



fedea

Fundación de
Estudios de
Economía Aplicada

**Una Propuesta para el Seguimiento
Puntual de la Coyuntura de la
Economía de Madrid**

Mario Alloza Frutos

FEDEA

Jesús Fernández-Villaverde

Universidad de Pennsylvania y Fedea

**Colección Estudios Económicos
07-2010**

ISSN 1988-785X
www.fedea.es

Una Propuesta para el Seguimiento Puntual de la Coyuntura de la Economía de Madrid*

Mario Alloza Frutos
Fedea

Jesús Fernández-Villaverde
Universidad de Pennsylvania y Fedea

Junio 2010

*Documento elaborado para la cuarta edición de *Estructura Económica de Madrid*, dirigido por José Luis García Delgado.

1. Introducción

La creciente complejidad de las sociedades modernas acentúa, más que nunca, la necesidad de conocer de manera fiable el estado de la economía en tiempo real. Tanto las Administraciones Públicas como las familias y empresas tienen constantemente que tomar decisiones económicas y, para que estas sean correctas, es fundamental disponer de información precisa acerca de la situación concreta del ciclo económico. Este requerimiento se cumple tanto a nivel nacional (e incluso a niveles más agregados como el del conjunto de la Unión Europea o del mundo) como a niveles más desagregados, en especial el que corresponde en España a las Comunidades Autónomas.

Por ejemplo, los encargados de la política económica de la Comunidad de Madrid (CM) precisan de información sobre la coyuntura para poder evaluar correctamente un amplio abanico de posibles medidas que abarcan desde las consideraciones presupuestarias (como un cambio impositivo o decisiones de gasto) a la introducción de ayudas para el crédito al consumo y la inversión o los planes de infraestructuras públicas. Las empresas con una fuerte implantación en la Comunidad necesitan saber cómo marcha la economía regional para determinar sus decisiones de inversión y de creación de empleo. Las familias pueden aprovechar el conocimiento de la situación del ciclo para decidir acerca de cambios de su puesto de trabajo o de vivienda o, simplemente, para gestionar sus ahorros de manera más fructífera.

Sin embargo, el panorama estadístico español -y regional- ofrece posibilidades limitadas para este objetivo: los datos de Contabilidad Nacional Regional se publican la sexta/séptima semana posterior al final del trimestre de referencia (aunque se ofrece una estimación preliminar del PIB con algunos días de antelación). Este retraso hace que, por ejemplo, si una empresa está evaluando la situación económica a principios de Enero, tenga que esperar hasta mediados de Mayo para conocer el PIB en ese momento. Dado que este retraso es muchas veces excesivo desde el punto de vista de la decisión a tomar, los agentes prestan especial atención a indicadores con una mayor periodicidad en su publicación (también llamados de alta frecuencia). Entre estos indicadores se encuentran los índices mensuales de producción industrial, ventas minoristas, afiliados a la seguridad social o matriculaciones de vehículos, solo por enumerar cuatro casos concretos particularmente populares.

Desafortunadamente, por su misma naturaleza, estos indicadores conllevan dos problemas. Primero, los indicadores de mayor periodicidad sólo arrojan información sobre una parcela concreta de la actividad económica y no ofrecen una imagen general del estado de la economía en su momento de publicación. Segundo, es complejo para los observadores de la coyuntura el valorar y agregar la información de los distintos indicadores sin caer en sesgos como la sobrevaloración del dato más reciente o la autoconfirmación de ideas preconcebidas. Por todo ello, incluso con indicadores de alta frecuencia, el conocimiento real de la actividad económica es limitado, y en la mayoría de los

casos, tardío.

Motivados por estas observaciones, los economistas se han concentrado en los últimos años en la elaboración de índices sintéticos que, con una metodología más formal, recojan y agreguen la información disponible en tiempo real sobre la situación económica de un país o región. Además muchos investigadores han argumentado que una consideración adicional de importancia es que la producción del índice sea lo más automática posible, tanto para facilitar su elaboración sistemática como para convertirlo en un instrumento aplicable al mayor número de situaciones posibles. De esta manera, el índice permitiría ser actualizado de forma contemporánea inmediatamente después de la publicación de un dato que intervenga en su elaboración.

Entre los objetivos buscados es que el índice sintético pueda utilizar información a distinta frecuencia, es decir, que permita incluir tanto la información trimestral existente como la información mensual, semanal o incluso diaria. Por ejemplo, existe información pública detallada y de fácil acceso de índices bursátiles, diferenciales de tipos de interés, o consumo de electricidad diario. Más interesante, las empresas gestoras de tarjetas de crédito y las grandes cadenas comerciales disponen de datos, lamentablemente no de acceso general, sobre el número y cuantía de transacciones realizadas también día a día (una interesante ilustración de cómo utilizar estos datos de altísima frecuencia es el análisis de John Taylor, 2009, de la evolución de las ventas diarias de *Target*, una gran cadena de hipermercados en Estados Unidos en las semanas anteriores y posteriores al colapso de Lehman).

Estas nuevas herramientas metodológicas reciben, en un juego de palabras en inglés que resulta difícil de traducir al castellano, el nombre de *nowcast*, en contraposición con el término de *forecast* o predicción. La idea clave es que cuando se efectúa un *nowcasting*, lo importante es saber dónde está la economía en este momento preciso en contraposición con los ejercicios de *forecasting*, que quieren saber dónde se encontrará la economía en el futuro.

La literatura de índices sintéticos es muy amplia y por consideraciones de espacio este capítulo no puede entrar en una discusión detallada de las distintas alternativas. Sin embargo cabe resaltar trabajos particularmente influyentes como el índice para la economía de Estados Unidos elaborado por Aruoba, Diebold y Scotti (2009), y en el que se basará el índice propuesto en las secciones siguientes, o el índice descrito en Camacho y Perez-Quiros (2009).

En la elaboración de estos índices existen al menos dos decisiones de importancia para el investigador:

1. Determinar la cantidad de teoría económica empleada en la elaboración del índice. Existe una tradición que empieza con Sargent y Sims (1977) que enfatiza la posibilidad de elaborar

índices sintéticos basados, de manera casi exclusiva, en una construcción estadística, por lo que no dependen de ninguna teoría económica concreta. En contraposición, se pueden utilizar modelos mucho más estructurales (como el propuesto en Burriel, Fernández-Villaverde y Rubio-Ramírez, 2010). La ventaja de los enfoques más estadísticos es que son robustos a los errores de especificación introducidos por la teoría económica. La desventaja es que, al tener menos estructura, en general serán menos eficientes en extraer información de los datos y pueden estar sometidos a movimientos en los parámetros que definen el modelo (esto es un ejemplo paradigmático de la tensión entre robustez y eficiencia constantemente presente en econometría).

2. Determinar el número de series temporales empleadas. Una posibilidad es emplear todas las series temporales existentes y dejar que el índice decida qué peso dar a cada una de ellas (Stock y Watson, 2002). Una segunda posibilidad (Aruoba, Diebold y Scotti, 2009) consistente en emplear información a priori para seleccionar un número reducido de series temporales que incluya a las más informativas acerca de la situación cíclica de la economía. La elección entre estas dos alternativas se mueve en unos criterios muy similares a los del párrafo anterior. La selección de series evita tener que estimar un gran número de parámetros, con la consiguiente pérdida de eficiencia que esto acarrea, pero a la vez también expone al investigador a los posibles errores inducidos por una elección incorrecta de las series.

