

**UN INDICADOR DEL NIVEL DE ACTIVIDAD DE LA ECONOMIA
PERUANA EN BASE A UN MODELO DE VARIABLES LATENTES**
Javier Escobal y Marco Castillo

**EFFECTOS DE LA POLITICA MONETARIA
SOBRE EL MERCADO DE ACCIONES**
Armando Cáceres y Javier Nagamine

**LAS DIFERENCIAS EN TASAS DE INTERES COMO
DETERMINANTES DEL TIPO DE CAMBIO LIBRE**
Armando Cáceres y Javier Nagamine

**UN INDICADOR DEL NIVEL DE ACTIVIDAD DE LA ECONOMIA
PERUANA EN BASE A UN MODELO DE VARIABLES LATENTES**
Javier Escobal y Marco Castillo

**EFFECTOS DE LA POLITICA MONETARIA
SOBRE EL MERCADO DE ACCIONES**
Armando Cáceres y Javier Nagamine

**LAS DIFERENCIAS EN TASAS DE INTERES COMO
DETERMINANTES DEL TIPO DE CAMBIO LIBRE**
Armando Cáceres y Javier Nagamine



1ra. Edición: Lima, 1993
Impreso en el Perú
(c) Grupo de Análisis para el Desarrollo, GRADE
Av. del Ejército 1870, San Isidro, Lima

Edición: José Peláez Cáceres

CENDOC-BIBLIOTECA-GRADE: Catalogación en la fuente

Escobal, Javier; Castillo, Marco; Cáceres, Armando; Nagamine, Javier

Un indicador del nivel de actividad de la economía peruana en base a un modelo de variables latentes. Efectos de la política monetaria sobre el mercado de acciones. Las diferencias en tasas de interés como determinantes del tipo de cambio libre. -- Lima: GRADE; Consorcio de Investigación Económica, 1993. -- (Notas para el debate, 7).

<METODOS ECONOMETRICOS> <ANALISIS DE ACTIVIDAD> <MODELOS DE SERIES DE TIEMPO> <MERCADO FINANCIERO> <TIPO DE CAMBIO> <POLITICA MONETARIA> <PERU>

ISBN 84-89305-34-X

A través de las publicaciones de la serie Notas para el Debate, el Grupo de Análisis para el Desarrollo -GRADE- busca difundir los resultados de algunas de sus actividades, como seminarios o conferencias, o productos preliminares de sus estudios. Su propósito es contribuir a la discusión pública sobre alternativas de política económica y social, desde la perspectiva de quienes vienen investigando y reflexionando sobre estos temas.

Las opiniones y recomendaciones vertidas en este documento son responsabilidad de sus autores y no representan necesariamente el punto de vista de GRADE, ni el de las instituciones a las que pertenecen.

Contenido

Presentación	7
--------------	---

UN INDICADOR DEL NIVEL DE ACTIVIDAD DE LA ECONOMIA PERUANA EN BASE A UN MODELO DE VARIABLES LATENTES

Javier Escobal y Marco Castillo	9
---------------------------------	---

Resumen	11
Introducción	12
1. Algunos conceptos sobre la representación dimensional de variables de estado	14
2. Especificación del modelo para el caso peruano	16
3. Resultados	18
3.1 Propiedades estocásticas de las series	18
3.2 Determinación de la estructura del modelo	21
Bibliografía	26
Anexo	28

EFFECTOS DE LA POLITICA MONETARIA SOBRE EL MERCADO DE ACCIONES

Armando Cáceres y Javier Nagamine	31
-----------------------------------	----

Resumen	33
Introducción	34
1. El mercado de valores en el Perú	35
2. Política monetaria y mercado de acciones	38
3. La hipótesis de mercados eficientes en el mercado de acciones	39
4. Resultados econométricos	40
4.1 Prueba de raíces unitarias	42
4.2 Regresiones correspondientes al Índice General Bursátil	43
4.3 Regresiones correspondientes al índice minero	44
4.4 Regresiones correspondientes al índice industrial	46
5. Conclusiones	48
Bibliografía	51
Anexo	53

LAS DIFERENCIAS EN TASAS DE INTERES COMO DETERMINANTES DEL TIPO DE CAMBIO LIBRE

Armando Cáceres y Javier Nagamine

55

Resumen	57
Introducción	58
1. Los mercados financiero y cambiario peruanos	60
1.1 Evolución de la política cambiaria	60
1.2 Evolución de la política de tasas de interés	64
2. Determinación del tipo de cambio y de la prima por riesgo	65
2.1 Un modelo econométrico	66
3. Resultados de la estimación	67
3.1 Estimación según el método MCO	67
3.2 Aplicación del método de Filtros de Kalman	69
3.3 Estimación según el método ARCH-M	72
4. Conclusiones	78
Bibliografía	80
Anexo	82

Presentación

Esta nueva entrega de la serie *Notas para el Debate* recoge tres ensayos elaborados por investigadores del Grupo de Análisis para el Desarrollo (GRADE) sobre tres distintos tópicos macroeconómicos: la evolución del nivel de actividad económica, los efectos de la política monetaria sobre el mercado de valores, y las relaciones entre tasa de interés y mercado cambiario. El propósito último de estos ensayos es aportar algunos criterios metodológicos que pueden ser de mucha utilidad para el análisis y seguimiento de las variables macroeconómicas involucradas.

El ensayo que abre el documento, elaborado por Javier Escobal y Marco Castillo, presenta un método econométrico para la elaboración de un indicador del nivel de actividad de la economía peruana. Dicho indicador permite predecir el comportamiento del nivel de actividad en base a series de producción, empleo, ingresos y ventas, datos de relativamente fácil acceso. El diseño y construcción del indicador es una tarea emprendida en GRADE como parte de su proyecto permanente de Seguimiento de la Coyuntura Macroeconómica, y fue presentado en setiembre de 1992 en el XI Encuentro Latinoamericano de la Sociedad Econométrica, en México.

El primero de los dos ensayos de Armando Cáceres y Javier Nagamine incluidos en este documento, se ocupa de evaluar en detalle cómo las decisiones de emisión y en general la política monetaria influyen en la determinación de los precios de las acciones en el mercado bursátil. El segundo analiza los efectos que las diferencias entre la tasa de interés interna y la externa tienen sobre la determinación del tipo de cambio en el mercado libre, y las razones por las cuales estas diferencias no siempre se traducen completamente en una devaluación. Ambos estudios forman parte de los resultados del proyecto «Política Monetaria e Inflación», orientado a realizar un análisis econométrico de los mercados de dinero y crédito en el Perú. El primer ensayo fue también presentado en el XI Encuentro Latinoamericano de la Sociedad Econométrica, mientras que el segundo fue una ponencia expuesta en el X Encuentro, realizado en Montevideo en setiembre de 1991.

