

# COMBINACIÓN DE PERSPECTIVAS DE INFLACIÓN EN ESPAÑA BASADA EN MEDIDAS DE INFORMACIÓN

*Blanca Moreno Cuartas*

*Ana Jesús López Menéndez*

**Departamento de Economía Aplicada**

**Universidad de Oviedo**

## RESUMEN

La diversidad de organismos dedicados a los análisis de prospectiva hace posible disponer de un amplio abanico de predicciones para una misma variable económica, sugiriendo la conveniencia de llevar a cabo un resumen de las mismas mediante algún mecanismo de combinación.

Generalmente, la predicción combinada suele obtenerse a partir de un vector de ponderaciones que calibra las predicciones individuales según su capacidad predictiva, procedimiento que exige disponer de información sobre las predicciones individuales, resultando más fiable a medida que aumenta el volumen de información disponible.

Dado que en algunas ocasiones no es posible acceder a la información requerida, la alternativa más habitualmente empleada consiste en efectuar la combinación a partir de la media aritmética de las predicciones individuales.

En este trabajo desarrollamos un nuevo procedimiento de combinación de predicciones a partir de las herramientas proporcionadas por la teoría de la información, proponiendo la aplicación del principio de maximización de entropía a la medida de Incertidumbre de Shannon (1948) y la medida de Incertidumbre cuadrática (Pérez, 1985).

Como aplicación empírica analizamos la evidencia disponible sobre perspectivas de inflación en España elaboradas por distintos organismos en diferentes estadios de predicción. En este contexto estudiamos la variabilidad de las predicciones individuales según el estadio de predicción y obtenemos combinaciones basadas en la aplicación del principio de máxima entropía a las predicciones individuales, estudiando si los resultados son significativamente distintos de los asociados al empleo de la media aritmética.

## 1. INTRODUCCIÓN

El indudable interés que tiene para los distintos agentes económicos disponer de predicciones sobre la inflación ha impulsado a numerosos organismos e instituciones a elaborar periódicamente informes de prospectiva cuya difusión ha mejorado sustancialmente gracias al desarrollo de las Tecnologías de la Información y la Comunicación (TIC).

Así, en la actualidad existe un amplio abanico de organismos tanto públicos como privados que elaboran regularmente predicciones cuantitativas sobre el crecimiento de los precios de consumo en España. A modo de ejemplo, la tabla 1 recoge algunas de las principales instituciones nacionales e internacionales que realizan estos informes de prospectiva.

**Tabla 1:** Instituciones que realizan predicciones sobre inflación en España.

	<b>NACIONALES</b>	<b>INTERNACIONALES</b>
<b>Organismos oficiales</b>	Ministerio de Economía y Hacienda, Banco de España, Instituto de Crédito Oficial (ICO)	Comisión Europea (CE), Organización para la Cooperación y el Desarrollo Económico (OECD), Fondo Monetario Internacional (FMI)
<b>Instituciones financieras</b>	La Caixa, Banesto, Grupo Santander, Caja Madrid, InfoAnalistas (AFI), BBVA	ING Financial Markets, UBS Warburg, JP Morgan, Morgan-Stanley, Goldman-Sachs, Intermoney, Merrill Lynch
<b>Centros vinculados a universidades y organizaciones</b>	Centro de Predicción Económica (CEPREDE), Instituto L.R. Klein, Instituto Complutense de Análisis Económico, Confederación Española de Organizaciones Empresariales (CEOE), Instituto de Estudios Económicos (IEE)	Naciones Unidas-Proyecto Link
<b>Paneles de predicción</b>	Fundación Cajas de Ahorros Confederadas para la Investigación Económica y Social (FUNCAS)	The Economist, Consensus Economics

Teniendo en cuenta que los distintos organismos difieren en lo que se refiere a la información disponible, las variables consideradas y las técnicas de predicción empleadas, existirá heterogeneidad entre los resultados obtenidos, pudiendo ser conveniente efectuar una

combinación de los mismos que conduzca a una predicción de síntesis. De hecho, la idea de la combinación asume que cada predicción es capaz de capturar diferentes aspectos de la información, y que por tanto una combinación de todas ellas contribuirá a aumentar la precisión y mejorar el aprovechamiento de la información disponible.

Los trabajos pioneros en esta línea corresponden a Bates y Granger (1969), quienes proponen técnicas para obtener una predicción de síntesis a partir de combinaciones lineales de predicciones individuales, cuyos pesos se obtienen a partir de las correspondientes varianzas de los errores de predicción. Granger y Ramanathan (1984) ponen de manifiesto que los métodos convencionales de combinación pueden ser estudiados desde la óptica de la regresión, donde la predicción combinada es la variable dependiente y las predicciones individuales efectuadas por distintos métodos y/o individuos son las variables explicativas.

El cálculo de las ponderaciones utilizadas en la combinación exige disponer de información suficiente para calibrar a cada predictor según sus predicciones pasadas, aumentando la fiabilidad del procedimiento con el volumen de información disponible. Sin embargo, dado que en ocasiones no es posible acceder a la información requerida, la alternativa más habitual consiste en efectuar la combinación a partir de la media aritmética de las predicciones individuales.

Teniendo en cuenta que la teoría de la información estadística proporciona un marco de análisis muy adecuado para el tratamiento de problemas en diversos ámbitos económicos como la desigualdad de la renta, la concentración industrial o la dependencia estadística, en este artículo estudiamos las posibilidades de las medidas de incertidumbre para la combinación de predicciones en un contexto de insuficiencia de información. Más concretamente, empleamos el principio de Maximización de Entropía para la medida de incertidumbre de Shannon (1948) y la medida cuadrática [Pérez (1985)] y utilizamos la evidencia empírica disponible para comparar las ponderaciones estimadas mediante este procedimiento con los pesos habitualmente empleados (uniformes en el caso de la media aritmética).

Como consecuencia de estas consideraciones el resto del trabajo se estructura en cuatro secciones adicionales. En la segunda se realiza un breve estudio de algunas de las medidas de entropía y divergencia que proporciona la teoría de la información y de los principios y requisitos que conducen a maximizar el valor de las primeras y a minimizar el de las segundas. En la sección tercera se describen brevemente los métodos de combinación tradicionales para posteriormente detallar la metodología propuesta en el trabajo. La cuarta sección contiene una aplicación de esta metodología a las predicciones que numerosos organismos facilitan sobre la inflación en España. En este contexto analizamos la variabilidad de las predicciones individuales según el estadio de predicción y obtenemos combinaciones basadas en la aplicación del principio de máxima entropía a las predicciones individuales, estudiando si los resultados son significativamente distintos de los asociados al empleo de la media aritmética. Finalmente, las conclusiones se resumen en la quinta y última sección.

## 2. MEDIDAS DE INFORMACIÓN Y PRINCIPIOS DE OPTIMIZACIÓN

Las medidas de entropía cuantifican la cantidad de incertidumbre asociada a un experimento aleatorio. En concreto, dada una variable aleatoria  $X$  con valores  $x_i$  y distribución de probabilidad  $P = (p_1, \dots, p_n)$  tal que  $p_i \geq 0$  ( $i=1, \dots, n$ ) y  $\sum_{i=1}^n p_i = 1$ , la medida de entropía de

Shannon (1948) se define como:

$$H_s(X) = H_s(p_1, \dots, p_n) = -\sum_{i=1}^n p_i \log p_i \quad (1)$$

A partir de esta expresión han aparecido en la literatura diversas medidas de entropía generalizadas entre las que destacan por su interés las medidas de orden  $\beta$  introducidas por Havrda y Charvat (1969), definidas como:

$$H_{HC}^\beta(X) = \frac{1}{1-\beta} \left( \sum_{i=1}^n p_i^\beta - 1 \right), \beta > 0, \beta \neq 1 \quad (2)$$

