

COMBINACIÓN DE PREDICCIONES SOBRE EL CRECIMIENTO ECONÓMICO EN ESPAÑA: UNA PROPUESTA BASADA EN MEDIDAS DE INFORMACIÓN

Blanca Moreno Cuartas

Departamento de Economía Aplicada

Universidad de Oviedo

E-mail: morenob@uniovi.es

Ana Jesús López Menéndez

Departamento de Economía Aplicada

Universidad de Oviedo

E-mail: anaj@uniovi.es

Rigoberto Pérez Suárez

Departamento de Economía Aplicada

Universidad de Oviedo

E-mail: rigo@uniovi.es

Resumen

La diversidad de fuentes de predicción existentes hace posible disponer de un amplio abanico de predicciones para el crecimiento económico, sugiriendo la conveniencia de efectuar un resumen de las mismas mediante algún mecanismo de combinación.

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 varias medidas de incertidumbre.

Como aplicación empírica utilizamos la evidencia disponible sobre perspectivas de crecimiento económico en España elaboradas por distintos organismos en diferentes estadios de predicción para analizar la idoneidad de la técnica de combinación propuesta comparando sus resultados con los asociados a la media aritmética habitualmente empleada.

Palabras Clave: Combinación de predicciones, Medidas de incertidumbre, Maximización de Entropía.

Área temática: Métodos Cuantitativos, Economía Nacional e Internacional.

1. Introducción.

El indudable interés que tiene para los distintos agentes económicos disponer de predicciones sobre el crecimiento económico futuro 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, existen numerosos organismos tanto públicos como privados que realizan regularmente predicciones cuantitativas sobre el crecimiento económico en España. Entre las instituciones públicas oficiales se encuentran el Ministerio de Economía, la Dirección General de Previsión y Coyuntura o el Banco de España, mientras los centros privados abarcan tanto entidades financieras nacionales (FUNCAS, BBVA, Banesto, La Caixa, Caja Madrid, etc) como otras vinculadas a organizaciones financieras internacionales (J.P. Morgan, Morgan-Stanley, Goldman-Sachs) o empresariales (CEOE, Consejo Superior de Cámaras de Comercio, IEE, AFI, ...). A ambos tipos de organismos debemos añadir también algunas iniciativas ligadas a centros universitarios o de investigación como CEPREDE, Hispalink, ICAE o el Instituto Flores de Lemus e instituciones internacionales como el FMI, la Comisión Europea, la OCDE o el Proyecto Link de Naciones Unidas. Algunas recopilaciones de estas predicciones se recogen en publicaciones como *The Economist* y *Consensus Economics*.

Teniendo en cuenta que los distintos organismos dedicados a la prospectiva difieren en lo que respecta a la información disponible, las variables consideradas y las técnicas de modelización, las predicciones económicas elaboradas pueden ser muy heterogéneas, pudiendo ser conveniente realizar una combinación de todas ellas.

La idea de la combinación de predicciones asume que cada predicción es capaz de capturar diferentes aspectos de la información, y por tanto una combinación de todas ellas resultará más precisa y aprovechará mejor la información disponible.

Los trabajos pioneros en este sentido corresponden a Bates y Granger (1969), quienes proponen técnicas para obtener una predicción de síntesis mediante combinaciones lineales de predicciones individuales, cuyos pesos se obtienen a partir de las correspondientes varianzas. Por su parte, 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 las variables explicativas.

El cálculo eficiente de las ponderaciones de la combinación exige disponer de suficiente información sobre las predicciones individuales que permita calibrar a cada predictor, aumentando la garantía de la estimación a medida que esta información aumenta. Sin embargo, en ocasiones no es posible disponer de suficiente información y en ese caso la alternativa más habitual es efectuar la combinación a partir de la media aritmética sobre los predictores individuales.

Puesto que la Teoría de la información estadística proporciona un marco de análisis muy adecuado para el tratamiento de problemas en ámbitos muy diversos de la economía (concentración industrial, desigualdad de la renta, dependencia estadística, etc), en este artículo estudiamos las posibilidades que las medidas de información ofrecen para la combinación de predicciones en un contexto de insuficiencia de información.

En concreto, empleamos el principio de Maximización de Entropía para las medidas de incertidumbre de Shannon (1948) y cuadrática (Pérez, 1985) y utilizamos la evidencia empírica disponible para comparar las ponderaciones estimadas por este procedimiento con las asociadas a la media aritmética.

Como consecuencia de estas consideraciones el resto del artículo se estructura en cuatro secciones. En el apartado que sigue 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, de los que haremos uso en posteriores apartados.

En la sección 3 estudiamos brevemente los métodos de combinación tradicionales con el propósito de justificar el interés de la metodología que proponemos basada en medidas la información. La cuarta sección contiene el desarrollo de esta metodología y los resultados obtenidos a partir de las predicciones que numerosos organismos realizan sobre el crecimiento económico a corto y medio plazo 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 aparecen en la sección 5.

