

# UN MODELO ECONOMETRICO PREDICTIVO PARA CASTILLA Y LEON(\*)

(\*\*) *Jesús Cavero Alvarez*  
*Carmen Lorenzo Lago*  
*Beatriz Rodríguez Prado*  
*José Luis Rojo García*

RESUMEN.—Este artículo describe el primer modelo econométrico elaborado para Castilla y León. Se plantean en el mismo, tanto los problemas surgidos en la elaboración de las series utilizadas como la especificación y estimación del modelo. Finalmente, se evalúa la capacidad predictiva del modelo y se representan las predicciones de las principales variables endógenas para el período 1987 a 1991.

## Introducción

El trabajo que a continuación se presenta es un modelo econométrico para la predicción, a medio y largo plazo, de los valores añadidos, el empleo y la renta en Castilla y León, con un cierto grado de desagregación sectorial.

En general, en los modelos tratamos de recoger de forma simplificada las relaciones existentes entre conjuntos de variables en un sistema de referencia y para un período determinado.

En la terminología econométrica, diferenciamos entre variables endógenas o explicadas por el modelo y exógenas, que son explicativas de las anteriores.

(\*) Modelo realizado con una ayuda de la Junta de C-L.

(\*\*) En orden alfabético.

Hemos de señalar que cuando el objetivo fundamental de un modelo es la predicción de las variables endógenas en función de unos valores «supuestamente conocidos» de las exógenas, la especificación de las relaciones entre las variables se ve condicionada por los resultados más que por la Teoría Económica.

Pero además, si un modelo ha demostrado capacidad predictiva, también puede utilizarse para medir los efectos de distintas políticas económicas, es decir, lo que podría ocurrir en el futuro si las variables exógenas tomasen unos valores determinados.

En España, el desarrollo del Estado de las Autonomías ha planteado la posibilidad de que los Gobiernos regionales esbocen sus propias políticas económicas; en consecuencia, necesitan contar con medios propios de análisis regional que les permitan predecir los valores futuros de las principales magnitudes y evaluar el impacto de políticas alternativas.

Obviamente, estas necesidades no tienen por qué satisfacerse únicamente con la ayuda de los modelos econométricos. Pero, dentro de las técnicas de estimación más ampliamente utilizadas, estos modelos ocupan un lugar relevante junto con los modelos de Economía de Base y los que se derivan del análisis Input-Output.

Aunque el estudio de los modelos econométricos regionales se inicia en la década de los setenta en los Estados Unidos, sobre todo en el análisis de ciertas áreas metropolitanas, hemos de reconocer que actualmente no se dispone de una teoría generalmente aceptada sobre cómo elaborar un modelo regional, y que las debilidades que muestran estos modelos se derivan esencialmente del estado en que se encuentran las estadísticas regionales. Con todo, las experiencias del pasado ayudan a orientar y encauzar la labor investigadora y han proporcionado al equipo la necesaria dosis de humildad a la hora de plantearse los objetivos del trabajo.

Existen una serie de características que son comunes a la mayoría de los modelos econométricos regionales elaborados hasta el momento; esencialmente serían las siguientes:

- Los modelos estimados son anuales ya que, para la mayoría de las macromagnitudes regionales las estadísticas no suministran datos con periodicidad inferior a la anual.
- Al ser anuales, el número de observaciones para cada serie será reducido, y difícilmente rebasará la veintena.
- La periodicidad anual de las series, junto con las pocas disponibles, obligan a que los modelos sean pequeños en su número de ecuaciones y simples en la explicación de las variables endógenas.
- La cortedad de las series obliga asimismo a evitar en lo posible la inclusión de retardos de variables que facilitarían cierta explicación dinámica del comportamiento de las magnitudes.
- Finalmente, la estructura de los modelos viene limitada en una doble dirección por la escasez en el número de series disponibles; en primer lugar, obliga

a utilizar como variables exógenas las variables nacionales, lo que limita la posibilidad de explicar internamente ciertos procesos; en segundo, limita el número de variables a predecir, ya que no se puede predecir una variable sin contrastar empíricamente su relación con otras y, por tanto, sin conocer su historia.

En España, la escasez y poca calidad de las estadísticas regionales hasta el presente ha impedido formalizar modelos econométricos regionales de cierta importancia, siendo el que presentamos uno de los pioneros. El modelo está enmarcado en el proyecto HISPALINK que reúne a diferentes equipos de carácter y procedencia universitaria con el objetivo de desarrollar modelos econométricos para sus comunidades respectivas.

## Base estadística del modelo

### *Desagregación por ramas*

El modelo se plantea desagregar los datos referentes a la producción en nueve ramas. Estas ramas han sido seleccionadas de acuerdo con el proyecto HERMES (modelo integrado para la CEE) y son las siguientes:

A: Agricultura, silvicultura y pesca.

E: Energía y agua.

Q: Productos manufacturados intermedios.

K: Productos manufacturados de equipo.

C: Productos manufacturados de consumo.

B: Obras públicas y Construcción.

Z: Transportes y comunicaciones.

L: Otros servicios designados a la venta excepto alquileres de inmuebles.

G: Servicios no destinados a la venta.

La utilización en los modelos nacionales de esta desagregación resulta coherente con la CNAE-74, ya que no rompe el esquema de agrupaciones de ésta. No obstante, crea problemas cuando se intenta plasmar en la modelización regional. Como luego comentaremos ha sido necesario agregar, en esta fase de desarrollo del modelo, las tres últimas ramas en una sola, a la que genéricamente hemos denominado de «Servicios».

La utilización de esta desagregación plantea problemas en el sector industrial derivados de la diversidad de fuentes estadísticas que ha sido necesario utilizar; en efecto, como el modelo exige series desde 1970 a 1986, los datos de valores añadidos de las ramas industriales desde 1970 a 1977 provienen de las antiguas publicaciones del Servicio Sindical de Estadística (Estadísticas de Producción Industrial), con una desagregación acorde con la CNAE-52, mientras que los posteriores (1978 a 1984) provienen de la Encuesta Industrial del INE, que está de acuerdo

con la CNAE-74, y de la que se obtiene por agregación la clasificación del HERMES.

### *Series utilizadas*

#### *Empleo (población activa ocupada).*

En general, se ha utilizado la EPA entre los años 1973 y 1986 (mediadas anuales). Los datos correspondientes a los tres primeros años se han estimado en función de ciertos indicadores de actividad y del comportamiento de la misma magnitud a nivel nacional.

Se han planteado, con todo, varios problemas: en primer lugar, el empleo de «Servicios no destinados a la venta» no se ha conseguido aislar en esta fase de desarrollo del modelo lo que, unido a las dificultades encontradas en la estimación del empleo en «Transportes y Comunicaciones» (por su particularidad) ha obligado a unir ambos sectores con «Servicios destinados a la venta» en uno solo.

En las ramas industriales, aún con la diversidad de datos que ofrecen la Encuesta Industrial y la Estadística de Producción Industrial, la E.P.A. permite obtener el empleo de forma suficientemente razonable.

En definitiva, se han realizado estimaciones del empleo regional para los años 1970 a 1986 y para las siete ramas mencionadas. Las variables obtenidas aparecen en la relación final de este apartado.

#### *Valores añadidos brutos a precios de mercado*

Se han estimado series de valores añadidos brutos a precios de mercado (VABpm) en pesetas constantes de 1980 para las siete ramas.

Para los sectores industriales, se han utilizado los datos de la Encuesta Industrial del INE para el subperiodo 1978 a 1984. Para los años 1970 a 1977 la única fuente disponible es la Encuesta de Producción Sindical de Estadística. Su utilización, con todo, plantea problemas importantes:

1. Las diferencias metodológicas de las subseries. Hubiera sido razonable elaborar al menos un año ambas estadísticas de forma que, al menos para la industria en su conjunto, el empalme fuera posible.

2. El desglose provincial y no regional de la vieja Estadística de Producción Industrial no hace sino añadir problemas. Obviamente, la responsabilidad no es del Servicio Sindical de Estadística, ya que la configuración autonómica actual es posterior a dicha publicación. Con todo, téngase en cuenta que dicha estadística excluye los datos provinciales cuando existen menos de tres establecimientos para salvaguardar el secreto estadístico, lo que obliga a repartir entre ellas la diferencia entre el total nacional y la suma de las magnitudes correspondientes a las provincias para las que no se ha aplicado el secreto estadístico, utilizando como criterio

de reparto el empleo en cada provincia (la única variable, junto con el número de establecimientos, no afectada por dicha cautela).