En ocasiones el coeficiente  $(1-\beta)^{-1}$  se sustituye por  $(2^{1-\beta} - 1)^{-1}$  y entonces la medida de orden  $\beta$  de Havrda y Charvat (1969) queda definida como:

$$H_{HC}^{\beta}(X) = \frac{1}{2^{1-\beta} - 1} \left( \sum_{i=1}^n p_i^{\beta} - 1 \right) \quad (3)$$

Un caso particular de esta familia, cuando  $\beta=2$ , da lugar a las medidas cuadráticas de incertidumbre cuyas propiedades -estudiadas por Pérez (1985)- las hacen adecuadas para numerosos análisis económicos, y cuya expresión es:

$$H^2(X) = H^2(p_1, \dots, p_n) = 2 \left( 1 - \sum_{i=1}^n p_i^2 \right) \quad (4)$$

La entropía de una variable aleatoria discreta  $X$  alcanza su valor máximo cuando todos estos valores presentan la misma probabilidad (distribución uniforme), lo que justifica de algún modo el “principio de razón insuficiente” de Laplace, para el caso de no disponer de ninguna información sobre los resultados. Sin embargo, en ocasiones el desconocimiento sobre la distribución de probabilidad de la variable  $X$  no es absoluto sino que se dispone de alguna información parcial sobre la distribución, basada en supuestos sobre determinados valores, momentos o características de la distribución que pueden ser formulados como restricciones de igualdad.

En ese caso es posible ajustar una distribución  $P$  mediante la aplicación del principio de entropía máxima teniendo en cuenta toda, y solamente la información disponible sobre la distribución. Así, si sobre la variable aleatoria  $X$  se conocen ciertos valores  $a_r$  ( $r=1, \dots, m$ ) asociados con ciertas funciones  $g_r(X)$  de los valores de la variable, aunque no su distribución, el problema consiste en llevar a cabo su determinación mediante una distribución no negativa que, cumpliendo las condiciones conocidas ( $p_i \geq 0$  para  $i=1, \dots, n$  y  $\sum_{i=1}^n p_i = 1$ ), haga máxima la entropía. Formalmente, el problema consiste en:

Maximizar la medida de entropía  $H(X) = H(p_1, \dots, p_n)$

$$\text{sujeta a las restricciones: } \left. \begin{array}{l} \sum_{i=1}^n p_i = 1 \text{ con } p_i \geq 0 \quad i=1, \dots, n \\ \sum_{i=1}^n p_i g_r(x_i) = a_r \quad r=1, \dots, m \end{array} \right\} \quad (P.1)$$

Mediante la resolución del problema de optimización obtenemos la distribución de probabilidad  $\hat{P} = \{\hat{p}_i\}$ . Conviene tener presente que la distribución de máxima entropía no tiene una solución cerrada, siendo necesario emplear técnicas de optimización numérica para calcular las probabilidades.

Además de las medidas de entropía, cuando a la variable aleatoria  $X$  que toma valores  $x_i$  ( $i=1, \dots, n$ ) con un sistema de probabilidades  $P=\{p_i\}$  se le puede asignar otro sistema de probabilidad  $Q=\{q_i\}$ , la teoría de la información proporciona un conjunto de medidas de divergencia. Así, se define la divergencia dirigida de Kullback-Leibler (1951) entre las dos distribuciones (o “información para discriminar a favor de  $P$  contra  $Q$ ”) como el valor de la expresión:

$$D_{KL}(P, Q) = D_{KL}(p_1, \dots, p_n; q_1, \dots, q_n) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i} \quad (5)$$

La idea que subyace en una medida de divergencia es la de cuantificar la cantidad de información proporcionada por los datos para discriminar a favor de una distribución y en contra de otra; es decir, medir el grado de discrepancia entre dos poblaciones caracterizadas por sus correspondientes distribuciones de probabilidad, lo que permite interpretar la divergencia dirigida como “una distancia” entre dos distribuciones.

En este caso, si se supone que además de una información previa de la distribución  $P$  (o una distribución a priori  $Q$ ) existe cierta información parcial sobre tal distribución en estudio, la idea de la optimización consiste en obtener la distribución  $P$  que, cumpliendo las restricciones, se separa menos (es menos divergente) de la estimación disponible mediante la aplicación del principio de divergencia mínima. Formalmente, el problema consiste en:

$$\text{Minimizar la medida de divergencia: } D(P, Q) = D(p_1, \dots, p_n; q_1, \dots, q_n)$$

sujeta a las restricciones:

$$\left. \begin{aligned} \sum_{i=1}^n p_i &= 1 \text{ con } p_i \geq 0 \quad i=1, \dots, n \\ \sum_{i=1}^n p_i g_r(x_i) &= a_r \quad r=1, \dots, m \end{aligned} \right\} \quad (\text{P.2})$$

La resolución del problema de optimización nos permite obtener la distribución de probabilidad  $\hat{P} = \{\hat{p}_i\}$ .

Las similitudes entre los dos procedimientos de optimización, maximización de entropía y minimización de divergencia, hacen que, bajo determinadas condiciones, las probabilidades de P obtenidas minimizando la divergencia de esa distribución con respecto a otra Q, sean las mismas que las obtenidas al maximizar la entropía. Este es el caso cuando la distribución Q es la uniforme.

### 3. COMBINACIÓN DE PREDICCIONES: UNA PROPUESTA BASADA EN MEDIDAS DE INFORMACIÓN

En general designaremos por Y a la variable para la cual, en el momento actual t, queremos anticipar su comportamiento futuro en un horizonte temporal h de amplitud T ( $h=1, \dots, T$ ).

Denotaremos la predicción efectuada por  $\hat{y}_{t+h,t}$  y los valores verdaderos (desconocidos) por  $y_{t+h}$ .

Dependiendo de la base informativa considerada en el momento de predicción t y del procedimiento seguido para procesarla, es posible obtener distintas predicciones para  $y_{t+h}$ , que denotaremos con el vector  $\hat{Y}_{t+h,t} = (\hat{y}_{t+h,t}^1, \hat{y}_{t+h,t}^2, \dots, \hat{y}_{t+h,t}^F)$ , donde cada elemento  $\hat{y}_{t+h,t}^i$  es la predicción efectuada a través de cada predictor i ( $i=1, \dots, F$ ).

Los métodos de combinación asumen la idea de que a partir de las predicciones individuales se puede obtener una combinación de predicciones, que llamaremos  $y_{t+h,t}^c$ , que combina toda la información contenida en las predicciones individuales. Es decir,  $y_{t+h,t}^c = f(\hat{Y}_{t+h,t}(\alpha))$ , donde  $\alpha$  es un vector de ponderaciones de las predicciones individuales. Dependiendo de cómo se formalice la relación causal, y se obtengan las estimaciones de  $\alpha$ , surgen una serie de métodos que tienen como objetivo generar una predicción combinada:  $\hat{y}_{t+h,t}^c = f(\hat{Y}_{t+h,t}(\hat{\alpha}))$ .

Los pioneros en el estudio teórico de la combinación de predicciones son Bates y Granger (1969), quienes proponen técnicas para obtener una predicción de síntesis a partir de combinaciones lineales de predicciones individuales, donde los pesos se obtienen a partir de la varianza del error de cada predicción individual. Posteriormente este planteamiento es ampliado para el caso de más predicciones por Newbold y Granger (1974).

