

*Construcción de un Índice Coincidente para la Economía
Colombiana: Una Aplicación de los Factores Comunes
Dinámicos*

NICOLÁS CHUDT RINCÓN
ESTADÍSTICO, MASTER(C)
CÓDIGO: 2950306475



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE ESTADÍSTICAS
BOGOTÁ, D.C.
ABRIL DE 2018

*Construcción de un Índice Coincidente para la Economía
Colombiana: Una Aplicación de los Factores Comunes
Dinámicos*

NICOLÁS CHUDT RINCÓN
ESTADÍSTICO, MASTER(C)
CÓDIGO: 2950306475

DISERTACIÓN PRESENTADA PARA OPTAR AL TÍTULO DE
MAGISTER EN CIENCIAS ESTADÍSTICAS

DIRECTOR
FABIO HUMBERTO NIETO SANCHEZ, PH.D.
DOCTOR EN ESTADÍSTICA

LÍNEA DE INVESTIGACIÓN
INVESTIGACIÓN



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE ESTADÍSTICAS
BOGOTÁ, D.C.
ABRIL DE 2018

Título en español

Construcción de un Índice Coincidente para la Economía Colombiana: Una Aplicación de los Factores Comunes Dinámicos.

Title in English

Construction of a Coincident Index for the Colombian Economy: An Application of the Common Dynamic Factors.

Resumen: El presente documento es un trabajo que busca obtener un índice coincidente, a través de los factores comunes de un conjunto de variables macroeconómicas y aplicando el enfoque de perfil coincidente basado en diferentes investigaciones. Además, se usa la prueba de Moses como herramienta para revalidar la coincidencia con el estado de la economía.

Abstract: This paper is designed to develop a proposal of coincident index for Colombian economic, using common factors from a set of macroeconomic variables and select one of them from the coincident profile approach. In addition, this study applies the Moses test in order to confirm the coincident feature of index as the closest to the state of the economy.

Palabras clave: Factores comunes dinámicos, Índice coincidente, Perfil coincidente, Prueba de Moses, Actividad Económica.

Keywords: Dynamic common factors, Coincident Index, Coincident Profile, Moses Test, Economic Activity.

Nota de aceptación

Trabajo de tesis

Aprobado

Jurado

Sergio Alejandro Calderon Villanueva

Jurado

Hector Manuel Zarate Solano

Director

Fabio Humberto Nieto Sanchez

Bogotá, D.C., Agosto 3 de 2018

Dedicado a

A mi madre, mi padre y Jacobo. A Claudia por su paciencia y amor. Una dedicación especial a mi amigo Diego Corredor donde quiera que Dios lo tenga.

Agradecimientos

Primero que todo quiero agradecer al profesor Fabio Humberto Nieto quien con su experiencia y sabiduría me acompañó y me guió en este importante proyecto académico. También quiero agradecer al grupo técnico de cuentas Nacionales Trimestrales del Departamento Administrativo Nacional de Estadística, encabezado en su momento por Giovanni Buitrago, quienes me dieron la asistencia y la disposición necesaria para el desarrollo de la investigación. Adicional, mi gratitud a la Unidad Técnica Especializada de Logística del Departamento Nacional de Planeación por su apoyo para este trabajo.

No puede dejar de lado un reconocimiento especial por los comentarios y aportes técnicos de Jovanna Elizabeth Palacios y Juan Fernando Plazas durante el desarrollo del documento. Para Natalia Arteaga, Lizet Romero y Mario Arrieta quienes de alguna manera ayudaron en el proceso de la investigación hasta su sustentación. Por último, un especial reconocimiento al profesor Norberto Rodríguez por sus comentarios y aportes a esta tesis.

Índice general

Índice general	I
Índice de tablas	III
Índice de figuras	IV
Introducción	V
1. Marco Metodológico	1
1.1. Factores Comunes de Peña y Poncela (2006)	1
1.2. Modelo de factores comunes de Lam & Yao (2012)	2
1.3. Perfil Coincidente de Martínez, Nieto y Poncela (2016)	3
1.4. Prueba de Moses	6
1.5. Seguimiento a la Economía Colombiana	8
2. Especificación de la Metodología	11
2.1. Prueba de Moses para el análisis de perfil coincidente	11
2.2. Proceso de Obtención del Índice Coincidente	12
2.3. Actualización del Índice Coincidente	16
3. Construcción del Índice Coincidente	18
3.1. Etapa 1 - Selección de Variables	18
3.2. Etapa 2 - Tratamiento de los indicadores seleccionados	23
3.3. Etapa 3 - Cálculo de los Factores Comunes	24
3.4. Etapa 4 - Perfil Coincidente y prueba de Moses	27
3.5. Etapa 5 - Obtención del ciclo	29
3.5.1. Análisis del Índice Coincidente	30
3.5.1.1. Análisis de Coyuntura	30

3.5.1.2. Contexto Regional	30
3.5.1.3. Análisis de los indicadores	31
3.6. Actualización del índice coincidente	31
Conclusiones	34
Trabajo futuro	35
Bibliografía	36
Apéndice 1: Soporte formal a la prueba de permutación	39

Índice de tablas

3.1. Conjunto de Variables.	19
3.2. Identificación de outliers.	24
3.3. Variables Transformadas.	24
3.4. Ajuste modelo estacional.	24
3.5. Valles y picos de los ciclos del ISE y del índice coincidente.	29
3.6. Crecimientos Anuales del PIB Fuente: DANE	30

Índice de figuras

1.1. Variaciones Anuales del ISE 2001-ene:2017-jun Fuente: DANE.	10
3.1. Series temporales seleccionadas.	20
3.2. Perfil coincidente de la Industria.	21
3.3. Perfil coincidente de la producción de Energía.	21
3.4. Perfil coincidente de la producción de caña.	22
3.5. Perfil coincidente del Comercio.	22
3.6. Perfil coincidente de la producción de Cemento.	23
3.7. Perfil coincidente de la Tasa de Desempleo.	23
3.8. Series Finales.	25
3.9. Comparación entre el ISE y el Factor 1 estimado.	26
3.10. Comparación entre el ISE y el Factor 2 estimado.	26
3.11. Perfil coincidente entre el ciclo del ISE y del factor 1.	27
3.12. Perfil coincidente entre el ciclo del ISE y del factor 2.	27
3.13. Boxplot de la prueba de Wilcoxon para el Índice Coincidente y el ISE	28
3.14. Ciclo del índice de seguimiento para la actividad económica de Colombia . .	29
3.15. Evolución de las variaciones anuales hasta noviembre, 2017	31
3.16. Evolución de las variaciones anuales regionales hasta noviembre, 2017. . . .	32
3.17. Perfiles coincidentes para julio, septiembre y noviembre 2017.	32

Introducción

Muchos investigadores han estudiado la manera de identificar los cambios en la tendencia de la dinámica económica, especialmente, para proveer información a los diferentes agentes económicos como los hacedores de política pública, empresas, entre otros. Desde hace un tiempo, la academia, así como las oficinas de estadística alrededor del mundo, han desarrollado diferentes estudios, enfocándose en la construcción de índices compuestos¹ que permitan explicar el comportamiento de lo que estadísticamente se conoce como el *estado de la economía*. Producto de estos estudios encontramos los *Índices Líderes* los cuales se caracterizan por anticipar el comportamiento futuro de la economía; también están los *Índices Coincidentes* que describen los patrones actuales del estado de la naturaleza y por último, los *Índices Rezagados* los cuales reflejan el comportamiento pasado de la economía. De acuerdo a lo anterior, este trabajo dará continuidad al estudio de [23] desarrollando un *Índice Coincidente* de seguimiento a la economía con una aplicación de la metodología al caso colombiano.

En varias oportunidades, el Producto Interno Bruto (PIB) ha sido uno de los índices coincidentes con mayor confiabilidad para el seguimiento de las condiciones económicas de un país; sin embargo, la consolidación de un PIB en frecuencias altas² no solamente requiere de un esfuerzo significativo para su elaboración, sino que además está sujeto a revisiones las cuales implican ajustes severos en cada reporte de las series. En otros escenarios, se ha tenido la producción industrial como variable de monitoreo sobre la tendencia económica; pero en este caso, también se presenta dificultades en la medida que los niveles de volatilidad, no permiten una coincidencia apropiada de la tendencia con miras al análisis que se busca.

Las experiencias muestran que la tarea de encontrar un índice apropiado de seguimiento a la dinámica de la economía no ha sido fácil, dado que la información requerida puede que sea oportuna pero, a la luz de la realidad, no cuenta con una amplia cobertura. Adicional, en algunos casos, los datos utilizados para la coyuntura en el análisis económico no tienen relación directa al contexto del *estado de la economía*. Sin embargo, existen trabajos importantes que han ayudado a encontrar índices que pueden contribuir a este estudio.

¹Según Guerrero en [15], los índices compuestos son útiles y necesarios porque incorporan varias variables para su construcción. Al identificar que cada una de estas variables posee una característica única en su ciclo, hace pensar que el ciclo del índice compuesto no tiene una causa única. Por ello, se hace relevante incorporar diversos indicadores que sirven como elementos descriptivos individualmente y pueden ser potenciados en un solo índice agregado.

²Los países construyen su PIB en frecuencia trimestral; es escasa la cantidad de países que construyen un PIB en frecuencia mensual y la mayoría de casos mensuales se aproximan a comportamientos del valor agregado.

Por ejemplo, en el trabajo realizado en [32], se desarrolla un índice por medio de un modelo probabilístico, encontrando un elemento común de los movimientos simultáneos de algunos agregados macroeconómicos.

Otro estudio interesante es el desarrollado en [7], el cual construye un índice líder para la economía española que permite **i.** identificar signos de recesión para las actividades económicas, **ii.** captar más preciso el mensaje del ciclo económico con solo cuatro variables, **iii.** tener la capacidad de anticipar los ciclos del PIB y **iv.** obtener mejores resultados que otras alternativas frente a los pronósticos. La metodología está sustentada en los modelos de extracción de *factores comunes*, apoyados en un modelo Dinámico de Regresión Lineal Armónico o DHLR.

Por otro lado, en el estudio realizado por [5] se analiza la idea de, no solo construir un índice, sino también de identificar la significancia estadística de este. El modelo está diseñado para *Índices Líderes* dejando de lado los *Índices Coincidentes* y *Rezagados*. Otros trabajos como el trabajo de [31], desarrollaron métodos econométricos en los cuales construyeron modelos que producen elementos comunes a través de factores dinámicos para vectores de series temporales que no requieren ser estacionarias. Para el caso colombiano, está el trabajo en [23] el cual incorpora el método desarrollo en [5] pero enfocando el análisis de *Perfil Líder* a *Perfil Coincidente* y apoyado en el trabajo realizado en [31] para construir un *Índice Coincidente Dinámico*.

Adicional a los trabajos ya mencionados, existen otros autores quienes han desarrollado investigaciones también relacionadas al *Índice Coincidente*, entre los que se destacan los trabajos mostrados en [1] y [4] quienes construyen un índice para la medición en tiempo real del estado actual de la economía en la Euro-zona. El método utilizado para la obtención de cada índice es un filtro de *Band-pass*. Complementando un poco lo anterior, están los estudios de [12] el cual crea un índice para los 50 estados de los Estados Unidos y, para el caso colombiano, los trabajos de [27] y [24].

El objetivo principal de esta investigación es extender el estudio de [23], el cual se basa en la selección de un factor común que represente de la manera más precisa y coincidente el estado de la naturaleza. La idea es involucrar un nuevo conjunto de variables relacionadas a las actividades económicas de Colombia, aplicar de nuevo el modelo de factores comunes de [31] pero ahora complementándolo con [30], formalizar la *Prueba de Permutación* utilizada en [23], incluir el análisis de la *Prueba de Moses* desarrollada en [26] y aplicada en [5] para complementar el análisis de perfil coincidente, y plantear un mecanismo recurrente para computar el índice, en un periodo de tiempo actual, garantizando la solidez y evitando la re-estimación de sus valores pasados.

El trabajo se divide en los siguientes capítulos: el primero es una descripción del marco teórico donde se muestran los métodos de [31], el trabajo de [30], el desarrollo metodológico de [23] y la prueba de Moses. También se realiza un pequeño análisis de la economía colombiana en los últimos años. En el segundo capítulo se explica la metodología aplicada a la obtención de un índice coincidente donde se plantea una nueva concepción de la prueba de Moses y una nueva propuesta al procedimiento hecho en [23]. El tercer capítulo es un ejercicio aplicado a la economía colombiana. Por último se plantean unas conclusiones y recomendaciones sobre el estudio

Marco Metodológico

1.1. Factores Comunes de Peña y Poncela (2006)

El objetivo principal del estudio en [31] se basa en construir un modelo de factores dinámicos para series multivariadas (Estacionarias o no-estacionarias). El método identifica el número de factores comunes a través de la estructura de valores propios de la matriz muestral de autocovarianzas generalizadas (SGCM por sus siglas en inglés) y de las matrices de Correlación Canónica.

El método supone un proceso multivariado $\{y_t\}$ m -dimensional, el cual es generado por un conjunto de $r < m$ factores comunes no observados, vía el modelo

$$y_t = Pf_t + e_t, \quad t \in \mathbb{Z}, \quad (1.1)$$

donde \mathbb{Z} denota el conjunto de números enteros, f_t representa el vector de los factores comunes en t de dimensión r , P es la matriz $m \times r$ de pesos de los factores comunes y $\{e_t\}$ es un ruido blanco multivariado Gaussiano con media 0 y una matriz diagonal de covarianzas de rango completo Σ_e . Por otro lado, el proceso $\{f_t\}$ sigue un vector autoregresivo de media móvil **VARMA(p,q)**,

$$\Phi(B)f_t = a + \theta(B)\nu_t, \quad (1.2)$$

donde $\Phi(B) = I + \Phi_1 B + \Phi_2 B^2 + \dots + \Phi_p B^p$ y $\Theta(B) = I + \Theta_1 B + \Theta_2 B^2 + \dots + \Theta_q B^q$ con B el operador de rezago y con las raíces de los polinomios, tanto del lado auto-regresivo como del lado media móvil, fuera o sobre el círculo unitario. Así mismo, a es un vector de constantes $r \times 1$ y $\{\nu_t\}$ es un ruido blanco multivariado Gaussiano con media 0 y una matriz diagonal de covarianzas de rango completo Σ_ν .

