge cuando el error es suficientemente pequeño y no existen ganancias al seguir aplicando el algoritmo. Es entonces cuando el proceso de entrenamiento habrá finalizado.

Las principales desventajas de este algoritmo son que el tiempo de entrenamiento crece exponencialmente a medida que se incrementa el número de nodos en la capa oculta y que no existe ninguna seguridad ni forma de verificar que hemos alcanzado el mínimo global.

1.5.3. Algunas Cuestiones Prácticas: Métodos de Entrenamiento de un MLP

En principio, existe un teorema fuerte que demuestra que un MLP con un número de neuronas suficiente en la capa oculta puede calcular cualquier función con el grado de aproximación que sea necesario [Funahashi, 1989, Hornik et al., 1989 y 1990; White, 1990]. Para funciones muy complejas o con discontinuidades, haría falta un MLP con dos capas ocultas [Cybenko, 1988]. En este sentido, el MLP ha sido considerado como un aproximador universal de funciones⁸. En la práctica, el MLP es especialmente útil para la clasificación y aproximación de datos y funciones que toleran cierta imprecisión, para los que existen muchos pares de entrenamiento y donde se dispone de información de las variables explicativas pero *carecemos a priori de un modelo matemático o económico preciso*; es decir, desconocemos la forma funcional que relaciona los inputs con los outputs.

El entrenamiento del MLP es un proceso de ensayo-error en el que no existen reglas generales. En este proceso debemos fijar una serie de parámetros y decidir la arquitectura de la red. Existen problemas frecuentes que aparecen al entrenar un MLP. Veamos algunos comentarios de interés acerca de su implementación práctica:

- *Coefficiente de aprendizaje*: Normalmente este coeficiente es fijo y con un valor reducido para asegurar que la red alcanzará una solución. Por otro lado, un valor muy reducido supone incrementar el número de iteraciones necesarias hasta obtener una solución.
- *Momento*: A menudo la ecuación para la modificación de los pesos es modelada de la siguiente forma:

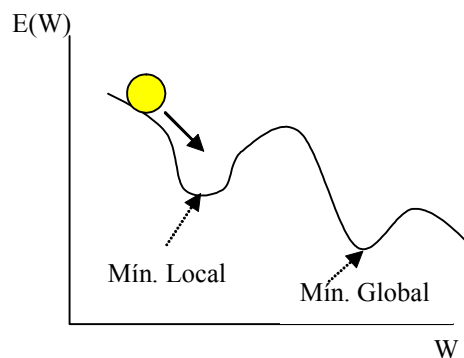
$$\Delta_t w_{ji} = -\alpha \frac{\partial E_t}{\partial w_{ji}} + \beta \Delta_{t-1} w_{ji}$$

donde β es el vector momento. Con esta formulación introducimos un *recuerdo* capaz de generar cierta inercia en el sistema ya que en el paso $t+1$ se conserva algo del cambio hecho en el paso t . El momento hace que no se produzcan saltos bruscos en los valores de los pesos de tal forma que el algoritmo descienda suavemente a través de la superficie del error.

⁸ Una revisión actual de la propiedad de aproximación universal de las RNAs se recoge en Scarselli y Chung (1998).

- *Mínimos Locales*: El algoritmo backpropagation busca minimizar el error del sistema. Para ello se moverá por la superficie de la función de error dando pasos descendentes. Debido a este método de aprendizaje, el sistema podría eventualmente quedar atrapado en un mínimo local y alcanzar una solución que no fuese óptima, es decir con un error elevado (figura 1.8).

Figura 1.8: Problemas en el aprendizaje. Mínimo local en la superficie del error.



Existen diversos criterios heurísticos diseñados a fin de evitar este problema:

- Modificar el número de nodos en la capa(s) oculta(s).
- Comenzar de nuevo el entrenamiento con un conjunto distinto de pesos.
- Cambiar los parámetros de aprendizaje.
- Añadir capas ocultas.
- Añadir ruido a las conexiones.

En problemas aplicados no podemos saber si el mínimo alcanzado es o no global. Debemos por tanto conformarnos con que el mínimo alcanzado responda de forma satisfactoria a los requerimientos del problema planteado.

- *Arquitectura de la red*: La elección de la arquitectura óptima de la red no tiene ninguna regla (únicamente heurísticos) y es por tanto un arduo proceso de ensayo-error. Respecto al número de capas ocultas, si bien existen resultados teóricos que apuntan a que en redes *feed-forward* la existencia de una capa oculta es suficiente, [véase Hornik, 1989 y

1990]; en muchas aplicaciones para problemas complejos es conveniente el uso de más de una capa oculta.

Además, debemos elegir el número de nodos en la capa oculta teniendo en cuenta que si la red es muy pequeña no será capaz de formar un buen modelo del problema. Por otro lado, si la red es muy grande corre el peligro de sobreentrenarse, esto es, la red se ajusta a los datos de entrenamiento e incluso a su ruido en vez de al patrón general de comportamiento de los datos, produciendo de este modo una mala generalización que detectaremos en la fase de validación. Normalmente se empieza con una red con pocas unidades en la capa oculta y se van añadiendo nodos en función de los resultados observados.

- *Entrenamiento:* Para entrenar la red con vistas a la predicción o clasificación existen distintas metodologías. Veamos a continuación algunas de las que son comúnmente utilizadas en los trabajos empíricos.

1.5.3.1. Empleo de una Submuestra Independiente o de Test

Es el método más utilizado en la práctica empírica. Para llevar a cabo su implementación, se divide la muestra en tres submuestras que denominaremos: de entrenamiento, de validación y de test. Cada una de estas tres submuestras debe incluir vectores de toda la distribución de la función que se quiere aproximar.

El entrenamiento de la red es un proceso de ensayo-error que básicamente consiste en la construcción de distintas arquitecturas de redes, que a partir de la muestra de entrenamiento generalicen la relación entre variables explicativas y explicadas. Para conseguir esta generalización, variaremos el número de neuronas en la(s) capa(s) oculta(s) e inicializaremos los pesos aleatoriamente para evitar caer en mínimos locales de la superficie del error. Al mismo tiempo usaremos la submuestra de validación para elegir el vector de pesos que mejor aproxima los datos.

Una vez construidas distintas RNAs con distintas arquitecturas utilizaremos la que obtiene la suma de errores al cuadrado más reducida en la muestra de validación, (Bishop, 1995; Fine, 1999). Con la RNA definitiva, se utilizan los datos de comprobación o de test, que nos servirán para obtener el error final de predicción de la red de forma no sesgada, ya que hasta este momento *nunca antes habían sido utilizados*. Por tanto, la característica crucial de una submuestra de test es que guarde una distribución similar a los datos de entrenamiento y validación y que no haya sido usada para elegir entre dos RNAs.

Para entrenar una red neuronal que generalice de forma adecuada, se debe tener en cuenta, como ya se ha indicado anteriormente, que un MLP con pocas neuronas tenderá a no aprender la estructura y generalidades de los datos, mientras que una red con demasiadas neuronas tiene el riesgo de quedar sobreentrenada. Es decir, el MLP no sólo aprende los patrones generales, sino que también ajusta el ruido de los datos de entrenamiento “aprendiéndolos casi de memoria”, por lo que su capacidad a la hora de predecir o clasificar datos nuevos quedará seriamente dañada. Es por ello por lo que en este método dividimos la muestra en tres submuestras

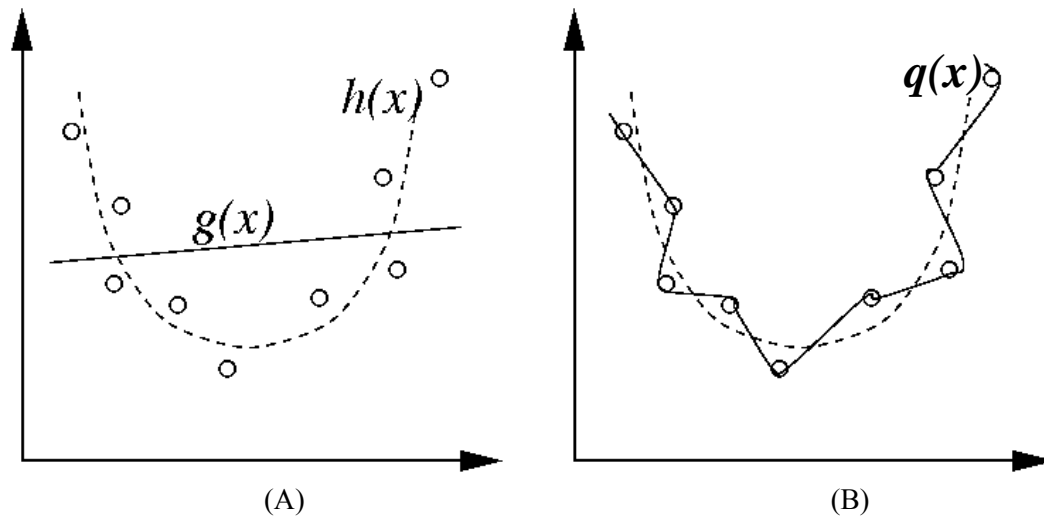
Así, al entrenar un MLP se plantea el problema conocido en estadística como el *dilema sesgo-varianza*, [German et al., 1992]. En el ejemplo de la figura 1.9, los datos representados por círculos han sido generados a partir de la función $h(x)$ a la que además se ha añadido cierto ruido. Nuestro objetivo sería en este problema aproximar $h(x)$ a partir del conjunto de datos disponible $y(x)$.

$$y(x) = h(x) + \varepsilon$$

En la figura 1.9.A, se ha aproximado $h(x)$ a partir de una función $g(x)$ con muy pocos parámetros obteniendo una línea recta. El modelo tiene la virtud de ser extremadamente simple, sin embargo no es capaz de ajustar correctamente los datos disponibles y no predecirá bien nuevos puntos. Se dice que se trata de un modelo con un alto sesgo.

Por el contrario, en la figura 1.9.B, se ha aproximado $h(x)$ a partir de la función $q(x)$ que utiliza muchísimos parámetros. La ventaja de este modelo es que ajusta casi de forma perfecta los datos originales de la muestra, sin embargo esta función no serviría para predecir $h(x)$ a partir de un nuevo vector de puntos x . En este caso el modelo presenta una varianza elevada.

Figura 1.9. El dilema sesgo-varianza.



Claramente lo que se busca es el caso intermedio, representado en la figura 1.9 por la curva de trazos discontinuos. Un modelo suficientemente robusto para ajustar los datos pero que no sea tan potente como para ajustar también el ruido asociado a la muestra de datos disponible.

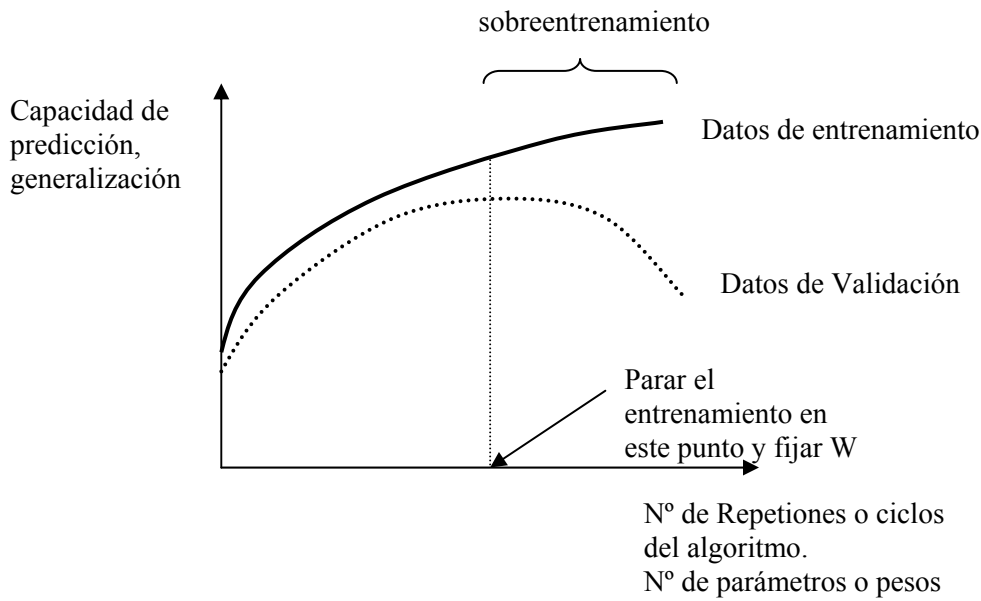
El dilema sesgo-varianza es más probable que sea un problema a medida que tenemos una muestra de datos más reducida. Si por el contrario disponemos de una base de datos continua, que tienda a infinito, el peligro de ajustar el ruido disminuye.

Existen distintos métodos para controlar el *sobreentrenamiento*, o ajuste excesivo de la red neuronal a los datos de entrenamiento, cuando la muestra de la que disponemos es limitada.

El primero es el *método de parada temprana* [Hecht-Nielsen, 1990]. Supone que durante el proceso de entrenamiento la red consigue un ajuste óptimo después del cual, si dejamos que el algoritmo continúe actuando, se produciría el sobreentrenamiento. Así, a la vez que entrenamos la red, los resultados deben ser validados con otra matriz de datos, que llamamos de validación, con una distribución muestral similar a la de la muestra de entrenamiento.

Tras un número de repeticiones del algoritmo detendremos el proceso de aprendizaje cuando la capacidad de generalización del MLP comience a deteriorarse, figura 1.10.

Figura 1.10: Método de parada temprana para evitar el entrenamiento excesivo de un MLP.



El segundo método sería el de *regularización*, y está basado en una perspectiva estadística. Apunta a que el aumento en el número de parámetros o pesos de la red, aumenta el grado de sobreentrenamiento. Así, deberemos probar la capacidad de generalización de varias redes neuronales entrenadas con diversas arquitecturas hasta encontrar la óptima. A menudo, para evitar un MLP sobreparametrizado se utiliza el *criterio de información de Schwartz (SIC)*, Schwartz (1978).

$$SIC(M) = N \cdot \log(MSE) + M \cdot \log(N)$$

Donde M es el número de parámetros o pesos del modelo, N es el tamaño de la muestra y MSE es el error cuadrático medio⁹. Otra solución similar sería añadir un término que penalice el exceso de parámetros en la función de error del MLP tal y como muestra la ecuación 1.5.

$$\vec{E} = E + \nu\Omega \quad (\text{ec. 1.5})$$

Siendo,

$$\Omega = \frac{1}{2} \sum_i w_i^2$$

⁹ Existen además otros métodos de regularización para evitar un número excesivo de parámetros como el AIC o Criterio de Información de Akaike, [Akaike, 1973].

Usando esta función de error modificada, los pesos se modificarían ahora de la siguiente forma,

$$\Delta w_{ij} = -\mu \frac{\partial \vec{E}}{\partial w_{ij}} = -\mu \frac{\partial E}{\partial w_{ij}} - \mu \nu w_{ij}$$

Donde el término final del lado derecho fuerza a que los pesos se modifiquen en función de su propio tamaño.

El tercer y último método que se describirá es el *entrenamiento con ruido (jitter)*, [Sarle, 1997]. Este método consiste, como su propio nombre indica, en añadir una pequeña cantidad de ruido blanco a cada uno de los inputs y/o outputs utilizados. Así, cada vez que presentemos el vector entrada-salida deseada al MLP le añadiremos un pequeño ruido aleatorio.

Si disponemos de una muestra pequeña, una manera alternativa de introducir ruido en el proceso de entrenamiento sería ajustar los pesos después de la presentación de un pequeño subconjunto de datos reordenando a continuación los casos aleatoriamente. De esta forma cada adaptación del vector de pesos estaría basada en una estimación *ruidosa* del verdadero gradiente.

Otro capítulo en el proceso de entrenamiento sería la cuestión de cuantas redes deben ser entrenadas antes de elegir la correcta. No existen reglas a la hora de realizar la elección del número de neuronas en la capa oculta, parámetros o si utilizar una o varias capas. Esto supone que habitualmente el entrenamiento de una red neuronal es un arduo proceso de ensayo y error y, en general, se busca la obtención de una solución plausible al problema planteado. Un buen objetivo a priori sería el de alcanzar al menos los resultados predictivos de un método econométrico lineal, aunque en muchas ocasiones ni siquiera esto será posible de obtener en un tiempo razonable debido al problema de los mínimos locales.

Gorr et al. (1994) proponen como estrategia operativa para el entrenamiento de un MLP, llevar a cabo una búsqueda exhaustiva del MLP más adecuado. Para ello estos autores proponen variar el coeficiente de aprendizaje entre 0.2 y 0.8 en intervalos de 0.1 y el número de neuronas en la capa oculta entre dos números definidos a priori, por ejemplo entre 8 y 15, si bien este número dependerá del problema planteado. Con ello tendríamos en este ejemplo que entrenar una parrilla de 7 x 7 MLPs durante un número de iteraciones predeterminado.

1.5.3.2. Validación Cruzada y Bootstrap

La principal limitación del método anterior es la necesidad de disponer de una muestra de datos extensa para dividirla en tres partes. En algunos problemas no disponemos de muchos datos por lo que debemos buscar otras alternativas para el entrenamiento del MLP. Existen dos métodos que mantienen la independencia en la validación usando al mismo tiempo casi la totalidad de la muestra para el entrenamiento que son la validación cruzada y el bootstrap.

El objetivo de estos métodos es reducir el sesgo implícito que supone usar los mismos datos para estimar conjuntamente los parámetros y comprobar su generalización. Sin embargo actualmente estos métodos son muy costosos computacionalmente y poseen una alta probabilidad de alcanzar distintos mínimos locales.

Validación Cruzada

La validación cruzada es una técnica para estimar la capacidad de generalización de un estimador a partir de una sola muestra. Esta técnica fue introducida por Stone (1974, 1977) y discutida en numerosas ocasiones a lo largo de la literatura¹⁰. En la validación cruzada se extraen repetidamente s submuestras del total de los datos y los resultados obtenidos por cada red se ponderan con las distintas submuestras extraídas. En la práctica, este proceso se repite k veces, de tal forma que su tamaño sea $s = n / k$, siendo n el número de elementos muestrales. Otra opción sería dejar cada vez un dato fuera y evaluar la red en función del dato que borramos del entrenamiento cada vez (*Jackknife*).

Mediante esta metodología el error de la red que normalmente es

$$\varepsilon_{pk} = (d_{pk} - y_{pk})^2$$

Sería

$$\varepsilon_{CV} = \frac{1}{n} \sum_{\alpha=1}^n \varepsilon_{\alpha}$$

¹⁰ Para una revisión puede acudir a Hastie y Tibshirani (1990), Sec 3.4..

Este error promedio va a permitir reducir el sesgo de la estimación (Fine, 1999, pp 252). Al modificar la muestra antes de cada entrenamiento, existe una mayor probabilidad de caer en distintos mínimos locales. Plutowski et al. (1994) demuestran que el error obtenido ε_{CV} es insesgado.

Además, necesitamos conocer las fluctuaciones respecto a la media. A menos que la varianza sea pequeña, el estimador no será robusto.

$$VAR(\varepsilon_{cv}) = \frac{1}{2} VAR(\varepsilon_1) + \left(1 - \frac{1}{n}\right) COV(\varepsilon_1, \varepsilon_2)$$

Teniendo en cuenta que $VAR(\varepsilon_\alpha)$ es independiente de la elección de α y que $COV(\varepsilon_\alpha, \varepsilon_\beta)$ para $\alpha \neq \beta$ no depende de la elección de índices distintos. Así, podemos concluir que $COV(\varepsilon_\alpha, \varepsilon_\beta)$ tiende asintóticamente a cero. Aunque esto no puede ser demostrado para muestras pequeñas o no muy grandes, precisamente esta metodología se propone en el caso de disponer de este tipo de muestras.

Bootstrap

Efron et al. (1993) definen un estimador bootstrap $\hat{\theta}$ para un parámetro θ , que es función de la verdadera distribución P y la muestra de entrenamiento T , $\theta = \theta(P, T)$, dado en términos de una estimación de la distribución empírica \hat{P}_T de P basada en T . La distribución empírica se describe como una función de probabilidad $\hat{P}_T(x, t)$ para el resultado observado (x, t) que viene dado por la frecuencia relativa de (x, t) en T .

$$\hat{P}_T(x, t) = \frac{1}{n} [\#(x_i, t_i) = (x, t)] \quad \hat{\theta} = \theta(\hat{P}_T, T)$$

Así, a partir de una forma explícita de θ , podemos calcular $\hat{\theta}$ sustituyendo \hat{P}_T por P en θ , tal y como hemos llevado a cabo en la expresión anterior. Sin embargo si carecemos de una forma explícita para $\theta(P, T)$ y siguiendo a Efron et al. (1993, Sec. 5.4), tendremos que la expresión $\theta(P, T) = P(\eta \neq t|T)$ es el error en la probabilidad de η .

En primer lugar, tomaremos una muestra bootstrap iid $T^{(b)} \left\{ (X_1, t_1)^{(b)}, \dots, (X_n, t_n)^{(b)} \right\}$, de tamaño $n = \|T\|$, distribuida igual que la distribución empírica de \hat{P}_T y evaluaremos $\eta^{(b)}$ entrenado con los datos de entrenamiento bootstrap $T^{(b)}$. A continuación evaluamos el error cometido usando $\hat{P}_T(\eta^{(b)} \neq t)$ y no la frecuencia relativa del error de la muestra de entrenamiento seleccionada aleatoriamente $T^{(b)}$. Este proceso debe ser repetido un elevado número de veces m y podemos obtener una estimación bootstrap de $\theta(P, T)$ a través de:

$$\tilde{\theta} = \frac{1}{m} \sum_{i=1}^m \hat{P}_T(\eta^{(b)} \neq t)$$

Desde aquí es inmediato que,

$$E\left(\tilde{\theta} - \theta(\hat{P}_T, T)\right)^2 \leq \frac{1}{4m}$$

Esta diferencia puede hacerse tan pequeña como permitan los recursos informáticos. A partir de aquí, podemos estimar $\theta(\hat{P}_T, T)$ con el grado de proximidad deseado y la única discrepancia vendrá dada por la diferencia entre $\theta(\hat{P}_T, T)$ y la deseada $\theta(P, T)$. Por otro lado, se sabe a partir de los teoremas de Glivenko-Cantelli y Kolmogorov-Smirnov [Vapnik, 1998, Sec. 3.9], que la función de distribución empírica converge a la verdadera función de distribución y por tanto la diferencia entre ambas tendería a cero.

Paass (1993) ilustra la aplicación de la metodología bootstrap en el uso de las RNAs. Esta requiere entrenar un amplio conjunto de RNAs usando distintas muestras obtenidas mediante bootstrap. A continuación, se generaría la función de distribución del output a partir de las respuestas de todas las RNAs, construyendo así los intervalos de confianza. Esta metodología es, sin embargo, muy costosa a la hora de llevar a cabo su implementación informática pero sin duda parece una buena opción para aplicaciones futuras.

1.5.3.3. Combinación de distintas RNAs

Tal y como se ha visto, la elección de la RNA óptima se hace después de entrenar múltiples RNAs con múltiples arquitecturas y parámetros. Normalmente, una de las RNAs entrenadas es seleccionada como la mejor y el resto son descartadas. Hashem (1993) y Hashem

y Schmeiser (1995) proponen el uso de combinaciones lineales de las RNAs entrenadas en lugar de una única RNA. La combinación de las RNAs es un método que integra el conocimiento adquirido por las distintas RNAs y por ello mejora el ajuste del modelo.

Hashem y Schmeiser (1995) derivan dos expresiones para la obtención de la combinación lineal óptima de MLPs, a partir de un conjunto de p MLPs entrenados para un problema multi-input con un solo output. Un MLP entrenado, transforma un vector input \vec{x} en un escalar $y(\vec{x})$. El error en la aproximación sería $\delta(\vec{x}) = t(\vec{x}) - y(\vec{x})$ donde $t(\vec{x})$ es la respuesta real del input \vec{x} . Una combinación lineal de los outputs de los p MLPs entrenados sería $\bar{y}(\vec{x}; \vec{\alpha}) = \sum_{j=1}^p \alpha_j y_j(\vec{x})$ siendo el error asociado $\bar{\delta}(\vec{x}; \vec{\alpha}) = t(\vec{x}) - \bar{y}(\vec{x}; \vec{\alpha})$, donde $y_j(\vec{x})$ es el output del MLP j -ésimo y α_j es el peso asociado para la combinación de $y_j(\vec{x})$, siendo $j = 1, \dots, p$.

El problema es encontrar los valores del vector $\vec{\alpha}$. Una solución sería elegir una sola red o ponderar a todas ellas por igual. Sin embargo ambas soluciones, a pesar de su simplicidad, no son óptimas. Sea el input \vec{x} una observación de la variable aleatoria \vec{X} que tiene una función de distribución aleatoria $F_{\vec{X}}$. Entonces la respuesta deseada sería $t(\vec{X})$, el output del MLP j -ésimo sería $y_j(\vec{X})$ y el error de aproximación asociado $\delta_j(\vec{X})$ siendo $j = 1, \dots, p$. La combinación lineal de outputs sería la variable aleatoria $\bar{y}(\vec{X}; \vec{\alpha}) = \sum_{j=1}^p \alpha_j y_j(\vec{X})$ siendo el error asociado la variable $\bar{\delta}(\vec{X}; \vec{\alpha}) = t(\vec{X}) - \bar{y}(\vec{X}; \vec{\alpha})$. La combinación lineal óptima vendrá definida así por el vector $\vec{\alpha}^* = \alpha_1^*, \dots, \alpha_p^*$ que minimize:

$$\int_S \ell(\bar{\delta}(\vec{X}; \vec{\alpha})) dF_{\vec{X}}$$

Donde S es el espacio en donde se encuentra definida $F_{\vec{X}}$ y ℓ es una función de pérdida que en nuestro caso se define como el error cuadrático $\ell(\bar{\delta}) = \bar{\delta}^2$. El objetivo es así minimizar el error cuadrático medio:

$$MSE(\bar{y}(\vec{x}; \vec{\alpha})) = E\left(\left(\bar{\delta}(\vec{X}; \vec{\alpha})\right)^2\right)$$

Donde E es la esperanza matemática con respecto a $F_{\bar{X}}$. Diferenciando el MSE respecto a $\bar{\alpha}$ obtendremos el vector de combinación de pesos óptimo no restringido.

$$\bar{\alpha}^* = \Phi^{-1}\bar{\Theta}$$

Donde $\Phi = [\phi_{ij}] = [E(y_i(\bar{X})y_j(\bar{X}))]$ es una matriz $p \times p$ y $\bar{\Theta} = [\theta_i] = E[t(\bar{X})y_i(\bar{X})]$ es un vector $p \times 1$.

Consideraremos también el caso de una combinación de pesos óptima restringida donde tendremos $\sum_{j=1}^p \alpha_j = 1$. En este caso, si las y_j 's son insesgadas, entonces \bar{y} será también insesgado. Bajo la condición de que la suma de pesos sea igual a uno, la combinación de pesos óptima es:

$$\bar{\alpha}^* = \Omega^{-1}\bar{1} / (\bar{1}'\Omega^{-1}\bar{1})$$

Donde $\Omega = [\omega_{ij}] = [E(\delta_i(\bar{x})\delta_j(\bar{x}))]$ es una matriz $p \times p$ y $\bar{1}$ es un vector $p \times 1$ con todos sus componentes iguales a 1.

En la práctica, rara vez se conoce $F_{\bar{x}}$ y por ello Φ , Θ y Ω deben ser estimados. Si D es un conjunto de observaciones independientes obtenidas de $F_{\bar{x}}$, entonces $\bar{x} \in D$ puede ser tratado con la misma probabilidad. Así, Φ , Θ y Ω serían estimados usando las siguientes expresiones:

$$\hat{\Phi}_{ij} = \sum^{D} (y_i(\bar{x}_k)y_j(\bar{x}_k)) / |D| \quad \forall i, j$$

$$\hat{\Theta}_i = \sum^{D} (y_i(\bar{x}_k)y_j(\bar{x}_k)) / |D| \quad \forall i$$

$$\hat{\omega}_{ij} = \sum^{D} (y_i(\bar{x}_k)y_j(\bar{x}_k)) / |D| \quad \forall i, j$$

1.6. Recapitulación: Principales Ventajas y Desventajas de las RNAs.

Tras esta revisión del funcionamiento de la técnica de RNAs, parece conveniente recapitular y enumerar las principales ventajas e inconvenientes de esta herramienta:

- *Método de aprendizaje*: No necesitan ser programadas. Son capaces de extraer sus propias reglas a partir de ejemplos reales mediante la adaptación de la matriz de pesos. Esas reglas quedan almacenadas y extendidas a lo largo de las conexiones.
- *Tolerancia al ruido*: Son capaces de abstraer las características esenciales de los datos y de esta manera pueden generalizar correctamente, aún en presencia de datos incompletos o distorsionados.
- *No paramétricas*: Si bien la técnica de RNAs puede ser considerada por algunos autores como paramétrica, ya que su implementación implica el cálculo de distintos pesos (parámetros), las RNAs no hacen supuestos acerca de la forma funcional de la función que van a aproximar, ni sobre la distribución de las variables independientes. Por otro lado, aunque podemos saber el valor de cada peso, no sabemos la influencia exacta o efecto marginal de cada peso en el resultado final. Es por ello que la denominación de técnica no paramétrica parece la más adecuada.
- *No lineal*: Las RNAs permiten realizar a través de sus funciones de activación transformaciones no lineales de los datos, lo cual supone una ventaja frente a los métodos tradicionales de regresión y análisis discriminante múltiple.
- *Cantidad de datos suficiente*: Dado que la aplicación de las RNAs requieren el uso de tres submuestras de datos, la obtención de una base de datos lo suficientemente grande puede ser una limitación para su implementación práctica. En este contexto el uso de la validación cruzada, las técnicas de bootstrap o la introducción de ruido, se han demostrado que son útiles para superar este problema.
- *Caja no del todo negra*: La principal desventaja de las RNAs son, por un lado, la imposibilidad de determinar cómo se procesa internamente la información y, por otro, su incapacidad para contrastar estadísticamente si los inputs utilizados son significativos. Por tanto, las RNAs no poseen propiedades paramétricas deseables para contrastar hipótesis equivalentes a los tests de la t o la F en los modelos econométricos. Sin embargo, podemos llevar a cabo análisis de sensibilidad de los modelos [Sarle, 2000; Baxt, W. G. y White, H., 1995; Gallant, 1987; Garson, 1991; y Gedeon, 1997]. Estos análisis se realizan calculando derivadas parciales locales y en diferencias en determinados puntos de la distribución, ya que a diferencia de los métodos lineales, la importancia de un factor puede ser distinta en dos puntos de la distribución. Estos

análisis nos permiten llevar a cabo cierta inferencia acerca de la influencia de cada variable explicativa, pero sin duda con unos resultados mucho menos robustos que los procedentes de test paramétricos. Ésta es actualmente la principal limitación de las RNAs.

- *No existen reglas para su implementación*: Como señalan Coakley y Brown (2000). “Determinar el número de capas ocultas o el número de neuronas en cada capa forma todavía parte del *arte* del uso de las redes neuronales”. A pesar de que los numerosos heurísticos existentes pueden ser puntos de partida para la construcción de las RNAs, lo cierto es que actualmente no existe una metodología clara y rigurosa a la hora de aplicar esta técnica. Es más bien la experiencia del investigador en su uso, a través de un proceso ensayo-error, la que en última estancia se aplica en cada caso.
- *No son mágicas*: El desconocimiento acerca de las posibilidades de esta técnica, ha contribuido a generar la falacia de que las RNAs eran una nueva técnica que iba a permitir resolverlo casi todo y además mejor que los métodos tradicionales. Las RNAs no son capaces de crear información si ésta no está contenida en los datos inputs que manejemos. La existencia de un modelo en sentido amplio, esto es, desconociendo la forma funcional pero sí conociendo las variables implicadas en el proceso, sigue siendo una condición necesaria para el éxito en la aplicación de esta herramienta. Debe quedar claro, por tanto, que por ejemplo una RNA no es capaz de adivinar números aleatorios o de descifrar un mensaje cifrado con un algoritmo potente.

1.7. Principales Aplicaciones de las RNAs en Economía.

El principal rasgo de esta tecnología es que ha sido aplicada a un inmenso conjunto de problemas, pertenecientes a diversas áreas del conocimiento, muchos de los cuales parecían ser demasiado complicados de resolver en modelos teóricos o para los cuales se carecía de un modelo matemático apropiado.

Las RNAs han sido aplicadas con éxito, por ejemplo, en el campo de la inteligencia artificial, para el reconocimiento de voz y de escritura; para la clasificación de imágenes de radar, sonar o infrarrojos; la detección de células cancerosas, lesiones neurológicas y cardíacas; prospecciones geológicas; predicción de audiencias televisivas; para jugar a diferentes juegos (ajedrez, backgammon, bridge); compresión de datos; predicciones meteorológicas; detección de explosivos en aeropuertos; control de robots o comprobación de fallos en equipos. A esta lista podríamos añadir un largo etcétera.

Asimismo, durante la última década, las RNAs han experimentado un desarrollo exponencial en diferentes campos de la economía. Existen así, cientos de libros y de publicaciones en revistas prestigiosas que recogen aplicaciones de RNAs en este ámbito, además de comparaciones con otras técnicas¹¹. Ni mucho menos todas las aplicaciones de RNAs resultan superiores a los métodos tradicionales, pero sin duda, el hecho de que se planteen como una alternativa plausible a las metodologías tradicionales constituye por sí mismo un avance prometedor.

Sin duda alguna, las aplicaciones más difundidas han sido en el área financiera y de empresa. Algunos ejemplos son: su uso en modelos de inversión a fin de optimizar carteras; en el mundo del marketing, como instrumento para vender productos a clientes potenciales; en entidades financieras, para predecir el riesgo potencial de los créditos que le son solicitados; análisis de información contable; predicción de indicadores económicos; calificación del riesgo de bonos y entidades; predicción de quiebras bancarias, etc.

Existen así muchos trabajos que recopilan resultados de RNAs en el ámbito financiero. Apostolos et al. (eds.). (1995) recoge numerosos artículos que aplican RNAs en tipos de cambio, comercio de bienes, modelos de riesgo e indicadores macroeconómicos Wong et al. (1997) recoge 207 artículos que aplicaban RNAs entre 1988 y 1995 en distintos ámbitos financieros y contables. O'Leary (1998) analiza 15 artículos donde se aplican RNAs para predecir bancarrotas. Zhang et al. (1998) resume 21 artículos donde las RNAs se usaban como herramientas de predicción. Más recientemente Vellido et al. (1999), analiza 123 estudios entre 1992 y 1998 en muy diversos ámbitos, todos ellos relacionados con el mundo financiero. Zapranis y Refenes (1999) revisan numerosos ejemplos de economía financiera presentando como resumen un completo caso de estudio totalmente desarrollado.

Otra aplicación económica de las RNAs es la de resolver el problema de los valores perdidos. Así, el Ministerio de la Seguridad Social en el Reino Unido utiliza una RNA para rellenar a posteriori el 70% de los 45.000 campos en blanco de su encuesta anual de recursos familiares. Dicha encuesta, con una muestra de 25.000 entrevistas, estaría llena de sesgos si los campos dejados en blanco por los entrevistadores se incluyeran en el análisis de los resultados. Una aplicación futura de las RNAs es la de rellenar los valores perdidos también en el censo del Reino Unido, [Clark, 2001].

¹¹ En torno al 90% de los trabajos en economía con RNAs aplican el MLP tal y como ha sido descrito con anterioridad, si bien las variantes en torno a la arquitectura de la red, algoritmos de aprendizaje o parámetros utilizados en los modelos son múltiples. Teniendo en cuenta este hecho mostraremos a continuación las referencias de algunos de estos libros y publicaciones.

En España, las RNAs están empezando a recibir un creciente interés en la literatura económica a medida que se comprueban las ventajas que este tipo de herramientas pueden presentar frente al tradicionalmente usado tratamiento estadístico y econométrico. Así, podemos ejemplificar este desarrollo en el ámbito de la investigación económica con las recientes tesis doctorales de Galache (1994), que trata sobre las ventajas que presentan las redes neuronales en su aplicación para la investigación económica; Serrano-Cinca (1994), para el análisis de información contable y para la clasificación de bancos en zonas de bancarrota, solvencia o dudosa solvencia, respectivamente; González (1996), para la valoración de opciones financieras o Núñez (1998), en el análisis de riesgos en entidades financieras. Asimismo, Alonso-Rodríguez (1999) muestra como las RNAs superan a los métodos de regresión tradicionales en la predicción del consumo de la economía española.

Dentro del ámbito de la economía pública el MLP ha sido aplicado, y comparado con otras técnicas, en la medición de la eficiencia del transporte público (metro de Londres) en Costa y Markellos (1997). Asimismo, Baker et al. (1999) utiliza redes neuronales para la predicción del gasto público educativo en los Estados Unidos. Gorr et al. (1994) utilizan un MLP para predecir el resultado académico en el primer año de estudios universitarios a partir de resultados en cursos previos y otras variables. Caulkins et al. (1993) emplean redes neuronales para la predicción de criminales reincidentes y Cushing et al. (1997) las aplican para la clasificación de declaraciones (Private Letter Rulings) del Internal Revenue Service en las dos categorías que la ley recoge, empleados y trabajadores independientes. Igualmente, Romay y Santín (2002) aplican RNAs para la predicción del tiempo de baja por incapacidad temporal laboral, en función de la enfermedad y otras variables del trabajador, con el objetivo de detectar posibles fraudes a la Seguridad Social y así ahorrar costes.

En este trabajo de Tesis se aplicarán RNAs a dos problemas de economía pública concretos. En primer lugar, en el capítulo 2 se desarrollarán las ventajas que aporta esta herramienta a la hora de estimar funciones de producción con el objetivo de medir la eficiencia económica. Así, se demostrará que las RNAs son una alternativa válida a los métodos de análisis envolvente de datos y de frontera estocástica en determinados procesos productivos caracterizados por relaciones no lineales entre los inputs y los outputs productivos. En segundo lugar, en los capítulos 3 y 4 se aplicarán RNAs para predecir el rendimiento académico de los alumnos de 13 años a partir de sus variables individuales y contextuales. Se sabe que la extensa información relativa al alumno está relacionada con su resultado académico y que posiblemente esta relación sea no lineal. Sin embargo, se carece de un modelo que indique como interactúan los inputs con los outputs educativos. Nuestro objetivo será usar las predicciones estimadas para calcular mediante un modelo en dos etapas el valor añadido aportado por la escuela al individuo.

Esta metodología propuesta, basada en el uso de las RNAs, es nueva en la literatura y es por tanto la principal aportación de esta tesis doctoral.

1.8 Consideraciones Finales

Existen otros métodos estadísticos que pueden ser usados para la construcción de modelos de clasificación o predicción con funciones flexibles y no lineales. Algunos de ellos serían los siguientes:

- Regresión polinómica [Eubank, 1999]. Consiste en encontrar los coeficientes de las ecuaciones normales que representarán finalmente un polinomio.
- Regresión con series de Fourier [Haerdle, 1990]. Utiliza la transformada de Fourier para la predicción en series temporales.
- *Wavelets* [Donoho y Johnstone, 1995]. Método no paramétrico para el ajuste de series temporales.
- Modelos basados en árboles de decisión [Haerdle, 1990]. Sistema de inducción de reglas y patrones en los datos.
- *Splines regression* [Eubank, 1999]. Ajusta los datos mediante distintos segmentos lineales.
- *Projection Pursuit Regression*. [Friedman y Stuetzle, 1981]. Ajusta los parámetros de un modelo aditivo mediante un algoritmo iterativo
- *Group Method of Data Handling*.. [Ivakhnenko, 1995 y Liao, 1992]. Basado en algoritmos genéticos, busca el polinomio que más se ajusta a los datos proporcionados.

¿Por qué usar entonces un MLP en lugar de alguno de estos métodos?. El MLP tiende a ser útil en los mismos casos que la técnica de *projection pursuit Regression*. La principal ventaja del MLP sobre *projection pursuit regression* [Sarle, 2000] es que el cálculo de los valores de las predicciones es más simple y rápido. Además el MLP obtiene mejores resultados en la aproximación de funciones flexibles y no lineales pero sin demasiadas “rarezas” (por ejemplo discontinuidades).

Respecto a los estudios que comparan diversas técnicas los resultados, como ya se ha apuntado, no son definitivos. Por cada resultado que señale la superioridad de una técnica, se puede encontrar otro que muestre el resultado contrario y muchos otros en los que no se puede demostrar la superioridad de un método frente a otros. Este hecho ha producido cierto *pesimismo* en estas técnicas, frente al desmedido optimismo inicial que surgió a principios de

los años 90, entre aquellos que pensaban que las nuevas técnicas de minería de datos solucionarían cualquier problema mejor que las técnicas tradicionales.

En general la superioridad de una técnica sobre otra estará condicionada al problema planteado, la existencia o no de un modelo económico teórico y las características de la base de datos y la disponibilidad de variables que empleemos. Se debe, por tanto, abandonar la idea de que existe una única técnica que genere resultados superiores ante todo problema económico al que nos enfrentemos. Se considera, sin embargo, que es importante el conocimiento de distintas alternativas estadísticas a la hora de abordar un mismo problema.

Por tanto, y en conclusión, el investigador económico, una vez tenga planteado un problema, parece adecuado que acudiera para resolverlo a la técnica estadística o econométrica tradicional. No obstante si los resultados obtenidos no son satisfactorios y/o si el problema reúne las características necesarias para la aplicación de RNAs, principalmente posibles no linealidades entre las variables del modelo, podrá usar esta herramienta u otra como alternativa. Ello nos va a permitir comparar distintos resultados dando mayor robustez a los resultados obtenidos por la técnica que mejor sirva para los propósitos de la investigación.

**CAPITULO 2: LA MEDICIÓN DE LA EFICIENCIA TÉCNICA
MEDIANTE REDES NEURONALES**

2.1. Introducción

El objetivo de este capítulo es mostrar como las RNAs son una alternativa no paramétrica a las técnicas tradicionales para la medición de la eficiencia técnica. Su ventaja comparativa se hace patente en contextos donde existan relaciones no lineales entre las variables y la cantidad de ruido en los datos es moderada, de tal forma que se pueda detectar la estructura de la función de producción que subyace bajo los datos observados.

La estimación de la función de producción de un bien o servicio público es un problema frecuente en el ámbito de la economía pública que tiene fundamentalmente dos aplicaciones económicas. La primera sería la predicción de resultados a partir de distintos factores productivos o inputs con el objeto de planificar e implementar políticas públicas; y la segunda, estrechamente relacionada con la anterior, es la medición de la eficiencia productiva. Ambos objetivos pretenden, como fin último, ofrecer una valiosa información al gestor para la posterior toma de decisiones.

Sin embargo, por sus especiales características, el análisis de la eficiencia con la que actúa el sector público presenta mayores dificultades que para el ámbito privado. No obstante, los objetivos macroeconómicos de equilibrio presupuestario, crecimiento económico y contención de precios, han llevado a los economistas a interesarse detalladamente por la gestión del gasto público, desarrollándose así en los últimos años campos como la economía de la defensa, de la sanidad, del transporte o de la educación.

En cuanto a las dificultades en la estimación de funciones de producción pública y siguiendo a Albi (1992), en primer lugar los objetivos públicos son múltiples, complejos y en ocasiones difusos y no claramente definidos. Además, los bienes y servicios que produce el sector público se valoran generalmente sólo por sus costes, ya que en bastantes casos la actuación pública carece de mercados en competencia perfecta que determinen los precios. Por otro lado, la utilización de los precios de mercado no sería siempre la adecuada en el sector público. En un análisis de la eficiencia pública, los precios deben reflejar el “valor social” de los recursos y la producción. Aunque teóricamente el precio de un mercado competitivo debe aproximarse al coste marginal del bien y éste reflejar su valoración social, en la práctica cabe encontrar divergencias a este principio.

Ejemplos de estas divergencias son los efectos externos, los bienes públicos, las imperfecciones del mercado o los impuestos. El criterio de rentabilidad privada no se puede usar sin tener en cuenta otros parámetros en la actividad pública. Una empresa rentable que

contamina puede dejar de obtener beneficios si se considera un coste el efecto externo negativo que origina. Una empresa municipal de transporte urbano que aplique sus tarifas de acuerdo a sus costes medios, puede presentar pérdidas financieras que se tornen en beneficios desde el punto de vista de la sociedad y de los efectos externos positivos obtenidos, en la descongestión del tráfico y el ahorro de tiempo y aumento de horas de trabajo que genera.

De igual manera y tal y como señala Albi (2000, pág. 56) “Las políticas públicas tienen procedimientos de aplicación complejos y plantean dificultades políticas, económicas y burocráticas. La intervención pública se desarrolla a lo largo del tiempo, dentro de una realidad social, y su gestión, desde la toma de decisiones hasta la ejecución de las mismas, presenta complicaciones que ponen en peligro la eficiencia y la equidad necesarias para el bienestar social”.

El objetivo de este capítulo es el de presentar a las RNAs como una alternativa para el problema empírico de la medición de la eficiencia, comparando sus posibles ventajas e inconvenientes con las técnicas tradicionales.

2.2. El Concepto Económico de Eficiencia

En general, los economistas muestran un elevado grado de consenso en considerar que una asignación es eficiente en sentido de Pareto, cuando no es posible reasignar los recursos existentes de tal forma que alguno (algunos) mejoren sin que otro (otros) empeoren. La situación de óptimo se garantiza si se cumplen las condiciones que caracterizan la eficiencia en la producción, la eficiencia en el intercambio y la eficiencia global.

Suele identificarse a Debreu (1951) como el primer autor en plantearse el problema de la construcción de una función de producción empírica basada en los datos observados y medir la eficiencia de cada unidad productiva en relación con aquella. Debreu sugirió una medida consistente de eficiencia con su concepto de “coeficiente de utilización del recurso”. Debreu apunta que una medida natural de la eficiencia productiva es la ratio del coste del nivel óptimo de *inputs* y de *outputs* observados.

$$\frac{px^*}{px^0}$$

Donde p es el precio de los *inputs* usados en el proceso productivo, x^* es la cantidad óptima de *inputs* y x^0 es el nivel de *inputs* observado.

Existen dos ventajas derivadas al emplear la medida de eficiencia de Debreu. La primera es que es independiente de la escala de los datos. Su valor estará comprendido entre cero, para el caso más ineficiente, y uno, para el caso más eficiente. La segunda ventaja es la introducción explícita de los precios de los factores productivos. Así, en la ecuación anterior, px^* es el coste asociado a la combinación de inputs eficientes mientras que px^0 es el coste observado.

Al medir las ratios de costes productivos, el concepto de Debreu mide indirectamente el coste de la ineficiencia productiva. Así, uno menos la ratio de Debreu es la fracción de costes observados que podrían ser reducidos si la ineficiencia productiva fuese eliminada. El principal problema de este indicador es el de obtener, a partir de una observación empírica, el nivel del vector de *inputs* óptimo x^* .

Posteriormente, el artículo seminal de Farrell (1957) retomó el trabajo de Debreu y definió una medida alternativa de eficiencia que ha influido notablemente a lo largo de la literatura. El propósito de Farrell era, en sus propias palabras, [Farrell, 1957, pág. 253] “suministrar una medida satisfactoria de la eficiencia productiva y mostrar como aquella puede ser calculada en la práctica (...) para cualquier organización productiva, desde un taller a la economía global”. Farrell observó que podía obtenerse una medida análoga de eficiencia productiva comparando los *outputs* óptimos observados frente a los *inputs* óptimos observados.

$$\frac{wy^*}{wy^0}$$

Dada una función productiva, podemos derivar el nivel de *output* óptimo predicho por la función y^* , que junto al *output* observado y^0 , y el precio del *output* w , definen el concepto de eficiencia de Farrell. Esta medida tiene las mismas ventajas que el concepto de Debreu pero la construcción empírica de la medida es una aplicación directa de la frontera productiva.

Por otro lado, la mayor parte de los economistas entienden que la ineficiencia tiene básicamente dos dimensiones:

- La ineficiencia técnica, como su propio nombre indica, es un concepto matemático-tecnológico que intenta analizar los procesos productivos y la organización de tareas, fijando su atención en las cantidades de *inputs* y *outputs* utilizadas y no en los costes o precios de los mismos. Puede expresarse por tanto en términos de *outputs* como en

términos de *inputs*. En el primer caso, indicaría el logro del máximo producto o servicio posible para una combinación de factores. En el segundo, indica la cantidad mínima de inputs requerida combinados en una determinada proporción para un nivel dado de producto o de servicio.

- La eficiencia asignativa implica alcanzar el coste mínimo de producir un nivel dado de producto cuando se modifican las proporciones de los factores de producción utilizados de acuerdo con sus precios y productividades marginales. Alternativamente, se puede definir como la obtención de una cantidad máxima de producto manteniendo el coste a través del reajuste de los factores de producción según sus costes de uso. La eficiencia asignativa supone que se ha logrado la eficiencia técnica.

En definitiva, en la medición de la eficiencia técnica se parte de una proporción concreta de factores. Esta proporción puede variar si, por ejemplo, se utiliza una tecnología distinta. Pero no por precios o productividades como ocurre con la eficiencia asignativa.

Una forma alternativa de definir la eficiencia técnica y asignativa sería la siguiente: en el caso de la eficiencia técnica, supongamos que nos enfrentamos ante un proceso de producción caracterizado por la utilización de dos inputs para la obtención de un único output, considerando la tecnología fija, estaríamos actuando eficientemente desde el punto de vista técnico cuando nos encontráramos en un punto sobre la isocuanta unitaria que caracteriza la tecnología frontera.

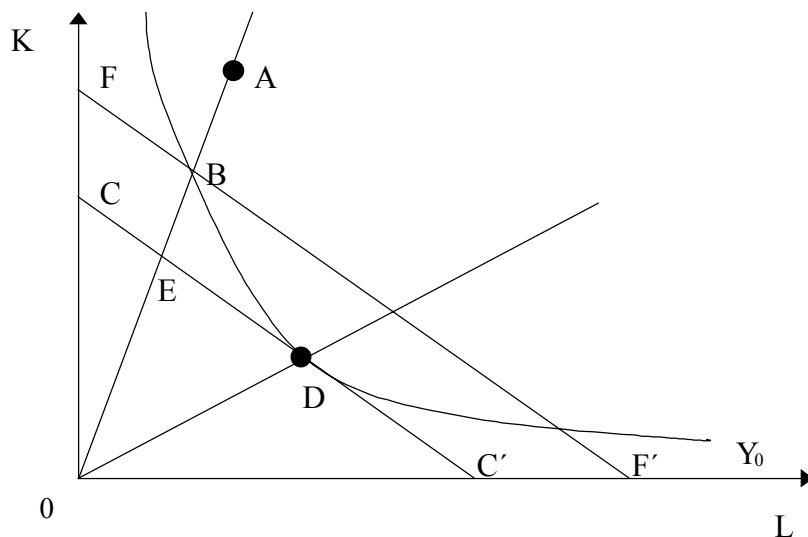
Por otro lado, si suponemos ahora que la tecnología puede variar, o lo que es lo mismo podemos utilizar diferentes combinaciones de inputs, se define una situación como asignativamente eficiente, cuando siéndolo desde el punto de vista técnico, estamos empleando la menor cantidad de recursos posibles, dado los precios de los factores. Nos encontramos así sobre la curva isocoste más baja.

En la figura 2.1, se recoge la situación de una planta productiva que para producir una unidad de output (Y) utiliza la combinación de dos factores productivos, trabajo (L) y capital (K), representada por el punto A . Suponiendo que la función de producción es conocida y que existen rendimientos constantes a escala, la curva Y_0 es lo que Farrell denomina isocuanta unitaria, que no es más que el conjunto de combinaciones de factores productivos necesarios para producir una unidad de producto.

Así, dada la situación descrita en la figura anterior podemos deducir que la unidad productiva A no es técnicamente eficiente en sentido de Farrell, ya que la isocuanta Y_0 indica

que el output producido por la entidad A podría ser obtenido utilizando una proporción OB/OA de los inputs que realmente utiliza y sin variar la combinación de los mismos.

Figura 2.1: Eficiencia Técnica y Eficiencia Asignativa.



Fuente: Albi et al. (2000), pág. 244.

Farrell define el cociente OB/OA como el índice de eficiencia técnica de la unidad productiva A. Se observa como la medida de eficiencia técnica de Farrell tomará el valor 1 si la entidad es técnicamente eficiente y valores más próximos a cero cuanto más ineficiente sea la entidad valorada.