En este caso la predicción combinada se expresa como,  $y_{t+h,t}^c = \hat{Y}_{t+h,t} \alpha$  donde  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_F)'$ ,  $1$  es un vector  $(F \times 1)$  de unos  $1' a = 1$  y  $0 \leq a_i \leq 1$  para todo  $i$ . La varianza del error de la predicción combinada es minimizada considerando:

$$\hat{\alpha} = \frac{(\sum^{-1} 1)}{(1' \sum^{-1} 1)} \text{ donde } \sum = E(e_{t+h} e_{t+h}') \text{ y } e_{t+h} = y_{t+h} 1' - \hat{Y}_{t+h,t} \quad (6)$$

Granger y Ramanathan (1984) ponen de manifiesto que las ponderaciones obtenidas con métodos convencionales de combinación tienen una interpretación como vector de coeficientes de la proyección lineal de  $y_{t+h}$  a partir de las predicciones de los  $F$  métodos:

$$y_{t+h} = \alpha_1 \hat{y}_{t+h,t}^1 + \dots + \alpha_F \hat{y}_{t+h,t}^F + u_{t+h} = \hat{Y}_{t+h,t} \alpha + u_{t+h} \quad (7)$$

Al desconocer el verdadero valor de  $y_{t+h}$  los pesos se obtienen a partir de las  $n$  predicciones pasadas. Si  $Y = (y_1, y_2, \dots, y_n)'$  recoge la evolución pasada de la serie y  $\hat{Y} = (\hat{Y}^1, \hat{Y}^2, \dots, \hat{Y}^F)$  es la matriz  $n \times F$  de predicciones pasadas que se realizaron sobre  $Y$  mediante las diferentes técnicas y/o individuos, con  $\hat{Y}^i = (\hat{y}_{1,0}^i, \hat{y}_{2,1}^i, \dots, \hat{y}_{n, n-1}^i)'$ , el objetivo será estimar  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_F)'$ , en la regresión  $Y = \hat{Y} \alpha + u$ . De este modo se estima  $Y$  como proyección lineal de  $\hat{Y}$ , para emplearla después en la predicción, siendo  $u$  un vector  $n \times 1$  de errores.

El cálculo eficiente de las ponderaciones  $\alpha$  exige como requisito mínimo que el número de observaciones ( $n$ ) sea superior al número de predictores individuales. Sin embargo en ocasiones no es posible disponer de suficiente información para su estimación, bien porque la situación a predecir sea nueva (o se trate de una situación con poca historia) o bien porque gradualmente se vayan incorporando nuevos predictores y no se disponga de información suficiente sobre ellos que permita calibrar su calidad, siendo por tanto  $\hat{Y}$  una matriz  $n \times F$  ( $F > n$ ) no invertible. En este

caso, la práctica más habitual es obtener el vector de ponderaciones a partir de la media aritmética de las predicciones individuales, que no aprovecha la información disponible sobre los predictores.

Una solución consiste en desarrollar un procedimiento de inversión no lineal, que es una generalización de las reglas de estimación más tradicionales, y que requiere la aplicación de las herramientas que proporciona la teoría de la información estadística en un contexto en el que la información observada está disponible en forma de medias o agregados donde, como resultado, se debe usar una distribución de probabilidad  $P$ , para representar la información parcial sobre las observaciones individuales.

El problema al que nos enfrentamos en la combinación de predicciones se puede formular, siguiendo la notación general de Golan, Judge y Miller (1996), de la siguiente manera: dada la información disponible sobre las predicciones individuales  $\hat{Y}$  y las realizaciones  $Y$ , nuestro interés se centra en el vector desconocido e inobservable  $\alpha$  y, puesto que no podemos cuantificar directamente  $\alpha$ , debemos de emplear medidas indirectas a partir de la información observada.

Cuando la información de  $Y$  es especificada sin error y en términos agregados nos referiremos al problema inverso como “inverso puro”,  $Y = \hat{Y}\alpha$ , mientras que si consideramos además un término de error,  $Y = \hat{Y}\alpha + u$  hablaremos de problemas de inversión “general”.

En un contexto como el descrito, supongamos que tenemos una versión lineal, discreta y finita de un problema inverso puro:

$$Y = \hat{Y}\alpha = \hat{Y}P \quad (8)$$

donde se desea determinar las frecuencias desconocidas e inobservables  $P = (p_1, \dots, p_F)'$ , que representan el proceso generador de datos. Entonces dentro de los posibles conjuntos de probabilidades que cumplen  $\sum_{i=1}^F p_i = 1$  con  $p_i \geq 0$ , debemos de escoger o asignar un único vector, objetivo que sin embargo no resulta alcanzable con la información disponible.

Para resolver este problema Jaynes (1957) propone el empleo de las medidas de entropía para escoger la distribución de probabilidad desconocida P. Bajo el principio de Máxima entropía elegimos la distribución P para la que el conjunto de predicciones individuales disponibles sean suficientes para determinar una única distribución.

Así, escogemos P que maximiza la expresión:  $H(\hat{Y}) = H(p_1, \dots, p_F)$

sujeto a las restricciones de consistencia con la información y requisitos de normalización y aditividad de P:

$$\left. \begin{array}{l} P'1 = 1 \text{ con } p_i \geq 0 \text{ para } i=1, \dots, F \\ \hat{Y}P = Y \end{array} \right\} \quad (P.3)$$

Si el problema inverso puro es definido de nuevo considerando una variable de error  $Y = \hat{Y}\alpha + u$ , el problema podría entonces ser reformulado como:  $Y = \hat{Y}P + uW$ , siendo W el vector  $1 \times n$  de ponderaciones del error de cada año que aparece con el objetivo de corregir el sesgo en la predicción y donde  $p_i$  y  $w_t$  son positivas para  $i=1, \dots, F$  y  $t=1, \dots, n$  respectivamente. El programa de optimización será:

$$\left. \begin{array}{l} \text{Maximizar } H(P, W) = H(P) + H(W) \\ \\ \text{Sujeto a: } \left. \begin{array}{l} Y = \hat{Y}P + uW \\ P'1 = 1 \text{ donde } 1 \text{ es un vector } (F \times 1) \text{ de unos} \\ W'1_n = 1 \text{ donde } 1_n \text{ es un vector } (n \times 1) \text{ de unos} \end{array} \right\} \end{array} \right\} \quad (P.4)$$

La resolución del problema de optimización nos permitirá conocer  $\hat{P}$  y  $\hat{W}$ .

Dado que el conjunto de medidas de entropía es muy numeroso, los resultados de las estimaciones dependerán de la medida considerada en cada caso.

En este sentido, si bien la expresión más habitualmente empleada es la medida de entropía de Shannon (1948), las propiedades de la entropía cuadrática  $H^2(X)$ , estudiadas por Pérez (1985), han mostrado la idoneidad de esta expresión en diversas aplicaciones económicas<sup>1</sup>; de ahí que nos planteemos su empleo en el problema de la combinación de predicciones.

---

<sup>1</sup> La medida cuadrática resulta adecuada conceptualmente para numerosos análisis económicos, entre los que destacan los relativos a la desigualdad [Pérez (1985); Pérez, Caso y Gil (1986); López y Pérez

La medida de incertidumbre cuadrática es una función cóncava de  $P$ , lo que garantiza que la solución que se obtiene en el problema de maximización sea un máximo global. Sin embargo, al contrario de lo que ocurre con la medida de Shannon donde cada  $\hat{p}_i$  es una función exponencial, no se garantiza que las probabilidades estimadas sean siempre positivas, de ahí que sea necesario imponer la restricción de no negatividad en  $P$  en el programa de maximización ( $p_i \geq 0$  para todo  $i=1, \dots, F$ ).

La resolución del problema de optimización con restricciones no lineales resulta complicada. No obstante, el empleo de paquetes de software específicos que permitan una optimización no lineal facilita la obtención de la solución numérica.

#### **4. COMBINACIÓN DE PREDICCIONES SOBRE LA INFLACIÓN EN ESPAÑA**

En esta sección presentamos un análisis empírico de las predicciones que distintas instituciones realizan sobre el crecimiento de los precios de consumo en España en el periodo 1994-2005, ilustrando a partir de ellas la metodología propuesta para la combinación de predicciones.