2. Medidas de información y principios de optimización.

En 1948 Shannon introdujo una medida de entropía que cuantifica la cantidad de incertidumbre asociada a un experimento aleatorio. Así, dada una variable aleatoria X que toma n valores x_i con una 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 incertidumbre de Shannon viene dada por la expresión:

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

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

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

Un caso particular de esta medida, 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 que vienen dadas por la expresión:

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

La entropía de una variable aleatoria discreta X alcanza su valor máximo cuando todos sus valores adoptan la misma probabilidad (y entonces P es una distribución uniforme). Ello 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 conocen determinados valores o características (que recogen información parcial disponible sobre la distribución) que son impuestos en forma de restricciones de igualdad. En ese caso es posible ajustar una distribución P mediante la aplicación del principio de entropía máxima teniendo en cuenta toda 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 determinar ésta 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: $H(X) = H(p_1, \dots, p_n)$

sujeta a las restricciones: $\sum_{i=1}^n p_i = 1$ con $p_i \geq 0$ $i= 1, \dots, n$

$$\sum_{i=1}^n p_i g_r(x_i) = a_r \quad r= 1, \dots, m$$

Mediante la resolución del problema de optimización obtenemos la distribución de probabilidad $\hat{P} = \{\hat{p}_i\}$. Hay que señalar que la distribución de máxima entropía no tiene una solución cerrada y se deben emplear técnicas de optimización numérica para calcular las probabilidades.

Además de las medidas de entropía, la teoría de la información proporciona un conjunto de medidas de divergencia. Así, 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 $Q=\{q_i\}$, es posible definir 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:

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

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 “distancia” entre dos distribuciones.

En este caso, si se supone que además de una cierta información parcial sobre la distribución en estudio P existe una información previa de tal distribución, o una distribución a priori Q , la idea de la optimización consiste ahora 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:

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

sujeta a las restricciones: $\sum_{i=1}^n p_i = 1$ con $p_i \geq 0 \quad i=1, \dots, n$

$$\sum_{i=1}^n p_i g_r(x_i) = a_r \quad r=1, \dots, m$$

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 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 predicción de consenso, 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 α es un vector de ponderaciones de las predicciones individuales. Dependiendo de cómo se formalice la relación causal, y se obtengan las

estimaciones de α , 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 fueron 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 de las predicciones individuales. Posteriormente este planteamiento es ampliado para el caso de más predicciones por Newbold y Granger (1974),

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

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

Granger y Ramanathan (1984) ponen de manifiesto que las ponderaciones obtenidas con métodos convencionales de combinación admiten una interpretación como vector de coeficientes de la proyección lineal de y_{t+h} a partir de las predicciones de los F métodos: $y_{t+h} = \alpha_1 \hat{y}_{t+h,t}^1 + \dots + \alpha_F \hat{y}_{t+h,t}^F + u_{t+h} = \hat{Y}_{t+h,t} \alpha + u_{t+h}$.

Al desconocer el verdadero valor de y_{t+h} los pesos se obtienen a partir de las n predicciones pasadas. Así, 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 por 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 α 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 es nueva (o se trata de una situación con poca historia) o bien porque gradualmente se van incorporando nuevos predictores y no se dispone de información suficiente sobre ellos que permita calibrar su calidad de

forma eficiente. 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 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 puede ser formulado, 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 α y, puesto que éste no puede ser cuantificado directamente, debemos de emplear medidas indirectas a partir de la información observada.

Cuando la información de Y es especificada sin error y en términos de 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: $Y = \hat{Y}\alpha = \hat{Y}P$, 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 seleccionar un único vector, objetivo que sin embargo no resulta alcanzable con la información disponible¹.

¹ Si queremos realizar inferencias, debemos de escoger a partir de la información disponible una estimación de los parámetros P tal que $\hat{P} = AY$, donde A es un operador desconocido que debe cumplir: $Y = \hat{Y}\hat{P} \Rightarrow \hat{Y}AY = \hat{Y}A\hat{Y}P$, lo cual significa que $\hat{Y}A\hat{Y} = \hat{Y}$ y entonces, de acuerdo con Kalman (1960), A es una inversa generalizada. Esta restricción es de gran ayuda, ya que permite identificar la clase de soluciones entre las que debe estar la verdadera distribución P . Sin embargo, no nos da un criterio para escoger A o \hat{P} entre esa clase, puesto que \hat{P} no se puede distinguir del verdadero P . Por lo tanto existen un número infinito de operadores A que satisfacen $Y = \hat{Y}\alpha = \hat{Y}P$ y que conducen a estimaciones de \hat{P} dentro de la clase de posibles soluciones que contienen $(F-n)$ parámetros arbitrarios.