La metodología estadística para determinar el número de factores comunes r y estimarlos en cada tiempo t , se puede consultar en los artículos de [31] y de [30], entre muchos otros. Siguiendo a [31], la especificación del número total r de factores comunes, se hace a través de una prueba estadística basada en los valores propios de las matrices de autocorrelación canónica. La idea es: la matriz de correlación canónica de rezago k está dada

por

$$\hat{M}_1(k, N) = \left[\sum_{t=k+1}^N y_t y_t^T \right]^{-1} \sum_{t=k+1}^N y_t y_{t-k}^T \left[\sum_{t=k+1}^N y_{t-k} y_{t-k}^T \right]^{-1} \sum_{t=k+1}^N y_{t-k} y_t^T,$$

para cada $k = 1, \dots, N - 1$, donde N denota el tamaño de muestra. Sea

$$S_{m-r,k} = -(N - k) \sum_{j=1}^{m-r} \ln(1 - \hat{\lambda}_j),$$

donde $\hat{\lambda}_1 \leq \dots \leq \hat{\lambda}_m$ son los valores propios ordenados de $M(k, N)$. Entonces, $S_{m-r,k}$ es una estadística de prueba para examinar la hipótesis nula de que a lo más $m - r$ valores propios, de la matriz límite de la sucesión de matrices $M(k, N)$ cuando N tiende a ∞ , son iguales a cero. La distribución límite de $S_{m-r,k}$ bajo esta hipótesis nula es $\chi_{(m-r)^2}^2$, para todo $k = 1, 2, \dots$.

En el trabajo de [31], además de determinar el número total de factores comunes, también identifican el número de factores comunes no-estacionarios, a través de la matriz muestral de autocovarianzas generalizadas (SGCM), dadas por

$$C_y(k, N) = \frac{1}{N^{2d+d'}} \sum_{t=k+1}^N (y_{t-k} - \bar{y})(y_t - \bar{y})', \quad k = 0, 1, \dots, N - 1 \quad (1.3)$$

donde $\bar{y} = \frac{1}{N} \sum_{t=1}^N y_t$, d denota el grado de integración de los factores comunes no estacionarios y d' puede tomar el valor de 0 ó 1¹.

Por último, las ecuaciones 1.1 y 1.2 pueden ser transformadas en un *modelo Espacio-Estado*, lo que permite estimar los parámetros con el método de *Máxima Verosimilitud* y, usando *Filtro de Kalman*, se pueden estimar los factores comunes.

1.2. Modelo de factores comunes de Lam & Yao (2012)

La metodología diseñada en [23] utiliza el método de [31] para la obtención del número de factores comunes identificado. Como fue visto en la sección 1.1, se definió $\{y_t\}$ como un proceso estocástico observable de dimensión m , el cual es generado por un proceso latente r -dimensional $\{f_t\}$, con $r \leq m$, como es descrito en la ecuación (1.1). Los procesos componentes de $\{f_t = (f_{1t}, \dots, f_{rt})'\}$, donde el símbolo $'$ significa transposición de matrices, se llaman los factores comunes y ellos pueden ser no estacionarios o estacionarios. En este trabajo se asumirá además que no son estacionales (para un estudio de factores comunes estacionales se puede consultar el trabajo de [29]). También fue definida la matriz muestral de autocovarianzas generalizada - SGCM ($C_y(k)$) y la matriz de correlación canónica ($\hat{M}_1(k, N)$); sin embargo, también existen otras metodologías para complementar el cálculo.

En este trabajo se va a utilizar como criterio para determinar el número de factores comunes, un procedimiento análogo al descrito en [30]. Esencialmente, este consiste en lo siguiente: dada la matriz $C(k, N)$, para $k_0 \geq 1$ conocido, sean

¹El criterio para la definición de los valores d' está sujeto a la inclusión del *drift*. Para mayor profundidad revisar los teoremas 1 y 2 de artículo [31].

$$M = \sum_{k=1}^{k_0} C(k, N)C(k, N)' , \quad (1.4)$$

y $\lambda_1 \geq \dots \geq \lambda_p$ los valores propios de M ordenados en forma descendente de magnitud. Entonces, un estimador de r está dado por

$$\hat{r} = \operatorname{argmin}\{\lambda_{i+1}/\lambda_i | 1 \leq i \leq R\} ,$$

donde R es un valor conocido tal que $r < R < m$. En la práctica, para un valor de m moderadamente grande, se toma $R = [m/2]$, donde $[*]$ se describe como la parte entera de un número real.

1.3. Perfil Coincidente de Martínez, Nieto y Poncela (2016)

El trabajo para la construcción de un índice coincidente en [23], además de adoptar la metodología de factores comunes creada en [31], también involucra el método de identificación del *Perfil Coincidente* a partir del *Perfil Líder* hecho en [5]. El método se basa en una prueba de permutaciones la cual fue desarrollada por Fischer en [14] y lleva el nombre de *Prueba de Permutación para Datos Pareados*. Para el caso del método descrito en [23], la prueba consiste inicialmente en calcular las *diferencias* entre los *puntos de quiebre* de una serie de referencia de datos no observados, C_t (la cual representa el *estado de la naturaleza*), y los *puntos de quiebre* de una serie observada, I_t , que sirve para el seguimiento de la economía, es decir, para el seguimiento de C_t .

Siguiendo el enfoque de Martínez, Nieto y Poncela en [23], pero precisando los conceptos y formalizando el método estadístico utilizado por esos autores, se presenta a continuación el enfoque del perfil coincidente. Los planteamientos aquí expuestos extienden su método y son, por lo tanto, una contribución teórica y metodológica al tema de construcción de índices coincidentes. En particular, de actividad económica y, por esta razón, a partir de ahora solo nos referiremos al contexto económico.

Definición 1. Sean $\{C_t\}$ e $\{I_t\}$ procesos estocásticos que representan, respectivamente, el estado de la economía y un estimador de $\{C_t\}$. Se dice que $\{I_t\}$ es l -coincidente con $\{C_t\}$ para algún $l \in \{0, \pm 1, \pm 2, \dots, \pm c\}$, con c número entero positivo conocido, si para cada trayectoria de $\{C_t\}$ e $\{I_t\}$, cada uno de los puntos de giro de la trayectoria de $\{I_t\}$ está rezagado o adelantado con respecto a cada uno de los correspondientes de la trayectoria de $\{C_t\}$, por $|l|$ periodos de tiempo.

Observaciones: 1. Estrictamente, los puntos de giro (o de inflexión) hacen referencia a valores máximos y mínimos locales de las trayectorias bajo consideración.

2. Si $l < 0$, se tiene un rezago. Si $l > 0$, se tiene un adelanto. El caso $l = 0$ será interpretado como coincidencia exacta o, simplemente, coincidencia.

3. En economía, y para variables mensuales, usualmente es razonable tomar $c \leq 3$. Este valor de c ha sido utilizado antes en la literatura, por ejemplo, por [24].

De acuerdo con la definición anterior, el objetivo es tener un índice 0-coincidente para $\{C_t\}$. Sin embargo, un índice l -coincidente con l muy cerca a cero podría ser muy útil para estimar el estado actual de la economía.

Para determinar el valor de l , se propone juzgar la hipótesis nula $H : l = l_0$ v.s. $K : l \neq l_0$, para l_0 un valor conocido en $\{0, \pm 1, \dots, \pm c\}$. Inicialmente, se supone $l_0 = 0$. Sean $\{c_t\}$ e $\{i_t\}$ series de tiempo (trayectorias) de $\{C_t\}$ e $\{I_t\}$, respectivamente, en el período muestral $1, \dots, N$. Bajo coincidencia exacta, habría Q puntos de giro en los puntos temporales t_1, \dots, t_Q , para algún $Q \geq 1$, en las series de tiempo $\{c_t\}$ e $\{i_t\}$. Pero, en virtud de la variabilidad muestral, los puntos de giro en $\{i_t\}$ podrían ocurrir en t'_1, \dots, t'_Q . Así, se tiene un pareamiento natural de los puntos de giro, digamos $\{(t_j, t'_j) : j = 1, \dots, Q\}$. Es de esperar, que $d_j = t_j - t'_j \approx 0$ para todo $j = 1, \dots, Q$ y, en consecuencia, $\sum_{j=1}^Q d_j \approx 0$. Obviamente, esta suma puede ser positiva o negativa.

Con el objeto de construir una estadística que permita decidir si la suma observada es un valor extremo y, en consecuencia, tener evidencia empírica para rechazar la hipótesis nula, se considera el siguiente procedimiento estadístico basado en las pruebas de permutación (un marco teórico formal aparece en el Apéndice).

Sean d_1, \dots, d_Q las diferencias observadas en los puntos de giro de las series de tiempo $\{c_t\}$ e $\{i_t\}$ y sea s_0 su suma. Las diferencias son reorganizadas de la siguiente forma: inicialmente, se considera el valor absoluto de las mismas, esto es, $|d_1|, \dots, |d_Q|$. Luego, se obtienen las 2^Q permutaciones de longitud Q de los signos $+, -$ y cada una de ellas es asignada a la sucesión $|d_1|, \dots, |d_Q|$. A manera de ejemplo, una de las sucesiones de signos es $+, +, \dots, +$, otra $-, -, \dots, -$ y otra la de las diferencias observadas mismas.

Ahora, sea s_j el valor de la suma para la j -ésima sucesión signada de las diferencias absolutas, $j = 1, \dots, Q$. Entonces, la suma observada coincide con s_{j_0} para algún $j_0 = 1, \dots, Q$. Sea

$$p_0 = \frac{\#\{s_j : |s_j| \geq |s_{j_0}|, j = 1, \dots, Q\}}{2^Q},$$

donde “#” denota el cardinal del conjunto en referencia. Este número es el p -valor de la sucesión signada observada, bajo la distribución de permutación para la variable aleatoria S , cuyas realizaciones son todas las posibles sumas consideradas en las permutaciones.

Al nivel de significancia α , $0 < \alpha < 1$, la regla de decisión es rechazar la hipótesis nula y aceptar la alterna si $p_0 \leq \alpha$. Un rechazo de la hipótesis nula significa que s_{j_0} es un valor extremo de S bajo su distribución de permutación o que la sucesión observada d_1, \dots, d_Q es un resultado extremo bajo una distribución equivalente. Estos hechos están formalizados en el Apéndice.

Suponiendo que H es rechazada, significa que $l_0 \neq 0$, luego $l_0 < 0$ ó $l_0 > 0$ y el interés ahora es identificar el valor de l_0 . Con ese fin, se consideran las hipótesis condicionales sucesivas

$$H^l : l_0 = l | H^{l'} \text{ ha sido rechazada}$$

v.s.

$$K^l : l_0 \neq l | H^{l'} \text{ ha sido rechazada,}$$

donde $l = -l'$ si $l' > 0$ o $l = -l' + 1$ si $l' < 0$. Nótese que $l \in \{\pm 1, \dots, \pm c\}$.

Equivalentemente, las hipótesis anteriores pueden ser replanteadas como

$$H^l : l_0 - l = 0 | H^{l'} \text{ ha sido rechazada}$$

v.s.

$K^l : l_0 - l \neq 0 | H^l$ ha sido rechazada,

con l como antes.

Dado $l \in \{\pm 1, \dots, \pm c\}$, se propone como estadística de prueba a $S^l = \sum_{i=1}^Q d_i^l$, con $d_i^l = d_i - l$, $i = 1, \dots, Q$. Luego, reorganizando las diferencias d_i^l , se procede como en el caso $l = 0$, definiendo

$$p_l = \frac{\#\{s_j^l : |s_j^l| \geq |s_{j_0}^l|, j = 1, \dots, Q\}}{2^Q},$$

donde s_j^l denota la suma de la j -ésima sucesión signada de $d_1 - l, \dots, d_Q - l$.

Definición 2. Al $\{p_l : l = 0, \pm 1, \dots, \pm c\}$ se le denomina el *perfil coincidente* de la serie de tiempo $\{i_t\}$ con respecto a $\{c_t\}$.

Definición 3. Se dice que los procesos $\{C_t\}$ e $\{I_t\}$ son l_m -coincidentes al nivel α , $0 < \alpha < 1$, si $l_m = \operatorname{argmax}\{p_l : l = 0, \pm 1, \dots, \pm c\}$ y $p_{l_m} > \alpha$.

Un gráfico cartesiano de $\{p_l : l = 0, \pm 1, \dots, \pm c\}$, donde l es la abcisa, ilustra visualmente el significado y utilidad en la práctica del perfil coincidente. Nótese que para l dado en el $\{0, \pm 1, \dots, \pm c\}$, si $p_l \geq \alpha$, entonces la hipótesis nula H^l no es rechazada. En este caso se puede interpretar p_l como un nivel de confianza para un intervalo aleatorio que incluya a l .

En la vida real, la serie de tiempo $\{c_t\}$ no es observada y esto es un inconveniente para hallar un perfil coincidente. Para dar una solución a este problema, en el trabajo de [23] proponen utilizar una variable observable que permita rastrear la dinámica del estado de la economía $\{C_t\}$. Esa serie de tiempo se considera una estimación preliminar del estado de la economía, en el período muestral considerado, y con ella se construye el perfil coincidente para un índice propuesto.

Finalmente, el procedimiento en [23] para obtener uno o varios factores comunes, que reflejen el *estado de la naturaleza*, no incorpora esta formalización del perfil coincidente sino que describe la prueba y continua con un proceso sistémico que seguía las siguientes etapas:

ETAPA 1. Desestacionalizar las series de tiempo cuando todas o algunas de ellas sean estacionales y ajustar por datos atípicos e intervenciones si es del caso. Transformaciones para ajustar por escala (incluyendo logaritmos) también son apropiadas.

ETAPA 2. Determinar el número de factores comunes de las variables utilizadas y obtener sus estimaciones, usando la metodología de [31] o la de [30].

ETAPA 3. En el entendido que el grado de coincidencia entre dos procesos lo determina el valor l_m , se obtienen los perfiles coincidentes al nivel α de cada factor común. Luego se escoge el factor común con grado de coincidencia más cercano a cero y mayor p -valor.

ETAPA 4. Asignar la base del índice y obtener los puntos de giro del mismo, con el fin de determinar el ciclo económico.

En definitiva, y volviendo a lo mencionado previamente, con la formalización de la prueba de perfil coincidente expuesta en este capítulo, se busca extender el método como

parte de la contribución teórica y metodológica al contexto de los índices coincidentes, principalmente, de la actividad económica.

1.4. Prueba de Moses

Las pruebas estadísticas que están basadas en contrastar medias muestrales para dos variables, no son útiles para dar conclusiones sobre sus poblaciones cuando sus valores extremos tienen grandes diferencias. No obstante, existen pruebas que permiten distinguir si dos muestras pertenecen a la misma población, a partir de la evaluación de la dispersión de sus datos. En particular, se tiene la prueba de Moses la cual es un método no-paramétrico para estudiar si hay diferencias en el grado de dispersión entre un grupo control y otro grupo experimental de datos *ordenados*.

