

Diseño de índices sintéticos de la coyuntura inflacionaria mediante análisis factorial ¹

**Enrique M. Quilis
Silvia Relloso**

**Instituto Nacional de Estadística
Subdirección General de Cuentas Nacionales**

¹ Las opiniones expresadas en este trabajo corresponden a los autores y no reflejan necesariamente las del Instituto Nacional de Estadística.

1. Introducción

El objetivo de este trabajo es la elaboración de un conjunto reducido de índices sintéticos que reflejen, desde una perspectiva estadística, los factores comunes que inciden sobre un amplio conjunto de series del sistema de Índices de Precios de Consumo (IPC). De esta forma, se dispone de una representación empírica de la pauta de movimientos conjuntos de dichos índices que será empleada dentro del informe que acompaña a los datos de la Contabilidad Nacional Trimestral (*El Momento Económico*).

Desde un punto de vista técnico, la metodología empleada se basa en la estimación de un modelo factorial estático. Dicho modelo permite tanto la exploración de la estructura latente de covariación como la elaboración de índices sintéticos para estimar los factores sistémicos del conjunto de series considerado.

2. Metodología

En este trabajo se asume que las tasas de variación interanual de las k rúbricas del IPC son generadas mediante un modelo factorial de la forma:

$$[1] \quad \pi_{i,t} = \mu_i + \lambda_{i,1}f_{1,t} + \dots + \lambda_{i,r}f_{r,t} + \varepsilon_{i,t}$$

siendo:

$$[2] \quad \pi_{i,t} = \frac{P_{i,t} - P_{i,t-12}}{P_{i,t-12}} * 100$$

donde p_{it} es el IPC correspondiente a la rúbrica i -ésima.

Las r variables $f_{j,t}$ representan los r factores comunes e inobservables que afectan a la evolución conjunta de las k series consideradas. Asimismo, se supone que el número de factores es inferior al de variables: $r < k$. Los parámetros λ_{ij} cuantifican la sensibilidad del comportamiento inflacionario de la rúbrica i respecto a cambios en el factor j .

$\varepsilon_{i,t}$ denota el componente idiosincrásico del crecimiento interanual de p_{it} esto es, la porción de variabilidad que no obedece a elementos compartidos con las demás rúbricas. μ_i refleja el nivel de $\pi_{i,t}$, una vez tenidos en cuenta los elementos comunes e idiosincrásicos.

Expresando el modelo [1] en notación matricial se obtiene:

$$\Pi_t = \Gamma F_t + E_t$$

donde Π_t : $k \times 1$, Γ : $k \times r$, F_t : $r \times 1$ y E_t : $k \times 1$. La primera característica notable de la ecuación [3] es que todo lo que se encuentra en su parte derecha es inobservable. Naturalmente, esto plantea un problema de equivalencia observacional que ha de ser resuelto incorporando una serie de hipótesis *a priori* acerca de la naturaleza de los elementos involucrados en [1]. Estas son:

$$[4a] \quad E(F_t) = 0$$

$$[4b] \quad E(F_t F_t') = I$$

Esta hipótesis establece la ortogonalidad de los factores comunes, de manera que la información contenida en cualquiera de ellos no está comprendida en alguno de los demás.

restantes factores. Asimismo, en [4a] se establece la nulidad del vector de esperanzas.

$$[5a] \quad E(E_t) = 0$$

$$[5b] \quad E(E_t E_t') = \Sigma_E = \text{diag}(\psi_1 \cdots \psi_r)$$

Con esta condición se asume que cada componente ε_{it} contiene exclusivamente información referida a la parte de π_{it} no explicada por los factores comunes. Se admite que la variabilidad de este componente "residual" sea también específica. Por estas razones, como ya se ha comentado, E_t recibe la denominación de "factor idiosincrásico". Por otra parte, en [5a] se establece la nulidad del vector de esperanzas.

Finalmente, se establece una condición que asegura que los factores comunes y los idiosincrásicos carecen de elementos informativos comunes:

$$[6] \quad E(F_t E_t') = 0$$

La aplicación del modelo factorial estático al vector de variables Π_t exige resolver, de forma secuencial, los siguientes problemas:

¿cuál es el número apropiado de factores, r ?

¿cómo estimar la matriz de cargas Γ a partir de la muestra sobre Π_t ?

¿cómo recuperar los factores F_t condicionados a la estimación de Γ y a la información muestral sobre Π_t ?