Las perspectivas de inflación nacional (tasas de variación anual de los precios de consumo) elaboradas por distintos organismos aparecen recopiladas en los *Informes Semestrales de Perspectivas Económicas y Empresariales* publicados por CEPREDE y el Instituto L.R. Klein. En concreto, en cada informe semestral se publican las predicciones que las instituciones elaboran para el año en curso y el siguiente así como la media de todas ellas. De este modo, para cada período considerado dispondremos de un total de cuatro predicciones, dos de ellas elaboradas el año anterior (Junio y Diciembre), y otras dos publicadas en los informes semestrales del año en curso (Junio y Diciembre).

Por lo que se refiere a las realizaciones de la variable investigada, adoptamos como referencia las cifras oficiales de inflación en España, que se corresponden con las tasas de variación interanual del Índice de Precios al Consumo (IPC) publicadas por el Instituto Nacional de Estadística (INE).

---

(1991)], la concentración industrial [Río y Pérez (1987)] y la asociación de caracteres [Alvargonzález y Pérez (1987)].

Como paso previo al cálculo de las ponderaciones de los organismos, analizamos la variabilidad que existe entre las predicciones individuales, y entre éstas y la predicción combinada, así como la volatilidad asociada a las predicciones realizadas en los diferentes estadios de la información disponible. Este análisis previo puede aportarnos valiosa información para la segunda parte del estudio en la que ya se estiman las ponderaciones mediante la maximización de medidas de entropía, puesto que de corroborarse una disminución en la dispersión entre organismos, estas predicciones deberían ser más homogéneas en el último estadio de información que en el primero. La aplicación incluye también un análisis de coherencia entre los pesos obtenidos con las medidas consideradas (la entropía de Shannon y la incertidumbre cuadrática) y la evaluación de la capacidad predictiva de la técnica de combinación propuesta.

#### **4.1. La variabilidad de las predicciones según el estadio de predicción**

De acuerdo con Gallo, Granger y Jeon (1999) a medida que se incorpora información más reciente o avanzamos en el estadio de predicción es previsible una disminución tanto en los sesgos de las predicciones individuales (y por tanto de la combinación) como en la dispersión de los diferentes organismos. En este apartado intentamos mostrar evidencia empírica sobre la convergencia en las predicciones individuales sobre la inflación, y analizar en qué medida se tiende a una predicción media. Denotaremos por  $\hat{y}_{t+h,t}^i$  la predicción realizada en t para un instante t+h, por un organismo  $i=1, \dots, F$ . Dado que dicha predicción puede revisarse a medida que se dispone de nueva información, entonces para un instante t+h podremos disponer de distintas predicciones de acuerdo con los estadios de la información s.

La información disponible para un año t consiste en cuatro predicciones que se corresponden con otros tantos estadios de información s, que denotaremos por  $s= t-IV, t-III, t-II$  y  $t-I$  y que se corresponden respectivamente con las predicciones realizadas en Junio de t-1, Diciembre de t-1, Junio de t y Diciembre de t. Por tanto, para cada año t cada institución i dispondrá de predicciones  $\hat{Y}_{t+h,t}^i = (\hat{y}_{t,t-IV}^i, \hat{y}_{t,t-III}^i, \hat{y}_{t,t-II}^i, \hat{y}_{t,t-I}^i)$  correspondientes a los cuatro estadios y en cada uno de esos estadios (s) habrá un conjunto de predicciones realizadas por las F

instituciones  $\hat{Y}_{t,t+s} = (\hat{y}_{t,t+s}^1, \hat{y}_{t,t+s}^2, \dots, \hat{y}_{t,t+s}^F)'$ .

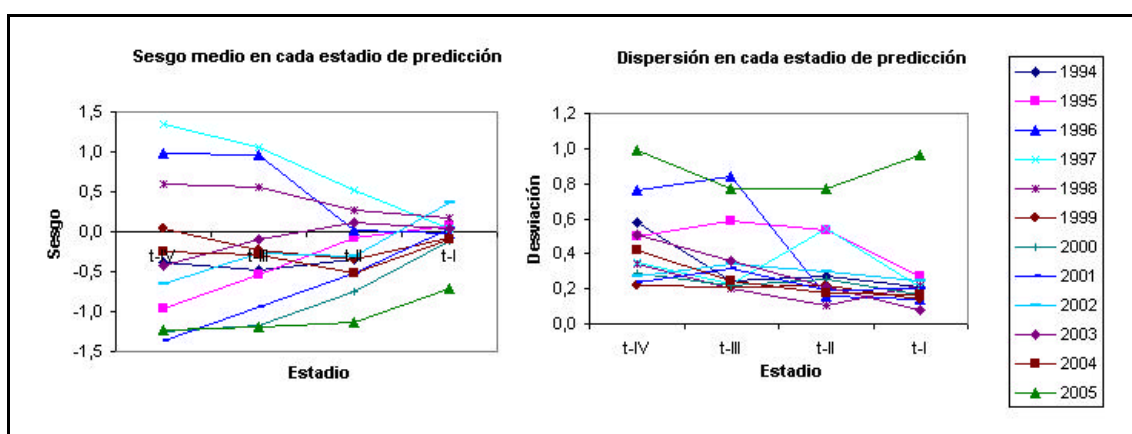
Además de las predicciones individuales, en cada informe se publica el promedio de todas ellas que denotamos por  $\hat{y}_{t,t+s}^c$ . Por lo tanto, si tenemos en cuenta la información publicada en su conjunto, para cada año  $t$  tendremos un panel formado por las predicciones de todas las instituciones y la predicción media en cada uno de los cuatro estadios  $(\hat{y}_{t,t+s}^1, \hat{y}_{t,t+s}^2, \dots, \hat{y}_{t,t+s}^F, \hat{y}_{t,t+s}^c)$  que es una matriz  $(F+1) \times 4$ .

Con la información proporcionada por cada panel es posible estudiar la revisión de las predicciones individuales y la convergencia de los distintos resultados a la predicción combinada, así como los sesgos observados al comparar las predicciones con las correspondientes realizaciones, esto es, las tasas interanuales del IPC publicadas por el INE.

Es de esperar que, a medida que se avanza en el proceso de predicción, los organismos tiendan a converger en sus predicciones, de una parte por el efecto a imitar al resto y de otra porque es previsible que al disponer de más información aumente la precisión.

Para comprobar esta idea que Gallo, Granger y Jeon (1999) modelizan teóricamente, disponemos de 12 paneles<sup>2</sup> correspondientes a los años del periodo 1994-2005. Para cada panel anual se ha calculado en cada estadio de la información el sesgo medio cometido por las instituciones y la dispersión entre instituciones. Una representación de los resultados obtenidos aparece recogida en el gráfico 1.

**Gráfico 1:** Sesgo medio y dispersión entre predicciones individuales sobre inflación.



**Fuente:** Elaboración propia a partir de datos CEPREDE, L.R. Klein e INE.

<sup>2</sup> No existe homogeneidad en las instituciones consideradas en cada panel, por lo que la base informativa para el periodo 1994-2005 incluye un número de instituciones comprendido entre 4 y 26.

Los resultados del análisis realizado muestran que, tal y como cabía esperar, el sesgo disminuye a medida que avanzamos en el estadio de predicción y también la dispersión entre las predicciones de los diferentes organismos se reduce cuando nos acercamos al instante para el que se predice, en la práctica total de los años analizados. Este comportamiento sugiere que el empleo de la media aritmética en la combinación resulta más adecuado en los últimos estadios de predicción que en los primeros.

#### 4.2. Combinación de predicciones con medidas de entropía

Los paneles considerados en los *Informes Semestrales de Perspectivas Económicas y Empresariales* van incorporando cada vez mayor número de instituciones (F) que aportan sus predicciones sobre la inflación en España. En la tabla 2 recogemos el número de predictores considerados en cada panel para el período investigado.