La figura anterior permite también establecer la medida de la eficiencia asignativa. Si se supone que los precios de los factores productivos están representados por la pendiente de la combinación lineal de puntos CC' , el punto D es aquel en el que se minimiza el coste. Tanto B como D son eficientes técnicamente por estar situados sobre la isocuanta, pero los costes de producción en D son tan sólo una fracción OE/OB de los costes de producción en D y por eso Farrell considera al cociente OE/OB como la eficiencia asignativa de B. Esta medida de eficiencia asignativa, que Farrell refiere al punto B, mide también la eficiencia asignativa de la entidad A que estamos evaluando. En efecto, la eficiencia precio de A mide exclusivamente el exceso de costes en que se está incurriendo por combinar los *inputs* de una forma diferente a la óptima. Esto indica que la valoración de la eficiencia asignativa pura precisa la eliminación de la eficiencia técnica lo que, en nuestro caso, equivale a situar la entidad A en el punto B y valorar el exceso de costes en esa situación, representada por el salto de B a D.

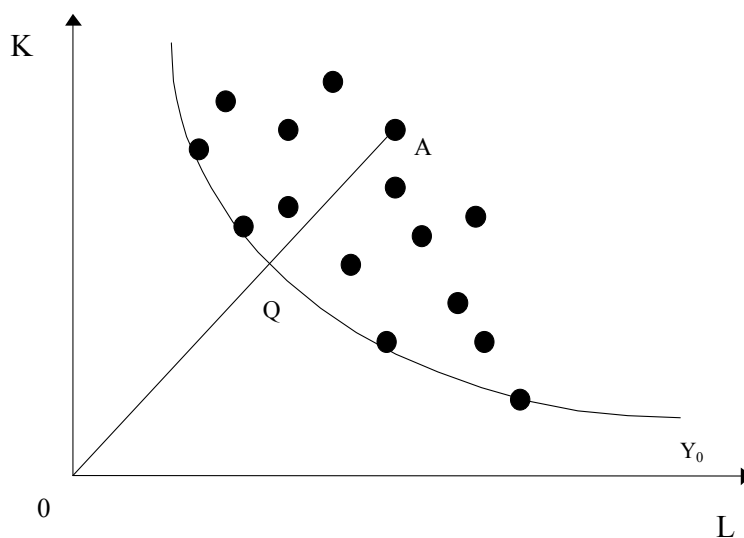
Farrell continua su trabajo seminal definiendo una medida de eficiencia global como el producto de la eficiencia técnica y eficiencia asignativa.

$$\text{Eficiencia Global (OE/OA)} = \text{Eficiencia Técnica (OB/OA)} * \text{Eficiencia Asignativa (OE/OB)}$$

El análisis efectuado a partir de la figura 2.1 anterior ha partido del conocimiento de la función de producción, representada por medio de la isocuanta Y_0 , la cual constituye, como se ha visto, el punto de referencia para llevar a cabo las estimaciones. Las situaciones reales, sin embargo, no se suelen caracterizar por el conocimiento de esas relaciones técnicas, tal y como comprobaremos posteriormente ocurre para el caso de la educación no universitaria, lo cual complica la medición operativa de la eficiencia. Este problema es considerado también por Farrell, quien propone un método original de estimación de la función de producción a partir de los datos de las entidades implicadas en la valoración. Es justamente la estimación de lo que él denomina una *frontera de producción empírica* lo que ha tenido una mayor repercusión en la literatura posterior sobre la medición de la eficiencia en aplicaciones reales.

En la figura 2.2 se representan, mediante puntos, las combinaciones de factores inputs utilizadas por diferentes unidades productivas para obtener una unidad de output.

Figura 2.2: La medición empírica del concepto de eficiencia.



Farrell impone dos condiciones a la isocuanta que va a estimar: que sea convexa y que no tenga en ningún punto pendiente positiva. La primera, normalmente planteada en teoría microeconómica, significa que si dos puntos se pueden alcanzar en la práctica, entonces también se podrá obtener cualquier otro que sea una combinación ponderada de aquellos. La segunda,

por su parte, es necesaria para asegurar que el aumento de los factores utilizados no implicará nunca una reducción en la cantidad de producto¹².

A partir de estas condiciones es fácil deducir que la isocuanta eficiente está representada por la curva Y_0 ; es decir, por el conjunto de puntos más próximos al origen (las unidades más eficientes en términos relativos) que puedan ser unidos a través de una curva convexa que no tenga en ningún punto una pendiente positiva. Determinada la isocuanta eficiente, el proceso de medir la eficiencia de cualquier unidad productiva es el que hemos especificado en el comentario de la figura anterior. Como allí destacábamos, se trata de comparar cada entidad que no pertenece a la isocuanta frontera con otra entidad eficiente que utilice los factores productivos en la misma proporción (esto es, que se encuentre en el mismo radio vector desde el origen).

En general, esa comparación se realizará con unidades hipotéticas que, empleando los mismos factores en la misma proporción, se encuentran sobre la isocuanta eficiente, pero que no se corresponden con ninguna observación real.

Así, en la figura anterior, la eficiencia del punto A se mide comparando los factores que utiliza con los que usa la unidad ficticia Q. Farrell señala que la esencia de su propuesta radica precisamente en la construcción de esas unidades hipotéticas y no en la representación de la isocuanta.

2.3. La Medición Empírica de la Eficiencia.

Para la medición empírica de la eficiencia técnica en educación se han aplicado fundamentalmente, desde el trabajo seminal de Farrell, dos tipos de técnicas que podemos clasificar a grandes rasgos en paramétricas y no paramétricas. A continuación comentaremos los principales rasgos de ambas metodologías. Para una extensa revisión de los trabajos de eficiencia en educación puede acudir a Worthington (2001).

2.3.1. Técnicas Paramétricas

Supongamos que la función de producción puede ser representada por una función lineal, normalmente una función Cobb-Douglas (ecuación 2.1) o Translog¹³:

¹² Este supuesto impide la existencia de posibles *áreas de congestión o saturación*, donde un aumento en un input supondría una disminución del output, a lo largo de la función de producción. Como veremos, a lo largo de este trabajo relajaremos este supuesto ya que estas zonas son posibles empíricamente.

$$\ln y_i = \beta_0 + \sum_i \beta_i \ln x_i + \varepsilon \quad (\text{ec. 2.1})$$

Para medir las ineficiencias productivas mediante la metodología econométrica la especificación de la estructura del error tomaría una forma ligeramente distinta (ecuación 2.2):

$$\ln y_i = \beta_0 + \sum_i \beta_i \ln x_i - (v - \varepsilon) \quad (\text{ec. 2.2})$$

Donde y_i es la producción de la firma i , x_i el vector de inputs y β sería un vector de parámetros desconocidos. El error que recoge el término v representa la ineficiencia productiva mientras que el término ε representa ruido blanco. De esta forma la frontera productiva tiene un carácter estocástico.

El problema del análisis empírico consiste en averiguar en cuánto desplazar hacia arriba la estimación media hasta alcanzar una frontera productiva. Además del uso de mínimos cuadrados ordinarios (OLS) Aigner *et al.* (1977) y Meeusen y van den Broeck (1977) sugirieron simultáneamente una función de máxima verosimilitud (SFA) de la forma general:

$$L = l(\alpha_0, \beta_i, \lambda, \sigma \mid y, x_i) \quad (\text{ec. 2.3})$$

Donde y , x_i serían los factores output e input respectivamente, α_0 y β_i los parámetros a estimar de la ecuación (2.3) y $\lambda = \sigma_v / \sigma_e$ y $\sigma = \sigma_v + \sigma_e$. Siendo σ_v el componente del error estándar que recoge la ineficiencia y σ_e el componente que recoge el ruido en los datos. La forma funcional específica de (2.3) depende de los supuestos que hagamos acerca de la distribución de v y e ; por ejemplo¹⁴ suponer $v \sim \text{iid } N(0, \sigma_v^2)$ y $e \sim \text{iid } N(0, \sigma_e^2)$.

El error aleatorio recogería por una parte todos aquellos factores ajenos al control de la organización, incluyendo tanto factores directamente relacionados con el entorno en el que se desarrolla el proceso, conocidos como variables ambientales, errores de especificación del modelo y por último el error estadístico.

¹³ Sin duda la forma paramétrica de función de producción más empleada en la construcción de fronteras ha sido la *translog* que fue propuesta originalmente por Heady y Dillon (1961, 205-208) y popularizada posteriormente por Christensen *et al.* (1971). La especificación de su forma funcional sería:

$$\ln y = \beta_0 + \sum_{n=1}^N \beta_n \ln(x_{ni}) + \frac{1}{2} \sum_{m=1}^N \sum_{n=1}^N \beta_{mn} \ln(x_{mi}) \ln(x_{ni})$$

¹⁴ Para extensiones de este modelo puede acudir a Battese y Coelli (1988, 1992 y 1995).

2.3.2. Técnicas No Paramétricas.

En segundo lugar, los modelos de Análisis Envolvente de Datos¹⁵ (DEA), fueron desarrollados, a partir del trabajo seminal de Farrell veinte años después por Charnes, Cooper y Rhodes (1978). Mediante la utilización de técnicas de programación matemática lineal, el DEA compara la eficiencia de un conjunto de unidades que producen outputs similares a partir de un conjunto común de inputs.

En su versión inicial más básica la eficiencia de la unidad que se pretende evaluar se define como la ratio de la suma ponderada de *outputs* con respecto a la suma ponderada de *inputs*:

$$\text{Max } h_0 = \frac{\sum_{r=1}^s U_r Y_{r0}}{\sum_{i=1}^m V_i X_{i0}}$$

Sujeto a:

$$\frac{\sum_{r=1}^s U_r Y_{rj}}{\sum_{i=1}^m V_i X_{ij}} \leq 1$$

$U_r, V_i \geq 0 ; r = 1 \dots s ; i = 1 \dots m$

Donde:

Y_{r0} = Cantidad de output r producido por la unidad evaluada.

X_{i0} = Cantidad de input i consumido por la unidad evaluada.

Y_{rj} = Cantidad de output r producido por la unidad j.

X_{ij} = Cantidad de input i consumido por la unidad j.

U_r = Ponderación asignada al output r.

V_i = Ponderación asignada al input i.

El problema fraccional formulado consiste en encontrar el conjunto de ponderaciones que maximizan el valor de los outputs de la unidad analizada con respecto a los inputs empleados, con la restricción de que aplicando estas mismas ponderaciones a las restantes unidades, ninguna debiera tener una relación output/input mayor que uno. Si, sujeto a esta restricción, fuera factible encontrar un conjunto de ponderaciones con las que la ratio de eficiencia de la unidad productiva evaluada sea igual a 1, entonces será considerada eficiente.

¹⁵ Utilizaremos para referirnos a esta técnica las siglas inglesas de *Data Envelopment Analysis*, DEA.

En caso contrario la unidad será considerada ineficiente, ya que incluso con el conjunto de ponderaciones más favorable puede encontrarse una ratio de eficiencia mayor.

El DEA presenta una serie de características que le hacen especialmente útil para la evaluación de la eficiencia de las unidades productivas. La principal ventaja de esta técnica es que a diferencia de los métodos paramétricos¹⁶ no impone ninguna forma funcional *a priori* sobre los datos. Esto es, la eficiencia no es evaluada sobre la base de una función de producción ideal; los índices de eficiencia se calculan a partir de las mejores prácticas observadas.

En segundo lugar, el modelo DEA se adapta al carácter multidimensional de determinadas actividades productivas así como a la ausencia de precios de mercado para determinados inputs. Además, es un método sumamente flexible que establece pocos supuestos a la hora de trazar la frontera productiva¹⁷ y permite introducir factores que están fuera del control de las unidades analizadas, [Banker et al., 1986]. Ofrece además objetivos a las unidades evaluadas ineficientes bien a través del aumento de outputs o la reducción de los inputs bajo el control de la unidad productiva y mediante la eliminación de las holguras existentes¹⁸.

Una vez indicadas las principales ventajas del modelo, se señalan a continuación alguna de sus debilidades. En general el DEA, dado su carácter determinista, estima la frontera productiva sin realizar ningún tipo de supuesto acerca de la forma de la distribución del término de error. Por otro lado, el carácter no paramétrico del modelo hace que las variables incluidas no posean propiedades estadísticas y que la contrastación de hipótesis resulte imposible.

Un aspecto que es importante resaltar es que las medidas de eficiencia derivadas del DEA pueden ser muy sensibles al número de observaciones y al número de variables incluidas en el modelo. A medida que aumenta la relación variables (explicativas más explicadas)/tamaño de la muestra¹⁹, la habilidad del DEA para discriminar entre las firmas disminuye significativamente, ya que como señalan Yunos et al. (1997) se vuelve más probable que una determinada firma encuentre algún conjunto de ponderaciones de outputs e inputs que la haga

¹⁶ Para una excelente discusión comparativa entre las ventajas e inconvenientes que plantean las técnicas paramétricas y no paramétricas así como las fronteras deterministas y estocásticas puede acudir a Lovell (1993).

¹⁷ Para una revisión de los beneficios que implica imponer restricciones sobre las ponderaciones en el DEA puede acudir a Pedraja et al. (1997).

¹⁸ Aunque como señala Alvarez (2001) págs. 36-37, “las reducciones radiales de inputs no convierten por sí mismas a las unidades productivas en eficientes ya que normalmente detrás de las causas de la ineficiencia no suelen estar las cantidades de inputs que se usan sino cómo se usan”.

¹⁹ Un conocido heurístico señala que el número de observaciones debería ser mayor al número de variables input más el número de variables output multiplicado por cuatro. Si bien estos heurísticos carecen de toda justificación teórica, al menos nos dan una idea de los grados de libertad necesarios para aplicar el DEA.

aparecer como eficiente. Esto es, muchas firmas pueden ser consideradas 100% eficientes no porque dominen a otras firmas, sino simplemente porque no hay otras firmas o combinaciones de firmas con las que puedan ser comparadas en tantas dimensiones. La limitación del número de variables puede llegar a suponer una mala especificación del modelo productivo.

En resumen, y como señala el profesor Lovell (1993); “Obviamente lo deseable sería convertir los modelos de programación lineal en estocásticos y hacer que los modelos econométricos fueran más flexibles en su estructura paramétrica”.

2.4. Las Redes Neuronales Artificiales como Herramienta para la Medición de la Eficiencia Técnica.

En el contexto de la medición de la eficiencia técnica, las RNAs son una herramienta más flexible que podemos sumar al conjunto de alternativas disponibles actualmente, descritas previamente. Se puede clasificar esta técnica como no paramétrica, dado que no se asume a priori ninguna restricción en torno a la forma de la función de producción que se quiere aproximar, estocástica, y con la posibilidad de llevar a cabo inferencias como la construcción de intervalos de confianza²⁰ para los índices de (in)eficiencia.

El MLP es, de todas las alternativas, la técnica más flexible y podría ser considerada desde un punto de vista econométrico como una regresión no lineal. Según Cheng et al. (1994), las redes neuronales son modelos matemáticos generales que en muchas ocasiones son equivalentes en su planteamiento a modelos habituales en econometría, como el modelo lineal general, modelos logit o probit, modelos AR de series temporales, etc., pero con una terminología y métodos de estimación específicos.

Una visión comparada de las alternativas para medir eficiencia se sintetiza en la tabla 2.1, de la que se obtiene la lógica conclusión de que ninguna de las alternativas es globalmente superior al resto, de forma que las características de cada problema particular determinarán la elección de la técnica a utilizar.

²⁰ Para la construcción genérica de intervalos de confianza y de predicción en RNA puede acudir a Hwang y Ding (1997), De Veaux *et al.* (1998) y Rivals y Personaz (2000).

Tabla 2.1: Comparación de los distintos enfoques para la medición de la eficiencia.

	OLS	SFA	DEA	MLP
Supuestos estadísticos	Altos	Altos	Medios	Ninguno
Flexibilidad	Baja	Media	Media	Alta
Estudios de eficiencia	Alta	Alta	Alta	Baja
Significación estadística	Sí	Sí	No	No
Interpretabilidad	Alta (con r.c.e)	Media	Media	Media
Proyección, Generalización	Media	Media-Alta	Ninguna	Alta
Multi-input, Multi-output	No	No	Sí	Sí
Coste de análisis	Bajo	Bajo	Bajo	Alto
Tipo de Frontera obtenida	Estoc/Determ	Estocástica	Determinista	Estocástica
Tipo de Modelo	Paramétrico	Paramétrico	No paramétrico	No paramétrico

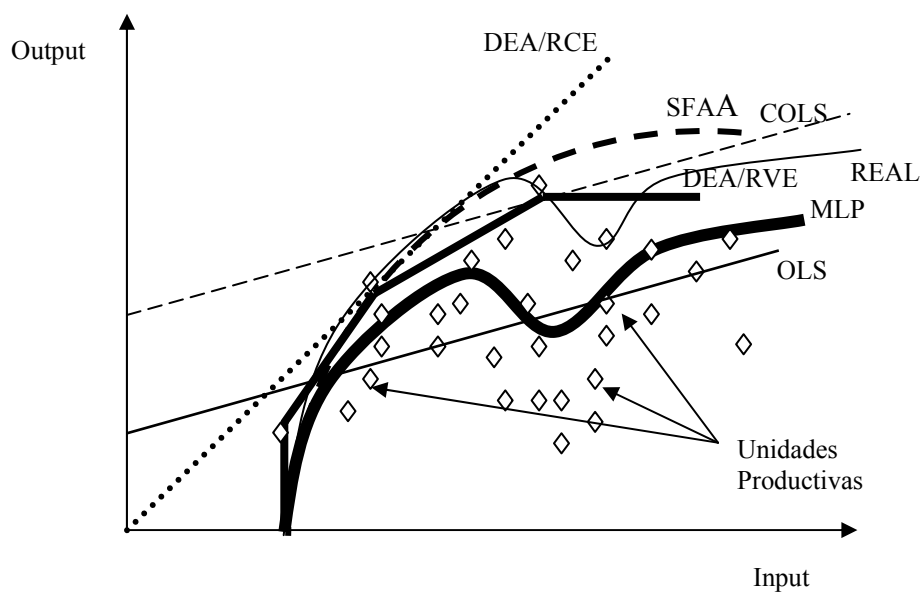
Fuente: Elaboración propia.

Los métodos de regresión son los únicos que permiten realizar inferencia estadística y sus resultados son más fácilmente interpretables (bajo rendimientos constantes a escala), si bien ya hemos comentado que el MLP es también susceptible de cierto análisis de sensibilidad de las variables explicativas o inputs utilizados. Por otra parte, los estudios de eficiencia con regresiones, SFA y DEA han sido aplicados en multitud de campos y son bien conocidos, mientras que el MLP está aún introduciéndose como herramienta alternativa y por tanto sus resultados son todavía mucho menos conocidos. Sin embargo el MLP es la única técnica completamente flexible para aproximar funciones de producción y que no realiza supuestos a priori acerca de la distribución de los datos.

Además, dado que el MLP no envuelve completamente los datos sino que nos permite generalizar la relación entre múltiples inputs y outputs, nos va a posibilitar, al igual que los modelos econométricos, realizar predicciones para fuera de la muestra, esto es, para futuros períodos. Se alcanzaría un resultado semejante al de los modelos DEA, para medir la eficiencia en un único período, a través del ajuste de los datos de la muestra, esto es, con una red neuronal sobreentrenada, [Costa et al., 1997]. Por último, el coste en tiempo de implementación de un MLP es superior al de los otros métodos convencionales

En la figura 2.3 se representan esquemáticamente las posibilidades de estimación de la eficiencia técnica. En ella se trata de recoger el carácter lineal de la aproximación econométrica y la gran flexibilidad de las RNAs. Debe quedar claro que la función de producción real propuesta es tan solo un ejemplo que trata de ilustrar como un MLP es capaz de aproximar cualquier forma funcional que en teoría pudiera existir en la realidad.

Figura 2.3: Fronteras productivas: Posibilidades teóricas de estimación.



Así, se puede contemplar como todos los modelos econométricos trazan como frontera líneas rectas o hiperplanos, asumiendo así rendimientos constantes a escala. El método de estimación por máxima verosimilitud, donde se asume una forma funcional como la translog, permite una aproximación de frontera estocástica más flexible aunque paramétrica. Si bien el modelo DEA puede considerar rendimientos variables a escala a la hora de estimar la frontera productiva, no nos proporciona información ni sobre los parámetros ni sobre la pendiente trazada. Además, no puede ser usado para realizar estimaciones de producción en períodos futuros.

El MLP es la única técnica totalmente flexible y que podría incluso detectar zonas productivas congestionadas en algún factor productivo en la función de producción real (REAL en la figura 2.3), donde un incremento del mismo supondría un empeoramiento en el resultado²¹. En la figura 2.3 este suceso se representa con un tramo de la pendiente de la frontera real con signo negativo.

²¹ Este caso no es tan extraño como en un primer momento podríamos pensar. Podemos señalar como ejemplo intuitivo de este tipo de no linealidades productivas el caso de la educación. En el aprendizaje no universitario dedicar un número de horas, por ejemplo a estudiar y aprender con la ayuda de un ordenador, puede ser mejor que no dedicar nada de tiempo. No obstante, podría llegar un momento en el que más horas dedicadas a esta actividad acapare la atención del alumno y le distraiga del estudio de otras asignaturas suponiendo un perjuicio en el rendimiento global del mismo. Parece, por tanto, de suma importancia detectar estas saturaciones, nada triviales, en los procesos productivos. En otro caso podríamos caer en conclusiones erróneas que supongan tomar decisiones de asignación de recursos nefastas.

A continuación, se recogen diversos trabajos en esta materia, con un resumen de los planteamientos y resultados alcanzados:

- Costa y Markellos (1997) proponen por primera vez el uso del MLP para la construcción de funciones de producción y la medición de la eficiencia técnica y lo aplican empíricamente para medir la eficiencia en el servicio público de metro de Londres con datos anuales sobre 2 inputs, flota y trabajadores, y 1 output, total de kms. recorridos. Estos autores crean una muestra sintética para la estimación de la frontera, añadiendo ruido blanco iid $N(0, \sigma^2)$ a los datos iniciales. Concluyen que las RNAs presentan resultados similares a COLS y DEA, pero presentan ventajas como facilitar la toma de decisiones, estudiar el impacto de rendimientos variables a escala versus rendimientos constantes o detectar áreas de congestión.
- Guermat y Hadri (1999) estudian la bondad de las distintas técnicas de medición de eficiencia mediante simulación de Monte Carlo. Los datos (2 inputs y 1 output) son generados a partir de diversas funciones: Cobb-Douglas, CES (Constant Elasticity Substitution) y generalizada Leontief, y consideran varias funciones: RNA, Cobb-Douglas, translog, CES y generalizada de Leontief. Las RNAs superan a las especificaciones translog y Cobb-Douglas, si bien los resultados no son igual de satisfactorios cuando los datos se generan a partir de funciones de Leontief o CES por problemas de especificación. Además, la incorrecta especificación de la forma funcional (con RNA y translog) no afecta significativamente a la eficiencia media, máxima y mínima, pero sí a ineficiencias puntuales y ordenaciones.
- Santín y Valiño (2000) estudian la eficiencia en educación estimando, a nivel de estudiante, la función de producción a través de una red neuronal. En su aplicación con datos de 7454 estudiantes y 12 inputs, las RNAs superan los resultados predictivos de la aproximación econométrica en la estimación de la función de producción.

A continuación se desarrollan dos experimentos donde se comprueba el ajuste que alcanzan las distintas técnicas a las funciones teóricas construidas para tal propósito.

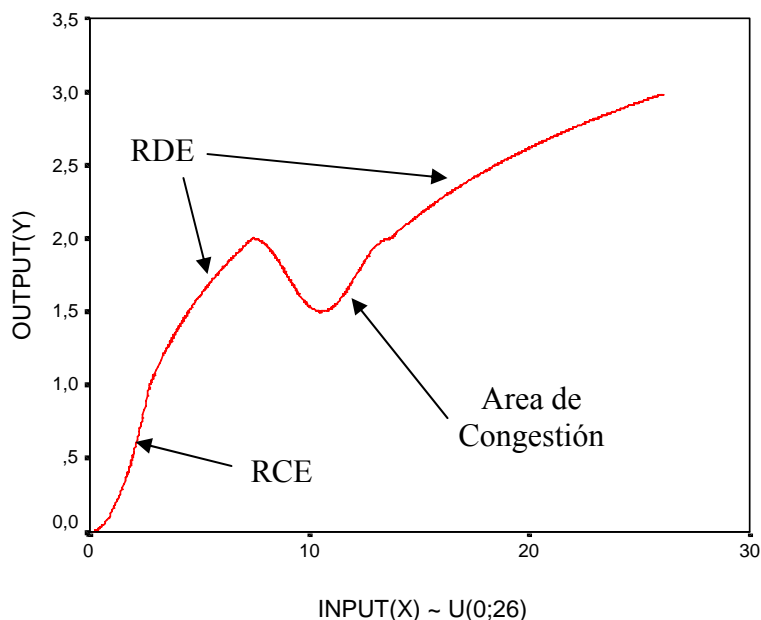
2.4.1. Ejemplo 1: La Medición de la Eficiencia Técnica mediante RNAs.

Sea $F(x)$ la siguiente función de producción continua con un input y un output:

$$F(x) = \begin{cases} \left(\frac{x}{e}\right)^2 & \text{if } x \in [0, e] \\ \text{Ln}(x) & \text{if } x \in [e, e^2] \\ A * \text{COS}(x - e^2) + 2 - A & \text{if } x \in [e^2, e^2 + \pi], \text{ donde } A = 0,25 \\ \text{Ln}(x - 2\pi) & \text{if } x \in [e^2 + \pi, 26] \end{cases}$$

En esta función (figura 2.4), incorporamos todas las posibilidades de rendimientos a escala: en el primer tramo, rendimientos crecientes (RCE); en el segundo y cuarto, rendimientos decrecientes (RDE); y en el tercero, una situación en la que aumentos de input significan reducciones de output, interpretado como área congestionada.

Figura 2.4: La función de producción no lineal.



Sin embargo, el objetivo de esta función es tratar de ilustrar lo que ocurre con la estimación de la eficiencia en el caso de que la función de producción no tuviera una especificación paramétrica habitual. En este punto, se piensa que en la vida real pueden existir

múltiples no linealidades, además de las que introduce la teoría económica con los rendimientos variables a escala. Estas no linealidades pueden existir no solo entre un input y un output sino entre distintos inputs. ¿Debemos considerar como factible la presencia de este tipo de fenómenos?

Costa et al. (1997), analizando la función de producción en el metro de Londres desde 1970 a 1994 con un MLP, encontraron este tipo de no linealidad. Así muestran la existencia de una pendiente negativa entre distintos inputs (tamaño de la flota y número de trabajadores) y entre un input y un output (millones de km. anuales recorridos en relación a la flota). Por otro lado, Baker (2001) concluye, al analizar la función de producción educativa con diferentes tipos de RNAs, que se pueden obtener ganancias en los resultados a partir del tamaño de las clases, reduciendo su tamaño de 14 a 10 alumnos, pero también incrementando el número de 18 a 20. Un modelo lineal sólo habría detectado una suave pendiente negativa.

Además, muchos artículos que analizan la educación publicados en revistas prestigiosas²² encuentran coeficientes significativos con el signo contrario al deseado (por ejemplo, más gasto por alumno asociado a peores resultados). Eide et al. (1998) y Figlio (1999) concluyen que las tradicionales especificaciones restrictivas de la función de producción educativa fracasan a la hora de encontrar posibles no linealidades de los recursos escolares. Estos autores emplean una metodología más flexible, como funciones translog o regresiones cuartílicas, para aproximar esta compleja función con buenos resultados sobre modelos lineales y homotéticos. ¿Por qué no explorar la existencia de otras posibles relaciones no lineales?

Volviendo al experimento propuesto, se van a considerar cuatro escenarios diferentes con 50, 100, 200 y 300 unidades de decisión (DMUs). A continuación se generan números pseudoaleatorios, distribuidos uniformemente a lo largo del input, para cada escenario.

$$X \sim U(0, 26)$$

El cálculo del verdadero output permite obtener así la verdadera frontera, (figura 2.4.) para, a continuación, generar ineficiencias inyectando diferentes cantidades de ruido. El ruido es inyectado en el output real de la siguiente forma.

$$y^* \sim U(y+ay; y-by)$$

²² Véase Hanushek (1986) para una revisión de la literatura.

Donde y^* es el output observado distribuido uniformemente entre $y+ay$; $y-by$, sabiendo que, $a = 0.05$ si $b = 0.1, 0.2, 0.3$; y $a=0.15$ si $b = 0.35, 0.6$. Mediremos la eficiencia técnica (te) de la siguiente forma:

$$te \sim (y^*/y) \quad (\text{se permite que } te > 1)$$

Con el fin de simplificar, se asume que los datos no contienen ruido y que toda la diferencia entre el output real y el observado son ineficiencias²³. Sin embargo se permite que $te > 1$ con el objetivo de simular la posible existencia de puntos extremos.

Para cada escenario se ha calculado la eficiencia técnica utilizando mínimos cuadrados ordinarios (OLS), mínimos cuadrados ordinarios corregidos (COLS), análisis de frontera estocástica (SFA), DEA con rendimientos constantes y variables a escala y un MLP.

Antes de entrenar el MLP, se dividieron los datos en dos partes, asignando el 80% de los datos a entrenamiento y el 20% restante a validación. El modelo es desarrollado a partir de los datos de entrenamiento y probado en los datos de validación. Después de un análisis exploratorio, se pudo comprobar como las diferencias en los errores de las submuestras de entrenamiento y validación eran prácticamente idénticas, por lo que se decidió juntar la estimación de la red para ambos conjuntos de datos a la hora de calcular el output. El proceso de entrenamiento consistió en una búsqueda del MLP con mejor ajuste, variando el número de unidades en la capa oculta entre 3 y 8 neuronas. El coeficiente de aprendizaje y el valor del momento fueron fijados en los valores 0,5 y 0,001 respectivamente.

Las estimaciones del MLP que más se ajustaban al output deseado (MLP Best) fueron seleccionadas para su comparación con el resto de técnicas. A continuación se calcularon los coeficientes de correlación de Pearson entre la eficiencia estimada y la eficiencia real para todas las técnicas utilizadas y para cada uno de los escenarios simulados (tabla 2.2)

Tal y como muestran los resultados que muestran la tabla 2.2, el ajuste alcanzado por el MLP supera al del resto de técnicas en todos los casos menos en uno. Y en este último, la diferencia a su favor respecto del ajuste alcanzado por el MLP es muy baja; mientras que en el resto de los casos la diferencia a favor del MLP es bastante alta. Se puede observar además como las estimaciones obtenidas por el MLP son más robustas que para el resto de técnicas

²³ Zhang and Bartels (1998) también suponen datos sin perturbaciones aleatorias. No obstante, obtendríamos conclusiones similares en este experimento si descompusiéramos el término de error en una parte de ruido blanco iid $u \sim N(0, \delta^2)$ y una parte debida a eficiencia iid $v \sim N(0, \delta_v^2)$

cuando varían la cantidad de ruido inyectado o el número de unidades productivas. El MLP supera a las técnicas tradicionales cuando los datos observados tienen una cantidad de ruido moderado y disponemos de suficientes unidades productivas.

Tabla 2.2. Coeficientes de correlación de Pearson entre la eficiencia estimada y la eficiencia real para las distintas técnicas usadas y según número de unidades productivas y ruido inyectado.

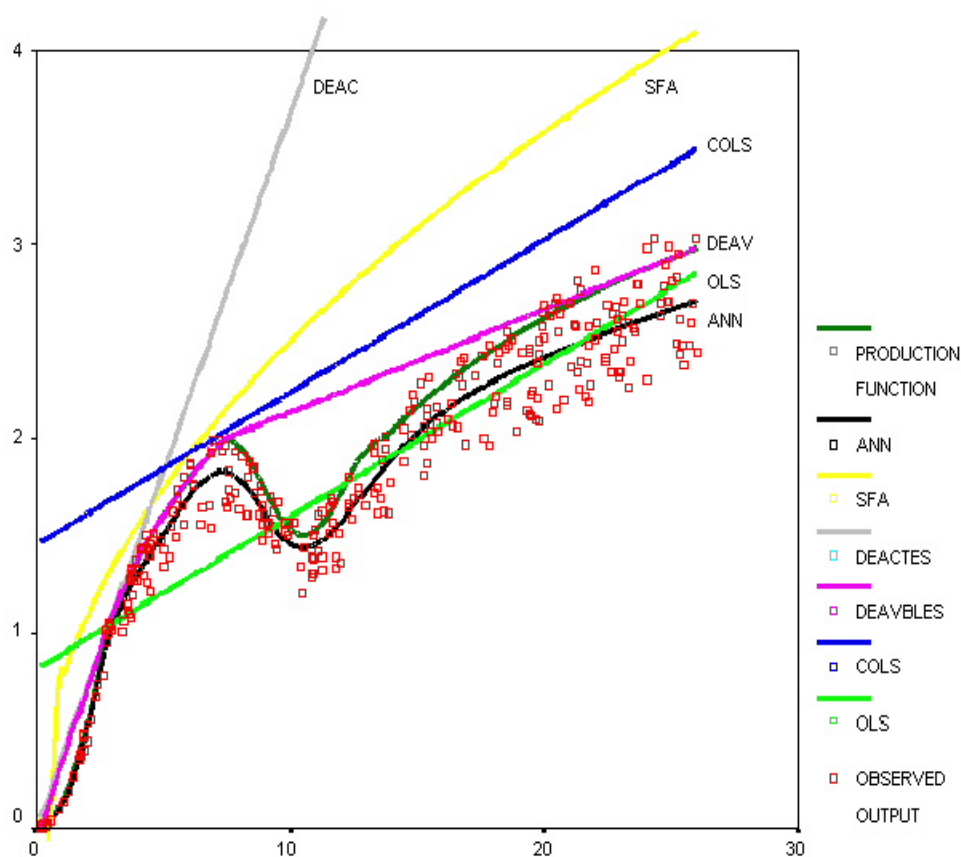
		Técnicas de Eficiencia					
Número de DMUs & Porcentaje de ruido inyectado	50 DMUs	OLS	COLS	SF	DEArce	DEArve	MLP_BEST
	50(15)	0,180	0,104	0,441	0,297	0,431	0,788
	50(25)	0,230	0,249	0,294	0,119	0,296	0,838
	50(35)	0,464	0,405	0,581	0,419	0,714	0,804
	50(50)	0,584	0,575	0,630	0,378	0,798	0,873
	50(75)	0,608	0,520	0,443	0,473	0,895	0,887
	100 DMUs	OLS	COLS	SF	DEArce	DEArve	MLP_BEST
	100(15)	0,145	0,146	0,096	0,090	0,183	0,897
	100(25)	0,255	0,211	0,239	0,286	0,293	0,751
	100(35)	0,297	0,237	0,332	0,357	0,498	0,919
	100(50)	0,496	0,490	0,321	0,345	0,661	0,951
	100(75)	0,557	0,517	0,474	0,543	0,728	0,855
	200 DMUs	OLS	COLS	SF	DEArce	DEArve	MLP_BEST
	200(15)	0,184	0,205	0,139	0,076	0,249	0,816
	200(25)	0,326	0,322	0,258	0,187	0,439	0,961
	200(35)	0,377	0,329	0,280	0,348	0,479	0,947
	200(50)	0,554	0,557	0,331	0,365	0,686	0,924
	200(75)	0,685	0,705	0,337	0,483	0,794	0,934
	300 DMUs	OLS	COLS	SF	DEArce	DEArve	MLP_BEST
	300(15)	0,214	0,248	0,029	0,026	0,302	0,887
300(25)	0,374	0,332	0,388	0,280	0,457	0,935	
300(35)	0,447	0,409	0,417	0,316	0,587	0,975	
300(50)	0,606	0,607	0,663	0,319	0,736	0,935	
300(75)	0,759	0,722	0,804	0,541	0,857	0,973	

En la figura 2.5 se presenta un caso de ejemplo para cuando se tienen 300 unidades de decisión y hemos inyectado un 25% de ruido uniformemente distribuido en el output. Después de trazar la verdadera función y todas las estimaciones proyectadas por todas las técnicas, se observa como el MLP es la única técnica capaz de encontrar la no linealidad contenida en los datos. Vemos además como el MLP da una estimación media del resultado. En este punto podríamos convertir la estimación en una frontera a través de un desplazamiento de la función en la cuantía del residuo positivo de mayor valor, tal y como hacemos en COLS.

A través de la figura 2.5 se observa igualmente, tal y como señala Lee et al. (1993), como el MLP es una buena herramienta para realizar una búsqueda de posibles no linealidades en nuestros datos antes de aplicar una técnica convencional, evitando así posibles errores de especificación. Debemos tener en cuenta que la posibilidad de existencia de no linealidades se

incrementa exponencialmente a medida que aumenta el número de inputs, outputs y variables ambientales implicadas en el proceso productivo.

Figura 2.5: Funciones de producción estimadas por distintas técnicas para medir la eficiencia técnica.



2.4.2. La Medición de la Eficiencia Técnica mediante RNAs a través de un Experimento de Monte Carlo: Ejemplo 2.

En economía normalmente asumimos una serie de supuestos en torno a la función de producción de cualquier proceso transformador de inputs en outputs²⁴, tales como la imposibilidad de producir alguna cantidad de output sin factores productivos, libre disposición de inputs, posibilidad de rendimientos a escala crecientes, decrecientes y constantes, aditividad, continuidad o doble diferenciabilidad de la función de producción.

Sin embargo, la función de producción del ejemplo 1 violaba el supuesto de ser monótona creciente, esto es, que cualquier aumento de cualquier factor productivo no se

²⁴ Para un estudio detallado de todas las condiciones puede acudir a Mas-Colell et al. (1995).

traduzca en una disminución de alguno de los outputs. Con ello se pretendía demostrar como la flexibilidad de un MLP es capaz de ajustar cualquier tipo de no linealidad que presentaran los datos, dado que no realizábamos ningún supuesto en torno a la forma que ésta puede adoptar.

A continuación se va a estudiar otra función de producción que sí cumple todos los supuestos microeconómicos tradicionales con el fin de comprobar el comportamiento de las RNAs a la hora de medir la eficiencia técnica. Para ello se lleva a cabo un experimento de Monte Carlo²⁵. Esta técnica de cálculo numérico, permite analizar múltiples escenarios de un modelo utilizando diferentes muestras de datos, que se generan a partir de una distribución de probabilidad previamente definida.

Típicamente, el experimento de Monte Carlo en los trabajos de eficiencia se lleva a cabo a partir del siguiente diseño básico²⁶:

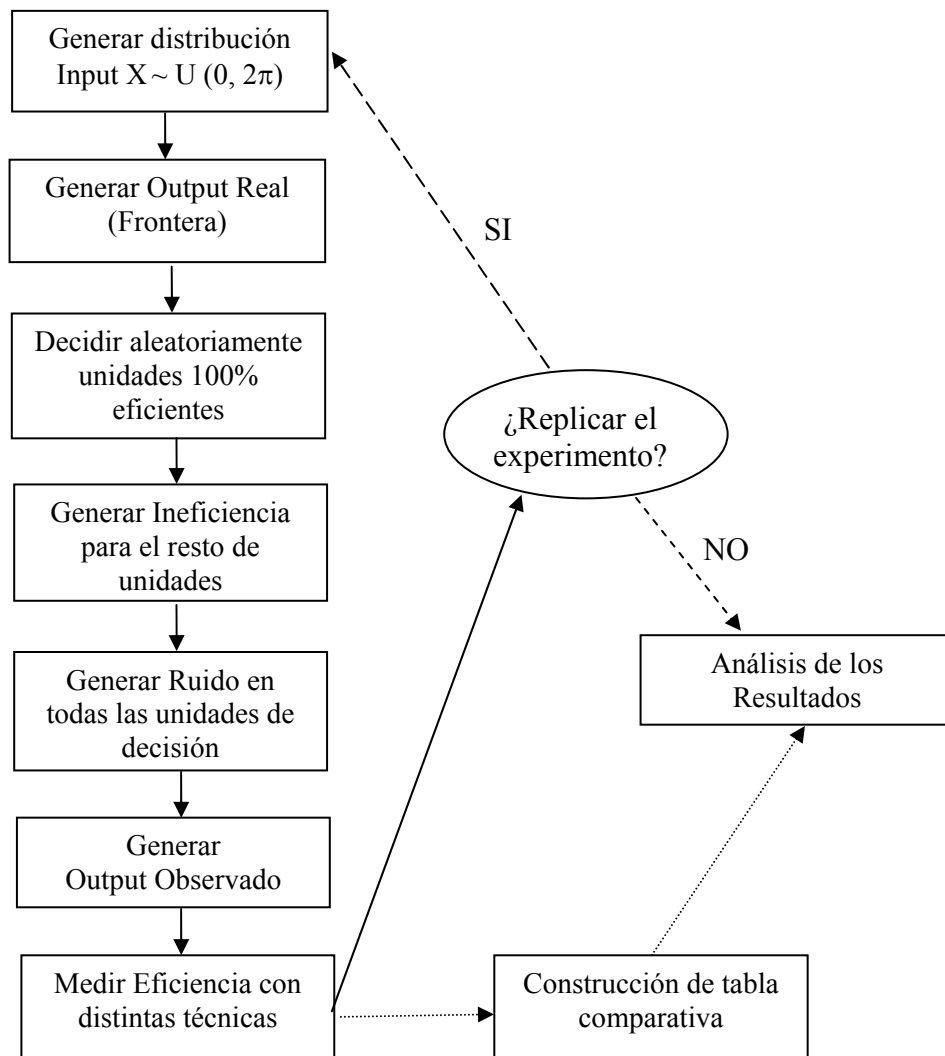
1. Asumir una distribución para el vector input y proceder a su generación.
2. Generar el output a partir de la función de producción definida. En este punto del análisis todas las unidades decisoras (DMUs) estarán sobre la frontera.
3. Se generan los valores de ineficiencia para todas las firmas. Para ello, se asume una distribución de ésta que suele ser la semi-normal o la exponencial.
4. A continuación se genera el output observado para cada DMU a partir del nivel del grado de ineficiencia y el vector de inputs.
5. Utilizamos una o distintas metodologías para comparar resultados o tratar de resolver el problema que hayamos planteado

La figura 2.6 muestra el esquema del experimento de Monte Carlo que llevaremos a cabo en este ejemplo.

²⁵ Los métodos de Monte Carlo están asociados a la generación de números aleatorios y su nombre se inspira en los juegos de azar de los casinos que existen en el Principado de Mónaco.

²⁶ Algunos de los numerosos trabajos que utilizan simulaciones de Monte Carlo en el ámbito de la medición de la eficiencia serían Coelli (1995), Zhang y Bartels (1998) o Bardhan et al. (1998)

Figura 2.6: Esquema del Experimento de Monte Carlo para la comparación de técnicas de medición de la eficiencia.



Así, sea $G(x)$ la siguiente función de producción continua, monótona creciente, doblemente derivable y diferenciable con un único input y un único output:

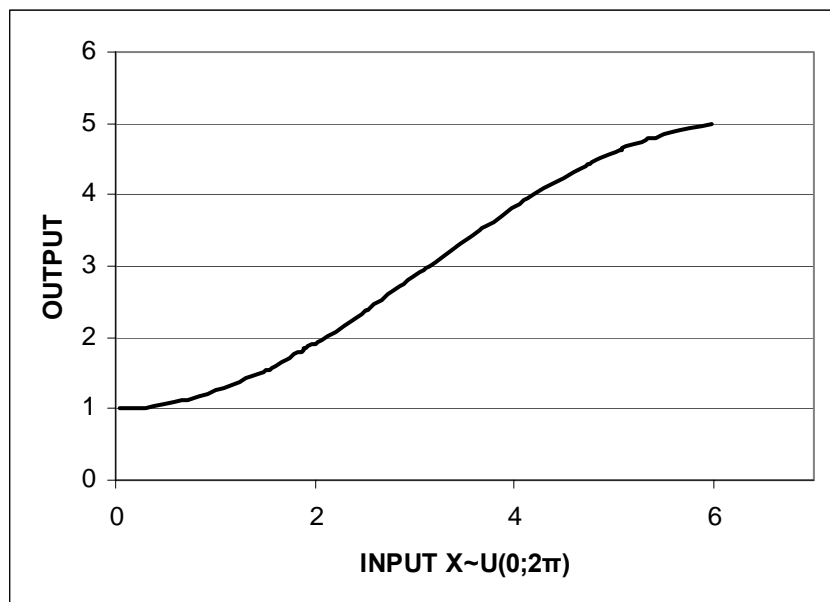
$$G(x) = 2 * \left[\text{seno} \left(\left(\frac{x}{2} \right) - \left(\frac{\pi}{2} \right) \right) + 1 \right] + 1$$

Donde el input se distribuye,

$$X \sim U(0, 2\pi)$$

Tal y como se puede observar a través de la figura 2.6 la función definida cumple todos los supuestos económicos tradicionales definidos en teoría microeconómica.

Figura 2.6: La función de producción seno.



A partir de esta distribución, llevaremos a cabo 100 replicaciones del experimento, todas ellas con 50 unidades productivas (DMUs). Este valor es sustituido en la función de producción para generar el output real que estará situado sobre la frontera. A continuación, se genera la ineficiencia que hemos supuesto se distribuye $v \sim |N(0,75; 0,0225)|$, teniendo en cuenta que en todos los casos hemos seleccionado aleatoriamente un número de DMUs como técnicamente eficientes a partir de una distribución Bernoulli (0,2). Estas unidades permanecen así sobre la frontera y serían 100% eficientes. Seguidamente se ha introducido ruido blanco en todas las DMUs $\varepsilon \sim N(0; 0,01)$ y calculamos el output observado.

Una vez que se dispone del input y del output observado, se calcula la eficiencia técnica real, antes de introducir el ruido, así como la eficiencia técnica observada, que es la que se hubiera obtenido si conociéramos la verdadera función de producción pero ignorando la distorsión que el ruido introduce en los valores reales de cada DMU.

A continuación, se comparan estos valores con los obtenidos por OLS, COLS (desplazando la frontera hasta el residuo positivo más alejado de la recta de regresión), SFA (suponiendo una función Cobb-Douglas y estimando mediante máxima verosimilitud), DEA asumiendo rendimientos constantes a escala, DEA asumiendo rendimientos variables a escala y

MLP. Para este fin, se calculan en cada replicación las medias y desviaciones típicas de los coeficientes de correlación medio entre los valores de eficiencia reales, observado y los estimados por cada técnica para cada DMU.

Tras un análisis exploratorio, se entrena en los 100 casos un MLP con 5 neuronas en la capa oculta y con coeficiente de aprendizaje y vector momento ambos con valores 0,1. La función de transferencia usada en las neuronas de la capa oculta y de salida es la sigmoide logística, mientras que en las capas de entrada se han usado pesos lineales en el intervalo [-1, 1]. Los pesos entre las conexiones son inicializados pseudo-aleatoriamente en el intervalo $\pm 0,3$. Los resultados muestran, al igual que en el experimento anterior, que el ajuste en las muestras de entrenamiento y validación es prácticamente idéntico. Por ello, aunque separamos la muestra en dos submuestras (de entrenamiento y de validación) con el objetivo de evitar el sobreentrenamiento, se evalúan los resultados obtenidos sobre el total de los datos.

Los resultados finales del experimento de Monte Carlo son los que muestra la tabla 2.3.

Tabla 2.3. Correlaciones medias con los valores de eficiencia reales obtenidos por los valores observados y estimados por las distintas técnicas en el experimento de Monte Carlo.

ESTIMACIONES	CORRELACIONES
Val. observados	0,8428 (0,0276)
MLP	0,7725 (0,0668)
DEArve	0,7270 (0,0506)
SFA	0,7231 (0,0545)
OLS	0,7011 (0,1240)
COLS	0,5821 (0,1146)
DEArce	0,2284 (0,1674)

La cifra entre paréntesis es la desviación típica.

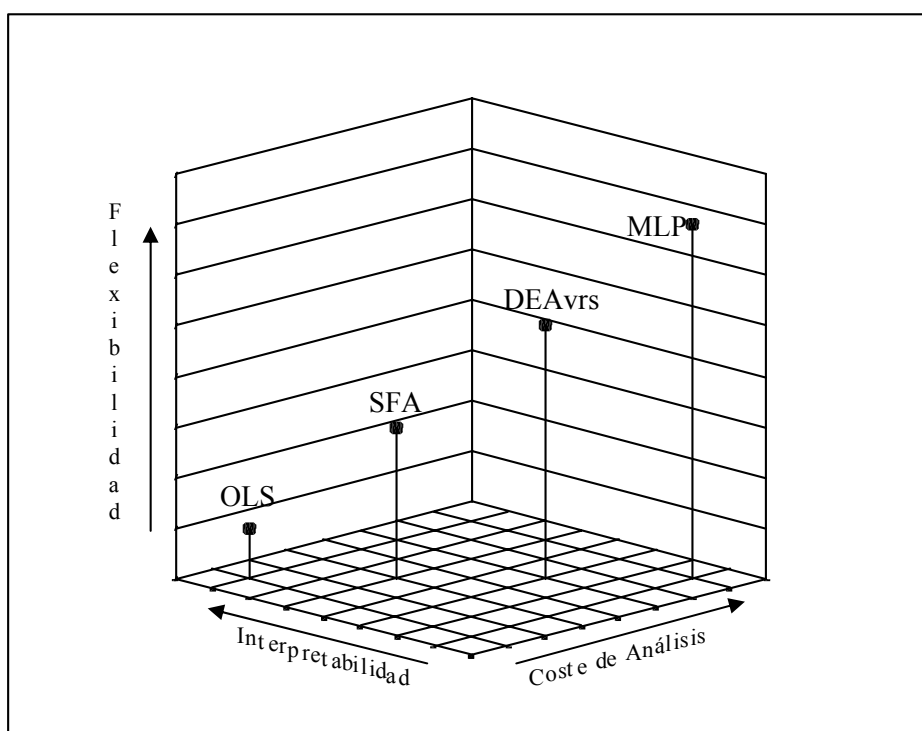
Como se observa en la tabla 2.3, la correlación más alta con los valores de eficiencia reales es la obtenida por los valores outputs observados una vez introducido el ruido estadístico. El MLP es la técnica de medición de eficiencia que obtiene una correlación más elevada con los valores de eficiencia reales. A continuación, irían las técnicas de DEA con rendimientos variables a escala y de frontera estocástica seguidas muy de cerca por mínimos cuadrados ordinarios. Las técnicas de COLS y DEA con rendimientos constantes a escala obtendrían, por último, los peores resultados.

Los resultados obtenidos en ambos experimentos son consistentes con los obtenidos por otros estudios de simulaciones como el de Banker et al. (1988). Éstos señalan que el análisis DEA con rendimientos variables a escala es, generalmente, una herramienta más flexible al aproximar una función de producción que cualquiera de las formas paramétricas más flexibles.

2.5. Recapitulación

Los resultados obtenidos por estas simulaciones muestran como las RNAs son una alternativa a las técnicas tradicionales para la medición de la eficiencia técnica. Otra de las conclusiones que podemos obtener es que ninguna de las herramientas disponibles es superior al resto de metodologías en todas las cualidades. La técnica de medición debería elegirse en función del tipo de problema que se pretenda analizar. El gráfico 2.7 ilustra el *trade-off* que existe entre las cualidades deseables de las técnicas de medición de la eficiencia.

Gráfico 2.7. Trade-off existente entre las cualidades deseables de las técnicas de medición de la eficiencia técnica.



Fuente: Elaboración propia.

Los beneficios de las RNAs son su elevada flexibilidad a la hora de descubrir posibles no linealidades en el proceso productivo, que difícilmente serían detectadas por otros métodos.

Sin embargo, al ser una herramienta no paramétrica la función obtenida es difícil de interpretar. Además el entrenamiento de un MLP es un proceso más costoso en términos de tiempo que el resto de herramientas. En el extremo opuesto tendríamos la regresión lineal cuyos parámetros son fáciles de interpretar pero a cambio se sacrifica la posible detección de no linealidades en el modelo. En un término intermedio se encuentran técnicas más flexibles, como los modelos de SFA y DEA con rendimientos variables a escala, donde sacrificamos más o menos flexibilidad a cambio de más o menos interpretabilidad.

A pesar de que la técnica de MLP ha sido utilizada, como se ha visto anteriormente, en distintas ocasiones para la medición de la eficiencia, sin duda parece razonable llevar a cabo más aplicaciones empíricas en distintos ámbitos así como más estudios comparativos de técnicas. El uso simultáneo de distintas técnicas es otra de las alternativas disponibles que deberá ser evaluada.

**CAPÍTULO 3: LA ESTIMACIÓN DE LA FUNCIÓN DE
PRODUCCIÓN EDUCATIVA: LOS MODELOS EN VALOR
AÑADIDO**

3.1. Introducción

En este capítulo se propone, a partir de un análisis crítico de los modelos de función de producción educativa actuales, un modelo para la estimación de la importancia de alumno y escuela en la explicación del rendimiento académico. Para ello se calcula el output escolar en valor añadido. Este concepto, clave a lo largo de este trabajo, puede ser definido como el porcentaje del rendimiento total del alumno que la escuela aporta independientemente de las características individuales del mismo. A su vez, el valor añadido escolar estaría compuesto por las características cuantificables de la escuela, la eficiencia escolar y el ruido estadístico.