Esta prueba fue empleada por [5] como parte de su análisis sobre los puntos de giro, con el fin de evaluar los ciclos económicos. Dado los anticipos calculados para dos índices líderes (uno control y otro experimental) potenciales, los cuales poseen las mismas características², la prueba busca detectar si existen diferencias poblacionales. Después que los puntos de quiebre son *ranqueados* y mezclados en una sola muestra, se calcula un estadístico el cual mide el grado de concentración de los anticipos del índice control. La hipótesis nula verifica si los anticipos de los dos índices están mezclados para determinar que no existe concentración por parte de alguno de los grupos. De lo contrario, los anticipos del índice líder experimental pueden estar más concentrados hacia los meses iniciales o hacia los meses finales del periodo de estudio en relación a los anticipos del índice control³. A continuación se describe con mayor precisión la prueba de Moses.

Definición 4: Sean C y E variables aleatorias discretas que representan los tiempos de ocurrencia en meses (o en trimestres) de los anticipos del índice líder control y los tiempos de ocurrencia en meses (o en trimestres) de los anticipos del índice líder experimental, respectivamente. Sean n_C el número de realizaciones de la variable aleatoria C y n_E el número de realizaciones de la variable aleatoria E . $n_C + n_E$ es el número de realizaciones de ambas variables, ordenadas de forma ascendente, con un rango asignado a cada realización de la nueva muestra incrementada⁴.

Si la hipótesis nula, H_0 , afirma que los tiempos de ocurrencia, tanto en C como en E , vienen de una misma población, entonces estos tiempos de ocurrencia están “mezclados” en el nuevo arreglo ordenado. Para el caso de la hipótesis alterna, H_A , si los anticipos de E están más dispersos entonces H_0 se rechaza en favor de H_A bajo las siguientes situaciones:

- la mayoría de realizaciones ordenadas de E con valores menores tenderían a estar ubicados en el extremo izquierdo (al inicio) de la escala, ó
- la mayoría de realizaciones ordenadas de E con valores mayores tenderían a estar ubicados en el extremo derecho (al final) de la escala, ó

²Los dos conjuntos de anticipos tienen como característica que no tienen diferencias significativas relacionadas a su tendencia central lo que, inicialmente, les permite anticipar los comportamientos de ciclos económico de la misma manera.

³El planteamiento de las hipótesis está relacionado a los casos de estudio en Psicología sobre *Comportamientos defensivos* los cuales pueden caracterizarse por puntajes muy grandes o puntajes pequeños. Es a lo que en el estudio [26] referencia como **reacciones extremas**

⁴Para los empates se asocia el rango promedio.

- c. un número razonable de realizaciones ordenadas de E con valores menores estarían ubicados en el extremo izquierdo de la escala y otro número razonable de realizaciones ordenadas de E con valores mayores estarían ubicados en el extremo derecho de la escala.

Desde otra perspectiva, [26] plantea que si los tiempos de ocurrencia en el grupo experimental están ubicados de acuerdo a lo indicado en la hipótesis alterna, entonces los puntos del grupo control estarán muy concentrados a la derecha (para el caso a.), o muy concentrados a la izquierda (para el caso b.) o en el medio (para el caso c.). Por tanto, la prueba determina si los puntos de ocurrencia en C están lo suficientemente concentrados en algún extremo como para rechazar H_0 en favor de H_A .

La construcción de una prueba se basa esencialmente en encontrar su *estadístico* y su respectiva *distribución* bajo la aceptación de la hipótesis nula. Para encontrar el *estadístico*, en el presente contexto, se calcula un rango s' dentro de la muestra ordenada de *rankings* la cual tiene como condición abarcar todas las realizaciones del índice C . Este rango es calculado tomando el mayor rango del índice C menos el menor rango del mismo índice C más 1. Este rango se interpreta como una medida de concentración pues describe el número más pequeño de rangos consecutivos necesarios para incluir todas las realizaciones del índice C . s' será considerado, entonces, como una propuesta inicial de estadístico de la prueba. Considérese el siguiente caso hipotético: sea $n_C = 9$, $n_E = 8$, con los puntos de ocurrencia dados en la siguiente tabla:

C		E	
jul-00	may-01	mar-00	ene-02
ago-00	jun-01	may-00	abr-02
sep-00	sep-01	jun-00	nov-02
dic-00	feb-02	dic-01	sep-04
	mar-02		

La unión de los dos conjuntos de realizaciones forma una nueva muestra ordenada con $n_C + n_E = 17$ dada por la siguiente estructura:

Mes	mar-00	may-00	jun-00	jul-00	ago-00	sep-00	dic-00	may-01	jun-01	sep-01	dic-01	ene-02	feb-02	mar-02	abr-02	nov-02	sep-04
Índice	E	E	E	C	C	C	C	C	C	C	E	E	C	C	E	E	E
Rango	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

Acá la primera fila son los puntos de ocurrencia tanto del índice C como del índice E ordenados de modo ascendente; la segunda fila corresponde a la variable aleatoria al que pertenecen (C ó E) y la tercera fila corresponde al rango asignado para cada elemento. Como el menor valor en C es el punto **julio de 2000 (jul-00)**, ranqueado en la posición 4, y el mayor valor es el punto **marzo de 2002 (mar-02)**, ranqueado en la posición 14, entonces $s' = 11$, que es el menor valor de rangos consecutivos necesarios para incluir todos las realizaciones del índice C .

Puesto que s' depende fuertemente de los valores extremos del índice control, para Moses, este estadístico es inestable⁵. En el trabajo de [26] se sugiere calcular una nueva *amplitud* la cual es denomina la *rango recortado* t_h . La idea es encontrar una nueva

⁵Según Moses, el principal problema del rango es que depende de los valores extremos del grupo denotado como “control”. Esta condición puede generar dos inconvenientes: el primero es que los valores del grupo “control” puedan estar contaminados de elementos con propiedades del grupo “experimental”. Ello induciría resultados erróneos por culpa de estos elementos. Segundo, ante un tamaño grande del grupo control, el rango calculado no sería un índice eficiente para medir el nivel de concentración de los elementos. Para mayor detalles, revisar el estudio de [26].

amplitud la cual se obtiene contando los n_C puntos de ocurrencia después de omitir h valores del grupo control por arriba y por debajo, para $h \in \{1, 2\}$. En el ejemplo anterior, si $h = 2$ entonces el rango recortado $t_2 = 10 - 6 + 1 = 5$, donde 10 es el mayor rango asignado después de omitir los dos mayores valores de C y 6 es el menor rango después de omitir los dos menores valores de C . En este caso, el estadístico t_h ha sido definido como el menor número de rangos consecutivos necesarios para incluir todas las realizaciones de C , omitiendo su(s) h mínima(s) y su(s) h máxima(s) realizaciones y, como caso general, manteniendo al menos las $n_C - 2h$ restantes.

Por lo anterior, si t_h es adoptado como el *estadístico* de prueba, entonces es importante identificar cual es su probabilidad asociada. Dado que t_h satisface la condición,

$$n_C - 2h \leq t_h \leq n_E + n_C - 2h,$$

entonces, el problema es determinar un valor para t_h más grande que $n_C - 2h$ pero, en realidad, manteniendo su magnitud pequeña. Por ello, se busca un valor de g , entero positivo, el cual determina el nivel de significancia α , en el sentido,

$$P\{t_h \leq n_C - 2h + g\} \leq \alpha. \quad (1.5)$$

Finalmente, dado un valor de g donde $t_h > n_C - 2h$, el criterio será rechazar la hipótesis nula cuando $t_h \leq n_C - 2h + g$. Por lo tanto, la distribución de t_h se obtiene a partir de,

$$P_{(t_h \leq n_C - 2h + g)} = \frac{\sum_{i=0}^g \binom{i+n_C-2h-2}{i} \binom{n_E+2h+1-i}{n_E-i}}{\binom{n_C+n_E}{n_C}}. \quad (1.6)$$

Si el valor de $p_{(t_h \leq n_C - 2h + g)}$ es igual o menor que el nivel de significancia α , $0 < \alpha < 1$, entonces la hipótesis nula es rechazada en favor de la hipótesis alterna. El rechazo de la hipótesis nula es confirmar que los puntos de quiebre del índice experimental tienen mayor dispersión que los del índice control, es decir, que existe evidencia estadística suficiente para encontrar diferencias significativas entre los ciclos del índice control y el índice experimental. Luego, debido a la menor dispersión, se escogerá el índice control como el índice líder seleccionado.

1.5. Seguimiento a la Economía Colombiana

Aunque los esfuerzos del documento se enfocan en la descripción metodológica para obtener un índice coincidente, es importante también explicar el contexto de la economía colombiana, en el periodo de estudio. Este entorno económico no solo se hace con miras a caracterizar el *estado de la naturaleza*, sino que también busca responder a la justificación de los indicadores seleccionados que van a emplearse en el proceso de construcción del índice coincidente. El comportamiento de la actividad productiva nacional ha obedecido a factores tanto exógenos como propios de su naturaleza; el alto nivel de dependencia de la economía con el acontecer internacional, los factores de la política interna, la relevancia del consumo de los hogares y la dinámica de los capitales nacional y extranjeros, por

⁶Detalles precisos de esta expresión se pueden encontrar en el anexo de [26].

mencionar algunos, son determinantes a la hora de analizar la tendencia productiva del país.

Para la economía colombiana, la década del sesenta se caracterizó por concentrar la política pública en la generación de empleo a través de la estimulación del sector industrial, el cual se pensaba como un actor trascendental para impulsar la economía doméstica. El gobierno promovió el plan Vallejo el cual fue una adaptación del *modelo cepalino* que buscaba motivar a la industria nacional para producir los bienes y servicios importados y llegar a un equilibrio en la balanza de pagos.

Los años setenta marcaron algunos hitos en la economía nacional: en 1975 las heladas del principal exportados mundial de café, Brasil, disparó sus precios lo que indujo a incrementar la producción del grano en Colombia, trayendo consigo una bonanza de precios. A pesar de esto, el plan vallejo no estaba dando los dividendos que se esperaban lo que suscitaba no solo cuestionamientos al modelo sino al planteamientos de nuevas alternativas frente a la política económica.

Posteriormente, los desordenes fiscales de los países latinoamericanos y la poca eficiencia del modelo cepalino desencadenan, para la década de los ochenta, la crisis de la deuda. Colombia empezó un programa de reajuste fiscal para no entrar en cesación de pagos y empieza una reestructuración de su deuda[25]. Adicional, el gobierno incorporó algunas reformas fiscales como la creación de la ley 9 de 1983 la cual introduce por primera vez a la economía el Impuesto al Valor Agregado - IVA[17].

Para la década de los noventa tuvieron lugar las primeras señales de liberación de la economía colombiana. Aunque desde la década anterior se había pensado en políticas que bajaran gradualmente los aranceles, estas iniciativas fueron postergadas por el descubrimiento de los yacimientos Cusiana y Cupiagua; estos *Booms petroleros* permitieron subsanar, en cierta medida, algunos desequilibrios fiscales del momento. Con la llegada de la *Apertura* llegaron varios eventos: gracias al descubrimiento de los yacimientos petroleros y las privatizaciones, incrementaron los ingresos de capital extranjero y sumado a la liberación de importaciones empezaron los efectos sobre la dinámica en la década([3], [23] y [11]).

El incremento en la inversión extranjera directa (IED) así como la entrada en vigencia de políticas de ajuste en el marco fiscal, en el mercado laboral, en el régimen pensional y las reformas tributarias, también incidieron en el crecimiento económico, lo que sumado a la revaluación de la tasa de cambio fueron características predominantes de los años 90. Para los siguientes años, el país experimentó una fuerte desaceleración económica por cuenta de una de las más fuertes crisis financieras de su historia, la cual trajo altas restricción de recursos externos de inversión y altas tasas de endeudamiento generando un déficit de cuenta corriente financiado con reservas internacionales ([23], [25] y [11]).

Para inicios del presente siglo, la dinámica del gasto público se vio impulsada por la política de seguridad democrática [2] donde se destacó el incremento en los rubros de defensa y seguridad, estimulando las cuentas fiscales de la nación; así mismo, esta política aumentó los flujos de inversión extranjera y activó sectores como la banca nacional y la explotación de hidrocarburos [11] generando un panorama expansionista en el país. Sin embargo, para 2008 el mundo enfrentaba una de las más complejas crisis económicas con la bancarrota del banco de inversiones Lehman Brothers y sus pérdidas por el mal manejo de los fondos subprime (créditos hipotecarios con mayor riesgo respecto al resto de los créditos); en el medio local hubo reducción en sus tasas de crecimiento lo que impactó en la

emisión de créditos y el aumento del desempleo, principalmente, en el sector real. Posterior a 2015, la economía del país ha tenido tasas de crecimiento moderadas acompañadas de programas macroeconómicos mucho más ajustados producto de la alarmante baja en el precio internacional del crudo y de otros hidrocarburos.

A pesar de la contextualización descrita, es difícil identificar los determinantes de las fluctuaciones del ciclo económico a través de una relación estrecha entre los hechos estilizados de la economía y las tasa de crecimiento de la actividad productiva. No obstante, se cuenta con índices que miden la actividad económica y su tendencia en el corto plazo. Tal es el caso del Indicador de Seguimiento a la Economía - **ISE**, publicado por el DANE, el cual está diseñado para medir la evolución de la economía a partir de un conjunto de indicadores mensuales que representan las diferentes actividades económicas. Los detalles de su construcción se pueden consultar en el sitio web <http://www.dane.gov.co/> [13].



FIGURA 1.1. Variaciones Anuales del ISE 2001-ene:2017-jun **Fuente:** DANE.

A partir de las cifras del **ISE**, también pueden indentificarse algunos patrones de la economía; tal como lo sugieren las cifras disponibles de este indicador en la Figura 1.1, la evolución de la tasa de crecimiento anual entre el periodo enero 2000 - noviembre 2017 muestra diferentes escenarios de la actividad económica; es de resaltar que en los primeros años, a partir del 2001, el ritmo de la tasa de crecimiento fue moderado debido a factores muy influyentes como la situación del mercado internacional, los precios relativos de los *commodities*, entre otros.

Hasta 2007, la economía del país alcanzó valores cercanos al 8%, lo cual puede atribuirse a un buen comportamiento de la demanda externa ante precios muy favorables de productos nacionales lo cual beneficio la balanza comercial. Después de 2008, se experimentaron algunas tasas menores; la Figura 1.1 evidencia una fase recesiva en este periodo la cual no es muy prolongada pero, sin duda, puede estar relacionada al shock de la crisis internacional. Para años posteriores a 2015 los hechos de coyuntura económica están enmarcados por escenarios de tipo político y económicos como las reformas tributarias y de tipo internacional como las dinámicas comerciales con Estados Unidos y los precios de los *commodities*.