GRADE desea agradecer al Centro Internacional de Investigación para el Desarrollo y a la Agencia Canadiense de Desarrollo Internacional por el apoyo continuo a la realización de los diversos estudios macroeconómicos de los cuales se derivan estos trabajos, y al Consorcio de Investigación Económica, auspiciado por esas mismas instituciones, por la ayuda financiera brindada para la publicación de este documento.

Lima, enero de 1993.

Un indicador del nivel de actividad de la economía peruana en base a un modelo dinámico de variables latentes

Javier Escobal D'Angelo
Marco Castillo del Castillo

Resumen

Se plantea aquí la construcción de un indicador del nivel de actividad económica para el Perú, a partir de un conjunto de series de tiempo de fácil acceso (series de producción, empleo, ingresos, etc.). El modelo supone que la dinámica de dichas series responde a la evolución de un sólo factor común no observable directamente y que subyace a todas ellas (variable latente), en este caso el nivel de actividad. El primer paso consiste en efectuar pruebas estadísticas para determinar si es posible especificar una sola variable latente común en el sistema. Luego se construye un modelo en el cual la evolución de las series de tiempo elegidas es explicada por la dinámica de la variable latente, la cual, a su vez, sigue un proceso autorregresivo de primer orden. Bajo la representación dimensional de variables de estado (*State Space Form*) es posible estimar los parámetros del sistema (la determinación de los valores óptimos de los parámetros se realiza a partir del algoritmo Esperanza-Maximización/EM). Por último, utilizando el algoritmo del filtro de Kalman, se identifica la evolución de la variable de estado, que representa al nivel de actividad de la economía.

Introducción

En la década de los treinta, el *National Bureau of Economic Research* (NBER) emprendió la labor de construir indicadores del nivel de actividad económica en base a las investigaciones que sobre el ciclo económico norteamericano realizaran Burns y Mitchel (1946). Como resultado de estos trabajos se implementó un sistema de monitoreo del nivel de actividad de la economía norteamericana, cuya principal finalidad era detectar la inminencia de un ciclo de expansión o de una depresión.

El mencionado sistema de estimación recurre al seguimiento de algunos agregados macroeconómicos que, teóricamente, resumen buena parte de la información relevante sobre la evolución del nivel de actividad económica. Estas variables-objetivo del sistema de predicción son, entre otras, el producto bruto interno, el nivel de empleo, el consumo agregado y el ingreso personal.

La metodología de construcción de los indicadores de la actividad económica se basa en la identificación de las variables correlacionadas con más fuerza con las variables-objetivo. El requisito es que la relación sea estable a lo largo del tiempo (es decir, que sea constante aun en periodos de expansión y de depresión). La correlación detectada con la variable-objetivo puede ser de rezago o adelanto, clasificándose a las variables como *adelantadas* si su correlación con periodos posteriores de la variable-objetivo es alta, *coincidentes*, si la correlación es alta en el periodo en curso, y *rezagadas* si su correlación mayor es mantenida con periodos pasados de la variable-objetivo.

A partir de las variables correlacionadas se elabora un índice que resume toda la información provista por ellas. Dicho índice, conocido como indicador líder, es un promedio ponderado de las variables seleccionadas -previamente diferenciadas o estandarizadas¹ - y sirve como un predictor de la evolución de la variable objetivo (v.g., la producción, el empleo o los ingresos).

Mucho del trabajo estadístico realizado para elevar la calidad predictiva de los indicadores líderes se ha centrado en evaluar las ponderaciones utilizadas, apuntando a lograr los pesos óptimos. Esta labor de optimización depende en gran medida del criterio empleado, que hace que los pesos relativos de cada variable en el indicador agregado tomen valores diferentes. El NBER ha generado una serie de criterios a los cuales se asocian diversos puntajes que, sumados, determinan el peso de cada variable en el índice. Dichos criterios son, entre otros, el grado de correlación de la serie de referencia (o indicador) y la variable-objetivo, su significancia económica, su eficiencia en reproducir los puntos de inflexión de la variable-objetivo, la consistencia metodológica de las series, la prontitud con que éstas son publicadas, etc. Como es de suponer, cada criterio tiene un alto

1. Las variables son sometidas a transformaciones conducentes a lograr su estacionaridad, lo cual permite realizar el trabajo estadístico sin riesgo de obtener correlaciones espurias.

grado de arbitrariedad, aun cuando han sido evaluados permanente y consistentemente durante décadas.

Contador (1977) intentó implementar la metodología para el caso brasileño, lo cual lo llevó a la larga a criticar la metodología de construcción de los ponderadores y a optar por criterios alternativos. Dicho autor consideró que el peso que cada variable correlacionada tiene en la determinación de la variable-objetivo estaba completamente resumido en su grado de correlación simple con ésta: así, el criterio de agregación resulta estadístico (poblacional) antes que arbitrario. Sin embargo, es claro que la elección del grado de correlación simple como criterio de ponderación ya es de por sí arbitraria, teniendo en cuenta, como el mismo Contador admite, que este estadístico no es siempre el óptimo².

En todo caso, la construcción de estos «pesos relativos» para cada variable correlacionada resulta una tarea importante, en tanto ellos determinan la forma en que se ha de discriminar la calidad y cantidad de información proporcionada por cada serie. Las aproximaciones tradicionales a la construcción de los índices de actividad económica suponen que éstos son resultado de la agregación de información proveniente de diversos mercados y actividades productivas; ello supone que la relación del indicador con sus variables componentes es lineal y fija.

Sin embargo, es posible enfocar el problema desde otra óptica, que consiste en suponer que las variables correlacionadas están generadas por un mismo proceso estocástico o que existe correlación entre ellas debido a que están afectadas por el mismo proceso económico. De ese modo, el indicador no resulta de la simple agregación de un número arbitrario de variables componentes, sino que es un factor que expresa su estructura común o revela la fuente común de correlación. Obviamente, esta fuente común es una variable no observable, requiriendo su estimación de una metodología acorde con este supuesto.