3. El problema tal vez más importante que plantea la utilización de la Encuesta de Producción Industrial es el hecho de que se utiliza la CNAE-52 en lugar de la de 1974, como antes mencionamos. Las agrupaciones de ambas clasificaciones no coinciden, se entremezclan, y ello ocurre no sólo para agrupaciones marginales en cuanto a su producción en la región, sino también para aquéllas de gran peso en el aparato industrial de Castilla y León. No vamos a descender en este artículo al desglose de las diferencias, pero digamos que ello ha obligado a repartos del valor añadido estimados, y a la introducción en la modelización de una variable ficticia, D7778, que salve la discrepancia metodológica entre ambas subseries.

Por lo que se refiere a los Servicios, ha resultado imposible, hasta el momento, estimar el valor añadido del sector y más aun, imputar el correspondiente a los servicios centrales de las instituciones financieras por lo que, aunque se dispone en estos momentos de buenas estimaciones para la rama de Transportes y Comunicaciones, ha resultado imposible separar las tres ramas de Servicios. El equipo trabaja actualmente en dicha separación, y creemos que es posible alcanzarla en la segunda fase del modelo. En su lugar ha habido que recurrir, provisionalmente, a repartir las cifras de Contabilidad Nacional en función de la información que suministra el Banco de Bilbao en sus estudios «Renta Nacional de España y su distribución provincial».

#### Relación alfabética de series del Banco de datos

(Acompañadas de un asterisco aparecen las variables tomadas del modelo WHARTON-UAM)

CLN	Empleo Sectores Clave Nacionales
CLR	Empleo Sectores Clave Regionales
DBAG	Variable ficticia para años climatológicamente buenos en agricultura
DMAG	Variable ficticia para años climatológicamente malos en agricultura
D7677	Variable ficticia que recoge cambio de metodología en la obtención de las series nacionales
D7778	Variable ficticia que recoge cambio de metodología en la Encuesta Industrial
D8586	Variable ficticia que recoge cambio de metodología en la EPA
EAGR	Empleo regional en agricultura
EBCR	Empleo regional en bienes de consumo
EBER	Empleo regional en bienes de equipo
EBIR	Empleo regional en bienes intermedios
ECOR	Empleo regional en construcción
EENR	Empleo regional en energía
ESRR	Empleo regional en servicios

ETR	Empleo regional total
FCLN	Empleo estimado en sectores clave nacional
FCLR	Empleo estimado en sectores clave regional
FEAGR	Empleo estimado regional en agricultura
FEBCR	Empleo estimado regional en bienes de consumo
FEBER	Empleo estimado regional en bienes de equipo
FEBIR	Empleo estimado regional en bienes intermedios
FECOR	Empleo estimado regional en construcción
FEENR	Empleo estimado regional en energía
FESRR	Empleo estimado regional en servicios
FNCLR	Empleo estimado sectores no clave regional
FVAGR	Valor añadido estimado regional en agricultura
FVBCR	Valor añadido estimado regional en bienes de consumo
FVBER	Valor añadido estimado regional en bienes de equipo
FVBIR	Valor añadido estimado regional en bienes intermedios
FVCOR	Valor añadido estimado regional en construcción
FVENR	Valor añadido estimado regional en energía
FVSRR	Valor añadido estimado regional en servicios
FVTR	Valor añadido estimado total
FYDR	Renta disponible regional estimada
LCLR	Logaritmo empleo sectores clave regional
LNCLR	Logaritmo sectores no clave regional
NCLR	Empleo en sectores no clave regional
PBAOAG	Población activa ocupada nacional en agricultura(*)
PBAOBC	Población activa ocupada nacional en bienes de consumo(*)
PBAOBE	Población activa ocupada nacional en bienes de equipo(*)
PBAOBI	Población activa ocupada nacional en bienes intermedios(*)
PBAOCO	Población activa ocupada nacional en construcción(*)
PBAOEN	Población activa ocupada nacional en energía(*)
PBAOSR	Población activa ocupada nacional en servicios(*)
PBAOTT	Población activa ocupada nacional total(*)
PIBR	Producto interior bruto real nacional(*)
T	Tiempo
VAGR	Valor añadido regional en agricultura
VARAG	Valor añadido nacional en agricultura(*)
VARBC	Valor añadido nacional en bienes de consumo(*)
VARBI	Valor añadido nacional en bienes intermedios(*)
VARCO	Valor añadido nacional en construcción(*)
VAREN	Valor añadido nacional en energía(*)
VARSR	Valor añadido nacional en servicios(*)
VARTT	Valor añadido nacional total(*)
VBCR	Valor añadido regional en bienes de consumo

VBER	Valor añadido regional en bienes de equipo
VBIR	Valor añadido regional en bienes intermedios
VCOR	Valor añadido regional en construcción
VENR	Valor añadido regional en energía
VSRR	Valor añadido regional en servicios
VTR	Valor añadido regional total
YDR	Renta familiar disponible regional
ZBC = VENR + VBER + VBIR	
ZBE = VENR + VBCR + VBIR	
ZBI = VENR + VBCR + VBER	
ZEN = VBCR + VBER + VBIR	

## Descripción del modelo

Como ya hemos comentado, los modelos econométricos regionales básicamente se encuentran condicionados, tanto en lo que respecta a su modelización como a los objetivos que persiguen, por la cantidad de la información estadística disponible.

En este primer intento para realizar proyecciones a corto y medio plazo de la economía Castellano-Leonesa mediante un modelo econométrico, se ha fijado, como objetivo principal, para las siete ramas de actividad mencionadas en el capítulo anterior, la obtención de predicciones del empleo y los valores añadidos en el período 1987-91.

Una primera cuestión que debemos destacar es que el funcionamiento de cualquier economía regional guarda una relación muy estrecha con el de la economía nacional, y por tanto, el modelo econométrico regional ha de estar relacionado con el marco nacional. Puesto que aceptamos la conexión de nuestro modelo con el modelo WHARTON-UAM, que proporciona predicciones de las variables nacionales más significativas, estas variables las consideramos siempre como exógenas, es decir que influyen sobre las variables regionales pero no al contrario.

Por otro lado, dada nuestra intención de que el modelo pueda servir para una posible planificación regional, se ha tratado de introducir variables exógenas regionales que puedan utilizarse como instrumentos en la política económica regional. No obstante, hemos de reconocer que la falta de series históricas, por el momento, de las macromagnitudes que hacen referencia al sector público regional, ha limitado este objetivo.

Así mismo, algunas relaciones económicas entre variables que a priori parecen tener un claro significado, en nuestro caso no se han mostrado significativas debido fundamentalmente a las imperfecciones de las series históricas regionales utilizadas.

La mayoría de los autores reconocen que a la hora de elaborar un modelo econométrico nacional, el paso de la teoría a la práctica es una arte en el que, desgraciadamente, las normas de actuación no están siempre claras. En el caso de los modelos regionales el proceso es más bien «artesanal», debido a que prácticamente no existe una teoría en que basarse. No obstante, en la convicción de que no hay modelos correctos, sino útiles, aceptamos de antemano que el modelo que presentamos a continuación, tal vez no sea el mejor que se pueda realizar para cubrir el objetivo descrito, pero supone un primer paso que conducirá, en un futuro próximo, a modelos más completos y precisos.

El modelo se ha estructurado en tres bloques; en los dos primeros se determinan el empleo y el valor añadido para cada rama de actividad y en el tercero, la renta regional disponible.

Las variables monetarias están medidas en términos reales, es decir, en pesetas constantes del 80 y por tanto, su predicciones vendrán expresadas en las mismas unidades.

A continuación pasamos a describir brevemente las ecuaciones que componen cada uno de los bloques.

### *Bloque 1 (Empleo)*

Lo primero que hemos de señalar, es que en este bloque hacemos referencia únicamente a la demanda de mano de obra.