La descripción de estos hechos económicos tienen como objetivo mostrar el comportamiento de la economía colombiana. Sin embargo, con esto, también se espera que la construcción del nuevo índice sea útil no solamente para alcanzar un índice robusto con la metodología de este documento sino que el resultado del nuevo índice responda a los hechos estilizados de la economía colombiana descritos en este capítulo y los hechos que a futuro se desarrollen.

Especificación de la Metodología

2.1. Prueba de Moses para el análisis de perfil coincidente

Como fue mencionado en el capítulo 1, la prueba de Moses es el método adaptado en [5] para el análisis de los puntos de quiebre y, en consecuencia, de los ciclos económicos. Es importante destacar que en el estudio de [5], la prueba de permutación de datos pareados es la herramienta que permite calcular el perfil líder para los factores calculados mientras que la prueba de Moses es un ejercicio complementario al análisis de los índices líderes obtenidos. Por lo anterior, la prueba de Moses no es un método para definir si el(os) factor(es) es(son) líder(es). Esta prueba tiene como finalidad medir el desempeño del índice líder a través de la dispersión de los respectivos puntos de quiebre.

Para [5], los índices líderes son una herramienta poderosa para estimar los ciclos del crecimiento. Sin embargo, estos índices carecen de evaluaciones estadísticas profundas y detalladas¹. Por ello, existen dos aspectos que deben ser explorados para la verificación de los ciclos: el primero es la identificación de los puntos de quiebre para medir el desempeño de los índices; en particular, el método realizado por [9] hace posible evaluar el desempeño en términos de calcular los puntos de quiebre y la duración de los ciclos asociados. El segundo aspecto está relacionado al número de ciclos en un periodo de tiempo; sin embargo, es muy común obtener índices que incluyen pocos ciclos lo que obliga que para su estudio sean usados métodos no-paramétricos para evaluar el desempeño.

La propuesta de [5] se basa en dos pruebas no-paramétricas para evaluar los ciclos de los índices: **1)** una gráfica de los puntos de quiebre, en términos de probabilidad, para identificar el perfil líder² y poder seleccionarlo y **2)** haciendo inferencia estadística, se evalúa la dispersión de los puntos de quiebre de dos índices líderes con la prueba de Moses. Con estas herramientas se busca estudiar adecuadamente los ciclos de la actividad económica a través de las diferencias entre los puntos de quiebre de dos índices y, posteriormente, con sus respectivas dispersiones.

Contrario al tratamiento de [5], de comparar los índices candidatos a ser *índices líderes* para seleccionar uno de ellos, el presente estudio ha enfocado la prueba de Moses de una

¹Al respecto, Banerji destaca la fuerza que ha tomado la prueba de Causalidad de Granger pero resalta, también la dificultad de evaluar estos índices; esta misma problema también la destaca el trabajo de [16]

²Para el caso del índice coincidente se tiene el estudio de [23].

forma diferente dentro del análisis de perfil coincidente. Dado que tanto el índice propuesto como el índice preliminar son índices que hacen seguimiento, de manera coincidente, al estado de la naturaleza, en la hipótesis nula, los tiempos de ocurrencia calculados para el índice propuesto y los calculados para el índice preliminar tienen el mismo desempeño y, por lo tanto, pertenecen a la misma población. Por consiguiente, los dos índices rastrean de modo simultáneo el estado de la economía. En la hipótesis alterna, los puntos de quiebre del índice propuesto no hacen parte de la misma población que los puntos de quiebre del índice preliminar. Por lo tanto, dada su coincidencia, los ciclos del índice propuesto no tendrían el mismo desempeño que los ciclos del índice preliminar.

La presente investigación mantiene el mismo proceso de cálculo para la prueba de Moses pero las interpretaciones del investigador están enfocadas a no rechazar la hipótesis nula: se tiene en cuenta el número de puntos de ocurrencia del índice preliminar $\{C_t\}$, n_C , y el número de puntos de ocurrencia de un índice propuesto $\{I_t\}$, n_I . La hipótesis nula plantea que los puntos de ocurrencia tanto de $\{C_t\}$ como de $\{I_t\}$, vienen de la misma población. Para la hipótesis alterna, los puntos de ocurrencia del índice $\{I_t\}$ estarán más concentrados, sea al comienzo, al final o distribuidos en los extremos del periodo muestral con relación a los puntos de ocurrencia de $\{C_t\}$. El cálculo del *Estadístico* y de su *distribución* están descritos en la sección 1.4 del capítulo 1.

2.2. Proceso de Obtención del Índice Coincidente

Habiendo precisado conceptos, formalizado el método estadístico de [23] en la sección 1.3 y explicado la aplicación de la prueba de Moses, la nueva metodología para la construcción del índice coincidente consiste en desarrollar las mismas etapas planteadas inicialmente en [23] pero con modificaciones en algunas de ellas. La propuesta es:

ETAPA 1

Una primera fase, es la exploración de un conjunto de indicadores económicos que sirvan como materia prima para la construcción del *índice coincidente*. Dentro del conjunto de indicadores seleccionados, se tendrán series económicas con periodicidad mensual que a su vez serán analizadas con el fin de remover cualquier tipo de ruido que interfiera con el proceso del cálculo del índice.

Las variables de estudio entran al análisis bajo el criterio económico descrito en los trabajos de [7] y [8]. Como bien lo dice Guerrero en [15], la construcción de un índice viene de acuerdo a la estructura de la economía en estudio sujeto a la características y coyunturas propias de la misma. Además, las variables seleccionadas no solo son representativas de los sectores más importantes de la economía sino que también contribuyen de manera notoria en las fases del ciclo económico.

Para esta primera parte, se determina el grado de coincidencia de cada una de las variables con respecto al “índice preliminar”, con el método de perfil coincidente de [23] como se describe en la sección 1.3. Simultáneamente, se calcula la función muestral de correlación cruzada entre cada variable y el índice preliminar (en primeras diferencias) y se examina la significancia del coeficiente de correlación cruzada en los rezagos $0, \pm 1, \pm 2, \pm 3$. Una variable con correlación cruzada significativa en alguno de esos rezagos, es considerada “coincidente” con respecto al estado de la economía.

Para este caso, el *estado de la naturaleza* es un proceso no observable el cual puede ser rastreado a través del “índice preliminar”. En el artículo de [27], el PIB es men-

sualizado y es usado como “índice preliminar”. Sin embargo, tiene como desventaja que los tiempos de publicación son muy rezagados y la serie tiene que ser desagregada mensualmente para poder ser trabajada. Aunque en trabajos anteriores ha sido utilizado el Índice de la Producción Industrial (IPI) como índice preliminar, no se descarta la posibilidad de incluir otro tipo de índices que incorpora características más acordes al desempeño de la actividad económica.

ETAPA 2

Este paso parte de aplicar varios procedimientos sobre las series de tiempo seleccionadas en la **ETAPA 1**, con el fin de remover ruidos o anomalías (datos atípicos, intervenciones, etcétera.) que tengan incidencias negativas sobre los resultados y lograr obtener los factores comunes con un proceso más *limpio*. La etapa se desarrolla en los siguientes pasos:

- i) Ajuste de los indicadores por datos atípicos (o Variables de intervención) con el paquete estadístico TSW desarrollado por [10].

Como lo describe [20], cualquier identificación y/o ajuste de un modelo adecuado para una serie de tiempo y, en especial, para lograr obtener los componentes relacionados al ciclo de negocios, puede requerir correcciones previas, debido a que se podrían incluir distorsiones no-deseadas en los resultados finales. Los valores extremos son un obstáculo inicial; ellos representan “*observaciones atípicas que no pueden ser explicadas por ningún modelo o ajuste y que de un modo violan algunos principios básicos de la modelación como pueden ser la normalidad o la homocedasticidad*” ([20]). En series de tiempo, usualmente se analizan tres³ tipos de datos atípicos: outlier aditivo (AO), cambio en nivel (LS) y cambio temporal (TC).

Como bien es remarcado en [20], el principal objetivo de la corrección de datos atípicos es modificar las series temporales de tal manera que las estimaciones estén alejadas de cualquier distorsión o conclusión errónea permitiendo la comparabilidad entre diferentes series. Usando el programa TSW, se procede al ajuste de los indicadores seleccionados que lo requieran.

- ii) Transformación de las series.

Primordialmente, se busca decidir si todas las series (tanto las ajustadas como las no intervenidas en el paso anterior) requieren transformaciones, como por ejemplo de la transformación logarítmica. Teniendo en cuenta el procedimiento automático de ajuste de TSW, se realiza la prueba y se define las series que requieren de aplicar la transformación logarítmica.

- iii) Desestacionalizar - Efecto Calendario

Las series ajustadas en los dos pasos anteriores son sometidas al proceso de desestacionalización. Este ajuste no solo involucra los ajustes por efectos estrictamente estacionales, sino también por efectos del calendario, que pueden incidir en la dinámica real de las series. Para identificar las series que requieren de variables regresoras para desestacionalizar por efectos estacionales y/o del calendario recurrimos al algoritmo del programa TSW.

- iv) Eliminación del ruido remanente de las series.

³En general se consideran 4 tipos pero para el actual caso de estudio no será incluido el Outlier innovacional

Este ejercicio pretende eliminar el ruido que no ha sido removido con los anteriores procedimientos y se enfoca principalmente en la transformación de la serie a través de filtros que extraigan componentes remanentes de alta frecuencia⁴ de la serie temporal. Con este propósito, las series son sometidas a dos tipos de filtros: un filtro de bandpass aplicado por [6] y el filtro de Hodrick y Prescott [19].

Un filtro debe cumplir con ciertos requerimientos⁵ que le permitan ser verdaderamente útil. En el estudio de [6], se construye un filtro “bandpass” de orden K el cual busca cumplir con estos requerimientos. En el contexto de las series temporales, el impacto de un filtro sobre una serie y_t está dado por un promedio móvil simétrico $\tilde{y}_t = \sum_{j=-K}^K a_j L^j y_t$ con L como el operador de rezago tal que $L^K x_t = x_{t-K}$ y los pesos $a_k = a_{-k}$ para $k = 1, \dots, K$. En el contexto frecuentista, el filtro está caracterizado por la transformación de Fourier, $\alpha(\omega)$. Para encontrar los valores de a_j , se resuelve el siguiente problema de minimización,

$$\begin{aligned} \text{Min}_{a_j} \int_{-\pi}^{\pi} |\beta(\omega) - \alpha(\omega)|^2 d\omega \\ \text{s.a. } \alpha(0) = 0 \end{aligned} \quad (2.1)$$

donde $|\beta(\omega)|$ es el filtro “ideal” con corte entre las frecuencias ω_1 y ω_2 . Resolviendo el problema de minimización se tiene:

$$\begin{aligned} a_j &= b_j + \theta, j = 0, \pm 1, \pm 2, \dots, \pm K; \\ b_j &= \begin{cases} \frac{\omega_2 - \omega_1}{\pi}, & \text{si } j = 0. \\ \frac{1}{\pi_j} (\sin \omega_2 j - \sin \omega_1 j), & \text{si } j = \pm 1, \pm 2, \dots \end{cases} \\ \theta &= \frac{-\sum_{j=-K}^K b_j}{2K + 1} \end{aligned}$$

La construcción de este filtro está basado en promedios móviles que aíslan tanto ruidos de baja frecuencia como de alta frecuencia. Por otro lado, se tiene el filtro Hodrick y Prescott el cual parte de minimizar la siguiente función:

$$\text{Min}_{\tau_t} \text{HP}(\lambda) = \text{Min}_{\tau_t} \sum_{t=1}^N (y_t - \tau_t)^2 + \lambda \sum_{t=3}^N (\tau_t - 2\tau_{t-1} + \tau_{t-2})^2, \quad (2.2)$$

para $t = -1, \dots, N$, τ_t como el componente de tendencia y y_t la serie original. También existe un valor $\lambda > 0$ el cual establece un balance entre la “Fidelidad” (**F**) y su “suavidad” (**S**). La fidelidad a los datos es el resultado de tener una serie filtrada con la misma conducta de los datos originales, y esta dado por,

⁴En otras palabras sería una superposición de ciclos de alta frecuencia

⁵Son seis requerimientos los cuales se centran en i) el filtro debe extraer el componente cíclico dentro del rango de tiempo especificado dejando las características de este componente sin distorsiones. ii) No debe haber cambio de fase. El filtro no debe cambiar el “ritmo” de los puntos de quiebre de la serie original. iii) El filtro debe ser una aproximación óptima del filtro ideal. iv) El filtro debe tener las propiedades de una tendencia-reducida. v) el filtro debe generar componentes del ciclo de los negocios no-relacionados con la duración del período. vi) El método debe ser operacional. Para mayor información puede consultarse en [6]

$$F = \sum_{t=1}^N (y_t - \tau_t)^2.$$

Por otro lado, su suavidad es el resultado de tener una serie filtrada de la forma,

$$S = \sum_{t=3}^N (\tau_t - 2\tau_{t-1} + \tau_{t-2})^2.$$

Si $\lambda \rightarrow 0$ entonces $HP(\lambda) \rightarrow Y_t$ es decir que $HP(\lambda)$ será cercano a (F). Así mismo, si $\lambda \rightarrow \infty$ entonces $HP(\lambda)$ será la función lineal $\tau_t - 2\tau_{t-1} + \tau_{t-2}$. Como es mencionado en [15], aunque este método ha sido bastante criticado por su uso excesivo y su tendencia a generar ciclos espureos⁶, es el método de preferencia por su desmepeño frente a otras alternativas.

ETAPA 3

Determinar el número de factores comunes. El proceso se enfoca en calcular el número de factores con las variables ya tratadas en la **ETAPA 1** y **2**, y obtener sus estimaciones usando la metodología de [31] y complementando con la de [30].

ETAPA 4

La definición de perfil coincidente dada en la Sección 2, está basada, filosófica y conceptualmente hablando, en la presencia de máximos y mínimos locales en los valores en niveles de las variables bajo estudio. En ocasiones, no es fácil detectar estos puntos extremos locales, debido a series de tiempo muy suaves o presencia de demasiado ruido en ellas. La sugerencia es entonces, transformar las series originales por medio de los filtros de primera diferencia finita (cambios absolutos) y, con base en los nuevos datos, encontrar el perfil coincidente. Al hacerlo así, el procedimiento de detección de puntos de giro en primeras diferencias, produce aproximadamente los denominados valles y picos del ciclo económico. Y el perfil coincidente que así se obtenga, mostrará solo la coincidencia entre estos puntos de las variables transformadas.