Para resolver este problema Jaynes (1957) propuso 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 :

$$\sum_{i=1}^F p_i = 1 \quad p_i \geq 0 \quad i=1, \dots, F$$

$$\sum_{i=1}^F \hat{y}_t^i p_i = y_t \quad t=1, \dots, n$$

Si el problema inverso puro es de nuevo definido considerando una variable de error $Y = \hat{Y}\alpha + u$ hablamos de un problema de inversión “general” que podrían entonces ser reformulado como: $Y = \hat{Y}\alpha + u = \hat{Y}P + uW$, siendo W el vector $1 \times n$ de ponderaciones del error de cada año, donde cada $w \in [0,1]$ y el problema de optimización será:

$$\text{Maximizar } H(P, W) = H(P) + H(W)$$

$$\text{Sujeto a: } Y = \hat{Y}P + uW$$

$$P' \ell = 1 \quad \text{donde } \ell \text{ es un vector } (F \times 1) \text{ de unos}$$

$$W' \ell_n = 1 \quad \text{donde } \ell_n \text{ es un vector } (n \times 1) \text{ de unos}$$

Dado que el conjunto de medidas de entropía es muy numeroso, los resultados de las estimaciones \hat{P} y \hat{W} podrían depender de la expresión considerada en cada caso. En este trabajo utilizaremos la entropía de Shannon habitualmente empleada y también la entropía cuadrática $H^2(X)$, estudiada por Pérez (1985) cuyas propiedades la hacen adecuada para diversos análisis económicos.

Más concretamente, entre las ventajas de la medida cuadrática podemos destacar las siguientes:

- (i). Cumple las principales propiedades que hacen a la entropía de Shannon adecuada como medida de incertidumbre.
- (ii). La familia de medidas de incertidumbre que incluyen en su expresión logaritmos presentan como principal limitación su escasa operatividad. Así, las medidas tipo Shannon presentan limitaciones de carácter inferencial, al no ser posible obtener estimadores insesgados basados en los analógicos (ya que

no se puede establecer una relación funcional entre la esperanza del logaritmo y el logaritmo de la esperanza). En cambio, las medidas cuadráticas presentan un buen comportamiento inferencial, permitiendo construir estimadores insesgados y deducir la varianza de los mismos².

- (iii). La incertidumbre cuadrática está acotada superiormente por 2, lo que supone la ventaja frente a la entropía de Shannon de disponer de una valoración indicativa del grado de incertidumbre.
- (iv). 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 (1991)], la concentración industrial [Río y Pérez (1987)] y la asociación de caracteres [Alvargonzález y Pérez (1987)].

La medida de incertidumbre cuadrática es función cóncava de P, lo que garantiza que la solución que se obtiene es 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, por lo que será 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 software específicos que permitan una optimización no lineal facilita la obtención de la solución numérica.

² Pérez (1985) obtuvo estimadores insesgados de la incertidumbre cuadrática y de la inquietud cuadrática para el muestreo con y sin reposición y también la varianza de los estimados (y un estimador insesgado de la misma) en los dos tipos de muestreo, observando una ganancia de precisión en el muestreo sin reposición. Posteriormente Caso (1988) desarrolló el estudio de estimadores de las medidas cuadráticas en el muestreo estratificado y analizó las varianzas y los estimadores de dichas varianzas, comprobando que se produce una ganancia de precisión en el muestreo estratificado respecto al simple.

4. Combinación de predicciones sobre crecimiento económico en España

En esta sección presentamos varias aplicaciones con el objetivo de obtener predicciones combinadas basadas en la maximización de entropía y adoptando como referencia las predicciones que distintas instituciones realizan sobre el crecimiento económico en España para el periodo 1994-2003.

Las perspectivas de crecimiento nacional aparecen en los *Informes Semestrales de Perspectivas Económicas y Empresariales* publicados por CEPREDE y el Instituto L.R. Klein. En concreto, cada año se publican dos informes semestrales donde se recogen las predicciones que las instituciones elaboran para el año en curso y el siguiente y la media de todas ellas.

Como paso previo al cálculo de las ponderaciones de los organismos empleando la maximización de medidas de entropía, analizamos la variabilidad que existe entre las predicciones individuales y entre éstas y la predicción de consenso, así como la volatilidad asociada a las predicciones realizadas con los diferentes estadios de la información disponible. De acuerdo con el estudio de Gallo, Granger y Jeon (1999), es previsible una disminución tanto en los sesgos de las predicciones individuales (y por tanto en la de consenso) como en la dispersión de los diferentes organismos a medida que se incorpora información más reciente o avanzamos en el estadio de predicción.

Este análisis previo puede aportarnos valiosa información para la segunda parte del estudio, donde 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 de ser más homogéneas en el último estadio de información que en el primero. Se analiza también la coherencia entre los pesos obtenidos con las medidas consideradas: la entropía de Shannon y la incertidumbre cuadrática.

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

Las perspectivas de crecimiento nacional en España (tasas reales de crecimiento interanual del PIB) son elaboradas por distintos organismos y aparecen recopiladas en los *Informes Semestrales de Perspectivas Económicas y Empresariales* publicados por CEPREDE y el Instituto L.R. Klein. En este apartado intentamos mostrar evidencia empírica sobre la convergencia en las predicciones individuales sobre crecimiento, y analizar en qué medida se tiende a una predicción media.

En concreto analizaremos el periodo 1994-2003, y para la tasa de crecimiento nacional de cada año 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).

Comenzando por la evaluación de predicciones cuantitativas según el estadio de predicción, si denotamos por $\hat{y}_{t+h,t}^i$ a la predicción realizada en t para un instante t+h, por un organismo ($i=1, \dots, F$), y 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 que denominaremos s ($s=1, \dots, h-2, h-1, h$).