El equipo que ha realizado este trabajo, está efectuando asimismo un estudio demográfico sobre la población de Castilla y León, que proporcionará proyecciones de población por grupos de edad y sexo, y que, al aplicar posteriormente las tasas de actividad a cada grupo, permitirá deducir la población económicamente activa, es decir, la oferta de mano de obra.

La razón fundamental para estudiar por separado la demanda y la oferta de mano de obra es que, mientras en la determinación de la demanda influyen gran cantidad de variables de diferentes tipos, en la determinación de la oferta influyen esencialmente variables demográficas (mortalidad, fecundidad y movimientos migratorios) que, dado su comportamiento más o menos regular, permiten la utilización de modelos específicos más generalizables a la mayoría de las regiones.

La comparación entre las proyecciones obtenidas para la demanda y oferta de mano de obra, proporcionará una visión global sobre la futura evolución del mercado de trabajo que permita la adopción de las medidas más convenientes para evitar o paliar los desajustes que se puedan presentar; más concretamente, para afrontar el principal problema con que se encuentra la economía española en general, y la castellano-leonesa en particular: el paro.

Para obtener las previsiones del empleo regional, en nuestro caso, después de ensayar varios procedimientos, se ha optado por seguir el desarrollado por un-gru-

po de investigadores de la Universidad de Carolina del Norte en Greensboro (Barnes y otros (1977)).

En este procedimiento, los sectores de la región sujetos a estudio son clasificados en sectores clave y no clave. Los sectores clave son aquellos cuyo crecimiento está principalmente asociado al crecimiento del mercado nacional; generalmente, abarcan a una gran parte de la actividad industrial de la región; en nuestro modelo, hemos incluido también en este grupo al sector agrícola por su importancia en nuestra región. Los sectores no clave son aquéllos cuyo nivel de actividad depende básicamente de la actividad dentro de la región. Como la producción de los sectores no clave es ya un producto final para los consumidores locales o un insumo para los sectores clave locales, los niveles de empleo en dichos sectores se supone que están determinados por el empleo en los sectores clave regionales.

En este modelo, se tienen en cuenta explícitamente las interdependencias que existen entre el mercado de trabajo regional y el nacional, tanto dentro de los sectores como entre ellos. Esto se lleva a cabo explicando los niveles de empleo regionales a través de un simple modelo econométrico en dos fases.

En la primera fase, el empleo regional total está determinado por las siguientes ecuaciones:

- (a)  $CLR = f(CLN)$
- (b)  $NCLR = f(CLR)$
- (c)  $ETR = CLR + NCLR$

donde CLR es el empleo regional en el conjunto de todos los sectores clave, CLN es el empleo nacional para ese mismo grupo de sectores clave, NCLR es el empleo regional en el conjunto de los sectores considerados como no clave y ETR es el empleo regional total.

La ecuación (a) plantea que el empleo regional en el llamado sector clave es una función del empleo en el sector clave a nivel nacional. La ecuación (b) plantea que el empleo en el sector no clave regional en una función del empleo en el sector clave regional. La ecuación (c) es una identidad que plantea que el empleo total en la región es la suma del empleo en los sectores clave y no clave.

En la segunda fase, para repartir el empleo regional entre diferentes sectores, se utiliza una serie de ecuaciones similar, aunque más detallada:

- (d)  $ECLR_i = f(ECLN_i, ETR)$
- (f)  $ENCLR_j = f(\sum_i ECLR_i)$
- (g)  $ETR = \sum_i ECLR_i + \sum_j ENCLR_j$

donde  $ECLR_i$  es el empleo regional en el sector clave  $i$ ,  $ENCLR_j$  es el empleo regional en el sector no clave  $j$ ,  $ECLN_i$  es el empleo nacional en el sector clave  $i$ .

La ecuación (d), acepta que el empleo regional en el sector clave  $i$ , es una función del empleo nacional en ese mismo sector y del empleo total regional. La

ecuación (f) plantea que el empleo regional en el sector no clave  $j$  es una función de la suma del empleo en todos los sectores clave.

El modelo anteriormente expuesto es esencialmente un modelo recursivo, esto es, la cadena causal se toma como unidireccional, pudiéndose estimar por mínimos cuadrados ordinarios (LS) (Ver, por ejemplo J. Johnston: «Método de Econometría». Ed. Vicens-Vives, 1975).

Asimismo, este bloque de empleo, en conjunto, también es recursivo respecto a los otros dos, es decir, las variables determinadas en este bloque intervienen como explicativas en alguna de las ecuaciones restantes del modelo, pero las variables endógenas de los otros bloques no intervienen en la explicación del empleo. Esto permite afrontar la estimación de este bloque de forma independiente a la del resto del modelo y evitar así los problemas que surgirían al aplicar los métodos de estimación apropiados para el caso de variables interdependientes, al no disponer de muestras de tamaño suficiente.

En la región considerada, Castilla y León, hemos clasificado como sectores clave a la agricultura, bienes intermedios y bienes de consumo (ramas A, Q y C). Las tres ramas de actividad regional se encuentran fuertemente asociadas a su comportamiento en el mercado nacional. En el caso de la agricultura, no es necesario justificarlo, puesto que es evidente; y respecto a bienes intermedios y bienes de consumo, la razón fundamental es la fuerte implantación en la región de empresas dedicadas a fabricación de vehículos unas y de productos alimenticios otras.

Las otras cuatro ramas de actividad (energía, bienes de equipo, construcción y servicios, se han clasificado como sectores no clave.

El bloque de empleo se encuentra, de esta forma, compuesto por 14 ecuaciones de las cuales 9 son estocásticas, o de comportamiento, y cuya estimación se presenta al final de este apartado (ecuaciones 3, 5, 8, 9, 10, 11, 12, 13, y 14). En las ecuaciones 5 y 13, correspondientes a la determinación del empleo en el sector no clave regional y en servicios, se ha incorporado una variable ficticia (D8586) para recoger los efectos del cambio metodológico de la Encuesta de Población Activa producido en el año 1986.

## *Bloque 2. Valores Añadidos*

Una vez descrito el procedimiento que proporciona los empleos regionales en función de las respectivas series nacionales, nos ocuparemos ahora de describir el bloque de ecuaciones que permiten obtener los valores añadidos para cada rama de actividad.

La total inexistencia tanto de series, a nivel regional, como incluso de variables «próximas» del grado de utilización del empleo, así como de las series de capital y su grado de utilización, obliga a enfocar las ecuaciones de producción desde el lado de la demanda.

En este sentido, las variables más utilizadas en experiencias internacionales para explicar el valor añadido en cada sector, hacen referencia a la comparación entre los precios de mercado regional y nacional, la diferencia entre los costes de producción, la renta familiar disponible regional y los respectivos valores añadidos nacionales. En nuestro caso, sólo se puede disponer de información sobre las dos últimas variables, por lo que ha sido necesario buscar enfoques alternativos en la especificación de las ecuaciones, basándonos en variables específicas para cada sector, en el marco de la información disponible.

En concreto, los sectores de agricultura, construcción y servicios se han explicado de forma independiente entre sí, mientras que las cuatro ramas de actividad industrial (energía, bienes de equipo, bienes intermedios y bienes de consumo) forman un conjunto de ecuaciones interdependientes unas de otras.

El valor añadido regional en la agricultura (VAGR), ecuación 15, se hace depender del empleo regional en el sector (EAGR), de la evolución de la variable a nivel nacional (VARAG), del tiempo (T), que recoge los efectos de los cambios tecnológicos, y de dos variables ficticias, que explican la influencia climatológica sobre las cosechas anuales: DBAG, para los años excepcionalmente buenos y DMAG, para los excepcionalmente malos.

La ecuación 16 explica el valor añadido en la construcción (VCOR) a través del propio valor observado en el periodo anterior (VCOR(-1)), por motivo de la influencia que tienen los retardos en esta actividad, el empleo regional del sector (ECOR) y el tiempo, que en este caso, recoge las tendencias de variables demográficas y sociales.

En el sector servicios (VSRR), ecuación 17, las variables que se han mostrado más significativas son el empleo total regional (ETR), como variable «vicaria» a la actividad económica regional, y la renta familiar disponible regional (YDR); también se ha incorporado la variable ficticia D8586, que recoge el cambio metodológico de la EPA, como ya se explicó al describir el bloque del empleo.