Stock y Watson (1988) han avanzado en esta dirección proponiendo que se dé al sistema una representación dimensional de variables de estado (*state space form*), en la cual el **indicador coincidente** representa a una única variable no observable (variable latente) que explica los cambios en el corto plazo en las **variables componentes** de dicho indicador (que son variables coincidentes correlacionadas), es decir, busca aproximar la evolución del factor común del cual se habló en el párrafo anterior. Dicho indicador coincidente es a su vez generado por un proceso aleatorio perfectamente representable mediante una

2. El autor señala que el grado de correlación simple puede indicar estructuras de rezago o adelanto entre la variable-indicador y la variable-objetivo, que no siempre se mantienen cuando se descompone la serie en sus componentes cíclicos. Las variables pueden estar fuertemente correlacionadas sólo en ciclos de determinada frecuencia, por lo que dicha asociación puede no reflejarse en la información aportada por la correlación simple. Para descubrir esta asociación se requiere analizar el espectrograma de la serie.

estructura autorregresiva de orden p , $AR(p)$, donde p es conocido y finito. De esa manera, las ecuaciones que representan a las variables componentes y las ecuaciones del factor común registran procesos estocásticos estacionarios. El recurrir a la representación dimensional de variables de estado permite hacer uso del algoritmo conocido como **Filtro de Kalman** para la estimación de los valores del indicador.

El planteamiento de un factor común del cual dependen las variables componentes intenta dar un sustento conceptual a la existencia de correlación entre este grupo de variables, siendo tal factor común un resumen del **estado general de la economía**. Existen diversos modelos económicos que pueden dar cabida a representaciones como la planteada aquí. Un caso es la versión del modelo del ciclo real de negocios de King, Plosser y Rebello (1988) en la cual la evolución del nivel real de actividad se ve reflejada en las modificaciones de un conjunto de variables, cuya dinámica responde básicamente a un factor común que sintetiza los cambios en la tecnología y en las preferencias³. El esquema es compatible con planteamientos en los cuales variables nominales también permiten determinar el factor común (o los factores comunes): tal es el caso del trabajo de Singleton (1982), que emplea modelos factoriales dinámicos para explicar la conducta coincidente de las tasa de interés a diferentes plazos, al igual que Watson y Kraft (1982) para descomponer los efectos anticipados y no anticipados de la demanda agregada sobre la producción.

Este estudio discute la aplicación de esta metodología para la construcción de un indicador del estado general de la economía peruana. Se inicia con una breve exposición sobre cómo aplicar una representación dimensional de variables de estado (*state space form*) a un modelo de variables latentes, y describe la forma en que puede emplearse el filtro de Kalman para la obtención de estimados óptimos de la variable no observable. Luego de describir los algoritmos usados para identificar los parámetros de la forma funcional utilizada, se procede a estimar un indicador del nivel de actividad para la economía peruana. Por último, el desempeño del indicador que aquí se propone es evaluado comparando su evolución con la del producto bruto interno.

1. Algunos conceptos sobre la representación dimensional de variables de estado

La representación dimensional de variables de estado (RDVE) permite relacionar un grupo de variables no observables (consideradas *inputs*) con otro grupo

3. Sin embargo, que dicho planteamiento teórico permita la determinación de un único factor común no quiere decir necesariamente que el modelo aquí planteado se identifique con él: para ello se tendría que imponer restricciones adicionales al modelo.

de variables que son generadas por aquéllas (consideradas *outputs*). Concretamente, la RDVE es una forma compacta de expresar un sistema de ecuaciones en diferencia o diferenciales de orden finito, como un proceso markoviano de primer orden.

La RDVE, en su forma más general, se compone de dos grupos de ecuaciones. El primer grupo es el de las **ecuaciones de transición**, que resumen el modo en que se determinan las variables de estado (no observables). Dichas ecuaciones son representadas mediante un vector de variables que es generado por un proceso autorregresivo de primer orden, un conjunto de variables exógenas y un vector aleatorio (que representa el término de perturbación). Formalmente se representa del siguiente modo:

$$x_t = \Theta_t x_{t-1} + \Upsilon_t z_t + G v_t \quad (1)$$

donde x , el vector de variables de estado, es j -dimensional; z , el vector de variables exógenas, es de dimensión k ; y v , el término de perturbación, es m -dimensional. La matriz G (de dimensión $m \times m$) permite por su parte una estructura de errores más flexible.

El segundo grupo de ecuaciones, llamadas **ecuaciones de medida**, relaciona las variables observadas con el vector de variables de estado, otras variables explicativas y un término de perturbación. La representación formal es como sigue:

$$y_t = \alpha_t x_t + \beta_t z_t + e_t \quad (2)$$

donde y , el vector de variables observadas, es de dimensión p ; x , el vector de variables de estado, es j -dimensional; z , el vector de variables exógenas, es k -dimensional; y e , el término de perturbación, tiene dimensión p .

Adicionalmente se considera que v_t y e_t se distribuyen como una normal multivariada, siendo independientes entre ellas:

$$\begin{bmatrix} v_t \\ e_t \end{bmatrix} \sim N \left[0, \begin{bmatrix} Q & O \\ O & R \end{bmatrix} \right] \quad (3)$$

donde Q y R representan las matrices de varianzas y covarianzas de v_t y e_t , respectivamente.

Si se supone que $\gamma = \beta = 0$ y que R y Q son matrices diagonales, las variables observadas pueden considerarse como simples señales o efectos de las variables de estado, que son las que finalmente generan el proceso. Este caso específico será utilizado más adelante.

Esta representación del sistema de ecuaciones permite hacer uso del algoritmo del Filtro de Kalman para estimar el valor de las variables de estado. El filtro de Kalman es un sistema de ecuaciones que hace uso de la información contenida

en las matrices de parámetros de las ecuaciones de transición y medida y en la matriz de varianzas de los términos de perturbación⁴. Dichas ecuaciones, llamadas de predicción, pueden presentarse de la siguiente forma:

$$x_t |_{t-1} = \phi_t x_{t-1} + \gamma_t z_t \quad (4)$$

$$P_t |_{t-1} = \phi_t P_{t-1} \phi_t' + G Q_t G' \quad (5)$$

$$y_t |_{t-1} = \alpha_t x_t |_{t-1} + \beta_t z_t \quad (6)$$

$$H_t = \alpha_t P_t |_{t-1} \alpha_t' + R_t \quad (7)$$

donde $x_t |_{t-1} = E[x_t | I_{t-1}]$, P_t es la matriz de varianzas correspondiente al estimado de x_t , y H_t la matriz de varianzas correspondiente al estimado de y_t . El sistema de ecuaciones de predicción está acompañado por otro grupo de ecuaciones (llamadas de actualización), las cuales permiten la acción iterativa del algoritmo. Estas ecuaciones son las siguientes:

$$x_t |_{t-1} = x_t |_{t-1} + P_t |_{t-1} \alpha_t' H_t^{-1} (y_t - y_t |_{t-1}) \quad (8)$$

$$P_t |_{t-1} = P_t |_{t-1} + P_t |_{t-1} \alpha_t' H_t^{-1} \alpha_t P_t |_{t-1} \quad (9)$$