El objetivo de este capítulo es esbozar cómo se puede crear un índice sintético del ciclo aplicado a la Comunidad de Madrid y una muestra preliminar de los resultados que un ejercicio de este estilo produce. Con ello se pretende permitir, con la mayor información estadística disponible hasta la fecha, un juicio sobre la situación actual de la coyuntura económica regional. Con respecto a las dos decisiones anteriormente señaladas, el índice propuesto apuesta por un enfoque puramente estadístico y empleando un número relativamente reducido de series temporales. Si bien estas decisiones requerirían una justificación más detallada que no se puede ofrecer en este capítulo por consideraciones de espacio, el lector obtendrá en todo caso una visión panorámica de las principales ideas detrás de este programa de investigación.

Para ello, en la sección 2 se describen los datos utilizados para la elaboración del índice; en la sección 3 se expone la metodología en la que se basa el indicador; la sección 4 resume los principales resultados obtenidos acerca de la evolución de la economía en la Comunidad de Madrid, mientras que la sección 5 los compara con la economía nacional. En último lugar, la sección 6 concluye el capítulo y señala posibles ampliaciones futuras.

2. Datos utilizados

En esta sección se presentan las variables utilizadas en la elaboración del índice, así como su horizonte temporal. La selección de una adecuada combinación de series es un elemento clave en la elaboración del índice. El objetivo es lograr un balance entre (i) la simplicidad y la omisión de variables redundantes y (ii) la inclusión de datos representativos de todas las parcelas de la economía. Además, esta selección de datos debe realizarse aprovechando una de las ventajas que ofrece la metodología aquí propuesta: la utilización de variables que tienen distinta frecuencias (por ejemplo trimestral, mensual, semanal, etc.).

Para el índice de la Comunidad de Madrid, las variables seleccionadas son:

1. Producto Interior Bruto (PIB), desde 1993-2002 en frecuencia anual y desde el primer trimestre de 2002 en frecuencia trimestral. Fuente: Instituto de Estadística de la Comunidad de Madrid.
2. Clima Industrial (CLI), desde Enero-1995. Fuente: Instituto de Estadística de la Comunidad de Madrid.
3. Matriculaciones de turismos (MT), desde Enero-1977. Fuente: Dirección General de Tráfico.
4. Consumo de Energía Eléctrica (CEE), desde Enero-1995. Fuente: Instituto de Estadística de la Comunidad de Madrid, con datos de Unión Fenosa e Iberdrola.
5. Afiliados a la Seguridad Social, último día de mes (ASS), desde Enero-1982. Fuente: Ministerio de Trabajo e Inmigración.
6. Índice de Comercio Minorista (ICM), desde Enero-2003. Fuente: Instituto Nacional de Estadística.
7. Índice de Producción Industrial (IPI), desde Enero-1993. Fuente: Instituto Nacional de Estadística.

Si bien se considera que las variables previamente señaladas son suficientemente representativas de la economía madrileña, puede ser objeto de discusión en futuras ampliaciones la inclusión de variables referentes al sector construcción (como puede ser el número de viviendas acabadas) y al sector exterior (el volumen de importaciones y exportaciones).

Todas las variables seleccionadas son desestacionalizadas (si no han sufrido este ajuste en la fuente de origen) y se introducen en el modelo en forma de tasa de variación logarítmica interanual

(*TVLI*), es decir:

$$TVLI = \log(y_t^i) - \log(y_{t-12}^i)$$

dónde y_t es la variable i observada en el momento t (referido a meses). En la figura 1 se muestra la evolución de todas las variables utilizadas en la elaboración del índice (transformadas en *TVLI*). Con la excepción del Consumo de Energía Eléctrica (serie de la que sólo se dispone de datos hasta diciembre de 2009), todas las variables están actualizadas hasta el mes de abril de 2010 (en el caso del PIB hasta el primer trimestre de 2010).

[Figura 1 aquí]

Una rápida inspección de los datos en la figura 1 revela el patrón principal de la evolución de la economía madrileña desde 1996, en especial la larga expansión de 1996 al 2007 (sólo con una cierta desaceleración en el 2001-2002 coincidente con la recesión de otras economías occidentales que España esquivó en gran medida), la profunda crisis del 2008-2009 y una cierta recuperación en el 2010.

3. Metodología

La metodología utilizada para la elaboración del índice se basa en técnicas estadísticas estándar de modelos de factores dinámicos (más información en el apartado dedicado a la Bibliografía). La idea es extraer de las series observadas uno o varios factores comunes que capturen el comovimiento de las series. En este caso, se asumirá que sólo existe un factor X_t que define las condiciones económicas de la Comunidad de Madrid que, si bien no puede observarse, está relacionado con el resto de variables (los indicadores descritos en la sección 2). Una de las principales ventajas de la definición de una variable latente es la facilidad de formalizar el modelo de interrelación entre ésta y el resto de variables mediante una representación en el espacio de los estados y la consiguiente posibilidad de estimar los parámetros que componen este modelo mediante la evolución de la verosimilitud que produce el filtro de Kalman.

El filtro de Kalman también permite:

1. Incorporar de manera sencilla indicadores económicos que están expresados en distinta frecuencia. Así, el PIB se conoce de forma trimestral mientras que el resto de indicadores tiene una publicación mensual.
2. Tratar algunas irregularidades en los datos como observaciones perdidas o variables flujo que muestran distinta frecuencia a lo largo del periodo analizado. Es este el caso del PIB para la

Comunidad de Madrid: durante 1993 y hasta 2002 los datos disponibles son de carácter anual, y desde este punto, se conocen con periodicidad trimestral. Para tratar este problema (y en general, el hecho de que el modelo contenga variables flujo), se crean variables que acumulan dichos flujos en los puntos no observados y las hace comparables con las variables stock.

Formalmente, el modelo econométrico se puede representar en forma de un sistema en el espacio de estados:

$$\tilde{X}_t = A_t \tilde{X}_{t-1} + B e_t \quad (1)$$

$$Y_t = C \tilde{X}_t + \Gamma w_t + u_t^* \quad (2)$$

dónde (1) se denomina ecuación de transición y (2) ecuación de medida.

La ecuación de transición (1) especifica cómo el vector de estados del sistema, $\tilde{X}_t$, evoluciona a lo largo del tiempo. Como se explicará a continuación, el factor o variable latente X_t , que es el índice sintético que resume el estado de la economía en el periodo t , es una de las variables que entra en $\tilde{X}_t$. Esta ecuación de transición es un proceso autorregresivo de orden 1 con matrices A_t (indiciada por el periodo t para permitir la incorporación de observaciones a distintas frecuencias), B y una innovación $e_t \sim \mathcal{N}(0, 1)$ que no muestra correlación serial.

Antes de continuar es importante resaltar que la restricción de un proceso autoregresivo de orden 1 es únicamente por conveniencia notacional. Si, por ejemplo, el modelo econométrico implicara que la evolución de $\tilde{X}_t$ se describe mejor por un proceso de orden 2, simplemente se puede reescribir el estado como:

$$\hat{X}_t = \begin{pmatrix} \tilde{X}_t \\ \tilde{X}_{t-1} \end{pmatrix}$$

y el proceso así transformado será autoregresivo de orden 1 en $\hat{X}_t$. Una transformación similar de los estados del sistema se puede realizar con los componentes de medias móviles o con la correlación serial de la innovación. Por tanto, (1) es una especificación a la vez sencilla y general que ahorra en notación.