**Tabla 2:** Número de predictores para cada año y cada estadio de predicción

Año de predicción t	Estadios de predicción			
	$\hat{Y}_{t,Junio\ t-1}^F$	$\hat{Y}_{t,Diciembre\ t-1}^F$	$\hat{Y}_{t,Junio\ t}^F$	$\hat{Y}_{t,Diciembre\ t}^F$
<b>1994</b>	4	9	12	11
<b>1995</b>	8	11	13	11
<b>1996</b>	12	11	18	16
<b>1997</b>	15	16	17	17
<b>1998</b>	16	17	18	22
<b>1999</b>	17	22	23	22
<b>2000</b>	23	22	21	21
<b>2001</b>	21	21	20	20
<b>2002</b>	17	20	26	26
<b>2003</b>	21	25	25	24
<b>2004</b>	20	24	24	24
<b>2005</b>	19	23	24	23

Si deseamos obtener una predicción combinada mediante técnicas de regresión, con el objetivo de garantizar un proceso de estimación adecuado necesitaríamos disponer de suficiente información, esto es  $n > F$  para cada estadio de predicción. Sin embargo, el rápido incremento del número de organismos dedicados a la prospectiva hace que tengamos muchos agentes individuales y poca información pasada sobre los mismos. Una solución sería limitar el panel de predictores a un número reducido de instituciones con suficiente tradición en la elaboración de predicciones, si bien deberíamos plantearnos si se desaprovecha la información suministrada por

el resto de organismos, y si los de más tradición predictiva son también aquéllos que realizan predicciones de mayor calidad.

La alternativa más habitual consiste en efectuar la combinación a partir de la media aritmética sobre las predicciones individuales, lo que implica asumir que todos los individuos que efectúan la predicción tienen la misma importancia, supuesto que podría justificarse en el caso de no disponer de ninguna información sobre ellos que permita calibrarlos de modo desigual. El problema que trataremos a continuación es el de transformar los pesos a priori (que asociamos con los de la media aritmética) en otros a posteriori mediante la aplicación del principio de entropía máxima teniendo en cuenta la información disponible sobre cada predictor.

Así, en un contexto como el descrito de escasez de información, es necesario resolver el cálculo de ponderaciones como un “problema inverso”. Si suponemos que tenemos una versión lineal, discreta y finita de un problema inverso puro  $Y = \hat{Y}\alpha$  entonces, a partir de  $Y = \hat{Y}P$ , maximizamos la entropía de  $P$  para obtener indirectamente  $\alpha$ , tal y como ya explicamos en la sección anterior.

Además de las observaciones pasadas, podemos tener una información inicial o a priori sobre las ponderaciones de los organismos  $\alpha^0$ . En ese caso debemos de obtener  $P$  minimizando la divergencia dirigida  $D(P,Q)$ , donde  $Q$  es la distribución asociada a los pesos iniciales, consistente con toda esa información.

Si deseamos justificar el uso que a priori se hace de la media aritmética en la combinación, el principio de minimizar la divergencia nos permitirá obtener los pesos  $\alpha$  que, cumpliendo las restricciones, divergen menos de los asociados a la media aritmética, y nos dará “información para discriminar a favor de  $\alpha$  contra  $\alpha^0$ ”. En ese caso los pesos son los asociados a la media aritmética  $\alpha_i^0 = q_i = 1/F$ , y la distribución  $Q$  se corresponde con la uniforme  $Q=U$ , con lo que la divergencia que debemos minimizar es  $D(P,U)$ , lo cual equivale a maximizar la entropía  $H(P)$ .

Como consecuencia de estas consideraciones, nos planteamos varios programas de maximización con las medidas de entropía de Shannon  $H_s(P)$  y la incertidumbre cuadrática  $H^2(P)$ .

En este caso, trabajamos separadamente con los paneles de información de cada estadio, con el objetivo de comparar las ponderaciones estimadas en cada uno de ellos, que denotamos por  $\hat{P}_{t-IV}, \dots, \hat{P}_{t-1}$  y que corresponden respectivamente a  $\hat{\alpha}_{t-IV}, \dots, \hat{\alpha}_{t-1}$ . En concreto, nos centramos en el análisis de los estadios más extremos donde cabe esperar mayor divergencia en los pesos estimados y realizamos un análisis recursivo en el que para la estimación de ponderaciones de cada año, se considera toda la información anterior y así sucesivamente.

Para la resolución de los programas de optimización empleamos el lenguaje de programación GAMS (*General Algebraic Modeling System*)<sup>3</sup>, que permite la resolución de diversos problemas de optimización.

Con esta propuesta de trabajo, al abordar los problemas de maximización nos encontramos con que en la mayoría de los casos considerados no resulta posible obtener una solución numérica de las ponderaciones, problema que no es debido a los programas implementados sino a la información empírica recogida en los paneles. La justificación es que puede haber casos en los que todos los organismos cometan un sesgo en la misma dirección (por ejemplo positivo), con lo cual para que se cumpla la restricción  $\hat{Y}P = Y$  sería necesario que la ponderación de alguno de ellos fuera negativa, y de modo análogo si por el contrario todos los organismos predicen a la baja será necesario que al menos la ponderación de uno de ellos supere la unidad.

Puesto que este tipo de situaciones se produce de forma recurrente, nuestros programas de maximización pueden no tener solución cuando se considera un problema inverso puro, con pesos positivos y de suma unitaria. Como consecuencia, resulta aconsejable una segunda alternativa consistente en plantear un problema de inversión general, que incorpora en las restricciones un término de error que puede emplearse para corregir el sesgo global tal y como plantea Theil (1971). En este caso, el problema es de nuevo definido como  $Y = \hat{Y}\alpha + u$ , siendo  $u$  un vector de errores, donde cada  $u_t$  se supone que tiene características de variable aleatoria con  $2 \leq J \leq \infty$  posibles resultados.

---

<sup>3</sup> La versión del programa que hemos empleado para esta aplicación es GAMS 21.3. Una descripción detallada del programa y sus actualizaciones puede verse en <http://www.gams.com/>.

Así, se asume que cada  $u_t$  está acotado por un intervalo  $(v_{t1}, v_{tJ})$ , tal que  $\Pr(v_{t1} < u_t < v_{tJ})$  pueda hacerse tan pequeño como se quiera. Por ejemplo para  $J=2$ , el error puede describirse como:  $u_t = w_t v_{t1} + (1 - w_t) v_{tJ}$  donde cada  $w_t \in [0,1]$  es un vector de ponderaciones del error. Los  $J \geq 2$  pueden emplearse para asumir ciertas características de simetría y apuntamiento que se deseen y así, si asumimos que el error es simétrico y centrado en 0, entonces  $-v_{t1} = v_{tJ}$  para cada  $t$ .

Los  $t$  errores desconocidos pueden expresarse en forma matricial como:

$$u = VW = \begin{bmatrix} V_1' & & & \\ & V_2' & & \\ & & \ddots & \\ & & & V_n' \end{bmatrix} \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_n \end{bmatrix} \quad (9)$$

donde  $V$  es una matriz  $(n \times nJ)$  y  $W$  es un vector de pesos de dimensión  $nJ$ , que se restringen a que sumen la unidad y sean positivos. Entonces la ecuación de restricción será:

$Y = \hat{Y}\alpha + u = \hat{Y}P + VW$  y el problema de maximización vendrá dado por:

Maximizar  $H(P, W) = -P' \ln(P) - W' \ln(W)$

$$\text{Sujeto a: } \left. \begin{array}{l} Y = \hat{Y}P + VW \\ P'1 = 1 \\ W'(I_n \otimes 1_n) = 1 \end{array} \right\} \quad (P.5)$$

donde  $1$  es un vector  $(F \times 1)$  de unos,  $1_n$  es un vector  $(n \times 1)$  de unos y  $\otimes$  es el producto de Kroneker

Las ponderaciones de los organismos y de los errores para cada año se estiman a partir de las soluciones obtenidas en la maximización:  $\hat{\alpha} = \hat{P}$  y  $\hat{u} = V\hat{W}$ .

Con este planteamiento general desarrollamos varios programas en GAMS, considerando distintos vectores  $V$  para el error<sup>4</sup> y diferentes niveles de información disponible.