El análisis de los autovalores y autovectores de la matriz de correlaciones contemporáneas de Π_t ofrece una vía para resolver (a). El problema de estimación planteado en (b) se resuelve mediante el método de los componentes principales y, por último, la determinación de los valores de los factores (*scoring*) se realiza mediante una estimación por mínimos cuadrados generalizados del modelo [3] en el que se ha sustituido Γ por su estimador obtenido por el método de los componentes principales. Un tratamiento completo de estas cuestiones se encuentra en Mardia *et al.* (1979) y aplicaciones económicas en Litterman y Scheinkman (1988), Knez *et al.* (1994) y Bechikh (1998), entre otros.

3. Datos

Los datos utilizados para examinar la evolución inflacionaria conjunta de los IPC han sido obtenidos de la base de datos Tempus del INE y comprenden sus 57 rúbricas desagregadas durante el período que abarca desde enero de 1987 hasta diciembre de 2000.

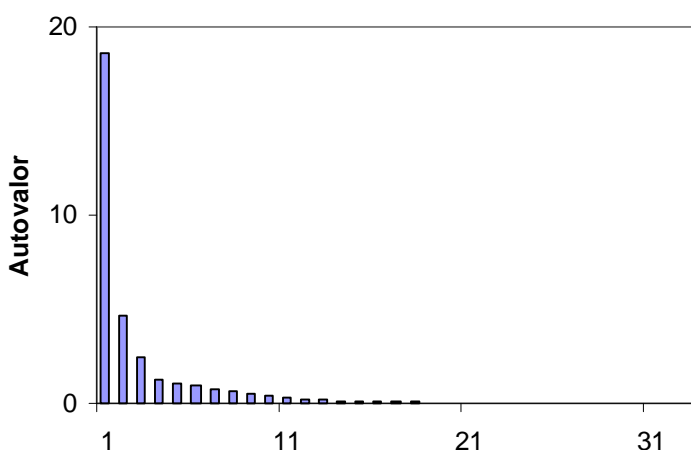
La señal inflacionaria contenida en dichas series ha sido estimada a través de su crecimiento interanual (véase [2]). Con el fin de mantener el nivel de generalidad del análisis multivariante lo más alto posible, no se ha efectuado corrección alguna de (posibles) valores anómalos o efectos de intervención.

La aplicación de la metodología expuesta en la sección anterior al conjunto de las 57 series dio como resultado un modelo de cuatro factores de compleja interpretación. La estructura de cargas y los niveles de comunalidad alcanzados sugirieron la exclusión del análisis de los IPC de los alimentos, las bebidas y el tabaco. De esta manera, el estudio final se centra en las 33 componentes no alimenticios del IPC.

4. Resultados empíricos

En esta sección se presentan los resultados del análisis factorial. La información proporcionada por el gráfico *Scree* y por los autovalores de la matriz de correlaciones conduce a la elección de un modelo de dos factores. El primer factor recoge el 56 por ciento de la varianza y al añadir el segundo factor se explica aproximadamente el 71 por ciento. La introducción de un tercero, que hubiera aumentado el porcentaje de varianza explicada en un 4 por ciento, no se consideró apropiada, tanto por el resultado mostrado por el gráfico *Scree*, como por el hecho de buscar un modelo que sintetice de forma sencilla la información recogida en los 33 componentes del IPC utilizados, lo que recomienda la elección de un modelo con un número de factores reducido.

Gráfico 1: Autovalores de la matriz de correlaciones



La estimación se ha efectuado por el método de los componentes principales, que arroja resultados muy similares a los proporcionados por el de máxima verosimilitud. La tabla 1 muestra las cargas asociadas a los dos factores estimados, el porcentaje de varianza (o comunalidad) de cada componente del IPC explicado por el modelo y el porcentaje de dicha varianza explicada que recoge cada factor. La información presentada corresponde a los factores rotados por el método Varimax, rotación efectuada para facilitar la interpretación del modelo. Aparecen sombreados los valores de las cargas cuyo valor iguala o supera a 0,70. Asimismo, se presentan en cursiva los casos en los que el total de varianza explicada por el modelo es inferior al 34 por ciento.