La ecuación de medida (2) relaciona el estado $\tilde{X}_t$ con las variables observadas Y_t dado, potencialmente, covariables w_t que afectan al observable pero no al estado (por ejemplo, si una de las variables observadas es la electricidad, w_t puede ser la temperatura media del mes) y un error $u_t^* \sim \mathcal{N}(0, \Sigma)$ que puede corresponder tanto a un error de medida como a desviaciones aleatorias de las variables con respecto al indicador. En este capítulo, se especificará que w_t es una variable que contiene un retardo de Y_t , para capturar así la persistencia en las variables observables que no puede ser completamente explicada por la persistencia en el factor. Además, se impondrá que la matriz de varianzas-covarianzas Σ es diagonal, con entrada genérica σ_i en la fila correspondiente a la variable i .

La construcción de la ecuación de transición (1) requiere cierto cuidado ya que, además de incluir el factor X_t , es preciso introducir dos variables auxiliares llamadas acumuladores que permiten combinar las observaciones a distintas frecuencias. El nombre de acumulador viene de la observación que estas variables “acumulan” el valor del factor en cada periodo en el que no existen nuevas observaciones.

En concreto (1) tiene el siguiente aspecto:

$$\underbrace{\begin{pmatrix} X_t \\ C_t^Y \\ C_t^Q \end{pmatrix}}_{\tilde{X}_t} = \underbrace{\begin{pmatrix} \rho & 0 & 0 \\ \rho & \xi_t^Y & 0 \\ \rho & 0 & \xi_t^Q \end{pmatrix}}_{A_t} \underbrace{\begin{pmatrix} X_{t-1} \\ C_{t-1}^Y \\ C_{t-1}^Q \end{pmatrix}}_{\tilde{X}_{t-1}} + \underbrace{\begin{pmatrix} \sqrt{1-\rho^2} \\ \sqrt{1-\rho^2} \\ \sqrt{1-\rho^2} \end{pmatrix}}_B e_t$$

La primera línea es la especificación de la evolución del factor:

$$X_t = \rho X_{t-1} + \left(\sqrt{1-\rho^2} \right) e_t$$

que es un simple proceso autorregresivo de orden 1. La elección de esta especificación está motivada desde la óptica de que un autorregresivo de orden 1 es una especificación natural para la mayoría de las series temporales. Sin embargo, nada en la metodología propuesta en este capítulo depende de esta especificación. Con cambios menores en la notación, se pueden especificar cualquier modelo $ARMA(m, n)$ para la evolución de X_t . Como se explicará en unos párrafos, el índice se va a estimar mediante métodos basados en la función de verosimilitud. Por tanto, se pueden emplear métodos como los ratios de verosimilitud, el criterio de información Bayesiana o la verosimilitud marginal para determinar m y n .

La especificación elegida implica que el factor tiene media cero. Esto es importante cuando, posteriormente en la sección de resultados, haya que interpretar los valores del índice. Un valor de X_t de cero denota que la economía de la región está creciendo a su media histórica, que es obviamente superior a cero. Esto es, por supuesto, nada más que una normalización que permite eliminar el parámetro que de otra manera controlaría la media de X_t . Una segunda normalización se produce al escalar las innovaciones por $\sqrt{1-\rho^2}$. De esta manera, el índice tiene varianza 1 y sus valores numéricos son de fácil interpretación con la distribución normal.

Otra consideración es que la persistencia del factor, ρ , potencialmente puede cambiar a lo largo del tiempo (por ejemplo, ya que los servicios suelen ser menos volátiles que la industria, un cambio en la estructura productiva de una región de la industria a los servicios incrementará la persistencia del factor). En el índice propuesto en este trabajo, y para simplificar la exposición, se omitirá esta posibilidad pero esta consideración sería una interesante extensión en el medio plazo.

En la segunda línea aparece C_t^Y , una variable que acumula los valores del factor X_t , expresado

en frecuencia mensual, para construir una variable agregada de flujo con frecuencia anual (los datos del PIB entre 1993 y 2002 tienen frecuencia anual), tal que

$$C_t^Y = \xi_t^Y C_t^Y + X_{t-1}$$

dónde:

$$\xi_t^Y = \begin{cases} 0 & \text{si es el primer mes del año en curso} \\ 1 & \text{en cualquier otro caso} \end{cases}$$

De manera análoga, en la tercera línea, C_t^Q acumula los valores del factor para construir una variable flujo agregada que tenga en cuenta la publicación trimestral del PIB desde el primer trimestre de 2003. Así,

$$C_t^Q = \xi_t^Q C_t^Q + X_{t-1}$$

dónde:

$$\xi_t^Q = \begin{cases} 0 & \text{si es el primer mes del trimestre en curso} \\ 1 & \text{en cualquier otro caso} \end{cases}$$

La ecuación de medida (2) queda por tanto configurada como

$$\underbrace{\begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \\ y_t^4 \\ y_t^5 \\ y_t^6 \\ y_t^7 \end{pmatrix}}_{Y_t} = \underbrace{\begin{pmatrix} 0 & \beta_1 & \beta_1 \\ \beta_2 & 0 & 0 \\ \beta_3 & 0 & 0 \\ \beta_4 & 0 & 0 \\ \beta_5 & 0 & 0 \\ \beta_6 & 0 & 0 \\ \beta_7 & 0 & 0 \end{pmatrix}}_C \underbrace{\begin{pmatrix} X_t \\ C_t^Y \\ C_t^Q \end{pmatrix}}_{\tilde{X}_t} +$$

$$\underbrace{\begin{pmatrix} 0 & \dots & 0 & \gamma_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & \gamma_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & \gamma_3 & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & \gamma_4 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & \gamma_5 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & \gamma_6 & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \gamma_7 \end{pmatrix}}_{\Gamma} \underbrace{\begin{pmatrix} 1 \\ \dots \\ 1 \\ y_{t-1}^1 \\ y_{t-1}^2 \\ y_{t-1}^3 \\ y_{t-1}^4 \\ y_{t-1}^5 \\ y_{t-1}^6 \\ y_{t-1}^7 \end{pmatrix}}_{w_t} + \begin{pmatrix} u_t^{*1} \\ u_t^{*2} \\ u_t^{*3} \\ u_t^{*4} \\ u_t^{*5} \\ u_t^{*6} \\ u_t^{*7} \end{pmatrix}$$

dónde los superíndices de las variables $(1, 2, \dots, 7)$ se corresponden con el orden en el que se presentaron las variables en la sección anterior.

Por consiguiente los parámetros que definen esta representación son:

$$\theta = \{\rho, \beta_1, \dots, \beta_7, \gamma_1, \dots, \gamma_7, \sigma_1, \dots, \sigma_7\}$$

El tratamiento del modelo en el espacio de estados fijados unos valores para θ se realiza por medio del filtro de Kalman. Dada la estructura lineal del modelo y la distribución normal de todas las innovaciones, todas las distribuciones condicionales del estado $\tilde{X}$ son también normales y por tanto se pueden describir de manera completa con la media $\tilde{X}_{t|j}$ y matriz de varianza condicional $P_{t|j}$ en el periodo t . Todos los condicionamientos son con respecto al conjunto de información disponible hasta el periodo j .