Un valle puede ser interpretado como el punto de fondo de una recesión, mientras que un pico señalaría el punto de techo de una expansión. Los puntos de giro de la variable original, corresponderían a aquellos en donde la primera diferencia es cero o cambia de signo. Así, de negativo a positivo, tendríamos un mínimo local y, al contrario, un máximo local.

Identificados los picos y valles y dado que el grado de coincidencia entre dos procesos lo determina el valor l_m , se procede a obtener los perfiles coincidentes al nivel α de cada factor común identificado. Luego, se selecciona el factor común con grado de coincidencia más cercano a cero y de mayor p -valor respecto al índice preliminar escogido. Esto se considera una *condición de coincidencia deseada*.

Únicamente al factor común que alcance las condiciones de coincidencia deseadas le es aplicado la prueba de Moses, también en comparación al índice preliminar escogido. Siguiendo la metodología descrita en la sección 2.1 y el trabajo de [5], si la

⁶Limitaciones tan sencillas, como por ejemplo la selección arbitraria del λ , son la causa de obtener resultados no esperados, como es el caso de los ciclos espureos, o de tener dificultades en el análisis de los valores extremos.

hipótesis nula no es rechazada, los puntos de quiebre del índice seleccionado tienen la misma dispersión que los puntos de quiebre del índice preliminar. Esto permite concluir que, bajo evidencia estadística, los ciclos del índice seleccionado tienen las condiciones ideales de desempeño en su labor de rastrear la dinámica de la actividad económica al igual que el índice preliminar. Si la hipótesis nula es rechazada en favor de la hipótesis alterna, los puntos de quiebre del índice propuesto no pertenecen a la misma población del índice preliminar.

ETAPA 5

La última etapa asigna la base del índice y obtiene los puntos de giro del mismo con el método de [9]. Para determinar el(os) ciclo(s) económico(s), se adopta la metodología desarrollada en [7], donde se comparan las tasas anuales de crecimiento del PIB y los periodos de contracción y expansión⁷ de las desviaciones (pendiente) de la tendencia obtenida para el índice propuesto. Este proceso permite analizar e identificar las coincidencias entre las tasas de crecimiento anual que son observadas en la economía y las fases del ciclo de los negocios o *business cycle*⁸ del índice. El presente estudio también incluye, en esta comparación, los puntos de giro obtenidos para confirmar los periodos de recesión y expansión.

2.3. Actualización del Índice Coincidente

Es importante mencionar, que existe un proceso para actualizar el índice coincidente el cual tendrá periodicidad mensual y que la solidez del índice dependerá, en gran medida, de los indicadores seleccionados (teniendo en cuenta que su criterio de selección no solo es a través de herramientas estadísticas sino de un juicio económico) y de su disponibilidad una vez termine el periodo de referencia o, lo que se entiende como la oportunidad de sus indicadores. El mecanismo recurrente para actualizar el índice es el siguiente:

- 1) Aplicando la **Etapa 1** hasta la **Etapa 5** se obtiene el índice coincidente para el periodo muestral el cual va desde el periodo inicial t_0 hasta el periodo de referencia T . El índice permite establecer un conjunto de variables consolidadas en este periodo muestral las cuales serán los indicadores definitivos para continuar con la actualización del índice.
- 2) Para obtener el índice coincidente para el periodo muestral que va desde el periodo inicial t_0 hasta el periodo de referencia $T + 1$, los indicadores, consolidados en el paso anterior, son tomadas con el mismo periodo muestral y es aplicada la metodología mencionada en este capítulo desde la **Etapa 2** hasta la **Etapa 5**.

Dado que los indicadores definitivos tienen un nuevo valor hasta el periodo de referencia $T + 1$ (comparada a las series reportadas hasta T) y, adicional, estos indicadores están sujetos a correcciones por parte de las fuentes, es posible que las estimaciones de los parámetros se vean afectados y, por ende, los valores pasados del índice. Por

⁷Según Bujosa y García en [7], el(os) periodo(s) de recesión están definidos como el punto donde la derivada del índice propuesto se vuelve negativa y se mantiene así por al menos seis periodos consecutivos para series mensuales. Del mismo modo, los periodos de expansión son definidos como el punto donde la derivada del índice propuesto se vuelve positiva y se mantiene así por al menos nueve periodos consecutivos para series mensuales

⁸También es definido como el *ciclo del crecimiento* en [15]

ello, si la selección de las variables fue adecuada, entonces estos cambios no deben afectar las propiedades estadísticas del índice en lo que respecta a su desempeño. En otras palabras, los resultados en la **Etapa 4** en lo que respecta al perfil coincidente y la prueba de Moses deben llegar a las condiciones de coincidencia deseables. Así, se obtiene el índice coincidente para este periodo.

- 3) Para actualizar el índice coincidente en el periodo muestral que va desde el periodo inicial t_0 hasta un nuevo periodo de referencia $T + i$, con i número entero positivo, los indicadores definitivos son tomados con el mismo periodo muestral y es aplicada la metodología desde la **Etapa 2** hasta la **Etapa 5** y así periódicamente (mensualmente).

El conjunto de variables o indicadores definitivos puede mantenerse por varios periodos hasta el momento en que el índice pierda sus propiedades deseadas de coincidencia y desempeño, esto es, cuando el perfil coincidente indique que el mayor p -valor respecto al índice preliminar deje de ser cercano a cero y la prueba de Moses muestre rechazada la hipótesis nula. Bajo esta premisa, se recurrirá al paso 1) del mecanismo recurrente para identificar el nuevo conjunto de variables definitivas para obtener el nuevo índice coincidente.

Construcción del Índice Coincidente

3.1. Etapa 1 - Selección de Variables

Una fase muy importante en la construcción de un índice coincidente de actividad económica, es la selección de los indicadores (variables) observables que lo constituirán, los cuales deben, en tiempo real, señalar razonablemente bien la evolución dinámica del estado de la economía. Para nuestro ejercicio, y siguiendo el enfoque desarrollado en [7] y [8], se tomaron inicialmente un total de 29 variables o indicadores. De ellas, finalmente se seleccionaron seis variables. Como fue comentado con el ISE en la sección 1.5 previamente, el período muestral a utilizar es Enero, 2000 - Noviembre, 2017.

La Tabla 3.1¹ presenta la lista de las 29 variables utilizadas inicialmente, junto con los sectores de la economía a los que pertenecen. Se tiene una primera sección donde están las variables asociadas a las actividades económicas de la producción; en ella están la producción de café y caña de azúcar, representando la actividad agrícola; el índice de la producción industrial representando el sector manufacturero; la demanda de energía del Sistema Interconectado Nacional (SIN) caracterizando el sector de servicios públicos; la producción de crudo en miles de barriles días calendario (BPDC), como variable del sector de minas y canteras; la producción de cemento como variable del sector de la construcción; el comercio minorista total sin combustibles, el comercio al por menor sin los combustibles ni los vehículos y la venta de vehículos, los cuales hacen parte del sector comercial.

En la segunda sección, se tiene dos tipos de variables relacionadas al sector gobierno: las primeras están asociadas al funcionamiento operacional del gobierno, entre las cuales se encuentran la remuneración a los asalariados, gasto general del gobierno central y las transferencias del sistema general de participaciones; las segundas corresponden al recaudo, y son el impuesto al valor agregado (IVA), otros tipos de impuestos (aranceles) y el gravamen al margen financiero (conocido como 4 por mil).

¹Aunque la mayoría de variables provienen de la dirección de cuentas nacionales y de la dirección de metodología y producción estadística del DANE, también se obtuvo información de de otras fuentes como la dirección de impuestos y aduanas nacionales-DIAN, la Asociación Colombiana de Petróleo, FEDECAFE, la empresa ISA especializada en la gestión de sistemas de información de tiempo real XM, el Portal de transparencia económica de la Presidencia de la República, el Sistema Electrónico de Transacciones e Información del mercado de divisas-SET-FX y el Sector Agroindustrial de la caña-ASOCAÑA, entre otros.

Para la tercera sección, Comercio Exterior, se incluyen las exportaciones en cuatro diferentes unidades de medida: precios FOB dólares, precios FOB pesos, volumen peso neto kilo y volumen peso bruto kilo. Así mismo, se incluyen las importaciones en precios CIF en dólares, precios CIF en pesos, volumen peso neto kilo y volumen peso bruto kilo.

En la última sección, Mercado Laboral y Precios, son incluidos la tasa de desempleo a nivel nacional, que hace parte del sector de mercado laboral, y las variables tasa de cambio, índice de precios a las exportaciones e índice de precios a las importaciones.

TABLA 3.1. Conjunto de Variables.

SECTOR	VARIABLES
I. ACTIVIDAD ECONÓMICA	
(a) AGROPECUARIO	Producción de miles de sacos de 60 Kg de café verde Producción de Caña de Azúcar Índice de Producción Real
(b) INDUSTRIA	Índice de Producción Real
(c) SERVICIOS PÚBLICOS	Demanda Energía (SIN) medido en kWh
(d) MINAS Y CANTERAS	Producción de barriles de Crudo en miles BPDC
(e) CONSTRUCCIÓN	Producción de Cemento
(f) COMERCIO	Total Comercio Minorista sin Combustibles Total Comercio Minorista sin Combustibles ni Vehículos Vehículos automotores y motocicletas
II. GOBIERNO	
(g) ADMINISTRACIÓN PÚBLICA	Remuneración a los salarios Gastos Generales del Gobierno Central Sistema general de participaciones
(h) IMPUESTOS	Impuesto al Valor Agregado Otros Impuestos Gravamen al Margen Financiero
III. COMERCIO EXTERIOR	
(i) EXPORTACIONES	Valores FOB Dólares y Valores FOB Pesos Cantidades Peso Neto Kilo y Peso Bruto Kilo
(j) IMPORTACIONES	Valores CIF Dólares y Valores CIF Pesos Cantidades Peso Neto Kilo y Peso Bruto Kilo
IV. MERCADO LABORAL Y PRECIOS	
(k) MERCADO LABORAL	Tasa de Desempleo
(l) INDICE DE PRECIOS	Tasa Representativa del Mercado Índice de Precios a las Importaciones Índice de Precios a las Exportaciones

En Colombia, actualmente, el Departamento Administrativo Nacional de Estadística (DANE) produce el Índice de Seguimiento a la Economía - **ISE** desde Enero del año 2000, el cual es una estimación *mensual* de la actividad económica de Colombia. Como tal, se puede considerar al ISE un “índice preliminar” y por lo tanto una estimación razonable del estado de la economía casi en tiempo real, pues su cálculo tiene un rezago de dos meses, aproximadamente. Ese punto de inicio del ISE es la razón por la cual, el rango muestral en este trabajo comienza en Enero, 2000.

La ventaja que tiene el ISE como índice preliminar, es que al ser armonizado con el valor agregado de la economía tiene un alto grado de correlación con el PIB Trimestral y el estado real de la economía. Para Colombia, se tenían procedimientos previos para la construcción de un índice coincidente. Es el caso, por ejemplo, de los trabajos de [24], [27] y [28] entre muchos otros. En particular, [27] y [28] obtiene una estimación óptima del PIB mensual la cual se considera un índice coincidente.

Para la selección de las variables finales se determinó el grado de coincidencia de cada una de ellas con respecto al ISE. Paralelamente, se calculó la función muestral de correlación cruzada entre cada variable y el ISE (en primeras diferencias) y se examinó la significancia del coeficiente de correlación cruzada en los rezagos $0, \pm 1, \pm 2, \pm 3$. Luego, tomando sector por sector, se hizo una evaluación de la importancia económica de cada variable en la dinámica del sistema económico. De esta manera, se llegó a un conjunto final de 6 variables, cuya gráfica aparece en la Figura 3.1.

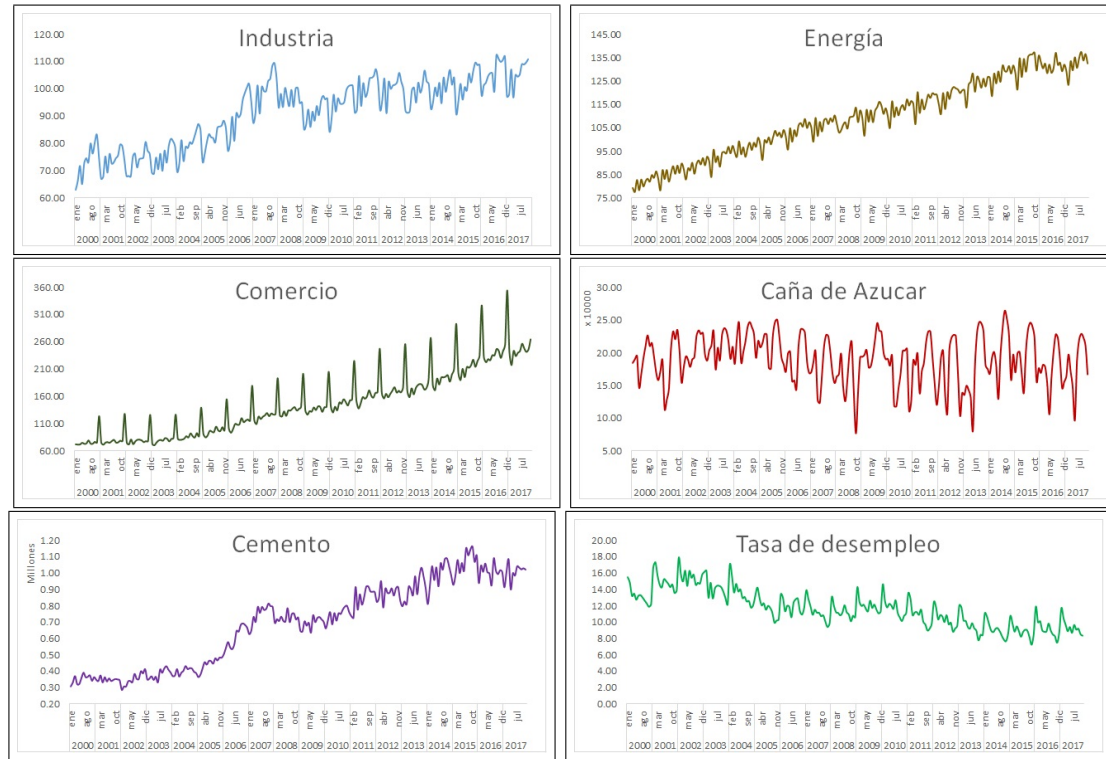


FIGURA 3.1. Series temporales seleccionadas.