Para la obtención de las ecuaciones correspondientes a las ramas de actividad industrial, hemos considerado entre otras cosas:

- Que los sectores son mixtos, es decir, que producen tanto para la «exportación» como para la demanda local. Esto se refleja en la incorporación en cada sector, de variables específicas tanto nacionales como regionales.
- Que las actividades industriales se encuentran fuertemente relacionadas, de forma que los incrementos o disminuciones de producción de un sector influyen en los demás sectores. Por este motivo, incorporamos como variables explicativas las «Z» (ZBE, ZBI, ZEN y ZBC) que representan la suma de los valores añadidos de los sectores industriales no explicados en la ecuación correspondiente.

Con todo ello, las ecuaciones quedan como sigue:

El valor añadido regional en energía (VENR), ecuación 19, depende de la variable ZEN (suma de VBER, VBIR y VBCR), el empleo regional (EENR) y la renta familiar disponible (YDR).

En el caso de los bienes intermedios, VBIR, que aparece en la ecuación 21, es función de la evolución de las restantes ramas industriales (ZBI) y de su correspondiente valor nacional (VARBI).

El valor añadido en bienes de consumo (VBCR), ecuación 23, asimismo depende de ZBC, suma de valores añadidos de las otras tres ramas, y del producto interior bruto nacional (PIBR).

Por último, VBER, valor añadido en bienes de equipo, se explica en la ecuación 25 por ZBE, el valor nacional (VARBE), y el tiempo.

En las ecuaciones 19, 23 y 25, aparece también la variable ficticia D7778, que recoge la sustitución de una encuesta industrial por otra entre los años 1977 y 1978.

La estimación de estas últimas cuatro ecuaciones junto a las del valor añadido en servicios (26) y de la renta familiar disponible (27), se han realizado por el procedimiento de mínimos cuadrados en dos etapas.

### *Bloque 3. Renta Familiar Disponible Regional*

Este bloque tiene como objetivo endogeneizar la Renta Familiar Disponible Regional. Para ello, se ha incorporado la ecuación 27 en la que la renta (YDR) se expresa en función del valor añadido total regional (VTR) y del tiempo; también se incorpora la variable ficticia D7778 por los motivos anteriormente expuestos.

El modelo se compone, por tanto, de 27 ecuaciones, en las cuales se explica el comportamiento de 23 variables endógenas mediante 19 exógenas.

### *Especificación del modelo*

- 1:  $clr = eagr + ebir + ebcr$
- 2:  $cln = pbaoag + pbaobi + pbaobc$
- 3:  $CLR = -85936.648 + .10576847 * CLN$
- 4:  $lclr = \log(clr)$
- 5:  $LNCLR = 9.6902752 + .25354037 * LCLR + 4.2122136D - 2 * D7677 + .14336068 * D8586 + [AR(1) = .6925417]$
- 6:  $nclr = \exp(lnclr)$
- 7:  $etr = clr + nclr$
- 8:  $EAGR = -157926.03 + 9.7856007D - 2 * PBAOAG + .23806659 * ETR$
- 9:  $EBIR = -276.2125 + 4.6412679D - 2 * PBAOBI + 1.2923463D - 3 * ETR$
- 10:  $EBCR = 6914.578 + 5.2134392D - 2 * PBAOBC + 6.1713205D - 3 * ETR$
- 11:  $EENR = 550.06827 - 1.4242318D - 2 * CLR + .1714414 * PBAOEN + [AR(1) = .65149603]$

- 12:  $EBER = 12024.88 - 3.6191468D - 2 * CLR + 5.5440762D - 2 * PBAOBE + [AR(1) = .38569497]$
- 13:  $ESRR = 252462.8 + .11572143 * CLR + 56896.51 * D8586 + [AR(1) = .64510711]$
- 14:  $ECOR = 22698.738 - 3.8301931D - 2 * CLR + 5.7183956D - 2 * PBAOCO$
- 15:  $VAGR = -11206805. + .43901499 * EAGR + 5579.5156 * T + .15176989 * VARAG + 15666.763 * DBAG - 1951.134 * DMAG$
- 16:  $VCOR = -3132647.4 + .34052596 * VCOR(-1) + 1572.1385 * T + 1.0201336 * ECOR$
- 17:  $VSRR = 185779.86 - .13941414 * ETR + .49498641 * YDR + 33491.924 * D8586$
- 18:  $zen = vber + vbir + vbcr$
- 19:  $VENR = -315484.41 + .43123861 * ZEN + 6.4608125 * EENR + .18840914 * YDR + 26993.21 * D7778$
- 20:  $zbi = vber + venr + vbcr$
- 21:  $VBIR = 7505.0717 + 6.6480646D - 2 * ZBI + 1.1428326D - 2 * VARBI$
- 22:  $zbc = venr + vbir + vber$
- 23:  $VBCR = -27050.29 + .54942679 * ZBC + .003152 * PIBR - 21114.035 * D7778$
- 24:  $zbe = venr + vbir + vbcr$
- 25:  $VBER = 5462502.7 - .39279065 * ZBE + 7.6905123D - 2 * VARBE + 88903.568 * D7778 - 2778.0804 * T$
- 26:  $vtr = vber + vbcr + vbir + venr + vsrr + vcor + vagr$
- 27:  $YDR = -6398669.2 - 121046.78 * D7778 + 3318.4587 * T + .75135118 * VTR$

## Evaluación del modelo

Aunque, por motivos de espacio, no se presentan todos los resultados estadísticos de las estimaciones, podemos afirmar que las medidas de bondad para cada una de las ecuaciones, así como la significación de los parámetros es bastante aceptable. Pero el hecho de que todas las ecuaciones, de forma aislada, se ajusten bien a los datos no garantiza que el modelo en su conjunto, cuando sean simuladas, reproduzca aquellas mismas series en forma ajustada, ya que el modelo tiene una estructura dinámica más rica que cada una de las ecuaciones individuales que lo componen.

Para conocer la eficacia del modelo, es necesario efectuar una simulación del mismo sobre datos pasados, y obtener alguna medida sobre la precisión de la simulación realizada.

La simulación del modelo no sólo es importante en sí misma, puesto que proporciona los valores de las variables endógenas (las que pretende explicar) en fun-

ción de las exógenas (explicativas), sino que puede ayudar a localizar los puntos más débiles del modelo y a diseñar un posible plan de mejora en el proceso de predicción. En este sentido, hemos de señalar que los resultados que hemos presentado son la consecuencia de múltiples ensayos en los que hemos utilizado diferentes especificaciones para las ecuaciones, basadas en teorías diferentes; así como de un largo proceso de selección de variables en el que siempre se ha tenido presente la finalidad predictiva del modelo.

Entre los diferentes criterios que hemos utilizado para evaluar la capacidad predictiva de un modelo, presentamos dos de ellos: El Error Cuadrático Medio y el estadístico U de Theil.

### *Errores cuadráticos medios*

Una gran parte de las medidas de precisión se basan en el valor del error cuadrático medio (E.C.M.). Sea:

$A_{t-1}$ , el valor observado de una variable en el período t-1

$A_t$ , el valor observado de la variable en el (año objetivo de predicción).

$F_t$ , la predicción de la variable, para el año objetivo obtenida con el modelo.

Llamemos «cambio observado relativo» a  $a_t = (A_t - A_{t-1})/A_{t-1}$  y «cambio proyectado relativo» a  $f_t = (F_t - A_{t-1})/A_{t-1}$ .

Se define, para un conjunto de n predicciones, la raíz del error cuadrático medio, R.E.C.M., para esa variable, como:

$$\text{RECM} = \sqrt{(1/n) \sum_{i=1}^n (f_t - a_t)^2} = \sqrt{(1/n) \sum_{i=1}^n ((F_t - A_t)/A_{t-1})^2}$$

La opinión general es considerar que el resultado es satisfactorio si esta medida tiene un valor inferior a 3.5, si bien este valor puede variar según el tipo de problema que se considere y la antelación con que se efectúen las predicciones.

En la tabla que aparece al final de este apartado presentamos los valores de RECM de las variables endógenas más representativas del modelo, obtenidos en la simulación dinámica, para el período 1970-86. Esta simulación en la que sólo se suponen conocidos los valores de las variables exógenas y el valor inicial de las endógenas, es la que permite contrastar la estabilidad del modelo y la calidad de sus predicciones.