Así pues, para t+h dispondremos de distintas predicciones, tanto individuales ($\hat{y}_{t+h,t+s}^i$) como de consenso ($\hat{y}_{t+h,t+s}^c$), de acuerdo con el instante s en el que nos situemos, siendo esperable que la predicción mejore a medida que se acerca el evento a predecir y que cuando $s=h$ la predicción coincida con la realización. En este sentido O'Connor, Remus y Griggs (2000) tratan de separar la reacción derivada de la información pasada de la derivada de la nueva información, analizando en qué medida la reacción también puede ser debida a información contextual.

Por otra parte, tal y como recogen Gallo, Granger y Jeon (1999), cuando se lleva a cabo una combinación de predicciones efectuadas por distintos organismos la correlación entre expertos hace que sea muy difícil apreciar las predicciones individuales (se converge a una predicción media). Para corroborar este hecho, estos autores proponen un modelo que explica cómo cada individuo u organismo genera su predicción cuando hay más expertos y canales de información entre ellos:

$$\hat{y}_{t+h,t+s}^i = \alpha + w_1^i \hat{y}_{t+h,t+s-1}^i + w_2^i \hat{y}_{t+h,t+s-1}^c + w_3^i \sigma_{t+h,t+s-1} + u_{t+h,t+s}^i$$

donde la predicción del organismo i en un momento t+s ($\hat{y}_{t+h,t+s}^i$) dependerá de la convicción que éste tiene en su predicción previa ($\hat{y}_{t+h,t+s-1}^i$), de la predicción de consenso anterior ($\hat{y}_{t+h,t+s-1}^c$) y la dispersión entre el resto de organismos observada en el estadio anterior $\sigma_{t+h,t+s-1}$. El último término de la regresión corresponde al error de predicción $u_{t+h,t+s}^i$, siendo de suponer 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.

La información disponible para un año t consiste en cuatro predicciones que se corresponden con otros tantos estadios de información, que denotaremos por 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 cuatros 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 de consenso en 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 dichas predicciones con la de consenso, así como sus sesgos al compararlas con las realizaciones, dadas por las tasas publicadas por el INE.

Para nuestra aplicación disponemos de 10 paneles³ correspondientes a los años del periodo 1994-2003. Para cada panel anual se ha calculado en cada estadio de la información la predicción de consenso, el sesgo cometido por cada institución y por el consenso y la dispersión entre instituciones. Un resumen de los resultados aparece recogido en la tabla 1.

³ No existe homogeneidad en las instituciones consideradas en cada panel, por lo que la base informativa para el periodo 1994-2003 incluye un número de instituciones comprendido entre 9 y 21.

Tabla 1: Sesgo medio y dispersión entre predicciones individuales

Año	Sesgo		Dispersión (Desviación típica)	
	Predicción de consenso inicial	Predicción de consenso final	Predicciones iniciales	Predicciones finales
1994	-0,89	-0,88	0,61	0,35
1995	-0,33	0,29	0,32	0,16
1996	0,98	-0,17	0,44	0,12
1997	-0,95	-0,75	0,2	0,16
1998	-1,29	-0,51	0,18	0,04
1999	-0,64	-0,59	0,15	0,12
2000	-1,04	-0,17	0,39	0,08
2001	0,59	-0,06	0,27	0,13
2002	1,17	-0,09	0,19	0,14
2003	0,70	0,15	0,30	0,20

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

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. Este comportamiento sugiere que el empleo de la media aritmética en la combinación resulta más adecuado en los últimos estadios de predicción.

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 el crecimiento económico en España, tal y como puede observarse en la tabla 2.

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

Año de predicción t	Estadios de predicción			
	$\hat{Y}_{t,Junio\ t-1}^F$	$\hat{Y}_{t,Diciembre\ t-1}^F$	$\hat{Y}_{t,Junio\ t}^F$	$\hat{Y}_{t,Diciembre\ t}^F$
1994	7	12	22	15
1995	7	13	22	13
1996	17	12	15	17
1997	13	17	20	20
1998	19	20	19	20
1999	19	20	23	23
2000	23	23	22	22
2001	21	22	21	23
2002	18	23	27	27
2003	21	26	26	25

Si deseamos obtener una predicción de consenso de los predictores individuales mediante técnicas de regresión, con el objetivo de garantizar un proceso de estimación adecuado necesitaríamos que la información disponible fuera suficiente, esto es $n > F$, y por tanto que para cada estadio en los diez años considerados (1994-2003) tuviésemos menos de 9 predicciones individuales. Sin embargo, el rápido incremento del número de organismos que realizan previsiones 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 que tuviesen mayor tradición en la elaboración de predicciones, si bien deberíamos plantearnos si de este modo 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.

Como ya hemos señalado, la alternativa más habitual es efectuar la combinación a partir de la media aritmética sobre los predictores individuales, lo que implica que se asigna la misma importancia a todos los individuos que efectúan la predicción, decisión que podría justificarse para el caso de no disponer de ninguna información sobre ellos que permita calibrarlos de modo desigual. El problema que trataremos en este apartado es el de ajustar los pesos a priori (que asociamos con los de la media) 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 α , tal y como explicamos en la sección anterior.