Se puede observar, en términos muy generales, la presencia de dos movimientos secuenciales comunes. Uno de ellos es sugerido por la variable Caña de Azúcar y el otro, por las restantes 5 variables. Desde el punto de vista económico y estadístico su descripción es la siguiente.

1. Producción industrial. En muchos países, la industria es el sector de mayor calidad y confiabilidad para captar los ciclos de la actividad económica [8]. En Colombia, la actividad industrial representa uno de los componentes que más contribuye a la evolución del PIB. Para el año 2000, la actividad manufacturera participaba con el 13,93 % del PIB en términos corrientes. Esta cifra cambió para el año 2016, donde su participación ahora es 11,6 %. La Figura 3.2 muestra su perfil coincidente con respecto al ISE para series desestacionalizadas, indicando 0-coincidencia con un p -valor (nivel de confianza) del 100 %. Este es un resultado muy deseable.

2. Energía eléctrica. La generación, administración y distribución de energía eléctrica está directamente relacionado con el funcionamiento de la actividad económica. Su importancia se manifiesta en la contribución que esta actividad tiene sobre el crecimiento económico y su influencia en los diferentes sectores de la economía. En Colombia, la Generación, captación y distribución de energía eléctrica ha tenido en promedio una participación del 2,41 % del PIB desde el año 2000. En el análisis de coincidencia para series desestacionalizadas, la Figura 3.3 nos muestra -1 -coincidencia con un p -valor del 85 %. Este hecho lo podemos interpretar como un rezago de un mes entre los valles y picos del ciclo de esta variable y los del ISE. Cabe destacar que a pesar de la menor participación de esta actividad comparada con la industria, es imperante su inclusión pues es un soporte trascendental de la economía.

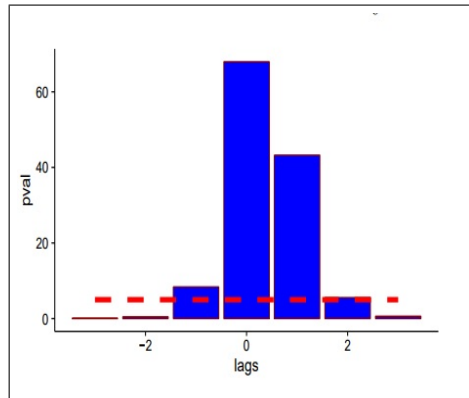


FIGURA 3.2. Perfil coincidente de la Industria.

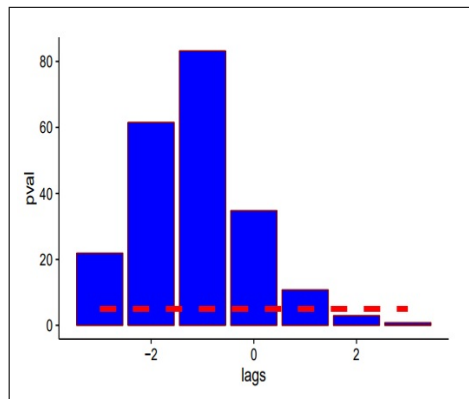


FIGURA 3.3. Perfil coincidente de la producción de Energía.

3. Producción de caña de azúcar. La inclusión de esta actividad puede generar cierta discusión, dado que los niveles de producción son bajos con respecto a otras variables de interés y, adicional, muestra una volatilidad muy alta. Sin embargo, durante los últimos años, el sector pasó de abastecer la demanda de los hogares, a canalizar sus recursos hacia la acumulación de materias primas para la producción de biocombustibles y otros derivados manufacturados. Si bien es una serie con presencia de heterocedasticidad y valores atípicos, su dinámica puede ser muy útil para reflejar los ciclos de la producción. Esto se demuestra en la Figura 3.4 para series desestacionalizadas, donde se observa que el grado de coincidencia es 0 con un nivel de confianza del 75 %.

4. Comercio al por menor sin combustible y sin vehículos. En general, las ventas comerciales al por menor sin combustible y sin vehículos tienen la característica de reflejar una relación directa con el comportamiento del gasto de los hogares. Esta variable muestra la evolución de las líneas de mercancía que los hogares más consumen, como los alimentos, las bebidas alcohólicas, las prendas de vestir, los equipos electrónicos y electrodomésticos, entre otros. Este sector permite ver, por ejemplo, el comportamiento de los ingresos reales de los hogares y su relación con el gasto y, por ende, con el comportamiento de los precios. En Colombia, la actividad comercial representó 8,84 % del PIB en precios corrientes en 2000 y disminuyó a 7,79 % en 2016. Aunque esta contribución puede verse disminuida, se espera que su participación siga siendo importante manteniendo su relación

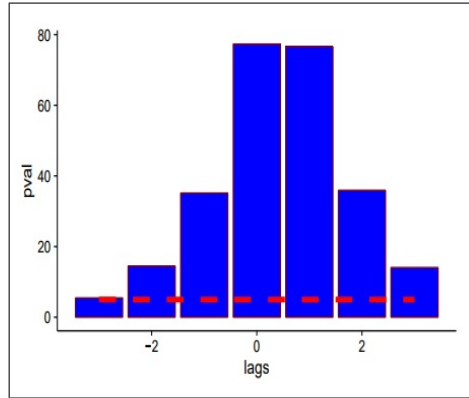


FIGURA 3.4. Perfil coincidente de la producción de caña.

con la dinámica de las importaciones y de la industria nacional. Para las series desestacionalizadas, la Figura 3.5 indica -2 -coincidencia con un nivel de confianza del 63,28 %, lo que implica que su ciclo se rezaga dos meses con respecto al ISE.

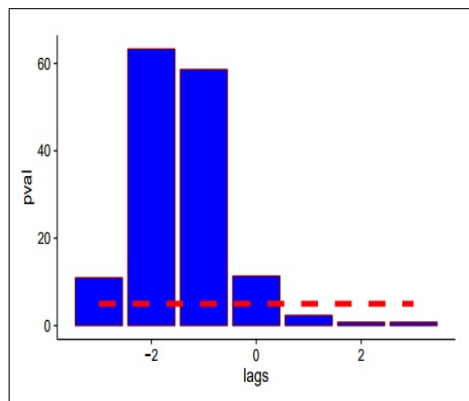


FIGURA 3.5. Perfil coincidente del Comercio.

5. Producción de cemento. Dos sectores, que tienen incidencias importantes en la actividad económica, son la construcción en obras civiles y las cuales son de las actividades que más repercuten en el crecimiento económico. En estas actividades son medidas, no solamente el grado de inversión por parte del sector oficial y privado, sino también su contribución a la dinámica de la infraestructura. La producción de cemento representa un insumo de vital importancia para la construcción de las edificaciones residenciales, no residenciales y obras civiles. Observando la Figura 3.6, para las series desestacionalizadas, la variable es 1-coincidente con un nivel de confianza del 95 % con respecto al ISE, lo cual revela un adelanto promedio de un mes de sus valles y picos con respecto a los del ISE.

6. Tasa de desempleo. Esta variable es un indicador muy importante para el análisis macroeconómico, ya que tiene la característica de indicar las principales señales del estado de la economía [21]. Por ejemplo, una fuerte alza en la tasa de desempleo muestra un claro síntoma de recesión. Así mismo, cuando la fase de recesión culmina, el nivel de desocupados se mantiene y se toma cierto tiempo para que empiece una fase de recuperación, pues, aunque las condiciones sean favorables, las empresas prefieren no comprometerse aún en contratar empleados. Ya para los primeros periodos de la recuperación, la tasa de

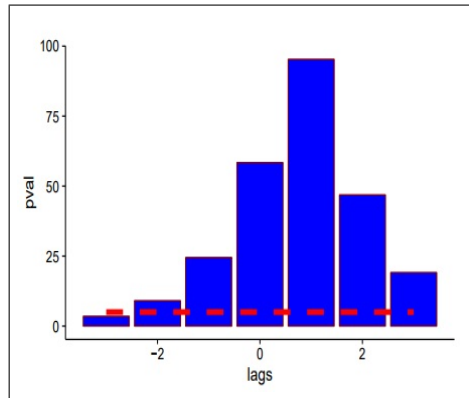


FIGURA 3.6. Perfil coincidente de la producción de Cemento.

desocupación puede disminuir y la economía puede repuntar. Como se puede apreciar en la Figura 3.7, las serie desestacionalizada es 0-coincidente con un nivel de confianza del 80%; en consecuencia, el indicador está sincronizado con la evolución del ciclo para el ISE.

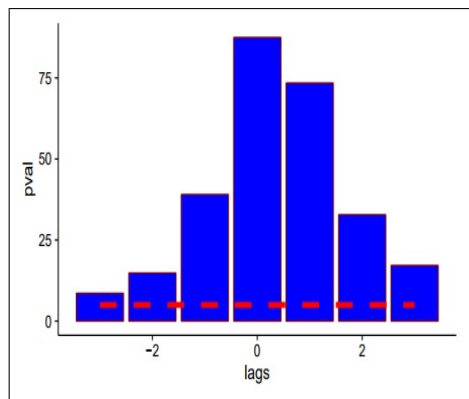


FIGURA 3.7. Perfil coincidente de la Tasa de Desempleo.

Con las variables coincidentes seleccionadas en la sección anterior, procedemos a encontrar un índice para la economía Colombiana.

3.2. Etapa 2 - Tratamiento de los indicadores seleccionados

Como se mencionó previamente, una de las etapas de *transformación* de cada variable es la de remover todas aquellas intervenciones o ruidos que afecten los resultados correspondientes a la obtención del número de factores y la estimación de los mismos. A continuación se presenta el resultado de este análisis.

- Para cinco de las seis variables seleccionadas hubo corrección de las series por identificación de valores extremos, excepto el indicador de energía eléctrica. Los resultados son presentados en la Tabla 3.2.

TABLA 3.2. Identificación de outliers.

Variable	Aditivo	Cambio de Nivel	Cambio Temporal
Industria	Julio-2016	-	-
Energía	-	-	-
Comercio	Abril-2002, Abril-2011	-	-
Caña de Azucar	Octubre-2008	-	Septiembre-2008
Cemento	Julio-2016	Enero-2006	-
Tasa de Desempleo	-	Enero-2003	-

- Cinco series fueron transformadas por el logaritmo natural, excepto el indicador de caña de azucar. En la Tabla 3.3 son presentados los seis indicadores con su resultado respectivo para decidir acerca de esta transformación.

TABLA 3.3. Variables Transformadas.

Variable	Decisión
Industria	Si
Energía	Si
Comercio	Si
Caña de Azucar	No
Cemento	Si
Tasa de Desempleo	Si

- Las variables transformadas y corregidas por valores atípicos ahora son desestacionalizadas usando TSW. Para este procedimiento se tuvo en cuenta que el programa realizara pruebas sobre la identificación de efectos relacionados al calendario , así como correcciones por días bisiestos y días calendario. Para todas las series, el programa identificó efectos del calendario menos para la serie tasa de desempleo. Los modelos SARIMA se presentan en la tabla 3.4.

TABLA 3.4. Ajuste modelo estacional.

Variable	Modelo
Industria	(2,2,1)(0,1,1)
Energía	(0,1,1)(0,1,1)
Comercio	(0,1,1)(0,1,1)
Caña de Azucar	(0,1,1)(0,1,1)
Cemento	(0,1,1)(0,1,1)
Tasa de Desempleo	(0,1,1)(0,1,1)

- Por último, las variables transformadas con los métodos anteriores fueron filtradas por el método de [6] y [19]. Para el caso de los dos filtros, la identificación del número de factores comunes se realizó con las series transformadas, en uno y otro caso. Desafortunadamente, los resultados obtenidos con series filtradas por los dos métodos no permitieron la selección del(os) factor(es). Por ello se llevo a la conclusión de mantener las series sin ningún tipo de filtro.

Después de realizado todo el proceso, las series finales son mostradas en la Figura 3.8.

3.3. Etapa 3 - Cálculo de los Factores Comunes

Usando el método de [30] y [31] se identificaron 2 factores comunes para las variables ajustadas. Luego, se obtienen las estimaciones preliminares de los dos factores usando

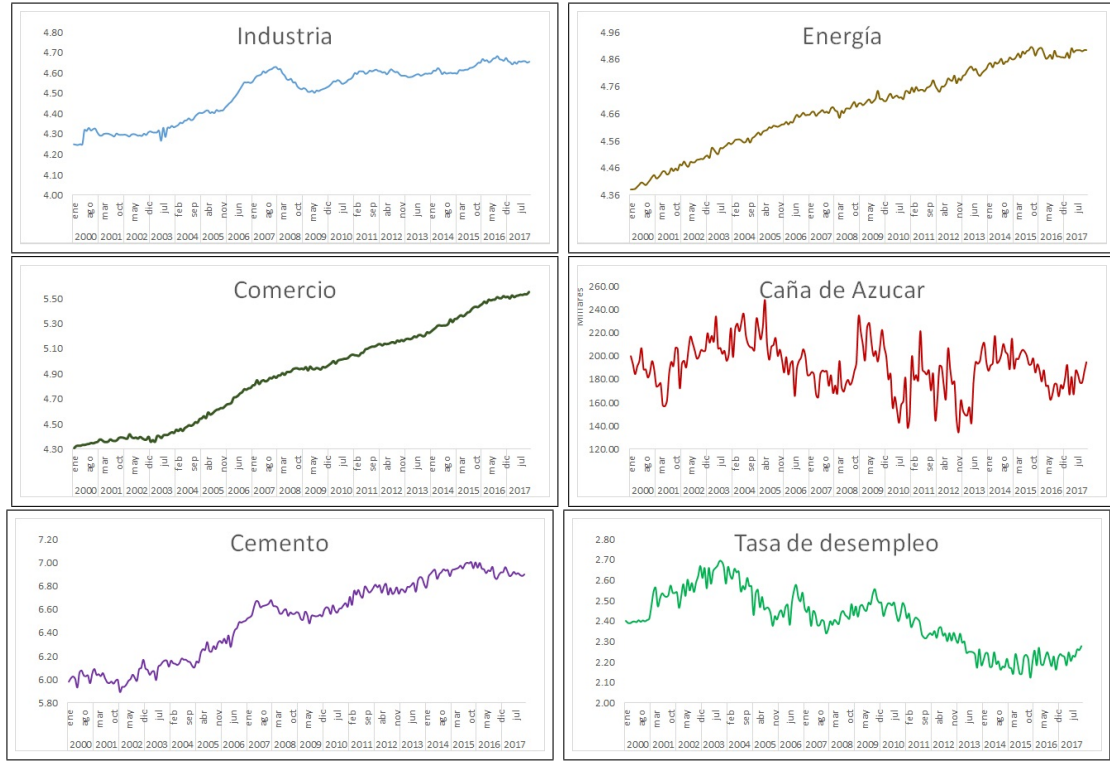


FIGURA 3.8. Series Finales.