2. Especificación del modelo para el caso peruano

Empleando la RDVE es posible especificar un modelo que permite obtener un indicador del nivel de actividad para el caso peruano, de acuerdo a los planteamientos hechos por Watson y Engle (1981) y por Stock y Watson (1989). Dicho indicador es construido a partir de la información que proporcionan las series de cuatro variables coincidentes: el Índice de Volumen Físico de la Producción Fabril, el Índice de Empleo, el Índice de Sueldos y Salarios y el Valor de las Ventas Minoristas en Lima Metropolitana⁵. A continuación se detalla la metodología empleada en la construcción de dicho indicador.

Considérese que y_t es un vector que contiene las cuatro variables coincidentes, expresadas como tasas de variación porcentual. Las series de dichas variables siguen el siguiente comportamiento:

$$y_{it} = \alpha_i C_t + e_{it} \quad (10)$$

4. Se supone que esta información es conocida *a priori* sin ella es imposible usar el algoritmo. Como se verá en la siguiente sección, la necesidad de especificar la forma de las ecuaciones y los valores de los parámetros exige disponer de una metodología adicional.

5. Las tres primeras series están expresadas con base 1979=100 y la cuarta está expresada en miles de intis de 1979. Para fines de comparación se usará luego el PBI, en miles de millones de intis de 1979. Todas las series son trimestrales y van desde el periodo 1976:1 hasta el periodo 1991:4.

donde $E[e_{it} e_{jt}] = 0$, $E[e_{it} C_t] = 0$, para todo t , i diferente de j . Por su parte, C_t es el indicador coincidente expresado como tasa de variación porcentual, y e_t es un vector de cuatro términos de perturbación. Se considera que C_t es el único factor que genera a las series líderes, siendo a su vez generado por un proceso AR(p), mientras que el término e_t es generado por un proceso AR(q):

$$C_t = \sum_{i=0}^{p-1} \phi_i C_{t-i} + v_t \tag{11}$$

$$e_{it} = \sum_{j=1}^q \Gamma_{ij} e_{it-j} + \mu_{it} \tag{12}$$

Para presentar este modelo según la RDVE, es preciso determinar las ecuaciones de transición y las ecuaciones de medida. Si X_t , el vector de variables de estado, incluye tanto al indicador coincidente como a los términos de perturbación, entonces la ecuación de transición puede escribirse del siguiente modo:

$$X_t = \begin{bmatrix} C_t \\ e_{1t} \\ e_{1t-1} \\ e_{2t} \\ e_{2t-1} \\ e_{3t} \\ e_{3t-1} \\ e_{4t} \\ e_{4t-1} \end{bmatrix} = \begin{bmatrix} \phi & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \rho_{11} & \rho_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \rho_{21} & \rho_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_{31} & \rho_{32} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_{41} & \rho_{42} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} (X_{t-1}) + v_t \tag{13}$$

y la ecuación de medida se puede escribir de la siguiente forma:

$$Y_t = \begin{bmatrix} \alpha_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \alpha_2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \alpha_3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \alpha_4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} (X_t) \tag{14}$$

$$Y_t = [y_t^1 \ y_t^2 \dots \ y_t^4]'$$

donde:

$$\begin{bmatrix} v_t \\ e_t \end{bmatrix} \sim N \left[\emptyset, \begin{bmatrix} Q & \emptyset \\ \emptyset & R \end{bmatrix} \right]; \quad Q = \begin{bmatrix} \sigma_c^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \sigma_1 & 0 \\ 0 & 0 & \sigma_4^2 \end{bmatrix} \quad R = \emptyset_{4 \times 4} \quad (15)$$

Esta forma de presentar el modelo permite usar las ecuaciones del filtro de Kalman, y mediante ellas calcular los errores de predicción del modelo. El algoritmo Esperanza-Maximización (EM), descrito en el anexo de este documento, permite usar dichos errores de predicción para estimar el valor de los parámetros que optimizan la función de verosimilitud.

3. Resultados

Toca ahora presentar los resultados de la estimación del modelo para series económicas peruanas. El primer paso para ello es analizar las propiedades estadísticas de las series utilizadas, para determinar la transformación óptima a la cual deben ser sometidas para hacerlas estacionarias. Luego se determina la estructura del modelo, previa revisión del procedimiento y los criterios empleados para elegir la estructura de rezagos y una especificación con un único factor común.

3.1 Propiedades estocásticas de las series

Para caracterizar estocásticamente las series se aplican pruebas conducentes a determinar si es necesario transformar las series antes de iniciar las estimaciones. Dado que las fuentes más usuales de no estacionaridad de las series económicas son su fuerte correlación serial, se puede aplicar las pruebas Dickey-Fuller para detectar la existencia de raíces unitarias, y el test Hylleberg-Engle-Granger-Yoo para detectar el mismo problema pero con series que tienen componentes estacionales.

La aplicación del test de Dickey-Fuller a las series de las cuatro variables empleadas (el Índice de Volumen Físico de la Producción Fabril, el Índice de Empleo, el Valor de las Ventas Minoristas de Lima Metropolitana y un índice de sueldos y salarios llamado Índice de Ingreso)⁶, sugiere que todas las series con-

6. Las pruebas han sido realizadas sobre el logaritmo natural de la series; si éstas tenían algún

tienen una raíz unitaria cuando están expresadas en niveles, pero rechaza la hipótesis cuando las series son presentadas en primeras diferencias⁷ (ver el cuadro 1). El contraste, sin embargo, supone que los factores estacionales de las series son determinísticos (lo que posibilita el retiro de éstos con *dummies*

Cuadro 1
Contraste para detectar raíces unitarias

Variable	Valor del estadístico			
	Niveles	Diferencias	Niveles y Tendencia	Diferencias y Tendencia
Ventas Minoristas	-1.02577	-4.39596	-2.00461	-4.43089
Ingresos	-1.01820	-3.33805	-2.14249	-3.30505
Empleo Manufacturero	-0.67196	-3.35455	-2.19048	-3.47663
IVF Total	-2.59195	-5.36744	-2.60994	-5.32169
PBI	-1.41593	-3.64400	-1.22702	-3.68558
Valores Críticos				
Nivel de significancia:				
al 1%	-3.43000	-3.43000	-3.96000	-3.96000
al 5%	-2.86000	-2.86000	-3.41000	-3.41000
al 10%	-2.57000	-2.57000	-3.12000	-3.12000