Además de las observaciones pasadas, podemos tener una información inicial o a priori sobre las ponderaciones de los organismos α^0 . En este 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 α que, cumpliendo las restricciones, divergen menos de los asociados a la media aritmética, y nos dará “información para discriminar a favor de α contra α^0 (media aritmética)”. En ese caso los pesos son $\alpha_i^0 = q_i = 1/F$, y la distribución Q considerada

corresponde con la distribución 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 la medida 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 GAMS (*General Algebraic Modeling System*)⁴, que es un lenguaje de programación 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.

El problema se debe no a los programas implementados, sino a la información empírica recogida en los paneles. Así, cuando todos los organismos cometen un sesgo en un estadio en la misma dirección (por ejemplo positivo), para que se cumpla la restricción $\hat{Y}P = Y$ es necesario que la ponderación de alguno de ellos sea negativa. Si por el contrario todos los organismos predicen a la baja es necesario que al menos la ponderación de uno de ellos supere la unidad. Puesto que este tipo de situaciones se produce de forma recurrente, al considerar un problema inverso puro, y restringir a la unidad los pesos que a priori fijamos como positivos, nuestros programas de maximización podrían no tener solución.

Como consecuencia, resulta aconsejable una segunda alternativa consistente en replantear el problema de inversión puro y considerar 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).

⁴ 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/>.

Entonces, el problema inverso puro 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.

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. 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:

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

donde V es una matriz $(nxnJ)$ 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:

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

$$\text{Sujeto a: } Y = \hat{Y}P + VW$$

$$P' \ell = 1, \text{ donde } \ell \text{ es un vector } (Fx1) \text{ de unos}$$

$$W'(I_n \otimes \ell_n) = 1, \text{ donde } \ell_n \text{ es un vector } (nx1) \text{ de unos y } \otimes \text{ es el producto de}$$

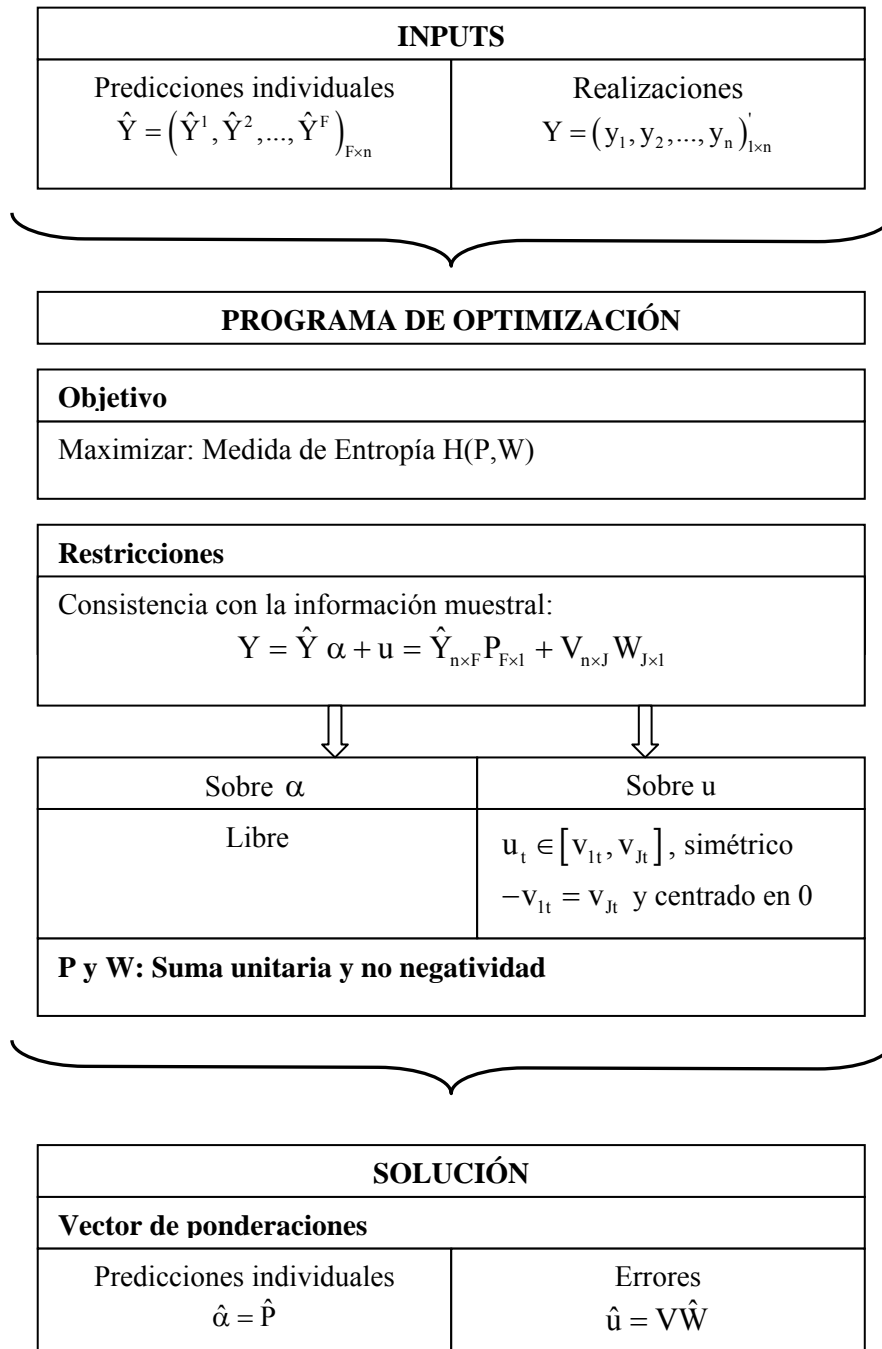
Kronecker.