Como podemos observar, en conjunto, el poder predictivo del bloque que proporciona el empleo en la diferentes ramas de actividad, es bastante superior al del bloque de los valores añadidos.

Se puede afirmar, en general, que el modelo proporciona a priori buenas predicciones del empleo. En cambio, tiene más dificultades para reproducir las series de los valores añadidos de algunas ramas de actividad, obteniéndose los peores re-

sultados en el caso del valor añadido en energía y en bienes de consumo. No obstante, los resultados no pueden clasificarse como malos si los comparamos con otros modelos, pues siempre hay partes o ecuaciones más difíciles de modelizar por diversas causas.

En lo que respecta a la simulación de la renta familiar disponible regional, puede considerarse aceptable, y su precisión mejorará sustancialmente cuando se atenúen los errores en los valores añadidos.

TABLA

Raíz de los Errores Cuadráticos Medios (en porcentaje) Estadístico  $U_{66}$  de Theil

Variable	denominación	RCEM	$U_{66}$
Clave Regional	CLR	2,04	0,38
No Clave Regional	NLCR	1,49	0,31
Empleo Total Reg.	ETR	1,25	0,31
Emp. Agricultura	EAGR	2,73	0,40
Emp. Energía	EENR	2,82	0,45
Emp. Bienes Interm.	EBIR	2,11	0,27
Emp. Bienes Equipo	EBER	2,54	0,24
Emp. Bienes Consumo	EBCR	3,27	0,59
Emp. Servicios	ESRR	1,54	0,27
Emp. Construcción	ECOR	3,78	0,47
Valor Añad. Agricul.	VAGR	4,94	0,27
Valor Añad. Construc.	VCOR	5,58	0,63
Valor Añad. Servicios	VSRR	2,96	0,82
Valor Añad. Energía	VENR	37,78	0,29
Valor Añad. B. Inter.	VBIR	5,37	0,35
Valor Añad. B: Consumo	VBCR	10,65	0,14
Valor Añad. B. Equipo	VBER	7,84	0,32
Renta Familiar Disponible	YDR	3,04	0,87

### Estadísticos $U$ de Theil

En muchas ocasiones, para determinar la validez predictiva de un modelo, es aconsejable comparar sus resultados con los que obtienen con otros modelos, en particular, con el modelo más simple que se puede utilizar, como es el basado en el supuesto de inexistencia de cambios, es decir, aquel en el que la predicción es igual a la última observación de la variable (modelo «ingenuo»).

En este sentido, se define el estadístico  $U$  Theil (1966) como:

$$U_{66} = \sqrt{\frac{\sum_{t=1}^n \left[ \frac{F_t - A_t}{A_{t-1}} \right]^2}{\sum_{t=1}^n \left[ \frac{A_t - A_{t-1}}{A_{t-1}} \right]^2}} = \sqrt{\frac{\text{ECM}}{\sum a_t^2}}$$

que puede tomar valores entre cero e infinito. Cuanto más pequeño sea el valor de  $U$  mejor será el poder predictivo del modelo.

Este índice posee además la siguiente propiedad:

- Si  $U = 1$ , el modelo «ingenuo», donde  $F_t$  sería igual a  $A_{t-1}$ , es tan bueno como el modelo que se evalúa.
- Si  $U$  es menor que 1, el modelo predictivo utilizado es mejor que el modelo «ingenuo».
- Si  $U$  es mayor que 1, entonces es preferible no utilizar el modelo evaluado, ya que el modelo «ingenuo» proporciona mejores resultados.
- Si  $U = 0$ , la predicción es perfecta.

Atendiendo, pues, a este criterio del estadístico  $U$ , comprobamos en la tabla anterior que, para todas las variables endógenas de nuestro modelo, el valor de  $U$  es inferior a la unidad y, en casi todos los casos, también es inferior a 0.5.

Curiosamente, para los valores añadidos de la energía y bienes de consumo, que obtenían los peores resultados respecto al criterio del ECM, se obtienen, ahora, los valores más bajos de la  $U$ , mostrándose el modelo, para estas variables, con gran capacidad predictiva.

La razón de esta aparente contradicción es que el estadístico  $U$  tiene en cuenta, de alguna forma, las dificultades «a priori» de la serie sobre la que se quieren efectuar predicciones, de forma que el denominador toma valores relativamente más altos para aquellos casos en que la variable presenta más oscilaciones; por lo que puede constituir, a veces, un criterio más adecuado.

De todo lo anterior, podemos concluir que, aún con todas las limitaciones ya comentadas, a la hora de efectuar conjeturas sobre estas variables, con vistas a lograr determinados objetivos de política económica regional, es preferible disponer del modelo que hemos presentado, a tener que estar condicionados por intuiciones más o menos acertadas.

### **Predicción del empleo, valores añadidos y renta regional para el período 1987-1991**

A continuación, pasamos a presentar las predicciones obtenidas con el modelo para las variables endógenas. Estos valores son los que proporciona directamente el modelo, es decir, no se han incorporado factores de ajuste, ya que los errores observados en la simulación no parecen tener un comportamiento sistemático. De hecho, el modelo tiene un comportamiento más aceptable, para los datos históricos, sin incorporar en el mismo los factores de ajuste más socorridos en la práctica, los últimos errores observados.

Para poder realizar la predicción son necesarios los valores, para el periodo considerado, de las variables exógenas del modelo. Las predicciones de los valores

nacionales han sido facilitadas por CEPREDE (Centro de Predicción Económica) y corresponden a las que proporciona el modelo WHARTON-UAM de la economía española. En el caso de algunas variables, ha sido necesario recurrir a un criterio de reparto, ya que de momento, las predicciones nacionales no se obtienen al mismo nivel de desagregación que nosotros utilizamos.

Las otras variables exógenas son el tiempo y las correspondientes ficticias.

La precisión de nuestras predicciones está condicionada por la bondad del modelo, por el hecho de que en el futuro no se produzcan fuertes cambios estructurales en la economía regional y por la propia precisión de los valores nacionales utilizados.

Hemos supuesto, asimismo, que no se van a producir condiciones meteorológicas excepcionalmente buenas o malas que provoquen cambios extraordinarios en los valores añadidos de la agricultura, es decir, consideramos que DBAG y DMAG no van a tener influencia y sus valores, por tanto, serán iguales a cero.

Se ha de señalar que la variable VAGR, en análisis de sensibilidad realizados, se ha mostrado con una influencia notable sobre los resultados que se obtienen para otras ramas de actividad. No obstante, siempre es posible especificar escenarios alternativos para las condiciones climatológicas.

De forma resumida, podemos describir nuestras predicciones como sigue:

El empleo total regional continuará descendiendo, pero sin llegar a los ritmos de decrecimiento de los últimos años. Este descenso se deberá, sobre todo, a las pérdidas de empleo en agricultura y servicios. El nivel se mantendrá más o menos constante en las ramas industriales. Destaca por el contrario, el crecimiento del empleo que experimentará el sector de la construcción.