La evolución de estos momentos condicionales viene descrita, cuando se observa un nuevo dato, por las siguientes ecuaciones:

$$\begin{aligned}\tilde{X}_{t|t} &= \tilde{X}_{t|t-1} + P_{t|t-1} C' F_t^{-1} v_t \\ P_{t|t} &= P_{t|t-1} - P_{t|t-1} C' F_t^{-1} C P_{t|t-1} \\ \tilde{X}_{t+1|t} &= A_{t+1} \tilde{X}_{t|t} \\ P_{t+1|t} &= A_{t+1} P_{t|t} A_{t+1}' + B B'\end{aligned}$$

dónde:

$$v_t = y_t - C \tilde{X}_{t|t-1} - \Gamma w_t$$

es un vector 7×1 de ruido blanco generado por la diferencia entre la observación, y_t , y la predicción, $C \tilde{X}_{t|t-1} + \Gamma w_t$, mientras que:

$$F_t = C P_{t|t-1} C' + H$$

dados valores iniciales $\tilde{X}_{0|0}$ y $P_{0|0}$ (que son fáciles o bien de estimar como parámetros adicionales del problema o fijar a los valores del estado estacionario implicado por el modelo estadístico, que es el procedimiento seguido en este capítulo).

En los casos en los que no se observa un nuevo dato, sólo se obtiene una predicción, pero no actualización, por lo que las anteriores ecuaciones se reducen a:

$$\begin{aligned}\tilde{X}_{t|t} &= \tilde{X}_{t|t-1} \\ P_{t|t} &= P_{t|t-1} \\ \tilde{X}_{t+1|t} &= A_{t+1} \tilde{X}_{t|t} \\ P_{t+1|t} &= A_{t+1} P_{t|t} A_{t+1}' + B B'\end{aligned}$$

Dado que v_t es ruido blanco, se puede emplear la función de densidad de la distribución normal y el output del filtro de Kalman para evaluar la función de verosimilitud de la secuencia de observables condicionado en unos valores para los parámetros θ :

$$\log p\left(\{y_t\}_{t=1}^T \mid \theta\right) = -\frac{1}{2} \sum_{t=1}^T (7 \log 2\pi + \log |F_t| + v_t' F_t^{-1} v_t)$$

donde T es el total de observaciones.

4. Resultados

Para estimar los parámetros del modelo se acude a técnicas bayesianas. Éstas permiten incorporar información premuestral en los valores de los parámetros por medio de distribuciones a priori, lo que resulta particularmente útil en el índice propuesto por la corta longitud de los datos existentes. En concreto, la distribución a posteriori de los parámetros, que recoge la incertidumbre existente sobre sus valores una vez dadas las observaciones viene dada por el teorema de Bayes:

$$p\left(\theta \mid \{y_t\}_{t=1}^T\right) = \frac{p\left(\{y_t\}_{t=1}^T \mid \theta\right) p(\theta)}{\int p\left(\{y_t\}_{t=1}^T \mid \theta\right) p(\theta) d\theta}$$

donde $p(\theta)$ es la distribución a priori de los parámetros y $p\left(\{y_t\}_{t=1}^T \mid \theta\right)$ es la función de verosimilitud encontrada en la sección anterior. Para todos los parámetros excepto ρ se utiliza una distribución a priori uniforme para los que reduce el impacto de la información premuestral en la inferencia y que facilita la comparación de los resultados bayesianos con una estimación frecuentista. El soporte de la uniforme se acota en el caso de los parámetros γ_i en $[0, 1]$ para garantizar estacionariedad y en los números positivos para las σ_i . Finalmente, para ρ se emplea una Beta(18,2) que centra la estimación en las zonas altas de la persistencia.

Desafortunadamente, la distribución de $p\left(\theta \mid \{y_t\}_{t=1}^T\right)$ no pertenece, en general, a ninguna familia de distribuciones conocida. Por tanto, ha de ser evaluada numéricamente. Para ello, se utiliza el algoritmo de *Metropolis-Hastings*, que pertenece al conjunto de métodos conocidos como *Monte Carlos por Cadenas de Markov*. El algoritmo de *Metropolis-Hastings* permite extraer muestras de la distribución $p\left(\theta \mid \{y_t\}_{t=1}^T\right)$ sin más requerimiento que ser capaz de evaluar numéricamente $p\left(\{y_t\}_{t=1}^T \mid \theta\right) p(\theta)$. Por una simple aplicación de un teorema en la clase de Glivenko-Cantelli, la distribución empírica de n iteraciones del algoritmo, $p_n\left(\theta \mid \{y_t\}_{t=1}^T\right)$ convergerá de manera casi segura a la distribución $p\left(\{y_t\}_{t=1}^T \mid \theta\right) p(\theta)$ cuando $n \rightarrow \infty$ y, con ella, todos sus momentos.

La generación con este procedimiento de la distribución a posteriori permite la descripción de toda la incertidumbre existente con respecto a los parámetros pero también, en conjunción con una

función de pérdida del investigador (es decir, el coste derivado de las diferencias entre el parámetro estimado y su valor real), la estimación puntual de θ y la elaboración de conjuntos de probabilidad. En el índice propuesto la estimación puntual reportada será la que maximiza la verosimilitud para facilitar la labor del lector en su interpretación (una interesante característica del algoritmo de *Metropolis-Hastings* es que la distribución empírica que genera se puede emplear también para maximizar la verosimilitud).

En la estimación del índice sintético se itera el algoritmo, tras un largo periodo de rodaje de la cadena de Markov para eliminar la dependencia de las condiciones iniciales, 50.000 veces. En la figura 2 se muestra la evolución del valor de la función de verosimilitud a lo largo de estas últimas 50.000 simulaciones.

[Figura 2 aquí]

Por su parte, las figuras 3 y 4 recogen la evolución de ρ (el parámetro asociado al primer retardo de la variable latente o factor) y del resto de parámetros, respectivamente. Estas figuras muestran que la cadena se comporta de manera suficientemente estacionaria, con lo cual la convergencia numérica es altamente probable.

[Figura 3 aquí]

[Figura 4 aquí]

Las figuras 5 y 6 muestran la distribución a posteriori de ρ y del resto de parámetros, respectivamente. En el caso de ρ , la figura 5 ilustra que la variable latente tiene una alta persistencia (el parámetro ρ se concentra en valores en torno a 0,92–0,99; el valor de ρ en el que la función de verosimilitud es máxima es 0,958).

[Figura 5 aquí]

[Figura 6 aquí]

Una vez obtenidas las estimaciones de los parámetros del modelo, se utiliza de nuevo el alisado de Kalman para generar el factor $X_{t|T}$, es decir, la estimación que maximiza la verosimilitud acerca de la media del índice de actividad económica en la Comunidad de Madrid condicionando en toda la muestra y no sólo en las observaciones anteriores al periodo t . Los resultados de este índice se muestran en la figura 7, acompañado por una estimación de la tendencia del mismo (la tendencia

se estima utilizando el filtro de Hodrick-Prescott con un coeficiente de alisamiento, λ , de 2). En esta figura se puede observar que el índice recoge los ciclos económicos de la economía madrileña, mostrando las dos etapas expansivas en 1997–2001 y 2003–2007. Sin embargo, el hecho que más llama la atención es la magnitud de la actual recesión (desde finales de 2007), que relativiza la desaceleración del crecimiento registrado entre 2001–2003.