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 distintos programas en GAMS, considerando distintos vectores V para el error⁵ y diferentes niveles de información disponible, y que siguen el esquema representado en la figura 1.

⁵ 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.

Figura 1: Esquema de programas de optimización



En la tabla 3 recogemos las ponderaciones estimadas en los estadios de predicción extremos (t-IV y t-I), y considerando los vectores V de la tabla 4, para algunas instituciones con las medidas de entropía de Shannon y la cuadrática, obtenidas incorporando de forma recursiva la información disponible para el período 1998-2001.

Tabla 3: Ponderaciones estimadas para estadios extremos según medidas de entropía

Información disponible	$\hat{\alpha}_{t-IV}$				$\hat{\alpha}_{t-I}$			
	1998		1998-2001		1998		1998-2000	
ORGANISMOS	Hs(P)	H ² (P)	H(P)	H ² (P)	Hs(P)	H ² (P)	Hs(P)	H ² (P)
Analistas Financieros Internac.	0,011	0,001	0,011	0,001	0,083	0,083	0,083	0,083
BSCH	0,011	0,001	0,011	0,001	0,083	0,083	0,083	0,083
Comisión Europea	0,056	0,117	0,056	0,117	0,083	0,083	0,083	0,083
Ceprede	0,005	0,001	0,005	0,001	0,083	0,083	0,083	0,083
Consensus Economics	0,024	0,001	0,024	0,001	0,083	0,083	0,083	0,083
FMI	0,697	0,597	0,697	0,597	0,083	0,083	0,083	0,083
FUNCAS	0,13	0,277	0,13	0,277	0,083	0,083	0,083	0,083
Goldman-Sachs	0,002	0,001	0,002	0,001	0,083	0,083	0,083	0,083
J.P. Morgan	0,005	0,001	0,005	0,001	0,083	0,083	0,083	0,083
Merril Lynch	0,024	0,001	0,024	0,001	0,083	0,083	0,083	0,083
OCDE	0,024	0,001	0,024	0,001	0,083	0,083	0,083	0,083
The Economist	0,011	0,001	0,011	0,001	0,083	0,083	0,083	0,083

Tabla 4: Vectores V

Estadios	J=5	$-v_{t1} = v_{t5}$
t-IV	(-1, -0.5, 0, 0.5, 1)	
t-I	(-0.6, -0.3, 0, 0.3, 0.6)	

Los resultados obtenidos permiten apreciar que en el estadio t-I las ponderaciones son más homogéneas que en el estadio inicial t-IV, por lo que si bien en ese caso existe evidencia a favor de la aplicación de la media aritmética para obtener la predicción combinada, no ocurre lo mismo para el estadio t-IV.

Si realizamos la comparación de las ponderaciones en cada estadio, se puede observar que en el estadio t-IV las ponderaciones obtenidas para 1998 no varían cuando aumentamos la información que se considera en la estimación (de 1998-2001), mientras que en el estadio t-I ocurre lo contrario. Una explicación a este hecho (corroborada mediante simulaciones)⁶ es que cuando la información suministrada de un año a otro no resulta significativa, las ponderaciones no varían.

⁶ Estas simulaciones se han llevado a cabo generando 10 predictores y analizando 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 hacen en pequeña medida.

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

Tabla 5: Ponderaciones \hat{W}_t de los errores en t según diferentes medidas de entropía

	Sesgo medio	ESTADIO t-IV	$V = (-1, -0.5, 0, 0.5, 1)$
		$H_s(P)$	$H^2(P)$
1998	-1.32	(0, 0, 0, 0, 1)	(0, 0, 0, 0, 1)
1999	-0.91	(0.15, 0.11, 0, 0, 0.74)	(0.15, 0.11, 0, 0, 0.74)
2000	-0.61	(0, 0, 0, 0.42, 0.58)	(0, 0, 0, 0.31, 0.69)
2001	0.65	(0.32, 0.63, 0, 0.02, 0.03)	(0.31, 0.63, 0, 0.02, 0.04)
	Sesgo medio	ESTADIO t-1	$V = (-0.6, -0.3, 0, 0.3, 0.6)$
		$H_s(P)$	$H^2(P)$
1998	-0.5	(0, 0, 0, 0.32, 0.68)	(0, 0, 0, 0.32, 0.68)
1999	-0.61	(0, 0, 0, 0, 1)	(0, 0, 0, 0, 1)
2000	-0.16	(0.36, 0, 0, 0, 0.64)	(0.36, 0, 0, 0, 0.64)

Cabe destacar que todas las 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.

Los resultados obtenidos⁷ para la medida de entropía de Shannon se resumen en la tabla 6.