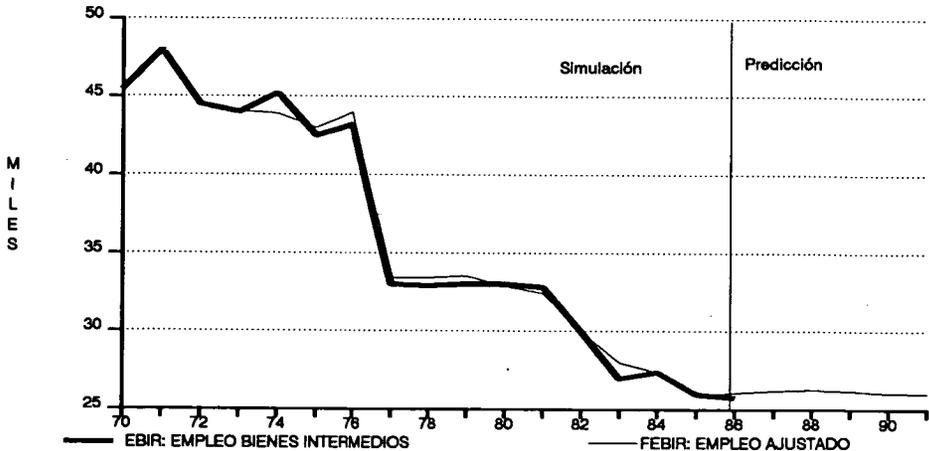
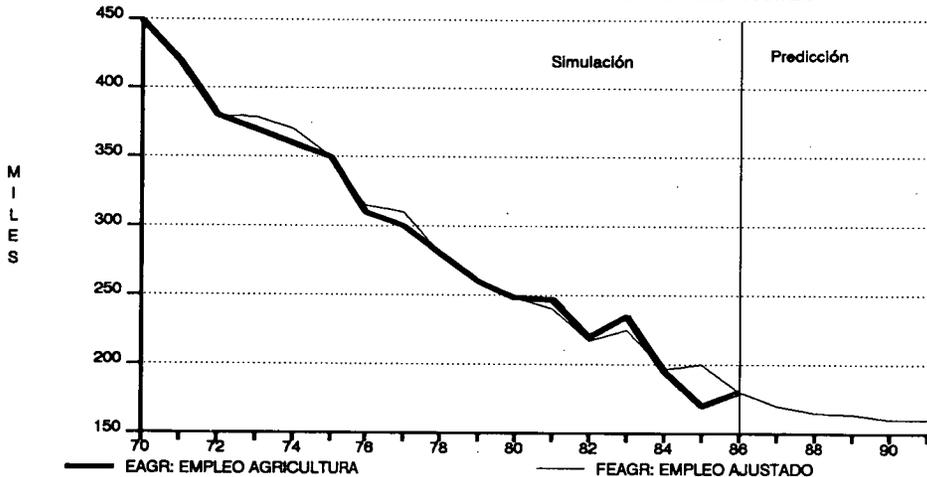
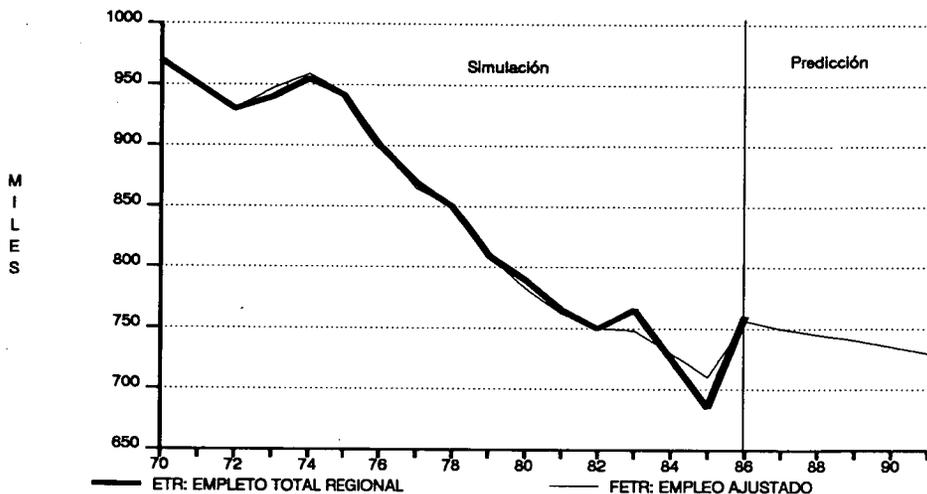
Hemos de recordar que el aumento del empleo recogido para los servicios en el año 86, es debido, únicamente, a un cambio de criterio del INE.

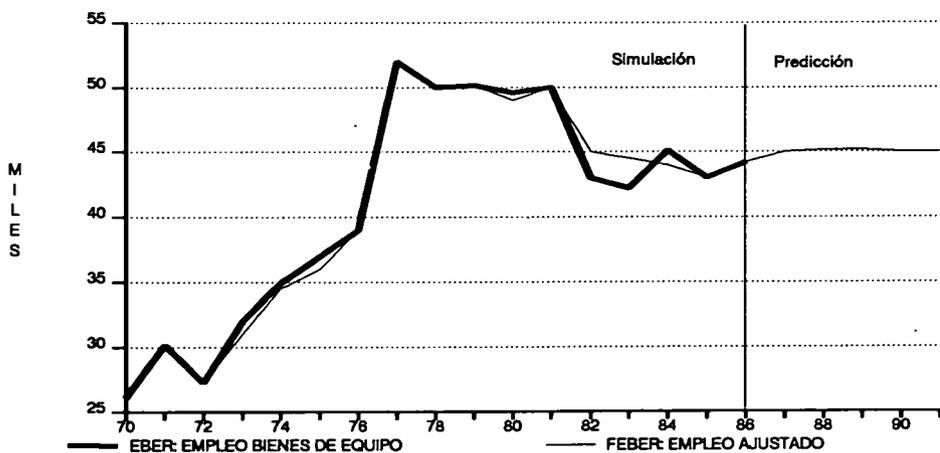
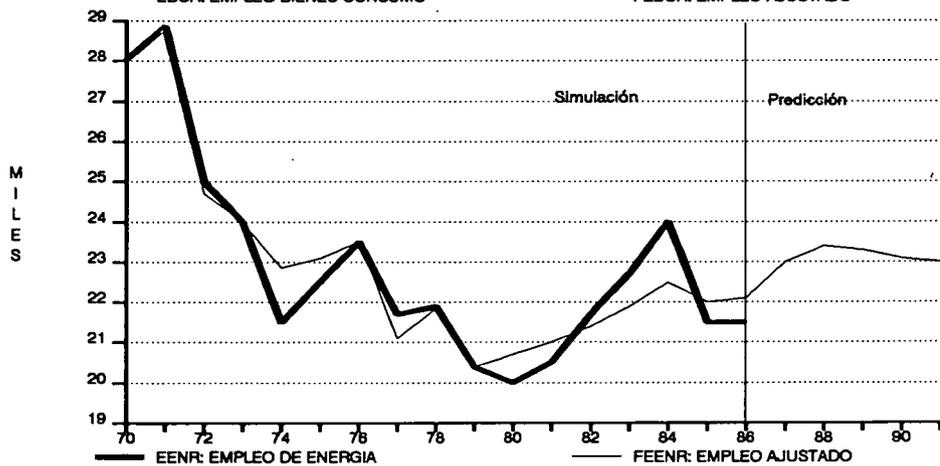
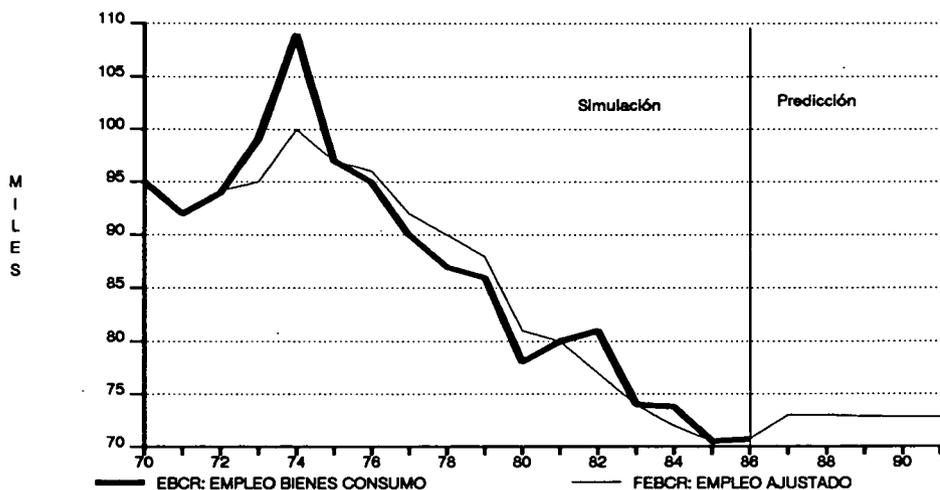
Las predicciones de los valores añadidos, medidas en pesetas constantes del 80, para las diferentes ramas de actividad, muestran una tendencia creciente. Destaca en este caso, el fuerte crecimiento del sector de la construcción.