Tabla 1: Estructura de cargas del modelo factorial

IPC: componentes	Cargas asociadas al:		Varianza Explicada (%)	Porcentaje de varianza explicada por	
	Factor 1	Factor 2		Factor 1	Factor 2
Prendas de vestir de hombre	0,165	0,960	94,8	2,9	97,1
Prendas de vestir de mujer	0,078	0,948	90,6	0,7	99,3
Prendas de vestir de niño y bebé	0,179	0,939	91,4	3,5	96,5
Complementos y reparaciones	0,559	0,755	88,2	35,4	64,6
Calzado de hombre	0,209	0,950	94,7	4,6	95,4
Calzado de mujer	0,245	0,927	92,0	6,5	93,5
Calzado de niño y bebé	0,060	0,963	93,1	0,4	99,6
Reparación de calzado	0,848	0,178	75,0	95,8	4,2
En alquiler	0,810	0,162	68,2	96,2	3,8
Calefacción, alumbrado y distr. Agua	0,795	-0,164	65,9	95,9	4,1
En propiedad	0,670	0,234	50,3	89,2	10,8
Muebles y revestimientos de suelo	0,699	0,597	84,5	57,9	42,1
Textiles y accesorios para el hogar	0,493	0,769	83,5	29,1	70,9
Electrodomésticos	0,542	0,803	93,8	31,3	68,7
Menaje	0,616	0,452	58,4	65,0	35,0
Artículos no duraderos para el hogar	0,369	0,072	14,1	96,3	3,7
Servicios para el hogar	0,768	0,433	77,8	75,9	24,1
Servicios médicos y similares	0,771	0,491	83,5	71,1	28,9
Medicamentos y otros productos farmacéuticos	-0,076	0,363	13,7	4,2	95,8
Transporte personal	0,752	0,131	58,3	97,1	2,9
Transporte público urbano	0,711	0,306	59,9	84,4	15,6
Transporte público interurbano	0,802	0,235	69,9	92,1	7,9
Correo y comunicaciones	0,180	0,452	23,6	13,7	86,3
Objetos recreativos	0,462	0,565	53,2	40,1	59,9
Publicaciones	0,737	0,337	65,8	82,7	17,3
Esparcimiento	0,646	0,651	84,0	49,6	50,4
Educación Infantil y Básica Obligatoria	0,705	-0,286	57,9	85,9	14,1
Enseñanza Secundaria	0,890	0,345	91,1	86,9	13,1
Enseñanza Universitaria	0,582	-0,005	33,9	100,0	0,0
Otros gastos de enseñanza	0,705	0,553	80,3	62,0	38,0
Artículos de uso personal	0,854	0,388	88,0	82,9	17,1
Turismo y hostelería	0,749	0,495	80,5	69,6	30,4
Otros bienes y servicios	0,661	0,564	75,5	57,9	42,1
Varianza explicada por cada factor				56	15

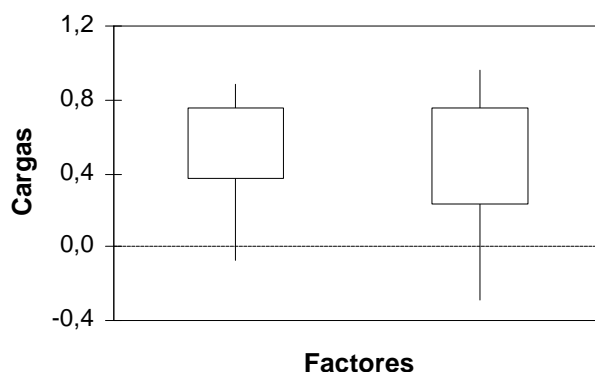
La tabla 2 recoge, para cada uno de los siete grupos de componentes del IPC considerados, el valor medio y mediano de las cargas así como los porcentajes de varianza explicada por cada factor. Los siete grupos considerados para los 33 componentes del IPC son: "Vestido y Calzado", "Vivienda", "Menaje", "Medicina", "Transporte y Comunicación", "Cultura" y "Otros".

Tabla 2: Estructura simplificada de cargas del modelo factorial

		Media y Mediana para cada grupo de:				
		Cargas asociadas al		Varianza	Porcentaje de varianza	
				Explicada	debida al	
		Factor 1	Factor 2		Factor 1	Factor 2
Vestido	Media	0,293	0,828	90,0	18,7	81,3
	Mediana	0,194	0,944	91,7		
Vivienda	Media	0,758	0,077	61,5	93,7	6,3
	Mediana	0,795	0,162	65,9		
Menaje	Media	0,581	0,521	68,7	59,3	40,7
	Mediana	0,579	0,525	80,6		
Medicina	Media	0,347	0,427	48,6	37,7	62,3
Transporte	Media	0,611	0,281	52,9	71,8	28,2
	Mediana	0,732	0,270	59,1		
Cultura	Media	0,675	0,309	66,6	72,4	27,6
	Mediana	0,705	0,345	65,8		
Otros	Media	0,755	0,482	81,4	70,1	29,9
	Mediana	0,749	0,495	80,5		

Adicionalmente, el gráfico 2 presenta un diagrama de caja que recoge la dispersión de las cargas estimadas para cada factor y que puede ayudar en la interpretación de ambos factores.