[Figura 7 aquí]

Conviene recordar que el índice, por construcción, está normalizado con media 0 y desviación típica 1. Así, valores del índice por debajo de 0 pueden interpretarse como desarrollos económicos por debajo de la media histórica (en nuestro caso, desde 1996 hasta la actualidad). Por el contrario, cuando el índice toma valores por encima de 1 se refiere a una evolución de la actividad económica por encima de la media. Cuando el índice registra un nivel de 0, la economía crece a su ritmo medio. El índice también permite comparaciones entre distintos puntos, por ejemplo, se puede hablar de que la actividad económica madrileña era sustancialmente menor entre 2008–2009 en comparación con 2001–2002.

¿Qué puede decir el índice sobre la evolución de la actividad económica en la Comunidad de Madrid en los últimos años? Teniendo en cuenta el periodo de tiempo considerado (desde enero de 1996 hasta la actualidad), el índice marcó un máximo global en abril de 2000, cuando registró un valor de 0,95 puntos. Dada la serie histórica del indicador, podemos decir que este evento tiene una probabilidad de ocurrencia cercana al 17 %. Por el contrario, el mínimo histórico ocurrió en mayo de 2009, cuando el índice se situó en $-2,99$ puntos, es decir, se trata de una situación que dista por tres desviaciones típicas respecto a los desarrollos económicos medios; esto es equivalente a decir, que según la serie histórica observada, un evento de esta magnitud ocurre en menos del 1 % de los casos. Este dato es especialmente ilustrativo si se compara con el suelo alcanzado en la recesión de 2001-2003, dónde el valor mínimo registrado ($-0,40$ puntos en mayo de 2002) puede ocurrir con el 35 % de probabilidad, según la serie histórica analizada.

¿Qué interpretación ofrece el índice sobre el actual periodo de recesión? Observando de nuevo la figura 7, se comprueba que el índice alcanzó un máximo local en febrero de 2007 (0,49 puntos). En ese momento el índice comienza una tendencia decreciente. En octubre del mismo año, la actividad empezó a crecer a un ritmo inferior a la media al situarse en valores inferiores a 0. Después de sucesivos meses de caída el índice alcanzó un suelo en mayo de 2009 (en $-2,99$ puntos). Desde entonces, la economía ha suavizado rápidamente la caída y se ha ido acercando a la media histórica. Recientemente, en abril de 2010, el índice alcanzaba un nivel aún negativo de $-0,35$ puntos.

Como se ha comentado anteriormente, el índice pretende ser una representación numérica de lo

que se ha definido como variable latente en la economía. ¿Qué grado de relación existe entre esta variable latente –es decir, el indicador– y la actividad económica tal y como se mide estadísticamente por medio del PIB (que es, no hay que olvidar, también una medida imperfecta de la actividad)?. Para realizar esta comparación, la figura 8 representa la evolución conjunta del índice en medias trimestrales y el PIB normalizado (restando su media y dividiendo por su desviación típica). Como se puede apreciar en esta figura, existe un alto grado de similitud entre ambas series. El índice propuesto es capaz de identificar los ciclos de expansión y recesión registrados por el PIB. Desde 2003 (fecha en la que se dispone de datos trimestrales para el PIB) y hasta la actualidad, el coeficiente de correlación entre ambas variables alcanza el 92 %.

[Figura 8 aquí]

5. Madrid y España, una perspectiva comparada

En esta sección se realiza una comparación entre la actividad económica de Madrid y la media española. Para llevar a cabo esta tarea se utiliza, por un lado, el índice propuesto en este artículo como representación de la coyuntura económica madrileña. Por otro lado, como indicador de la actividad en España, se utiliza el Índice Fedea de Actividad Económica (disponible en <http://crisis09.es/indice>). Ambos indicadores comparten un marco metodológico muy similar, lo que su facilita su comparación.

[Figura 9 aquí]

El primer hecho que salta a la vista es el grado de similitud entre ambas variables: tanto la economía madrileña como la española muestran similares periodos de recesión y expansión económica (el coeficiente de correlación entre ambas variables es de más del 96 %). Respecto a la coyuntura actual, la caída de la actividad registrada en 2008-2009 es comparable entre ambas economías, aunque ligeramente más aguda en el caso del total de España.

6. Conclusiones

Ante la escasez de herramientas que permitan un seguimiento cercano de la actividad económica en la Comunidad de Madrid, en este artículo se propone una alternativa que permite medir la coyuntura de la economía madrileña con una mayor periodicidad. El indicador de actividad económica aquí presentado se enmarca dentro de la metodología denominada *nowcasting* y presenta ciertas

ventajas técnicas en cuanto a la especificación del modelo y las variables incorporadas. Así, además de ofrecer la posibilidad de combinar indicadores expresados en distinta frecuencia, permite ser actualizado en cuanto se publica alguno de los datos que lo integran.

El análisis del índice propuesto muestra con claridad los ciclos de expansión y recesión de la economía madrileña. Concretamente, la evolución del índice es similar a la del PIB de la Comunidad de Madrid; cabe destacar que la correlación entre ambas variables es del 92 %.

Una de las ventajas que ofrece esta propuesta, consiste en que su estructura está abierta a futuras modificaciones o ampliaciones que supongan una mejora respecto a su construcción original. En primer lugar, y dada la estructura económica de la Comunidad de Madrid, podría considerarse la incorporación de indicadores que enfatizan el sector exterior y de la construcción.

Otra mejora, siempre que el catálogo estadístico así lo permita, puede ser la inclusión de variables de muy alta frecuencia, por ejemplo, de publicación diaria (como puede ser el consumo de energía eléctrica, la matriculación de turismos, o algunas variables financieras). Este último aspecto puede tener un gran interés en los casos en los que se desee utilizar el índice como una herramienta que permita tomar decisiones sobre política económica en tiempo real, es decir, cuando se necesita un juicio preciso y contemporáneo de la actividad económica en Madrid.

Por último, pueden considerarse mejoras relacionadas con el marco metodológico en el que se basa el índice. De esta manera, podría considerarse introducir una mayor dinámica en el caso de la variable latente (suponiendo, por ejemplo, que esta sigue un proceso $AR(2)$ en vez del $AR(1)$ propuesto).

Todas estas extensiones ilustran el potencial de los instrumentos presentados en este capítulo para todos aquellos agentes interesados en un seguimiento detallado de la actividad económica de la Comunidad de Madrid.

Orientación Bibliográfica

El planteamiento seguido en este artículo se basa principalmente en los desarrollos contenidos en:

- Aruoba, S. Borağan, Diebold, Francis X. y Scotti, Chiara (2009). “Real-Time Measurement of Business Conditions,” *Journal of Business and Economic Statistics* 27, 417-427.

Otras fuentes que utilizan una metodología similar (un modelo de factores dinámicos con la presencia de una variable latente) o se basan en aspectos relacionados:

- Camacho, Máximo y Perez-Quiros, Gabriel (2008). “Introducing the EURO-STING: Short Term Indicator of Euro Area Growth,” *Banco de España Working Papers* 0807, Banco de España.
- Stock, James H. y Watson, Mark W. (1989). “New Indexes of Coincident and Leading Economic Indicators,” en *NBER Macroeconomics Annual 1989*, Volume 4, 351-409. National Bureau of Economic Research.
- Stock, James H. y Watson, Mark W. (1991). “A Probability Model of the Coincident Economic Indicators.” En K. Lahiri y G. Moore (Editores), *Leading Economic Indicators: New Approaches and Forecasting Records*. Cambridge, Cambridge University Press, 63-89.
- Stock, James H. y Watson, Mark W. (1991). “Macroeconomic Forecasting Using Diffusion Indexes,” *Journal of Business and Economic Statistics* 20, 147-162.