combinaciones lineales de las variables. Los coeficientes de estas combinaciones lineales son los componentes de los vectores propios, asociados a los dos primeros valores propios de la matriz M (ecuación 1.4). En seguida, fueron especificados los modelos ARIMA para cada uno de los factores estimados preliminarmente y ellos fueron ARIMA(0,1,1) para el primer factor (valor propio mayor) y ARIMA (0,2,0) para el segundo (segundo mayor valor propio). Nótese que este último factor es un proceso de caminata aleatoria integrado (IEW por sus siglas en inglés), el cual es el modelo básico en el trabajo de [7] para la actividad económica de España. Utilizando la forma de espacio-estado del modelo factorial de [31], explicado en la sección 1.1, y descrito completamente por el trabajo en [29], se estimaron los hiperparámetros del modelo factorial y los factores. La estimación de la transpuesta de la matriz de pesos, P^T , y de las varianzas de los ruidos intrínsecos fue la siguiente:

$$\hat{P}^T = \begin{bmatrix} 0,41 & 0,43 & 0,19 & 0,42 & 0,58 & 0,24 \\ 0,00 & 7,43 \times 10^{-5} & -7,31 \times 10^{-4} & 1,46 \times 10^{-3} & 8,28 \times 10^{-4} & -1,11 \times 10^{-3} \end{bmatrix},$$

y $\hat{\Sigma}_e = \text{diag}\{2,23 \times 10^{-3}, 1,24 \times 10^{-4}, 0,03, 6,12 \times 10^{-5}, 0,01, 4,53 \times 10^{-3}\}$. Todos estos últimos parámetros son significativos al 1%. Además, la estimación del parámetro de promedio móvil del factor 1 es $\hat{\theta} = 0,96(1,63 \times 10^{-3})$ (error standard en paréntesis), el cual es significativo al 5%. Según estas estimaciones, el primer factor es un promedio ponderado de las variables, mientras que el segundo refleja un contraste entre el grupo de variables conformado por energía, comercio sin vehículos y cemento y el grupo integrado

por caña de azúcar y tasa de desempleo². En la estimación de la matriz P se fijó el valor de la entrada (1, 2) en cero, con el objeto de completar las condiciones de identificabilidad del modelo factorial, como lo proponen [31].

Teniendo en cuenta que hemos seleccionado al ISE como la estimación preliminar del estado de la economía, en las Figuras 3.9 y 3.10 es presentada una comparación visual de este con cada factor común estimado, con el fin de tomar una idea acerca de las dinámicas del ISE y los factores comunes. Se observa en la Figura 3.9 que la dinámica del ISE y del factor 1 es similar, mientras que hay mucha disimilitud entre las dinámicas del ISE y del factor 2, como se aprecia en la Figura 3.10.

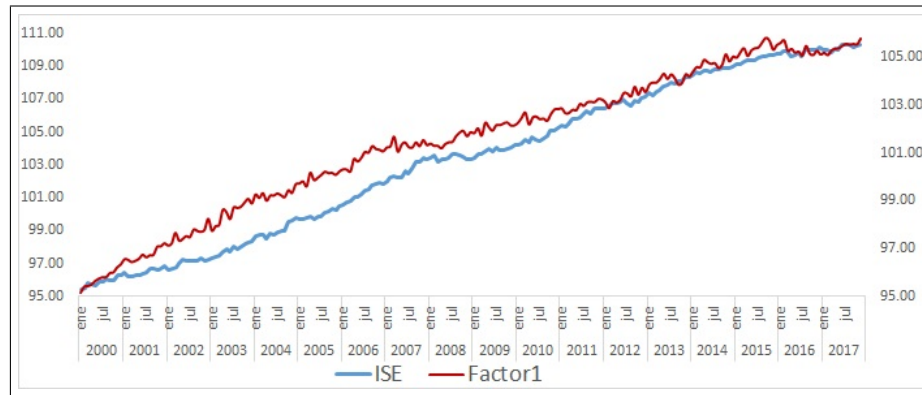


FIGURA 3.9. Comparación entre el ISE y el Factor 1 estimado.

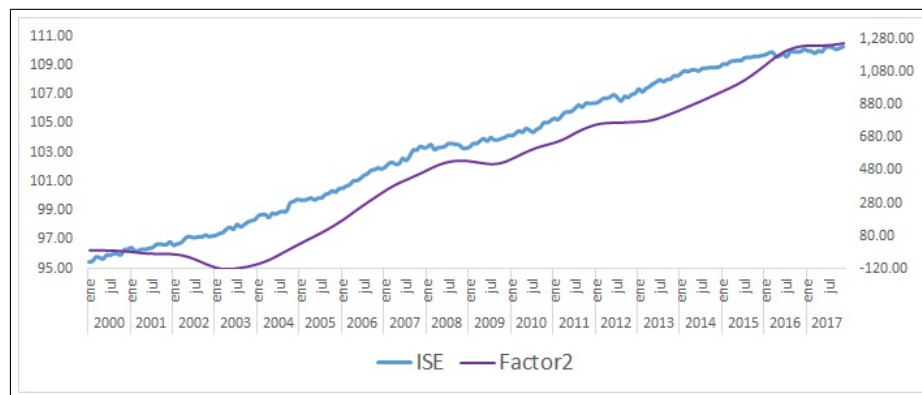


FIGURA 3.10. Comparación entre el ISE y el Factor 2 estimado.

²El primer grupo responde al tipo de variables procíclicas donde se evidencia una dinámica en la misma dirección con respecto a la actividad económica; de hecho varias propiedades como la tendencia, la estacionalidad y los mismos patrones cíclico responden de forma similar, particularmente, al PIB. En el segundo grupo las variables son de tipo anticíclico; la tasa de desempleo es el ejemplo más sobresaliente en macroeconomía: niveles bajos en la tasa de desempleo, responden a un mejor comportamiento de la producción. La producción de caña de azúcar, la cual hace parte de este segundo grupo, también muestra comportamiento anticíclicos debido a que los efectos climáticos, como el fenómeno del Niño, generan cambios en periodos posteriores en los calendarios de cosecha. Esto generalmente sucede en los cultivos de tipo permanente (es decir, cultivos que toman por lo menos un año de preparación para su cosecha) donde los efectos climáticos muestran efectos rezagado sobre el comportamiento de la producción.

3.4. Etapa 4 - Perfil Coincidente y prueba de Moses

Para escoger el factor común con más coincidencia con el ISE, se cálculo el perfil coincidente para cada uno de ellos a partir de los puntos de giro calculados por el método de [9]. Siguiendo la literatura sobre el tema, específicamente el método de [7], primero se obtiene la primera derivada (diferencia finita) de las *tendencias* del ISE y de los factores comunes. Estas tendencias fueron extraídas usando el filtro de Hodrick y Prescott con una constante de suavizamiento $\lambda = 700$. Como lo explican [7], se busca extraer el ciclo de los negocios y, de esta manera, los perfiles coincidentes se obtienen para los ciclos. Este procedimiento es una modificación al de [23], quienes usaron las series temporales en *niveles* (datos originales), para el cálculo de los perfiles coincidentes. En ese trabajo, la estimación preliminar del PIB mensual fue obtenida con el método de [27].

En las Figuras 3.11 y 3.12 son presentados los perfiles coincidentes obtenidos para los ciclos del ISE y de cada factor. Como el ciclo del factor 1 es 0-coincidente y el del factor 2 no lo es, se toma el primero como el índice coincidente. Nótese que, en términos de las estimaciones de las filas de P^T , el primer factor tiene más interpretabilidad que el segundo, en términos económicos.

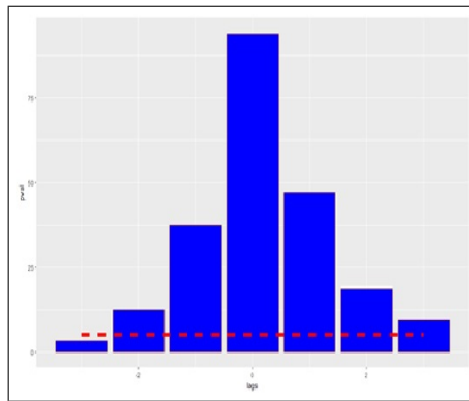


FIGURA 3.11. Perfil coincidente entre el ciclo del ISE y del factor 1.

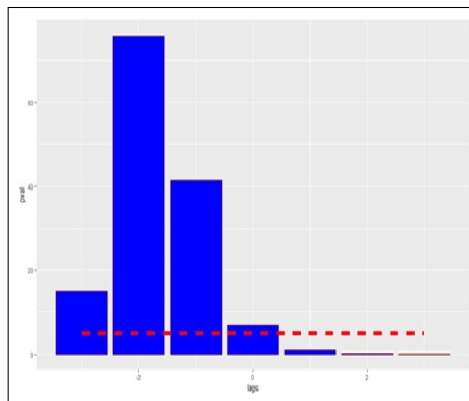


FIGURA 3.12. Perfil coincidente entre el ciclo del ISE y del factor 2.

Para completar el ejercicio del desempeño del índice coincidente obtenido, los puntos de giro calculados del índice coincidente y del ISE que fueron usados para obtener el perfil

coincidente por el método de [23], ahora serán usados para realizar la prueba de Moses. Cuando se aplican pruebas de medias, es posible observar que los puntos de quiebre para cada índice son estadísticamente similares. Por ejemplo, La prueba de *Wilcoxon* comprueba que a un nivel de significancia del 5% se obtiene un *p-Valor* de 67.37% luego la hipótesis nula no se rechaza y la diferencia mediana en la población entera es cero. En la Figura 3.13 se puede observar un *boxplot* para cada uno de los índices.

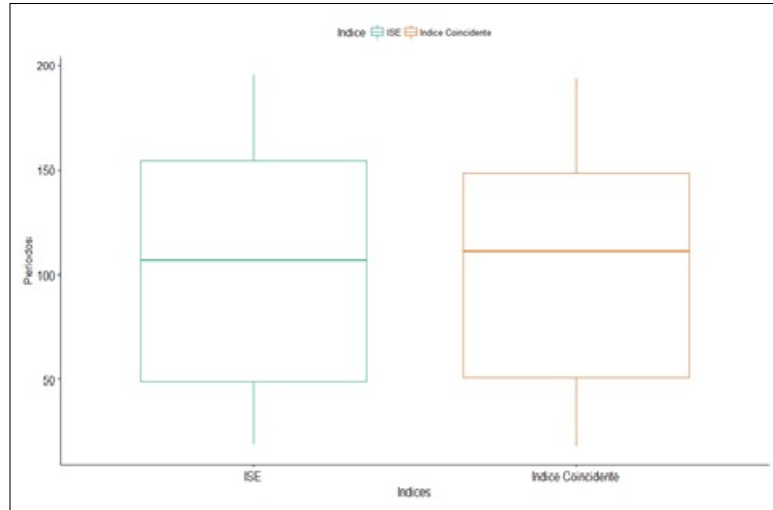


FIGURA 3.13. Boxplot de la prueba de Wilcoxon para el Índice Coincidente y el ISE

Sin embargo, esta prueba no otorga un indicio claro sobre como se comportan los valores extremos de cada una de las muestras. Ahora bien, teniendo el número de puntos de quiebre del ISE n_C , y el número de puntos de quiebre de un índice coincidente n_I ; la hipótesis nula, H_0 , afirma que los puntos de quiebre muy amplios, muy cortos y los moderados, tanto los n_I como los n_C , vienen de la misma población y se encuentran mezclados después de haber asignado un *ranking* a cada posición. Para la hipótesis alterna, H_A , los n_I están más concentrados, sea por arriba o por abajo o por ambas situaciones con respecto a los n_C puntos de quiebre.

El cálculo del *Estadístico* es $t_h = 11$, para $h = 1$, y su *p-Valor* de 69.58%, lo que indica que no existe suficiente evidencia estadística para rechazar la hipótesis nula al nivel de confianza de 5%; esto permite inferir que los puntos de giro calculados tanto del ISE como del índice coincidente pertenecen a la misma población y que, bajo evidencia estadística, los ciclos del índice coincidente propuesto tienen las condiciones ideales de desempeño, en su labor de rastrear la dinámica de la actividad económica colombiana, al igual que los ciclos del ISE.

En síntesis, el perfil coincidente y la prueba de Moses corrobora que el índice propuesto, no solo es un índice coincidente apropiado para hacerle seguimiento al estado de la economía, sino que además los ciclos obtenidos tienen las condiciones ideales de desempeño para rastrear la dinámica de la actividad económica colombiana.

3.5. Etapa 5 - Obtención del ciclo

Con el fin de facilitar la comparación con respecto al ISE, el factor seleccionado es transformado a un índice con base promedio año 2005 y, con base en él, se hace el siguiente análisis para la actividad económica colombiana. En la Tabla 3.5 son presentados los picos y valles del ciclo del índice coincidente y del ISE, obtenidos con el procedimiento de [9]. Resultados similares fueron obtenidos con el enfoque de [18]. En la segunda y tercera columnas, se presentan las localizaciones temporales de ellos (año-mes) y en la cuarta, se incluye la diferencia en meses entre las localizaciones temporales del índice y el ISE.

Se puede observar que las diferencias en valor absoluto son menores o iguales a 7 meses y que hay más coincidencia en los puntos temporales (izquierda para el ISE y derecha para el índice) (septiembre, 2004; diciembre, 2004), (septiembre, 2010; diciembre, 2010) y (Marzo, 2016; Febrero, 2016). En la Figura 3.14 se presenta la gráfica del ciclo del índice de seguimiento para la actividad económica de Colombia.

TABLA 3.5. Valles y picos de los ciclos del ISE y del índice coincidente.