La educación es uno de los bienes que el Sector Público provee y produce gratuitamente por considerarla esencial para el desarrollo de una vida digna. Es por ello por lo que los estados de las economías modernas obligan a sus ciudadanos a consumir un mínimo de este bien. En España, este derecho básico queda recogido a través del artículo 27.4 de la Constitución Española: “la enseñanza básica es obligatoria y gratuita”, artículo que podemos identificar con el concepto económico de equidad categórica. Por otro lado, el artículo 31.2 apunta que, “el gasto público realizará una asignación equitativa de los recursos públicos y su programación y ejecución responderán a los criterios de eficiencia y economía”. Quedan reconocidos, por tanto, constitucionalmente los objetivos de eficacia, eficiencia y equidad en la educación.

Desde la elaboración de la Teoría del Capital Humano por Gary Becker (1964), objetivos como la reducción del fracaso escolar, el aumento generalizado del rendimiento académico, la mejora en la calidad de la enseñanza y la igualdad de oportunidades educativas para todos los alumnos, engloban, a grandes rasgos, el deseo general de todos los colectivos que componen el mundo educativo. Pero en una economía de recursos escasos, todo esfuerzo de mejora en ésta o en cualquier otra actividad se hace, obviamente, a costa de reducir las inversiones en usos alternativos. De ahí la importancia de actuar con la máxima eficiencia en toda actividad económica, o lo que es lo mismo, produciendo el máximo posible, en nuestro caso educación, con las dotaciones de trabajo y capital disponibles.

Durante la década de los 80, los objetivos prioritarios en el campo de la educación no universitaria española fueron conseguir la plena escolarización y la gratuidad de la enseñanza. Sin embargo, tal y como señala San Segundo (1991), “una vez que se consideran cumplidos los objetivos básicos de escolarización obligatoria y gratuita, todos los sistemas educativos se preocupan por ofrecer mejoras cualitativas, aunque no siempre haya acuerdo sobre cómo medir la calidad de la enseñanza”.

Por sus especiales características, el análisis de la eficiencia y la valoración de los resultados del sector público en educación presenta mayores dificultades que para el ámbito privado. No obstante, los objetivos macroeconómicos de reducción del déficit y sobre todo de crecimiento económico, [Lee y Barro, 1998], han llevado a los economistas a interesarse detalladamente por la gestión del gasto público educativo, desarrollándose así, sobre todo desde las tres últimas décadas, el campo de la economía de la educación.

En el caso específico de la educación no universitaria, el principal objetivo a lo largo de la literatura ha sido la búsqueda de una función de producción educativa que relacionara los factores productivos, el contexto y los resultados educativos, a fin de obtener una herramienta útil para la toma de decisiones. En el siguiente apartado se verá cómo se ha desarrollado hasta la actualidad la búsqueda de esta función de producción que ha servido, entre otras cosas, para la medición de la eficiencia técnica de las escuelas.

3.2. La Función de Producción Educativa

Se pueden situar los orígenes de los estudios de función de producción educativa en el trabajo de Carroll (1963). John Carroll propuso un modelo para explicar el aprendizaje en las escuelas. Su hipótesis de partida fue definir el aprendizaje como una función del tiempo. Para ser más específico Carroll propuso:

$$\text{Rendimiento} = f(\text{tiempo utilizado} / \text{tiempo necesitado})$$

Donde:

$$\text{Tiempo utilizado} = \text{oportunidad} \times \text{perseverancia}$$

$$\text{Tiempo necesitado} = h(\text{aptitud, conocimientos previos, calidad de la clase})$$

Rendimiento: Resultado en un test homogéneo sobre conocimientos básicos.

Tiempo Utilizado: Número de minutos por día escolar en los que los alumnos están realmente trabajando. Está construido a partir de dos variables.

- *Oportunidad*: Tiempo que realmente es usado en la clase para tareas de aprendizaje.
- *Perseverancia*: Tiempo que realmente es usado en clase por el alumno.

Tiempo Necesitado: Es el tiempo que un alumno necesita para alcanzar un resultado en función de sus circunstancias.

- *Aptitud*: Capacidad de aprender el temario escolar. Se mediría con el coeficiente de inteligencia.
- *Conocimientos previos*: Conocimientos de partida del alumno.
- *Calidad de la clase*: Adecuación de las técnicas de enseñanza.

La principal debilidad en el modelo de Carroll es que se omite de forma explícita las variables de contexto educativo. Para Carroll las variables escolares están directamente relacionadas con el aprendizaje mientras que las variables familiares y de la comunidad estarían indirectamente relacionadas con el rendimiento. Vemos por tanto como Carroll detecta la presencia de dos tipos de variables que influyen en el resultado del alumno que se podrían dividir, a grandes rasgos, en variables escolares y no escolares o de contexto.

Desde mediados de los sesenta y a lo largo de los setenta, numerosos autores comenzaron a investigar los factores que estaban relacionados con los resultados escolares basados en su mayoría en el esquema de función de producción desarrollado por Carroll. Dos de los estudios que más atención han recibido posteriormente a lo largo de la literatura fueron los realizados por Coleman (1966) y Jenks (1972).

James Coleman y su equipo publican en 1966 “Igualdad de Oportunidades Educativas”, conocido posteriormente como “El Informe Coleman”. El objetivo inicial de este trabajo, era el estudio de la distribución de los recursos educativos en los Estados Unidos para comprobar²⁷ si éstos se aplicaban por igual sin distinción de razas o minorías. Sin embargo, el estudio fue más allá al recopilar una extensa base de datos acerca de 600.000 estudiantes, 60.000 profesores y de los recursos escolares de los 4.000 colegios a los que asistían.

El trabajo concluía que las tasas de escolarización, la educación y el salario del profesorado, así como el currículo, era relativamente igual entre los niños de raza blanca y los de otras razas. También se detectó entre otros resultados que los estudiantes de raza negra estaban segregados en escuelas separadas de los alumnos de raza blanca. Sin embargo, el hallazgo clave del “Informe Coleman” por el que adquirió relevancia posterior en la investigación económica de la educación aparentaba demostrar que, teniendo en cuenta los antecedentes familiares y personales de los alumnos, su rendimiento escolar apenas guardaba relación con los recursos escolares.

²⁷ Este estudio fue financiado para demostrar que la calidad y la cantidad de los recursos de la escuela eran de forma significativa los determinantes del éxito o del fracaso escolar, principalmente de las clases minoritarias en los Estados Unidos, y no los aspectos familiares. Sin embargo, como en tantas ocasiones, el resultado obtenido fue justamente el contrario a la hipótesis de partida.

La crítica más extendida al “Informe Coleman”, que concluyó que las características socioeconómicas del alumno y el entorno en el que se ubica la escuela explican la mayor parte del rendimiento académico, se refiere a la metodología en la que se apoyan sus conclusiones, (Bowles y Levin 1968a y b). Esta se basaba en regresiones aplicadas a una muestra extensa de centros en las que las variables independientes, (servicios escolares, influencia familiar y composición escolar), eran introducidas sucesivamente analizándose su influencia sobre el resultado escolar en función de su contribución al incremento del coeficiente de determinación.

El problema es que este método sólo es válido si las variables exógenas son independientes, ya que si existe multicolinealidad, los resultados obtenidos dependerán del orden de entrada de las variables. En el “Informe Coleman” las primeras variables introducidas fueron los antecedentes familiares del alumno y éstos pueden estar correlacionados con el gasto del distrito escolar. Así, los escasos efectos atribuidos a la escuela pueden ser consecuencia del sesgo implícito en el procedimiento, de forma que parte de los efectos familiares recogerían también la influencia de los recursos escolares.

A pesar de las críticas metodológicas, el Informe Coleman tuvo una enorme difusión y dejó abierto un debate, que actualmente todavía no está cerrado, que podríamos plantear con la siguiente pregunta: ¿Importan las escuelas?²⁸

Siguiendo el orden cronológico, el influyente artículo de Summers y Wolfe (1977) vino a contestar de forma positiva la cuestión planteada anteriormente. Destacamos este trabajo sobre otros porque introdujo además nuevas consideraciones teóricas en el debate. Anita Summers y Bárbara Wolfe definieron la siguiente función de producción educativa:

$$\Delta A = F(GSES, TQ, SQ, PG)$$

Donde ΔA es el incremento en el rendimiento de un alumno que se ha producido entre dos cursos diferentes, GSES trata de medir la dotación genética y el contexto socioeconómico del alumno, TQ mide la calidad del profesorado, SQ mediría otros recursos escolares distintos del profesor y PG recoge el efecto compañeros.

²⁸ Esta es una traducción literal de las frases en inglés que más abundan en los artículos científicos que tratan sobre producción educativa. Does School Matter?, Does Money Matter?, Does School Quality Matter?. Una traducción menos literal, aunque con un significado más duro, sería ¿sirven realmente para algo las escuelas?.

Estas dos autoras argumentan las siguientes ideas en torno a la función de producción educativa y la importancia de la escuela.

- Esta función de producción no es como la de una fábrica ya que hay factores productivos que el gestor de los recursos no controla (GSES, y parte del PG) y sin embargo otros sobre los que si posee capacidad de gestión.
- En este artículo se critica el uso de los datos agregados en vez de usarlos a nivel individual. Igualmente se critica el hecho de usar un indicador de resultados y no el valor añadido, entendido como la diferencia en rendimiento entre dos cursos y por tanto descontado el efecto de la familia, que el centro es capaz de aportar al alumno. Se concluye que los datos desagregados son capaces de arrojar mayor luz acerca de los factores productivos que resultan significativos.
- El objetivo de muchas escuelas no termina con el aumento del rendimiento académico; otros objetivos tales como la actitud frente a individuos de otras razas y religiones, sentido de participación y de reconocimiento de los valores democráticos son igualmente parte de los logros educativos.
- Se encuentra que muchos inputs escolares importan y que su impacto varía considerablemente con el tipo de estudiante. Determinados inputs ayudan más a los alumnos más desaventajados mientras que otros inputs parecen favorecer en mayor medida a los alumnos aventajados. Detrás de este resultado se encuentra la creencia de padres y profesores de que lo mejor es una educación mucho más individualizada.
- Una posible explicación a que muchos estudios encuentren que determinados inputs no son significativos se encuentra en que no se observa si estos inputs son relevantes para determinados alumnos como consecuencia de la agregación de los datos.

Tras este trabajo de Summers y Wolfe, que valoraba positivamente la escuela, se han ido publicando cientos de investigaciones sobre este tema a lo largo de las dos últimas décadas con resultados contradictorios a la hora de señalar si la escuela importa o no en la producción educativa. Se resaltan, a continuación, las dos líneas principales en las que el debate se encuentra actualmente.

Por un lado, el influyente trabajo de Eric Hanushek²⁹ defiende en sus numerosas publicaciones la tesis de que más dinero no implica por sí solo mejores resultados. Hanushek demuestra que, a lo largo de la literatura en economía de la educación, no existe una relación significativa clara que ratifique de forma estadística que más factores productivos escolares,

²⁹ Para una revisión de su influyente literatura en el campo de la economía de la educación podemos destacar: Hanushek [1971, 1979, 1981, 1986, 1997].

tales como el menor número de alumnos en las clases, el mayor sueldo del profesorado o la incorporación de más servicios escolares, supongan un aumento, *per se*, en los resultados de los alumnos.

El mismo Hanushek apunta que esta tesis no significa que aumentar los recursos escolares no sea positivo para la educación, ni que existan multitud de ejemplos de escuelas donde más recursos suponen mejoras educativas. Este autor defiende básicamente que los mayores recursos deben de ir acompañados de cambios institucionales como la competencia entre escuelas, la medición sistematizada de los resultados y los recursos o la introducción de incentivos al profesorado.

La segunda línea de pensamiento sería la defendida originalmente por Summers y Wolfe (1977) que argumenta que la escuela sí importa. Pero en buena medida desde 1986 ha existido una reacción al trabajo de Hanushek por autores como Card y Krueger (1992), Figlio (1999), Dewey (2000) y Krueger (1999, 2000). Este último autor incluso analizó de nuevo los resultados que obtuvo Hanushek a lo largo de la revisión de la literatura concluyendo que si se tienen en cuenta determinadas características de esos trabajos³⁰, los resultados apuntan que sí existe una relación positiva entre los recursos dedicados a la educación y los resultados de los estudiantes. El debate sigue abierto.

3.3. Problemas en la Estimación de la Función de Producción Educativa

Un alto porcentaje de la controversia y el debate que existe en torno a la importancia y el papel de la escuela en la producción educativa es debido a las peculiaridades de este proceso productivo, muy diferente de las funciones de producción tradicionales en economía. Estas peculiaridades condicionan la estimación de esta función de producción.

Así, dentro del ámbito de la economía de la educación, el estudio de la eficiencia técnica de los centros escolares no universitarios ha sido abordado fundamentalmente a través de la búsqueda de una *función de producción educativa* teórica. Esta herramienta nos permitiría así relacionar los factores productivos educativos y los outputs escolares, medidos estos últimos generalmente, a partir de algún indicador del rendimiento académico como pruebas objetivas de matemáticas, ciencia o lenguaje.

³⁰ En particular Krueger pondera los trabajos en función del impacto de la revista en la que fue publicado.

3.3.1 Características del Proceso Educativo.

Una de las funciones de este tipo más aceptada y buscada empíricamente fue la propuesta ya en los orígenes de esta disciplina por autores como Levin (1974) o Hanushek (1979) que recoge la ecuación 3.1.

$$A_{it} = f(B_i^{(t)}, P_i^{(t)}, S_i^{(t)}, I_i) \quad (\text{ec.3.1})$$

A_{it} = Rendimiento del alumno i en el instante t

$B_i^{(t)}$ = Vector de las características familiares del alumno i acumuladas hasta el instante t

$P_i^{(t)}$ = Vector de las influencias de los compañeros en el alumno i acumuladas hasta el instante t

$S_i^{(t)}$ = Vector de los inputs escolares acumulados hasta el instante t

I_i = Vector de la capacidad innata del alumno i

Este esquema para modelar el proceso productivo educativo ha sido empleado en numerosos estudios como en Burkhead et al. (1967), Bowls y Levin (1968a), Debertin (1976), o Sengupta y Sfeir (1986) entre otros.

Mediante esta función, se asume que existe una variable dependiente A_{it} , medida a nivel individual, sobre la que influyen dos tipos de variables explicativas correspondientes a dos niveles distintos de actuación. El primer tipo de variables son aquellas, medidas a nivel de alumno, sobre las que el centro posee escasa o nula influencia y que condicionan el rendimiento académico del alumno, $(B_i^{(t)}, I_i)$. Desde el Informe Coleman existe un amplio acuerdo en la literatura respecto a la importancia de estas variables en el resultado educativo.

El segundo conjunto de variables $(P_i^{(t)}, S_i^{(t)})$ estarían medidas a nivel de escuela y podríamos clasificarlas en dos grupos. Por un lado están aquellas variables sobre las que el centro o el organismo encargado de la asignación de los recursos puede intervenir mediante el aumento o la reducción física de recursos. Por otro lado, están aquellas variables contextuales sobre las cuales el centro educativo no puede intervenir, al menos en el corto plazo, tales como las dimensiones del centro, las características del profesorado disponible o las características socioeconómicas del grupo de alumnos matriculados en el centro.

Este modelo de producción educativa podría ser fácilmente ampliado a más niveles, tomando por ejemplo el entorno sociodemográfico y socioeconómico de la región donde se encuentran ubicados distintos grupos de colegios. A pesar de que el proceso educativo se

desarrolla a través de varios niveles, esto es, se encuentra estructurado de forma jerárquica, la forma más habitual de estimar esta función para posteriormente medir la eficiencia técnica en los estudios económicos ha sido con los datos agregados a nivel escolar³¹. En este contexto, la medición de la eficiencia en el ámbito educativo plantea una serie de dificultades.

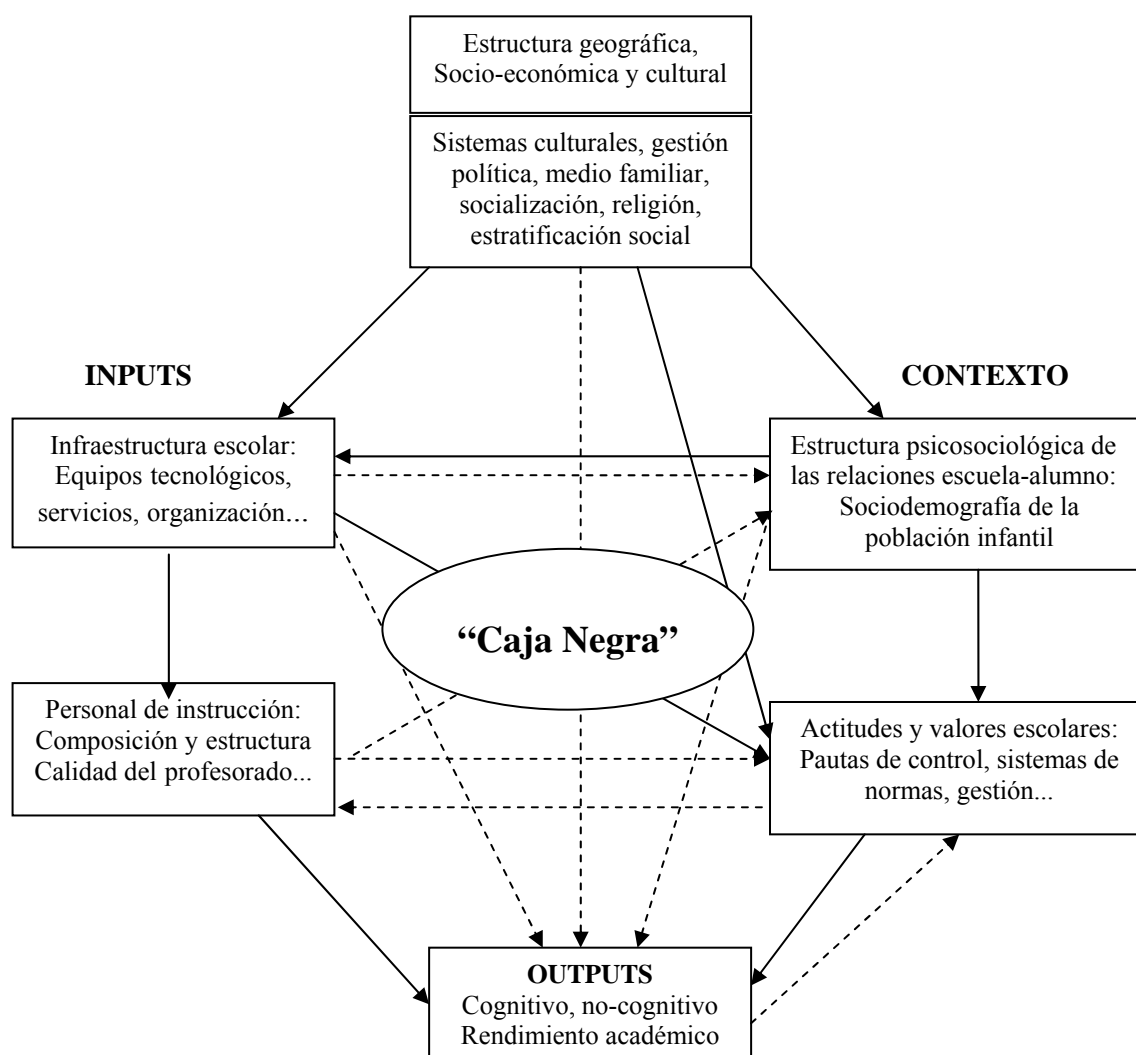
En primer lugar, la forma funcional de (3.1) sigue siendo, desde las primeras investigaciones sobre producción educativa, completamente desconocida, lo cual limita la toma de decisiones a la hora de asignar los recursos productivos.

Este desconocimiento es debido, tal y como señalan Tejedor et al. (1988), a la multitud de factores productivos implicados en el proceso educativo, formados a grandes rasgos por el entorno socioeconómico y regional del alumno, dotaciones escolares y gestión del centro, calidad del profesorado y métodos pedagógicos, nivel aptitudinal y clima en el aula, etc. Estos factores contextuales no son siempre fácilmente cuantificables y además parecen estar fuertemente interrelacionados, aunque no siempre siguiendo una misma lógica, (gráfico 3.1).

Como se observa en el gráfico 3.1 además de las múltiples interrelaciones que se producen en la educación no universitaria, éstas a su vez son de “caja negra”. Es decir, aunque se pueden intuir y conocer las variables que influyen en el resultado de los estudiantes, existe un gran desconocimiento en torno a los procesos o tecnología que transforma los recursos en resultados. Así, un aumento de los mismos recursos en dos escuelas se podrían traducir en mejores y peores resultados para cada una de ellas respecto a los resultados de partida. ¿Por qué determinados países con mayor inversión en educación obtienen peores resultados académicos que otros con menor inversión? La respuesta a esta pregunta no es fácil y debe ser buscada en variables difícilmente cuantificables tales como los métodos pedagógicos, factores culturales hacia la educación y otros factores contextuales.

³¹ A pesar de que en economía abundan en mayor medida los estudios a nivel escolar, existen trabajos como el de Jenks et al. (1972), Madaus et al. (1980) o Castejón (1994) en los que la unidad de análisis a la hora de estimar la función de producción educativa es el alumno. Estos estudios que resuelven el problema de la “unidad de análisis” con la desagregación, pueden caer en la llamada *falacia individualista o atomística*, Riley (1963) y Grunfeld et al. (1960), es decir, en extraer conclusiones para los niveles agregados a partir exclusivamente de las características de los individuos.

Gráfico 3.1: Estructura relacional de las variables contextuales en el proceso educativo.



Fuente: Tejedor et al. (1988) y elaboración propia.

En esta misma dirección apunta el trabajo de Vandenberghe (1999). Este autor señala cómo las diferencias existentes en la calidad de las escuelas no sólo se corresponden con diferencias de gasto por alumno, tamaño de la clase u otras variables tradicionales, sino en variables como la capacidad del profesor o el clima escolar. Sin embargo, el esquema tradicional de la función de producción educativa difícilmente recoge estos factores que se asume forman parte de la caja negra en la que la escuela queda convertida. Además, las relaciones tecnológicas entre inputs y outputs estarían condicionadas por la presencia de los factores relacionados con la organización o también llamada eficiencia-x (Leibenstein, 1966).

Estas características contextuales no pueden ser compradas en el mercado como los profesores con mayor formación, mejores laboratorios, ordenadores o libros. El control de la organización escolar no parece por tanto una cuestión trivial. Levin (1994) argumenta que las condiciones necesarias para una buena organización de la escuela se corresponderían básicamente con las de las empresas eficientes y, por tanto, uno de los mecanismos más poderosos para la mejora de la eficiencia de los colegios sería la competencia entre escuelas.

Para Hanushek (1986) las escuelas más productivas son aquellas capaces de combinar satisfactoriamente los siguientes cinco elementos:

1. Tener unos objetivos a alcanzar claramente definidos y aceptados ampliamente por todos los miembros implicados en el proceso.
2. Tanto directivos como profesores deben de tener un esquema de incentivos orientado al éxito de sus estudiantes. Los incentivos pueden ser tanto intrínsecos: tales como la vocación, sentido del deber o satisfacción por el trabajo bien hecho; como extrínsecos: en forma de complementos salariales, premios, avance en una jerarquía o carrera profesional y reconocimiento del resto de profesores. Los esquemas de incentivos podrían introducirse de forma individual o colectiva pero su objetivo final debe ser maximizar el esfuerzo para lo que es preciso la recogida sistemática de información.
3. La información acerca de los métodos pedagógicos exitosos y su evolución debe estar disponible permanentemente.
4. La escuela y la educación se desarrollan en un entorno cambiante y su estructura y forma de funcionar debe evolucionar en función de las demandas tanto individuales como sociales.
5. Las escuelas deben elegir la tecnología más eficiente según sus restricciones presupuestarias.

Por otra parte, el verdadero producto de la educación recibida por un individuo es muy difícil de medir empíricamente debido a su intangibilidad inherente. La educación no se compone únicamente de la capacidad de repetir información y resolver cuestiones sino que comprende además la capacidad de razonar, el comportamiento social y el pensamiento crítico, todos ellos de muy difícil cuantificación. A pesar de la naturaleza multi-producto de la educación, la mayoría de los trabajos han utilizado los resultados de rendimiento en pruebas objetivas por distintas razones. Entre ellas quizás la más importante sea que tanto políticos como padres de alumnos usan este criterio para evaluar el producto educativo, por encima de

otros outputs, y utilizan la información que se deriva para elegir incluso lugar de residencia, escuela y tomar decisiones de inversión [Hoxby, 1999].

Además de los aspectos referentes a la organización y la medición de los inputs y los outputs escolares existen otras variables que influyen significativamente en el éxito o fracaso del comportamiento de la escuela y complican la estimación de la función de producción educativa. Sin duda uno de los factores más nombrados a lo largo de la literatura es el llamado *efecto compañeros*³². La idea que subyace en este concepto es que el nivel de conocimientos que un alumno asimila a lo largo de los años en la escuela depende directamente de las características de sus compañeros y en definitiva de distintas iteraciones y externalidades sociales.

Este hecho supone que un alumno de alto rendimiento que asiste a una escuela deja de generar una externalidad positiva en otra escuela. Así la asignación óptima de alumnos a lo largo de las escuelas, así como que padres y alumnos aceptaran esa asignación sería una cuestión tampoco resuelta por el mercado y que no se incorpora claramente en los modelos productivos. El objetivo social educativo puede alcanzarse en menor o mayor medida en función del grado de heterogeneidad que exista dentro del aula. Sin embargo, los alumnos mejor dotados tanto en habilidad como en recursos económicos tenderán a segregarse de los peor dotados que a su vez tratan de buscar las externalidades positivas que les proporcionan los alumnos de este primer grupo.

3.3.2. Algunas Dificultades en la Estimación Econométrica de la Función de Producción Educativa mediante el Esquema Input-Output.

El elevado desconocimiento acerca de las interrelaciones de todos estos factores implicados junto a la excesiva simplicidad del esquema input-output de la función de producción educativa supone que en la investigación aplicada se obtengan resultados contradictorios. Por ejemplo, en el artículo ya citado, Hanushek (1986) recoge 147 prestigiosos³³ trabajos, (tabla 3.1), que estudian la escuela pública y que cubren todas las regiones de los Estados Unidos, diferentes niveles de enseñanza, diferentes tipos de medidas de actuación y diferencias en los procedimientos estadísticos y analíticos que se han utilizado. A la luz de estos datos se observa que, a la hora de estimar empíricamente la función de producción educativa, no se obtienen conclusiones del todo claras, sino más bien contradictorias, en cuanto

³² Para una revisión véase, Betts (2000).

³³ Hanushek señala en su meta-análisis que estos estudios reúnen 3 características: 1) publicados en libro o revista prestigiosa. 2) relaciona los resultados del alumno con características de la familia. 3) proporciona información sobre si la variable es o no estadísticamente significativa.

a la relevancia de uno u otro input en el rendimiento del alumno. Se hace por tanto evidente la carencia de un modelo de producción educativa claramente definido.

Tabla 3.1: Resumen de los coeficientes estimados en 147 estudios de funciones de producción educativas.

Input	Número de Estudios	Estadísticamente Significativo		Total	Estadísticamente no significativo		Signo Desconocido
		+	-		+	-	
Ratio prof/alumno	112	9	14	89	25	43	21
Educac. del profesorado	106	6	5	95	26	32	37
Experiencia del profesor	109	33	7	69	32	22	15
Salario del profesorado	60	9	1	50	15	11	24
Gasto por alumno	65	13	3	49	25	13	11

Fuente: Hanushek, E.A. (1986). pág. 1161.

Este resultado que cuestiona la influencia de la escuela en los resultados de rendimiento académico, también obtenido por otros autores, [Deller et al., 1993; Pritchett et al., 1999], ha sido objeto de diversas explicaciones que lo justifican. Según Worthington (2001), la primera explicación sería la metodología input-output de estos modelos, en la cual se tiende a tomar la información de inputs escolares que existen en otros trabajos o la única que esté disponible y a partir de ella se hace el análisis sin observar las características particulares del contexto educativo.

Una segunda explicación, realmente pesimista, sería que la escuela no importe a la hora de explicar el resultado de los alumnos. Según esta explicación la capacidad innata y el origen socioeconómico del individuo serían determinantes del éxito o fracaso académico. La *teoría de la señalización*, propuesta por economistas y sociólogos como Berg (1970), Spence (1973), Arrow (1973) o Stiglitz (1975), defendería que la educación es un filtro para señalar la capacidad innata de la persona de cara al oferente del trabajo y, por tanto, la escuela no añadiría capital humano al individuo ni le haría más productivo, tan solo le destacaría del resto a la hora de alcanzar los mejores puestos de trabajo.

Finalmente, los estudios de producción educativa asumen que todas las escuelas dispondrían de la misma tecnología lineal, aditiva y homotética transformadora de inputs

escolares en resultados académicos cuando en absoluto parece obvio, a tenor de la revisión de la literatura, que este esquema tan sencillo sea el caso real.

A este respecto, Eide y Showalter (1998) señalan, a partir de los resultados obtenidos en una regresión por cuantiles de inputs escolares sobre el resultado académico, el siguiente comentario: “nuestros resultados sugieren que pueden existir diferentes efectos de los inputs escolares sobre las ganancias en resultados dependiendo del punto de la distribución condicional de resultados donde nos encontremos. Este hecho es muy útil para la toma de decisiones del sector público, ya que algunos recursos escolares que parecen no tener ningún efecto sobre la media de rendimiento del centro pueden tener gran influencia en distintos tramos de la distribución condicional de las ganancias de rendimiento”.

Como hemos visto anteriormente, Summers y Wolfe (1977) obtuvieron conclusiones similares: diferentes inputs escolares son importantes para diferentes alumnos, usando como variable dependiente el rendimiento en valor añadido, esto es, la diferencia entre el nivel de conocimientos de un grupo de alumnos en los niveles tercero y sexto.

Todos estos resultados de diversos autores apuntan otra de las dificultades que existen cuando los datos son agregados al nivel escolar, es el problema conocido como *falacia ecológica*, término acuñado por Robinson (1950). Este problema consiste en suponer que las correlaciones existentes a nivel agregado son las mismas que a nivel individual. William Robinson demostró que las correlaciones que existen a nivel individual pueden ser mayores, menores o incluso con el signo contrario que las encontradas a nivel agregado debido a que la variabilidad existente a nivel individual desaparece al usar los valores medios³⁴.

Por otro lado, el uso habitual del resultado escolar promedio de la escuela como output educativo cuando se estima la función de producción educativa para medir la eficiencia técnica presenta también otras debilidades:

1. No recoge la variabilidad de los resultados dentro del aula. Tal y como señalan Hanushek et al. (1996), la probabilidad de que un input escolar aparezca como significativo aumenta con el grado de agregación de los datos pudiendo este hecho conducirnos a resultados engañosos.

³⁴ Para una extensión de los posibles errores que se pueden producir en los estudios agregados puede acudir a Theil (1954).

2. El resultado en un test de rendimiento depende de la educación recibida por un alumno en el curso presente, pero dado el proceso acumulativo de la educación también depende de la educación recibida en años anteriores. Hanushek (1986).
3. No suele distinguir entre diferentes aulas dentro de un colegio o diferentes profesores que enseñan en distintas aulas de un mismo nivel educativo. Meyer (1997).
4. En distritos con una alta movilidad de estudiantes los resultados podrían quedar sesgados por la incorporación o la fuga de alumnos. Meyer (1997).
5. No recoge el valor que la escuela añade al estudiante independientemente de lo que el alumno *consigue* por sus propias características individuales. Si reconocemos que el alumno es también materia prima para su educación, parte del resultado que éste obtiene es debido a su propia dotación individual y por tanto no todo el resultado es debido a la escuela. Así, deberíamos comparar alumnos con la misma *materia prima* en distintas escuelas de tal forma que aquellas escuelas que consiguen en sus alumnos resultados de rendimiento reales mayores que el potencial esperado para escuelas con alumnos de parecidas características serían eficientes. La diferencia entre el resultado real de un alumno y el que debería obtener según sus características individuales sería el verdadero output educativo, el valor añadido que la escuela aporta al alumno.

Un acertado intento de solucionar todos estos problemas constituyen los llamados *modelos multinivel*. A pesar de que las proposiciones teóricas de estos modelos datan de los años setenta, como es el caso del trabajo de Lindley et al. (1972), no es hasta finales de los años 80 y primeros 90 cuando estos modelos comienzan a aplicarse a partir del desarrollo de la computación moderna y de las exigencias de ofrecer soluciones a problemas sociales caracterizados por estar anidados. Así, los modelos multinivel³⁵ son utilizados en contextos donde los datos poseen una estructura claramente de tipo jerárquica o anidada³⁶. En nuestro caso, los alumnos estarían anidados o pertenecerían a distintos contextos más amplios como son el aula en el que reciben las clases, su colegio, el barrio, la ciudad, la región de residencia, etc.

A pesar de que estos modelos han supuesto indudablemente un avance en la estimación de los problemas con datos multinivel, esta metodología asume determinadas restricciones estadísticas. En particular la linealidad del modelo, los supuestos acerca de la distribución de las variables aleatorias y la forma paramétrica de la función de producción educativa, que puede no corresponderse exactamente con un proceso tan complicado a priori como el educativo en el que

³⁵ Para una revisión de la literatura de los modelos multinivel puede acudir a Bryk y Raudenbush (1992), Goldstein (1987, 1995) y Morris (1995).

³⁶ Por ejemplo los enfermos estarían anidados en hospitales y éstos a su vez en áreas de salud; los encuestados, en los entrevistadores; o las muestras de rocas, en sus localizaciones geográficas.

influyen tanto variables contextuales como escolares con distintos efectos sobre distintos tipos de alumnos. Además no permite la introducción de un contexto multioutput como el educativo.

Podemos concluir esta breve revisión acerca de la función de producción educativa apuntando que, actualmente, si el Sector Público pretende aumentar el rendimiento académico y reducir el fracaso escolar a través de aumentos significativos del gasto público sobre determinados factores productivos escolares, sin tener en cuenta todos los condicionantes que actúan sobre los resultados de cada alumno, en el mejor de los casos no tendrá la seguridad de conseguirlo y en el peor, como Hanushek (1986) señala, “conseguirá el efecto contrario al deseado”.

3.4. Debilidades Metodológicas en el Modelo Input-Output para la Medición de la Eficiencia en las Escuelas. Medidas de Resultados, Clima Escolar y Modelos en Valor Añadido.

Para introducir el concepto de ineficiencia productiva comentado en el capítulo anterior, expresamos la ecuación (3.2) a partir de la siguiente desigualdad:

$$A_{it} \leq f(B_i^{(t)}, P_i^{(t)}, S_i^{(t)}, I_i) \quad (\text{ec. 3.2})$$

En este caso el rendimiento real de los estudiantes podría ser menor que el máximo obtenible a partir de las variables escolares. Una manera alternativa de expresar la ecuación (3.2) se puede observar en la ecuación (3.3).

$$A_{it} = f(B_i^{(t)}, P_i^{(t)}, S_i^{(t)}, I_i) - v \quad (\text{ec. 3.3})$$

Donde $v \geq 0$ representa el grado de ineficiencia de la unidad productiva. Si $v > 0$ existirían ineficiencias. La ecuación (3.3) es en la práctica una frontera productiva.

Se pueden clasificar las técnicas comúnmente utilizadas para estimar la eficiencia técnica a nivel escolar en dos grandes conjuntos: los modelos econométricos y los modelos matemáticos de programación lineal. En primer lugar, los estudios econométricos³⁷ han tratado de estimar, de forma paramétrica, la producción media esperada de todas las unidades de la muestra, de tal forma que, aquellas unidades por encima de la recta de regresión serían eficientes y aquellas por debajo ineficientes.

³⁷ Para una revisión de los métodos paramétricos aplicados en educación no universitaria puede acudir a numerosos trabajos, entre ellos destacamos los de Gray et al. (1986) y Callan et al. (1990).

Las técnicas econométricas pueden ser utilizadas como fronteras asumiendo una forma funcional más flexible, como la translog, y estimando bajo máxima verosimilitud. Sin embargo, la principal crítica a las herramientas paramétricas alude a que esta metodología supone la forma funcional de la función de producción educativa, cuando ésta es totalmente desconocida.

En segundo lugar, en la medición de la eficiencia técnica de las unidades productivas públicas, ha cobrado una fuerte relevancia, especialmente en el ámbito educativo³⁸, el análisis envolvente de datos (DEA). Se define el modelo DEA que generalmente es aplicado en educación, como una técnica frontera, no-paramétrica y determinista. Estos modelos de programación matemática son mucho más flexibles que los econométricos a la hora de trazar la frontera productiva a partir de los mejores comportamientos de escuelas individuales. Sin embargo, estos modelos son muy sensibles a la especificación del modelo y a la presencia de observaciones atípicas.

3.4.1. Inputs Individuales y Escolares, Medidas de Resultados y Clima Escolar.

Como se ha visto, la medición de la eficiencia en la educación no universitaria ha implicado en la literatura educativa la maximización del rendimiento educativo medio y la idea de construir la frontera productiva a partir de los inputs individuales, escolares y ambientales implicados en el proceso productivo. Sin embargo en los últimos años, tal y como se ha observado anteriormente con los modelos jerárquicos o multinivel, existe una nueva tendencia de intentar separar el *valor añadido* que la escuela aporta al alumno de la influencia en el resultado que ejercen las propias características personales y familiares del alumno. Ejemplos de esta propuesta mediante métodos paramétricos y no paramétricos son los trabajos de Meyer (1997), Thanassoulis et al. (1999) y Ladd et al. (2002).

El sector público difícilmente puede intervenir en las características personales y familiares del alumno, tales como su condición socioeconómica, nivel de estudios de los padres, recursos en el hogar o inteligencia innata. Sin embargo, se observa cómo estas variables no enteramente escolares, e intrínsecas al individuo, influyen sustancialmente en el resultado académico.

³⁸ Algunos trabajos que aplican DEA en educación serían los de Ray (1991), Mc Carty et al. (1993) o Thanassoulis y Dustan (1994). Para una revisión de los aspectos metodológicos de la técnica DEA aplicada en educación a lo largo de la literatura puede acudir a Worthington (2001). Para el caso particular de la aplicación empírica en España puede acudir a Pedraja y Salinas (1996) y los excelentes trabajos de tesis doctoral de Mancebón (1996) y Muñiz (2000).

Tal y como señala San Segundo (1998) para el caso español, “la educación de los padres³⁹ parece afectar de manera apreciable la demanda de educación de los hijos. En 1996, sólo un 12% de los jóvenes cuyos padres carecen de estudios acceden a la universidad, mientras que la probabilidad de que el hijo de un titulado superior acceda a la universidad es el 72%”. En esta línea, Salas et al. (1998) explican las elecciones educativas de los jóvenes como una función del logro educativo de sus padres, siendo los estudiantes de niveles educativos familiares más elevados los que mayor probabilidad tienen de cursar una carrera de ciclo largo.

Por tanto, a pesar de que actualmente los modelos económicos que analizan el proceso educativo consideran que todos los alumnos son iguales, ya que a todos ellos se les aplica los mismos recursos en aproximadamente la misma intensidad en una misma escuela, creemos que los alumnos no son unidades homogéneas sino más bien muy heterogéneas a la hora de enfrentarse a su educación. Cada alumno llega a la escuela con una dotación de inputs personal, distinta a la de sus compañeros, que le permitirá transformar el tiempo y los recursos que recibirá en la escuela en unos determinados resultados.

Así, pensamos que el resultado obtenido por un alumno sólo es comparable por el que obtiene otro alumno de sus mismas o parecidas características personales e intelectuales, en términos económicos diríamos *que utilizase su misma tecnología productiva*. Tratar de comparar el rendimiento de un alumno con el de otro sin fijarnos en la tecnología que utiliza para transformar los recursos sería equivalente a comparar *peras y manzanas*. Éste ha sido uno de los supuestos claves en el análisis de eficiencia, la homogeneidad de los alumnos a la hora de ser agregados a nivel de escuela. Si tal y como se ha comprobado a lo largo de la literatura, los alumnos son unidades productivas distintas, los resultados que se obtendrían una vez realizada la agregación serán engañosos.

En esta línea de superación del modelo input-output, asumiendo la misma tecnología para todos los protagonistas educativos, estarían los modelos contexto-input-proceso-output postulados entre otros por Scheerens (1992). Estas investigaciones acerca de la eficacia de las escuelas introducen variables relativas al clima escolar como claves del proceso educativo.

³⁹ El nivel de estudios está altamente correlacionado con el ingreso y los recursos en el hogar que son puestos a disposición del alumno así como en la motivación y la capacidad de seguir estudiando.

Se asume, en el modelo presentado a continuación, que las tecnologías que poseen los alumnos para transformar factores productivos en resultados son distintas como consecuencia de sus características personales⁴⁰.

No deben ser confundidos los recursos individuales, la capacidad de aprendizaje, los recursos escolares y el clima escolar:

- *Recursos Individuales*: Serían las características cuantificables del alumno y básicamente estarían formadas por su condición socioeconómica y su habilidad innata. Algunas de estas variables pueden cambiar con el tiempo, si bien sería lentamente y no a través de un proceso automático, pero otras no. Por ejemplo en principio la inteligencia puede ser entrenada y se pueden dar becas o ayudas para dotar el hogar del alumno de mayores recursos, o favorecer el acceso a determinados recursos en el hogar. Sin embargo, es más difícil y, a veces totalmente imposible, cambiar a corto plazo el nivel de estudios o de ingresos de los padres así como las variables de contexto como el número de hermanos o las dimensiones del hogar.
- *Tecnología transformadora individual o capacidad de aprendizaje*: Es la dimensión más difícilmente cuantificable, y serían aquellos inputs o recursos individuales que influyen en la F o tecnología individual de cada alumno. Cada ser humano es física, genética y emocionalmente distinto de otro y este hecho importa en la educación, tanto como factor productivo como factor que determina la propia tecnología productiva, y por tanto no puede escapar al modelo económico a la hora de explicar el rendimiento académico. Así la extraversión del alumno, su timidez, motivación, inteligencia, aspiraciones, capacidad de liderazgo y de esfuerzo, etc. varían sensiblemente de un individuo a otro. Esta complejidad nos hace asumir la hipótesis de no linealidad en el modelo. Esto explicaría porqué distintos alumnos con los mismos recursos individuales transforman la educación con distinta productividad.

Podemos identificar en parte esta tecnología transformadora definida anteriormente con el concepto de *Inteligencia emocional*⁴¹. Tras distintos intentos de definición más o menos exitosos, el término inteligencia emocional fue utilizado por

⁴⁰ Este es uno de los supuestos claves del modelo que proponemos y desarrollaremos a continuación. El objetivo es profundizar en el conocimiento de la *caja negra* y para ello es preciso relajar supuestos tradicionales como la consideración de que todos los alumnos son igual de productivos a la hora de recibir la educación

⁴¹ Para una discusión en torno a este concepto y su medición puede acudir a Mestre et al. (2000). Otro de los muchos términos que ha sido usado para nombrar este concepto es el de Educación Informal.

primera vez por Peter Salovey de la Universidad de Harvard y por John Mayer de la Universidad de New Hampshire, en 1990, como “la capacidad de controlar y regular los sentimientos de uno mismo y de los demás y utilizarlos como guía del pensamiento y de la acción”. La inteligencia emocional se concreta en un amplio número de habilidades y rasgos de personalidad: empatía, expresión y comprensión de los sentimientos, independencia, capacidad de adaptación, capacidad de decisión, simpatía, capacidad de resolver los problemas de forma interpersonal, habilidades sociales, persistencia, cordialidad, amabilidad o respeto.

Pero fue Goleman (1995) con su libro *Inteligencia Emocional* quien popularizó el concepto. Así, para Goleman inteligencia emocional sería una forma de interactuar con el mundo que tiene muy en cuenta los sentimientos, y engloba habilidades tales como el control de los impulsos, la autoconciencia, la motivación, el entusiasmo, la perseverancia, la empatía, la agilidad mental, etc. Ellas configuran rasgos de carácter como la autodisciplina, la compasión o el altruismo, que resultan indispensables para una buena y creativa adaptación social. Para este autor el rendimiento escolar del estudiante depende del más fundamental de todos los conocimientos, la capacidad de aprender a aprender.

Las dimensiones fundamentales a tener en cuenta en la inteligencia emocional en la educación no universitaria serían las siguientes:

1. Confianza. La sensación de controlar y dominar el propio cuerpo, la propia conducta y el propio mundo. La sensación de tener muchas posibilidades de éxito en la tarea emprendida y que los adultos pueden ayudar en esa tarea.
2. Curiosidad. La sensación de que el hecho de descubrir algo es positivo y placentero.
3. Intencionalidad. El deseo y la capacidad de lograr algo y de actuar en consecuencia. Esta habilidad está ligada a la sensación y a la capacidad de sentirse competente, de ser eficaz.
4. Autocontrol. La capacidad de modular y controlar las propias acciones en una forma apropiada a la edad del estudiante; la sensación de control interno.
5. Relación. La capacidad de relacionarse con los demás, una capacidad que se basa en el hecho de comprenderles y de ser comprendido por ellos.
6. Capacidad de comunicar. El deseo y la capacidad de intercambiar verbalmente ideas, sentimientos y conceptos con los demás. Esta capacidad exige la confianza en los demás (incluyendo a los adultos) y el placer de relacionarse con ellos.

7. Cooperación. La capacidad de armonizar las propias necesidades con las de los demás en las actividades grupales.

Hasta hace poco, la habilidad del individuo era medida básicamente a través del coeficiente intelectual. La inteligencia, en esos términos, era entendida como un bloque monolítico de habilidades intelectuales; por eso, ante la evidencia de que hay algo más que nutre a ese cúmulo de saber, se comienza a buscar y tratar de medir en el individuo aquello que actualmente se denomina inteligencia emocional. Así, se observa que muchos niños que estaban por debajo del estándar aceptable de coeficiente intelectual, obtenían logros y resultados exitosos en su rendimiento escolar. A la vez, junto a estos alumnos se aprecia el otro extremo; aquellos individuos con altos coeficientes intelectuales que no lograban alcanzar los objetivos de éxito planteados.

En economía estas características sólo son cuantificables a través de variables proxy muy burdas tales como el tiempo dedicado a distintas tareas fuera de la escuela como estudiar, practicar deporte o leer libros; implicaciones en conflictos escolares, opiniones, relaciones con los compañeros, etc. La interrelación existente de estas variables con el resultado escolar así como con los inputs individuales y escolares es evidente. Es por ello que uno de los objetivos de la escuela sería tratar de informar a padres y enseñar a los alumnos el tipo de inteligencia emocional que a largo plazo garantizaría mejores resultados para cada tipo de alumno.

Todas las dimensiones de la inteligencia emocional que no se pueden medir cuantitativamente de forma directa, pero que la escuela y los padres pueden desarrollar, fomentando por ejemplo el compañerismo, la solidaridad, el hábito de estudio, la curiosidad y en definitiva el clima favorable para la obtención de buenos resultados deberían tenerse en cuenta a la hora de medir la eficiencia. Por otra parte, una buena escuela será aquella que consigue que sus alumnos actúen por encima de la media esperada para ellos, que estimaremos a partir de los resultados de otros alumnos con características innatas y socioeconómicas similares.

De aquí se deriva el concepto de valor añadido escolar. En el modelo propuesto en este trabajo como alternativa a los tradicionales se introduce, como novedad en la investigación educativa, todas las variables individuales y además se investigan las posibles relaciones no lineales existentes entre ellas. Se considera, tal y como se acaba de señalar, que un colegio que aporta más valor añadido es aquel capaz de hacer que un mayor porcentaje de alumnos actúe por encima de lo que de ellos sería esperable si

tenemos en cuenta como actuaron en media otros alumnos en otros colegios con parecidas características. Los inputs escolares y la eficiencia técnica que explican las diferencias en el valor añadido obtenidas por diferentes escuelas serán analizados en una segunda etapa.

- *Variables escolares:* Son los inputs escolares cuantificables. Básicamente estarían formados por los recursos escolares que tradicionalmente se incluyen en los estudios tradicionales de eficiencia y de producción educativa. Algunas de estas variables serían el número de alumnos por profesor, calidad del profesorado, recursos materiales tales como libros, ordenadores o instalaciones, etc. Estas variables pueden ser modificadas a partir de más recursos económicos y en principio estarían correlacionadas positivamente con los resultados aunque el resultado podría ser incluso negativo si no se recoge el resto de variables de clima escolar y el tipo de alumno implicado en el proceso educativo.

- *Clima Escolar:* Podemos definir el clima escolar como todos aquellos factores relacionados con la organización y el funcionamiento de la escuela. Los protagonistas de gestionar esta dimensión no sólo serían el director y los profesores de la escuela sino también los padres de los alumnos que asisten a la misma a través del grado de implicación que asuman. El buen entendimiento entre los profesores a la hora de compartir ideas y solucionar problemas, actualización de métodos pedagógicos, asistencia a cursos de reciclaje por parte del profesorado, ayuda personalizada a alumnos de riesgo, implicación de los padres en el funcionamiento de la escuela, control de conflictos internos, elaboración de los materiales educativos, liderazgo del equipo de dirección, etc. El clima escolar se corresponde por tanto con el concepto económico de eficiencia técnica, es decir, aquellas escuelas que actúan por debajo de sus posibilidades serían técnicamente ineficientes debido a que por distintos motivos, no siempre detectables, no están aprovechando los recursos el máximo de lo que sería posible.

A igualdad de variables escolares y tipología del alumnado entre dos escuelas lo que diferenciaría a una escuela con buenos resultados de otra con malos serían los factores relacionados con el la eficiencia escolar.

3.4.2. Indicadores y Modelos Educativos en Valor Añadido

Veáse el siguiente modelo. Sea una escuela que recibe a dos alumnos. El alumno tipo 1 tiene un status socioeconómico alto y el alumno tipo 2 bajo, siendo el resto de características homogéneas para ambos⁴². Este status es indicativo del tipo de tecnología que permite a los alumnos transformar el tiempo y los inputs recibidos en la escuela en rendimiento académico, de tal manera que la tecnología del individuo con nivel socioeconómico alto domina o es más productiva que la tecnología del individuo de nivel socioeconómico bajo, esto es, permite transformar la misma cantidad de recursos escolares en mejores resultados⁴³.

Supongamos además que el objetivo, tanto de los padres como de los profesores de ambos alumnos es maximizar un único output educativo que llamaremos resultado escolar (Y). Existen además dos factores productivos escolares para producir educación: gasto en material escolar por alumno (X_1) y gasto en atención personalizada por alumno (X_2). Cada uno de los alumnos es más intensivo en la utilización de cada uno de los factores productivos. El alumno 1 es más intensivo en X_1 y su productividad es además mayor en este factor que la del alumno 2. Por otro lado, el alumno 2 es intensivo en X_2 y su productividad es mayor en este factor que la del alumno 1. Un aumento equiproporcional de ambos inputs supondría un mayor incremento del resultado en el alumno 1. La función de producción satisface por tanto:

$$\frac{\partial F_i}{\partial x_i} > 0, \frac{\partial^2 F_i}{\partial x_i^2} < 0, \frac{\partial F_1}{\partial x_1} > \frac{\partial F_1}{\partial x_2}, \frac{\partial F_2}{\partial x_1} < \frac{\partial F_2}{\partial x_2}, \frac{\partial F_1}{\partial x_1} > \frac{\partial F_2}{\partial x_1}, \frac{\partial F_1}{\partial x_2} < \frac{\partial F_2}{\partial x_2}, \frac{\partial F_1}{\partial X} > \frac{\partial F_2}{\partial X}$$

donde $i = 1, 2$ representa a cada alumno. El coste de los factores productivos se determina exógenamente por el sector público en un nivel C . Además, el sector público decide asignar el mismo presupuesto educativo al alumno 1 y al alumno 2. Esta asignación se justifica en criterios de equidad ya que asumimos que el organismo planificador no tiene información a priori ni de la tecnología ni de la productividad de cada tipo de alumno.

⁴² El modelo podría ser fácilmente ampliado a más casos teniendo en cuenta por ejemplo además del nivel socioeconómico, el resultado en alguna prueba de inteligencia como proxy de la capacidad innata o alguna proxy de la inteligencia emocional del alumno. En definitiva el supuesto que hacemos es que los alumnos no son homogéneos, es decir sus funciones de producción son distintas.