---

<sup>4</sup> En el primer estadio de predicción los individuos comenten mayores sesgos que en el último (donde disponen de más información), por lo que los vectores  $V$  considerados deber ser de menor amplitud en este estadio que en el primero, de acuerdo además con la evidencia empírica suministrada en el apartado anterior.

Considerando los vectores  $V$  de la tabla 3 y utilizando las medidas de entropía de Shannon y cuadrática, en la tabla 4 recogemos las ponderaciones estimadas para algunas instituciones en los estadios de predicción extremos ( $t$ -IV y  $t$ -I), incorporando de forma recursiva la información disponible para el período 2000-2004.

**Tabla 3:** Vectores  $V$

Estadios	$J=5$	$-v_{t1} = v_{t5}$
<b>t-IV</b>	(-1,25, -0.5, 0, 0.5, 1,25)	
<b>t-I</b>	(-0.35, -0.15, 0, 0.15, 0.35)	

**Tabla 4:** Ponderaciones estimadas para estadios extremos según medidas de entropía

Información disponible	$\hat{a}_{t-IV}$				$\hat{a}_{t-I}$			
	2000		2000-2004		2000		2000-2004	
	$H_s(P)$	$H^2(P)$	$H(P)$	$H^2(P)$	$H_s(P)$	$H^2(P)$	$H_s(P)$	$H^2(P)$
<b>ORGANISMOS</b>								
<b>BSCH</b>	0,076	0,088	0,066	0,071	0,100	0,100	0,094	0,094
<b>Comisión Europea</b>	0,059	0,067	0,060	0,056	0,100	0,100	0,094	0,094
<b>Ceprede</b>	0,164	0,149	0,287	0,244	0,100	0,100	0,101	0,102
<b>Consensus Economics</b>	0,127	0,128	0,080	0,100	0,100	0,100	0,097	0,098
<b>FMI</b>	0,098	0,108	0,073	0,086	0,100	0,100	0,126	0,126
<b>FUNCAS</b>	0,164	0,149	0,194	0,201	0,100	0,100	0,094	0,094
<b>Goldman-Sachs</b>	0,098	0,108	0,015	0,001	0,100	0,100	0,094	0,094
<b>Instituto de Crédito Oficial</b>	0,059	0,067	0,008	0,001	0,100	0,100	0,097	0,098
<b>La Caixa</b>	0,027	0,006	0,099	0,097	0,100	0,100	0,097	0,098
<b>The Economist</b>	0,127	0,128	0,119	0,143	0,100	0,100	0,105	0,105

Los resultados obtenidos permiten apreciar que en el estadio  $t$ I las ponderaciones son más homogéneas que en el estadio inicial  $t$ -IV, confirmando así nuestras expectativas. Como consecuencia, la evidencia parece sugerir que, si bien en el estadio final sería asumible la aplicación de la media aritmética para obtener la predicción combinada, no ocurre lo mismo para el estadio inicial.

Al comparar las ponderaciones en cada estadio, se observa que en  $t$ IV las ponderaciones obtenidas para 2000 varían más que en el estadio  $t$ -I cuando aumentamos la información que se considera en la estimación (de 2000 a 2004). Una explicación a este hecho (corroborada

mediante análisis de simulación)<sup>5</sup> es que las ponderaciones no varían cuando la información suministrada de un año a otro no resulta significativa.

Los pesos  $\hat{W}$  obtenidos para los errores en cada año aparecen recogidos en la tabla 5 junto con el sesgo medio de las instituciones consideradas, y muestran que las ponderaciones están relacionadas de forma inversa con el sesgo medio correspondiente a cada año. Así, cuando el sesgo es negativo las mayores ponderaciones corresponden a los valores positivos del vector  $V$  mientras en caso de sesgo positivo se asignan las mayores ponderaciones a los valores negativos de  $V$ , tal y como cabía esperar teniendo en cuenta que  $\hat{e}$  corrige el sesgo global.

**Tabla 5:** Ponderaciones  $\hat{W}_t$  de los errores en  $t$  según las medidas de entropía  $H_s(P)$  y  $H^2(P)$ .

Año	ESTADIO t-IV		ESTADIO t-1	
	Sesgo medio	$V = (-1,25, -0.5, 0, 0.5, 1,25)$	Sesgo medio	$V = (-0.35, -0.15, 0, 0.15, 0.35)$
2000	-1,25	(0, 0, 0, 0, 1)	-0,12	(0,4, 0, 0, 0, 0,6)
2001	-1,36	(0, 0, 0, 0, 1)	0,03	(0,48, 0, 0, 0, 0,52)
2002	-0,66	(0,25, 0, 0, 0, 0,75)	0,36	(1, 0, 0, 0, 0)
2003	-0,41	(0,32, 0, 0, 0, 0,68)	0,04	(0,53, 0, 0, 0, 0,47)
2004	-0,26	(0,33, 0, 0, 0, 0,66)	-0,10	(0,0,32, 0,23, 0,26, 0,19)

Cabe destacar además que las dos medidas de entropía conducen a resultados similares, confirmándose la coherencia entre los pesos estimados.

A partir de estos primeros resultados nos planteamos una aplicación más completa en la que consideramos todo el panel de predicciones para el estadio  $t$ -IV. De acuerdo con el enfoque recursivo, para las ponderaciones de cada año  $t$  se emplea toda la información anterior  $t$ -1 y así sucesivamente. En el caso de que un organismo no realice la predicción en  $t$ , las ponderaciones calculadas para el mismo en  $t$ -1 se distribuyen entre el resto de organismos proporcionalmente a sus ponderaciones.

---

<sup>5</sup> Estos análisis se han llevado a cabo simulando 10 predictores y estudiando su comportamiento en distintos escenarios. Así, se observa que cuando todos los individuos predicen con total exactitud, la ponderaciones corresponden a la media aritmética; cuando sólo un individuo predice correctamente y el resto con mucha inexactitud la ponderación máxima corresponde al primero; cuando todos los años los individuos realizan la misma predicción (siendo la realización constante) las ponderaciones no varían y si la dispersión observada entre los predictores a lo largo del tiempo es reducida (esto es, si ninguno de ellos aporta información significativa sobre su capacidad predictiva) entonces las ponderaciones no varían o lo hacen en pequeña medida.

Los resultados obtenidos para la medida de entropía de Shannon se resumen en las tabla 6.

**Tabla 6:** Ponderaciones para t-IV ( $\hat{\alpha}_{t-IV}$ ) según la medida de Shannon  $H_s(P)$

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
<b>AFI</b>			0,003		0,003	0,003	0,003	0,003			0,003
<b>Argentaria</b>			0,003	0,003	0,003						
<b>Banesto</b>						0,003	0,003		0,003	0,003	
<b>BBVA</b>			0,003	0,003	0,003		0,003			0,003	
<b>BCH (BSCH)</b>		0,003	0,003	0,003		0,003	0,003	0,003	0,003	0,003	0,003
<b>Caja Madrid</b>									0,003		
<b>Carlos III</b>					0,003	0,003	0,003		0,003	0,003	0,003
<b>Comisión Europea</b>		0,030	0,029	0,029	0,029	0,029	0,029	0,031	0,031	0,031	0,031
<b>Ceprede</b>	0,582	0,579	0,571	0,570	0,568	0,563	0,561	0,609	0,597	0,603	0,601
<b>Commerzbank</b>									0,003		
<b>Cons.Cam.Comercio</b>											0,003
<b>Consensus Economics</b>		0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003
<b>Dresdner Bank</b>									0,003		
<b>Economist Intelligent Unit</b>											0,003
<b>FMI</b>		0,017	0,017	0,017	0,017	0,017	0,016	0,018	0,018	0,018	0,018
<b>FUNCAS</b>					0,003	0,003	0,003	0,003	0,003	0,003	0,003
<b>Goldman-Sachs</b>	0,020	0,049	0,048	0,048	0,048	0,048	0,047	0,052	0,050	0,051	0,051
<b>ICO</b>						0,003	0,003	0,003	0,003	0,003	0,003
<b>IEE</b>			0,003	0,003	0,003	0,003	0,003				
<b>Intermoney</b>									0,003	0,003	0,003
<b>J.P.Morgan</b>		0,003	0,003	0,003	0,003	0,003	0,003				
<b>La Caixa</b>						0,003	0,003	0,003	0,003	0,003	0,003
<b>Link</b>	0,020	0,066	0,065	0,065	0,065	0,064	0,064				
<b>Merril Lynch</b>					0,003	0,003	0,003	0,003			
<b>Morgan Stanley</b>									0,003	0,003	0,003
<b>OCDE</b>	0,382	0,189	0,187	0,186	0,186	0,184	0,184	0,199	0,195	0,197	0,197
<b>Presupuestos</b>				0,003		0,003	0,003	0,003	0,003		
<b>Santander Invst.</b>				0,003							
<b>The Economist</b>	0,020	0,052	0,051	0,051	0,051	0,051	0,051	0,055	0,054	0,054	0,054
<b>UBS Warburg</b>									0,003	0,003	0,003