El valor de cada prueba es el valor de la *t* de student para β en especificaciones alternativas. La primera y la segunda columnas para cada variable son las pruebas realizadas en niveles y primeras diferencias de la siguiente especificación:

$$dz = \alpha + \beta z_{t-1} + \sum_{i=1}^{i-4} dz_{t-i}$$

mientras que las dos últimas columnas muestran los resultados de las pruebas de una especificación que incluye una tendencia lineal:

$$dz = \alpha + \tau_t + \beta z_{t-1} + \sum_{i=1}^{i-4} dz_{t-i}$$

Las pruebas se realizaron considerando cuatro rezagos para evitar problemas de autocorrelación en los residuos.

componente estacional éste fue retirado con *dummies* estacionales. Los componentes estacionales sólo parecen existir en las series de PBI, ventas minoristas y producción industrial.

7. El test de Dickey-Fuller consiste en la estimación de la siguiente ecuación:

$$Dy_t = \alpha + \beta y_{t-1} + \theta(L)Dy_{t-1},$$

para su posterior contraste con la hipótesis que sostiene que β es significativamente diferente de cero. El test puede incluir una tendencia determinística.

estacionales) No obstante, si los componentes estacionales son estocásticos es posible que la no estacionaridad de las series provenga de esta fuente.

Dado que se emplean series trimestrales, pudiendo además algunas de ellas contener marcados patrones estacionales (PBI, producción fabril y ventas minoristas), es posible que la fuente de no estacionaridad de estas series radique en alguna frecuencia estacional. Si se supone la existencia de raíces unitarias estacionales el test de Dickey-Fuller estaría incorrectamente especificado, al no considerar otras frecuencias diferentes a la frecuencia cero⁸. Tal problema ha sido resuelto por Hylleberg-Engle-Granger-Yoo (1990) quienes sugieren un test que contrasta la hipótesis conjunta de existencia de raíces unitarias en la frecuencia cero y en alguna otra frecuencia estacional.

Cuadro 2
Contraste para detectar raíces unitarias estacionales

Variable	π_1	π_2	$\pi_3 \cap \pi_4$
1. PBI	-2.423923	-2.087921	5.83703
2. IVF total	-2.737003	-2.391702	6.30461
3. Ventas Minoristas	-2.258348	-2.172374	10.6370
<i>Dummies</i> estacionales y constante			
Valores criticos (N=48 y $\alpha = 10\%$)	-2.72	-2.69	5.50

El contraste se hace respecto a la hipótesis conjunta de existencia de raíces unitarias en la frecuencia cero (\bar{x}_t es $I(1)$) y en alguna frecuencia estacional. La hipótesis de raíz unitaria en la frecuencia cero se acepta si π_1 no es significativamente diferente de cero; la existencia de raíces unitarias estacionales se acepta si π_2 no es significativamente diferente de cero y además no puede rechazarse la hipótesis que $\pi_3 = \pi_4 = 0$ ⁹. Para verificar la hipótesis conjunta se ha añadido a la prueba rezagos de la variable dependiente, lo que permite evitar

8. Recuérdese que si existe una raíz unitaria en una serie, ésta presentará un máximo en la frecuencia cero de su espectograma; luego, si existen raíces unitarias en alguna frecuencia estacional, entonces el máximo se encontrará en la frecuencia correspondiente a la estación.

9. El contraste de las hipótesis se hace en base a la estimación de la siguiente ecuación:

$$\theta(L)y_{4t} = \pi_1 y_{1,t-1} + \pi_2 \bar{y}_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1},$$

donde los subíndices denotan que las series han sido transformadas a fin de poder identificar los componentes de cada frecuencia.

problemas de correlación de los errores, y *dummies* estacionales para saber si es suficiente usar éstas o resulta necesario realizar alguna otra transformación. El contraste confirma que no basta retirar el factor estacional de las series de Producto Bruto Interno, IVF Total y Ventas Minoristas. Para lograr que las series sean estacionarias es necesario además diferenciarlas, como se hizo con las series de ingresos y de empleo.

3.2 Determinación de la estructura del modelo

El siguiente paso es la determinación de la estructura del modelo. Ello requiere estimar el número de rezagos óptimos y el número de factores subyacentes al sistema. Para tal efecto se debe evaluar el ratio de verosimilitud entre especificaciones alternativas del modelo (Harvey 1990). Dicho ratio compara los valores de la función de verosimilitud de dos modelos, suponiendo que un modelo puede ser obtenido imponiendo restricciones lineales a los parámetros del otro. En caso que las restricciones sean de nulidad, la prueba tiene como hipótesis nula que los parámetros bajo restricción no son significativamente diferentes de cero. La prueba se escribe como $-2\ln(I_0/L_1)$ y tiene una distribución χ^2 con tantos grados de libertad como el número de restricciones. L_0 denota la función de verosimilitud del modelo restringido y L_1 denota la función de verosimilitud del modelo irrestricto.

Al estimar el indicador del nivel de actividad a partir de especificaciones distintas, se verifica que el desempeño de los modelos donde el componente no observado sigue un proceso mayor a AR(1) es inferior al de los modelos donde dicho componente sigue un proceso AR(1). De otro lado, la mejor especificación plantea que los errores asociados a las ecuaciones de medida siguen un proceso AR(2). En este caso el ratio de verosimilitud rechaza el modelo con un proceso AR(1). No obstante, se adopta finalmente la especificación mediante un modelo AR(1) para el factor común y AR(1) para los errores (excepto para la ecuación de empleo), porque su capacidad predictiva respecto al PBI es mayor, aun cuando el proceso AR(2) se adapta mejor para la estimación del factor común.

Una última prueba es necesaria a fin de asegurar que la especificación del modelo sea la óptima. La prueba consiste en confrontar el modelo AR(2) para los errores y AR(1) para el factor común, con un modelo idéntico pero que admita un segundo factor común y luego con un modelo AR(2) para todos los errores, el cual no admita factores comunes. Dadas estas especificaciones, el modelo sin factores comunes puede ser obtenido imponiendo restricciones al modelo de un sólo factor común; asimismo, el modelo de un sólo factor común puede derivarse imponiendo restricciones al modelo de dos factores comunes.