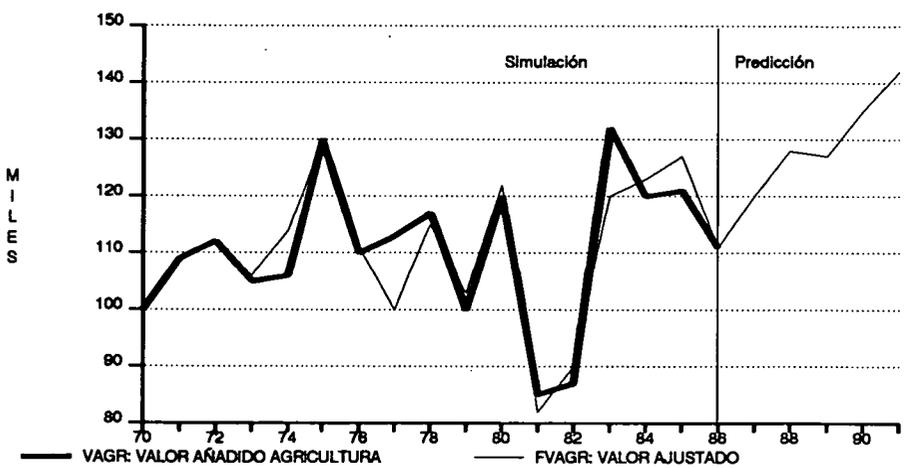
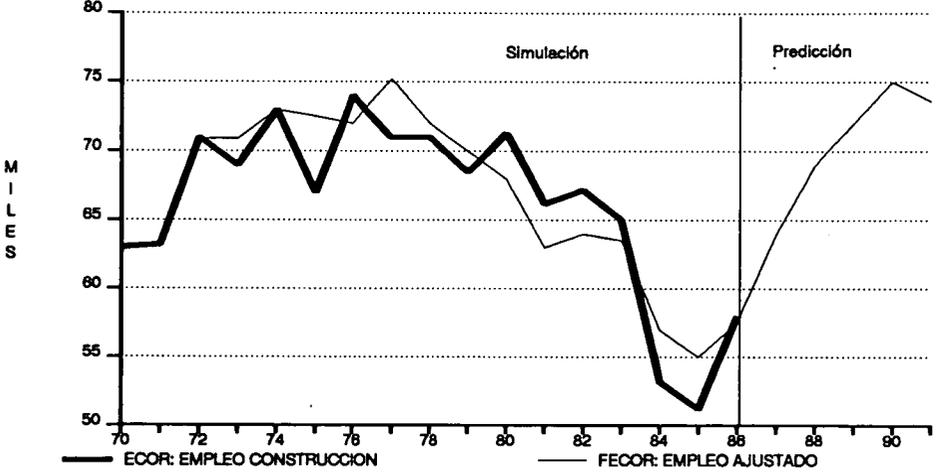
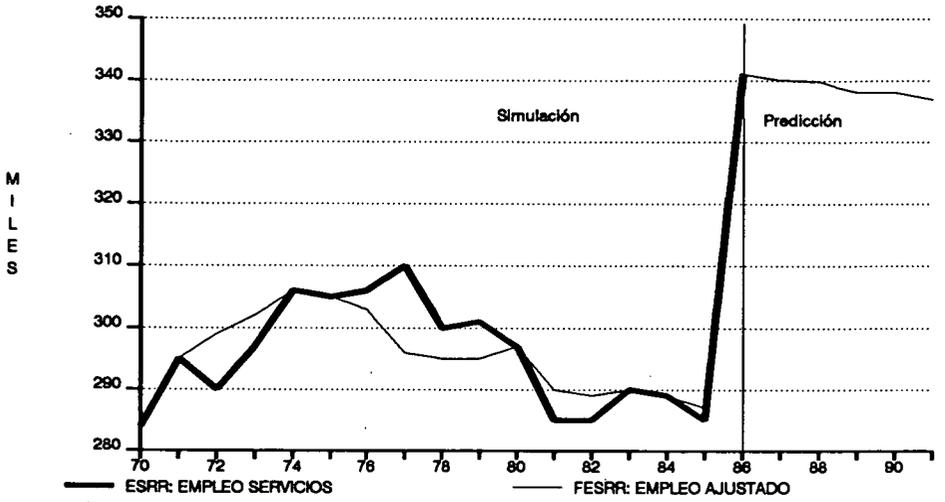
El único valor añadido que decrece a un ritmo apreciable corresponde el sector de bienes de equipo. Sin embargo, esta predicción no tiene en cuenta las recientes medidas tomadas por la Administración Autonómica para la minería de la comunidad castellano-leonesa.

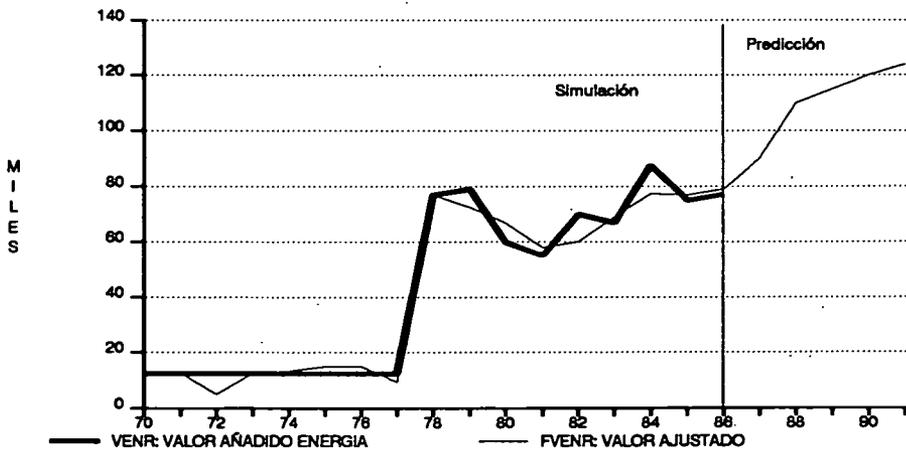
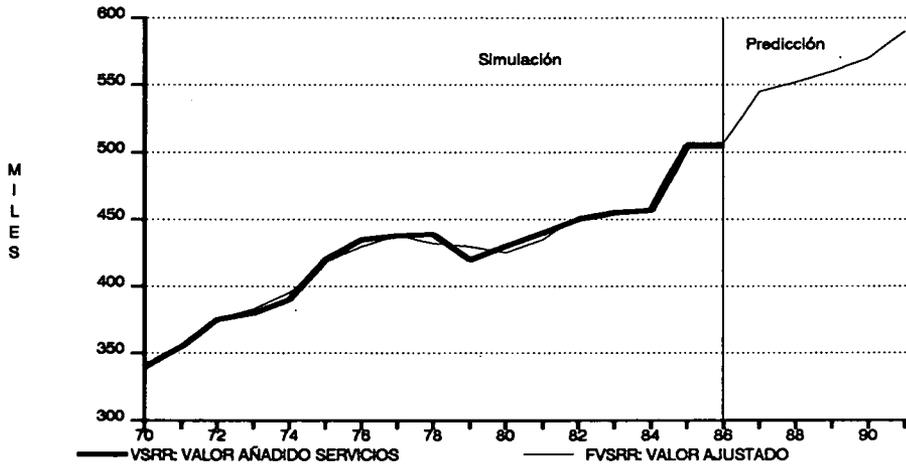
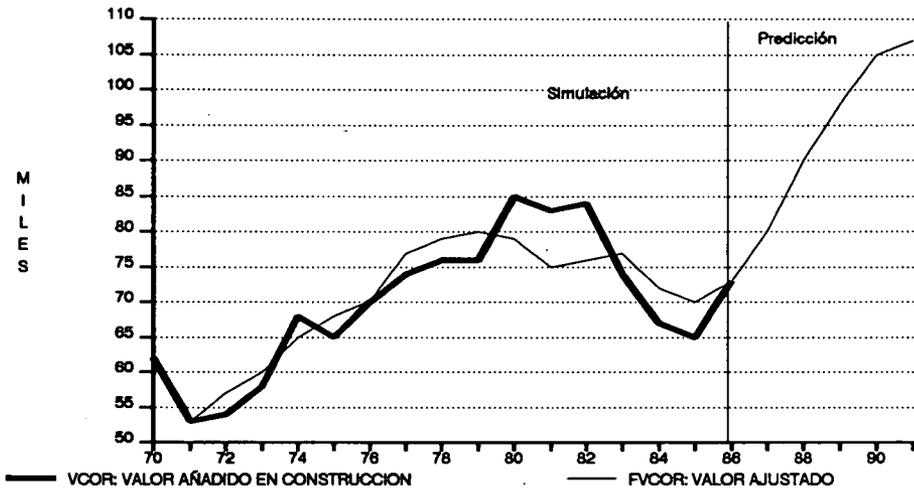
La renta regional disponible presenta también una tendencia creciente como consecuencia del crecimiento de los valores añadidos.