En términos generales, las ponderaciones estimadas (que muestran una elevada estabilidad temporal) indican de nuevo que no existe evidencia empírica que permita justificar el empleo de la media aritmética como sistema de ponderaciones en t-IV.

Además, si tenemos en cuenta la información que suministra la heterogeneidad de las ponderaciones, se puede apreciar que el mayor porcentaje de los pesos se concentra en un conjunto concreto de organismos, hecho que se repite, en el estadio I, aunque con ponderaciones más similares entre ellos, tal y como mostramos en la tabla 7.

**Tabla 7:** Ponderaciones para t-I ( $\hat{\alpha}_{t-I}$ ) según la medida de Shannon Hs(P)

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
<b>AFI</b>			0,004	0,004	0,004	0,004	0,005	0,006	0,006	0,007	0,007
<b>Argentaria</b>			0,004	0,004	0,004						
<b>Banesto</b>					0,004	0,004	0,005	0,006	0,006	0,007	
<b>BBVA</b>			0,004	0,004	0,004		0,005	0,006	0,006	0,008	0,009
<b>BCH (BSCH)</b>		0,004	0,004	0,004	0,004	0,004	0,005	0,006	0,006	0,007	0,009
<b>Caja Madrid</b>									0,004	0,004	0,007
<b>Carlos III</b>					0,004	0,004	0,005	0,006	0,006	0,008	0,009
<b>Comisión Europea</b>	0,152	0,151	0,149	0,149	0,145	0,145	0,167	0,156	0,178	0,202	0,273
<b>Ceprede</b>	0,176	0,175	0,173	0,172	0,168	0,168	0,194	0,188	0,213	0,233	0,301
<b>Commerzbank</b>									0,004		
<b>Cons.Cam.Comercio</b>									0,004	0,004	0,009
<b>Consensus Economics</b>	0,004	0,004	0,004	0,004	0,004	0,004	0,005	0,006	0,006	0,007	0,009
<b>Dresdner Bank</b>									0,004		
<b>Economist Intelligent Unit</b>											0,00
<b>FMI</b>	0,105	0,105	0,103	0,103	0,100	0,100	0,116	0,119	0,136	0,149	0,173
<b>FUNCAS</b>					0,005	0,005	0,006	0,006	0,006	0,007	0,009
<b>Goldman-Sachs</b>	0,152	0,151	0,149	0,149	0,145	0,145	0,167	0,164	0,187	0,201	
<b>ICO</b>					0,005	0,005	0,006	0,006	0,006	0,007	0,009
<b>IEE</b>			0,004	0,004	0,004	0,004	0,005	0,006		0,005	
<b>Intermoney</b>									0,006		0,003
<b>J.P.Morgan</b>			0,004	0,004	0,004	0,004	0,005				
<b>La Caixa</b>						0,004	0,005	0,006	0,006	0,007	0,009
<b>Link</b>	0,141	0,141	0,139	0,138	0,134	0,135					
<b>Merril Lynch</b>				0,004	0,004	0,004	0,005	0,006			
<b>Morgan Stanley</b>									0,004	0,004	0,009
<b>OCDE</b>	0,121	0,121	0,119	0,119	0,116	0,116	0,134	0,129		0,115	0,145
<b>Presupuestos</b>	0,152	0,151	0,149	0,149	0,145	0,145	0,167	0,173	0,197		
<b>The Economist</b>	0,004	0,004	0,004	0,004	0,004	0,004	0,005	0,006	0,006	0,007	
<b>UBS Warburg</b>								0,003	0,004	0,004	0,007

Si consideramos únicamente los organismos que de forma sistemática concentran el mayor porcentaje de los pesos (en concreto Comisión Europea, Ceprede, Goldman-Sachs, Link, OCDE y The Economist que representan el 95% de la ponderación total en el estadio tIV) los resultados conducen a algunas reflexiones adicionales que justifican el empleo de la técnica desarrollada: la consideración de un grupo de expertos selectos como el descrito permitiría el cálculo eficiente de las ponderaciones en  $Y = \hat{Y}\alpha + u$  con técnicas de regresión convencionales si la información estadística lo permite ( $n > F$ ), con la consiguiente ganancia de operatividad. Queda así justificada una ventaja de la técnica empleada, consistente en permitir calibrar a los organismos de forma heterogénea aún cuando la información sobre ellos sea escasa. Con el

objetivo de llevar a cabo un análisis más exhaustivo nos planteamos a continuación analizar la capacidad predictiva de la metodología propuesta, comparando los errores asociados a la predicción combinada mediante los procedimientos de maximización de entropía con los correspondientes a la predicción obtenida como media aritmética. Los cálculos han sido llevados a cabo tanto para el panel total de instituciones como para el subgrupo selecto (Comisión Europea, Ceprede, Goldman-Sachs, Link, OCDE y The Economist) y adicionalmente hemás hemos considerado también la predicción combinada mediante mínimos cuadrados ordinarios con un subgrupo reducido formado por la Comisión Europea y Ceprede, que acumulan el 70% de la ponderación, estimando los pesos iniciales para la predicción del año 2000 con la muestra de 1994-1999 y reestimándolos sucesivamente para las predicciones posteriores, posibilidad que como hemos justificado no resulta factible sobre el panel total. Las predicciones obtenidas según los distintos métodos de combinación de predicciones se resumen en la tabla 8, donde también se recoge la raíz del error cuadrático medio asociada a cada una de las alternativas.

**Tabla 8:** Predicciones combinadas y errores asociados

Año	Tasa IPC	PANEL				
		TOTAL		SUBGRUPO	SUBGRUPO Reducido	
		Media Aritmética	Media ponderada	Media ponderada	Media ponderada	Media ponderada pesos MCO
<b>1995</b>	4,64	3,68	3,69	3,69	3,68	
<b>1996</b>	3,56	4,54	4,35	4,36	4,20	
<b>1997</b>	2,01	3,36	3,30	3,30	3,23	
<b>1998</b>	1,81	2,41	2,44	2,45	2,35	
<b>1999</b>	2,34	2,38	2,31	2,31	2,27	
<b>2000</b>	3,39	2,12	2,24	2,25	2,23	2,27
<b>2001</b>	3,59	2,22	2,43	2,44	2,50	3,46
<b>2002</b>	3,07	2,41	2,55	2,56	2,60	3,37
<b>2003</b>	3,04	2,63	2,50	2,49	2,52	3,56
<b>2004</b>	3,04	2,71	2,45	2,42	2,40	3,42
<b>2005</b>	3,71	3,68	3,69	2,40	2,37	3,56
<b>ECM</b>		0,86	0,80	0,80	0,75	0,3

*Fuente:* Elaboración propia a partir IPC del INE (Instituto Nacional de Estadística) y CEPREDE, Instituto L.R. Klein

Como puede apreciarse, los resultados referidos al conjunto de instituciones presentan una ligera ventaja, en cuanto a su capacidad predictiva, a favor de las técnicas de maximización. No

obstante, si tenemos presente el tipo de variable sobre el que trabajamos (tasas de variación del IPC), cuyo recorrido es bastante reducido y para la que existe una considerable similitud entre las predicciones de las distintas instituciones, no es posible apreciar diferencias sustanciales en las predicciones combinadas con las nuevas ponderaciones. No obstante, los resultados obtenidos con la preselección de un grupo de organismos predictores muestran una mejora significativa de la capacidad predictiva, tal y como se aprecia en la reducción del error cuadrático medio obtenido tanto con la media ponderada como con la aplicación de pesos basados en la regresión por Mínimos Cuadrados Ordinarios. Esta ganancia de precisión sería por tanto una ventaja adicional que debe ser tenida en cuenta a la hora de valorar la combinación basada en maximización de entropías.