Una introducción a los métodos bayesianos y a los Métodos de Monte Carlo por Cadenas de Markov:

- Robert, Christian (2001). *The Bayesian Choice: from Decision-Theoretic Motivations to Computational Implementation*. Nueva York, Springer-Verlag, segunda edición.

El lector puede referirse a la siguiente fuente dónde se tratan posibles soluciones ante problemas de “irregularidades” en los datos (por ejemplo, la necesidad de crear una variable que acumule los flujos de una serie observada con distinta frecuencia):

- Harvey, Andrew C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Para el caso de un indicador de la actividad económica en España enmarcado en esta tipo de metodología:

- Camacho, Máximo y Perez-Quiros, Gabriel (2009). “Ñ-STING: España Short Term Indicator of Growth,” *Banco de España Working Papers* 0912, Banco de España.
- Fernandez-Villaverde, Jesús y Rubio-Ramírez, Juan F. (2010). “Índice Fedea de Actividad Económica.” Manuscrito, Fedea. <http://crisis09.es/indice>

Otros trabajos citados en este capítulo incluyen:

- Burriel, Pablo, Fernandez-Villaverde, Jesús y Rubio-Ramírez, Juan F. (2010). “MEDEA: A DSGE Model for the Spanish Economy,” *SERIES: Journal of the Spanish Economic Association* 1, 175-243.
- Sargent, Thomas J. y Sims, Chris .A. (1977). “Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory,” en Chris Sims (editor), *New Methods in Business Research*. Minneapolis: Federal Reserve Bank of Minneapolis.
- Taylor, John (2009). “Analysis of Daily Sales Data during the Financial Panic of 2008.” Manuscrito, Stanford University.