⁴³ De nuevo hemos supuesto, para simplificar, que los alumnos de alto status poseen la tecnología más productiva asumiendo que cuentan con más recursos en el hogar. Por eso, a iguales recursos escolares mejor rendimiento para el alumno con mejores recursos familiares. Sin embargo, como ya hemos comentado, las dimensiones referentes a la capacidad innata y a la inteligencia emocional pueden hacer que un alumno con menos recursos familiares obtenga resultados escolares significativamente mejores.

Así el problema de optimización que debería resolver la escuela sería:

$$\begin{aligned} \text{Max } y_i &= F_i(x_1, x_2) \\ \text{s.a. } C &= C(p_{x_i}, x_1, x_2) \end{aligned}$$

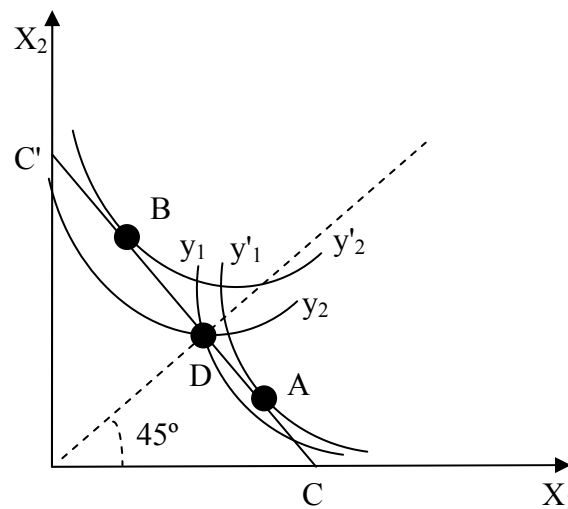
donde:

$$\frac{\partial F_1}{\partial x_1} = \frac{\partial C}{\partial x_1} = p_1^*, \quad \frac{\partial F_1}{\partial x_2} = \frac{\partial C}{\partial x_2} = p_2^*$$

$$\frac{\partial F_2}{\partial x_1} = \frac{\partial C}{\partial x_1} = p_1^*, \quad \frac{\partial F_2}{\partial x_2} = \frac{\partial C}{\partial x_2} = p_2^*$$

Si además se asume igualdad en el precio de los factores productivos, se tiene que el alumno 1 alcanzará el óptimo utilizando más factor productivo X_1 respecto a X_2 . Por el contrario el alumno 2 alcanzará el óptimo justamente a través de la asignación contraria, es decir, a través de un mayor empleo del factor X_2 respecto a X_1 .

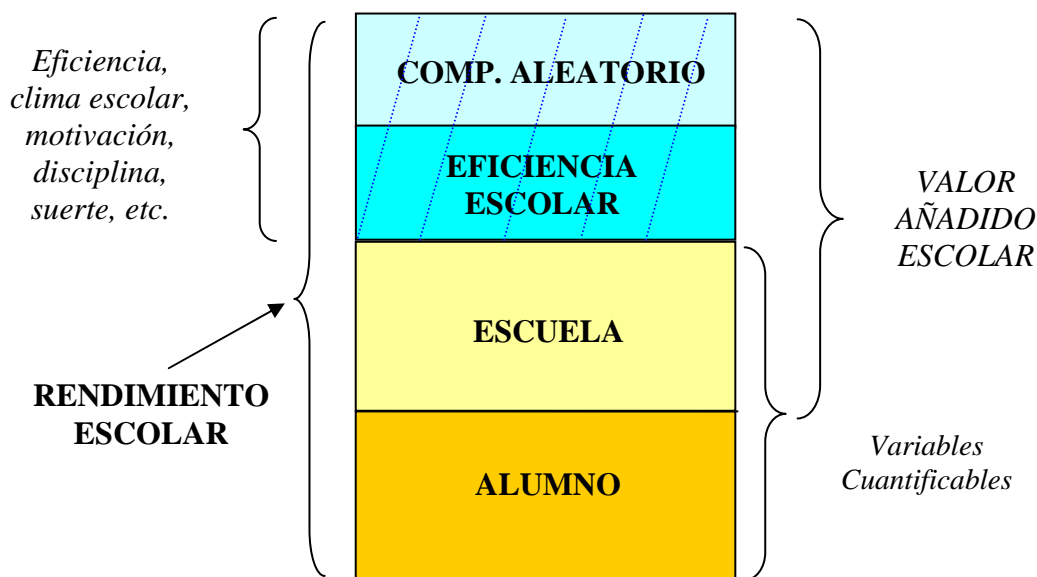
El gráfico 3.2 muestra la asignación de un planificador central que desconoce absolutamente las funciones de producción de los dos alumnos del modelo. En este caso optaría por asignar a ambos alumnos los mismos recursos escolares (D) independientemente de las productividades que presentan respecto a los inputs considerados. Los puntos A y B muestran los óptimos donde los alumnos maximizan su rendimiento a partir de los recursos disponibles dados por la restricción presupuestaria CC' a través de las isocuantas y'_1, y'_2 . El punto D muestra el resultado de la política conocida como *café para todos*, donde se observa claramente como las isocuantas que cortan el punto D, y'_1, y'_2 , son menores que en el punto óptimo D.

Gráfico 3.2: Comparación de las asignaciones óptimas y la asignación de *café para todos*.

Dado que en el mundo real se desconocen las distintas funciones de producción individuales, se pueden tratar de estimar a partir de los distintos tipos de alumno que existen tanto por sus características personales como socioeconómicas. Sin embargo, parte del resultado educativo sería debido a la escuela y por tanto se debe analizar en una segunda etapa la parte no explicada en la primera; esto es, como han influido los recursos educativos sobre distintos tipos de alumnos. La parte que no puede ser explicada ni por el propio alumno ni por la escuela será debida a la eficiencia de la escuela y a factores aleatorios. Como ya se ha comentado en el capítulo anterior existen distintos factores que hacen que una unidad productiva homogénea, en nuestro caso una escuela, produzca por debajo de sus posibilidades. Los factores de incentivos y organización suelen estar detrás de este resultado.

En el gráfico 3.3 se muestra de forma esquemática las partes, ya definidas en el punto anterior, a las que se puede atribuir el resultado de un alumno. La idea de intentar *separar estadísticamente* o atribuir el resultado escolar a distintas dimensiones de un mismo proceso es sin duda un artificio económico para tratar de medir la eficiencia escolar. El proceso educativo se desarrolla simultáneamente en un amplio conjunto de escenarios interdependientes y mediante el modelo que se propone a continuación tan solo se aspira a tener un conocimiento más profundo de este proceso, pero reconociendo en todo momento la limitación existente a la hora de cuantificar de forma precisa la producción educativa.

Grafico 3.3: Factores que explican el resultado escolar educativo



Fuente: Elaboración propia.

Así el gráfico 3.3 muestra cómo el rendimiento escolar sería debido fundamentalmente a cuatro características: el alumno, la escuela, la eficiencia escolar y factores aleatorios. Escuela y alumno tendrían dimensiones más o menos cuantificables y el sector público podría actuar sobre alguna de ellas, fundamentalmente sobre la escuela, tanto aumentando los recursos como reasignando los ya existentes. Sin embargo, para que esta actuación sea eficiente es preciso conocer la tecnología transformadora de cada alumno y no aplicar de forma indiscriminada los mismos recursos a todos los alumnos.

El concepto que se pretende medir en este trabajo es *el valor añadido escolar*. No se debe confundir este concepto con la escuela, tal y como está definida en el gráfico 3.3.

En definitiva, el hecho de asistir a una escuela con más recursos es una condición necesaria para recibir una mejor educación. Pero este concepto no es condición suficiente para que la escuela sea eficiente o para que la escuela aporte valor añadido al alumno. Un buen alumno que en una escuela obtiene buenos resultados no tendría por qué hacer eficiente a la escuela. Para saber si esa escuela es o no eficiente vamos a tener que comparar los resultados de este alumno con el de otros con similares características en otras escuelas. Aquellas escuelas que obtengan en sus alumnos un mejor resultado que el esperado, una vez tenidos en cuenta las dotaciones escolares, serían eficientes y aquellas que lo obtienen por debajo ineficientes. El concepto de valor añadido escolar sería por tanto el siguiente.

Valor añadido escolar: Este concepto recoge, además de la escuela, todos los otros factores escolares que no son fácilmente cuantificables y que podemos dividir en eficiencia técnica y factores aleatorios como la suerte u otros. La eficiencia técnica de la escuela se corresponde con el concepto de clima escolar y de organización y estaría formado, tal y como ya se ha señalado, por los métodos pedagógicos, la motivación del profesor y la que éste sepa transmitir a sus alumnos dotándoles de mayor inteligencia emocional. La disciplina, el compañerismo, la solidaridad, el respeto o la curiosidad por aprender junto a otros factores relacionados con lo que sucede en el aula serían también variables de clima escolar. A igualdad de factores productivos el objetivo de los padres sería enviar a sus hijos a las escuelas que maximicen la dimensión de eficiencia técnica que junto a otros factores aleatorios determinarán el valor añadido escolar.

Si tuviéramos un modelo que permitiera separar individuo y escuela, sería fácil evaluar los factores productivos escolares. Esta información sería utilizada por el gestor de la política educativa para asignar los recursos con el objetivo de maximizar el rendimiento académico de los alumnos teniendo en cuenta sus especiales características. En el siguiente punto se aporta un modelo que, basado en esta idea, trata de superar las deficiencias en los modelos actuales de producción educativa.

3.5. Un Modelo en Valor Añadido para la medición de la Eficiencia Escolar

3.5.1. El Nivel de Alumno

A continuación, se presenta el siguiente modelo en dos etapas para el cálculo de la función de producción educativa a partir del concepto de valor añadido escolar. Sea la siguiente función de producción explicativa del rendimiento académico medido a nivel de alumno. Una primera y simple ecuación es:

$$A_{is} = F(B_{is}, IE_{is}) + \eta_{is} + \varepsilon_{is} \quad (\text{ec. 3.4})$$

Donde i se refiere a estudiantes tomados de forma individual y s se refiere a colegios. A_{is} es el rendimiento alcanzado por el alumno, B_{is} representa el conjunto de inputs individuales y familiares presentes y acumulados que suponemos influyen en el rendimiento, IE_{is} hace referencia a la variable de inteligencia emocional del alumno, η_{is} es el efecto de las variables escolares sobre el rendimiento del alumno i que deberá ser estimado y ε_{is} recoge el ruido en los datos así como otros factores determinantes que también influyen en el rendimiento pero que no

han sido observados. A través de (ec. 3.4) se obtiene una estimación \hat{A}_{is} del rendimiento para cada alumno.

La principal característica de (ec. 3.4) es que la forma funcional F , que relaciona los inputs individuales y el output en forma de rendimiento académico, no está definida de forma paramétrica. Para proceder a su construcción se puede acudir a los métodos econométricos clásicos, pero a la luz de la literatura sobre producción educativa se puede intuir que existe una fuerte interrelación entre numerosos y distintos inputs individuales, con la posible presencia de no linealidades.

Figlio (1999) critica que, a pesar del conocimiento de estas interrelaciones, los modelos econométricos a nivel de alumno continúen siendo lineales y homotéticos o independientes de la escala. Al relajar estos supuestos y usar funciones más flexibles, por ejemplo funciones translog frente a las habituales Cobb-Douglas, este autor detecta relaciones que de otra manera no hubieran sido encontradas y que nos llevarían a conclusiones erróneas, conclusiones obtenidas como ya se ha visto en numerosos trabajos, como que el factor escuela no importa a la hora de explicar el resultado del alumno. Por otra parte Baker (2001) utiliza, también a nivel de escuela, diversos modelos de RNAs y encuentra relaciones no lineales entre los inputs escolares y el resultado medio de la escuela. Santín et al. (2000) encontraron como un modelo que aplicaba un MLP realizaba mejores predicciones del resultado académico a nivel de alumno que un modelo de regresión múltiple como consecuencia de tener en cuenta las relaciones no lineales.

Por tales motivos, parece óptimo relajar el supuesto de que todos los alumnos utilizan una tecnología transformadora uniforme y acudir a un análisis no lineal. Para ello se utiliza un tipo de RNAs, el MLP y su mecanismo de aprendizaje supervisado. Con esta herramienta se construye, a partir de datos individuales, el modelo predictivo que permite indagar la parte del resultado escolar que se puede explicar en función de variables no escolares.

La segunda característica de (ec. 3.4) es su componente estocástico y no envolvente. Al construir este modelo de actuación promedio a nivel de alumno, se evitan las desventajas derivadas de un modelo que trabajara con los puntos extremos. Un modelo de este tipo sería muy sensible al ruido de los datos ya que evaluaría el comportamiento de muchos individuos en función del de unos pocos⁴⁴. Un modelo frontera a nivel de alumno tipo DEA nos dificultaría

⁴⁴ Correríamos el peligro de identificar en la frontera como muy eficientes, alumnos que con muy poco esfuerzo o con un contexto familiar muy desfavorable logran resultados brillantes y evaluar el comportamiento del resto de alumnos respecto a estos casos que sin duda existen pero que podemos considerar como *extraños* o poco frecuentes. Parece necesaria más investigación acerca de cómo medir la inteligencia emocional para explicar este tipo de resultados.

además formular objetivos a priori, esto es, predecir resultados ex-ante en períodos futuros para cada alumno.

El objetivo en este primer nivel es múltiple. Por una parte, se trata de explicar qué parte del resultado académico del alumno está poco o nada influenciado por el centro académico al que acude, dadas unas condiciones homogéneas para todos ellos, (por ejemplo, residencia en una determinada región). Por otra parte, esta función, construida a partir de criterios individuales, servirá en un futuro para predecir, ex-ante, la probabilidad de fracaso académico de los alumnos a partir de su resultado esperado. Si se valoran los criterios de equidad, parece lógico que el sector público intervenga para tratar de minimizar este resultado, sobre todo en los colectivos más desfavorecidos socioeconómicamente, a fin de garantizar la igualdad de oportunidades educativas antes de que empiece el proceso de la educación.

Uno de los principales objetivos que la sociedad demanda para la mejora de la educación no universitaria es la reducción del fracaso escolar, como condicionante del nivel de estudios que finalmente alcanzará el alumno y por tanto, de sus ingresos futuros esperados. Reduciendo el fracaso escolar el sector público invierte por un lado en capital humano, contribuyendo a la redistribución de la renta, ya que el fracaso escolar aparece ligado frecuentemente al nivel socioeconómico. Por otro lado, mejora la eficiencia del sistema educativo en su conjunto al lograr una mejora en los resultados productivos con los mismos recursos y por tanto una situación superior en sentido de Pareto a la anterior.

Además, se utilizará el resultado de la predicción del modelo para posteriormente, al finalizar el curso académico, comparar los resultados realmente obtenidos por los alumnos con los que según sus características deberían obtener y construir los outputs en valor añadido de la escuela y medir así el valor añadido escolar.

Por otro lado, la función de producción estimada servirá para identificar a los alumnos con una alta probabilidad de rendimiento elevado, sobre los que se debería aplicar una política educativa de gasto distinta evitando⁴⁵, en la medida de lo posible, la política de *café para todos*

⁴⁵ Esta aplicación de recursos distinta no implica necesariamente dividir a los alumnos en *buenos o malos*. Los alumnos más aventajados pueden influir positivamente a través del llamado “efecto compañeros” en los alumnos más desaventajados. Políticas como la introducción de profesores de apoyo que ayudaran a cada alumno en las asignaturas con más dificultades, institucionalizar las clases de recuperación o dividir las clases en grupos de trabajo dentro de una misma aula parecen alternativas más adecuadas y sobre todo en las edades más tempranas. En cualquier caso, parece necesaria una mayor inversión de recursos en el estudio de los métodos pedagógicos más eficientes, así como la introducción de incentivos al profesorado que hagan que su objetivo sea la maximización de las capacidades de sus alumnos según las características de éstos.

a lo largo de la distribución de resultados en la función de producción educativa, que como ya se ha señalado es completamente ineficiente.

3.5.2. La Construcción de Outputs en Valor Añadido.

Una vez obtenida la RNA óptima que minimice la suma de errores absolutos al cuadrado y que mejor generalice la relación entre inputs y resultados individuales, se tiene una expresión como la ecuación (3.5), en la que explicaremos el porcentaje de la varianza no explicada en el primer nivel a partir de los inputs escolares que varían a lo largo de los distintos centros educativos.

$$O_{js} = g_j(A_{is}, \hat{A}_{is}) = \eta_{js} + \varepsilon_{is} \quad (\text{ec. 3.5})$$

Es necesario señalar en este punto, que se construyen j outputs en valor añadido para cada escuela, a partir de las g_j relaciones que se definen entre A_{is} y \hat{A}_{is} . Para ello se siguen distintos criterios correspondientes a los distintos grupos de alumnos dentro de cada centro sobre los que se quiere investigar, con el fin de analizar la influencia y la eficiencia de cada escuela en cada una de las dimensiones del output en valor añadido a estudiar. La consideración de un único resultado, por ejemplo la media del rendimiento académico en cada centro, no parece adecuada a tenor de los resultados obtenidos en los distintos estudios que se ha comentado con anterioridad.

Se justifica por tanto, a partir de los resultados empíricos previos, la utilización del concepto de output en valor añadido y por tanto se buscarán diferentes medidas de eficiencia para cada centro según la dimensión del output que analizado. Algunos ejemplos de outputs en valor añadido que se utilizan posteriormente en el trabajo empírico son:

- Porcentaje de alumnos en cada centro que obtienen un rendimiento mayor que el esperado por el modelo según sus características.
- Porcentaje de alumnos en cada centro que según el modelo deberían obtener una puntuación menor que el tercer cuartil de rendimiento y sin embargo obtienen un rendimiento mayor. Esta dimensión analizaría los factores que hacen que un alumno con un resultado medio pase a estar entre el 25% de los mejores estudiantes.
- Porcentaje de alumnos en cada centro que según el modelo deberían obtener una puntuación menor que el primer cuartil de rendimiento y sin embargo obtienen un rendimiento mayor. Este output en valor añadido analiza los recursos escolares que

influyen en que un alumno consiga escapar del 25% de alumnos con peores resultados y reduciendo así sus probabilidades de fracaso escolar.

Aunque de cara al ejercicio empírico desarrollado en el siguiente capítulo sólo se han definido estos tres outputs, el modelo permitiría aumentar notablemente el número de outputs definidos en valor añadido. Para ello habría que tener en cuenta otros factores, tales como la distancia entre los resultados esperados y alcanzados a lo largo de la distribución de resultados o la especial consideración de determinados grupos de estudiantes socioeconómicamente más desfavorecidos. A este último grupo pertenecerían, por ejemplo, algunos colectivos de la población inmigrante, el grupo de alumnos con rentas familiares bajas o niños con minusvalías, problemas de psicomotricidad o de salud, por citar sólo algunos casos.

Por tanto, en algunos trabajos recientes subyace la idea de que la forma tradicional de trabajar empíricamente a nivel de centro escolar, sin analizar por separado los distintos tipos de alumno que componen el centro, puede conducir a errores. El objetivo de esta primera parte del modelo es hacer explícita la complejidad del nivel individual y la construcción de indicadores output en valor añadido. El alumno, materia prima con la que la escuela trabaja, no es una “pasta homogénea” que pueda ser agregada y elaborada con los mismos inputs, sino que cada grupo de alumnos debería ser tratado en función de sus especiales características. Así, un centro puede ser eficiente en el trato con un grupo de alumnos, por ejemplo con el colectivo de los más aventajados, y ser menos eficiente en sus resultados, respecto a los obtenidos por otros centros, en la educación de sus alumnos más desaventajados.

3.5.3. El Nivel Escolar

Con el fin de desarrollar un indicador que recoja la contribución del colegio a los logros académicos se define, a partir del modelo de rendimiento académico en valor añadido de Meyer (1997), la siguiente ecuación, (3.6), de segundo nivel:

$$\eta_{js} = \beta_{j0} + \beta_{j1} \text{Externas}_{js} + \beta_{j2} \text{Internas}_{js} + u_{js} - v_s \quad (\text{ec. 3.6})$$

Esta ecuación recoge los factores escolares que contribuyen al aumento del rendimiento académico, donde η_{js} es el efecto del colegio s sobre el output en valor añadido j que recoge la ecuación 3.6.

Externas.- Incluye todas las características escolares observadas que podrían considerarse externas al centro. Recoge por tanto características sociales de la

comunidad y características del contexto socioeconómico de los alumnos del centro y en general todas aquellas variables que no dependen del gasto corriente de la escuela.

Internas.- Incluye todas las características escolares observadas que podrían considerarse internas al centro sobre las que la escuela y el sector público pueden actuar con una mayor facilidad. En particular inputs escolares, y políticas del centro.

u_{js} .- Es el término correspondiente a la perturbación aleatoria. $u \sim \text{iid } N(0, \delta^2)$

v_{js} .- Representa la ineficiencia técnica de la escuela s para el output j . $v \sim \text{iid } N(0, \delta_v^2)$

El concepto de eficiencia trataría de medir la dimensión de *clima escolar*, o por qué dos escuelas con los mismos recursos obtienen distintos resultados. La eficiencia no mejoraría mediante más o menos recursos sino a través de políticas que incentiven su búsqueda, fundamentalmente a través de la organización, la introducción de incentivos a profesores y directivos vinculados a objetivos, pero también a través de la implicación de los padres en el proceso educativo de sus hijos, que en parte depende también de las iniciativas del colegio.

Desagregando así el efecto de la escuela se explican los j outputs construidos a través de la ecuación 3.7:

$$O_{js} = \beta_{j0} + \beta_{j1} \text{Externas}_{js} + \beta_{j2} \text{Internas}_{js} + (u_{js} - v_{js}) \quad (\text{ec. 3.7})$$

De esta forma, la frontera productiva tiene un carácter estocástico. En este punto es necesario hacer supuestos acerca de la distribución de v ; por ejemplo la semi-normal o la normal truncada, [Aigner et al., 1977 y Meeusen et al., 1977]. El término de error final sería por tanto una mezcla de eficiencia, ruido aleatorio y variables no introducidas en el modelo.

A través de esta metodología para la medición de la eficiencia y en particular dentro de este segundo nivel conseguiremos alcanzar los siguientes objetivos:

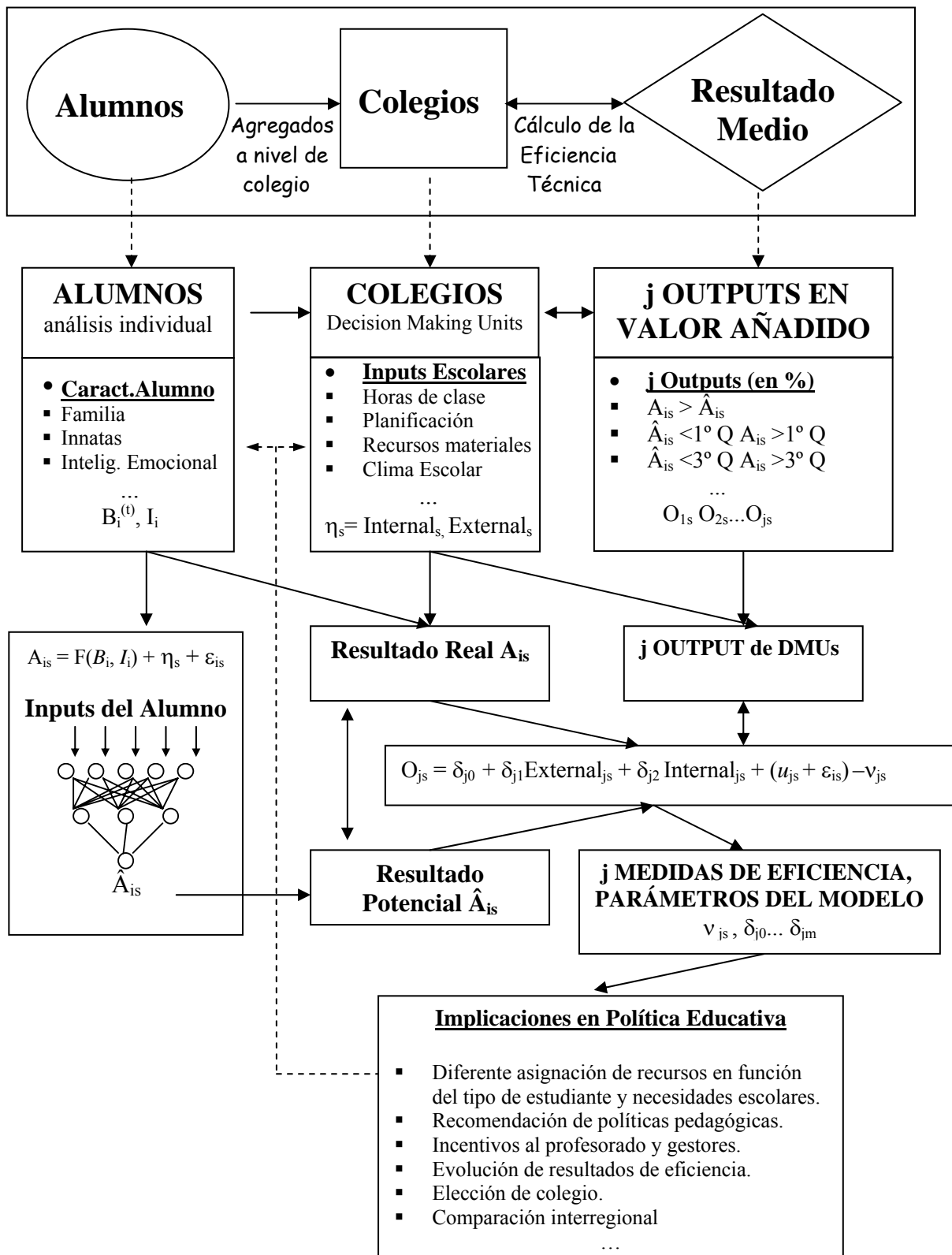
1. Cuantificar de forma paramétrica en qué medida influyen los distintos inputs escolares sobre los distintos outputs en valor añadido definidos correspondientes a los distintos colectivos de alumnos y, especialmente, en aquellos grupos con una probabilidad de fracaso escolar más elevada o con mayores limitaciones. Posible introducción de políticas de equidad en el sistema educativo según las preferencias sociales.
2. A partir del resultado anterior podremos asignar el gasto público educativo en aquellas escuelas que observemos peor dotadas en los inputs escolares que influyen de forma más significativa sobre el resultado en valor añadido. De igual manera, podremos

recomendar tanto la reducción o la reasignación de inputs en aquellas escuelas que así lo requiera, por ejemplo mediante métodos pedagógicos que en otras escuelas han contribuido a obtener buenos resultados.

3. Podremos ordenar a los centros por su grado de eficiencia-ineficiencia en los distintos outputs considerados. Estos niveles de eficiencia alcanzados deberían asimismo contribuir a la articulación de mecanismos de incentivos al profesorado a partir del establecimiento de objetivos a priori para cada centro.
4. La información que arroja el modelo construido puede ser utilizada, a principio de curso, para la elección de centro escolar según las características de cada alumno y de la eficiencia de los centros a los que pueda optar.

El gráfico 3.4 representa un resumen del modelo propuesto que será desarrollado empíricamente en el siguiente capítulo. Frente a la metodología tradicional de usar el resultado medio de la escuela como único output educativo, se propone completar este único resultado mediante la construcción de outputs en valor añadido. El objetivo final es obtener más información para la toma de decisiones a partir del modelo propuesto si se comprobase que distintos recursos influyen sobre distintos tipos de alumno.

Gráfico 3.4. Esquema del modelo en dos etapas propuesto para la medición de la eficiencia en las escuelas a partir de outputs en valor añadido.



Así, el objetivo es tratar de separar qué parte del resultado académico del alumno es debido a sus variables de contexto, a su inteligencia emocional, y por último a la escuela. La falta de un modelo de producción educativa junto a los resultados contradictorios que se obtienen en muchos trabajos es un indicador de que la tecnología que transforma inputs educativos en resultados puede ser bastante complicada debido a la multitud de factores productivos, que además están interrelacionados, que quedan implicados en este proceso. Por ello se justifica el uso de una metodología no lineal como las RNAs para aproximar dicha función de producción.

La parte del resultado que el modelo no explica con las variables del alumno deberá ser explicada en una segunda etapa con las variables escolares. Para este fin se construyen outputs en valor añadido, donde se considera que el aumento de un factor productivo es adecuado cuando es capaz de conseguir que los alumnos rindan por encima de lo esperado en función de sus características. En este trabajo se construyen tres outputs en valor añadido, si bien, como ha sido indicado anteriormente, podrían construirse otros en función de los objetivos educativos demandados por la sociedad.

Una vez construidos estos outputs en valor añadido, se trata de explicar las diferencias a lo largo de las escuelas con las variables propiamente escolares y del profesor que impartió la asignatura de matemáticas a los alumnos evaluados. El objetivo de esta segunda etapa es descubrir las variables escolares que propician la aportación de valor añadido sobre los distintos alumnos. Por otro lado, se mide la eficiencia de la escuela como la distancia entre el objetivo que debería haber alcanzado y el resultado que finalmente obtuvo. La suma de ambas dimensiones es lo que se ha denominado valor añadido escolar. Por último, toda la información que proporciona el modelo sería empleada para la toma de decisiones con el objetivo de introducir eficiencia en el sistema educativo. Así, el objetivo es tratar de separar qué parte del resultado académico del alumno es debido a sus variables de contexto, a su inteligencia emocional, y por último a la escuela. La falta de un modelo de producción educativa junto a los resultados contradictorios que se obtienen en muchos trabajos es un indicador de que la tecnología que transforma inputs educativos en resultados puede ser bastante complicada debido a la multitud de factores productivos, que además están interrelacionados, que quedan implicados en este proceso. Por ello se justifica el uso de una metodología no lineal como las RNAs para aproximar dicha función de producción.

3.6. Recapitulación

La literatura acerca de la función de producción educativa ha mostrado cómo el proceso educativo es muy distinto a otros bienes y servicios. Una mala y simple especificación de este modelo puede suponer la obtención de conclusiones erróneas a la hora de llevar a cabo recomendaciones de políticas educativas. En particular se considera que los alumnos no pueden ser agregados de forma simple ya que éstos son unidades heterogéneas debidos a su condición socioeconómica y su educación informal o inteligencia emocional.

Así la política de asignar los mismos factores productivos a individuos distintos con distinta tecnología es sin duda ineficiente. Para superar estas limitaciones se propone un modelo que calcule los outputs escolares en valor añadido que va a permitir estimar la importancia de cada factor productivo escolar sobre diferentes tipos de alumno.

La parte que no pueda ser explicada ni por el alumno ni por la escuela será la dimensión de eficiencia técnica escolar. El fin último del modelo es la introducción de un círculo virtuoso de mejora en el sistema educativo.

**CAPITULO 4: EL MODELO EN VALOR AÑADIDO PARA LA
MEDICIÓN DE LA EFICIENCIA ESCOLAR: UNA APLICACIÓN
AL CASO ESPAÑOL**

4.1. Introducción

En los albores del siglo XXI el progreso tecnológico está determinando sin ningún género de dudas el crecimiento económico y la riqueza futura de los países. Para asegurar una plena adaptación a los cambios tecnológicos, los países necesitarán ciudadanos más preparados que puedan participar en industrias relacionadas con todas las ramas de la ingeniería y la informática. Los alumnos de hoy buscarán empleos en un mundo cada vez más globalizado que requerirá de individuos con amplios conocimientos técnicos y con capacidad de adaptación. Para enfrentarse a todos estos cambios, nuestros alumnos necesitarán estar bien formados en todas las áreas pero especialmente en matemáticas y ciencias. Estas asignaturas permiten además una mejor comparación internacional por poder establecer pruebas homogéneas.

El hecho de que las matemáticas y las ciencias sean materias clave para el progreso económico, ha llevado a los países a buscar información acerca de lo que su población escolar sabe y puede hacer, con el objetivo de decidir qué políticas tomar para aumentar los conocimientos en estas asignaturas y mejorar la aptitud hacia el aprendizaje. Así, la preocupación por los resultados del Proyecto PISA⁴⁶ de la OCDE que situó a Alemania por debajo de la media internacional ha hecho que la revisión de la educación en este país se haya vuelto una prioridad electoral.

En este capítulo se lleva a cabo una aplicación empírica del modelo en valor añadido para el cálculo de la eficiencia en las escuelas presentado en el capítulo anterior, con el fin de ilustrar cómo debería ser empleado en la práctica. El objetivo que nos planteamos es tratar de buscar respuestas a las preguntas en torno a cómo mejorar la educación no universitaria así como proponer y aplicar una metodología para introducir eficiencia en el sistema educativo. Por último, se analizarán las políticas educativas que se derivarían del análisis efectuado para los datos españoles.

⁴⁶ El Proyecto Internacional para la Producción de Indicadores de Rendimiento de los Alumnos (PISA) es el resultado de un compromiso por parte de los gobiernos de los países miembros de la OCDE para establecer un seguimiento de los resultados de los sistemas educativos en cuanto al rendimiento de los alumnos, dentro de un marco internacional común.

4.2. ¿Qué es el TIMSS?

Como se acaba de comentar, en todos los países del mundo, las Matemáticas y las Ciencias son una parte importante del currículo escolar y se consideran materias fundamentales para la educación de los alumnos. Esto es así porque ambas materias son un pilar básico para la futura integración del individuo en el mercado de trabajo y le preparan para afrontar con éxito el reto tecnológico y científico que plantea el mercado de trabajo actual. De igual manera, el estudio de las Matemáticas y las Ciencias es considerado como un medio para desarrollar en la educación del alumno hábitos de razonamiento riguroso y crítico.

En consecuencia, hay un interés creciente por los proyectos de evaluación internacional del rendimiento en Matemáticas y Ciencias, a los que cada vez se adhieren más países. Algunas de las organizaciones que recogen información educativa son la OCDE (Organización para la Cooperación y el Desarrollo Económico), UNESCO (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura), Unión Europea, NCES (National Center for Educational Research), Australian Council for Educational Research (ACER, Consejo Australiano de Investigación Educativa), el Netherlands National Institute for Educational Measurement (CITO, Instituto Nacional Holandés de Evaluación Educativa), el Service de Pédagogie Expérimentale de l'Université de Liège (SPE, Servicio de Pedagogía Experimental de la Universidad de Lieja), etc.

Una de estas organizaciones es la IEA (The International Association for the Evaluation of Educational Achievement) de la que actualmente son miembros 58 países⁴⁷. Desde su fundación en 1959, la IEA ha desempeñado una única labor de especialización en la investigación comparativa internacional de la producción educativa en las etapas no universitarias. En sus más de 20 investigaciones de resultados escolares la IEA ha evaluado sobre todo materias básicas del currículo escolar como son las matemáticas y las ciencias a través de los IMSS (International Mathematics and Science Studies) y la comprensión lectora a través del PIRLS (Progress in International Reading Literacy Studies), siempre con datos internacionales de corte transversal.

⁴⁷ La agencia que coordina la participación española en los estudios de la IEA es el INCE (Instituto Nacional de Calidad y Evaluación) dependiente del MEC (Ministerio de Educación Cultura y Deporte). Para un resumen de como el TIMSS fue llevado a cabo en España puede acudirse a López et al.(1997a) y López et al. (1997b).

El primer estudio sobre el rendimiento en matemáticas de la IEA fue llevado a cabo durante 1963 y 1964 en once países. Más de 133.000 estudiantes de 13 y 17 años, 18.500 profesores y 5.450 escuelas participaron en este estudio, Husén (1967). El primer estudio en ciencias, Comber y Keeves (1973), fue llevado a cabo sobre 17 países y sobre alumnos de 10 y 14 años en el período comprendido entre 1966 y 1973. El segundo estudio en matemáticas, Robitaille et al. (1989), fue llevado a cabo de nuevo en 17 países, sobre alumnos de 10 y 13 años y sobre aquellos que cumplían el último curso de la secundaria en el período 1980-1982. Por otro lado, el segundo estudio en ciencias, Postlethwaite et al. (1992), se desarrolló para los mismos grupos de edad y sobre 23 países entre 1983 y 1986.

Posteriormente, durante el curso 94-95 tuvo lugar la aplicación por parte de la IEA de la primera oleada del TIMSS (The Third International Mathematics and Sciences Study). El TIMSS es hasta la fecha el estudio de evaluación de Matemáticas y Ciencias más extenso en recopilación de datos. En esta primera oleada del TIMSS participaron en total más de 500.000 alumnos de tres niveles no universitarios (4º, 8º y 12º nivel, que se corresponden respectivamente con las edades de 10, 13 y 17 años), de 15.000 escuelas en más de 40 países de todo el mundo y en más de 30 idiomas, siendo así la evaluación internacional con mayor número de países llevada a cabo hasta esa fecha. Además de la evaluación en matemáticas y ciencias se llevó a cabo una enorme recolección de información tanto personal, como de los métodos educativos referente a los alumnos, profesores y directores de las escuelas.

El objetivo del estudio es conocer el nivel de rendimiento de los alumnos, comparar los resultados entre países y tratar de explicar las diferencias observadas en función de las distintas características de los sistemas educativos. Por otro lado, se pretende dar a los encargados nacionales de asignar los fondos educativos una base científicamente sólida en torno a cómo funciona el proceso educativo para la toma de decisiones óptimas.

En España el estudio se llevó a cabo entre los meses de mayo y junio de 1995. Participaron 153 colegios con una muestra de 7596 alumnos, para el nivel de 8º grado, (único nivel en el que España participó de los tres evaluados), siendo 3855 de 8º de Educación General Básica (EGB) y 3741 de 7º de EGB. En esta aplicación empírica tan sólo se manejan los resultados correspondientes al nivel de 8º curso, que recoge mayoritariamente a la población de 13 años. La Tabla 4.1, contiene los resultados obtenidos por los países participantes en la prueba de matemáticas y ciencias. La puntuación media de los alumnos españoles de 8º de EGB es de 487 en matemáticas; esto representa el puesto 31º, (empatados con Islandia), sobre 41 países. En ciencias, el resultado medio es de 517; representando el puesto 26º, (empatados con Escocia), de nuevo sobre 41 países.

En la tabla 4.1 se observa además cómo Singapur es el país con mejores resultados, tanto en matemáticas como en ciencias. Otros países con buenos resultados fueron Corea del Sur, Japón, el grupo de alumnos flamencos de Bélgica y la República Checa. Entre los países con peores resultados destacan Sudáfrica, Colombia y Kuwait. Se observa además que las diferencias entre el grupo de países con buenos y con malos resultados son significativas. Junto a estos 41 países participaron además Argentina, Italia, Indonesia y México que no terminaron todos los pasos necesarios para presentar sus informes a tiempo y, por tanto, no aparecen en los resultados finales.

Tabla 4.1: Resultados medios en matemáticas y ciencias en 8º nivel (13 años) por país participante en el TIMSS de 1995.

Ranking	Países	Matemáticas	Ranking	Países	Ciencias
1	Singapur	643 (4,9)	1	Singapur	607 (5,5)
2	Corea del Sur	607 (2,4)	2	Rep. Checa	574 (4,3)
3	Japón	605 (1,9)	3	Japón	571 (1,6)
4	Hong Kong	588 (6,5)	4	Bulgaria	565 (5,3)
5	Bélgica (Fl)	565 (5,7)	5	Corea del Sur	565 (1,9)
6	Rep. Checa	564 (4,9)	6	Holanda	560 (5,0)
7	Eslovaquia	547 (3,3)	7	Eslovenia	560 (2,5)
8	Suiza	545 (2,8)	8	Austria	558 (3,7)
9	Holanda	541 (6,7)	9	Hungría	554 (2,8)
10	Eslovenia	541 (3,1)	10	Inglaterra	552 (3,3)
11	Bulgaria	540 (6,3)	11	Bélgica (Fl)	550 (4,2)
12	Austria	539 (3,0)	12	Australia	545 (3,9)
13	Francia	538 (2,9)	13	Eslovaquia	544 (3,2)
14	Hungría	537 (3,2)	14	Irlanda	538 (4,5)
15	Rusia	535 (5,3)	15	Rusia	538 (4,0)
16	Australia	530 (4,0)	16	Suecia	535 (3,0)
17	Irlanda	527 (5,1)	17	EE.UU	534 (4,7)
18	Canadá	527 (2,4)	18	Alemania	531 (4,8)
19	Bélgica (Fr)	526 (3,4)	19	Canadá	531 (2,6)
20	Israel	522 (6,2)	20	Noruega	527 (1,9)
21	Tailandia	522 (5,7)	21	N. Zelanda	525 (4,4)
22	Suecia	519 (3,0)	22	Tailandia	525 (3,7)
23	Alemania	509 (4,5)	23	Israel	524 (5,7)
24	N. Zelanda	508 (4,8)	24	Hong Kong	522 (4,7)
25	Inglaterra	506 (2,6)	25	Suiza	522 (2,5)
26	Noruega	503 (2,2)	26	Escocia	517 (5,1)
27	Dinamarca	502 (2,8)	27	España	517 (1,7)
28	EE.UU	500 (4,6)	28	Francia	498 (2,5)
29	Escocia	498 (5,5)	29	Grecia	497 (2,2)
30	Letonia	493 (3,1)	30	Islandia	494 (4,0)
31	Islandia	487 (4,5)	31	Rumania	486 (4,7)
32	España	487 (2,0)	32	Letonia	485 (2,7)
33	Grecia	484 (3,1)	33	Portugal	480 (2,3)
34	Rumania	482 (4,0)	34	Dinamarca	478 (3,1)
35	Lituania	477 (3,5)	35	Lituania	476 (3,4)
36	Chipre	474 (1,9)	36	Bélgica (Fr)	471 (2,8)
37	Portugal	454 (2,5)	37	Irán	470 (2,4)
38	Irán	428 (2,2)	38	Chipre	463 (1,9)
39	Kuwait	392 (2,5)	39	Kuwait	430 (3,7)
40	Colombia	385 (3,4)	40	Colombia	411 (4,1)
41	Sudáfrica	354 (4,4)	41	Sudáfrica	326 (6,6)

Fuente: TIMSS (1997) y elaboración propia.
Entre paréntesis figura la desviación estándar.

Además el TIMSS es un estudio que está siendo replicado en sucesivas oleadas con el objetivo de comprobar en el tiempo los progresos en las políticas educativas nacionales. En 1999 el TIMSS fue replicado para el nivel de 8º en 38 países de los cuales 26 participaron en 1995. Desafortunadamente España no participó en esta segunda oleada y es por ello por lo que utilizaré en esta aplicación empírica los datos correspondientes a 1995. De igual forma está previsto que se lleve a cabo una nueva oleada en 2003, la cual muy posiblemente no será la última, en la que España está previsto que vuelva a participar.

La información de los distintos países obtenida por el TIMSS ha sido analizada para buscar patrones comunes en la educación de los alumnos de 13 años. Algunas de las características más relevantes que se repiten son las siguientes, [Beaton et al., 1996].

- En todos los países se observa una relación clara entre una fuerte preferencia del alumno por las matemáticas y su resultado. Los países en los que existe un mayor desagrado por esta asignatura son la República Checa, Austria, Alemania, Hungría, Japón, Corea del Sur, Lituania y los Países Bajos.
- En ningún país les gusta en media más las matemáticas a las niñas que a los niños.
- Los recursos en el hogar resultan estar en todos los países fuertemente relacionados con los resultados en matemáticas. Los elementos más relacionados son la posesión de diccionario, ordenador, escritorio de estudio propio y número de libros.
- El nivel de estudios de los padres está positivamente relacionado con los mejores resultados en matemáticas del alumno.
- Ver televisión más de 1 o 2 horas diarias está negativamente asociado a malos resultados.
- En comparación con la fuerte relación encontrada entre recursos en el hogar y resultados escolares, la relación entre métodos de enseñanza y rendimiento en matemáticas está mucho menos clara. Se reconoce así que las interacciones entre todas las variables del contexto escolar pueden llegar a ser muy complejas y se concluye que más análisis es requerido para su comprensión.
- Existen diferencias considerables en los tamaños de las clases. En torno al 90% de los alumnos de estos países estaban en clases de menos de 30 alumnos. Sin embargo el 93% de los alumnos coreanos, uno de los países con resultados más brillantes, estaban en clases de más de 40 alumnos. La relación frecuentemente citada de que menos alumnos en clase supone más rendimiento no queda clara en absoluto.

- Uno de los resultados más observados es que la calculadora es muy frecuentemente usada en la clase de matemáticas casi todos los días. Una excepción es de nuevo Corea del Sur cuyos alumnos y profesores declaran que no usan habitualmente la calculadora en la enseñanza de las matemáticas.
- Si bien en torno a la mitad de países declaran que mandan deberes entre dos y tres horas diarias, no se encuentra ninguna relación entre la cantidad de deberes y los resultados en la prueba de matemáticas.

El resultado en matemáticas fue obtenido mediante la evaluación de seis áreas temáticas. Fracciones y sentido numérico, medidas, proporcionalidad, representación de datos y análisis de probabilidades, geometría y álgebra. De estas preguntas en torno a un cuarto eran de respuesta libre y requerían del estudiante generar su propia respuesta. Este grupo de preguntas tenía asignado un tercio del tiempo total.

Dado que el hogar, la escuela y el contexto son variables claves para comprender el mundo educativo, el TIMSS recogió una extensa base de datos en torno a estas variables. Los alumnos que fueron evaluados en matemáticas completaron también cuestionarios acerca de su entorno socioeconómico y sus experiencias, opiniones y hobbies dentro y fuera del ámbito escolar, etc.. De igual manera, los profesores y los directores del colegio rellenaron cuestionarios sobre las prácticas educativas que llevaban a cabo⁴⁸.

Para la aplicación del modelo propuesto en el capítulo anterior sólo se han usado los datos de los alumnos españoles de 8º de EGB y el output en la prueba de matemáticas. El objetivo es tanto buscar una muestra más homogénea como ilustrar el modelo de la forma más sencilla, teniendo presente que el mismo podría ser igualmente aplicado con distintos outputs. Muchas de las variables originales fueron eliminadas por distintas razones. La principal razón para su eliminación fue que muchas de ellas presentaban valores perdidos para todas las unidades. Una segunda razón es que otras muchas variables carecían de interés para este estudio.

Las variables restantes fueron recodificadas casi en su totalidad. Tanto el nombre como los nuevos valores de las variables codificadas en este trabajo proceden de la base de datos original del TIMSS. Tras un análisis exploratorio, las variables que se consideran relevantes para el análisis son un total de 168 a nivel de alumno, 22 a nivel de escuela y 48 variables

⁴⁸ No es el objetivo de este trabajo realizar una extensa revisión de como se llevó a cabo la recogida de datos del TIMSS y de como se derivó toda la información contenida en el mismo. Para un estudio detallado puede acudir a los manuales informativos elaborados por el TIMSS, González et al. (1997).

relativas al profesor y a los métodos pedagógicos en el aula. En la elección de las mismas hemos tenido en cuenta tanto las que habitualmente son usadas en los trabajos de economía de la educación como aquellas otras relacionadas con el contexto educativo y especialmente las relacionadas con la inteligencia emocional a nivel de alumno. En los tres niveles considerados, muchas de las variables son en realidad categorías de una misma variable original. Por otro lado, se han eliminado aquellas variables que estaban correlacionadas con otra que medía una dimensión muy parecida, en particular con un coeficiente de correlación mayor a 0.4, manteniendo en cada caso para el análisis aquella que resultaba más ilustrativa. Todas ellas se encuentran recogidas en el ANEXO 1, para el nivel de alumno y en el ANEXO 2 para el nivel escolar y del profesor.

En cuanto al tratamiento de los datos perdidos se procedió de la siguiente forma. A nivel de alumno se eliminaron aquellas observaciones que en las 168 variables tenían más de 15 datos perdidos. En aquellos alumnos con menos de 15 datos perdidos, cuyo total acumulaba más del 95% de la muestra, y siguiendo a Little et al. (1997) se reemplazaron los valores perdidos por el valor medio de la muestra en las variables continuas y por el valor de la mediana en las variables categóricas. Así, de los 3855 alumnos originales quedaron 3700 para el análisis. La tabla 4.2 muestra las frecuencias acerca de cómo se distribuyen los datos perdidos en los 3855 datos originales.

Tabla 4.2 Distribución de los datos perdidos en las variables para el nivel de alumno.

Datos Perdidos	Frecuencia	Porcentaje	Porcentaje acumulado
0	37	0,960	0,960
1	1687	43,761	44,721
2	769	19,948	64,669
3	381	9,883	74,553
4	301	7,808	82,361
5	170	4,410	86,770
6	107	2,776	89,546
7	71	1,842	91,388
8	51	1,323	92,711
9	27	0,700	93,411
10	19	0,493	93,904
11	14	0,363	94,267
12	14	0,363	94,630
13	18	0,467	95,097
14	34	0,882	95,979
15 o Más	155	4,021	100
Total Válidos	3700		
Total	3855	100	

En cuanto a los datos escolares, de las 155 escuelas⁴⁹ originales, se eliminan a las escuelas 48, 58 y 110 por no tener datos en ninguna variable. Además se eliminan los colegios 22, 49, 54, 69 y 114 porque todos ellos contaban con más de 40 datos perdidos. La tabla 4.3 muestra las estadísticas descriptivas acerca de cómo se distribuyen los datos perdidos a lo largo de las 147 escuelas en las variables escolares y de profesor.

Tabla 4.3 Distribución de los datos perdidos en las variables escolares y de profesor.

Nº de Datos Perdidos	Frecuencia	Porcentaje	Porcentaje Acumulado
0	27	18,367	18,367
1	24	16,327	34,694
2	23	15,646	50,340
3	13	8,844	59,184
4	7	4,762	63,946
5	5	3,401	67,347
6	5	3,401	70,748
7	7	4,762	75,510
8	2	1,361	76,871
9	1	0,680	77,551
10	3	2,041	79,592
11	1	0,680	80,272
12	2	1,361	81,633
13	1	0,680	82,313
15	1	0,680	82,993
17	1	0,680	83,673
30	2	1,361	85,034
31	8	5,442	90,476
32	6	4,082	94,558
33	6	4,082	98,639
37	1	0,680	99,320
38	1	0,680	100,000
Total	147	100	-

En la tabla 4.3 se observa cómo la mayor parte de los datos tienen menos de 3 datos perdidos. Con el objetivo de tener más unidades para el estudio se procedió a incluir otras escuelas con muchos más datos perdidos. Éstos fueron reemplazados igual que en el caso de los datos de los alumnos sustituyendo por el valor de la media las variables continuas y por el de la mediana las variables categóricas.

⁴⁹ Aunque el TIMSS recoge 153 escuelas en dos de ellas, la 108 y la 131 de la base de datos original, fueron evaluadas dos aulas distintas con distintos profesores y que para el análisis hemos considerado como si fueran dos escuelas distintas. Es por eso por lo que de inicio tenemos 155 escuelas.

4.3. El Nivel de Alumno

Siguiendo el modelo propuesto en el capítulo anterior se va a aproximar la siguiente función, (ecuación 4.1):

$$A_{is} = F(B_{is}, IE_{is}) + \eta_{is} + \varepsilon_{is} \quad (4.1)$$

Donde i se refiere a estudiantes tomados de forma individual y s se refiere a colegios. A_{is} es el rendimiento alcanzado por el alumno, B_{is} representa el conjunto de inputs individuales y familiares presentes y acumulados que influyen en el rendimiento, IE_{is} hace referencia a la variable de inteligencia emocional del alumno. η_{is} es el efecto de las variables escolares y del profesor sobre el rendimiento del alumno i que deberá ser estimado, y ε_{is} recoge el ruido en los datos así como otros factores determinantes que también influyen en el rendimiento pero que no han sido observados. A través de (4.1) se obtiene una estimación \hat{A}_{is} del rendimiento para cada alumno.

4.3.1 Modelo 1: Variables Individuales

Para este propósito se comparan los resultados predictivos que obtienen un modelo de regresión lineal y uno de RNAs. En este primer modelo se introducen sólo las variables relacionadas con el alumno relativas a su contexto pero no las relacionadas con la inteligencia emocional y los hábitos. El objetivo en este punto es, por tanto, tratar de separar qué parte del resultado educativo es debido a las variables individuales exógenas que no dependen en ningún momento del propio alumno y qué parte aportan las variables de inteligencia emocional relacionadas con el contexto educativo. Así, en primer lugar estimaremos.

$$A_{is} = F(B_{is}) + \xi_{is} \quad (4.2)$$

Donde $\xi_{is} = IE_{is} + \eta_{is} + \varepsilon_{is}$ es el término que recoge todos los otros factores que no hacen referencia a las características del alumno. La tabla 4.4 recoge los resultados de la regresión lineal llevada a cabo para este primer grupo de 47 variables.

En primer lugar, se comprueba cómo el modelo es significativo a la hora de explicar el resultado medio en matemáticas de cada alumno. SEXO indica que los chicos obtienen mejores resultados que las chicas en matemáticas, lo cual se repite en las pruebas del TIMSS que se han hecho en otros países. La variable CASTELLANO indica que aquellos alumnos que hicieron la

prueba de matemáticas en castellano obtienen resultados significativamente mejores que en el resto de lenguas en las que se pasó el examen. En cuanto a los familiares que conviven con el alumno, se observa cómo el resultado es independiente de la composición familiar. Sólo aquellos alumnos que viven habitualmente con otros individuos en casa que no son familiares tendrían peores resultados.