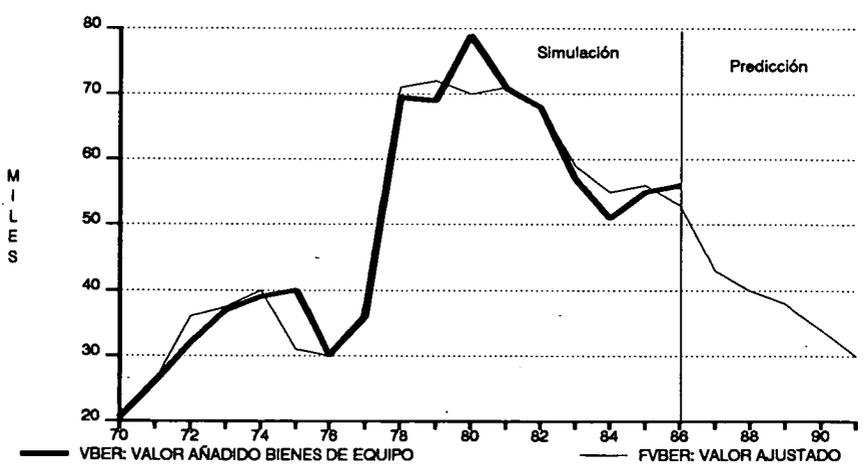
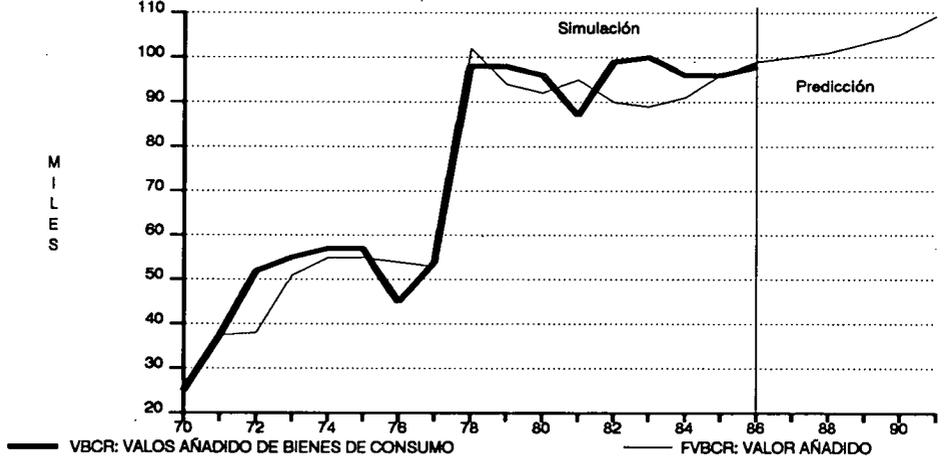
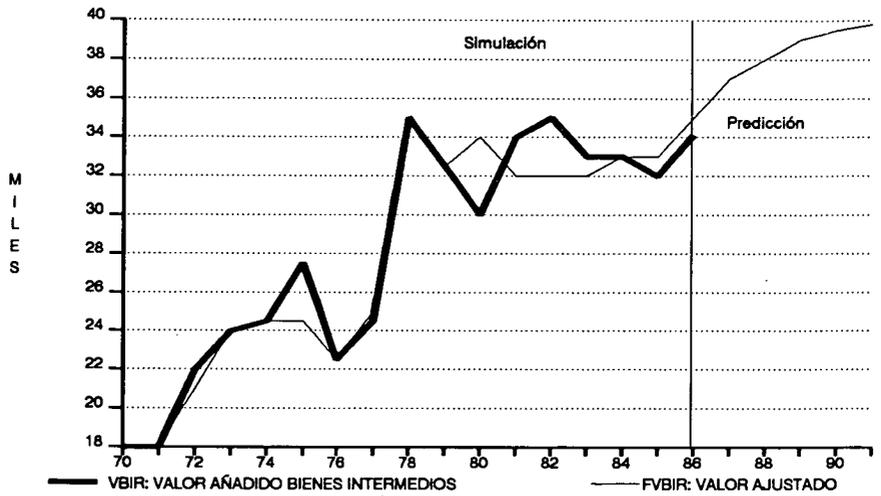
Se incluyen por último los gráficos que recogen la evolución de los valores observados, los de los valores simulados y las predicciones efectuadas, para las principales variables explicadas por el modelo.

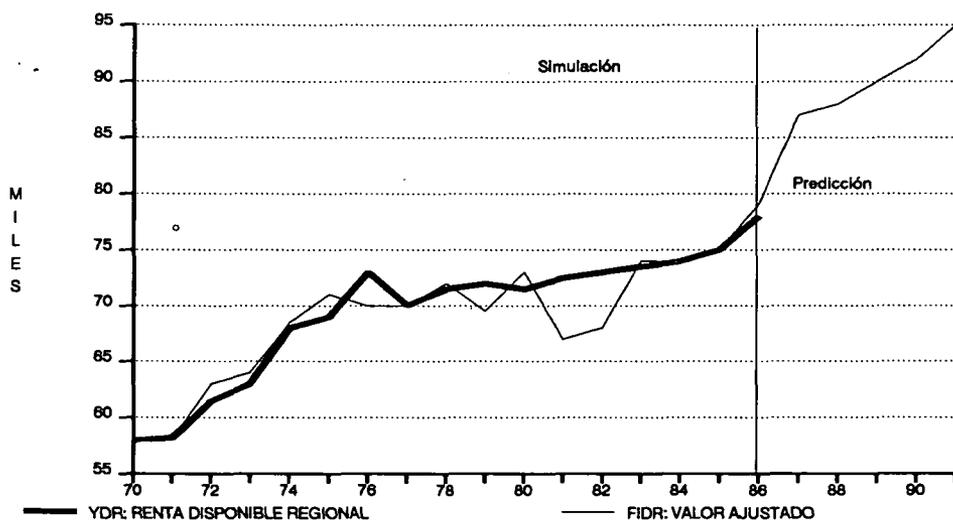












## Bibliografía consultada

- Adams, F.G., Brookman (1975) «On the Specification and Simulation of a Regional Econometric Model; A model of Mississippi», *Journal of Regional Science*, nº 57.
- Acaide, J. (1985) «Información estadística regional y local», *Situación*, nº 3.
- Barnes, G.T. y otros (1977) «Manpower Forecasting for small regions», *Growth and change*, pp. 15-23. Este procedimiento se encuentra también recogido en *El mercado trabajo. Modelos de previsión de oferta y demanda*. El caso español, Cavero Alvarez, J. Universidad de Valladolid 1984.
- Bell, F.W. (1967) «An Econometric Forecasting Model for a Region», *Journal of Regional Science*, nº 7.
- Courbis, R. (1986), «L'expérience française de nodélisation regional et multiregionale», *Coloquio sobre modelización regional*. Málaga, octubre 1986.
- Duobinis, S.F. (1979), «An Econometric Model of the Chicago Standars Metropolitan Statistical Area», *Journal of Regional Science*, nº 21.
- Dutta, M. and Su (1969) «An Econometric Model of Puerto Rico», *Review of Economic Studies*, Vol. XXXVI
- Glickman, N.J. (1976) «An Econometric Forecasting Model for the Philadelphia Region», *Journal of Regional Science*, nº 11.
- Hall, O.P. (1974) «Building small Region Econometric Models: Extension of Glickman's Structure to Los Angeles», *Journal of Regional of Science*, nº 14.
- Martinez, A. (1985) «La demanda de información estadística regional en España», *Situación*, nº 3.
- Pulido, A. (1983), *Modelos Económicos*, Ed. Pirámide.
- Rodriguez Feijóo, S. (1977) «Métodos de Análisis Regionales: Modelos Económicos Regionales», *II Jornadas de Modelos Económicos*, Barcelona, junio 1987.