Figura 1: Evolución de los indicadores en TVLI

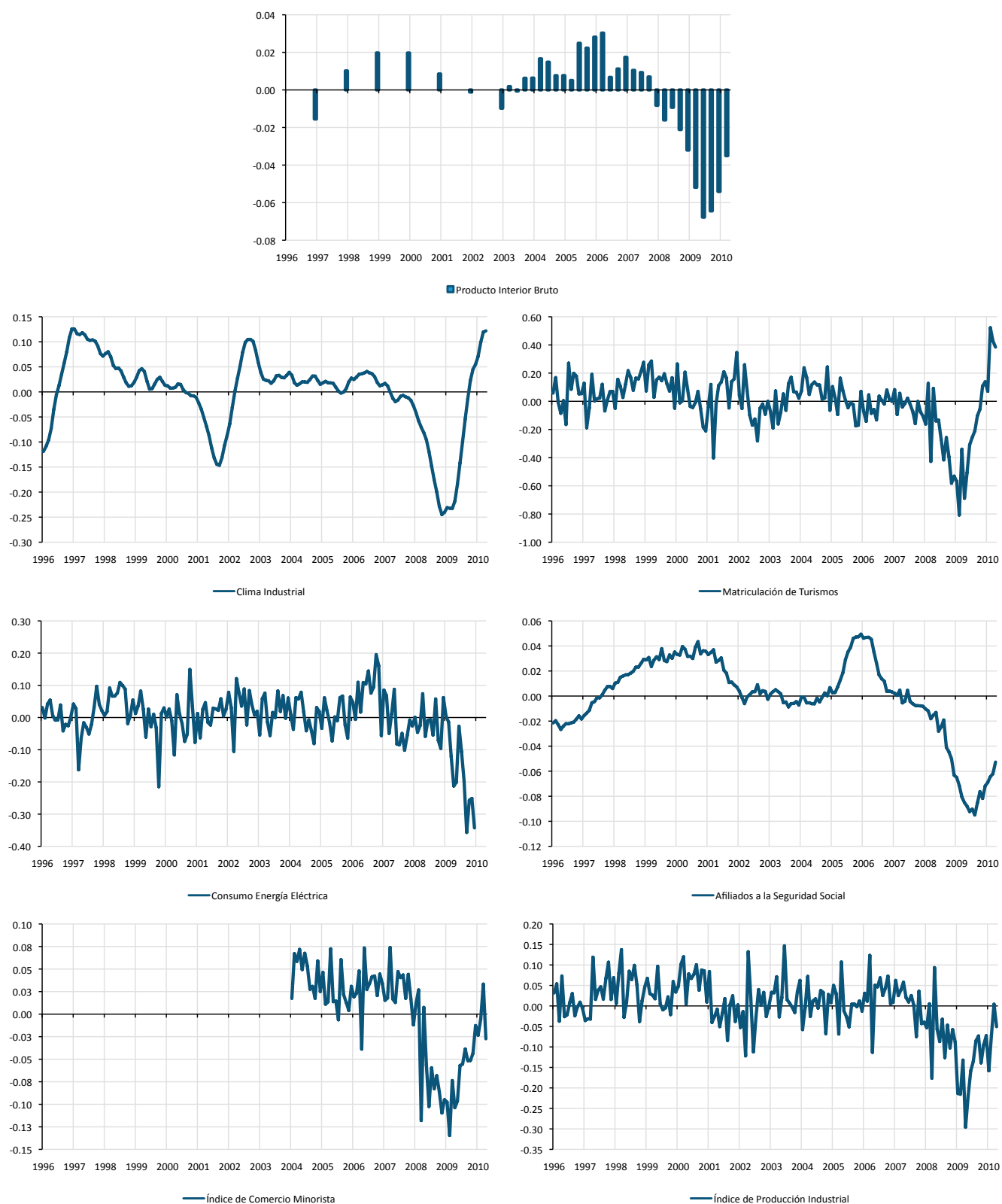


Figura 2: Evolución de la función de verosimilitud

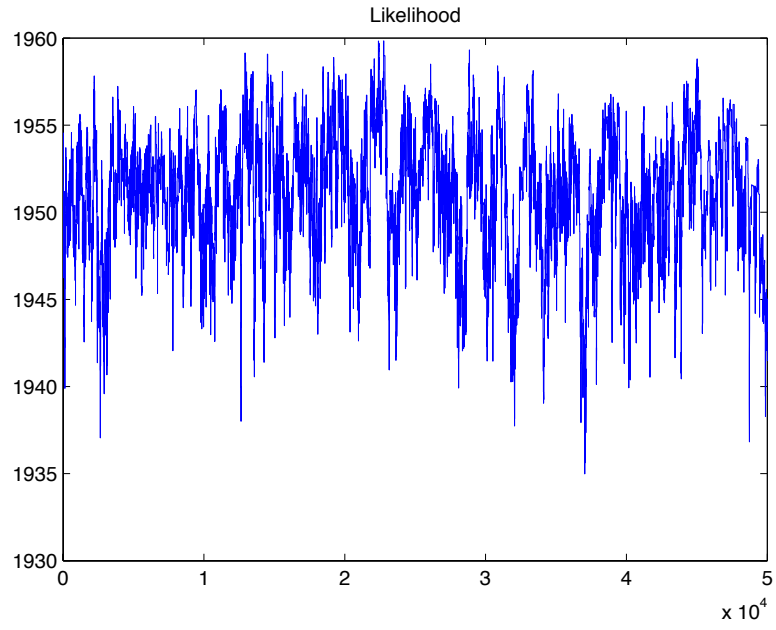


Figura 3: Evolución del valor del parámetro ρ

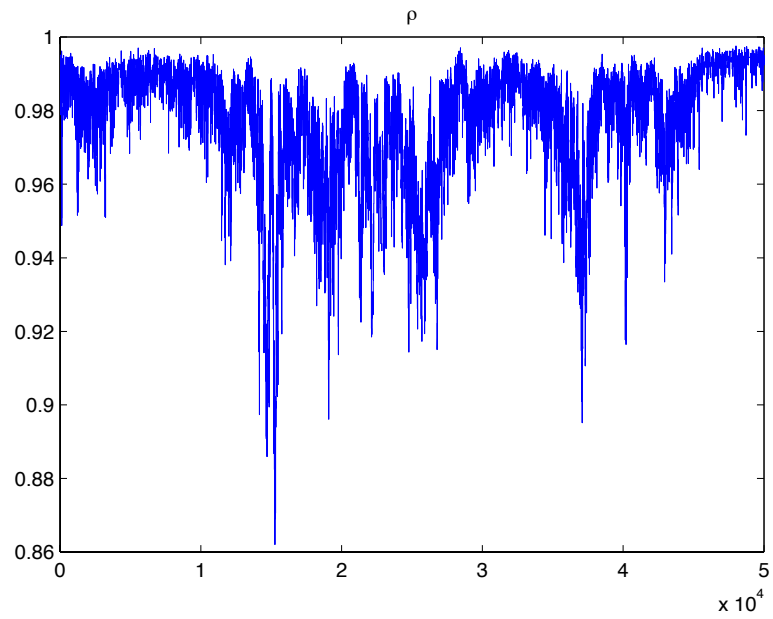


Figura 4: Evolución del resto de parámetros

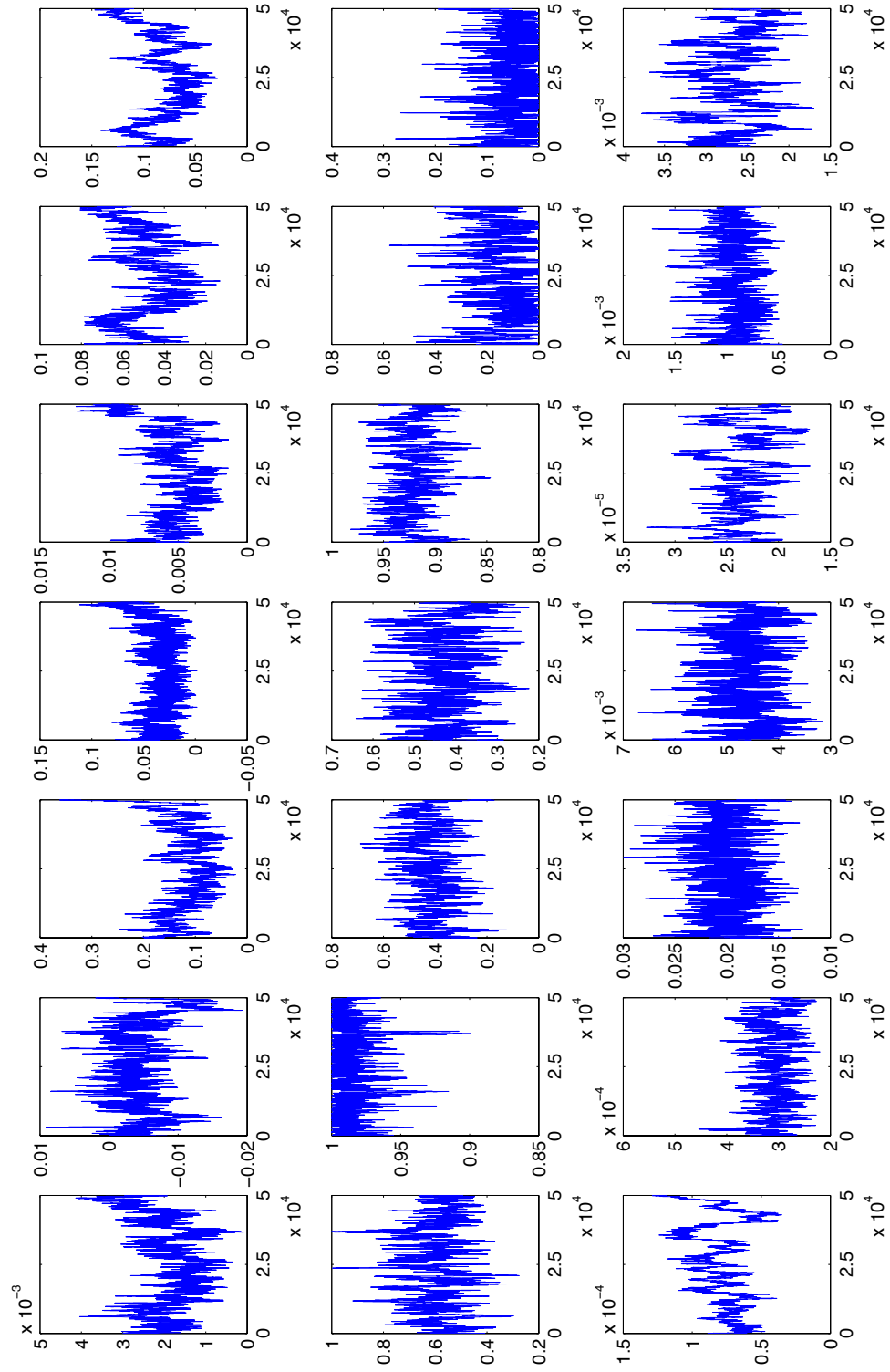


Figura 5: Histograma de ρ