Por otro lado, los recursos en el hogar aparecen como significativos para explicar el rendimiento en matemáticas. La posesión de más número de libros, ordenador y escritorio de estudio suponen mejores resultados que los alumnos que no poseen estos elementos. La explicación más plausible a este resultado es que estas variables son proxies de la renta familiar. Sin embargo, la posesión de libros específicos de matemáticas o de vídeos educativos aparecen negativamente correlacionados con el resultado académico. Este resultado que debe ser tomado con cautela puede responder a que recursos educativos extras en el hogar revelen en realidad que el alumno tiene mayores necesidades debidas a un bajo nivel.

La edad en años y meses se encuentra negativamente correlacionada con el resultado fundamentalmente por el efecto de los repetidores. En cuanto a los estudios de los padres observamos cómo padres con más estudios suponen mejores resultados académicos en sus hijos. En particular que la madre tenga estudios secundarios o universitarios terminados junto a estudios secundarios del padre supone una mejora estadísticamente significativa en la prueba de matemáticas. En cuanto al porcentaje de alumnos que trabaja, observamos cómo el compaginar estudios con trabajo a tan corta edad es perjudicial para el rendimiento académico.

Tabla 4.4: Coeficientes obtenidos por mínimos cuadrados ordinarios para las variables individuales del alumno.

	Coeficientes no estandarizados		Coeficientes estandarizados		
	B	Error típ.	Beta	t	Sig.
(Constante)	730,148	30,243		24,143	0,000*
ITSEX	11,036	2,296	0,075	4,806	0,000*
CASTELLANO	17,934	4,582	0,103	3,914	0,000*
GALLEGO	19,636	12,882	0,025	1,524	0,128
CATALAN	2,071	5,957	0,010	0,348	0,728
NACESPANA	-10,272	7,229	-0,022	-1,421	0,155
LENGUACASA	4,279	3,649	0,023	1,173	0,241
MADRE	0,842	7,055	0,002	0,119	0,905
PADRE	5,472	3,910	0,023	1,399	0,162
HERMANOS	3,111	2,549	0,020	1,221	0,222
HERMANAS	3,471	2,486	0,023	1,396	0,163
MADRASTRA	-0,927	11,890	-0,001	-0,078	0,938
PADRASTRO	12,910	8,550	0,024	1,510	0,131
ABUELO	1,647	2,861	0,010	0,576	0,565
PARIENTES	5,095	3,727	0,024	1,367	0,172
OTROS	7,729	4,103	0,031	1,884	0,060**
NUMTOTAL	-0,640	0,971	-0,012	-0,659	0,510
MADRESP	5,670	5,133	0,024	1,105	0,269
PADRESP	-1,219	5,308	-0,005	-0,230	0,818
LIBROSCASA	12,250	1,147	0,190	10,681	0,000*
CALCULADORA	1,160	9,980	0,002	0,116	0,907
ORDENADOR	-6,027	3,004	-0,040	-2,007	0,045*
ESCRITORIO	-12,215	4,921	-0,039	-2,482	0,013*
DICCIONARIO	4,330	15,741	0,004	0,275	0,783
ENCICLOPEDIA	-5,383	4,000	-0,022	-1,346	0,178
LIBRMATES	10,592	3,189	0,068	3,321	0,001*
LIBRCIENCIAS	-5,243	3,343	-0,033	-1,568	0,117
LABINSTRUM	-0,926	2,585	-0,006	-0,358	0,720
VIDEOSEDUC	6,233	2,518	0,040	2,475	0,013*
PROGORDEDUC	-3,571	3,379	-0,021	-1,057	0,291
LUGESTUDPROP	-0,635	3,985	-0,003	-0,159	0,873
EDAD	-23,541	1,733	-0,214	-13,585	0,000*
PRIMMAD	3,060	4,144	0,021	0,738	0,460
ALSECMAD	6,790	5,139	0,029	1,321	0,187
SECMAD	14,209	5,738	0,050	2,477	0,013*
VOCACMAD	9,960	6,734	0,027	1,479	0,139
ALUNIMAD	10,307	8,119	0,022	1,270	0,204
UNIMAD	15,250	5,746	0,061	2,654	0,008*
PRIMPAD	1,948	4,027	0,013	0,484	0,629
ALSECPAD	8,235	4,877	0,036	1,689	0,091**
SECPAD	10,634	5,277	0,040	2,015	0,044*
VOCACPAD	0,178	6,118	0,001	0,029	0,977
ALUNIPAD	8,656	7,673	0,019	1,128	0,259
UNIPAD	6,758	5,201	0,030	1,299	0,194
PAIDM1	-10,753	5,477	-0,030	-1,963	0,050*
PAID1Y2	-6,982	4,533	-0,024	-1,540	0,124
PAID3Y5	11,133	7,020	0,024	1,586	0,113
PAIDMAS5	-19,655	6,423	-0,047	-3,060	0,002*

* La variable es significativa al 95%,

**La variable es significativa al 90%

Resumen del Modelo 1 (47 variables)

Coefficiente de Correlación	R Cuadrado	R Cuadrado Corregida	Error Típico de la Estimación
,413	,170	,160	67,7972

Análisis de Varianza del Modelo 1 (47 variables)

Modelo	Suma de Cuadrados	Grados de Libertad	Media Cuadrática	F	Sig.
Regresión	3442941,477	47	73254,074	15,937	,000
Residual	16786289,359	3652	4596,465		
Total	20229230,836	3699			

* El modelo de regresión múltiple es estadísticamente significativo al 99%.

A continuación se procede a aproximar el mismo modelo con RNAs⁵⁰. Para ello se divide aleatoriamente la muestra en tres submuestras, de entrenamiento, de validación y de test⁵¹ y se realiza una búsqueda del MLP óptimo con las muestras de entrenamiento y validación tal y como se explica en el capítulo 1, (apartado, 1.5.3.1) Tras realizar una búsqueda exhaustiva ensayando con diversas arquitecturas y parámetros el MLP final tiene las siguientes características:

Nº Neuronas en la Capa Inicial (INPUTS): 47

Nº de Capas Ocultas: 1

Nº Neuronas en la Capa Oculta: 3

Nº Neuronas en la Capa de Salida (OUTPUT): 1

Coefficiente de Aprendizaje: 0,1

Momento: 0,1

Funciones de Transferencia: Sigmoide Logística en la capa del medio y de salida y función lineal en el intervalo ± 1 en la de entrada.

Valor de los pesos en la Inicialización Aleatoria: $\pm 0,3$

Para comparar la capacidad de predicción de los dos modelos se usan los 3 indicadores que ilustran a continuación, que además son los más utilizados a lo largo de la literatura económica. Junto a ellos, se calcula el coeficiente de correlación y de determinación para comparar los distintos modelos predictivos:

⁵⁰ Para ello hemos usado el software Neuroshell 2.0 con el algoritmo *backpropagation*.

⁵¹ En este trabajo de tesis todas las muestras han sido particionadas asignando un 80% de los datos a entrenamiento, un 10% a validación y el 10% restante a testear el modelo final.

- $RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (\rho_i - \hat{\rho}_i)^2}$: es la suma media de la diferencia entre el valor real y el ajustado (dentro de la muestra) o predicho (fuera de la muestra) por el modelo.
- $MAE = \frac{1}{T} \sum_{i=1}^T |\rho_i - \hat{\rho}_i|$: es el error absoluto medio entre el valor real y el ajustado (dentro de la muestra) o predicho (fuera de la muestra) por el modelo.
- $MAPE = \frac{1}{T} \sum_{i=1}^T \left| \frac{\rho_i - \hat{\rho}_i}{\rho_i} \right|$: es el error absoluto medio en porcentaje entre el valor real y el ajustado (dentro de la muestra) o predicho (fuera de la muestra) por el modelo.

Tabla 4.5. Comparación de la capacidad predictiva de los modelos de regresión múltiple y de RNAs para las variables exógenas. 47 variables.

	Unidades	CORR	R ²	RMSE	MAE	MAPE	MIN	MAX
Total	3700	0,4150	0,1722	67,3616	53,8039	0,1117	0,07	262,31
Entrenamiento	2960	0,4173	0,1741	66,7123	53,1800	0,1102	0,07	262,31
Validación	370	0,3778	0,1427	70,9725	56,0613	0,1167	0,319	257,512
Test	370	0,4359	0,19	68,8075	56,5377	0,1190	0,891	176,436
OLS	3700	0,4130	0,17	67,3560	53,9140	0,1127	0,07355	260,8529

Los resultados, tabla 4.5, muestran que ambos modelos consiguen unos resultados muy similares. Si bien el MLP alcanza en los datos test un mayor ajuste que el modelo de regresión en el coeficiente de determinación, el RMSE y el MAE alcanza valores más pequeños para el modelo de regresión simple.

A partir de los resultados obtenidos en la regresión podemos inferir que muchas de las variables utilizadas para entrenar el MLP no son relevantes y por tanto deterioran la búsqueda del mínimo en la superficie del error, aumentando las posibilidades de caer atrapados en mínimos locales como consecuencia del ruido que introducen las variables que no son relevantes. Para comprobar si esta posibilidad se está produciendo en nuestro análisis, procedemos a eliminar aquellas variables con un p-valor mayor a 0,5 en la regresión lineal con el fin de intentar conseguir que la red neuronal aproxime mejor el modelo sin renunciar a la búsqueda de posibles no linealidades en variables que no son significativas linealmente. El MLP óptimo presenta las siguientes características.

La red neuronal final es del tipo perceptrón multicapa con una sola capa oculta. Sus características son:

Nº Neuronas en la Capa Inicial (INPUTS): 35

Nº de Capas Ocultas: 1

Nº Neuronas en la Capa Oculta: 25

Nº Neuronas en la Capa de Salida (OUTPUT): 1

Coefficiente de Aprendizaje: 0,1

Momento: 0,1

Funciones de Transferencia: Sigmoide Logística en la capa del medio y de salida, lineal en el intervalo ± 1 en la de entrada.

Valor de los pesos en la Inicialización Aleatoria: $\pm 0,3$

Tabla 4.6. Comparación de la capacidad predictiva de los modelos de regresión múltiple y de redes neuronales artificiales para las 35 variables exógenas con un p-valor en la regresión lineal menor a 0,5.

	Unidades	CORR	R ²	RMSE	MAE	MAPE	MIN	MAX
Total	3700	0,4462	0,1991	66,2688	53,0884	0,1112	0,0290	263,8780
Entrenamiento	2960	0,4533	0,2055	65,4086	52,2832	0,1094	0,0290	254,9110
Validación	370	0,3957	0,1566	70,5447	56,1357	0,1179	0,0410	263,8780
Test	370	0,4421	0,1954	68,6492	56,4832	0,1197	0,1370	190,5080
OLS	3700	0,4130	0,1700	67,3560	53,9140	0,1127	0,07355	260,8529

La tabla 4.6 muestra unos resultados muy similares a los de la tabla 4.5. Sin embargo se puede comprobar cómo, a pesar de haber sacado del análisis 12 variables, los resultados del MLP mejoran levemente con un coeficiente de correlación en la muestra de test superior a la del caso anterior.

De esta forma se obtiene un modelo capaz de explicar en torno al 20% de la varianza del alumno a partir únicamente de sus características personales. Por otro lado, el ajuste similar del modelo lineal y no lineal hace pensar que a este nivel no existen fuertes no linealidades en el modelo y por tanto los coeficientes que hemos obtenido con el modelo de regresión lineal tendrían una interpretación válida y más robusta por no estar sesgada por la forma funcional o especificación del modelo.

4.3.2. Modelo 2: Variables Individuales y de Inteligencia Emocional

A continuación se procede a introducir las variables de inteligencia emocional y de hábitos del alumno relacionados con la escuela. Si bien estas variables tienen cierto grado de endogeneidad pueden ser conocidas a priori y por tanto podrían ser usadas para predecir a principio de curso el rendimiento del alumno en función de sus características personales. Tanto padres como escuela podrían en cierto grado influir en estas características mediante recomendaciones para mejorar el resultado del alumno.

La tabla 4.7 muestra los resultados de la regresión lineal múltiple para las variables individuales introducidas en el modelo 1 junto a las de inteligencia emocional y hábitos del alumno.

Tabla 4.7. Coeficientes obtenidos en la regresión múltiple para las 168 posibles variables explicativas seleccionadas en el modelo a nivel de alumno.

Variables	Coeficientes no estandarizados		Coeficientes estandarizados		
	Beta	Error típ.	Beta	t	Sig.
(Constante)	576,361	57,442		10,034	,000*
SEXO	16,054	2,595	,108	6,186	,000*
CASTELLANO	14,311	4,336	,083	3,300	,001*
GALLEGO	18,154	12,051	,023	1,506	,132
CATALAN	,265	5,606	,001	,047	,962†
NACESPAÑA	-9,896	6,713	-,022	-1,474	,141
LENGUACASA	1,432	3,423	,008	,418	,676†
MADRE	-3,640	6,588	-,009	-,552	,581†
PADRE	8,696	3,660	,037	2,376	,018*
HERMANOS	1,183	2,369	,008	,499	,618
HERMANAS	4,600	2,302	,030	1,998	,046*
MADRASTRA	-9,980	11,073	-,014	-,901	,368
PADRASTRO	11,911	7,913	,022	1,505	,132
ABUELO	,0285	2,659	,000	,011	,991†
PARIENTES	5,574	3,468	,026	1,607	,108
OTROS	5,950	3,833	,024	1,552	,121
NUMTOTAL	-,0065	,905	,000	-,007	,994†
MADRESP	8,020	4,751	,034	1,688	,091**
PADRESP	,202	4,922	,001	,041	,967†
LIBROSCASA	8,795	1,090	,136	8,070	,000*
CALCULADORA	2,909	10,196	,004	,285	,775†
ORDENADOR	-6,644	10,017	-,045	-,663	,507
ESCRITORIO	-14,852	6,675	-,048	-2,225	,026*
DICCIONARIO	10,132	15,326	,010	,661	,509
ENCICLOPEDIA	-5,898	5,429	-,024	-1,086	,277
LIBRMATES	2,621	4,885	,017	,536	,592†
LIBRCIENCIAS	-6,768	5,010	-,042	-1,351	,177
LABINSTRUM	-1,293	4,710	-,008	-,274	,784†
VIDEOSEDUC	-,268	4,584	-,002	-,058	,953†
PROGORDEDUC	-7,418	5,116	-,044	-1,450	,147
LUGESTUDPROP	-,636	5,457	-,003	-,117	,907†
MADMEJCLASE	3,332	3,681	,015	,905	,365
NOTRABAJAN	-1,329	2,220	-,009	-,599	,550
SEPORTANBIEN	2,048	2,404	,014	,852	,395
OBEDECEN	-,796	2,430	-,005	-,328	,743†

(Continuación)

AMIGMAT	8,213	6,338	,031	1,296	,195
AMIGLENG	-15,547	6,045	-,060	-2,572	,010*
AMIGDIVERS	9,588	11,374	,013	,843	,399
AMIGMEJCLASE	-6,460	2,934	-,039	-2,202	,028*
OPBUENOMAT	22,069	11,250	,030	1,962	,050**
OPBUENODIVERS	-,556	13,532	-,001	-,041	,967†
OPBUENOSPORT	-1,115	4,933	-,003	-,226	,821†
OPBUENOCLASE	-,915	3,661	-,005	-,250	,803†
BUENOMATES	21,946	2,735	,133	8,026	,000*
GENETICO	4,170	2,299	,027	1,813	,070**
SUERTE	-8,586	2,339	-,055	-3,671	,000*
ESTDURO	-13,026	3,447	-,056	-3,779	,000*
MEMORIZAR	-8,733	2,257	-,058	-3,870	,000*
GUSTANMATES	11,173	2,815	,073	3,969	,000*
ORDCLMAT	-11,486	2,786	-,060	-4,123	,000*
DIVERTAPMAT	-4,147	2,469	-,026	-1,680	,093**
ABURRAPMAT	-3,095	2,312	-,021	-1,339	,181
MATESFACILES	3,279	2,446	,021	1,341	,180
MATESIMPVIDA	-2,581	4,904	-,007	-,526	,599†
LIKETRABMAT	3,014	2,335	,020	1,291	,197
OPMATIMPJOB	-,841	2,789	-,005	-,302	,763†
BIENMATPADR	-3,117	3,076	-,015	-1,013	,311
BIENMATINSTI	4,126	3,820	,017	1,080	,280
BIENMATSELF	3,043	5,022	,009	,606	,545
SCIAIRPOLUT	1,810	4,479	,009	,404	,686†
SCIAGUAPOLUT	6,627	4,953	,032	1,338	,181
SCIDESTBOSQU	-1,077	3,874	-,006	-,278	,781†
SCIANIMEXTIN	-2,621	3,592	-,014	-,730	,466
SCIOZONO	5,645	3,899	,029	1,448	,148
SCIPOWNUCPL	,0997	3,116	,001	,032	,974†
FRECLEOLIBRO	,825	2,672	,005	,309	,758†
FRECISMUSEO	-15,277	6,764	-,036	-2,259	,024*
FRECCONCIER	6,119	6,235	,015	,981	,326
FRECTEATRO	4,995	7,574	,010	,659	,510
FRECCINE	-4,475	2,631	-,025	-1,701	,089**
FRECNOTICIAS	2,379	2,671	,014	,891	,373
FRECOPERA	-4,587	3,902	-,017	-1,175	,240
FRECATURE	-2,618	2,363	-,017	-1,108	,268
FRECPOPMUSIC	,0891	2,161	,001	,041	,967†
FRECSPORT	1,219	2,885	,007	,422	,673†
FRECVIDEOJUE	-3,962	2,676	-,025	-1,480	,139
FRECARTOONS	-,149	2,370	-,001	-,063	,950†
FRECAVCOSU	14,971	3,672	,061	4,077	,000*
EDAD	-13,010	1,721	-,118	-7,558	,000*
NUMEROITEMS	-2,759	4,030	-,070	-,685	,494
DESKCOMPDIC	-1,076	9,434	-,007	-,114	,909†
PRIMMAD	1,648	3,855	,011	,428	,669†
ALSECMAD	2,955	4,802	,013	,615	,538
SECMAD	10,079	5,353	,036	1,883	,060**
VOCACMAD	4,449	6,269	,012	,710	,478
ALUNIMAD	6,427	7,570	,014	,849	,396
UNIMAD	11,731	5,372	,047	2,184	,029*
PRIMPAD	1,260	3,742	,008	,337	,736†
ALSECPAD	7,466	4,522	,032	1,651	,099**
SECPAD	5,344	4,899	,020	1,091	,275
VOCACPAD	4,458	5,683	,014	,784	,433
ALUNIPAD	8,727	7,128	,020	1,224	,221
UNIPAD	1,861	4,838	,008	,385	,701†
PRIMSELF	-10,089	4,645	-,038	-2,172	,054**
ALCESEL	-10,533	5,456	-,031	-1,931	,030*
SECSELF	-6,063	4,866	-,021	-1,246	,213
VOCACSEL	-7,754	4,605	-,029	-1,684	,092**
ALUNISEL	-6,408	5,874	-,017	-1,091	,275
UNISELF	12,247	3,354	,083	3,651	,000*
EXTMATM1	-13,433	4,525	-,058	-2,968	,003*

(Continuación)

EXTMAT1Y2	-18,439	4,037	-,090	-4,568	,000*
EXTMAT3Y5	2,487	8,659	,005	,287	,774†
EXTMATMAS5	23,587	18,916	,023	1,247	,212
EXTSCIM1	-4,196	4,893	-,017	-,858	,391
EXTSCI1Y2	3,634	5,001	,014	,727	,467
EXTSCI3Y5	-10,005	11,662	-,015	-,858	,391
EXTSCIMAS5	-9,902	23,783	-,008	-,416	,677†
SCMATCLUBM1	-4,598	4,456	-,016	-1,032	,302
SCMATCLUB1Y2	-,749	4,895	-,002	-,153	,878†
SCMATCLUB3Y5	12,323	13,944	,013	,884	,377
CLUBMAS5	-12,504	19,351	-,010	-,646	,518
PAIDM1	-,686	5,167	-,002	-,133	,894†
PAID1Y2	1,546	4,270	,005	,362	,717†
PAID3Y5	15,968	6,585	,035	2,425	,015*
PAIDMAS5	-2,027	6,188	-,005	-,328	,743†
TVM1	2,828	5,319	,017	,532	,595
TV1Y2	8,988	5,279	,061	1,703	,089**
TV3Y5	12,986	5,721	,065	2,270	,023*
TVMAS5	5,948	7,525	,016	,790	,429
NINTM1	1,091	2,939	,006	,371	,710†
NINT1Y2	1,644	4,595	,006	,358	,721†
NINT3Y5	-26,262	9,271	-,042	-2,833	,005*
NINTMAS5	-3,806	12,027	-,005	-,316	,752†
AMIGM1	7,658	4,069	,050	1,882	,060**
AMIG1Y2	1,518	4,271	,009	,355	,722†
AMIG3Y5	-,414	4,780	-,002	-,087	,931†
AMIGMAS5	-5,250	5,722	-,019	-,917	,359
DOMESTM1	1,879	3,148	,013	,597	,551
DOMEST1Y2	-7,805	3,353	-,049	-2,328	,020*
DOMEST3Y5	-3,239	4,727	-,012	-,685	,493
DOMESTMAS5	-5,488	8,913	-,009	-,616	,538
SPORTM1	7,124	3,450	,041	2,065	,039*
SPORT1Y2	,764	3,283	,005	,233	,816†
SPORT3Y5	3,431	4,253	,016	,807	,420
SPORTMAS5	-7,570	5,747	-,023	-1,317	,188
DIVERLIBRM1	-,596	2,585	-,004	-,231	,818†
DIVERLIBR1Y2	3,915	3,386	,020	1,156	,248
DIVERLIBR3Y5	-,942	6,994	-,002	-,135	,893†
DIVERLIBRMAS5	-15,020	12,326	-,018	-1,219	,223
ESTMATM1	-3,814	7,406	-,026	-,515	,607†
ESTMAT1Y2	-12,758	7,680	-,086	-1,661	,097**
ESTMAT3Y5	-21,369	9,248	-,065	-2,311	,021*
ESTMATMAS5	-34,525	20,482	-,035	-1,686	,092**
ESTSCIM1	23,013	6,238	,156	3,689	,000*
ESTSCI1Y2	19,920	6,634	,133	3,003	,003*
ESTSCI3Y5	15,367	9,092	,040	1,690	,091**
ESTSCIMAS5	45,420	23,487	,040	1,934	,053**
ESTOTRASM1	-10,158	6,813	-,065	-1,491	,136
ESTOTRAS1Y2	-3,916	6,891	-,026	-,568	,570
ESTOTRAS3Y5	-,428	7,688	-,002	-,056	,956†
ESTOTRMAS5	-24,846	13,434	-,034	-1,850	,064**
SKIPCLASS12	-6,702	4,144	-,024	-1,617	,106
SKIPCLASS34	,638	11,054	,001	,058	,954†
SKIPCLASSM5	-4,540	8,522	-,008	-,533	,594†
ROBA12	-5,070	2,871	-,027	-1,766	,077**
ROBA34	-8,740	7,902	-,016	-1,106	,269
ROBAM5	,907	8,937	,002	,101	,919†
AMENAZA12	2,773	2,802	,015	,990	,322
AMENAZA34	4,207	8,030	,007	,524	,600†
AMENAZAM5	-5,873	6,689	-,013	-,878	,380
ASKIPCLASS12	6,114	2,527	,038	2,420	,016*
ASKIPCLASS34	5,101	4,110	,019	1,241	,215
ASKIPCLASSM5	3,163	3,520	,015	,899	,369
AROBA12	-2,700	2,623	-,016	-1,029	,303
AROBA34	-3,527	5,328	-,010	-,662	,508

(Continuación)

AROBAM5	-4,005	6,113	-,011	-,655	,512
ADAÑO12	2,587	2,446	,016	1,058	,290
ADAÑO34	5,875	4,980	,018	1,180	,238
ADAÑOM5	-,742	5,149	-,002	-,144	,885†

a Variable dependiente: BIMATSCR o resultado en la prueba de matemáticas.

* Variable Estadísticamente Significativa al 95%.

** Variable Estadísticamente Significativa al 90%.

Análisis de Varianza del Modelo 2 de Regresión Múltiple

Modelo	Suma de Cuadrados	Grados de Libertad	Media Cuadrática	F	Sig.
Regresión	6.711.897,63	168	39.951,772	10,436	,000*
Residual	13.517.333,2	3.532	3.828,188		
Total	20.229.230,8	3.699			

* El modelo de regresión múltiple es estadísticamente significativo al 99%.

Resumen del Modelo 2

Coefficiente de Correlación	R Cuadrado	R Cuadrado Corregida	Error Típico de la Estimación
,576	,332	,300	61,8724

El modelo es por tanto significativo a la hora de explicar el rendimiento de los alumnos. Añadir las variables relacionadas con el contexto escolar y la inteligencia emocional hace que ahora podamos explicar en torno al 33% de la varianza con las variables individuales, esto supone prácticamente el doble de lo que explicábamos anteriormente. La tabla 4.8 muestra un resumen de las variables que resultaron ser significativas.

Tabla 4.8. Variables significativas a la hora de explicar el rendimiento de los alumnos de 8º nivel en matemáticas

Variables	Coeficientes no estandarizados		Coeficientes estandarizados		
	Beta	Error típ.	Beta	t	Sig.
(Constante)	576,361	57,442		10,034	,000*
SEXO	16,054	2,595	,108	6,186	,000*
CASTELLANO	14,311	4,336	,083	3,300	,001*
PADRE	8,696	3,660	,037	2,376	,018*
HERMANAS	4,600	2,302	,030	1,998	,046*
MADRESP	8,020	4,751	,034	1,688	,091**
LIBROSCASA	8,795	1,090	,136	8,070	,000*
ESCRITORIO	-14,852	6,675	-,048	-2,225	,026*
AMIGLENG	-15,547	6,045	-,060	-2,572	,010*
AMIGMEICLASE	-6,460	2,934	-,039	-2,202	,028*
OPBUENOMAT	22,069	11,250	,030	1,962	,050**
BUENOMATES	21,946	2,735	,133	8,026	,000*
GENETICO	4,170	2,299	,027	1,813	,070**
SUERTE	-8,586	2,339	-,055	-3,671	,000*
ESTDURO	-13,026	3,447	-,056	-3,779	,000*
MEMORIZAR	-8,733	2,257	-,058	-3,870	,000*
GUSTANMATES	11,173	2,815	,073	3,969	,000*
ORDCLMAT	-11,486	2,786	-,060	-4,123	,000*
DIVERTAPMAT	-4,147	2,469	-,026	-1,680	,093**
FRECVMUSEO	-15,277	6,764	-,036	-2,259	,024*
FRECCINE	-4,475	2,631	-,025	-1,701	,089**
FRECAVCOSU	14,971	3,672	,061	4,077	,000*
EDAD	-13,010	1,721	-,118	-7,558	,000*
SECMAD	10,079	5,353	,036	1,883	,060**
UNIMAD	11,731	5,372	,047	2,184	,029*
ALSECPAD	7,466	4,522	,032	1,651	,099**
PRIMSELF	-10,089	4,645	-,038	-2,172	,030*
ALSECSELF	-10,533	5,456	-,031	-1,931	,054**
VOCACSELF	-7,754	4,605	-,029	-1,684	,092**
UNISELF	12,247	3,354	,083	3,651	,000*
EXTMATM1	-13,433	4,525	-,058	-2,968	,003*
EXTMATIY2	-18,439	4,037	-,090	-4,568	,000*
PAID3Y5	15,968	6,585	,035	2,425	,015*
TV1Y2	8,988	5,279	,061	1,703	,089**
TV3Y5	12,986	5,721	,065	2,270	,023*
NINT3Y5	-26,262	9,271	-,042	-2,833	,005*
AMIGM1	7,658	4,069	,050	1,882	,060**
DOMESTIY2	-7,805	3,353	-,049	-2,328	,020*
SPORTM1	7,124	3,450	,041	2,065	,039*
ESTMATIY2	-12,758	7,680	-,086	-1,661	,097**
ESTMAT3Y5	-21,369	9,248	-,065	-2,311	,021*
ESTMATMAS5	-34,525	20,482	-,035	-1,686	,092**
ESTSCIM1	23,013	6,238	,156	3,689	,000*
ESTSCIY2	19,920	6,634	,133	3,003	,003*
ESTSCI3Y5	15,367	9,092	,040	1,690	,091**
ESTSCIMAS5	45,420	23,487	,040	1,934	,053**
ESTOTRASMAS5	-24,846	13,434	-,034	-1,850	,064**
ROBA12	-5,070	2,871	-,027	-1,766	,077**
ASKIPCLASS12	6,114	2,527	,038	2,420	,016*

Variable dependiente: BIMATSCR

* Variable Estadísticamente Significativa al 95%.

** Variable Estadísticamente Significativa al 90%.

A continuación se descubren los efectos de distintos grupos de variables sobre el rendimiento en la prueba de matemáticas y en particular aquellas relacionadas con la inteligencia emocional.

INDIVIDUALES

Además de las variables comentadas con anterioridad referentes a las características familiares y de idioma utilizado, se observa cómo las variables PADRE y HERMANAS indican que los alumnos que no viven habitualmente con su padre ni con hermanas obtienen mejores resultados. MADRESP indica que los alumnos cuya madre no era de nacionalidad española obtuvieron resultados ligeramente mejores. ORDCLMAT indica que los alumnos que alguna vez usan ordenador en la escuela para estudiar matemáticas no obtienen resultados mejores que sus compañeros que no lo usan e incluso obtendrían resultados peores.

COMPAÑEROS

Por un lado las variables AMIGLENG y AMIGMEJCLASE obtienen un signo negativo es decir aquellos alumnos cuyos amigos no valoran como muy importante para ellos los resultados en lengua, o no dan importancia a estar entre los mejores de la clase, obtienen mejores resultados. Quizás detrás de este comportamiento se encuentre que es más positivo para el alumno que sus amigos valoren de él otras virtudes y características aparte de la de estar entre los que mejores resultados escolares obtienen. Incluso podría llegar a estar “mal visto”, por los alumnos de 13 años, pertenecer al colectivo de los mejores de la clase. Por otro lado AMIGM1 indica que hablar o jugar con amigos durante menos de una hora fuera del horario escolar es positivo para el resultado. Esta actividad puede ser realizada simultáneamente practicando otras actividades tales como deporte, ir al cine u otras.

OPINIONES

Las opiniones del alumno en torno a las matemáticas son variables muy predictivas de su resultado. Estas variables presentan un alto grado de endogeneidad que hace que su capacidad explicativa del resultado quede sesgada. Así, se debe ser cauteloso a la hora de interpretar este resultado, ya que además estas variables es muy posible que interaccionen con otras del modelo. En cualquier caso, las variables de opinión y actitud hacia las matemáticas podrían ser conocidas ex-ante al proceso educativo y ser utilizadas para realizar predicciones y ayudar al alumno a mejorar su percepción e interés en las mismas.

OPBUENOMAT, BUENOMATES y GUSTANMATES están fuertemente correlacionadas de forma positiva con un mejor rendimiento en matemáticas. Es decir, aquellos alumnos que consideran que las matemáticas serán importantes para su vida, les gusta la materia y que además se consideran buenos obtienen mejores resultados. Sin embargo DIVERAPMAT indica que existe una correlación negativa entre considerar que aprender matemáticas es divertido y los buenos resultados.

GENETICO, SUERTE, ESTDURO y MEMORIZAR hacen referencia a los factores que el alumno considera que sirven para aumentar el resultado en esta prueba. Los que piensan que el factor genético tiene importancia obtienen resultados mejores. Por otro lado los alumnos que consideran que los buenos resultados en matemáticas son debidos a factores como la suerte, estudiar duro o memorizar apuntes obtienen peores resultados que los que opinan lo contrario.

Resultan muy interesantes las variables relacionadas con las expectativas en la escuela. PRIMSELF, ALSECSELF y VOCACSELF indican frente a UNISELF que los alumnos que no tienen expectativas de terminar estudios universitarios obtienen peores resultados que los alumnos que declaran que sus expectativas son terminar la educación superior universitaria.

A pesar de su componente endógeno, ya que alumnos con mejores resultados mejoran sus expectativas, podría ser interesante llevar a cabo políticas que conciencien tanto a los alumnos como a los padres de la importancia de obtener buenos resultados en la escuela no sólo de cara al acceso a la universidad sino también para terminar con éxito cualquier tipo de estudios o formación profesional.

HÁBITOS FUERA DE LA ESCUELA

Las variables FRECVISMUSEO y FRECCINE indican que los alumnos que frecuentemente van al cine o a visitar museos, obtienen peores resultados en matemáticas. Por otro lado TV1Y2 y TV3Y5, indican que ver televisión a diario entre 1 y 2 horas o entre 3 y 5 horas es positivo para el rendimiento escolar, resultado que como ya hemos comentado anteriormente también se da en otros países, si bien en España el margen de horario se amplía más. Por otro lado, la variable FRECAVCOSU indica además que aquellos alumnos que frecuentemente ven películas en la televisión o el vídeo de aventuras, comedia y suspense obtienen mejores resultados.

SPORTM1 indica que practicar deporte durante 1 hora al día es positivo para la mejora del rendimiento en la prueba de matemáticas. Un resultado extraño es el de la variable PAID3Y5 que indica que el 2,65% de alumnos que llevaban a cabo un trabajo remunerado de tan larga duración obtenían buenos resultados. La explicación para este resultado es sin duda alguna controvertida y se requeriría más análisis de los datos para la explicación del mismo.

EXTMATM1 y EXTMAT1Y2 indica que aquellos alumnos que han recibido clases extraescolares de matemáticas obtienen peores resultados, lo cual es fácilmente explicable debido a que son los alumnos que van peor en la asignatura a los que se asigna horas extras dedicadas a las matemáticas.

HABITOS EN EL HOGAR

Una variable que requiere especial atención por lo extendido de su uso en los últimos tiempos es NINT3Y5. Esto es, jugar a videojuegos, consolas u ordenador más de 3 horas resulta muy perjudicial para el rendimiento en matemáticas, ya que llega a restar 26 puntos en la prueba de rendimiento. DOMEST1Y2 indica que realizar trabajos domésticos, tales como fregar los platos, hacer las camas o limpieza durante más de 1 hora diaria es negativo. Sin embargo tal y como muestra la tabla 4.5 el coeficiente de DOMESTM1 es positivo aunque no estadísticamente significativo, es decir, ayudar en las tareas domésticas un rato cada día sí sería positivo para la formación del alumno.

Las variables relacionadas con los hábitos de estudio presentan de nuevo un comportamiento un tanto extraño con lo que en principio cabría esperar. El tiempo dedicado a diario a estudiar matemáticas, ESTMAT1Y2, ESTMAT3Y5, ESTMATMAS5 al igual que estudiar otras asignaturas distintas de matemáticas y ciencias ESTOTRASMAS5, está negativamente correlacionado con el rendimiento en las mismas. Sin embargo estudiar ciencias ESTSCIM1, ESTSCI1Y2, ESTSCI3Y5 y ESTSCIMAS5 sí está correlacionado positivamente con el resultado en matemáticas. Quizás una explicación plausible a este fenómeno sería que los alumnos que van muy bien en matemáticas dedican más tiempo a estudiar otra asignatura en principio difícil como las ciencias, frente a los que van peor en matemáticas que dejarían poco tiempo para otras asignaturas.

COMPORTAMIENTOS CONFLICTIVOS

De los comportamientos relacionados con conflictos, la variable ROBA12, indica que los alumnos a los que durante el mes anterior a la encuesta del TIMSS les robaron una o dos cosas obtienen resultados ligeramente peores como consecuencia de encontrarse rodeados de un ambiente negativo. Sin embargo ASKIPCLASS12 indica que tener conocimiento de que determinados amigos se han saltado alguna clase durante el último mes está correlacionado positivamente con mejores resultados. La explicación a esta variable puede estar relacionada con que se trata de un alumno atento y pendiente del comportamiento de los demás.

Algunos de los coeficientes obtenidos pueden resultar extraños y difícilmente interpretables. Una de las posibles explicaciones a estas rarezas es que dada la cantidad de variables introducidas puede que alguna de ellas interactúe de forma no lineal con otras y que por tanto la especificación lineal no sea la más plausible y arroje así resultados engañosos. Una posibilidad sería comprobar en un modelo de regresión no lineal muchas de estas posibles combinaciones. Sin embargo, el desconocimiento del proceso educativo y el alto número de variables implicadas aconsejan un análisis exploratorio no lineal más potente. Es por ello que para saber si existen este tipo de no linealidades en los datos procedemos a aplicar RNAs con el fin de aproximar los inputs al resultado en matemáticas.

Igual que hicimos en el modelo 1, hemos suprimido para el análisis con RNAs aquellas variables que resultaron no tener ninguna importancia sobre el resultado en el modelo lineal y que sólo contribuirían, por tanto, a introducir ruido y deteriorar el proceso de minimización del algoritmo de retropropagación supervisada. Las variables suprimidas son aquellas que están marcadas con el símbolo † en la tabla 4.5 y que tienen un p-valor mayor que 0,6. En total utilizaremos 120 inputs para el análisis. Después de realizar la búsqueda del MLP óptimo, la RNA final elegida es de nuevo un MLP con las siguientes características.

Nº Neuronas en la Capa Inicial (INPUTS): 120

Nº de Capas Ocultas: 1

Nº Neuronas en la Capa Oculta: 25

Nº Neuronas en la Capa de Salida (OUTPUT): 1

Coefficiente de Aprendizaje: 0,1

Momento: 0,4

Funciones de Transferencia: Sigmoide Logística en la capa del medio y de salida, lineal en el intervalo ± 1 en la de entrada.

Valor de los pesos en la Inicialización Aleatoria: $\pm 0,3$

Tabla 4.9: Comparación de la capacidad predictiva de los modelos de regresión múltiple y de redes neuronales artificiales para las variables exógenas y de inteligencia emocional.

	Unidades	CORR	R ²	RMSE	MAE	MAPE	MIN	MAX
Total	3700	0,597	0,357	59,359	46,928	0,097	-190,347	229,595
Entrenamiento	2960	0,592	0,350	59,154	46,714	0,097	-190,347	226,595
Validación	370	0,587	0,345	62,061	49,564	0,103	-150,814	187,396
Test	370	0,656	0,431	57,943	46,088	0,097	-157,474	192,953
OLS	3700	0,576	0,332	60,423	47,953	0,100	-204,608	229,922

La tabla 4.9 muestra los resultados obtenidos con el MLP. Así, se puede comprobar que después del entrenamiento de la red neuronal, el ajuste predictivo obtenido por los datos test, nunca antes vistos por el modelo son mayores en todos los indicadores utilizados. En particular, el coeficiente de determinación indica que se puede explicar en torno al 40% del resultado de un alumno a partir de sus características individuales y de inteligencia emocional, independientemente de las variables escolares.

Así, se puede inferir que el mayor ajuste obtenido por el modelo no lineal implica que existen distintas interacciones entre las variables explicativas que hace que se deban interpretar los anteriores coeficientes con cautela. Por otro lado, el mayor inconveniente de las técnicas de RNAs es que por su naturaleza no paramétrica no permiten obtener coeficientes que cuantifiquen la importancia final de cada variable e independientemente del punto de la distribución de resultados donde se encuentre en el resultado final.

Sin embargo, se puede llevar a cabo cierto análisis de sensibilidad, calculando derivadas parciales en diversos puntos de la distribución de resultados, para obtener un indicador aproximado de la importancia de cada variable en el modelo. La tabla 4.10 muestra la importancia relativa de las 120 variables utilizadas en el análisis que se llevó a cabo mediante un análisis de las ponderaciones entre las conexiones. Hay que insistir en que este es un indicador muy burdo debido a que no recoge ni las interacciones entre las variables ni la importancia de éstas sobre el resultado en cualquier punto de la distribución.

Tabla 4.10. Ponderaciones asignadas por el programa de redes neuronales Neuroshell 2 en función de la contribución relativa de cada variable a la predicción de la variable dependiente.

Rank	Variables	Peso	Rank	Variables	Peso	Rank	Variables	Peso
1	BUENOMATES	0,01533	41	APEYAM5	0,00874	81	DAÑOMAS5	0,00731
2	LIBROSCASA	0,01512	42	ALSECSEL	0,00869	82	SCIANIMEXTIN	0,00728
3	EXMIY2	0,0147	43	ESTMMAS5	0,00865	83	SPORMAS5	0,00727
4	EDAD	0,01421	44	LIKETRABMAT	0,00858	84	OTROS	0,00726
5	UNISELF	0,01305	45	PADRE	0,00858	85	VOCACMAD	0,00724
6	SEXO	0,01201	46	ROBA34	0,00851	86	ALUNIPAD	0,00724
7	CASTELLANO	0,01138	47	DICCIONARIO	0,00851	87	HOME3Y5	0,00724
8	ESTCM1	0,011	48	BIENMATSELF	0,00848	88	CLUB3Y5	0,00723
9	PEYAMAS5	0,01061	49	ESTO1Y2	0,00842	89	NOTRABAJAN	0,00722
10	GUSTANMATES	0,01051	50	HERMANOS	0,0084	90	FRECOPERA	0,00719
11	FRECAVCOSU	0,01038	51	FRECCINE	0,00827	91	EXMM1	0,00715
12	MATFACILES	0,01034	52	AROBAM5	0,00826	92	AROBA34	0,00715
13	PAID3Y5	0,01027	53	FRECVIDEOJUE	0,00825	93	BIENMATPAD	0,0071
14	ESTDURO	0,01022	54	PROGORDEDUC	0,00817	94	TV1Y2	0,00703
15	MEMORIZAR	0,01011	55	EXMMAS5	0,00817	95	TVM1	0,007
16	SUERTE	0,00997	56	FRECNOTICIAS	0,00816	96	UNIMAD	0,00699
17	AMIGLENG	0,00994	57	PRIMSELF	0,00814	97	ESTOM1	0,00696
18	TV3Y5	0,00965	58	ADAÑO34	0,00812	98	AMIGM1	0,00693
19	ABURRAPMAT	0,00964	59	SECMAD	0,00801	99	ENCICLOPEDIA	0,00692
20	OPBUENOMAT	0,00959	60	BIENMATINSTI	0,00794	100	FRECTEATRO	0,0069
21	ESCRITORIO	0,00956	61	EXC1Y2	0,00793	101	AMIGMAT	0,00687
22	FRECVISMUSEO	0,00953	62	LIBR1Y2	0,00791	102	ESTC3Y5	0,00679
23	FRECNAURE	0,00949	63	HOMEMAS5	0,00789	103	PADRASTRO	0,00679
24	ESTM3Y5	0,00949	64	HOME1Y2	0,00787	104	AROBA12	0,00676
25	NACESPAÑA	0,00943	65	MADMEJCLASE	0,00785	105	AMIGMEJCLAS	0,00671
26	LIBRCIENCIAS	0,00932	66	HOMEM1	0,0078	106	PEYA12	0,00665
27	ESTOMAS5	0,00931	67	ESTC1Y2	0,00779	107	VOCACPAD	0,00664
28	NUMEROITEMS	0,00929	68	TVMAS5	0,00777	108	ALSECMAAD	0,0066
29	ADAÑO12	0,00922	69	DIVERAPMAT	0,00774	109	ROBA12	0,00649
30	SCIPOWNUCPL	0,00917	70	LIBRMAS5	0,00762	110	SEPORTANBIEN	0,00645
31	NINT3Y5	0,00915	71	SPORTM1	0,00761	111	DAÑO12	0,00634
32	AMIGMAS5	0,00906	72	MADRESP	0,00757	112	ALUNISEL	0,00628
33	SECPAD	0,00901	73	CLUBMAS5	0,00757	113	CLUBM1	0,00608
34	ESTMIY2	0,00898	74	VOCACSEL	0,00756	114	ALUNIMAD	0,00606
35	ORDENADOR	0,00893	75	SECSELF	0,00755	115	APEYA34	0,00605
36	PARIANTES	0,00887	76	HERMANAS	0,00748	116	AMIGDIVERS	0,00585
37	SCIAGUAPOLUT	0,00886	77	GENETICO	0,00745	117	ESTCMAS5	0,0058
38	FRECCONCIER	0,00883	78	ALSECPAD	0,00741	118	GALLEGO	0,00569
39	EXCM1	0,00879	79	MADRASTRA	0,0074	119	EXC3Y5	0,00563
40	ORDCLMAT	0,00879	80	APEYA12	0,00739	120	SPORT3Y5	0,0055

A continuación se comprueba si este ranking está o no cercano al de la importancia que asigna el modelo de regresión lineal. Para ello se calculan los coeficientes de correlación entre los valores de la t de Student de las variables en el modelo de regresión con el de los valores de la importancia relativa de las variables empleadas en el modelo de RNAs, tabla 4.11.

Tabla 4.11 Coeficiente de correlación entre la cuantía de la importancia explicativa asignada para cada variable en los métodos de regresión múltiple y de redes neuronales.

		t Student OLS	Importancia MLP
t Student OLS	Correlación de Pearson	1,000	,749**
	Sig. (bilateral)	,	,000
	N	120	120
Importancia MLP	Correlación de Pearson	,749**	1,000
	Sig. (bilateral)	,000	,
	N	120	120

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

El coeficiente de correlación obtenido es suficientemente alto como para afirmar que existe bastante similitud entre la importancia asignada a cada variable individualmente en ambos modelos. De igual manera se calcula el coeficiente de correlación no paramétrico de Spearman para la posición de cada variable en los rankings de importancia de ambos modelos, tabla 4.12.

Tabla 4.12. Coeficiente de correlación de rangos de Spearman entre los rankings de variables explicativas contruidos según su importancia en los métodos de regresión múltiple y de redes neuronales.

			Ranking OLS	Ranking MLP
Rho de Spearman	Ranking OLS	Coefficiente de correlación	1,000	,421**
		Sig. (bilateral)		,000
		N	120	120
	Ranking MLP	Coefficiente de correlación	,421**	1,000
		Sig. (bilateral)	,000	,
		N	120	120

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

El resultado obtenido indica que el coeficiente es de nuevo significativo, aunque más bajo, y por tanto, existe una correlación significativa no sólo en la cuantía sino también en el orden de importancia de las variables en ambos modelos. De nuevo señalar que el modelo esconde no linealidades e interacciones en las variables que pueden ser detectadas con las RNAs, pero que no se pueden cuantificar de forma paramétrica.

Una vez llevado a cabo el análisis a nivel de alumno se concluye que el modelo de RNAs consigue un mejor ajuste predictivo a los datos. Por tanto sus resultados serán más seguros a la hora de predecir el resultado de los alumnos e investigar la parte no explicada por las variables individuales a partir de las variables escolares.

4.4. La Construcción de los Outputs Escolares en Valor Añadido

Ahora que se dispone de un modelo capaz de realizar estimaciones acerca del resultado que un alumno obtendrá en la escuela, sin tener en cuenta las variables escolares, nuestro siguiente objetivo es tratar de explicar la diferencia entre el resultado escolar real y la estimación del modelo individual a partir de las variables escolares y las políticas educativas de cada centro. Para llevar a cabo este propósito se han construido los tres outputs en valor añadido definidos de la siguiente forma:

- *OUTPUT EN VALOR AÑADIDO 1*: Es el porcentaje de alumnos en cada centro escolar que obtiene un rendimiento real mayor que el esperado según sus características individuales utilizando el modelo 2 obtenido con RNAs. Este output, que tiene en cuenta a todos los tipos de alumno, considera eficiente a las escuelas que consiguen que un alumno obtenga un resultado mejor que el esperado, sea cual sea el punto de la distribución de resultados en el que se encuentre el alumno.
- *OUTPUT EN VALOR AÑADIDO 2*: Es el porcentaje de alumnos en cada centro escolar que según el modelo 2 debería obtener una puntuación menor que el primer cuartil de rendimiento y sin embargo obtienen un rendimiento mayor. Este output en valor añadido analiza los recursos escolares que influyen en que un alumno consiga escapar del conjunto del 25% de alumnos con peores resultados reduciendo por tanto sus posibilidades de fracaso escolar.
- *OUTPUT EN VALOR AÑADIDO 3*: Es el porcentaje de alumnos en cada centro escolar que según el modelo debería obtener una puntuación menor que el tercer cuartil de rendimiento y sin embargo obtienen un rendimiento mayor. Esta dimensión analizaría los factores que hacen que un alumno con un resultado medio pase a estar entre el 25% de los mejores estudiantes.

Además de analizar las variables escolares que son significativas para explicar estos outputs en valor añadido, se analizan también las variables que explican el resultado medio de la clase. El objetivo es comparar los tipos de recursos que son más productivos para distintos tipos de alumnos. La hipótesis de partida es que el modelo propuesto permitirá evitar asignar los mismos recursos a todos los alumnos en el caso de que distintos recursos importen sobre distintos tipos de alumnos. Así, el uso de este modelo sería más eficiente, que el uso de los modelos que sólo consideran el resultado medio de la clase ya que permite una asignación más óptima de los recursos escasos.

La tabla 4.13 contiene los estadísticas descriptivas de los 4 outputs, 3 en valor añadido y el resultado medio que serán analizados en la aplicación empírica.

Tabla 4.13: Estadísticas Descriptivas de los Outputs Utilizados.

		OUTPUT1	OUTPUT2	OUTPUT3	PROMEDIO
N	Válidos	147	142	147	147
	Perdidos	0	5	0	0
Media		48,1460	32,9938	22,4386	485,8593
Error típ. de la media		1,4468	2,1043	1,2512	2,7463
Mediana		48,4848	33,3333	20,8300	485,3900
Desv. típ.		17,5420	25,0758	15,1695	33,2975
Asimetría		-,074	,519	,658	,269
Error típ. de asimetría		,200	,203	,200	,200
Curtosis		-,166	-,085	,012	,163
Error típ. de curtosis		,397	,404	,397	,397
Rango		91,89	100,00	66,67	178,98
Mínimo		,00	,00	,00	407,00
Máximo		91,89	100,00	66,67	585,98
Percentiles	25	36,3636	14,2857	11,1100	459,2200
	50	48,4848	33,3333	20,8300	485,3900
	75	60,8696	50,0000	31,5800	506,0800

Los 5 valores perdidos existentes en el output en valor añadido 2 que refleja la tabla 4.11 se refieren a 5 escuelas para las cuales no existía ningún alumno que a priori estuviera en el grupo de riesgo de resultados previsibles entre el 25% de los peores resultados y por tanto no serán evaluados. Esta dimensión del output es sin lugar a dudas una de las que resultan más interesantes. El objetivo sería, por tanto, tomar decisiones y políticas que ayuden a reducir la probabilidad de fracaso escolar y sobre todo la asociada a los grupos de población con mayores desventajas socioeconómicas.

4.5. El Nivel Escolar

Para tratar de explicar qué variables influyen en las distintas dimensiones de outputs se han utilizado las variables escolares y de profesor que recogen el Anexo 2. Dado que sólo se disponen de 147 unidades y se carece de un modelo educativo que a priori indique cuáles son los inputs que son significativos, la metodología que se utiliza en este nivel para encontrar el modelo que mejor aproxime los datos, disponiendo de suficientes grados de libertad, ha sido la de llevar a cabo regresiones por pasos sucesivos o *stepwise regression*, [Darlington, 1990]. A partir de un modelo que se hace explícito se va alterando el mismo iterativamente, introduciendo y extrayendo variables, en base a algún criterio de entrada o salida definido previamente. La búsqueda se detiene cuando los criterios de entrada-salida se detienen o tras un número de pasos definido.

El criterio de entrada-salida seguido es el conocido como *hacia atrás*. Mediante este método se incluyen, antes del primer paso, todas las variables del modelo, e iterativamente se van borrando aquellas variables que no afectan de forma significativa al estadístico F del modelo. Una variable se introduce en el modelo si el nivel de significación de su valor F es menor que el valor de entrada, y se elimina si el nivel de significación de su valor F es mayor que el valor de salida⁵². La entrada debe ser menor que la salida y ambos valores deben ser positivos.

A continuación, se recogen las tablas con las variables que resultan ser explicativas para cada uno de los outputs. Los inputs en negrita son aquellos que únicamente resultaron ser significativos para el output evaluado, mientras que los inputs en cursiva se encuentran repetidos en otras dimensiones del output pero no en todas. Por último, los inputs en letra normal son aquellos que se repiten en todos y cada uno de los outputs. En primer lugar se indicará para cada output aquellos inputs que únicamente son significativos para ellos. Por último se analizan los inputs que son relevantes en varios y en todos los outputs.