Como consecuencia de los comentarios anteriores, podemos concluir que la metodología propuesta resultaría de utilidad al permitir la selección de un subgrupo de predictores y la correspondiente predicción combinada, con la consiguiente ganancia tanto en operatividad como en precisión. Cabe añadir además que, si bien el uso de la media aritmética se justifica a menudo por su sencillez de cálculo, el desarrollo de programas informáticos permite la aplicación de alternativas como la aquí presentada con un coste operativo perfectamente asumible, máxime si se tiene en cuenta que no sería necesario replicar este procedimiento en cada estadio de predicción sino que podría plantearse como una etapa preliminar para seleccionar un grupo de predictores de calidad.

## **5. CONCLUSIONES**

En este trabajo hemos investigado las posibilidades que las medidas de información ofrecen para la combinación de predicciones. En concreto, la aplicación del principio de máxima entropía nos permite, en un contexto de escasa información, estimar ponderaciones que calibran de modo desigual a los predictores individuales. Puesto que la práctica más habitual es efectuar la combinación a partir de la media aritmética sobre los predictores individuales, la propuesta se convierte en una alternativa para tratar de aprovechar la información disponible al mismo tiempo que contribuye a la solución de problemas como el reducido tamaño muestral o la existencia de colinealidad entre las predicciones individuales.

Los análisis aplicados sobre el panel de predicciones de crecimiento de los precios de consumo en España permiten apreciar que la utilización de las medidas de entropía ofrece la posibilidad de discriminar entre predictores, seleccionando un grupo cuyas ponderaciones son las más elevadas y estables en el tiempo. La consideración de este panel reducido de predictores aporta dos tipos de ventajas: por una parte, al disminuir el tamaño del panel proporciona ganancias operativas, permitiendo la aplicación de métodos de regresión; en segundo lugar, se obtienen ganancias significativas en la precisión de la predicción combinada, especialmente cuando se utiliza la media ponderada o pesos basados en la aplicación de Mínimos Cuadrados Ordinarios.

## REFERENCIAS

- ALVARGONZÁLEZ, M.; PÉREZ, R. (1989): “Información cuadrática e independencia en información”, *Actas XIV Jornadas Hispano-Lusas de Matemáticas*, Tenerife, II, p. 753-758.
- BATES, J.M.; GRANGER, C.W.J. (1969): “The Combination of Forecasts”, *Operational Research Quarterly*, Vol. 20, nº 4, p. 451-468.
- BROOKE, A.; KENDRICK, D.; MEERAUS, A.; RAMAN, R. (1998): *GAMS, A users guide*, Gams Development Corporation, Washington (USA), <http://www.gams.com/docs/gams/GAMSUsersGuide.pdf>.
- CEPREDE-INSTITUTO L.R.KLEIN (varios semestres): *Informes Semestrales de Perspectivas Económicas y Empresariales*.
- GALLO, G.M.; GRANGER, C.W.J.; JEON, Y. (1999): The impact of the use of forecast in information sets, *Discussion Paper 99-18, University of California*, San Diego.
- GOLAN, A.; JUDGE, G.; MILLER, D. (1996): *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, Jonh Wiley & Sons Ltd, London.
- GRANGER, C.W.J.; RAMANATHAN, C. (1984): “Improved Methods of Combining Forecasts”, *Journal of Forecasting*, Vol. 3, nº 2, p. 197-204.
- HAVRDA, J.; CHARVAT, F. (1967): “Quantification method of classification processes”, *Kybernetika*, nº 3, p. 30-35.
- INE (Varios años): *Boletín Mensual de Estadística*, <http://www.ine.es>
- JAYNES, E.T. (1957a): “Information Theory and Statistical Mechanics I”, *Physics Review*, Vol. 106, nº 4, p. 620-630.
- JAYNES, E.T. (1957b): “Information Theory and Statistical Mechanics II”, *Physics Review*, Vol. 108, nº 2, p. 171-190.
- KALMAN, R. (1960): “A New Approach to Linear Filtering and Prediction Problems”, *Journal of Basic Engineering, Transactions ASME, Serie D*, Vol. 82, nº 1, p. 35-45.
- KULLBACK, S; LEIBLER, R.A. (1951): “On information and sufficiency”, *Annals of Mathematical Statistics*, Vol. 22, nº 1, p. 79-86.
- LÓPEZ, A.J.; PÉREZ, R. (1991): Indicadores de Desigualdad y Pobreza. Nuevas alternativas, *Documento de trabajo nº37, Facultad de CC. Económicas y Empresariales, Universidad de Oviedo*.
- MANKIW, N.G.; SHAPIRO, M.D. (1986): “News or Noise? An analysis of GNP revisions”, *NBER Working Paper Series, Working Paper n.1939*.
- MORENO, B. (2005): *Combinación de predicciones y métodos de evaluación: Nuevas alternativas basadas en medidas de información*, Tesis Doctoral, Universidad de Oviedo.

- MORENO, B.; LÓPEZ, A.J.; PÉREZ, R. (2005): “Combining Forecasts through Information Measures”, *55th Session of the International Statistical Institute*, Sydney.
- NEWBOLD, P.; GRANGER, C.W.J. (1974): “Experience with Forecasting Univariate Time Series and the Combination of Forecasts”, *Journal of the Royal Statistical Society*, Serie A, nº 137. Part. 2, p. 131-165.
- O’CONNOR, M.; REMUS, W.; GRIGGS, K. (2000): “Does updating judgmental forecast improve forecast accuracy?”, *International Journal of Forecasting*, Vol.16, nº 1, p. 101-109.
- PÉREZ, R. (1985): “Estimación de la incertidumbre, la incertidumbre útil y la inquietud en poblaciones finitas: una aplicación a las medidas de desigualdad”, *Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales*, Tomo LXXIX, nº 4, p. 651-654.
- PÉREZ, R.; CASO, C.; GIL M.A. (1986): “Unbiased estimation of income inequality”, *Statistische Hefte*, nº 27, p. 227-237.
- PULIDO, A (1998): *Una apuesta por el futuro. Predicciones y profecías económicas*, Ed. Pirámide, Madrid.
- PULIDO, A. (2001): “Fuentes internacionales y nacionales en predicción económica”, *Revista Fuentes Estadísticas*, nº 55, <http://www.fuentesestadisticas.com>.
- PONS, J. (2003): “Obstinación, reputación y efecto rebaño”, *Revista de Economía Aplicada*, nº 12, Vol. XI, p. 97-114.
- RÍO, M. J.; PÉREZ, R. (1987): “El índice cuadrático como medida de la concentración industrial”, *XII Jornadas Luso-Espanholas de Matemática*, Braga (Portugal).
- SHANNON, C.E. (1948): “A Mathematical Theory of Communication”, *Bell System Technology Journal*, Vol. 27, nº 3, p. 379-423.
- THEIL, H. (1971): *Principles of Econometrics*, John Wiley and Sons, New York.