La primera prueba del ratio de verosimilitud muestra que el modelo de un factor común es preferible al modelo sin factores, a casi cualquier nivel de significancia ($\chi^2(5) = 28.41$). Sin embargo, al confrontar los modelos unifactorial y multifactorial, se acepta el modelo de un sólo factor a un nivel de significancia de 97.1% ($\chi^2(5) = .88$)¹⁰.

De acuerdo a la especificación establecida, en base a los datos de las series descritas se puede realizar la estimación del modelo (ver el cuadro 3). El factor común se obtiene luego de que el modelo converge en 48 iteraciones. La capacidad explicativa sobre las variables seleccionadas sobrepasa en todas las especificaciones estimadas el 72%, a excepción de la variable empleo, cuyo desempeño no supera el 50%.

Cuadro 3
Parámetros del modelo

	Producción Industrial	Indice de Empleo	Indice de Ingresos	Ventas Minoristas
α_i	0.897 (0.024)	0.310 (0.085)	0.711 (0.063)	0.816 (0.32)
ρ_{1i}	-0.428 (0.065)	0.324 (0.134)	-0.589 (0.104)	0.064 (0.026)
ρ_{2i}		0.287 (0.113)		
σ_i	0.157	0.602	0.520	0.378
$C_t =$	0.160 * C_{t-1} (0.094)			

Nota: Los números entre paréntesis son las desviaciones estándar de las estimaciones.

Una vez que se dispone de la estimación del indicador coincidente, es posible deducir los ponderadores correspondientes para poder calcular dicho indicador a partir de las variables componentes. Esto se consigue invirtiendo las ecuaciones

10. Las pruebas para determinar el número de factores comunes a un conjunto de series se hacen en el dominio de las frecuencias, debiéndose para ello transformar las series en estacionarias y luego en procesos "ruido blanco". Su ventaja sobre la metodología utilizada aquí reside en que no recae en la arbitrariedad de la estructura de rezagos del modelo que se está evaluando y adicionalmente en que puede identificarse las frecuencias en las cuales el modelo es válido. La derivación de estas pruebas se debe a Geweke (1977).

del sistema y utilizando las ecuaciones del Filtro de Kalman. Dado que las ecuaciones de actualización del filtro usan como información los errores de predicción del modelo ($v_t = Y_t - \alpha C_{t/t-1}$), es posible reemplazar en la ecuación de actualización a $C_{t/t}$ por $C_{t/t-1} + P_{t/t-1} \times \alpha \times \text{inv}(H_t) \times v_t$, de donde resulta:

$$C_{t/t} = (I - G_t \times \alpha) \times \varnothing_t \times C_{t-1/t-1} + G_t \times Y_t, \quad (16)$$

siendo $G_t = P_{t/t-1} \times \alpha \times \text{inv}(H_t)$, ecuación que se conoce como la ganancia de Kalman. Como se supone aquí que $\varnothing_t = \varnothing$ para todo t , la ganancia de Kalman tiende a un valor estacionario G^* . En estas condiciones, la ecuación para el componente no observado puede escribirse del siguiente modo:

$$(I - KL)C_{t/t} = G^*Y_t \quad (17)$$

donde $K = (I - G_t \times \alpha) \times \varnothing_t$. Si K tiene raíces características menores que uno, la ecuación puede escribirse así:

$$C_{t/t} = e_1 \sum_{j=0}^{j-\infty} K^j G^* Y_{t-j} \quad (18)$$

donde $e = (1, 0, \dots, 0)$.

Los ponderadores obtenidos son presentados en el cuadro 4, donde se muestran los resultados para los primeros cuatro rezagos de las series observadas (los ponderadores han sido normalizados para que su suma sea uno). Así, la última fila muestra el peso de cada variable observada en la ponderación para calcular el indicador coincidente, mientras la última columna presenta la participación de cada rezago en la construcción del índice.

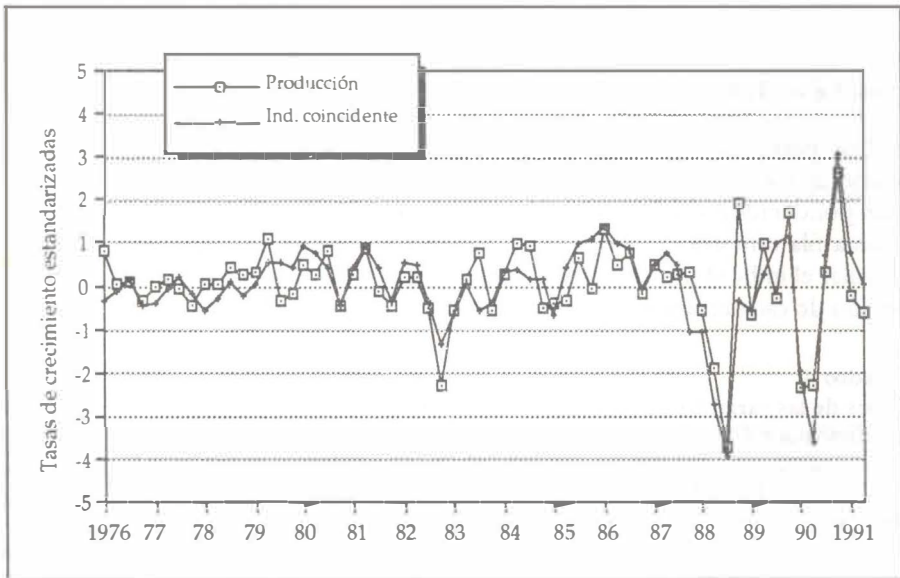
Cuadro 4
Pesos de las variables observadas en la construcción del Indicador Coincidente

Rezago	Producción Industrial	Indice de Empleo	Indice de Ingresos	Ventas Minoristas	TOTAL
0	0.7053	0.0754	0.1607	0.3028	1.2443
1	-0.1767	-0.0192	-0.0401	-0.0767	-0.3127
2	0.0482	0.0052	0.0169	0.0209	0.0853
3	-0.0131	-0.0014	-0.0030	-0.0057	-0.0232
4	0.0036	0.0004	0.0008	0.0015	0.0063
Total	0.5673	0.0604	0.1294	0.2429	1.0000

Nota: El R^2 ajustado respecto al PBI es de 77.7% y el coeficiente de correlación es 89.1%.