4.5.1. Los Inputs Significativos para Una Sola Dimensión de Output

En primer lugar, se calculan los inputs escolares que importan para explicar el output en valor añadido 1, es decir el porcentaje de alumnos en cada escuela con un resultado real mayor al esperado, y que recoge la tabla 4.14.

⁵² Los valores de entrada y de salida utilizados fueron 0,05 y 0,1 respectivamente.

Tabla 4.14. Coeficientes explicativos del VA en educación a nivel escolar: OUTPUT 1

Variable	Coefficiente	Coef. Estand.	t	P-Valor
(Constante)	51,584 (18,289)		2,82	0,006
dirhabpadres	-0,488 (0,178)	-0,2	-2,747	0,007
semanasaño	1,18 (0,342)	0,246	3,446	0,001
<i>dismaterial</i>	-8,754 (4,745)	-0,133	-1,845	0,067
reuniones	-17,786 (5,936)	-0,204	-2,996	0,003
<i>sexto</i>	9,557 (2,616)	0,266	3,654	0,000
planifimates	0,652 (0,248)	0,182	2,63	0,010
planifotras	-1,09 (0,351)	-0,22	-3,105	0,002
horascoopera	2,487 (1,315)	0,14	1,891	0,061
<i>librotexto</i>	5,906 (2,483)	0,168	2,379	0,019
gastomaterial	-13,377 (4,331)	-0,234	-3,089	0,002
<i>minsemana</i>	-0,102 (0,036)	-0,204	-2,813	0,006
<i>calclase</i>	1,396 (0,705)	0,14	1,979	0,050
debate	3,417 (2)	0,123	1,709	0,090
<i>profguia</i>	-4,044 (1,714)	-0,162	-2,36	0,020
porparejas	-7,414 (2,262)	-0,236	-3,277	0,001
profhpadres	-6,364 (2,218)	-0,203	-2,869	0,005
pizarra	-0,14 (0,052)	-0,187	-2,703	0,008
tareaclase	-0,219 (0,059)	-0,264	-3,72	0,000
<i>ratio</i>	0,43 (0,193)	0,179	2,232	0,027
<i>repetidores</i>	-4,464 (2,58)	-0,124	-1,73	0,086
<i>privado</i>	5,634 (3,057)	0,149	1,843	0,068

Variable dependiente: Output 1

Resumen del Modelo para el Output en Valor Añadido 1

R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
0,704	0,495	0,407	13,559

Análisis de Varianza para el Output en Valor Añadido 1

	Suma de cuadrados	gl	Media cuadrática	F	p-valor
Regresión	21647,153	21	1030,817	5,607	0,000
Residual	22062,173	120	183,851		
Total	43709,326	141			

La tabla 4.14 ilustra los inputs que finalmente resultaron ser explicativos para el output en valor añadido 1. Asimismo se comprueba cómo el modelo es significativo y explica en torno al 40% del resultado.

El coeficiente negativo de **Dirhabpadres** y **profhabpadres** parece indicar que aquellos colegios en los que tanto el director del mismo como el profesor de matemáticas dedican más tiempo a hablar con los padres tienen peores resultados. Se puede explicar este fenómeno debido a que en las escuelas más conflictivas tanto el director, los profesores, pero fundamentalmente los padres, que ven que sus alumnos tienen determinados problemas tienden a concertar más entrevistas para hablar de los problemas de sus hijos en la escuela.

Semanasaño indica que aquellos colegios que tienen más semanas escolares, en el modelo educativo español serían aquellos que empiezan en los primeros días de septiembre y terminan en los últimos días de junio, obtendrían mejores resultados. **Planifotras** indica que las horas que el profesor de matemáticas dedica a planificar otras asignaturas distintas a la de enseñar matemáticas influyen negativamente en el resultado. Esta variable parece indicar que sería más eficiente que los profesores de matemáticas estuvieran dedicados en exclusiva a esta labor y no preparen clases de forma simultánea para distintas asignaturas.

Horascoopera muestra que a medida que aumentan las horas dedicadas a reuniones de profesores para cooperar en la elaboración de actividades y ayuda a resolver problemas aumentan los resultados. **Gastomaterial** indica que parece más positivo que se centralicen las decisiones en torno al gasto en material y que no sea dejada esta decisión a los profesores de manera individual.

En cuanto a los comportamientos y metodologías de enseñanza que se siguen en clase, **Debate** parece indicar que la actividad de debatir o discutir las respuestas a cuestiones que se realizan en clase aumenta el número de alumnos que mejoran el resultado. Sin embargo **porparejas y pizarra**, parecen señalar que tanto que los alumnos trabajen a menudo por parejas sin control del profesor o que los alumnos copien directamente los apuntes de la pizarra son

metodologías discutibles en cuanto a los resultados que obtienen, ya que disminuye el output en valor añadido 1 de la escuela.

La tabla 4.15 recoge a continuación los inputs que resultaron ser significativos para explicar el output 2, es decir las políticas que hacen que un alumno que a priori estaba en la zona de riesgo consiga salir de ella.

Tabla 4.15: Coeficientes explicativos del VA en educación a nivel escolar: OUTPUT 2

Variable	Coeficiente	Coef. Estand.	t	p-valor
(Constante)	5,458 (17,742)		0,308	0,759
mismoprofesor	15,19 (8,005)	0,137	1,898	0,060
reuniones	-17,66 (8,414)	-0,148	-2,099	0,038
planifmates	1,458 (0,362)	0,291	4,032	0,000
<i>apoyo</i>	-5,041 (1,815)	-0,208	-2,777	0,006
unasemana	-10,47 (3,763)	-0,199	-2,783	0,006
temas	9,191 (3,878)	0,184	2,37	0,019
librosprofe	6,939 (2,665)	0,188	2,603	0,010
tiempotexto	4,714 (1,877)	0,184	2,511	0,013
razonar	8,232 (3,353)	0,178	2,455	0,016
ayudarespuesta	-7,009 (3,191)	-0,159	-2,197	0,030
<i>profguia</i>	-7,275 (2,576)	-0,204	-2,824	0,006
tareanotas	-5,85 (2,211)	-0,194	-2,646	0,009
tareaclase	-0,243 (0,09)	-0,205	-2,698	0,008
rural	6,277 (3,598)	0,129	1,745	0,084
<i>ratio</i>	0,834 (0,255)	0,243	3,27	0,001

Variable dependiente output 2

Resumen del Modelo para el Output en Valor Añadido 2

R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
0,662	0,438	0,369	19,447

Análisis de Varianza para el Output en Valor Añadido 2

	Suma de cuadrados	gl	Media cuadrática	F	p-valor
Regresión	35733,251	15	2382,217	6,299	0,000
Residual	45762,081	121	378,199		
Total	81495,332	136			

El signo positivo de **mismoprofesor** indica que para el progreso del grupo de alumnos más desaventajados resulta positivo permanecer durante varios años con un mismo profesor. Así, el profesor acabaría conociendo tras el primer curso al alumno y sabría como tratarle, cuales son sus áreas débiles y cuanto exigirle en función de sus posibilidades. Cambios anuales de profesor quizás vendrían a fomentar un nuevo período de adaptación tanto del niño al profesor como del profesor al niño.

En cuanto a los métodos pedagógicos y actividades en el aula, **unasemana** indica que en los colegios donde los profesores se reúnen una vez a la semana se obtienen peores resultados que en los que sus profesores se sitúan en otras categorías de la variable frecuencia de las reuniones. **Tópicos** indica que es positivo para el rendimiento de los alumnos desaventajados que el profesor tenga influencia en los tópicos y materias que serán enseñados. En este caso la explicación podría estar en que de esta forma el profesor puede adaptar el temario a las características del alumno.

Librosprofe indica que existe una correlación positiva entre el número de libros que el profesor tiene en casa y el progreso en los resultados académicos de sus alumnos más desfavorecidos. **Tiempotexto** indica que cuanto más porcentaje del tiempo de clase de matemáticas se dedica al libro de texto más fácil resulta para el alumno su progreso. Una posible explicación de este resultado sería que al alumno con dificultades le sea más fácil seguir un libro que copiar únicamente los apuntes de la pizarra. Igualmente, **razonar**, indica que es muy recomendable hacer que los alumnos desaventajados expliquen a menudo razonadamente las ideas relacionadas con el temario.

Ayudarespuesta indica sin embargo que no resulta positivo para este grupo de alumnos que un compañero corrija a otro después de una respuesta incorrecta. Así habría que ser muy cauteloso con esta metodología y aplicarla con cuidado de que los alumnos que no contestan correctamente no se sientan incómodos como para no responder en futuras ocasiones. **Tareanotas** señala que tener en cuenta las tareas de casa para evaluar a los alumnos más desfavorecidos tampoco parece la mejor metodología a seguir con este grupo de alumnos.

Por último, **rural** indica que los alumnos desfavorecidos que se encuentran en localidades rurales, es decir fuera de las grandes ciudades, consiguen escapar mejor de la zona de alto riesgo de fracaso escolar.

A continuación la tabla 4.16 recoge los inputs que resultan significativos para explicar el output 3, es decir las políticas que hacen que un alumno que, a priori, no pertenecía por sus características al grupo de alumnos con mejores resultados escolares si que pertenezca después de pasar por la escuela. ¿Qué factores hacen por tanto que un alumno con un rendimiento medio pase a estar entre los de mejores resultados?.

Tabla 4.16. Coeficientes explicativos del VA en educación a nivel escolar: OUTPUT 3

Variable	Coeficiente	Coef. Estand.	t	p-valor
(Constante)	-5,017 (11,268)		-0,445	0,657
mas5años	-0,117 (0,046)	-0,18	-2,551	0,012
<i>mas75%</i>	-0,194 (0,058)	-0,234	-3,318	0,001
<i>preescolar</i>	0,21 (0,066)	0,222	3,177	0,002
<i>dismaterial</i>	-7,404 (4,06)	-0,132	-1,824	0,071
reuniones	-9,855 (5,209)	-0,133	-1,892	0,061
educacion	2,548 (1,271)	0,144	2,006	0,047
<i>sexto</i>	6,114 (2,211)	0,2	2,765	0,007
planifmates	0,913 (0,216)	0,299	4,222	0,000
tareas	1,521 (0,85)	0,136	1,79	0,076
<i>apoyo</i>	-3,207 (1,121)	-0,222	-2,86	0,005
<i>librotexto</i>	8,707 (2,147)	0,29	4,056	0,000
<i>calclase</i>	1,681 (0,64)	0,197	2,626	0,010
preguntaotro	2,941 (1,641)	0,131	1,792	0,076
autocorregir	-2,713 (1,235)	-0,159	-2,197	0,030
tareadiscusion	2,753 (1,301)	0,158	2,117	0,036
tareaclase	-0,191 (0,051)	-0,27	-3,772	0,000
cenciuda	-5,128 (2,486)	-0,145	-2,063	0,041
<i>repetidores</i>	-4,548 (2,245)	-0,149	-2,026	0,045
<i>privado</i>	7,628 (2,469)	0,237	3,09	0,002

Variable dependiente output 3

Resumen del Modelo para el Output en Valor Añadido 3

R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
0,670	0,449	0,363	11,980

Análisis de Varianza para el Output en Valor Añadido 3

	Suma de cuadrados	Media cuadrática	F	Sig.	p-valor
Regresión	14270,154	19	751,061	5,233	0,000
Residual	17510,746	122	143,531		
Total	31780,900	141			

Mas5años señala que aunque muy ligeramente, un mayor porcentaje de profesores que llevan más de cinco años en la escuela reduce el número de alumnos que alcanza el grupo de los mejores. Puede por tanto que el grupo de profesores más jóvenes y motivados ayude a los alumnos de nivel intermedio a despertar un mayor interés por la asignatura que les haga mejorar sus resultados. Además, **educación** indica que el nivel de educación formal del profesor influye positivamente en el porcentaje de alumnos que consiguen entrar en el grupo de los alumnos más aventajados.

Tareas indica que para los alumnos más aventajados resulta óptimo dedicar más horas a supervisar sus tareas. De esta forma al alumno que va bien se le puede pedir un nivel de exigencia algo mayor a la hora de revisar su trabajo.

A diferencia de lo que se ha visto para otros outputs, **preguntaotros** indica que la actividad en clase que consiste en que el profesor pregunte a un alumno con más probabilidad de acertar después de la respuesta incorrecta de un compañero es positiva para los mejores alumnos. Por tanto, este método pedagógico fomentaría una sana competencia entre los mejores alumnos por ver quien es el que es capaz de responder a más cuestiones. Sin embargo, **autocorregir**, indica que no es positivo que los propios alumnos corrijan cada uno sus propios deberes frecuentemente. El alumno no se esforzaría o lo haría en menos medida ya que muchos de ellos tan sólo se limitarían a copiar las soluciones de la pizarra. Por otro lado, **tareadiscusion** indica que si resulta positivo realizar discusiones en torno a la tarea que se está resolviendo. Estas discusiones fomentaría el debate y que se planteen distintas preguntas al profesor que aumenten los conocimientos del tema tratado.

Por último, **cenciudad** indica que recibir la educación en un colegio ubicado en el centro de una ciudad supone cinco puntos porcentuales menos de alumnos que entran en el

grupo de los estudiantes con mejores resultados, frente a vivir en zonas rurales, que no son ciudad o en las afueras de una ciudad.

Finalmente la tabla 4.17 recoge los inputs que resultan significativos para explicar el resultado medio de la clase. Este es el output que más veces ha sido utilizado a nivel de escuela en los estudios de función de producción educativa. El objetivo es comparar si las variables que influyen en el resultado medio son las mismas que para distintos colectivos de alumnos dentro del aula. Por otro lado, se comparan los resultados de eficiencia en la obtención de outputs en valor añadido con los del resultado medio.

Tabla 4.17. Coeficientes explicativos del VA en educación a nivel escolar: MEDIA

Variable	Coeficiente	Coef. Estand.	t	p-valor
(Constante)	412,661 (21,91)		18,834	0,000
<i>mas75%</i>	-0,238 (0,122)	-0,131	-1,943	0,054
faltaclase	-2,1 (0,908)	-0,154	-2,312	0,022
<i>preescolar</i>	0,5 (0,136)	0,242	3,689	0,000
reuniones	-42,565 (10,736)	-0,261	-3,965	0,000
<i>sexto</i>	10,833 (4,517)	0,162	2,398	0,018
planifmates	1,781 (0,439)	0,266	4,053	0,000
burocraprof	-1,913 (0,975)	-0,131	-1,962	0,052
librotexto	11,045 (4,344)	0,168	2,542	0,012
<i>minsemana</i>	0,129 (0,064)	0,138	2,002	0,047
tareaclase	-0,407 (0,104)	-0,263	-3,909	0,000
<i>privado</i>	15,899 (4,796)	0,226	3,315	0,001

Variable dependiente: resultado medio

Resumen del Modelo para Resultado Medio

R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
0,678	0,460	0,415	25,142

Análisis de Varianza para el Resultado Medio de la clase

	Suma de cuadrados	gl	Media cuadrática	F	p-valor
Regresión	70080,932	11	6370,994	10,078	0,0000
Residual	82179,476	130	632,150		
Total	152260,407	141			

De los inputs introducidos únicamente dos explican las diferencias entre escuelas en el resultado medio. **Faltaclase** indica que el porcentaje de alumnos que a diario falta a clase influye negativamente en el resultado medio de la clase. De igual forma **burocraprof** señala que el número de horas que el profesor dedica a actividades burocráticas influye negativamente en el resultado medio de sus alumnos.

4.5.2. Los Inputs Significativos para Varias Dimensiones de Output

A continuación, se señalan las características de aquellos inputs que resultan ser significativos para explicar varios de los outputs considerados.

Dismaterial indica que aquellos colegios que no fomentan las reuniones y claustros de profesores para la discusión de material y métodos educativos obtienen peores resultados en los outputs en valor añadido 1 y 3. La variable **sexto** muestra que los profesores que no simultáneas las clases de octavo nivel con otros cursos, en particular con clases en sexto de primaria, obtuvieron mejores resultados en sus alumnos que aquellos profesores que si lo hacían. **Librotexto** indica que es positivo que el profesor tenga mucha influencia en la elección del libro de texto. Parece, por tanto, adecuado que cada profesor elija de un abanico de posibilidades la que mejor se adapte a los objetivos y preferencias del mismo.

Minsemana es una variable con un resultado controvertido ya que por una parte presenta un signo negativo para el output en valor añadido 1 y un signo positivo para el resultado medio de la escuela. Así, clases de matemáticas más largas aumentarían el resultado medio de la clase pero no fomentaría el aumento del valor añadido. **Calclase** indica que existe un coeficiente negativo asociado al uso de la calculadora en clase de matemáticas en los outputs en valor añadido 1 y 3. **Profguia** señala que tampoco parece adecuado que los alumnos trabajen en grupo con el profesor como guía. En muchas ocasiones este tiempo no es aprovechado por el alumno que lo dedica a hablar o a copiar todos la tarea a realizar del alumno con más conocimientos.

Ratio muestra como una reducción del número de alumnos en clase tiene un efecto negativo para el output 1 y el output 2. Este resultado, contrario al esperado por la lógica educativa de que menos alumnos en clase supondría que el profesor pueda dedicar más tiempo a cada uno de ellos, se ha encontrado en distintos trabajos y sin embargo se siguen destinando recursos a reducir el número de alumnos en clase. Parece por tanto que sería más eficiente llevar a cabo otras políticas más productivas para el aumento de resultados. La explicación al comportamiento de esta variable no es bien conocido y sin duda más investigación educativa parece necesaria para la búsqueda del número de alumnos óptimo en cada clase.

El número de repetidores en la clase viene expresado en la variable **repetidores**, que muestra una correlación negativa para los outputs 1 y 3. Parece claro que estos alumnos podrían retrasar el aprendizaje de los mejores alumnos de la clase. Por último la variable **privado** muestra que salvo para los alumnos más desaventajados que recoge el output 2, en los cuales la diferencia no es significativa, los centros privados concertados obtienen mejores resultados que los centros públicos. Sin ser definitivo, en este punto de nuevo parece necesaria más investigación en torno a los beneficios que reportaría un sistema educativo con provisión pública y producción privada, lo que se ha dado en llamar el *cheque escolar*.

La variable **apoyo**, que hace referencia al número de horas que los profesores dedican a tareas de apoyo a los alumnos, está correlacionada negativamente con el resultado en los outputs 2 y 3. Concluimos por tanto que aquellos profesores que detectan más deficiencias en el aprendizaje dedican más horas a intentar ayudar a los alumnos más desaventajados repercutiendo negativamente en el resto de alumnos. Parece por tanto más adecuado la introducción de profesores de apoyo específicos que complementen el trabajo normal del profesor para aquellos alumnos que más lo necesiten e incluso para aquellos que van muy bien.

Por otro lado **mas75%** indica que el porcentaje de profesores en la escuela con más del 75% de su docencia dedicada a las matemáticas obtienen peores resultados en el output 3 y en el resultado medio. Por tanto y junto con el resultado anterior de la variable **sexto** la política óptima sería que el profesor de matemáticas diera clase a un mismo curso sin concentrar toda la docencia en la asignatura de matemáticas. Tanto para el resultado medio como para el output 3 la variable **preescolar**, que hace referencia al porcentaje de alumnos en la escuela que asistió a preescolar, está correlacionada positivamente con el rendimiento académico. Sería, por tanto, muy eficiente las políticas de escolarización temprana de los alumnos.

4.5.3. Los Inputs Significativos para Todas las Dimensiones de Output

De todas las variables empleadas, encontramos dos que resultan ser significativas a lo largo de los cuatro outputs. Por un lado, la variable **reuniones**, indica muy claramente que el hecho de que los profesores no se reúnan regularmente para discutir distintas ideas acerca del material y de los métodos educativos es negativo tanto para los outputs en valor añadido como para el resultado medio de la clase. La cuantía del resultado es más negativa para los outputs en valor añadido 1 y 2.

Por otro lado, la variable **planifmates** muestra como el número de horas semanales que el profesor dedica a planificar la clase de matemáticas está positivamente correlacionado con los buenos resultados en matemáticas en los 4 outputs. De los outputs en valor añadido el output 2 es el que saldría más beneficiado de que los profesores dediquen más horas a la planificación de la asignatura de matemáticas.

Por tanto, parece evidente que ambas políticas deberían ser incentivadas por los centros a partir de los resultados del modelo. Además ambos inputs influyen de forma más incisiva sobre aquellos alumnos con más desventaja de partida. Asimismo, se obtiene cómo el modelo es capaz de dar recomendaciones a cada centro sobre como asignar los recursos tanto materiales, de tiempo y de actividades, para maximizar los outputs de los distintos tipos de alumnos, es decir, mediante el modelo se puede aumentar la eficiencia asignativa de cada escuela reordenando los recursos que se aplican a la educación de los alumnos.

4.5.4. La Medición de la Eficiencia en las Escuelas.

Una vez que se sabe cómo los inputs se han transformado en outputs se puede medir la eficiencia⁵³ técnica como la distancia entre la cantidad del output que cada colegio debería haber obtenido, teniendo en cuenta como asignó sus recursos, y la cantidad realmente alcanzada. Por tanto, el indicador obtenido sería la eficiencia técnica de la escuela ya que se considera que las cantidades de inputs utilizadas son fijas⁵⁴. Asimismo, con el fin de simplificar se considera que toda la distancia es debida a la eficiencia escolar, aunque este supuesto podría ser relajado descomponiendo la distancia entre la observación real y la predicción en ruido aleatorio y eficiencia tal y como quedó planteado en el apartado 3.4.3.

⁵³ Para medir la eficiencia técnica hemos tenido en cuenta los resultados reales y las predicciones recogidas en el Anexo 3.

⁵⁴ Sin embargo tal y como ya hemos apuntado podemos asignar, a partir de la información obtenida en el modelo, más eficientemente los recursos productivos en forma de políticas pedagógicas y educativas.

El anexo 4 muestra las medidas de eficiencia obtenidas por las escuelas. Para los tres outputs en valor añadido, aquellos centros con un valor mayor que uno serían eficientes, ya que consiguen que un mayor porcentaje de sus alumnos actúe por encima de lo esperado en la cuantía porcentual que supere la unidad. Por el contrario, aquellas escuelas con un valor de eficiencia por debajo de uno serían ineficientes, ya que un porcentaje de sus alumnos, menor al esperado, obtienen un mejor resultado que el esperado. Para el resultado medio la interpretación es similar, las escuelas cuyos alumnos en media obtienen un resultado mayor que el esperado serán eficientes en la cuantía que sobrepasen la unidad.

Gráfico 4.1: Representación gráfica de la eficiencia técnica para el output en valor añadido 1.

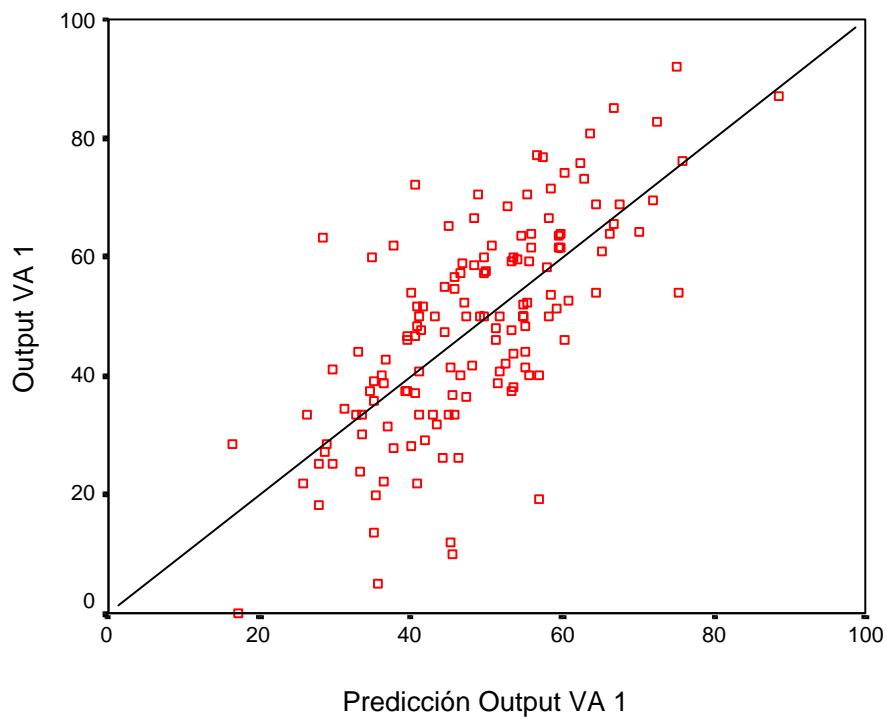


Gráfico 4.2: Representación gráfica de la eficiencia técnica para el output en valor añadido 2.

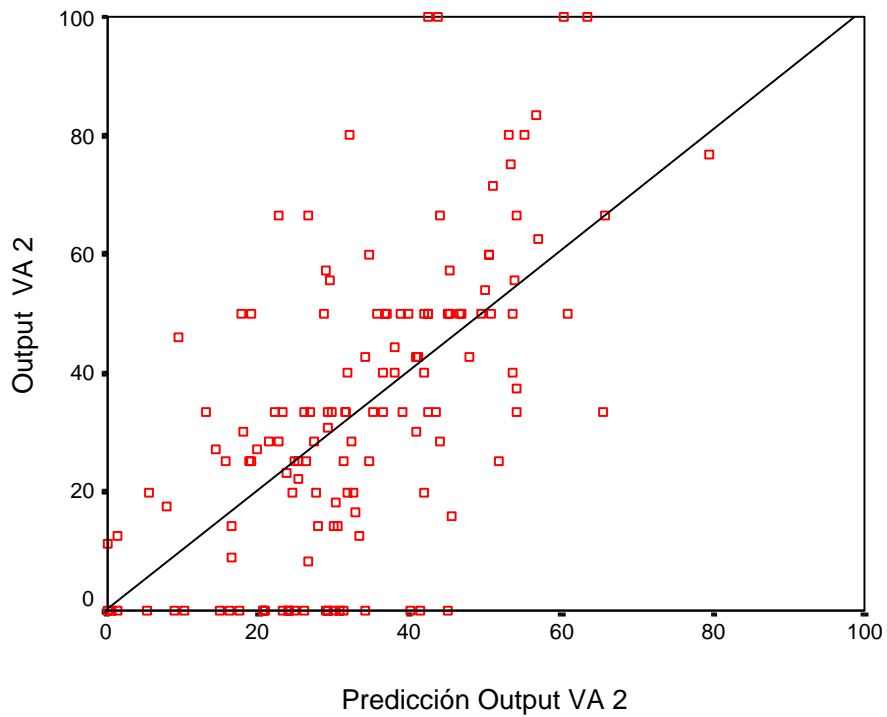


Gráfico 4.3: Representación gráfica de la eficiencia técnica para el output en valor añadido 3.

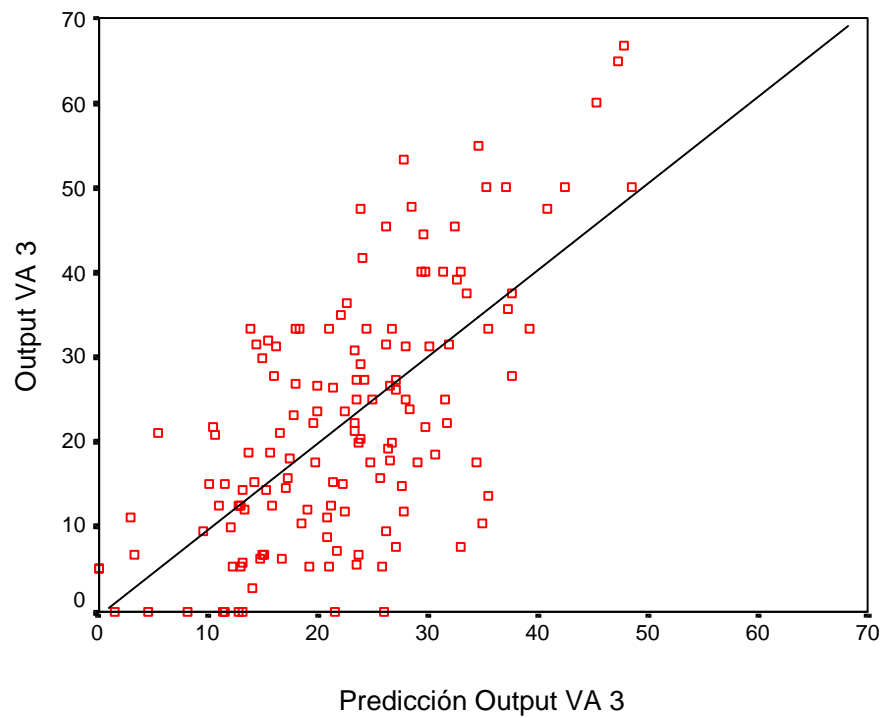
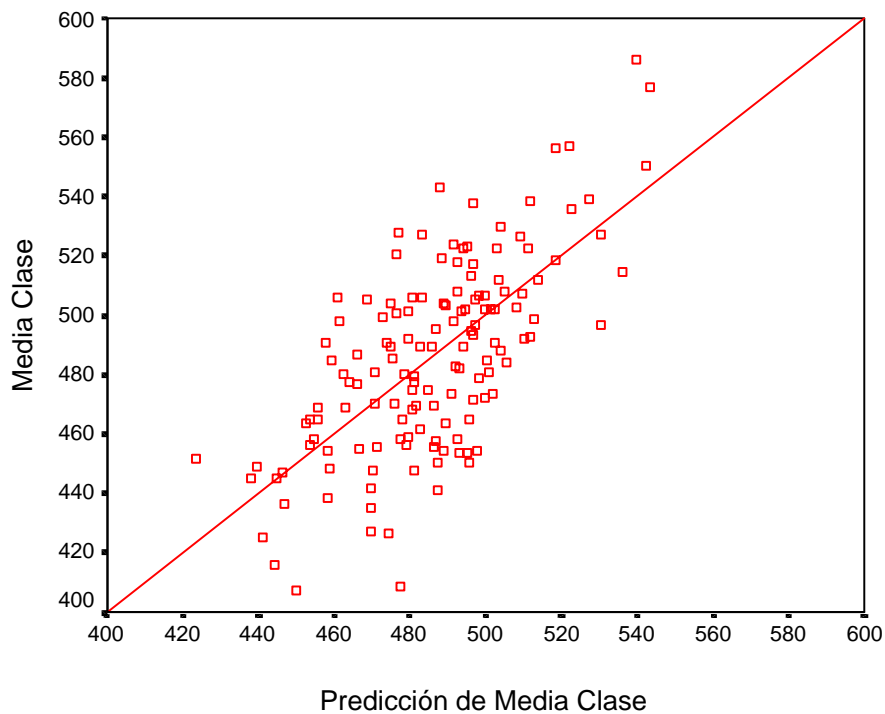


Gráfico 4.4: Representación gráfica de la eficiencia técnica para el output de la media de la clase.



Los gráficos 4.1, 4.2, 4.3 y 4.4 ilustran claramente la forma en la que se ha obtenido la eficiencia técnica de las escuelas. Aquellos colegios por encima de la línea de 45° serían eficientes ya que obtienen un resultado mejor que el esperado. A medida que una escuela se sitúa más a la derecha en el gráfico dispone de una mayor cuantía y una mejor asignación de recursos que le hace tener un potencial de resultados mayor. Por otro lado, una escuela es más eficiente a medida que se distancia por arriba de su objetivo de resultados inicial.

La tabla 4.18 muestra el coeficiente de correlación entre las medidas de eficiencia técnica para los cuatro outputs. Así, se comprueba cómo todas ellas están correlacionadas positivamente y en casi todos los casos la correlación es significativa. Por tanto, normalmente un centro eficiente lo será en todas las dimensiones. Sin embargo, el output en valor añadido 2, que hace referencia a los alumnos más desaventajados, presenta una correlación mucho más débil con el resto de outputs. La conclusión que se deriva es que una escuela puede tener un resultado medio aceptable y sin embargo no estar ayudando a sus alumnos más desaventajados. Las políticas a seguir con distintos alumnos son, por tanto, diferentes y así un colegio puede ser eficiente educando a determinado tipo de alumnos pero ineficiente con otros.

Tabla 4.18: Coeficiente de correlación de Pearson entre el valor absoluto de eficiencia de las escuelas en los distintos outputs.

		EFIC1	EFIC2	EFIC3	EFIMEDIA
EFIC1	Correlación de Pearson	1.000	.189*	.475**	.602**
	Sig. (bilateral)	.	.027	.000	.000
	N	142	137	142	142
EFIC2	Correlación de Pearson	.189*	1.000	.354**	.110
	Sig. (bilateral)	.027	.	.000	.199
	N	137	142	137	137
EFIC3	Correlación de Pearson	.475**	.354**	1.000	.453**
	Sig. (bilateral)	.000	.000	.	.000
	N	142	137	142	142
EFIMEDIA	Correlación de Pearson	.602**	.110	.453**	1.000
	Sig. (bilateral)	.000	.199	.000	.
	N	142	137	142	142

*. La correlación es significativa al nivel 0,05 (bilateral).

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

El anexo 5 presenta el ranking de eficiencia que ocupa cada colegio en cada output. Podemos comprobar cómo los outputs en valor añadido construidos en el modelo proporcionan una información mucho más rica que la del resultado medio, dado que podemos saber específicamente la cuantía de la eficiencia-ineficiencia de la escuela para cada tipo de alumnado.

Tabla 4.19: Coeficientes de correlación no paramétricos de Spearman entre los rankings de eficiencia alcanzados por cada escuela en los distintos outputs.

			RANKING1	RANKING2	RANKING3	RANKMED
Rho de Spearman	RANKING1	Coefficiente de correlación	1.000	.381**	.588**	.586**
		Sig. (bilateral)	.	.000	.000	.000
		N	142	137	142	142
	RANKING2	Coefficiente de correlación	.381**	1.000	.282**	.299**
		Sig. (bilateral)	.000	.	.001	.000
		N	137	142	137	137
	RANKING3	Coefficiente de correlación	.588**	.282**	1.000	.589**
		Sig. (bilateral)	.000	.001	.	.000
		N	142	137	142	142
	RANKMED	Coefficiente de correlación	.586**	.299**	.589**	1.000
		Sig. (bilateral)	.000	.000	.000	.
		N	142	137	142	142

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

Además, la tabla 4.19 muestra, al igual que hacía la tabla 4.18, cómo los rankings de eficiencia de las escuelas están correlacionados positivamente y de forma significativa. De nuevo las correlaciones más débiles se producen entre el output 2 y el resto. Sin embargo, una observación más detallada del anexo 4 muestra cómo todas las combinaciones de eficiencia son posibles para distintas escuelas. Algunas de ellas son buenas con todos sus alumnos otras obtienen malos resultados con todos ellos y la mayoría son buenas en alguna dimensión y sin embargo presentan deficiencias en otras. De nuevo queda patente cómo la eficiencia obtenida únicamente para el resultado medio puede ser completamente engañosa a pesar de que todas las dimensiones estén correlacionadas positivamente.

Por tanto, a partir de este modelo se puede evitar ocultar la información de los alumnos más desaventajados en el resultado medio. Una escuela eficiente será aquella que consiga que estos alumnos salgan del grupo de riesgo y para ello las escuelas deberán aplicar políticas de asignación de recursos distintas a las que aplican a otros alumnos.

4.6. La Eficiencia Regional

Las transferencias en materia educativa en España han constituido un proceso que se originó cuando el País Vasco y Cataluña asumieron sus competencias en educación no universitaria en 1980. Posteriormente Galicia y Andalucía lo hacen en 1982, Canarias y la Comunidad Valenciana reciben los traspasos en 1983, Navarra en 1990, Baleares en 1998 y el resto de Comunidades Autónomas reciben las competencias entre 1999 y 2001. Además, las competencias en enseñanza universitaria se traspasan, salvo en Navarra, posteriormente, entre 1985 y 1987. Durante los años 1995 y 1996 tienen lugar las transferencias en materia universitaria a las Comunidades Autónomas que aún no las habían asumido.

Dado que actualmente todo el gasto educativo está descentralizado, se deberían articular mecanismos que velen por la eficiencia y la equidad en la asignación de los recursos en todo el Estado. Especialmente este control parece más oportuno en las primeras etapas de la educación con el fin de poder asegurar la igualdad de oportunidades educativas. Los estudios comparativos de los recursos y resultados para la educación que se aplican y se obtienen en las distintas regiones deben servir para garantizar ayudas a las regiones más necesitadas, no necesariamente en forma de mayores recursos, sino a través de una mejor asignación de los mismos. Aprender de los métodos educativos de otras regiones y detectar las principales potencialidades y debilidades de cada región debe ser el motor que incentive la introducción de eficiencia en el sistema educativo y el objetivo que se pretende alcanzar a través del modelo propuesto.

Santín (2001) concluye que una vez clasificadas las Comunidades Autónomas (CC. AA) respecto al grado de necesidad en inversiones y gasto educativo se pueden obtener dos importantes conclusiones. En primer lugar, la carencia de recursos influye significativamente en uno de los principales outputs escolares, el rendimiento académico, en particular el rendimiento en matemáticas a los 14 años. En segundo lugar, las diferencias que actualmente existen en España en gasto por alumno no responden a un esfuerzo de las regiones más desfavorecidas en la búsqueda del objetivo de alcanzar la igualdad de oportunidades, sino más bien al contrario. Las regiones con más necesidad de recursos gastan menos por alumno que aquellas regiones con menor necesidad. Las diferencias en gasto por alumno responden, por tanto, a otros criterios: costes de producción, oferta instalada, dispersión, demanda de los ciudadanos o al propio sistema de financiación autonómica, pero no parece que estén basadas en la distinta necesidad relativa del alumnado para garantizar una producción similar del mismo servicio.

De no romperse esta tendencia, las diferencias con toda probabilidad se agravarán, pudiendo llegar a existir regiones de “primera” y de “segunda” en la producción educativa, lo que repercutirá a largo plazo en el crecimiento económico de las Comunidades Autónomas y del conjunto del estado español.

El modelo de evaluación permanente del sistema educativo que se propone en este trabajo debería servir entre otros objetivos para que los ciudadanos sepan las diferencias existentes en calidad y grado de necesidad de gasto entre las distintas CC. AA, a fin de que éstos puedan reclamar, fundamentalmente al gobierno regional, pero también al central, un mejor servicio en función de sus preferencias y restricciones presupuestarias.

Una vez que se dispone de los datos de la eficiencia de las escuelas es posible investigar la eficiencia que obtuvieron las CC.AA que participaron en el TIMSS. Una de las limitaciones de este estudio es que sólo tenemos información de 8 regiones españolas⁵⁵. La segunda limitación a la que hay que hacer frente es que las autoridades educativas españolas decidieron ocultar los datos de los nombres de las regiones a las que pertenecía cada escuela y sólo hemos podido inferir los datos de las CC.AA de Cataluña, Comunidad Valenciana y Galicia a partir del hecho de que algunos alumnos usaron el idioma de estas regiones para completar el test de la prueba de matemáticas.

⁵⁵ Posiblemente las 8 regiones educativas se corresponden con las 7 CC. AA que en 1995 tenían transferidas las competencias en materia educativa más el denominado “territorio MEC”, o resto de regiones con las competencias en manos del Estado Central dependientes del Ministerio de Educación.

Es por ello que los resultados que mostramos a continuación son únicamente ilustrativos de la utilidad que proporcionaría el modelo para evaluar las diferencias en eficiencia con la que actúan las distintas regiones.

Tabla 4.20: Eficiencia Media y rankings de las regiones españolas que participaron en el TIMSS

REGION	Eficiencia1	Eficiencia2	Eficiencia3	EficMedia	Rank1	Rank2	Rank3	RankM
Región 1	0,958 (0,304)	0,919 (0,716)	1,199 (0,888)	0,989 (0,049)	7	6	3	6
Región 2	0,955 (0,463)	1,206 (0,731)	1,265 (1,468)	0,987 (0,041)	8	2	2	7
Región 3	1,463 (0,673)	0,627 (0,545)	1,573 (0,627)	1,033 (0,064)	1	8	1	1
Región 4	1,02 (0,209)	0,723 (0,627)	0,833 (0,48)	0,987 (0,032)	4	7	8	8
Región 5	0,987 (0,298)	0,96 (0,687)	0,977 (0,565)	1,006 (0,052)	6	5	5	3
Cataluña	1,021 (0,198)	1,116 (0,857)	0,915 (0,537)	1,005 (0,047)	3	4	7	4
Galicia	1,041 (0,184)	1,532 (0,913)	0,979 (0,411)	1,011 (0,027)	2	1	5	2
C. Valenciana	1,016 (0,317)	1,182 (1,174)	1,078 (0,955)	0,989 (0,055)	5	3	4	5

En la tabla 4.20 se puede observar cómo de las tres regiones conocidas, Galicia es la que mejores resultados obtiene respecto a la Comunidad Valenciana y Cataluña, que obtienen resultados muy parecidos. Galicia es, además, la región que mejores resultados obtiene en sus alumnos más desaventajados, si bien Cataluña y la Comunidad Valenciana también obtienen buenos resultados para este tipo de alumnos.

La tabla 4.21 muestra cómo la correlación entre los rankings de las regiones no es significativa e incluso para determinadas dimensiones es negativo. Se concluye, por tanto, que este modelo proporciona una mayor riqueza informativa para asignar recursos en aquellos outputs en los que la región presente más problemas o necesidades. En este sentido, una buena aplicación de la descentralización educativa permitirá unos resultados más eficientes que en el anterior modelo centralizado.

Tabla 4.21: Coeficientes de correlación no paramétricos de Spearman entre los rankings de eficiencia regionales.

			RANK1	RANK2	RANK3	RANKMEDI
Rho de Spearman	RANK1	Coefficiente de correlación	1,000	-,214	-,168	,667
		Sig. (bilateral)	,	,610	,691	,071
		N	8	8	8	8
	RANK2	Coefficiente de correlación	-,214	1,000	-,072	,000
		Sig. (bilateral)	,610	,	,866	1,000
		N	8	8	8	8
	RANK3	Coefficiente de correlación	-,168	-,072	1,000	,252
		Sig. (bilateral)	,691	,866	,	,548
		N	8	8	8	8
	RANKMEDI	Coefficiente de correlación	,667	,000	,252	1,000
		Sig. (bilateral)	,071	1,000	,548	,
		N	8	8	8	8

4.7. Recapitulación

En este capítulo se ha implementado empíricamente el modelo que propuesto en el capítulo 3. Las principales conclusiones obtenidas a partir de los resultados generados son las siguientes.

- Como resultado principal del modelo se confirma la hipótesis de que la política más eficiente en la asignación de recursos educativos no universitarios no es asignar las mismas dotaciones a distintos tipos de alumnos. El componente socioeconómico marca unas distancias ex-ante al proceso educativo que deberían ser superadas poniendo más y distintos recursos sobre aquellos alumnos más desfavorecidos y sobre aquellos con más probabilidad de fracaso escolar a priori. Se confirma así que la política de *café para todos* es claramente ineficiente para todos los alumnos, si bien este resultado no supone la división de los alumnos en distintas clases sino el tratamiento distinto dentro de una misma aula. Hay que tener en cuenta en este punto que el efecto compañeros se muestra a lo largo de la literatura como una variable muy relevante para los alumnos con más probabilidad de fracaso para la mejora de sus resultados.
- Se confirma cómo el proceso de producción educativa es complicado, ya que un modelo no lineal basado en RNAs consigue aproximar los datos a nivel de alumno con mayor acierto que los de un modelo lineal. Los beneficios de un mayor ajuste son, sin duda, muy relevantes ya que las diferencias entre el objetivo de rendimiento del alumno y el

que obtenga en realidad deben responder con el mayor ajuste posible al modelo real a fin de evitar errores de especificación que lleve a la toma de políticas educativas erróneas.

- Los resultados de esta trabajo muestran cómo, a grandes rasgos, podemos explicar el resultado escolar en un 20% debido a las características personales exógenas. Se explicaría en torno al 40% si se introduce además las variables relacionadas con la inteligencia emocional del alumno. Por último, los recursos escolares explicarían en torno a otro 25% más del resultado. En definitiva, se puede explicar con este modelo un 65% del resultado del alumno, lo que dejaría un 35% del resultado final debido a la eficiencia escolar y a variables aleatorias. Este resultado confirma que la escuela sí importa, ratificando la Teoría del Capital Humano, y que la investigación de la eficiencia y las políticas educativas que la incentiven son sumamente importantes para la mejora de la educación no universitaria.
- Si bien el resultado medio del aula es un indicador completamente válido de los resultados de sus alumnos éste resulta incompleto. Con este indicador no se puede comprobar si los alumnos han progresado o no sobre lo que de ellos cabría esperar en función de sus características. Es por ello que resultará más informativo completar este indicador con outputs en valor añadido. La ventaja de este tipo de output es que permite distinguir distintos tipos de alumnos, evitando así los problemas de agregar alumnos con distintas características.
- Los centros privados concertados obtienen mejores resultados tanto medios como en valor añadido que los centros públicos. Parece por tanto necesaria una revisión en profundidad en torno a cuál es la forma más eficiente de llevar a cabo la producción educativa.
- Un resultado de difícil explicación teórica, aunque también obtenido en otros trabajos, es el del signo positivo de la variable escolar *ratio*. Seguir destinando recursos a la reducción del número de alumnos en cada clase por delante de otros objetivos sin una mayor investigación que avale científicamente la bondad de esta política no parece que esté justificado.
- Las diferencias entre los valores de eficiencia obtenidos entre las distintas escuelas para los distintos outputs son muy heterogéneas. Si bien existe una correlación positiva en el sentido de que una escuela eficiente lo será en mayor o menor medida en todas las dimensiones, se puede observar todas las combinaciones posibles de escuelas eficientes e ineficientes a un mismo tiempo en distintos outputs. El modelo proporciona así información sobre los puntos fuertes y débiles de cada escuela.

- Las regiones españolas tienen unos valores de eficiencia muy distintos, tanto dentro de cada región como comparando entre varias regiones. En tanto sea aceptada, tal y como se ha defendido en este trabajo, la vigencia de la teoría del Capital Humano, resulta relevante que las autoridades se preocupen del nivel educativo de su región respecto al del resto ya que éste determinará en el medio y largo plazo buena parte de su crecimiento económico.

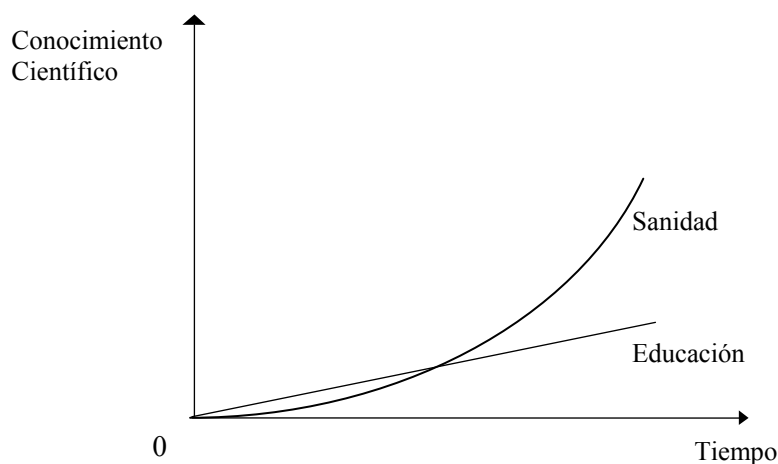
En definitiva parece necesaria una apuesta clara por la investigación educativa crítica y dinámica, que mejore el modelo año tras año, con el objetivo último de obtener la máxima eficiencia y por tanto aprovechar al máximo el uso de los recursos escasos con los que cuenta el sector público en general y el mundo educativo en particular.

**REFLEXIONES FINALES Y LÍNEAS FUTURAS DE
INVESTIGACIÓN**

A lo largo de esta tesis doctoral hemos ido recogiendo, al final de cada capítulo, las principales conclusiones derivadas en cada uno de ellos. En este último apartado pretendemos apuntar algunas reflexiones y cuestiones que quedan pendientes en cuanto a los temas desarrollados. El objetivo final es establecer un punto de partida para continuar investigando en aquellos temas relacionados con distintos apartados de esta Tesis que no han podido ser tratados totalmente en este trabajo.

Si estableciéramos una analogía entre el mundo de la sanidad y el de la educación descubrimos que ambos bienes preferentes son igualmente demandados por los ciudadanos para alcanzar una alta calidad de vida. Sin embargo, los avances que la sociedad ha experimentado en ambos campos podrían ser calificados de exponencial y lineal respectivamente, tal y como muestra la figura siguiente.

La evolución en el tiempo del conocimiento científico en educación y sanidad.



Cuando un individuo presenta una enfermedad, el sistema sanitario de los estados desarrollados tiene la capacidad, la mayoría de las veces, de darle un diagnóstico y un tratamiento para superar la patología. Muchos recursos han tenido que ser invertidos en investigación médica para alcanzar este objetivo. Sin embargo, las cosas en educación son bien distintas. Cuando un alumno presenta, por diferentes motivos, problemas de aprendizaje o de rendimiento en la escuela, no existe la mayoría de las veces la posibilidad de realizar lo que sería un “diagnóstico”, ni de ofrecer un “tratamiento”, según el problema del alumno. La misma solución, repetir curso o abandonar los estudios, suele ser la pésima fórmula o *tratamiento* que

se aplica a todos los alumnos con problemas. Esta solución parece tan ineficiente como sería recetar aspirinas para cualquier enfermedad.

Parece evidente que la educación debe ser objeto de una mayor inversión en investigación básica con el objeto de superar muchas de las ineficiencias del sistema educativo debidas al enorme desconocimiento del proceso productivo. Ésta sería una de las conclusiones principales de esta Tesis Doctoral. Sin embargo, muchas son las cuestiones que quedan planteadas y pendientes de desarrollo al finalizar este trabajo.

La principal pregunta que queda sin responder sería evaluar cómo funcionaría el modelo planteado en la realidad, cuantificando los beneficios y los costes que reportaría a la sociedad. En principio, la puesta en marcha del modelo requeriría de unos costes fijos en personal encargado de recoger los datos durante el primer año y analizarlos, así como en los equipos informáticos y software necesarios. Una vez informatizado el sistema, el funcionamiento del modelo sería relativamente fácil. A principio y a final de curso se pasarían cuestionarios a todos los protagonistas del proceso educativo; alumnos, padres, profesores y directivos. Con esta información se realizarían las predicciones de resultados o la *diagnosis* para los alumnos y se informaría de las mismas a padres y profesores. Se proporcionarían a continuación las recomendaciones en torno a las políticas, tanto dentro como fuera del aula, a seguir con cada tipo de alumno que contribuyan a aportar valor añadido al mismo.

Tal y como ha quedado reflejado a lo largo de esta Tesis *el mercado* no resuelve satisfactoriamente la asignación de recursos educativos, fundamentalmente debido a que la información disponible es imperfecta. El modelo aquí propuesto trataría de ser una solución, con todos sus problemas, a esta realidad. ¿Qué tipo de preguntas e información se recogería en los cuestionarios?, ¿Cómo mediríamos el rendimiento final y en qué pruebas?, ¿Deberían introducirse incentivos al profesorado en función de sus resultados?, ¿Podemos mejorar las predicciones que realizan las RNAs mediante otros métodos o combinaciones de métodos?, ¿Son beneficiosos los itinerarios o la separación de alumnos?, ¿Cuál es el número óptimo de alumnos por clase? son algunas de las cuestiones que quedan pendientes.

En cuanto a las RNAs ya hemos visto cómo su principal limitación radica, al igual que en todos los métodos no paramétricos, en la dificultad de interpretar la causalidad de los resultados predictivos obtenidos. A pesar de los recientes avances en inferencia llevados a cabo en esta técnica, debemos ser cautelosos a la hora de interpretar las predicciones de las RNAs y en cualquier caso, esta información debería ser complementada con análisis de tipo causal a nivel de alumno. En este último punto el análisis de búsqueda de reglas mediante otras de las

novedosas técnicas de *data mining*; en particular árboles de decisión y ecuaciones estructurales, parece una de las posibles líneas futuras que deberán ser evaluadas.

Por otro lado, tal y como quedó reflejado en el capítulo 2, las RNAs han demostrado que pueden superar a las técnicas tradicionales en determinados problemas de estimación de funciones productivas, y ser por tanto una opción más a la hora de evaluar la eficiencia técnica de unidades homogéneas en distintos sectores. Sin embargo, las aplicaciones empíricas reales son hasta la fecha muy escasas por lo que queda pendiente dar el salto definitivo de los resultados teóricos a las aplicaciones empíricas.

En definitiva, concluimos que el progreso en la investigación educativa debe ser uno de los objetivos prioritarios durante el siglo XXI con el fin de seguir avanzando hacia los objetivos de eficiencia y equidad de nuestra sociedad.