⁷ Puesto que en este caso el volumen de información es mayor, el vector de ponderaciones que adoptamos es ahora $V = (-1.2, -0.6, 0, 0.6, 1.2)$.

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

	1995	1996	1997	1998	1999	2000	2001	2002	2003
Analistas Financieros Int.		0,02		0,01	0,00	0,00	0,00	0,00	
Argentaria		0,02	0,02	0,03					
BBVA			0,02	0,04	0,01	0,01	0,01	0,01	
BSCH		0,02	0,02	0,03	0,00	0,00	0,00	0,00	0,00
Carlos III						0,00	0,00	0,00	0,00
Comisión Europea	0,21	0,14	0,14	0,12	0,29	0,26	0,26	0,27	0,31
Ceprede	0,12	0,10	0,11	0,09	0,04	0,03	0,03	0,03	0,04
Consensus Economics		0,02	0,02	0,03	0,01	0,01	0,01	0,01	0,01
FMI		0,09	0,09	0,08	0,23	0,19	0,19	0,20	0,24
FUNCAS				0,01	0,03	0,03	0,03	0,03	0,03
Goldman-Sachs	0,07	0,08	0,08	0,07	0,03	0,03	0,03	0,03	0,03
Instituto de Crédito Oficial					0,00	0,00	0,00	0,00	0,00
IEE		0,02	0,02	0,03	0,05	0,04	0,04	0,04	
J.P. Morgan		0,02	0,02	0,03	0,00	0,00	0,00	0,00	0,00
La Caixa				0,01	0,01		0,01	0,01	0,01
LINK	0,16	0,11	0,11	0,11		0,13	0,13	0,14	
Merril Lynch				0,01	0,01	0,01	0,01	0,01	0,01
OCDE	0,31	0,21	0,22	0,18	0,21	0,18	0,18	0,19	0,22
Presupuestos Generales		0,02		0,01	0,03	0,02	0,02	0,02	0,02
The Economist	0,12	0,10	0,10	0,09	0,07	0,06	0,06	0,06	0,08

En términos generales, las ponderaciones estimadas (que muestran una elevada estabilidad temporal) indican de nuevo que no hay evidencia empírica que permita justificar el empleo de la media aritmética como sistema de ponderaciones en t-IV. Además, se puede apreciar que el mayor porcentaje de los pesos se concentra en un conjunto concreto de organismos (Comisión Europea, Fondo Monetario Internacional, OCDE y The Economist).

Si bien las estimaciones de los pesos divergen de la media aritmética, nos planteamos a continuación analizar la capacidad predictiva de la metodología propuesta. Para ello, comparamos los errores asociados a la predicción combinada con los procedimientos de maximización de entropía con los errores cometidos en la predicción calculada con la media aritmética, obteniéndose los resultados de la tabla 7.

Tabla 7: Predicciones de consenso y errores de predicción

Año	Tasa PIB	Predicción			Error		
		Hs(P)	H ² (P)	Media	H(p)	H ² (P)	Media
1995	2,8	2,41	2,4	2,47	0,39	0,4	0,33
1996	2,4	3,40	3,3	3,35	-1,00	-0,9	-0,95
1997	4,0	3,02	3,0	3,05	0,98	1,0	0,95
1998	4,3	3,01	3,0	3,02	1,29	1,3	1,28
1999	4,2	3,53	3,6	3,56	0,67	0,6	0,64
2000	4,2	3,29	3,3	3,18	0,91	0,9	1,02
2001	2,8	3,44	3,5	3,39	-0,64	-0,7	-0,59
2002	2,0	3,22	3,2	3,21	-1,22	-1,2	-1,21
2003	2,4	3,16	3,2	3,16	-0,76	-0,8	-0,76

Se observan claras similitudes entre los resultados, por lo que la aplicación del método de maximización de entropía no conlleva una reducción significativa del error de predicción. No obstante, conviene tener presente que el tipo de información con el que trabajamos (tasas de variación del PIB) lleva asociada una considerable similitud entre las predicciones de las instituciones, limitando el recorrido de los errores e impidiendo que se aprecien diferencias sustanciales en las predicciones combinadas con las nuevas ponderaciones.

Sin embargo, es posible observar que varios organismos concentran una considerable proporción de la ponderación total (en concreto, Comisión Europea, Fondo Monetario Internacional, OCDE y The Economist cuyo peso conjunto representa un 70%). La comparación de las predicciones combinadas para este grupo reducido de instituciones (ponderada subgrupo) con las asociadas a la media aritmética y a la media ponderada del total de instituciones del panel aparece en la tabla 9, donde se aprecia que la preselección de un grupo de organismos predictores no sólo no supone ningún tipo de limitación a la hora de valorar la metodología en función de su capacidad predictiva sino que incluso la mejora.

Tabla 8: Predicciones de consenso e indicadores de error de predicción

Predicción Consenso	1999	2000	2001	2002	2003	SUMA RESIDUOS	ECM	RAIZ ECM
MEDIA	3,56	3,18	3,39	3,21	3,16	3,84	0,77	0,88
MEDIA PONDERADA	3,53	3,29	3,44	3,22	3,16	3,75	0,75	0,87
PONDERADA SUBGRUPO	3,58	3,47	3,44	3,20	3,19	3,41	0,68	0,83

Además, la selección de un grupo de expertos permitiría el cálculo eficiente de las ponderaciones en $Y = \hat{Y}\alpha + u$ mediante técnicas de regresión convencionales si la información estadística lo permite ($n > F$).