La capacidad predictiva del indicador coincidente se confirma al comparar su evolución con la del PBI (ver el gráfico 1). Vale la pena notar que al comparar la evolución del factor común con la del conjunto de series utilizadas (ver los gráficos 2 y 3), se deduce que es el Índice de Producción Fabril el que mayor información aporta al indicador. Asimismo, se hace evidente que la información contemporánea (rezago cero) brindada por las variables componentes debe ser corregida con la evolución pasada de las mismas variables. Las ponderaciones mostradas en el cuadro 4 indican que suponer que cada variable participa en el indicador en un solo periodo resta información al agregado (la contenida en los rezagos). No es extraño entonces que el desempeño del indicador aquí obtenido sea superior -en términos de seguir la evolución del PBI- que el indicador resultante del análisis factorial estático¹¹.

Gráfico 1
Indicador coincidente y PBI



11. Es posible, usando el análisis factorial, estimar una o más variables latentes a las series del modelo, que sirvan de indicadores del proceso común a dichas series. Sin embargo, dichos indicadores no tendrían una estructura de rezagos que permita recoger toda la dinámica de la serie. Al estimar un indicador usando esa metodología y compararlo con el proveniente de la especificación elaborada aquí (estableciendo el punto de comparación en relación a la capacidad predictiva de cada indicador sobre el PBI), el primero mostró un R^2 ajustado de 71% y una correlación de 84%, mientras la especificación mediante la RDVE mostró un mejor desempeño (R^2 ajustado de 77.7% y una correlación de 89.1%). Ello confirma la ventaja de una especificación dinámica del componente común.

Gráfico 2
Indicador y variables coincidentes

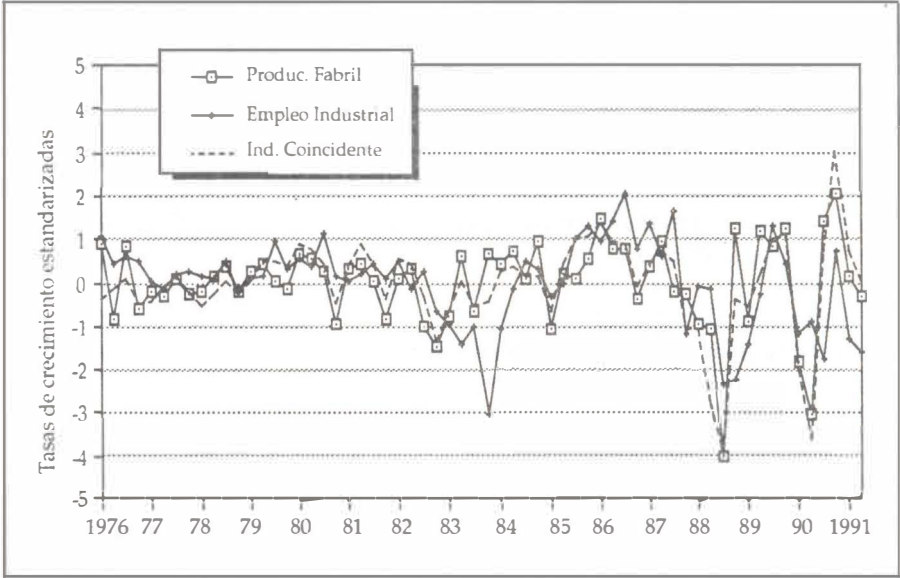
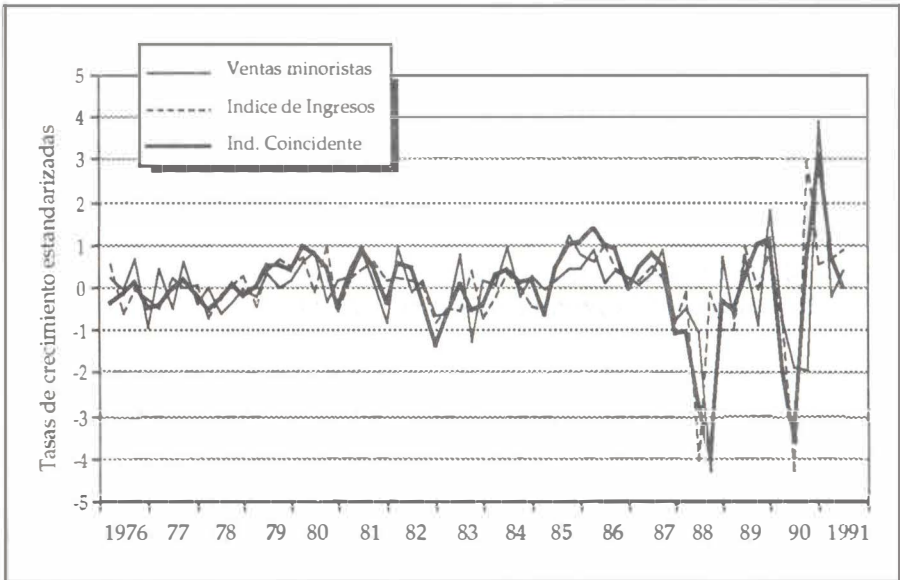


Gráfico 3
Indicador y variables coincidentes



Bibliografía

AUERBACH, A.

1982 «The Index of Leading Indicators: 'Measurement Without Theory', Thirty-five Years Later». En: **Review of Economics and Statistics** 4, pp.589-595.

BURMEISTER, E. y K. WALL

1982 «Kalman filtering estimation of unobserved rational expectations with an application to the German hyperinflation». En: **Journal of Econometrics** 20, pp.255-284.

BURNS, A.F. y W.C. MITCHELL

1946 **Measuring Business Cycles**. NBER. New York.

CONTADOR, C. R.

1977 **Ciclos económicos e indicadores de atividade no Brasil**. IPEA/INDES. Rio.

DIEBOLD, F.X. y G.D.RUDEBUSCH

1990 «A Nonparametric Investigation of Duration Dependence in the American Business Cycle». En: **Journal of Political Economy**, 98(3), pp. 596-616.

1989 «Scoring the Leading Indicator». En: **Journal of Business**, 62(3), pp. 369-391.

GEWEKE, J.

1977 «The dynamic factor analysis of economic time-series models». En: Aigner, D.J. y A.S., **Latent Variables in Socio-Economic Models**, North-Holland.

GHOSH, D.

1989 «Maximun Likelihood estimation of the Dynamic Shock-Error Model». En: **Journal of Econometrics** 41, pp. 121-143.

HARVEY, A.C.

1990 **Forecasting, structural time series models and the Kalman filter**. Cambridge University Press.