REFERENCIAS

- Aigner, D.J., Lovell, C.A.K. y Schmidt, P. (1977): "Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics* 6, 21-37.
- Akaike, H. (1973): "Information theory and an extension of the maximum likelihood principle. In Petrov, B. N. y Csaki, F., eds., *Proceeding 2nd International Symposium on Information Theory*. Budapest: Akademia Kiado. 267-281.
- Albi, E. (1992): "Evaluación de la eficiencia pública". *Hacienda Pública Española*, N° 120-121, págs. 299-316.
- Albi, E. (2000): *Público y privado: Un acuerdo necesario*. Ariel, Barcelona.
- Albi, E., González-Páramo, J. M. y Zubiri, I. (2000): *Economía Pública I*. Ariel, Barcelona.
- Alonso-Rodríguez, A. (1999): "Forecasting Economic Magnitudes with Neural Networks Models". *International Advances in Economic Research*, Vol. 5 N° 4. November.
- Alvarez, A. (2001): *La Medición de la Eficiencia y la Productividad*. Ed. Pirámide.
- Apostolos-Paul N. Refenes. (eds.). (1995). *Neural Networks in the Capital Markets*. John Willey & Sons. Chichester.
- Arbib, M.A. (1995): *The Handbook of Brain Theory and Neural Networks*. The MIT Press.
- Arrow, K. J. (1973): "Higher education as a filter". *Journal of Public Economics*, 2(3), págs. 193-216.
- Averch, H. A., Carroll, S. J., Donaldson, T. S., Kiesling, H. J. y Pincus, J. (1974): *How effective is schooling?. A critical review of research*. New Jersey. Educational Technology Publications. Englewood Cliffs.
- Banker, R. D. y Morey, R. C. (1986): "Efficiency Analysis for Exogenously Fixed Inputs and Outputs". *Operations Research*, Vol. 34 N° 4.
- Banker, R. D., Charnes, A., Cooper, W. W. y Maindiratta, A. (1988): "A comparison of DEA and translog estimates of production frontiers using simulated observations from a known technology" in Dogramaci, A. y Färe, R. (Eds.) *Applications of Modern Production Theory: Efficiency and Productivity* (Boston, MA, Kluwer).
- Baker, B. D. y Richards, C. E. (1999): "A comparison of conventional linear regression methods and neural networks for forecasting educational spending". *Economics of Education Review* Vol.18, pp. 405-415.
- Baker, B. D. (2001): "Can flexible non-linear modeling tell us anything new about educational productivity?". *Economics of Education Review* 20 (2001) 81-92.
- Bardhan, I. R., Cooper, W. W. y Kumbhakar, S. C. (1998): "A Simulation Study of Joint Uses of Data Envelopment Analysis and Statistical Regressions for Production Function Estimation and Efficiency Evaluation". *Journal of Productivity Analysis*, 9, 248-278.

- Barrow, M. y Wagstaff, A. (1989): "Efficiency measurement in the public sector: An appraisal". *Fiscal Studies*, Vol. 10, N°1, pp 72-97.
- Battese, G.E. and Coelli, T.J. (1988), "Prediction of Firm-Level Technical Efficiencies With a Generalised Frontier Production Function and Panel Data", *Journal of Econometrics*, 38, 387-399.
- Battese, G.E. and Coelli, T.J. (1992), "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India", *Journal of Productivity Analysis*, 3, 153-169.
- Battese, G.E. and Coelli, T.J. (1993), "A Stochastic Frontier Production Function Incorporating a Model for Technical Inefficiency Effects", *Working Papers in Econometrics and Applied Statistics*, No.69, Department of Econometrics, University of New England, Armidale, pp.22.
- Battese, G. y Coelli, T. (1995): "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data". *Empirical Economics*, 20, 325-332.
- Baxt, W. G. Y White, H. (1995): "Bootstrapping confidence intervals for clinical input variable effects in a network trained to identify the presence of acute myocardial infarction", *Neural Computation*, 7, 624-638.
- Beaton, A. E. y otros (1996): "Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS)". TIMSS International Study Center. Chestnut Hill, MA, USA.
- Becker, G. (1964): *Human Capital: a Theoretical and Empirical Analysis with Special Reference to Education*. Columbia University Press. New York.
- Beltratti, A., Margarita, S., Terna, P. (1996): *Neural Networks for economic and financial modelling*. International Thomson. Computer Press.
- Bessent, A. M. y Bessent, E. W. (1980): "Determining the comparative efficiency of schools through data envelopment analysis", *Educational Administration Quarterly*, 16, 57-75.
- Betts, J. R. (1995): "Does school quality matter? Evidence from the National Longitudinal Survey of Youth", *The Review of Economics and Statistics*, 77(2), pp. 231-247.
- Berg, I. (1970): *Education and jobs: The great training robbery*. NY: Praeger, 1970.
- Betts, J. R. y Shkolnik, J. L. (2000): "The effects of ability grouping on student achievement and resource allocation in secondary schools". *Economics of Education Review*, 19, pp. 1-15.
- Bifulco, R. y Bretschneider, S. (2001): "Estimating school efficiency: A comparison of methods using simulated data", *Economics of Education Review*, 20, 417-429.
- Bishop, C. M. (1996): *Neural Networks for Pattern Recognition*, Oxford: Oxford University Press.

- Bjurek, H. Hjalmarsson, L. y Forsund, F. R. (1990): "Deterministic parametric and nonparametric estimation of efficiency in service production: A comparison". *Journal of Econometrics*, 46 (1/2): 213-227.
- Bowles, S. y Levin, H. M. (1968a): "The determinants of scholastic achievement: an appraisal of some recent evidence". *The Journal of Human Resources* 3 (1), 3-24.
- Bowles, S. y Levin, H. M. (1968b): "More on multicollinearity and the effectiveness of schools". *The Journal of Human Resources* 3 (3): 393-400.
- Bryk, A. S., Raudenbush, S. W. (1992): *Hierarchical Linear Models: Applications and data Analysis Methods*. Newbury Park, CA: Sage.
- Burkhead, J., Fox, T. G. y Holland, J. W. (1967): *Input-Output in Large City High Schools*. Syracuse, NY: Syracuse University Press.
- Callan, S. y Santerre, R. (1990): "The production characteristics of local public education: a multi-product and input approach". *Southern Economic Journal*, 57, 468-480.
- Card, D. y Krueger, A. B. (1992): "Does school quality matter? Return to education and the characteristics of public schools in the United States. *Journal of Public Economics*, 100(1), pp. 1-40.
- Card, D. y Krueger, A. B. (1996): "Labor Market Effects of School Quality: Theory and Evidence," in Gary Burtless, ed., *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success* (Washington, DC: Brookings Institution, 1996).
- Carroll, J. (1963): "A model of school learning". *Teachers College Record*, 64 723-733.
- Castejón, J. L. (1994): "Estabilidad de diversos índices de eficacia de centros educativos". *Revista de Investigación Educativa*, 24, 45-60.
- Caulkins, J., Cohen, J. Gorr, W. L. y Wei, J. (1996): "Predicting Criminal Recidivism: Comparison of Neural Network Models with Statistical Methods" *Journal of Criminal Justice*, Vol. 24, No. 3, 1996.
- Clark, D. (2001): "Planning for the 2001 census of the United Kingdom". Dirección del censo del Reino Unido. Mimeo.
- Coakley, J. R. y Brown, C. E. (2000): "Artificial Neural Networks in Accounting and Finance: Modeling Issues". *International Journal of Intelligent Systems in Accounting, Finance & Management*, 9, 119-144.
- Cobb, C.W. y Douglas, P. H. (1928), "A Theory of Production", *American Economic Review*, 18, 139-172.
- Coelli, T. (1995): "Estimators and hypothesis tests for a stochastic frontier function: A Monte Carlo analysis". *Journal of Productivity Analysis*, 6, 247-268.
- Coelli, T. (1996a), *A Guide to DEAP Version 2.1: A Data Envelopment Analysis Program*, Centre for Efficiency and Productivity Analysis (CEPA), Working Paper 96/08.

- Coelli, T. (1996b), A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation, Centre for Efficiency and Productivity Analysis (CEPA), Working Paper 96/07.
- Coelli, T., D.S. Prasada Rao y G.E. Battese (1998), An introduction to efficiency and productivity analysis, Kluwer Academic Publishers.
- Coleman, J. S. et al. (1966): Equality of educational opportunity. Washington, DC: U.S. GPO.
- Comber, L. C. and Keeves, J. P. (1973): "Science education in nineteen countries: An empirical study". Stockholm: Almqvist and Wiksell.
- Costa, A y Markellos, R (1997): "Evaluating public transport efficiency with neural networks models". Transportation Research. Part C. Vol. 5C, N°5. October.
- Cushing, W. W. y Arguea, N. M. (1997): "Neural network Analysis of the Employee Clasification Problem for Tax Purposes". Documento de Trabajo N° 9701. Instituto Complutense de Análisis Económico.
- Cybenko, G. (1988): "Continuous valued neural networks with two hidden layers are sufficient". Technical Report, Dept. of Computer Science, Tufts University.
- Charnes, A., Cooper, W. y Rhodes, E. (1978): "Measuring the Efficiency of Decision Making Units". *European Journal of Operational Research*. Vol. 2 (6), pp. 429-444.
- Charnes, A., Cooper, W. y Rhodes, E. (1979): "Short communication: Measuring the Efficiency of Decision Making Units". *European Journal of Operational Research*. Vol. 3 (4), pp. 339.
- Charnes, A., Cooper, W. y Rhodes, E. (1981): "Evaluating program and managerial efficiency: An application of Data Envelopment Analysis to Program Follow Through". *Managerial Science*. 27 (6), pp. 668-697.
- Cheng, B. y D.M. Titterington (1994): "Neural Networks: a Review from a Statistical Perspective", *Statistical Science*, 9 (1), 2-54.
- Christensen, L. R., Jorgensen, D. W. y Lau, L. J. (1971): "Transcendental logarithmic production frontiers", *Review of Economics and Statistics*, 55, 28-55.
- Darlington, R. B. (1990): Regression and Linear Models. Mc Graw-Hill.
- De Veaux, R.D., J. Schumi, J. Schweinsberg y L.H. Ungar (1998), "Prediction Intervals for Neural Networks via Nonlinear Regression", *Technometrics*, 40 (4), 273-282.
- Debertin, D. L. (1976): "Estimating education production functions in rural and urban areas". *South Agricultural Economics*, 0, 31-35.
- Debreu, G., The Coefficient of Resource Utilization, *Econometrica* 19(3), July 1951.
- Deller, S. C. y Rudnicki, E. (1993): "Production Efficiency in Elementary Education: The Case of Maine Publics Schools". *Economics of Education Review* Vol.12. Pp. 45-57.
- Dewey, J., Husted, T. A. y Kenny, L. W. (2000): "The ineffectiveness of school inputs: a product of misspecification?". *Economics of Education Review* 19, 27-45.

- Donoho, D. L. y Johnstone, I. M. (1995): "Adapting to unknown smoothness via wavelet shrinkage", *Journal of the American Statistical Association*, 90, 1200-1224.
- Efron, B. y Tibshirani, R. (1993): *An introduction to the bootstrap*. Chapman & Hall, New York.
- Eide, E. y Showalter, M. H. (1998): "The effect of school quality on student performance: A quantile regression approach". *Economics Letters* 58, pp. 345-350.
- Eubank, R. L. (1999): *Nonparametric Regression and Spline Smoothing*, 2nd ed., Marcel Decker. New York.
- Färe, R., Grosskopf, S. y Lovell, C. A. K. (1994): *Production Frontiers*. Cambridge University Press. New York.
- Färe, R., Grosskopf, S. y Weber, W. L. (1989): "Measuring school district performance". *Public Finance Quarterly*, 17 (4): 409-428.
- Farrell, M.J. (1957): "The measurement of efficiency productive". *Journal of the Royal Statistical Society*, serie A, 120, pp. 253-266.
- Fausett, L. (1994): *Fundamentals of Neural Networks*, Englewood Cliffs, NJ: Prentice Hall.
- Figlio, D.N. (1999): "Functional form and the estimated effects of school resources". *Economics of Education Review* Vol.18. Pp. 241-252.
- Fine, T. L. (1999): *Feedforward Neural Networks Methodology*. Statistics for Engineering and Information Science. Springer. New York.
- Fried, H. O., Lovell, C. A. K. and Schmidt, S. S. (1993): *The Measurement of Productive Efficiency*. Oxford University Press, Oxford.
- Friedman, J. H. y Stuetzle, W. (1981): "Projection Pursuit Regression", *Journal of the American Statistical Association*, 76, 817-823.
- Funahashi, K. (1989): "On the approximate realization of continuous mappings by neural networks", *neural networks 2*: 183-192.
- Galache, T. (1994): "Redes neuronales artificiales: Análisis de un nuevo método de modelización aplicable en economía". Tesis Doctoral. Universidad de Málaga.
- Gallant, A. R. (1987): *Nonlinear Stastical Models*, NY: Wiley.
- Garson, G. D. (1991): "Interpreting neural network connection weights", *AI Expert*, April 1991, 47-51.
- Gedeon, T. D. (1997): "Data mining of inputs: Analysing magnitude and functional measures", *International Journal of Neural Systems*, 8, 209-218.
- Geman, S., E. Bienenstock y R. Doursat (1992), "Neural Networks and the Bias/Variance Dilemma", *Neural Computation*, 4, 1-58.
- Glenn, C. L. (1989): "Choice of School in Six Nations: France, Netherlands, Belgium, UK, Canada, West-Germany" (Washington, D. C., US Department of Education).

- Glennester, H. (1991): "Quasi-Markets for Education?". *The Economic Journal*, Vol. 101 (Sep.), pp. 1268-1276.
- Glewwe, P. (1997): "Estimating the Impact of Peer Group Effects on Socioeconomic Outcomes: Does the Distribution of Peer Group Characteristics Matter?". *Economics of Education Review* Vol.16, pp. 39-43.
- González, F. (1996): "Análisis de la volatilidad en opciones financieras: una variable fundamental". Tesis Doctoral. Universidad Autónoma de Madrid.
- González, E. J. y Smith, T. A. (eds.). (1997): *User Guide for the TIMSS International Database. Primary and Middle School Years*. IEA. Boston College. Chestnut Hill, MA, USA.
- Goldstein, H. (1987): *Multilevel Models in Educational and Social Research*. New York. Oxford University Press.
- Goldstein, H. (1995): *Multilevel Statistical Models*, (2ª Edición), London. Edward Arnold.
- Goleman, D. (1995): *Emotional Intelligence*. New York. Bantam Books.
- González, J. E. y Smith, T. A. (Eds.). (1997): *User Guide for the TIMSS International Database. Primary and Middle School Years. (Population 1 and Population 2). Suplementos incluidos*. TIMSS International Study Center. Boston College. Chestnut Hill, MA, USA.
- González Miranda, F. (1996): "Análisis de la volatilidad en opciones financieras: una variable fundamental". Tesis Doctoral. Universidad Autónoma de Madrid.
- Gorr, W. L., Nagin, D. y Szczypula, J. (1994): "Comparative study of artificial neural network and statistical models for predicting student grade point averages". *International Journal of Forecasting* 10. pp. 17-34.
- Gray, J., Jesson, D. y Jones, B. (1986): "The search for a fairer way of comparing schools examinations results", *Research Papers in Education*, 1, 91-122.
- Greenwald, R., Hedges, L. y Laine, R. D. (1996): "The effect of school resources on student achievement". *Review of Educational Research*, 66, 361-396.
- Greenwald, R., Hedges, L. y Laine, R. D. (1996): "Interpreting research on school resources and student achievement: A rejoinder to Hanushek". *Review of Educational Research*, 66, 411-416 .
- Grunfeld, Y. y Griliches, Z. (1960): "Is aggregation necessarily bad?". *Review of Economics and Statistics* 42, 1-13.
- Guermat, C. y Hadri, K. (1999): "Backpropagation Neural Network Versus Translog Model in Stochastic Frontiers: A Monte Carlo Comparison". Discussion Paper in Economics N° 99/16, University of Exeter.
- Haerdle, W. (1990): *Applied Nonparametric Regression*, Cambridge University Press.
- Hammond, C. (1986): "Estimating the statistical cost curve: An application of the stochastic frontier technique". *Applied Economics*, 18: 971-984.

- Hanushek, E. A. (1971): "Teacher characteristics and gains in student achievement: estimation using micro data". *American Economics Review*, 61, Mayo: 280-288.
- Hanushek, E. A. (1979): "Conceptual and Empirical issues in the estimation of educational production functions". *Journal of Human Resources*, 14, 351-388.
- Hanushek, E. A. (1981): "Throwing money at schools". *Journal of Policy Analysis and Management*, 1, 19-41.
- Hanushek, E. A. (1986): "The economics of Schooling". *Journal of Economic Literature*. Vol 24, n°3, pp. 1141-1171.
- Hanushek, E.A. (1997): "Assesing the effects of school resources on student performance: An update". *Educational Evaluation and Policy Analysis*. 19, 141-164.
- Hanushek, E. A., Rivkin, S. G. y Taylor, L. L. (1996): "Aggregation and the estimated effects of school resources". *The Review of Economics and Statistics*. November 1996. 78 (4), 611-627.
- Hashem, S. (1993): "Optimal Linear Combinations of Neural Networks". Doctoral Thesis. Purdue University.
- Hashem, S. y Schmeiser, B. (1995): "Improving model accuracy using optimal linear combinations of trained neural networks". *IEEE Transactions on Neural Networks*, 6(3):792-794.
- Hastie, T. R. y Tibshirani, R. (1990): *Generalized Additive Models*. Monographs on Statistics and Applied Probability 43, Chapman and Hall, New York.
- Hayken, S. (1999): *Neural Networks: A Comprehensive Foundation*. (2nd Edition), New York, Prentice Hall, 1999.
- Hebb, D.O. (1949): *The organization of behavior*. Science Editions, New York, NY.
- Hecht-Nielsen, R. (1989), "Kolmogorov's Mapping Neural Network Existence Theorem", *International Joint Conference on Neural Networks*, 3, 11-14, Washington.
- Hecht-Nielsen, R. (1990): *Neurocomputing*. Addison-Wesley, Reading, MA, 1990.
- Hertz, J., Krogh, A. y Palmer, R.G. (1991). *Introduction to the Theory of Neural Computation*. Addison Wesley.
- Hill, T., Marquez, L., O'Connor, M. y Remus, W. (1994): "Artificial neural network models for forecasting and decision making". *International Journal of Forecasting* 10. pp. 5-15.
- Hopfield, J. (1982): "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of the National Academy of Sciences*, Volume 79, 2554-2558. April 1982.
- Hornik, K., Stinchcombe, M. y White, H. (1989): "Multilayer Feed-forward Networks are Universal Approximators". *Neural networks* 2, pp. 359-66.

- Hornik, K., Stinchcombe, M. and White, H. (1990). "Universal Approximation of an Unknown Mapping and its Derivatives using Multilayer Feed-forward Networks". *Neural Networks* 3, pp. 551-60.
- Hoxby, C. M. (1999): "The productivity of schools and other local public goods producers". *Journal of Public Economics*, 74, 1-30.
- Hoxby, C. M. (2001): "School Choice and School Productivity (or Could School Choice be a Tide that Lifts all Boats?). Conference on the economics of school choice. February 22-24. Cheeca Lodge. Islamorada, FL.
- Husén, T. ed. (1967): "International study of achievement in mathematics". Vol. II. New York: John Willey and Sons.
- Hwang, J. T. y Ding, A. A. (1997): "Prediction Intervals for Artificial Neural Networks". *Journal of the American Statistical Association*, 92(438), 748-757.
- Ivakhnenko, A. G. y Ivakhnenko, G. A. (1995): "The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (GMDH)". *Pattern Recognition and Image Analysis*, Vol. 5, Nº 4, pp. 527-535.
- Jenks, C. S. (1972): "Inequality: A reassessment of the effects of family and schooling in America". NY: Basic Books, 1972.
- Jondrow, J., Lovell, C.A.K., Materov, I. S. y Schmidt, P. (1982): "On the estimation of technical inefficiency in the stochastic frontier production function model". *Journal of econometrics* 19 págs. 233-238.
- Jordan, M.L. (1986) "An introduction to linear algebra in Paralell Distributed Processing". En Rumelhart, DE. ,McClelland, JL. (eds.) (1986) "Parallel Distributed Processing. Explorations in the Microstructure of Cognition". Cambridge, MA: MIT Press.
- Kohonen, T. (1984): *Self-Organization and Associative Memory*, Berlin. Springer.
- Koopmans, T. C. (1951): "An analysis of production as an efficient combination of activities" in T. C. Koopmans, ed., *Activity Analysis of Production and Allocation*, Cowles Commission for Research in Economics, Monograph Nº 13. New York: John Willey and Sons, Inc.
- Krueger, A. B. (1999): "Experimental Estimates of Educational Production Functions". *Quarterly Journal of Economics*, 114, Nº 2, 497-532.
- Krueger, A. B. (2000): "Economic Considerations and Class Size". Working Paper Nº 447. Princeton University.
- Kuan, C. and White, H. (1994): "Artificial neural networks: an econometrics perspective". *Econometrics Review*, 13 (1): 1-91.
- Kuan, C.M. y T. Liu, (1995) "Forecasting Exchange Rates Using Feedforward and Recurrent Neural Networks", *Journal of Applied Econometrics*, 10, 347-364.
- Ladd, H. F. y Walsh, R. P. (2002): "Implementing value-added measures of school effectiveness: getting the incentives right". *Economics of Education Review*, 21, 1-17.

- Lee, J. y Barro, R. J. (1998): "Schooling quality in a cross section of countries". Working Paper 6198. NBER.
- Lee, T., White, H. and Granger, C. W. J. (1993): "Testing for neglected nonlinearity in time series models: a comparison of neural networks methods and alternative tests". *Journal of Econometrics*, 56 (3): 269-91.
- Leibenstein, H. (1966): "Allocative efficiency and x-efficiency". *American Economic Review*, 56, pp. 392-495.
- Levin, H. M. (1974): "Measuring efficiency in educational production". *Public Finance Quarterly*. 2. 3-24.
- Levin, H. M. (1994): "Economics of school reform for at risk students", in Hanushek et al (eds.): *Making Schools Work: Improving Performance and Controlling Costs* (Washington, D. C., The Brookings Institution).
- Levin, H. M. (1997): "Raising School Productivity: An X-Efficiency Approach". *Economics of Education Review* Vol.16. Pp. 303-311.
- Liao, T. F. (1992): "A modified GMDH approach for social science research: exploring patterns of relationships in the data". *Quality and Quantity*, 26, 19-38.
- Lindley, D. V. y Smith, A. F. M. (1972): "Bayes estimates for the linear model". *Journal Royal Statistics Society B*, 34, 1-44.
- Little, R. J. A. y Rubin, D. B. (1987): *Statistical Analysis with Missing Data*. John Wiley & Sons. New York.
- López Varona, J.A. y Moreno Martínez, M^a.L. (1997a): "Resultados en Matemáticas. Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS)". Secretaría General de Educación y Formación Profesional. MEC.
- López Varona, J.A. y Moreno Martínez, M^a.L. (1997b): "Resultados en Ciencias. Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS)". Secretaría General de Educación y Formación Profesional. MEC.
- Lovell, C. A. (1993): "Production frontiers and productive efficiency". In *The Measurement of Productive Efficiency*, ed. H. O. Fried, C. A. Lovell and S. S. Schmidt, pp. 3-67. Oxford University Press, Oxford.
- Madaus, G. F., Afrasian, P. W. Y Kellaghan, F. (1980). *School effectiveness: a reassessment of evidence*. New York. Mc Graw-Hill.
- Mas-Colell, A., Whinston, M., y Green, J. (1995): *Microeconomic Theory*, Oxford University Press.
- Mancebón, M.J. (1996): "La evaluación de la eficiencia de los centros educativos públicos". Tesis doctoral de la Universidad de Zaragoza.

- Mc Carty, T. y Yaisawarng, S. (1993): "Technical efficiency in New Jersey school districts". In Fried, H., Lovell, C. A. K. y Schmidt, S, *The measurement of productive efficiency* (pp. 271-287). Oxford: Oxford University Press.
- Mc Culloch, W.S. y Pitts, W.A. (1943): "A logical Calculus of the Ideas Immanent in Nervous Activity". *Boullletin of Mathematics and Biophysics*, 5, págs. 115-133.
- Meeusen, W. and van den Broeck, J. (1977), "Efficiency Estimation from Cobb-Douglas Production Functions With Composed Error", *International Economic Review*, 18, 435-444.
- Mestre, J. M., Guil, M^a. R, Carreras, M^a. R. y Braza, P. (2000): "Cuando los constructos psicológicos escapan del método científico: el caso de la inteligencia emocional y sus implicaciones en la validación y evaluación". *Revista Electrónica de Motivación y Emoción*, Vol. 3, N^o 4.
- Meyer, R.H. (1997): "Value-Added Indicators of School Performance: A Primer". *Economics of Education Review* Vol.16. Pp. 283-301.
- Minsky, M. y Papert, S. (1969): *Perceptrons: An introduction to Computacional Geometry*. The MIT Press. Cambridge, MA.
- Modrego, A.; San Segundo, M.J. (1988): "Ecuaciones de rendimiento escolar para la evaluación de las enseñanzas medias". *Revista de Educación*. Septiembre-Diciembre 1988. pp 147-179.
- Morris, C. N. (1995): "Hierarchical models for educational data: An overview". *Journal of Educational and Behavioral Statistics*, N^o 20, 201-204.
- Murillo, F. J. (1999): "Los Modelos Jerárquicos Lineales aplicados a la Investigación sobre Eficacia Escolar". *Revista de Investigación Educativa*, 17(2), 453-460.
- Muñiz, M. A. (2000): "Eficiencia técnica e inputs no controlables. El caso de los institutos asturianos de educación secundaria". Tesis Doctoral. Universidad de Oviedo.
- Núñez, C. (1998): "La construcción de una red neuronal para el análisis de riesgos en las entidades financieras". Tesis Doctoral. Universidad de Sevilla.
- O'Leary, D. E. (1998): "Using neural networks to predict corporate failure", *International Journal of Intelligent Systems in Accounting, Finance and Management*, 7: 187-197.
- Odom, M. D. y Sharda, R. (1993): "A Neural Network Model for Bankruptcy Prediction". In Trippi, R. y Turban, E. (eds.) (1993): *Neural Networks in Finance and Investing: using artificial intelligence to Real World Performance*. Probus Publishing Company, Chicago.
- Paass, G. (1993): "Assessing and improving neural networks predictions by the bootstrap algorithm", in Hanson, S., Cowan, C. L., Giles, eds., *Advances in Neural Information Processing Systems* 5, Morgan Kaufmann Pub., 196-203.
- Parra, E. (1994): "Predicción de audiencia de televisión mediante redes neuronales". Tesis Doctoral. Universidad de Málaga.

- Pedraja, F. y Salinas, J. (1996): "Eficiencia del gasto público en educación secundaria: una aplicación de la técnica envolvente de datos". Hacienda Pública Española Nº138, pp. 87-96.
- Pedraja, F., Salinas, J. y Smith, P. (1997): "On the role of weight restrictions in data envelopment analysis". Journal of Productivity Analysis, 8, 215-230.
- Plutowski, M., Sakata, S. y White, H. (1994): "Cross-validation estimates IMSE", in Cowan, J. D., Tesauro, G. y Alspector, J. (eds.). Advances in Neural Information Processing Systems 6, San Mateo, CA: Morgan Kaufman, pp. 391-398.
- Postlethwaite, T. N. y Wiley, D. E. (1992): "The IEA study of science II: Science achievement in twenty-three countries". Pergamon Press. Oxford.
- Pritchett, L. y Filmer, D. (1999): "What education production functions really show: a positive theory of education expenditures". Economics of Education Review Vol.18. Pp. 223-239.
- Quinlan,(1991). Connectionism and Psychology. Harvester Wheateaf. N.Y.
- Ramón y Cajal, S. (1906): The structure and connexions of neurons. Reprinted in Nobel Lectures: Physiology or Medicine, 1901-1921, New York: Elsevier, 1967, pp. 220-253.
- Ray, S. C. (1991): "Resource-use efficiency in public schools: A study of Connecticut data", *Management Science*, 37(12): 1620-1629.
- Riley, J. G. (1979): "Testing the Educational Screening Hypothesis". Journal of Political Economics, Oct. Pt II, 87(5), pp. 227-52.
- Rivals, I y Personnaz, L. (2000): "Construction of Confidence Intervals for Neural Networks Based on Least Squared Estimation". Neural Networks, 13, 463-484.
- Ripley, B. D. (1996): Pattern Recognition and Neural Networks, Cambridge: Cambridge University Press.
- Robinson, W. S. (1950): "Ecological correlations and the behaviour of individuals". American Sociological Review, Vol. 15, 351-357.
- Robitaille, D. and Garden, R. (eds.). (1989): "The IEA study of mathematics II: Context and outcomes of school mathematics". Vol. II. Pergamon. Oxford.
- Romay, R. y Santín, D. (2002): "El Control del Gasto Público por Incapacidad Temporal mediante Redes Neuronales". Hacienda Pública Española / Revista de Economía Pública. (En proceso de revisión).
- Roseblatt, F. (1958): "The Perceptron: A probabilistic model for information storage and organization in the brain". Psychological Review 65, págs 386-408.
- Rumelhart, D. E.; Mc Clelland, J.L.; and the PDP Research Group (Eds.) (1986): "Parallel Distributed Processing: Explorations on the Microstructure of Cognition, vol.I. Foundations", MIT Press, Cambridge, MA.

- Saeed, M. y Cameron, N. (2000): "Neural Network Versus Econometric Models in Forecasting Inflation". *Journal of Forecasting*, 19, 201-217.
- Salas, M. y Jiménez, J.D. (1998): "Análisis microeconómico de la elección de carrera universitaria". VII Jornadas de la Asociación de la Economía de la Educación. pp. 249 – 260. Universidad de Cantabria.
- Salovey, P. y Mayer, J. D. (1990): "Emotional Intelligence". *Imagination, Cognition and Personality*, 9, 185-211.
- Samdani, G. (1990): "Neural Nets: They Learn From Examples" *Chemical Engineering*, vol. 97(8), pp.37-45.
- San Segundo, M.J. (1991): "Evaluación del Sistema Educativo a Partir de Datos Individuales". *Economía Industrial*, Marzo - Abril 1991, pp. 23-37.
- San Segundo, M.J. (1998): "Igualdad de oportunidades educativas". *Ekonomiaz* N° 40 1.er Cuatrimestre.
- Santín, D, y Valiño, A. (2000): "Artificial Neural Networks for Measuring Technical Efficiency in Schools". II Oviedo Workshop on Efficiency and Productivity Analysis. Oviedo, Mayo 2000.
- Santín, D. Delgado, F. y Valiño, A. (2001): "Measuring Technical Efficiency with Neural Networks: A Review". *Efficiency Series Paper 09/2001*. Departamento de Economía. Universidad de Oviedo.
- Santín, D. (2001): "La necesidad de gasto en educación no universitaria en las Comunidades Autónomas: ¿existe igualdad de oportunidades educativas?". *Hacienda Pública Española* 159-4, 217-250.
- Sarle, W. S., ed. (2000): *How to measure importance of inputs?*. SAS Institute Inc., Cary, NC, USA. URL: <ftp://ftp.sas.com/pub/neural/importance.html>
- Sarle, W. S., ed. (2002): *Neural Network FAQ*, periodic posting to the Usenet newsgroup comp.ai.neural-nets, URL: <ftp://ftp.sas.com/pub/neural/FAQ.html>
- Scarselli, F. y A. Chung, (1998), "Universal Approximation Using Feedforward Neural Networks: a Survey of Some Existing Methods, and Some New Results", *Neural Networks*, 11 (1), 15-37.
- Scheerens, J. (1992): "Conceptual models and theory-embedded principles on effective schooling". *School Effectiveness and School Improvement*, 8(3), 269-310.
- Schwartz, G. (1978): "Estimating the dimension of a model". *Annals of Statistics*, 6; 461-464.
- Sengupta, J. K. y Sfeir, R.E (1986): "Production frontier estimates of scale in public schools in California". *Economics of Education Review*, 5, pp. 297-307.
- Serrano, C. (1994): "Las Redes Neuronales Artificiales en el Análisis de la Información Contable". Tesis del Departamento de Contabilidad y Finanzas de la Universidad de Zaragoza, noviembre de 1994.

- Sherrington, C. (1906): *The Integrative Action of the Nervous System*. New York: Oxford University Press.
- Spence, M. (1973): "Job Market Signaling". *Quarterly Journal of Economics*. Vol. 87, pp. 1-17.
- Stiglitz, J. E. (1975): "The Theory of Screening, Education and the Distribution of Income". *American Economic Review*. Vol. 65. N° 3.
- Stone, M. (1974): "Cross-validatory choice and assessment of statistical predictions". *Journal of Royal Statistical Society, B* 36, 111-147.
- Stone, M. (1977): "Asymptotics for and against cross-validation". *Biometrika*, 69, 29-35.
- Summers, A. A. y Wolfe, B. L. (1977): "Do Schools make a difference?". *American Economic Review* Vol. 67 N°4. (Sept.).
- Tejedor, F. J. y Caride, J. A. (1988): "Influencia de las variables contextuales en el rendimiento académico". *Revista de Educación*. Septiembre-Diciembre 1988. pp 112-146.
- Thanassoulis, E. y Dustan, P. (1994): "Guiding Schools to Improved Performance Using Data Envelopment Analysis: An illustration with Data from a Local Education Authority". *Journal of Operational Research Society*, Vol. 45, 1247-1262.
- Thanassoulis, E. y Silva Portela, M. C. (1999): *Decomposing pupil under attainment into pupil and school components*. Aston Business School Research Institute, Junio 1999.
- Theil (1954): *Linear aggregation of economic relations*. North-Holland. Amsterdam.
- Trippi, R. y Turban, E. (eds.). (1993): *Neural Networks in Finance and Investing: using artificial intelligence to Real-World Performance*. Probus Publishing Company. Chicago.
- Vandenberghe, V. (1999): "Economics of Education. The Need to go Beyond Human Capital Theory and Production-Function Analysis". *Educational Studies*, Vol. 25, N° 2.
- Vapnik, V. N. (1998): *Statistical Learning Theory*. Wiley, New York.
- Vellido, A., Lisboa, P. J. G. y Vanghan, J. (1999): "Neural networks in business: a survey of applications (1992-1998)". *Expert Systems with applications*, 17: 51-70.
- Waltz, D. y Feldman, J.A. (1988): "Connectionist Models and Their Implications". In Waltz, D. & Feldman, J.A. (eds.). *Connectionist Models and Their Implications* Norwood, NJ: Ablex Publishing.
- Wassermann, P.D. (1989). *Neural computing: Theory and Practice*. VNR. New York.
- Werbos, P.J. (1974): "Beyond Regression: New tools for Prediction and Analysis in the Behavioral Sciences". Unpublished Ph.D. Dissertation, Harvard University. Department of Applied Mathematics.
- White, H. (1990): "Connectionist Nonparametric Regression: Multilayer Feedforward Networks Can Learn Arbitrary Mappings". *Neural Networks*, 3, 535-550.
- Widrow, B. (1959): "Adaptative sampled-data systems, a statistical theory of adaptation". 1959 IRE WESCON Convention Record, part 4. New York: Institute of Radio Engineers.

-
- Wong, F. S., Bodnovich, T. A. y Selvi, Y. (1997): "Neural network applications in business. A review and analysis of the literature (1988-1995)", *Decision Support Systems*, 19: 301-320.
- Worthington, A. C. (2001): "An Empirical Survey of Frontier Efficiency Measurement Techniques in Education". *Education Economics*, Vol. 9, N° 3.
- Zapranis, A. y Refenes, A-P. (1999): *Principles of Neural Model Identification, Selection and Adequacy with Applications to Financial Econometrics*, Springer.
- Zhang, Y. and Bartels, R. (1998): "The Effect of Sample Size on the Mean Efficiency in DEA with an Application to Electricity Distribution in Australia, Sweden and New Zealand". *Journal of Productivity Analysis*, 9, 187-204.
- Zhang, G., Patuwo, B. E. y Hu, M. Y. (1998): "Forecasting with artificial neural networks: The state of the art", *International Journal of Forecasting* 14; 35-62.

ANEXOS

ANEXO 1: Etiquetas y Códigos de las Variables de Alumno creadas a partir del fichero original del TIMSS para los datos de España: bsgcode.cdt

NOMBRE UTILIZADO	NOMBRE ORIGINAL	CODIGO	SIGNIFICADO
MATEMÁTICAS (Variable Dep.) internacional*	BIMATSCR		*Resultado en matemáticas comparable a nivel internacional*
VALOR Número con dos decimales			
1 SEXO	ITSEX		*Sexo del Estudiante*
		0	FEMENINO
		1	MASCULINO
2 CASTELLANO	ITLANG		*Lengua utilizada en el Test*
		0	OTRA
		1	CASTELLANO
3 GALLEGO	ITLANG		*Lengua utilizada en el Test*
		0	OTRA
		1	GALLEGO
4 CATALAN	ITLANG		*Lengua utilizada en el Test*
		0	OTRA
		1	CATALAN
<i>*La Categoría de Referencia para ITLANG es el Valenciano</i>			
5 NACESPAÑA	BSBGBRN1		*¿Nació el alumno en España?*
		0	SI
		1	NO
6 LENGUACASA	BSBGLANG		*¿Con qué frecuencia hablas en casa la lengua en la que has hecho el test?*
		1	Siempre o casi Siempre
		0	A veces o nunca
7 MADRE	BSBGADU1		*¿Vive tu madre contigo todo o la mayoría del tiempo?*
		0	SI
		1	NO
8 PADRE	BSBGADU2		*¿Vive tu padre contigo todo o la mayoría del tiempo?*
		0	SI
		1	NO

9 HERMANOS	BSBGADU3	*¿Vives con hermano(s) todo o la mayoría del tiempo?*
	0	SI
	1	NO

10 HERMANAS	BSBGADU4	*¿Vives con hermana(s) todo o la mayoría del tiempo?*
	0	SI
	1	NO

11 MADRASTRA	BSBGADU5	*¿Vive tu madrastra contigo todo o la mayoría del tiempo?*
	0	SI
	1	NO

12 PADRASTRO	BSBGADU6	*¿Vive tu padrastro contigo todo o la mayoría del tiempo?*
	0	SI
	1	NO

13 ABUELO	BSBGADU7	*¿Vive(n) algún/varios abuelo(s) contigo todo o la mayoría del tiempo?*
	0	SI
	1	NO

14 PARIENTES	BSBGADU8	*¿Vive(n) otro(s) pariente(s) contigo todo o la mayoría del tiempo?*
	0	SI
	1	NO

15 OTROS	BSBGADU9	*¿Vive(n) otra(s) persona(s) no pariente(s) contigo todo o la mayoría del tiempo?*
	0	SI
	1	NO

16 NUMTOTAL	BSBGHOME	*¿Cuántas personas viven en tu casa?*
	VALOR	Núm. Personas

17 MADRESP	BSBGBRNM	*¿Nació tu madre en España?*
	0	SI
	1	NO

18 PADRESP	BSBGBRNF	*¿Nació tu padre en España?*
	0	SI
	1	NO

19 LIBROSCASA BSBGBOOK *¿Cuántos libros hay en tu casa aproximadamente?

- 1 0-10 libros
 - 2 11-25 libros
 - 3 26-100 libros
 - 4 101-200 libros
 - 5 Más de 200 libros
-

20 CALCULADORA BSBGPS01 *¿Tienes calculadora en casa?*

- 0 SI
 - 1 NO
-

21 ORDENADOR BSBGPS02 *¿Tienes ordenador en casa?*

- 0 SI
 - 1 NO
-

22 ESCRITORIO BSBGPS03 *¿Tienes escritorio en casa?*

- 0 SI
 - 1 NO
-

23 DICCIONARIO BSBGPS04 *¿Tienes diccionario en casa?*

- 0 SI
 - 1 NO
-

24 ENCICLOPEDIA BSBGPS05 *¿Tienes enciclopedia en casa?*

- 0 SI
 - 1 NO
-

25 LIBRMATES BSBGPS06 *¿Tienes libros de matemáticas en casa?*

- 0 SI
 - 1 NO
-

26 LIBRCIENCIAS BSBGPS07 *¿Tienes libros de ciencias en casa?*

- 0 SI
 - 1 NO
-

27 LABINSTRUM BSBGPS08 *¿Tienes instrumentos de laboratorio en casa?*

- 0 SI
 - 1 NO
-

28 VIDEOS EDUC	BSBGPS09	*¿Tienes videos educativos en casa?*	0 SI 1 NO
----------------	----------	--------------------------------------	--------------

29 PROGORDEDUC	BSBGPS10	*¿Tienes programas de ordenador educativos en casa?*	0 SI 1 NO
----------------	----------	--	--------------

30 LUGESTUDPROP	BSBGPS11	*¿Tienes un lugar de estudio propio en casa?*	0 SI 1 NO
-----------------	----------	---	--------------

31 MADMEJCLASE	BSBMIP6	*Mi madre piensa que es importante para mí estar entre los mejores de la clase	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
----------------	---------	--	--

32 NOTRABAJAN	BSBMCLS1	*En mi clase de matemáticas mis compañeros a menudo no hacen sus tareas*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
---------------	----------	--	--

33 SEPORTANBIEN	BSBMCLS2	*En mi clase de matemáticas mis compañeros están en orden y callados*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
-----------------	----------	---	--

34 OBEDECEN	BSBMCLS3	*En mi clase de matemáticas mis compañeros hacen lo que el profesor dice*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
-------------	----------	---	--

35 AMIGMAT	BSBMFIP2	*La mayoría de mis amigos piensan que es importante para mí ser bueno en matemáticas*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
------------	----------	---	--

36 AMIGLENG	BSBGFIP3	*La mayoría de mis amigos piensan que es importante para mí ser bueno en lenguaje*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
-------------	----------	--	--

37 AMIGDIVERS	BSBGFIP4	*La mayoría de mis amigos piensan que es importante para mí tener tiempo para divertirme*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
---------------	----------	---	--

38 AMIGMEJCLASE	BSBGFIP6	*La mayoría de mis amigos piensan que es importante para mí ser de los mejores de la clase*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
-----------------	----------	---	--

39 OPBUENOMAT	BSBMSIP2	*Creo que ser bueno en matemáticas es importante para mí*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
---------------	----------	---	--

40 OPBUENODIVERS	BSBGSIP4	*Creo que tener tiempo para divertirme es importante para mí*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
------------------	----------	---	--

41 OPBUENOSPORT	BSBGSIP5	*Creo que ser bueno en deportes es importante para mí*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
-----------------	----------	--	--

42 OPBUENOCLASE	BSBGSIP6	*Creo que estar entre los mejores de mi clase es importante para mí*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
-----------------	----------	--	--

43 BUENOMATES	BSBMGOOD	*Habitualmente soy bueno en matemáticas*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
---------------	----------	--	--

44 GENETICO	BSBMDOW1	*Para obtener buenos resultados en matemáticas es importante tener talento natural*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
-------------	----------	---	--

45 SUERTE	BSBMDOW2	*Para obtener buenos resultados en matemáticas es importante tener suerte*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
-----------	----------	--	--

46 ESTDURO	BSBMDOW3	*Para obtener buenos resultados en matemáticas es importante estudiar duro*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
------------	----------	---	--

47 MEMORIZAR	BSBMDOW4	*Para obtener buenos resultados en matemáticas es importante memorizar los apuntes de clase*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
--------------	----------	--	--

48 GUSTANMATES	BSBMLIKE	*¿Te gustan las matemáticas?*	1 Me gustan o me gustan mucho 0 No me gustan o no me gustan nada
----------------	----------	-------------------------------	---

49 ORDCLMAT	BSBMCMLK	*¿Usas alguna vez el ordenador en la clase de matemáticas?	0 No uso ordenador nunca 0 A veces uso ordenador
-------------	----------	--	---

50 DIVERAPMAT	BSBMENJY	*Aprender matemáticas es divertido*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
---------------	----------	-------------------------------------	--

51 ABURRAPMAT	BSBMBORE	*Aprender matemáticas es aburrido*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
---------------	----------	------------------------------------	--

52 MATFACILES	BSBMEASY	*Las matemáticas son fáciles*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
---------------	----------	-------------------------------	--

53 MATIMPVIDA	BSBMELIFE	*Aprender matemáticas es importante en la vida de toda persona*	1 Muy de acuerdo o de acuerdo 0 En desacuerdo o muy en desacuerdo
---------------	-----------	---	--

54 LIKETRABMAT BSBMWORK *Me gustaría tener un trabajo que implique el uso de las matemáticas*

1 Muy de acuerdo o de acuerdo
0 En desacuerdo o muy en desacuerdo

55 OPMATIMPJOB BSBMJOB *Necesito ser bueno en matemáticas para alcanzar el trabajo deseado*

1 Muy de acuerdo o de acuerdo
0 En desacuerdo o muy en desacuerdo

56 BIENMATPAD BSBMPRNT *Necesito ser bueno en matemáticas para agradar a mis padres*

1 Muy de acuerdo o de acuerdo
0 En desacuerdo o muy en desacuerdo

57 BIENMATINSTI BSBMSCHL *Necesito ser bueno en matemáticas para entrar en el instituto que deseo*

1 Muy de acuerdo o de acuerdo
0 En desacuerdo o muy en desacuerdo

58 BIENMATSELF BSBMSELF *Necesito ser bueno en matemáticas para mi propia satisfacción*

1 Muy de acuerdo o de acuerdo
0 En desacuerdo o muy en desacuerdo

Dada la siguiente lista de problemas medioambientales. ¿Crees que la aplicación de la ciencia puede ayudarnos a solucionarlos?.

59 SCIAIRPOLUT BSBGENV1 *La contaminación del aire*

1 Mucho o bastante
0 Nada o muy poco

60 SCIAGUAPOLUT BSBGENV2 *La contaminación del agua*

1 Mucho o bastante
0 Nada o muy poco

61 SCIDESTBOSQ BSBGENV3 *La destrucción de los bosques*

1 Mucho o bastante
0 Nada o muy poco

62 SCIANIMEXTIN BSBGENV4 *La protección de animales en extinción*

1 Mucho o bastante
0 Nada o muy poco

63 SCIOZONO	BSBGENV4	*La destrucción de la capa de ozono*	1 Mucho o bastante 0 Nada o muy poco
-------------	----------	--------------------------------------	---

64 SCIPOWNUCPL	BSBGENV5	*Los problemas con las centrales nucleares*	1 Mucho o bastante 0 Nada o muy poco
----------------	----------	---	---

Fuera de la escuela ¿Con qué frecuencia dedicas tiempo a las siguientes actividades?

65 FRECLEOLIBRO	BSBGACT1	*Leer un libro o revista*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
-----------------	----------	---------------------------	---

66 FRECVISMUSEO	BSBGACT2	*Visitar un museo o exhibición de arte*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
-----------------	----------	---	---

67 FRECCONCIER	BSBGACT3	*Asistir a un concierto*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
----------------	----------	--------------------------	---

68 FRECTEATRO	BSBGACT4	*Asistir al teatro*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
---------------	----------	---------------------	---

69 FRECCINE	BSBGACT5	*Asistir al cine*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
-------------	----------	-------------------	---

Fuera de la escuela ¿Con qué frecuencia dedicas tiempo a ver en televisión o video los siguientes tipos de programas?

70 FRECNOTICIAS	BSBGNEWS	*Las noticias o documentales*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
-----------------	----------	-------------------------------	---

71 FRECOPERA	BSBGOPER	*Opera, ballet o música clásica*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
--------------	----------	----------------------------------	---

72 FRECNATURE	BSBGNATR	*Programas sobre la naturaleza, la vida salvaje o la historia*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
---------------	----------	--	---

73 FRECPOPMUSIC	BSBGPOPU	*Programas sobre música pop*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
-----------------	----------	------------------------------	---

74 FRECSPORT	BSBGSPORT	*Deportes*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
--------------	-----------	------------	---

75 FRECVIDEOJUE	BSBGVIDE	*Juegos de video*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
-----------------	----------	-------------------	---

76 FRECARTOONS	BSBGCRTN	*Dibujos animados*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
----------------	----------	--------------------	---

77 FRECAVCOSU	BSBGCMDY	*Películas de comedia, aventuras o suspense*	1 Casi todos los días o al menos una vez a la semana 0 Al menos una vez al mes o muy raramente
---------------	----------	--	---

78 EDAD	BSDAGE	*La edad del estudiante en años y meses*	
	VALOR	Número con un decimal	

79 NUMEROITEMS	BSDGPS	*¿Cuántos de los 11 elementos adecuados para el estudio posees en total?*	
	VALOR	Número	

80 DESKCOMPDIC BSDGPSA *¿Posees en tu casa un escritorio para estudiar, ordenador y diccionario?*

0 SI
1 NO

¿Cuál es el nivel de estudios de tu madre?

*La categoría de referencia es la respuesta “No lo sé”

81 PRIMMAD BSBGEDUM *Primarios*
1 SI
0 Otros

82 ALSECMAD BSBGEDUM *Secundarios no terminados*
1 SI
0 Otros

83 SECMAD BSBGEDUM *Secundarios terminados*
1 SI
0 Otros

84 VOCACMAD BSBGEDUM *Estudios técnicos o vocacionales posteriores a la secundaria*
1 SI
0 Otros

85 ALUNIMAD BSBGEDUM *No terminó en la universidad*
1 SI
0 Otros

86 UNIMAD BSBGEDUM *Estudios universitarios completos*
1 SI
0 Otros

¿Cuál es el nivel de estudios de tu padre?

*La categoría de referencia es la respuesta “No lo sé”

87 PRIMPAD BSBGEDUF *Primarios*
1 SI
0 Otros

88 ALSECPAD BSBGEDUF *Secundarios no terminados*
1 SI
0 Otros

89 SECPAD	BSBGEDUF	*Secundarios terminados*	1 SI 0 Otros
-----------	----------	--------------------------	-----------------

90 VOCACPAD	BSBGEDUF	*Estudios técnicos o vocacionales posteriores a la secundaria*	1 SI 0 Otros
-------------	----------	--	-----------------

91 ALUNIPAD	BSBGEDUF	*No terminó en la universidad*	1 SI 0 Otros
-------------	----------	--------------------------------	-----------------

92 UNIPAD	BSBGEDUF	*Estudios universitarios completos*	1 SI 0 Otros
-----------	----------	-------------------------------------	-----------------

¿Cuál es el nivel de estudios que esperas conseguir durante tu vida?.

*La categoría de referencia es la respuesta “No lo sé”

93 PRIMSELF	BSBGEDUS	*Primarios*	1 SI 0 Otros
-------------	----------	-------------	-----------------

94 ALSECSELF	BSBGEDUS	*Secundarios no terminados*	1 SI 0 Otros
--------------	----------	-----------------------------	-----------------

95 SECSELF	BSBGEDUS	*Secundarios terminados*	1 SI 0 Otros
------------	----------	--------------------------	-----------------

96 VOCACSELF	BSBGEDUS	*Estudios técnicos o vocacionales posteriores a la secundaria*	1 SI 0 Otros
--------------	----------	--	-----------------

97 ALUNISELF	BSBGEDUS	*No terminar en la universidad*	1 SI 0 Otros
--------------	----------	---------------------------------	-----------------

98 UNISELF	BSBGEDUS	*Estudios universitarios completos*	1 SI 0 Otros
------------	----------	-------------------------------------	-----------------

¿Cuánto tiempo dedicas fuera de la escuela durante la semana a las siguientes actividades?