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 del PIB en España permiten apreciar que los resultados obtenidos con las medidas de entropía consideradas no presentan diferencias significativas y además ofrecen la posibilidad de discriminar entre predictores, seleccionando un grupo cuyas ponderaciones correspondan a las más elevadas y estables en el tiempo.

La posibilidad de reducir el grupo de predictores considerados aporta dos tipos de ventajas: en primer lugar reduce el error de la predicción combinada del grupo con respecto al asociado al panel global de predictores o a la media aritmética, y en segundo lugar, al disminuir el tamaño del panel, proporciona ganancias operativas, permitiendo la aplicación de métodos de regresión.

Bibliografía.

1. Alvargonzález, M. y Pérez, R. (1989): “Información cuadrática e independencia en información”, *Actas XIV Jornadas Hispano-Lusas de Matemáticas*, Tenerife, **II**, pp. 753-758.
2. Bates, J.M. y Granger, C.W.J. (1969): “The Combination of Forecasts”, *Operational Research Quarterly*, **20**, pp. 451-468.
3. Brooke, A.; Kendrick, D.; Meeraus. A.y Raman, R. (1998): *GAMS, A users guide*, Gams Development Corporation, Washington (USA), <http://www.gams.com/docs/gams/GAMSUsersGuide.pdf>.
4. Caso, C. (1988): *Inferencias sobre medidas de información en el muestreo estratificado*, Tesis doctoral, Universidad de Oviedo.

5. Ceprede-Instituto L.R.Klein (varios semestres): *Informes Semestrales de Perspectivas Económicas y Empresariales*
6. Gallo, G.M.; Granger, C.W.J. y Jeon, Y. (1999): The impact of the use of forecast in information sets, *Discussion Paper 99-18, University of California, San Diego*.
7. Golan, A.; Judge, G. y Miller, D. (1996): *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, Jonh Wiley & Sons Ltd, London.
8. Granger, C.W.J. y Ramanathan, C. (1984): "Improved Methods of Combining Forecasts", *Journal of Forecasting*, **3**, pp. 197-204.
9. Havrda, J. y Charvat, F. (1967): "Quantification method of classification processes", *Kybernetika*, **3**, pp. 30-35.
10. INE (Varios años): *Boletín Mensual de Estadística*, <http://www.ine.es>.
11. Jaynes, E.T. (1957a): "Information Theory and Statistical Mechanics I", *Physics Review*, **106**, pp. 620-630.
12. Jaynes, E.T. (1957b): "Information Theory and Statistical Mechanics II", *Physics Review*, **108**, pp. 171-190.
13. Kalman, R. (1960): "A New Approach to Linear Filtering and Prediction Problems", *Journal of Basic Engineering, Transactions ASME, Serie D*, **82**, pp. 35-45.
14. Kullback, S y Leibler, R.A. (1951): "On information and sufficiency", *Annals of Mathematical Statistics*, **22**, pp. 79-86.
15. López, A.J. y Pérez, R. (1991): Indicadores de Desigualdad y Pobreza. Nuevas alternativas, *Documento de trabajo n°37, Facultad de CC. Económicas y Empresariales*, Universidad de Oviedo.
16. Moreno, B. (2005): *Combinación de predicciones y métodos de evaluación: Nuevas alternativas basadas en medidas de información*, Tesis Doctoral, Universidad de Oviedo.
17. Moreno, B.; López, A.J.; Pérez, R. (2005): "Combining Forecasts through Information Measures", *55th Session of the International Statistical Institute*, Sydney.
18. O'Connor, M.; Remus, W. y Griggs, K. (2000): "Does updating judgmental forecast improve forecast accuracy?", *International Journal of Forecasting*, **16**, pp. 101-109.

19. 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*, **LXXIX**, pp. 651-654.
20. Pérez R.; Caso, C. y Gil M.A. (1986): “Unbiased estimation of income inequality”, *Statistiche Hefte*, **27**, pp. 227-237.
21. Pulido, A (1998): *Una apuesta por el futuro. Predicciones y profecías económicas*, Ed. Pirámide, Madrid.
22. Pulido, A. (2001): “Fuentes internacionales y nacionales en predicción económica”, *Revista Fuentes Estadísticas*, **55**, <http://www.fuentesestadisticas.com>.
23. Pons, J. (2003): “Obstinación, reputación y efecto rebaño”, *Revista de Economía Aplicada*, **12**, pp. 97-114.
24. Río, M. J. y Pérez, R. (1987): “El índice cuadrático como medida de la concentración industrial”, *XII Jornadas Luso-Espanholas de Matemática*, Braga (Portugal).
25. Shannon, C.E. (1948): “A Mathematical Theory of Communication”, *Bell System Technology Journal*, **27**, pp. 379-423.
26. Theil, H. (1971): *Principles of Econometrics*, John Wiley and Sons, New York.