Gráfico 2: Diagrama de caja de las cargas

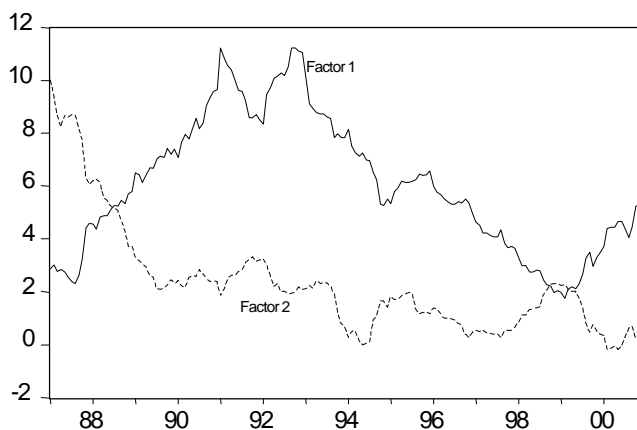


Con la información recogida en las dos tablas y en el gráfico 2, ambos factores se pueden interpretar como un promedio. El primer factor, que recoge la mayor parte de la varianza explicada por el modelo, es una combinación lineal de la mayoría de los componentes, con la clara excepción de casi todos los del grupo de "Vestido y Calzado". Los componentes que forman parte del promedio están asociados, principalmente, a los servicios. El segundo factor, que recoge un porcentaje sensiblemente menor de la varianza explicada, se puede asimilar a un promedio del grupo "Vestido y Calzado" (menos el IPC de "Reparación de calzado") junto con una parte del grupo "Menaje". En este caso, prácticamente todos los componentes que forman parte del promedio son bienes. En ambos casos, se aprecia una dispersión importante en los valores de las cargas asociadas a los distintos componentes, más acusada en el segundo factor.

Por último, hay cuatro componentes para los que las cargas (en valores absolutos) y las comunalidades son bastante reducidas: "Artículos no duraderos para el hogar", "Medicamentos y otros productos farmacéuticos", "Correo y comunicaciones" y "Enseñanza Universitaria", los tres últimos afectados por un grado notable de regulación. Esto puede indicar que el comportamiento de dichos componentes no está muy vinculado con el del resto.

Hay que señalar que no existe correspondencia alguna entre los grupos en los que se distribuyen los componentes del IPC y la estructura de cargas de los dos factores estimados. Es más, en algunos grupos la heterogeneidad entre los componentes, en términos de las cargas asociadas a los mismos, es bastante elevada. El gráfico 3 muestra los dos factores estimados.

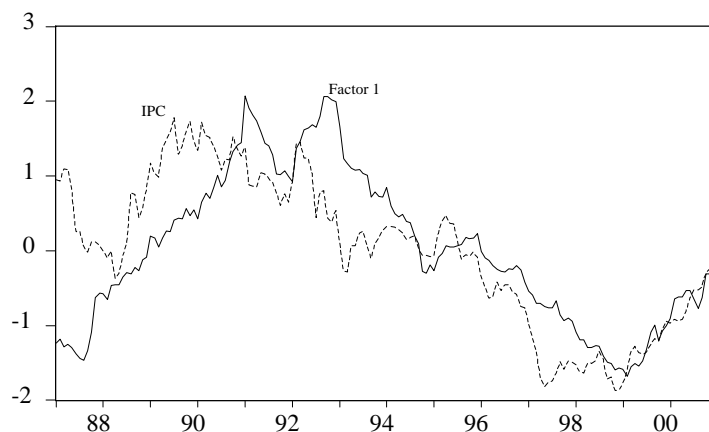
Gráfico 3: Factores estimados



Se aprecia, en primer lugar, el proceso de desinflación seguido por ambos factores aunque con pautas diferentes. Así, el primer factor ha registrado un acusado descenso durante el período 1993-1998 mientras que el segundo lo venía mostrando desde 1987, si bien de forma más moderada. En ambos casos, este perfil descendente se ha visto truncado por eventuales repuntes, siendo el que muestra el primer factor desde 1999 uno de los más notables tanto por su intensidad como por su duración. Asimismo, en la medida en que la pauta de concentración de los precios de los servicios es más acusada en el primer factor y la de los bienes en el segundo, el diferencial entre ambos constituye una medida aproximada del proceso de unificación dual de la economía española. De esta manera, dicho diferencial se ha venido reduciendo desde 1993 como un rasgo adicional del proceso de desinflación, para truncarse abruptamente a partir de 1999, debido a la divergencia creciente entre ambos factores.