La variable de referencia es no dedicar nada de tiempo

99	EXTMATM1	BSBMEXTR	*Recibo clases particulares de matemáticas durante menos de 1 hora*
		1 SI	
		0 Otras	

100	EXTMAT1Y2	BSBMEXTR	*Recibo clases particulares de matemáticas entre 1 y 2 horas*
		1 SI	
		0 Otras	

101	EXTMAT3Y5	BSBMEXTR	*Recibo clases particulares de matemáticas entre 3 y 5 horas*
		1 SI	
		0 Otras	

102	EXTMATMAS5	BSBMEXTR	*Recibo clases particulares de matemáticas por más de 5 horas*
		1 SI	
		0 Otras	

103	EXTSCIM1	BSBSEXTR	*Recibo clases particulares de ciencias durante menos de 1 hora*
		1 SI	
		0 Otras	

104	EXTSCI1Y2	BSBSEXTR	*Recibo clases particulares de ciencias entre 1 y 2 horas*
		1 SI	
		0 Otras	

105	EXTSCI3Y5	BSBSEXTR	*Recibo clases particulares de ciencias entre 3 y 5 horas*
		1 SI	
		0 Otras	

106	EXTSCIMAS5	BSBSEXTR	*Recibo clases particulares de matemáticas por más de 5 horas*
		1 SI	
		0 Otras	

107	SCMATCLUBM1	BSBGCLUB	*Asisto a clubes relacionados con las matemáticas o las ciencias durante menos de 1 hora*
		1 SI	
		0 Otras	

108 SCMATCLUB1Y2	BSBGCLUB	* Asisto a clubes relacionados con las matemáticas o las ciencias durante entre 1 y 2 horas*
	1 SI 0 Otras	

109 SCMATCLUB3Y5	BSBGCLUB	* Asisto a clubes relacionados con las matemáticas o las ciencias durante entre 3 y 5 horas *
	1 SI 0 Otras	

110 CLUBMAS5	BSBGCLUB	*Asisto a clubes relacionados con las matemáticas o las ciencias durante más de 5 horas*
	1 SI 0 Otras	

111 PAIDM1	BSBGPAID	*Trabajo de forma remunerada durante menos de 1 hora*
	1 SI 0 Otras	

112 PAID1Y2	BSBGPAID	*Trabajo de forma remunerada entre 1 y 2 horas*
	1 SI 0 Otras	

113 PAID3Y5	BSBGPAID	*Trabajo de forma remunerada entre 3 y 5 horas*
	1 SI 0 Otras	

114 PAIDMAS5	BSBGPAID	*Trabajo de forma remunerada durante más de 5 horas*
	1 SI 0 Otras	

En un día normal de colegio ¿Cuánto tiempo dedicas fuera del colegio a realizar las siguientes actividades?

La variable de referencia es no dedicar nada de tiempo

115 TVM1	BSBGDAY1	*Veo TV o videos durante menos de 1 hora*
	1 SI 0 Otras	

116 TV1Y2	BSBGDAY1	*Veo TV o videos entre 1 y 2 horas*
	1 SI 0 Otras	

117 TV3Y5	BSBGDAY1	*Veo TV o videos entre 3 y 5 horas*
	1 SI 0 Otras	

118 TVMAS5	BSBGDAY1 1 SI 0 Otras	*Veo TV o videos durante más de 5 horas*
119 NINTM1	BSBGDAY2 1 SI 0 Otras	*Juego con los videojuegos durante menos de 1 hora*
120 NINT1Y2	BSBGDAY2 1 SI 0 Otras	*Juego con los videojuegos entre 1 y 2 horas*
121 NINT3Y5	BSBGDAY2 1 SI 0 Otras	*Juego con los videojuegos entre 3 y 5 horas*
122 NINTMAS5	BSBGDAY2 1 SI 0 Otras	*Juego con los videojuegos durante más de 5 horas*
123 AMIGM1	BSBGDAY3 1 SI 0 Otras	*Juego o hablo con amigos durante menos de 1 horas*
124 AMIG1Y2	BSBGDAY3 1 SI 0 Otras	*Juego o hablo con amigos entre 1 y 2 horas*
125 AMIG3Y5	BSBGDAY3 1 SI 0 Otras	*Juego o hablo con amigos entre 3 y 5 horas*
126 AMIGMAS5	BSBGDAY3 1 SI 0 Otras	*Juego o hablo con amigos más de 5 horas*
127 DOMESTM1	BSBGDAY4 1 SI 0 Otras	*Realizo trabajos domésticos durante menos de 1 hora*

128 DOMEST1Y2	BSBGDAY4	*Realizo trabajos domésticos entre 1 y 2 horas*
	1 SI	
	0 Otras	

129 DOMEST3Y5	BSBGDAY4	*Realizo trabajos domésticos entre 3 y 5 horas*
	1 SI	
	0 Otras	

130 DOMESTMAS5	BSBGDAY4	*Veo TV o videos durante más de 5 horas*
	1 SI	
	0 Otras	

131 SPORTM1	BSBGDAY5	*Practico deporte durante menos de 1 hora*
	1 SI	
	0 Otras	

132 SPORT1Y2	BSBGDAY5	*Practico deporte entre 1 y 2 horas*
	1 SI	
	0 Otras	

133 SPORT3Y5	BSBGDAY5	*Practico deporte entre 3 y 5 horas*
	1 SI	
	0 Otras	

134 SPORTMAS5	BSBGDAY5	*Practico deporte durante más de 5 horas*
	1 SI	
	0 Otras	

135 DIVERLIBRM1	BSBGDAY6	*Leo un libro durante menos de 1 hora*
	1 SI	
	0 Otras	

136 DIVERLIBR1Y2	BSBGDAY6	*Leo un libro entre 1 y 2 horas*
	1 SI	
	0 Otras	

137 DIVERLIBR3Y5	BSBGDAY6	*Leo un libro entre 3 y 5 horas*
	1 SI	

		0 Otras
--	--	---------

138 DIVERLIBRMAS5	BSBGDAY6	*Leo un libro durante más de 5 horas*
	1 SI	
	0 Otras	

139 ESTMATM1	BSBGDAY7	*Estudio matemáticas durante menos de 1 hora*
	1 SI	
	0 Otras	

140 ESTMAT1Y2	BSBGDAY7	*Estudio matemáticas entre 1 y 2 horas*
	1 SI	
	0 Otras	

141 ESTMAT3Y5	BSBGDAY7	*Estudio matemáticas entre 3 y 5 horas*
	1 SI	
	0 Otras	

142 ESTMATMAS5	BSBGDAY7	*Estudio matemáticas durante más de 5 horas*
	1 SI	
	0 Otras	

143 ESTSCIM1	BSBGDAY8	*Estudio ciencias durante menos de 1 hora*
	1 SI	
	0 Otras	

144 ESTSCI1Y2	BSBGDAY8	*Estudio ciencias entre 1 y 2 horas*
	1 SI	
	0 Otras	

145 ESTSCI3Y5	BSBGDAY8	*Estudio ciencias entre 3 y 5 horas*
	1 SI	
	0 Otras	

146 ESTSCIMAS5	BSBGDAY8	*Estudio ciencias durante más de 5 horas*
	1 SI	
	0 Otras	

147 ESTOTRASM1	BSBGDAY9	*Estudio otras asignaturas distintas durante menos de 1 hora*
----------------	----------	---

1 SI
0 Otras

148 ESTOTRAS1Y2	BSBGDAY9	*Estudio otras asignaturas entre 1 y 2 horas*
-----------------	----------	---

1 SI
0 Otras

149 ESTOTRAS3Y5	BSBGDAY9	*Estudio otras asignaturas entre 3 y 5 horas*
-----------------	----------	---

1 SI
0 Otras

150 ESTOTRASMAS5	BSBGDAY9	*Estudio otras asignaturas durante más de 5 horas*
------------------	----------	--

1 SI
0 Otras

¿Cómo de a menudo sucedieron las siguientes cosas durante el último mes en el colegio?.

*La categoría de referencia para todas estas variables es “nunca”

151 SKIPCLASS12	BSBGSSKP	*Me salté una clase una o dos veces*
-----------------	----------	--------------------------------------

1 SI
0 Otras

152 SKIPCLASS34	BSBGSSKP	*Me salté una clase 3 o 4 veces*
-----------------	----------	----------------------------------

1 SI
0 Otras

153 SKIPCLASSM5	BSBGSSKP	*Me salté más de cinco clases*
-----------------	----------	--------------------------------

1 SI
0 Otras

154 ROBA12	BSBGSSTL	*Me robaron algo una o dos veces*
------------	----------	-----------------------------------

1 SI
0 Otras

155 ROBA34	BSBGSSTL	*Me robaron algo 3 o 4 veces*
------------	----------	-------------------------------

		1 SI 0 Otras
156 ROBAM5	BSBGSTL	*Me robaron algo más de cinco veces*
		1 SI 0 Otras
157 AMENAZA12	BSBGSHRT	*Pensé una o dos veces que otro alumno podría pegarme*
		1 SI 0 Otras
158 AMENAZA34	BSBGSHRT	*Pensé unas tres o cuatro veces que otro alumno podría pegarme*
		1 SI 0 Otras
159 AMENAZAM5	BSBGSHRT	*Pensé cinco o más veces que otro alumno podría pegarme*
		1 SI 0 Otras
160 ASKIPCLASS12	BSBGFSKP	*Alguno de mis amigos se saltó una clase una o dos veces*
		1 SI 0 Otras
161 ASKIPCLASS34	BSBGFSKP	*Alguno de mis amigos se saltó una clase 3 o 4 veces*
		1 SI 0 Otras
162 ASKIPCLASSM5	BSBGFSKP	*Alguno de mis amigos se saltó más de cinco clases*
		1 SI 0 Otras
163 AROBA12	BSBGFSTL	*Robaron algo a alguno de mis amigos una o dos veces*
		1 SI 0 Otras
164 AROBA34	BSBGFSTL	*Robaron algo a alguno de mis amigos 3 o 4 veces*
		1 SI 0 Otras

165 AROBAM5	BSBGFSTL	*Robaron algo a alguno de mis amigos más de cinco veces*
	1 SI	
	0 Otras	

166 ADAÑO12 veces*	BSBGFHRT	*Amenazon o pegaron a alguno de mis amigos una o dos veces*
	1 SI	
	0 Otras	

167 ADAÑO34	BSBGFHRT	*Amenazaron o pegaron a alguno de mis amigos 3 o 4 veces*
	1 SI	
	0 Otras	

168 ADAÑOM5	BSBGFHRT	*Amenazaron o pegaron a alguno de mis amigos más de cinco veces*
	1 SI	
	0 Otras	

ANEXO 2: Etiquetas y Códigos de las Variables del Colegio y el Profesor creadas a partir del fichero original del TIMSS para los datos de España: bcgcode.cdt y btmcode.cdt

VARIABLES DEL COLEGIO: Codebook, Date: 24.09.97 File: bcgcode.SDB

Var.Question Location/Format No.	Variable Label Name	Code	Option
--	------------------------	------	--------

Nº en TIMSS	NOMBRE EN EL TIMSS *Descripción*	Valores
-------------	-------------------------------------	---------

Las variables que no tienen un nombre descriptivo sólo han sido utilizadas para derivar otras variables

5 COLEGIO 10-15 / N 6.0	IDSCHOOL *SCHOOL ID*	VALUE	School ID
----------------------------	----------------------	-------	-----------

10 SCQ2-1 25 / C 1.0	BCBGCOMM GEN\TYPE OF COMMUNITY Localización del colegio		
	RURAL	1 Area rural	0 Otro
	CIUDAD	1 Cerca del centro de la ciudad	0 Otro

La variable de referencia es: Localización a las afueras o extrarradios de la ciudad

36 SCQ2-5 84-86 / N 3.0	BCBGTE5Y GEN\TEACHER AT SCHOOL FOR 5+ YEARS Porcentaje de profesores en la escuela durante más de 5 años	MAS5AÑOS	Valor en porcentaje
----------------------------	---	----------	---------------------

37 SCQ2-6A 87-89 / N 3.0	BCBMTEAC MAT\%TEACHER\75%+ IN MATHEMATICS Porcentaje de profesores con más del 75% de clases para matemáticas.	MAS75%	Valor en porcentaje
-----------------------------	---	--------	---------------------

43 SCQ2-7 105 / C 1.0	BCBGSAME GEN\TIME STUDENTS STAY WITH TEACHER Tiempo que los estudiantes permanecen con un mismo profesor MISMOPROFESOR		
		1	De uno a tres años escolares
		0	Cuatro años escolares

49 SCQ2-10B 115 / C 1.0	BCBGCOL2 GEN\COOPERATION\ENCOURAGED ¿Fomenta el colegio que los profesores discutan sobre distintas ideas acerca del material y los métodos educativos?	DISMATERIAL	
		0	Si
		1	No

50 SCQ2-10C	BCBGCOL3 GEN\COOPERATION\REGULAR MEETINGS ¿Se reúnen regularmente los profesores para hablar sobre objetivos y temas educativos?		
116 / C 1.0	REUNIONES	0 Si 1 No	
59 SCQ2-11I	BCBGAC09 GEN\PRINCIPAL\TALK WITH PARENTS Como director, ¿Cuánto tiempo dedicas al mes a hablar con los padres?		
141-143 / N 3.0	DIRHABPADRES	Valor	horas/mes
61 SCQ2-11K	BCBGAC11 GEN\PRINCIPAL\RESPONDING TO REQUEST Como director, ¿Cuánto tiempo dedicas a tareas administrativas con instituciones?		
147-149 / N 3.0	BUROCRADIR	Valor	horas/mes
65 SCQ2-12A	BCBGRP01 GEN\RESPONSIBLE\HIRE TEACHERS En el colegio, ¿Quién tiene la responsabilidad de contratar al profesorado?		
159 / C 1.0	PRIVADO	0 Colegio público 1 Colegio privado concertado	
96 SCQ2-15A	BCBGCOM1 GEN\COMPUTERS\AVAILABLE ¿Cuántos ordenadores hay en tu colegio por cada 100 alumnos		
190-192 / N 3.0	ORDEALUM	Valor	número de ordenadores por cada 100 alumnos
102 SCQ2-16B	BCBGST02 GEN\SHORTAGE\BUDGET FOR SUPPLIES ¿Está la capacidad de enseñar del colegio limitado por bajo presupuesto?		
206 / C 1.0	LIMPRESUP	1 Pocas 0 Bastantes	
120 SCQ2-17B	BCBGABST GEN\STUDENT\STUDENT\% ABSENT VALUE ¿Qué porcentaje de alumnos faltan a la escuela un día normal?		
230-232 / N 3.0	FALTA CLASE	Valor en Porcentaje	
122 SCQ2-17D	BCBGTNSF GEN\STUDENT\% TRANSFERRING INTO SCHL ¿Qué porcentaje de alumnos llegan a su escuela una vez empezado el curso		
236-238 / N 3.0	NUEVOS	Valor en porcentaje	
133 SCQ2-17K1	BCBGUBER GEN\U-GRADE\ENROLLMENT\BOYS Total número de niños en octavo		
269-271 / N 3.0	Valor	número de niños	

134 SCQ2-17K2	BCBGUGER GEN\U-GRADE\ENROLLMENT\GIRLS Total número de niñas en octavo
272-274 / N 3.0	Valor número de niñas

137 SCQ2-17M	BCBGUSIZ GEN\U-GRADE\AVERAGE CLASS SIZE Tamaño medio de la clase
281-283 / N 3.0	RATIOALU Número de alumnos en clase

Variable derivada a partir de las tres variables anteriores del TIMSS

181 SCQ2-19A2	BCBGUDYY GEN\U-GRADE\DAYS PER YEAR Número de días del año escolar
375-377 / N 3.0	DIASAÑO Valor días

187 SCQ2-19D2	BCBGUTHW GEN\U-GRADE\# TOTAL HOURS PER WEEK ¿Cuántas horas a la semana tienen de clase los alumnos de octavo?
384-385 / N 2.0	HORASEMANA Valor horas

194 SCQ2-20B2	BCBGUTMP GEN\U-GRADE\DURATION OF INSTR PERIOD ¿De cuántos minutos es una clase normal?
398-400 / N 3.0	MINUTOSCLASE Valor minutos

217 SCQ2-25B	BCBMUC2 MAT\U-GRADE\WEEKS PER YEAR ¿Cuántas semanas al año se dedican a clases de matemáticas?
425-426 / N 2.0	SEMANASAÑO Valor semanas

251 SCQ2-29A	BCBGSTD1 GEN\% FROM DISADVANTAGED BACKGROUND ¿Qué porcentaje de alumnos de tu colegio provienen de familias de nivel socioeconómico bajo?
485-487 / N 3.0	NIVELSOCBAJO Valor porcentaje

254 SCQ2-29D	BCBGSTD4 GEN\% ATTENDED PRESCHOOL ¿Qué porcentaje de alumnos en tu colegio asistieron a preescolar?
494-496 / N 3.0	PREESCOLAR Valor porcentaje

256 SCQ2-29F	BCBGSTD6 GEN\% HAVE LEARNING PROBLEMS ¿Qué porcentaje de alumnos de tu colegio tienen problemas de aprendizaje?
500-502 / N 3.0	PROBAPREND Valor porcentaje

VARIABLES DE PROFESOR: Codebook, Date: 24.09.97 File: btmcode.SDB

Var.Question Location/Format No.	Variable Label Name	Code	Option
6 TEACHER 16-23 / N 8.0	IDTEACH	*TEACHER ID*	VALUE Teacher ID
10 CLASS1 29-36 / N 8.0	ILCLASS1	*FIRST CLASS ID*	VALUE Class ID
19 TQM2A1 59-60 / C 2.0	BTBGAGE GEN\AGE OF TEACHER Edad del profesor EDADPROFESOR		1 25-49 0 50 o más
20 TQM2A2 61 / C 1.0	BTBGSEX GEN\SEX OF TEACHER Sexo del profesor SEXOPROF		0 Mujer 1 Hombre
21 TQM2A3 62-63 / C 2.0	BTBGEDUC GEN\LEVEL OF EDUCATION COMPLETED Nivel de educación más elevado EDUCACION		1 Sólo formación de profesor 2 Secundaria más 3 o 4 años de formación específica 3 Licenciado 4 Licenciado más formación específica 5 Magíster
30 TQM2A4H 72 / C 1.0	BTBMGR6 MAT\GRADE LEVEL\GRADE 6 ¿Enseñas matemáticas en sexto además de en octavo este año? SEXTO		0 Si 1 No
31 TQM2A4I 73 / C 1.0	BTBMGR7 MAT\GRADE LEVEL\GRADE 7 ¿Enseñas matemáticas en séptimo además de en octavo este año? SÉPTIMO		0 Si 1 No
54 TQM2A6 96 / C 1.0	BTBGPTFT GEN\TEACH PART-TIME OR FULL-TIME ¿Enseñas a tiempo parcial o a tiempo completo DEDICACION		

0 tiempo parcial
1 tiempo completo

71 TQM2A9A BTBMSUB1 MAT\SCHEDULED SUBJECTS\MATHEMATICS
¿Cuántas horas semanales dedicas a planificar la asignatura de matemáticas?
114-115 / N 2.0 PLANIFMATES
Valor Número de horas

72 TQM2A9B BTBSSUB2 SCI\SCHEDULED SUBJECTS\GENERAL SCI
¿Cuántas horas semanales dedicas a planificar la asignatura de ciencias?
116-117 / N 2.0 PLANIFCIENC
Valor Número de horas

79 TQM2A9I BTBGSUB9 GEN\SCHEDULED SUBJECTS\OTHER
¿Cuántas horas semanales dedicas a planificar otras asignaturas?
130-131 / N 2.0 PLANIFOTRAS
Valor Número de horas

80 TQM2A10A BTBGTSK1 GEN\SCHEDULED TASKS\SUPERVISION
¿Cuántas horas dedicas a supervisar las tareas de los alumnos?
132-133 / N 2.0 TAREAS
Valor Número de horas

81 TQM2A10B BTBGTSK2 GEN\SCHEDULED TASKS\COUNSELLING
¿Cuántas horas a la semana dedicas a apoyar a los alumnos?
134-135 / N 2.0 APOYO
Valor Número de horas

82 TQM2A10C BTBGTSK3 GEN\SCHEDULED TASKS\ADMINISTRATION
¿Cuántas horas a la semana dedicas a tareas burocráticas?
136-137 / N 2.0 BUOCRAPROF
Valor Número de horas

84 TQM2A10E BTBGTSK5 GEN\SCHEDULED TASKS\COOP PLANNING
¿Cuántas horas a la semana dedicas a tareas de cooperación con otros
profesores?
140-141 / N 2.0 HORASCOOPERA
Valor Número de horas

85 TQM2A10F BTBGTSK6 GEN\SCHEDULED TASKS\NON SPECIFIED
¿Cuántas horas a la semana dedicas a otras tareas?
142-143 / N 2.0 OTRAS
Valor Número de horas

91 TQM2A12E BTBGACT5 GEN\OUTSIDE SCHL\MEET PARENTS
¿Cuántas horas diarias no lectivas dedicas al encuentro con padres?
150 / C 1.0 PROFHABPADRES
Valor Número de horas

95 TQM2A13 BTBGMEET GEN\MEETING WITH OTHER TEACHERS
¿Cómo de a menudo te reúnes con otros profesores para discutir sobre el temario o el desarrollo de las clases?

154-155 / C 2.0

La categoría de referencia es “Nunca o una o dos veces al año”

PARMESES	1 Una vez cada par de meses	0 Otra
UNAMES	1 Una vez al mes	0 Otra
UNASEMANA	1 Una vez a la semana	0 Otra
VARIASEMANA	1 Varias veces a la semana	0 Otra

96 TQM2A14A BTBGINF1 GEN\INFLUENCE\SUBJECT MATTER
¿Cuánta influencia tienes en los tópicos que se enseñan en las distintas materias? 156 / C 1.0
TOPICOS
0 Algo
1 Mucha

97 TQM2A14B BTBGINF2 GEN\INFLUENCE\SPECIFIC TEXTBOOKS
¿Cuánta influencia tienes en la elección del libro de texto?
LIBROTEXTO
0 Algo
1 Mucha

98 TQM2A14C BTBGINF3 GEN\INFLUENCE\MONEY FOR SUPPLIES
¿Cuánta influencia tienes en la cantidad de dinero a gastar en material educativo?
158 / C 1.0
GASTOMATERIAL
0 Algo
1 Mucha

126 TQM2A22 BTBGBOOK GEN\NUMBER OF BOOKS IN HOME
186 / C 1.0
LIBROSPROFE
1 Suficientes para llenar una librería (26-100)
2 Suficientes para llenar dos librerías (101-200)
3 Suficientes para llenar tres o más librerías (>200)

136 TQM2B1/1 BTBMBOY MAT\CLSS\NUMBER OF BOYS IN CLASS
Número de alumnos en clase
205-206 / N 2.0
ALUMNOS
Valor Número de chicos

137 TQM2B1/2 BTBMGIRL MAT\CLSS\NUMBER OF GIRLS IN CLASS
Número de alumnas en clase
207-208 / N 2.0
Valor Número de chicas

141 TQM2B3 218-220 / N 3.0	BTBMTIME MAT\CLSS\LENGTH OF MATH CLASS ¿Cuántos minutos a la semana enseñas matemáticas a tu clase? MINSEMANA Valor Minutos a la semana
-------------------------------	---

154 TQM2B5 234 / C 1.0	BTBMTXBS MAT\CLSS\TEXTBOOK NAME ENTERED ¿Qué porcentaje del tiempo de la clase de matemáticas está basado en el libro de texto? TIEMPOTEXTO 1 0 – 25% 2 26 – 50% 3 51 – 75% 4 76 – 100%
---------------------------	---

172 TQM2B8 253 / C 1.0	BTBMCALC MAT\ACCESS TO CALC DURING MATH LESS ¿Cuántos de tus estudiantes tienen acceso a calculadora durante la clase de matemáticas? CALCLASE 1 Casi todos 2 Unos tres cuartos 3 La mitad 4 En torno a un cuarto 5 Nadie
---------------------------	---

388 TQM2B14B 510 / C 1.0	BTBMSGRP MAT\STUDENT WORK IN SMALL GROUPS ¿Trabajan los niños en grupos pequeños? GRUPOS 0 Nada 1 Algo de tiempo
-----------------------------	--

389 TQM2B15A 511 / C 1.0	BTBMASK1 MAT\TASKS\EXPLAIN REASONING ¿Con qué frecuencia preguntas a tus alumnos que expliquen las ideas de forma razonada? RAZONAR 1 En algunas lecciones 2 La mayoría de las lecciones 3 Cada una de las lecciones
-----------------------------	--

390 TQM2B15B 512 / C 1.0	BTBMASK2 MAT\TASKS\ANALYZE RELATIONSHIPS ¿Con qué frecuencia pides a tus alumnos que utilicen tablas, gráficos o mapas? GRAFICOS 1 Nunca o casi nunca 2 En algunas lecciones 3 La mayoría de las lecciones 4 Cada una de las lecciones
-----------------------------	--

392 TQM2B15D 514 / C 1.0	BTBMASK4 MAT\TASKS\USE COMPUTERS ¿Con qué frecuencia les pides a tus alumnos que utilicen el ordenador? ORDENADOR 0 Nunca o casi nunca 1 En algunas lecciones
-----------------------------	---

395 TQM2B16A 517 / C 1.0	BTBMDO1 MAT\REACTION\CORRECT STUDENT Después de una respuesta incorrecta ¿Con qué frecuencia corriges al alumno en frente del resto de la clase? CORREGIRALUM 1 Nunca o casi nunca 2 En algunas lecciones 3 La mayoría de las lecciones 4 Cada una de las lecciones
-----------------------------	---

396 TQM2B16B 518 / C 1.0	BTBMDO2 MAT\REACTION\HELP STUDENT Después de una respuesta incorrecta ¿Con qué frecuencia pides a otro alumno que le ayude a dar la respuesta? AYUDARESPUESTA 1 En algunas lecciones 2 La mayoría de las lecciones 3 Cada una de las lecciones
-----------------------------	---

397 TQM2B16C 519 / C 1.0	BTBMDO3 MAT\REACTION\OTHER\CORRECT ANSWER Después de una respuesta incorrecta ¿Con qué frecuencia pides a otro alumno que tiene más probabilidad de contestar bien que responda? PREGUNTAOTRO 1 Nunca o casi nunca 2 En algunas lecciones 3 La mayoría de las lecciones 4 Cada una de las lecciones
-----------------------------	---

398 TQM2B16D 520 / C 1.0	BTBMDO4 MAT\REACTION\OTHER\DISCUSS RESPONSE Después de una respuesta incorrecta ¿Con qué frecuencia pides otras respuestas y luego éstas son discutidas? DEBATE 1 Nunca o casi nunca 2 En algunas lecciones 3 La mayoría de las lecciones 4 Cada una de las lecciones
-----------------------------	---

399 TQM2B17A 521 / C 1.0	BTBMLES1 MAT\INDIVIDUAL WORK\NO ASSISTANCE En clase de matemáticas ¿Con qué frecuencia los alumnos trabajan individualmente sin ayuda? TRABINDIVSIN 1 Nunca o casi nunca 2 En algunas lecciones 3 La mayoría de las lecciones 4 Cada una de las lecciones
-----------------------------	---

400 TQM2B17B	BTBMLES2 MAT\INDIVIDUAL WORK\ASSISTANCE En clase de matemáticas ¿Con qué frecuencia los alumnos trabajan individualmente con ayuda?
522 / C 1.0	TRABINDCON 1 Nunca o casi nunca 2 En algunas lecciones 3 La mayoría de las lecciones 4 Cada una de las lecciones

401 TQM2B17C	BTBMLES3 MAT\WORK TOGETHER\TEACHER TEACHES En clase de matemáticas ¿Con qué frecuencia los alumnos hacen tareas en grupo con el profesor como guía?
523 / C 1.0	PROFGUIA 1 Nunca o casi nunca 2 En algunas lecciones 3 La mayoría de las lecciones 4 Cada una de las lecciones

402 TQM2B17D	BTBMLES4 MAT\WORK TOGETHER\STUDENTS INTERACT En clase de matemáticas ¿Con qué frecuencia los alumnos trabajan juntos respondiéndose las preguntas unos a otros?
524 / C 1.0	INTERACCION 1 Nunca o casi nunca 2 En algunas lecciones 3 La mayoría de las lecciones 4 Cada una de las lecciones

403 TQM2B17E	BTBMLES5 MAT\WORK IN GROUPS\NO ASSISTANCE En clase de matemáticas ¿Con qué frecuencia los alumnos trabajan en parejas sin ayuda?
525 / C 1.0	PORPAREJAS 1 Nunca o casi nunca 2 En algunas lecciones 3 La mayoría de las lecciones

405 TQM2B18	BTBMHMW2 MAT\HOMEWORK FREQUENCY ¿Con qué frecuencia pones tarea para casa?
527 / C 1.0	TAREACASA 1 Menos de una vez a la semana 2 Una o dos veces a la semana 3 Tres o cuatro veces a la semana 4 Todos los días

406 TQM2B19	BTBMHWT2 MAT\AVERAGE TIME NEEDED FOR HOMEWK ¿Cuántos minutos de tarea asignas normalmente a tus alumnos?
528-529 / C 2.0	MINUTOSTAREA 1 Menos de 15 minutos 2 15-30 minutos 3 31-60 minutos 4 61-90 minutos

417 TQM2B21A 540 / C 1.0	BTBMWHW1 MAT\WRITTEN HMWK\RECORD COMPLETION ¿Con qué frecuencia apuntas quien ha hecho o no la tarea? TAREAHECHA 1 A veces 2 Siempre 0 No mando tarea
-----------------------------	--

419 TQM2B21C 542 / C 1.0	BTBMWHW3 MAT\WRITTEN HMWK\RETURN ASSIGNMENTS ¿Con qué frecuencia recoges, corriges y devuelves las tareas? RECOGERTAREA 1 Nunca 2 Raramente 3 A veces 4 Siempre 0 No mando tarea
-----------------------------	---

421 TQM2B21E 544 / C 1.0	BTBMWHW5 MAT\WRITTEN HMWK\STUDNT CORRECT OWN ¿Con qué frecuencia los alumnos corrigen en clase sus propios deberes? AUTOCORREGIR 1 Nunca 2 Raramente 3 A veces 4 Siempre 0 No mando tarea
-----------------------------	--

422 TQM2B21F 545 / C 1.0	BTBMWHW6 MAT\WRITTEN HMWK\STUDNT CORRECT OTHR ¿Con qué frecuencia los alumnos intercambian las tareas para corregirlas? INTERCORREC 1 Nunca 2 Raramente 3 A veces 4 Siempre 0 No mando tarea
-----------------------------	---

423 TQM2B21G 546 / C 1.0	BTBMWHW7 MAT\WRITTEN HMWK\BASIS FOR DISCUSSN ¿Con qué frecuencia usas las tareas como base para discusiones en clase? TAREADISCUSSION 1 Nunca 2 Raramente 3 A veces 4 Siempre 0 No mando tarea
-----------------------------	---

424 TQM2B21H 547 / C 1.0	BTBMWHW8 MAT\WRITTEN HMWK\GRADE CONTRIBUTION ¿Con qué frecuencia usas las tareas como criterio para poner las notas? TAREANOTAS 1 Nunca 2 Raramente 3 A veces 4 Siempre 0 No mando tarea
-----------------------------	---

Porcentaje de alumnos que contestan en una clase que casi siempre o muy a menudo realizan las siguientes actividades en clase

BSBMNOTE: Copio los apuntes de la pizarra.

PIZARRA

BSBMHWCL: Empezamos las tareas de casa al final de la clase.

TAREACLASE

ANEXO 3: Valores de los outputs reales y las predicciones para los centros evaluados.

IDCOLEGIO	output1	output2	output3	promedio	pred1	pred2	pred3	premedia
1	36,84	33,33	10,53	438,71	45,33	29,05	18,44	457,95
2	13,64	0	0	448,32	35,16	10,16	1,38	458,75
3	50	25	26,32	454,33	54,80	34,53	21,23	458,31
4	25	11,11	5	415,76	29,50	0,00	0,00	444,24
5	35,71	0	12,5	481,08	35,04	29,09	11,00	470,89
6	53,85	25	27,27	480,26	40,04	19,01	24,08	478,39
7	37,5	9,09	14,81	455,78	39,43	16,29	27,59	486,13
8	50	40	14,29	469,4	58,16	31,82	13,05	481,73
9	33,33	25	36,36	458,53	33,52	26,14	22,55	477,56
10	52	33,33	13,64	481,02	54,89	65,34	35,39	500,97
11	64	40	29,17	482,02	59,85	53,61	23,85	492,78
12	76,92	80	47,83	522,46	57,52	54,96	28,51	493,84
13	52,38	20	12,5	454,3	46,97	32,44	15,68	488,64
14	41,94	.	17,65	502,07	52,54	64,83	24,74	501,34
15	40,91	14,29	15	444,89	29,60	16,29	11,52	437,92
17	21,74	33,33	5,26	453,98	40,69	53,99	12,09	497,67
18	35	0	27,78	469,39	-	30,53	-	-
19	0	0	0	407	17,20	14,71	13,03	450,09
20	34,48	30	11,11	451,54	31,06	18,05	2,87	423,51
21	71,43	50	23,81	496,83	58,41	53,39	28,22	530,50
23	38,89	50	5,88	455,18	51,46	44,99	13,01	466,60
24	86,96	.	60	585,98	88,59	44,39	45,22	539,76
25	46,67	30,77	6,67	425	40,65	29,00	3,23	441,24
26	57,58	0	23,08	501,95	49,90	23,81	17,71	499,98
27	40	0	30,77	518,47	56,79	44,83	23,19	518,68
28	53,85	50	19,35	492,73	75,24	36,51	26,36	511,62
29	59,46	66,67	47,62	539,24	53,93	43,84	40,82	527,01
30	77,14	50	55	538,51	56,64	41,90	34,55	511,49
31	75,76	100	60	566,78	-	42,28	-	-
32	50	100	50	537,7	41,05	43,70	37,05	496,47
33	33,33	0	14,29	464,69	40,92	0,00	15,14	478,16
34	61,54	57,14	0	475,11	59,44	28,82	12,76	480,42
35	19,05	25	10,53	458,53	56,82	51,62	34,85	492,53
36	31,82	25	6,67	469,56	43,49	25,24	15,03	485,98
37	37,5	18,18	5	456,23	34,56	30,23	0,00	453,76
38	60	.	30	517,57	34,77	32,85	14,88	492,69
40	68,42	71,43	31,25	503,31	52,81	50,97	30,13	489,59
41	31,58	22,22	5,26	435,16	36,94	25,32	19,24	469,53
42	47,83	25	21,05	477,1	41,18	31,08	16,44	465,99
43	61,54	44,44	32	489,32	55,93	37,99	15,46	474,89
44	50	28,57	18,75	468,47	49,06	44,00	15,65	480,71
45	52,38	66,67	15,79	478,65	55,24	26,48	25,67	497,95
46	65,52	53,85	20	489,2	66,82	49,87	26,62	482,65
47	70,59	55,56	21,74	511,65	55,31	53,85	29,63	503,52
51	21,74	0	0	448,99	25,70	20,62	-0,24	439,26
52	26,32	0	6,25	464,81	44,11	20,75	16,71	495,67
53	30	28,57	17,65	455,8	33,62	22,67	19,73	471,11
55	33,33	14,29	18,18	480,26	32,72	29,87	17,42	462,10
56	59,09	20	33,33	506,08	46,88	5,51	26,66	483,33

IDCOLEGIO	output1	output2	output3	promedio	pred1	pred2	pred3	premedia
57	47,22	27,27	14,71	471,93	44,43	19,76	17,05	499,75
59	59,26	42,86	45,45	506,28	53,18	40,95	26,17	461,00
60	58,62	50	40	529,89	48,31	35,54	32,98	504,14
61	31,43	28,57	6,06	443,19	-	27,37	-	-
62	68,75	60	40	495,57	64,43	50,44	31,35	486,96
63	85	50	37,5	517,48	66,65	28,52	37,60	496,81
64	61,54	55,56	11,76	506,04	59,64	29,45	22,42	480,32
65	50	12,5	28,57	460,16	-	1,40	-	-
66	68,75	60	37,5	501,83	67,50	50,29	33,39	494,60
67	57,14	20	17,65	485,39	46,37	27,43	34,44	475,55
68	40	50	27,78	479,74	46,58	39,78	15,92	481,02
70	39,13	75	15,79	484,45	35,02	53,20	17,12	500,19
71	48,48	66,67	14,29	503,51	-	22,57	-	-
72	40	0	27,78	482,72	55,56	41,17	37,55	492,01
73	63,16	0	53,33	519,35	28,29	26,10	27,77	488,47
74	64	50	33,33	514,43	66,15	50,57	39,27	536,21
75	50	33,33	7,69	497,8	43,20	31,34	26,96	491,60
76	33,33	0	27,27	463,39	42,79	17,31	26,99	489,59
77	60	50	39,13	507,2	53,49	45,12	32,52	509,65
78	55	33,33	23,53	474,86	44,47	23,06	19,92	484,49
79	63,64	33,33	22,22	491,02	59,50	29,64	19,43	473,78
80	47,83	16,67	9,52	450,54	53,34	32,73	26,13	487,35
81	37,5	12,5	6,67	440,92	53,21	33,38	23,71	487,12
82	66,67	46,15	21,05	464,77	48,35	9,22	5,41	455,69
83	24	0	8,7	445,02	33,34	28,76	20,78	444,78
84	46,67	25	22,22	493,42	39,44	18,64	23,30	496,48
85	42,86	33,33	20,83	465,04	36,62	38,93	10,59	453,42
86	70,59	42,86	21,43	500,77	48,71	34,02	23,36	476,37
87	56,67	0	31,58	499,07	45,66	24,02	14,34	472,66
88	10	33,33	0	436,71	45,54	31,40	8,03	446,92
89	75,86	50	27,27	505,33	62,28	49,42	23,39	468,83
90	51,72	25	12	461,41	41,57	24,75	13,22	482,38
91	27,78	17,65	2,86	426,64	37,58	7,78	13,93	474,42
92	53,66	16	20,51	457,52	58,43	45,35	23,73	486,63
93	41,67	50	25	484,02	47,99	38,60	31,58	505,34
94	40	.	50	512,17	36,11	53,39	42,48	514,02
95	20	0	25	427,48	35,34	1,21	24,88	469,79
96	73,33	80	35	527,85	62,88	52,93	22,08	476,68
97	69,57	33,33	25	505,36	71,94	12,96	28,01	496,89
98	44	14,29	17,65	502,91	55,05	30,32	28,93	508,08
99	58,33	30	5,26	477,69	57,89	40,74	25,84	481,02
100	46,15	42,86	21,74	477,29	51,28	40,73	10,32	463,98
101	76,19	33,33	50	522,97	75,87	35,15	35,20	495,07
102	60	0	25	484,68	49,58	0	23,45	459,47
103	29,17	40	22,22	489,56	41,80	38,00	31,73	485,64
104	57,14	33,33	44,44	501,61	49,53	21,99	29,62	493,75
105	33,33	0	5,26	441,43	26,21	34,14	12,94	469,70
106	28,57	50	10	446,78	16,31	18,85	12,06	446,43
107	27,27	0	0	468,8	28,70	0	11,20	463,05
108a	22,22	0	11,11	447,84	36,42	0,45	20,71	470,16
108b	50	25	33,33	447,84	47,31	15,65	18,31	480,81

IDCOLEGIO	output1	output2	output3	promedio	pred1	pred2	pred3	premedia
109	33,33	33,33	15,38	486,6	44,88	25,91	14,06	465,83
111	26,09	0	7,14	473,52	46,32	39,91	21,64	490,78
112	80,77	50	45,45	535,73	63,54	60,90	32,33	522,65
113	64,29	50	33,33	523,54	70,25	46,50	17,85	491,59
115	11,76	0	7,69	450,57	45,12	24,74	32,97	495,59
116	28	33,33	15	471,21	40,00	36,44	22,22	496,84
117	61,9	50	41,67	542,85	37,55	17,70	24,02	487,59
118	82,76	57,14	47,62	526,68	72,50	45,13	23,77	508,86
119	59,38	80	26,67	527,05	55,64	31,91	26,47	482,86
120	50	27,27	6,25	489,28	51,78	14,31	14,60	493,79
121	37,04	33,33	5,26	469,97	40,50	26,69	20,87	475,82
122	65,22	42,86	33,33	508,07	45,01	47,73	24,43	492,49
123	5	0	0	408,59	35,61	16,09	21,55	477,49
124	36,36	50	12,5	496,63	47,17	46,69	12,81	497,10
125	40,74	50	15	469,17	41,17	36,83	10,08	455,68
126	60,87	14,29	31,25	473,29	65,12	27,71	27,99	501,65
127	50	42,86	31,58	494,92	49,56	40,83	31,89	495,90
128	44	50	18,75	501,16	32,95	42,39	13,65	479,66
129	61,76	33,33	26,92	503,95	50,68	42,29	17,98	474,91
130	40,63	23,08	17,86	456,22	51,59	23,57	26,51	478,78
131	33,33	0	12,5	453,69	45,73	20,40	21,20	493,11
131	18,18	20	0	453,69	27,77	31,78	11,45	495,29
132	48	62,5	20	492,25	51,05	56,91	23,71	509,93
133	41,38	66,67	11,76	506,72	45,13	65,64	27,75	499,50
134	51,61	33,33	9,52	491,97	40,71	43,48	9,41	479,55
135	41,38	66,67	6,67	506,9	55,11	53,99	14,86	498,29
136	66,67	20	35,71	526,94	58,22	24,50	37,18	530,40
137	25	28,57	0	459,22	27,76	21,26	4,39	479,47
138	50	0	26,09	507,77	54,84	0	27,11	505,08
139	64	50	40	550,13	55,83	42,46	29,71	542,10
140	63,64	100	40	556,31	54,54	60,36	29,31	518,39
141	91,89	76,92	66,67	577,06	75,06	79,52	47,80	543,23
142	51,43	0	50	522,33	59,23	23,14	48,60	502,70
143	43,75	0	23,53	504,24	53,50	29,91	22,33	488,59
144	74,19	100	65	557,21	60,34	63,34	47,25	521,96
145	54,05	83,33	33,33	498,67	64,36	56,73	35,45	512,91
146	38,89	40	18,52	490,53	36,26	36,46	30,59	502,37
147	48,39	37,5	31,58	501,82	40,89	53,95	26,20	502,10
148	48,39	.	15,38	513,31	54,99	12,08	21,29	495,96
149	45,95	20	12	487,94	60,37	41,89	18,94	503,78
150	52,63	0	0	522,4	60,71	31,21	25,91	511,37
316	28,57	0	12,5	470,31	28,89	8,86	12,63	470,54
451	54,55	28,57	31,25	497,94	45,80	32,08	16,03	461,53
850	38,1	8,33	5,56	458,38	53,53	26,57	23,37	454,29
894	37,5	40	26,67	463,43	39,28	41,70	19,87	452,67
895	46,15	0	33,33	490,99	39,38	5,22	13,83	457,75
902	72,22	60	33,33	520,79	40,48	34,58	20,87	476,61

. → El colegio no tiene ningún alumno con un rendimiento esperado por debajo del primer cuartil del total.

- → Falta el dato de la titularidad de la escuela para calcular el output en estas dimensiones.

ANEXO 4: Eficiencia técnica de los centros evaluados para cada output.

IDCOLEGIO	Eficiencia 1	Eficiencia 2	Eficiencia 3	EficMedia
1	0,81	1,15	0,57	0,96
2	0,39	0,00	0,00	0,98
3	0,91	0,72	1,24	0,99
4	0,85	8,91	3,89	0,94
5	1,02	0,00	1,14	1,02
6	1,34	1,31	1,13	1,00
7	0,95	0,56	0,54	0,94
8	0,86	1,26	1,09	0,97
9	0,99	0,96	1,61	0,96
10	0,95	0,51	0,39	0,96
11	1,07	0,75	1,22	0,98
12	1,34	1,46	1,68	1,06
13	1,12	0,62	0,80	0,93
14	0,80	-	0,71	1,00
15	1,38	0,88	1,30	1,02
17	0,53	0,62	0,44	0,91
18	-	0,00	-	-
19	0,00	0,00	0,00	0,90
20	1,11	1,66	3,87	1,07
21	1,22	0,94	0,84	0,94
23	0,76	1,11	0,45	0,98
24	0,98	-	1,33	1,09
25	1,15	1,06	2,06	0,96
26	1,15	0,00	1,30	1,00
27	0,70	0,00	1,33	1,00
28	0,72	1,37	0,73	0,96
29	1,10	1,52	1,17	1,02
30	1,36	1,19	1,59	1,05
31	-	2,36	-	-
32	1,22	2,29	1,35	1,08
33	0,81	8,91	0,94	0,97
34	1,04	1,98	0,00	0,99
35	0,34	0,48	0,30	0,93
36	0,73	0,99	0,44	0,97
37	1,08	0,60	3,89	1,01
38	1,73	-	2,02	1,05
40	1,30	1,40	1,04	1,03
41	0,85	0,88	0,27	0,93
42	1,16	0,80	1,28	1,02
43	1,10	1,17	2,07	1,03
44	1,02	0,65	1,20	0,97
45	0,95	2,52	0,62	0,96
46	0,98	1,08	0,75	1,01
47	1,28	1,03	0,73	1,02
51	0,85	0,00	0,00	1,02
52	0,60	0,00	0,37	0,94
53	0,89	1,26	0,89	0,97
55	1,02	0,48	1,04	1,04
56	1,26	3,63	1,25	1,05
57	1,06	1,38	0,86	0,94
59	1,11	1,05	1,74	1,10
60	1,21	1,41	1,21	1,05
61	-	1,04	-	-

IDCOLEGIO	Eficiencia 1	Eficiencia 2	Eficiencia 3	EficMedia
62	1,07	1,19	1,28	1,02
63	1,28	1,75	1,00	1,04
64	1,03	1,89	0,52	1,05
65	-	8,91	-	-
66	1,02	1,19	1,12	1,01
67	1,23	0,73	0,51	1,02
68	0,86	1,26	1,74	1,00
70	1,12	1,41	0,92	0,97
71	-	2,95	-	-
72	0,72	0,00	0,74	0,98
73	2,23	0,00	1,92	1,06
74	0,97	0,99	0,85	0,96
75	1,16	1,06	0,29	1,01
76	0,78	0,00	1,01	0,95
77	1,12	1,11	1,20	1,00
78	1,24	1,45	1,18	0,98
79	1,07	1,12	1,14	1,04
80	0,90	0,51	0,36	0,92
81	0,70	0,37	0,28	0,91
82	1,38	5,00	3,89	1,02
83	0,72	0,00	0,42	1,00
84	1,18	1,34	0,95	0,99
85	1,17	0,86	1,97	1,03
86	1,45	1,26	0,92	1,05
87	1,24	0,00	2,20	1,06
88	0,22	1,06	0,00	0,98
89	1,22	1,01	1,17	1,08
90	1,24	1,01	0,91	0,96
91	0,74	2,27	0,21	0,90
92	0,92	0,35	0,86	0,94
93	0,87	1,30	0,79	0,96
94	1,11	-	1,18	1,00
95	0,57	0,00	1,00	0,91
96	1,17	1,51	1,59	1,11
97	0,97	2,57	0,89	1,02
98	0,80	0,47	0,61	0,99
99	1,01	0,74	0,20	0,99
100	0,90	1,05	2,11	1,03
101	1,00	0,95	1,42	1,06
102	1,21	1,00	1,07	1,05
103	0,70	1,05	0,70	1,01
104	1,15	1,52	1,50	1,02
105	1,27	0,00	0,41	0,94
106	1,75	2,65	0,83	1,00
107	0,95	1,00	0,00	1,01
108a	0,61	0,00	0,54	0,95
108b	1,06	1,60	1,82	0,93
109	0,74	1,29	1,09	1,04
111	0,56	0,00	0,33	0,96
112	1,27	0,82	1,41	1,03
113	0,92	1,08	1,87	1,06
115	0,26	0,00	0,23	0,91
116	0,70	0,91	0,68	0,95
117	1,65	2,82	1,73	1,11
118	1,14	1,27	2,00	1,04

IDCOLEGIO	Eficiencia 1	Eficiencia 2	Eficiencia 3	EficMedia
119	1,07	2,51	1,01	1,09
120	0,97	1,91	0,43	0,99
121	0,91	1,25	0,25	0,99
122	1,45	0,90	1,36	1,03
123	0,14	0,00	0,00	0,86
124	0,77	1,07	0,98	1,00
125	0,99	1,36	1,49	1,03
126	0,93	0,52	1,12	0,94
127	1,01	1,05	0,99	1,00
128	1,34	1,18	1,37	1,04
129	1,22	0,79	1,50	1,06
130	0,79	0,98	0,67	0,95
131	0,73	0,00	0,59	0,92
131	0,65	0,63	0,00	0,92
132	0,94	1,10	0,84	0,97
133	0,92	1,02	0,42	1,01
134	1,27	0,77	1,01	1,03
135	0,75	1,23	0,45	1,02
136	1,15	0,82	0,96	0,99
137	0,90	1,34	0,00	0,96
138	0,91	1,00	0,96	1,01
139	1,15	1,18	1,35	1,01
140	1,17	1,66	1,36	1,07
141	1,22	0,97	1,39	1,06
142	0,87	0,00	1,03	1,04
143	0,82	0,00	1,05	1,03
144	1,23	1,58	1,38	1,07
145	0,84	1,47	0,94	0,97
146	1,07	1,10	0,61	0,98
147	1,18	0,70	1,21	1,00
148	0,88	-	0,72	1,03
149	0,76	0,48	0,63	0,97
150	0,87	0,00	0,00	1,02
316	0,99	0,00	0,99	1,00
451	1,19	0,89	1,95	1,08
850	0,71	0,31	0,24	1,01
894	0,95	0,96	1,34	1,02
895	1,17	0,00	2,41	1,07
902	1,78	1,74	1,60	1,09

- → No se pudo calcular la medida de eficiencia para esta dimensión.

Las casillas sombreadas son colegios para los cuales la predicción era obtener un 0% de alumnos y obtuvieron un resultado positivo. A estos centros se les asignó el valor de eficiencia que obtuvo la escuela más eficiente.

ANEXO 5: Ranking de eficiencia técnica de cada escuela para cada output.

IDCOLEGIO	Ranking 1	Ranking 2	Ranking 3	RankMedia
1	110	52	105	113
2	137	137	138	93
3	92	98	43	85
4	105	1	1	128
5	66	122	54	49
6	11	36	55	70
7	82	105	106	126
8	102	42	58	98
9	73	81	19	111
10	85	107	118	110
11	59	95	44	92
12	12	26	18	18
13	50	103	89	131
14	112	-	96	71
15	8	88	39	58
17	136	102	113	136
18	-	116	-	-
19	142	140	140	140
20	52	18	2	13
21	27	83	86	127
23	117	54	110	96
24	76	-	36	6
25	44	63	7	107
26	42	121	38	69
27	128	132	37	74
28	125	32	93	108
29	54	22	51	47
30	10	46	21	23
31	-	10	-	-
32	30	11	33	7
33	109	1	76	100
34	64	13	131	88
35	138	109	122	130
36	121	76	112	104
37	56	104	1	67
38	4	-	8	26
40	14	30	63	41
41	104	87	125	132
42	40	92	40	45
43	55	51	6	38
44	67	100	48	97
45	84	8	101	109
46	77	58	91	62
47	15	69	94	56
51	106	126	137	48
52	133	134	119	125
53	97	40	81	103
55	69	110	62	31
56	20	3	42	27
57	62	31	84	121
59	51	67	16	3
60	31	29	45	25
61	-	68	-	-

IDCOLEGIO	Ranking 1	Ranking 2	Ranking 3	RankMedia
62	61	48	41	53
63	16	16	69	30
64	65	15	108	22
65	-	1	-	-
66	68	47	56	60
67	24	97	109	51
68	103	43	15	79
70	49	28	78	102
71	-	4	-	-
72	124	131	92	90
73	1	117	12	15
74	78	77	85	112
75	41	61	123	63
76	114	128	66	120
77	48	55	47	81
78	23	27	49	91
79	58	53	53	33
80	96	108	120	133
81	127	113	124	139
82	9	2	1	52
83	123	130	116	73
84	35	35	75	82
85	37	89	10	43
86	6	41	79	24
87	22	119	4	20
88	140	62	133	94
89	29	71	52	9
90	21	72	80	116
91	120	12	129	141
92	88	114	83	123
93	99	37	90	114
94	53	-	50	80
95	134	135	68	137
96	39	24	22	2
97	79	7	82	55
98	111	112	102	87
99	71	96	130	84
100	95	65	5	40
101	72	82	26	19
102	32	73	60	21
103	130	64	97	66
104	43	23	23	57
105	17	118	117	124
106	3	6	88	72
107	83	74	134	64
108a	132	133	107	118
108b	63	20	14	129
109	119	38	59	29
111	135	136	121	106
112	18	90	27	44
113	90	59	13	14
115	139	138	128	138
116	129	84	98	119
117	5	5	17	1
118	47	39	9	34

IDCOLEGIO	Ranking 1	Ranking 2	Ranking 3	RankMedia
119	60	9	67	5
120	80	14	114	86
121	91	44	126	89
122	7	85	31	37
123	141	139	139	142
124	115	60	72	77
125	74	33	25	39
126	87	106	57	122
127	70	66	70	78
128	13	49	30	28
129	28	93	24	17
130	113	78	99	117
131	122	129	104	134
131	131	101	135	135
132	86	56	87	105
133	89	70	115	61
134	19	94	65	42
135	118	45	111	54
136	46	91	74	83
137	94	34	132	115
138	93	75	73	68
139	45	50	34	59
140	38	19	32	10
141	26	79	28	16
142	100	124	64	32
143	108	127	61	36
144	25	21	29	12
145	107	25	77	99
146	57	57	103	95
147	34	99	46	76
148	98	-	95	35
149	116	111	100	101
150	101	125	136	50
316	75	123	71	75
451	33	86	11	8
850	126	115	127	65
894	81	80	35	46
895	36	120	3	11
902	2	17	20	4