HYLLEBERG, S., R.F ENGLE, C.W.J.GRANGER y B.S.YOO

1990 «Seasonal Integration and Cointegration». En: **Journal of Econometrics** 44, pp. 215-238.

KING, R., C. PLOSSER y S. REBELLO

1988 «Production, Growth and Business Cycles. The Basic Neoclassical Model». En: **Journal of Monetary Economics** 21, pp.195-232.

NORRBIN, S. y D. SCHLAGENHAUF

1988 «An inquiry into the sources of macroeconomic fluctuations». En: **Journal of Monetary Economics** 22, pp. 43-70.

SINGLETON, K.

1982 «A Latent Time Series Model of the Cyclical Behavior of Interest Rates». En: **International Economic Review** 21, 3.

STOCK, J. y M. WATSON

1988 **A probability model of the coincident economic indicators**. National Bureau of Economic Research, Working Paper No. 2772, Cambridge.

1989 «New Index of Coincident and Leading Economic Indicators». En: Blanchard, O. y S. Fischer (eds.): **NBER Macroeconomics Annual**. MIT Press, Massachusetts, pp. 351-394.

WATSON, M. y R. ENGLE

1983 «Alternative algorithms for the estimation of Dynamic Factor, MIMIC and Varying Coefficients Regressions Models». En: **Journal of Econometrics** 23, pp. 385-400.

1981 «A one factor multivariate time series model of metropolitan wage rates». En: **Journal of American Statistical Association** 76, pp. 774-781.

WATSON, M. y D.F. KRAFT

1982 **Testing the Interpretation of Indices in a Macroeconomic Index Model**. Harvard Institute of Economic Research, Discussion Paper No. 953.

Anexo

Método de estimación del Indicador Coincidente Compuesto y de los parámetros del modelo

Se expone aquí los métodos de estimación del modelo en su representación dimensional de variables de estado. El filtro de Kalman estima las variables de estado en base a las matrices de parámetros, siendo necesario para aplicarlo conocer previamente dichos parámetros. Precisamente, el algoritmo Esperanza-Maximización (EM) se utiliza para calcular conjuntamente los parámetros del modelo y las variables de estado.

La estimación de los parámetros se inicia definiendo la función de verosimilitud en términos de los errores de predicción de las variables de la ecuación de medida (Watson y Engle 1981, 1983; Burmeister y Wall 1982). Estos están definidos por:

$$\begin{aligned}\eta_t &= y_t - \hat{y}_t \\ &= y_t - \theta_t x_{t-1}\end{aligned}\tag{A.1}$$

Si el modelo está correctamente especificado, estos errores deben ser «ruido blanco» y no debe existir correlaciones cruzadas significativas con sus rezagos. La matriz de varianza-covarianza contemporánea de estos errores es denotada por H_t y es función de los parámetros del modelo. Con esta información puede construirse una función logarítmica de verosimilitud de la forma siguiente

$$\mathcal{L} = \sum \mathcal{L}_t = \sum - \frac{1}{2} (1/n) |H_t| + \eta_t^T H_t^{-1} \eta_t\tag{A.2}$$

Usualmente se estima estos modelos usando el algoritmo conocido como **método de puntuación** (*scoring method*) el que, sin embargo, ha sido criticado por ser ineficiente en términos de cómputo. Dicho método requiere la estimación de las derivadas respecto de los parámetros, lo que implica un cálculo numérico para cada uno de ellos; igualmente es necesario calcular la matriz de información.

Para solucionar tal problema se ha ideado otro método más eficiente en términos de cómputo, que es conocido como método de **Estimación-Maximización** (*EM Method*). El método consta de dos procedimientos que han de repetirse iterativamente. Al principio se atribuyen valores arbitrarios a los parámetros correspondientes a θ , τ , α , β , Q y R . Esta información, acompañada de valores iniciales para $x_{0/0}$ y $P_{0/0}$, permite el cálculo de $x_{t/T}$, que es el valor estimado (esperado) de la(s) variable(s) de estado. El valor de $x_{t/T}$ corresponde al resultado de suavizar el resultado a través del filtro de Kalman.

Este paso permite tener un valor estimado de $x_{t/T}$, y gracias a ello considerar el sistema de ecuaciones para x_t e y_t como si todas las variables fueran observadas. Mediante un arreglo de las ecuaciones es posible hacer la estimación por el método *SUR* (*Seemingly Unrelated Regressions*), generándose así un nuevo valor de los parámetros de θ , τ , α , β , Q y R . Estos resultados sirven para estimar nuevamente $x_{t/T}$ mediante las ecuaciones del filtro de Kalman. Los valores usados en la estimación requieren la obtención de las matrices de *momentos* que permitan considerar la presencia de los errores de predicción obtenidos por el filtro de Kalman.

Las ecuaciones de medida tienen la siguiente forma:

$$Y_t = \alpha x_t + e_t \quad (\text{A.3})$$

donde la variable observada se define como función de la variable no observada. Siendo el estimador MCO (Mínimos Cuadrados Ordinarios) de α igual a

$$\hat{\alpha} = (X'X)^{-1}XY \quad (\text{A.4})$$

donde $X = \{x_1, \dots, x_T\}'$ e $Y = \{y_1, \dots, y_T\}'$. Como $x_t = x_t + x_{t/T} - x_{t/T}$, se tiene que

$$\begin{aligned} E(X'X) &= E[(\hat{X} - X - \hat{X})'(\hat{X} + X - \hat{X})]' \\ &= \hat{X}'\hat{X} + \sum_{t=1}^{t=T} E[(x_t - x_{t/T})'(x_t - x_{t/T})] \end{aligned} \quad (\text{A.5})$$

dado que

$$E[(X - \hat{X})'(\hat{X} - X)] = 0 \quad (\text{A.6})$$

donde $\hat{X} = (x_{t/T}, \dots, x_{t/T})'$ y $E\{(x_t - x_{t/T})'(x_t - x_{t/T})\} = P_{t/T}$. Este nuevo estimado es más eficiente que el MCO evaluado sobre $x_{t/T}$ solamente.

El procedimiento anterior corresponde a la estimación. La etapa de maximización consistirá en repetir este procedimiento hasta satisfacer un criterio de convergencia. Según Watson y Engle (1983) este método logra converger en relativamente pocos pasos, aun cuando los valores iniciales para el vector de parámetros sean pobres. Contrariamente al método de valoración, éste está libre del cálculo de cualquier derivada o de la matriz de información.