El gráfico 4 recoge el factor 1 junto a la tasa interanual del IPC general (ambas series tipificadas).

Gráfico 4: Tasa interanual del IPC general y factor 1

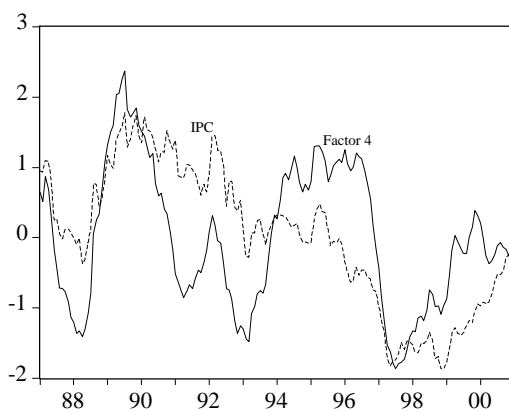


En el gráfico anterior, se observa que el factor 1 parece seguir una pauta similar a la de la tasa interanual del IPC general, aunque algo retrasada. Dado que no se han incluido en el modelo de dos factores presentado los componentes del grupo “Alimentación”, se ha analizado si en el modelo de cuatro factores con el que se inició el estudio, existe algún factor que muestre una evolución coincidente o adelantada con respecto a la de la tasa interanual del IPC general.

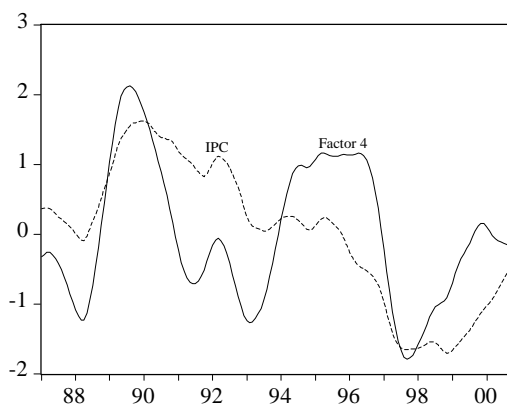
El gráfico 5(a) recoge la tasa interanual del IPC junto al factor 4 de dicho modelo, y el gráfico 5(b) presenta las mismas series suavizadas, donde es más fácil apreciar los puntos de giro. En ambos gráficos, los datos están estandarizados.

Gráfico 5: Tasa interanual del IPC general y factor 4 (modelo con 57 series)

Datos brutos (a) y suavizados (b)
Tasas tipificadas



(a)



(b)

Con el modelo de cuatro factores para los 57 componentes, se recoge un 68 por ciento de la varianza total, estimándose la aportación del cuarto factor en un 7 por ciento. Como se aprecia en el gráfico 5(b), parece que la pauta de comportamiento del factor 4, es similar a la de la tasa interanual del IPC General, mostrando una correlación de 0,54. Sólo ocho componentes están asociados al cuarto factor, y aunque todos pertenecen al grupo “Alimentación”, son bastante heterogéneos: carne de cerdo, huevos, productos lácteos, aceites y grasas, frutas frescas, conservas y preparados de frutas, preparados de hortalizas y bebidas alcohólicas. En consecuencia, un pequeño número de índices elementales de carácter alimenticio resulta muy importante en la determinación de los puntos de giro de la tasa agregada de inflación.

Referencias

Bechikh, Y. (1998) "On deformation of the yield curve", BNP-Paribas, Economic Notes, n. 1998-5.

Knez, J, Litterman, R.B. y Scheinkman, J. (1994) "Explorations into factors explaining money market returns", *Journal of Finance*, vol. 49, n. 5, p. 861-1882.

Litterman, R. y Scheinkman, J. (1988) "Common factors affecting bond returns", Goldman, Sachs & Co., Financial Strategies Group, Technical Report n. 62.

Mardia, K.V., Kent, J.T. y Bibby, J.M. (1979) *Multivariate analysis*, Chapman and Hall, New York, U.S.A.