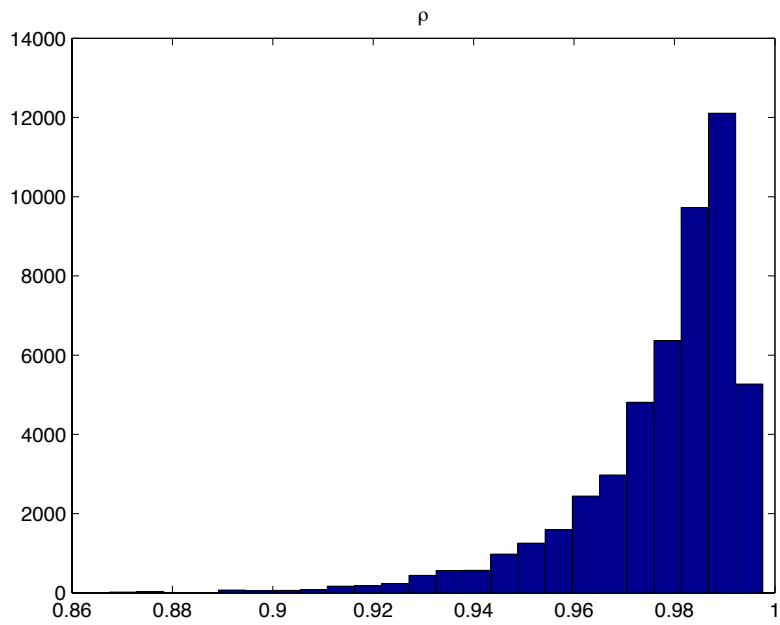


Figura 6: Histograma del resto de parámetros

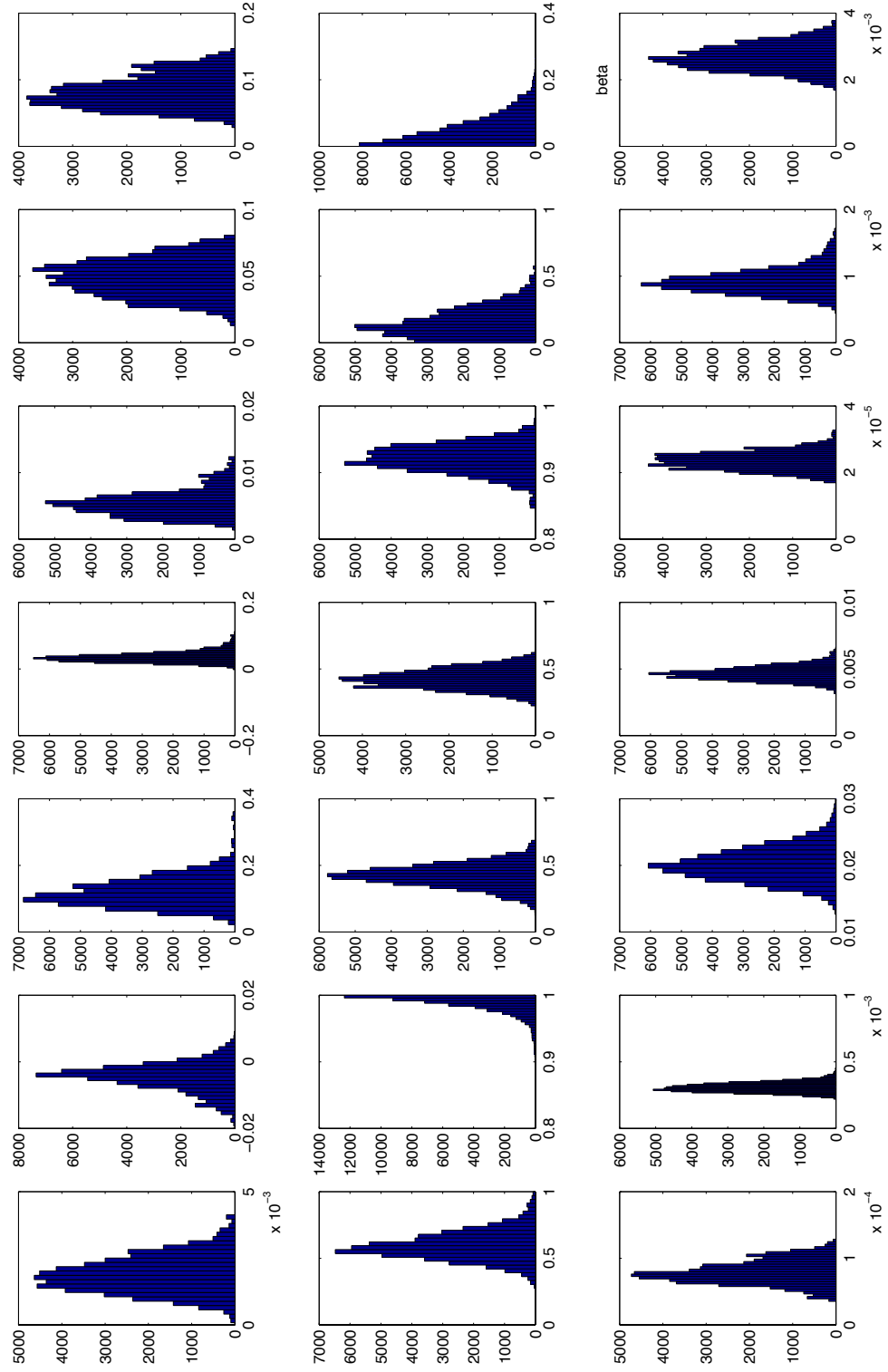


Figura 7: Índice y tendencia

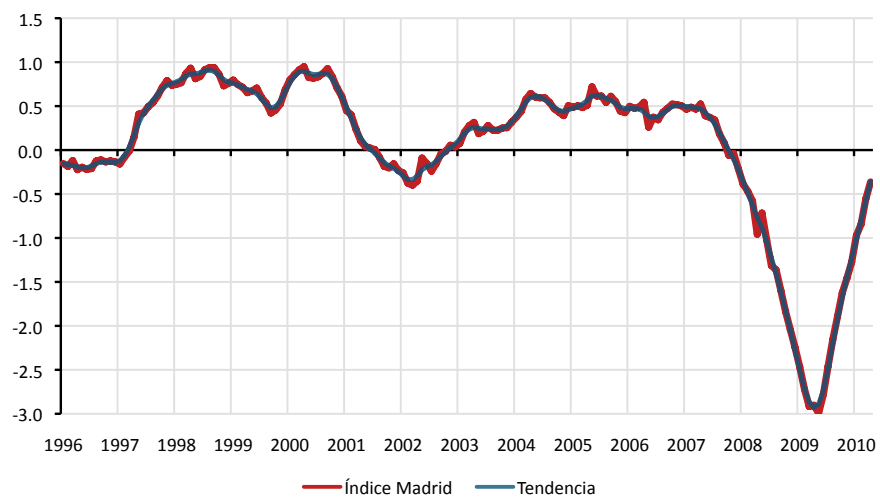


Figura 8: Índice en media trimestral y PIB normalizado

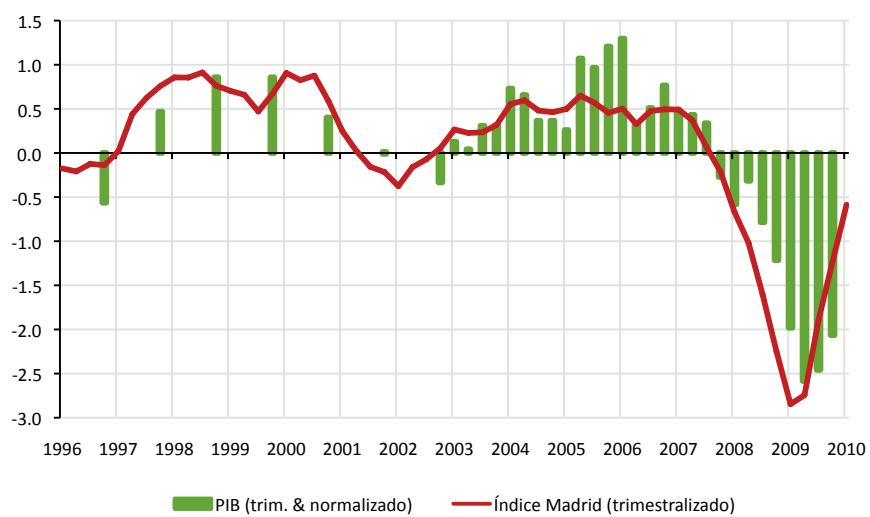


Figura 9: Índice para la Comunidad de Madrid e Índice para España