Puntos de Giro	Posición - ISE	Posición - Factor 1	Diferencia Meses
Max	2004,sep	2004,dic	-3
Min	2007,ene	2007,jun	-5
Max	2010,sep	2010,dic	-3
Min	2012,jun	2012,ene	5
Max	2013,oct	2013,mar	7
Min	2016,mar	2016,feb	1



FIGURA 3.14. Ciclo del índice de seguimiento para la actividad económica de Colombia

A partir de los criterios de duración de las fases de un ciclo del trabajo de [7] y comparando con los picos y valles de la Tabla 3.5, se identificó una señal para una *presunta*³ fase de recesión en el período abril, 2014 - noviembre, 2016, la cual se muestra en la Figura 3.14 usando un área sombreada. Adicional, estos puntos de la Tabla 3.5, junto con el ciclo en la Figura 3.14, permiten destacar varios periodos de pequeñas desaceleraciones, como es el caso de la fase (1) mayo, 2003 - junio, 2004; la fase (2) mayo, 2006 - octubre, 2007 y la

³A diferencia de Estados Unidos, Colombia no tiene instituciones establecidas u organizaciones de investigación encargadas de determinar o identificar los ciclos de la economía como lo es el National Bureau of Economic Research (NBER). Este escenario, impide tener un consenso unificado sobre las fases cíclicas de la economía colombiana entre los diferentes entes oficiales y privados. Esta falta de consenso lleva a que sea difícil confirmar la existencia de un ciclo de recesión oficialmente establecido para los periodos propuestos en este estudio. Sin embargo, esta misma situación no impide mostrar que, bajo la luz de los hechos estilizados, la economía vivió un escenario complicado de recesión.

fase (3) enero, 2009 - abril, 2010. La eventual fase de recesión y las fases de desaceleración son comparadas con los crecimientos anuales del PIB de Colombia para el periodo 2000 - 2017⁴, los cuales se muestran en la Tabla 3.6. Aunque no hay tasas de crecimiento negativas para este periodo, algunos años muestran tasas de crecimiento pequeñas como es el caso de 2001, 2009, 2016 y 2017. Se puede observar que la fase (3) de desaceleración coincide con el año 2009 y la *eventual* fase de recesión con los años 2016 y 2017 los cuales han mostrados tasas de crecimiento muy pobres respecto a otros años. También puede haber la salvedad que, aunque las fases (1) y (2) no coinciden con años de bajo crecimiento o crecimiento negativo, estas son fases de desaceleración de poca significancia en el índice coincidente.

TABLA 3.6. Crecimientos Anuales del PIB **Fuente:** DANE

Año	2001	2002	2003	2004	2005	2006	2007	2008	2009
%	1,7	2,5	3,9	5,3	4,7	6,7	6,9	3,5	1,7
Año	2010	2011	2012	2013	2014	2015	2016	2017	
%	4,0	6,6	4,0	4,9	4,4	3,1	2,0	1,3	

3.5.1. Análisis del Índice Coincidente

3.5.1.1. Análisis de Coyuntura

Además de las interpretaciones económicas del índice presentadas en la ETAPA 5 de la sección 3.5, el índice coincidente también se puede utilizar para un análisis de coyuntura, el cual está basado en las variaciones anuales del índice. Estas variaciones están definidas por la expresión

$$VA_{m,t} = \left(\frac{I_{m,t}}{I_{m,t-1}} - 1 \right) \times 100 ,$$

donde, en general, $I_{m,t}$ es el valor del índice en el mes m del año de referencia t , con $m \in \{Enero, \dots, Diciembre\}$. En la Figura 3.15 se observa un crecimiento en su variación anual de 0,5% para noviembre de 2017; mientras que para el mismo periodo del año 2016 no presentaba variación, lo que refleja un escenario de crecimiento parsimonioso de la economía colombiana. En un contexto general, desde el año 2015 la economía viene presentando malos síntomas que están afectando su dinámica. El índice coincidente refleja estos indicios y avisa sobre la preocupante situación en 2016 y todavía más complejo para el primer semestre de 2017 a pesar de una muy leve tasa de crecimiento para el mes de noviembre. Sin embargo, hay que destacar que el segundo trimestre de 2017 muestra un escenario más alentador pero a tasas de crecimiento lentas.

3.5.1.2. Contexto Regional

El ciclo del índice se puede utilizar en el análisis del contexto regional latinoamericano. Para el período Noviembre, 2017, las principales economías latinoamericanas mostraron un escenario parecido al de la economía nacional. En efecto, el Índice Coincidente para Colombia tuvo un crecimiento de 0.5%; el Índice Mensual de la Actividad Económica para

⁴La tasa de crecimiento para el año 2017 corresponde a una estimación que solo tiene en cuenta el periodo enero-noviembre 2016 y enero-noviembre 2017

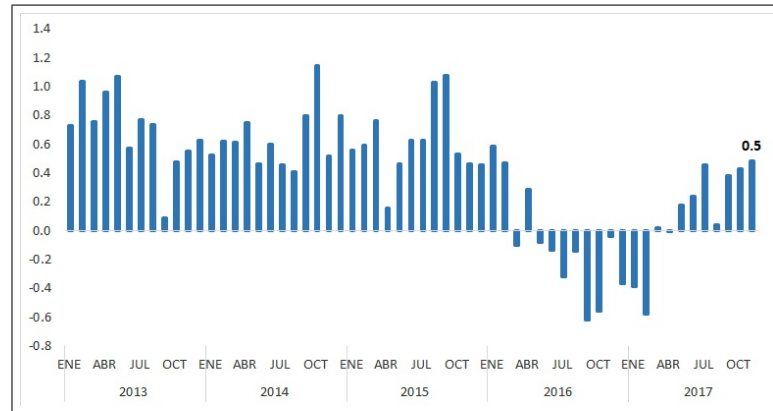


FIGURA 3.15. Evolución de las variaciones anuales hasta noviembre, 2017

Argentina (EMAE), producido por el INDEC, creció 3,7% mientras que el Índice de la Actividad Económica del Banco Central del Brasil (IBC) creció 1.4%. Se destaca el Índice Mensual de la Actividad Económica para Chile (IMAEC), el cual creció 6.4% para este periodo. Para el caso de México y Perú, se mantiene una estabilidad positiva donde el Índice Global de la Actividad Económica para México (IGAE) reportó un crecimiento de 1.5% y el Índice Mensual de la Producción Nacional de Perú (IMPAN) creció 1.4% para este mes en su variación anual. El análisis muestra que para este periodo, hay algunas similitudes entre las dinámicas de las economías de la región. La evolución de los índices es visualizada en la Figura 3.16.

3.5.1.3. Análisis de los indicadores

Con el ciclo del índice se puede realizar también un análisis sectorial de la actividad económica colombiana, a través de los indicadores utilizados para la construcción del índice. Por ejemplo, cabe destacar que, para el mes de noviembre de 2017, aunque el índice de producción industrial tuvo una tasa de crecimiento anual negativa (-0,1%), los demás sectores tuvieron un comportamiento favorable, lo que no permitió que la evolución para este periodo no tomara valores negativos.

3.6. Actualización del índice coincidente

En la medida que llega un nuevo periodo de referencia, se producen nuevos resultados de los indicadores seleccionados. La introducción de un nuevo dato en los indicadores induce a modificaciones en la estructura de los mismos; además los indicadores mensuales por su condición de corto plazo, están sujetos a ajustes de carácter provisional, principalmente en los últimos periodos, y a revisiones permanentes por parte de las fuentes; lo que, finalmente, provoca cambios en los resultados publicados. Esto obliga a que el índice coincidente deba ser actualizado en todo el periodo de muestral.

Dado que este documento construye el índice con el último periodo de referencia, no puede aplicarse el mecanismo recurrente descrito en la sección 2.3 para los meses posteriores hasta no tener estos datos disponibles. Para efectos de estudiar este mecanismo,

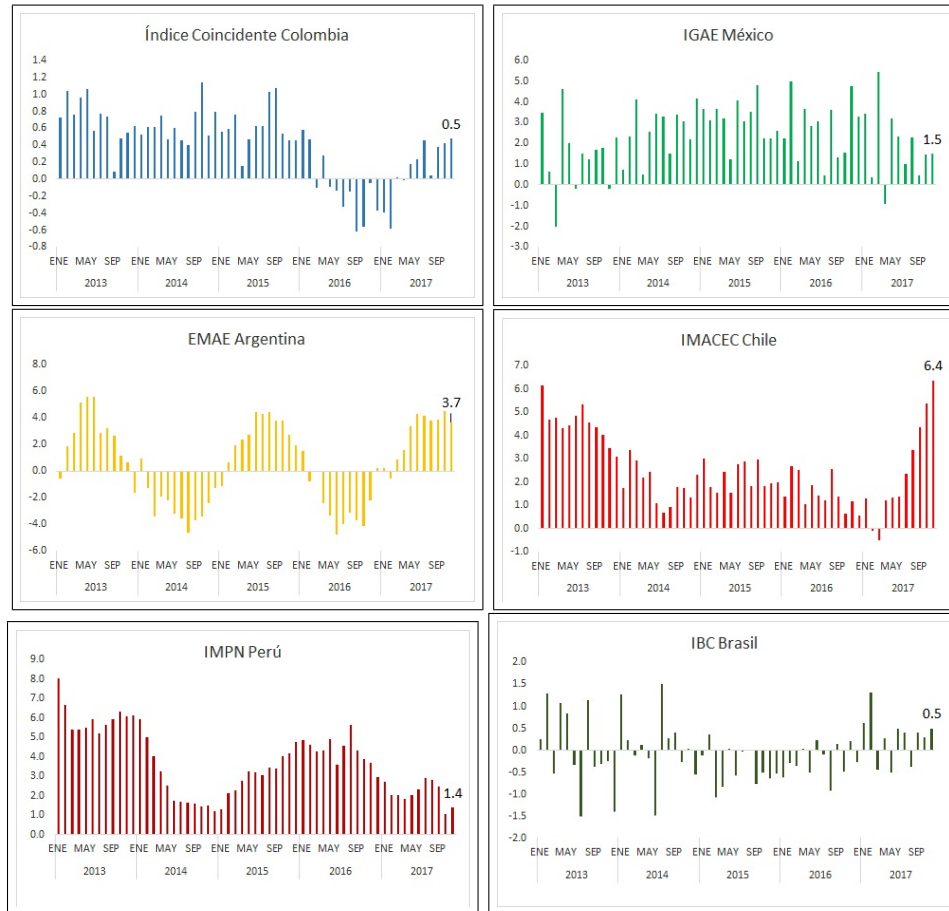


FIGURA 3.16. Evolución de las variaciones anuales regionales hasta noviembre, 2017.

serán evaluados periodos de referencia pasados con el fin de analizar el comportamiento del índice coincidente propuesto.

Después de realizado el paso 1) para **noviembre-2017**, como ejemplo, se toman los periodos de referencia **julio-2017** y **septiembre-2017**, con el mismo conjunto de variables identificadas en la sección 3.1, y se aplica la **Etapa 2** hasta la **Etapa 5**, como se sugiere en el mecanismo, para obtener el índice coincidente de estos dos periodos con sus respectivos resultados. A continuación, se presentan los gráficos de los perfiles coincidentes para cada periodo de referencia:

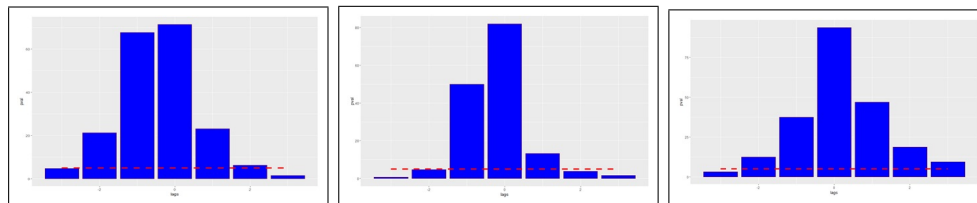


FIGURA 3.17. Perfiles coincidentes para julio, septiembre y noviembre 2017.

Como puede verse en el figura 3.17, para los tres meses de referencia, el cálculo de perfil coincidente alcanza l-coincidencia cercana a cero. Por otro lado, para los tres meses de

referencia, con un nivel de significancia $\alpha = 5\%$, se tiene que para julio 2017 el $p - value = 84.46\%$, septiembre 2017 el $p - value = 55.06\%$ y noviembre 2017 el $p - value = 69.58\%$. En otras palabras, la hipótesis nula no es rechazada luego en los tres casos los ciclos del índice coincidente hacen parte de la misma población de los ciclos del índice preliminar.

Con este análisis se demuestra que los resultados para diferentes periodos de referencia no solo alcanzan las propiedades de coincidencia y desempeño deseadas sino que estas propiedades se pueden mantener a lo largo de los periodos analizados.

Conclusiones

- En el presente trabajo, se ha construido un índice coincidente para la economía Colombiana, extendiendo el enfoque de perfil coincidente de [23] y teniendo en cuenta medidas de desempeño del índice, a través de la prueba de Moses, que corroboran sus propiedades como índice de seguimiento de la actividad económica. El estudio incluyó no solo la actualización del marco teórico y mejoras en la metodología sino que incorporó un ejercicio aplicado en el cual se obtiene un índice coincidente con alta oportunidad y puede mostrarse como es su relación con la coyuntura económica actual en Colombia.
- Se utilizó un conjunto nuevo de variables macroeconómicas y el período muestral analizado cubre hasta noviembre de 2017, mes para el cual se tienen los últimos datos de las variables involucradas. La selección de las variables no solo tuvo en cuenta el análisis estadístico sino que también tuvo en cuenta el análisis económico en el cual los indicadores miden, en cierta medida, sectores importantes de la economía nacional.
- Por último, fue diseñado un mecanismo recurrente para computar el índice coincidente de tal manera que actualizará el índice en tiempo real para cada mes de referencia y garantizará la solidez del mismo. Aunque no se puede evitar los ajustes en los valores pasados del índice, el mecanismo asegura que las propiedades estadísticas del mismo se mantengan gracias al juicio exhaustivo en la selección de sus variables macroeconómicas.

Trabajo futuro

- Como línea futura de investigación en el tema, se piensa en la inclusión de nuevas variables que pueden estar relacionadas a características climatológicas, financieras o relacionadas a la accidentalidad vial. Esta inclusión es relevante puesto que la actualización de los indicadores es de vital importancia no solo para el ajuste y calibración del índice coincidente sino también para incorporar las diferentes dinámicas que pueden tener incidencias en los ciclos económicos.
- Adicional, este estudio también considera importante realizar un trabajo con este índice coincidente el cual puede ser utilizado como el ciclo de referencia para construir un índice líder para la actividad económica de Colombia.
- Por último, está por ser definido un trabajo sobre las propiedades estadísticas de la prueba de permutación, que está detrás de la construcción de los perfiles coincidentes, en este caso de datos autocorrelacionados.

Bibliografía

- [1] Altissimo F., Cristadoro R., Forni M., Lippi M. y Varonese G. (2003), *New eurocoin: Tracking economic growth in real time.*, Review of Economics and Statistics, 92(4), pag 1024 - 1034.
- [2] Argáez, C. (2003), *La estrategia de la seguridad democrática y la economía colombiana: un ensayo sobre la macroeconomía de la seguridad.*, Borradores de Economía, Banco de la República., N 234, pag 61 - 111.
- [3] Aristizabal C., Lozano L. I., M.J.J. (2003), *Déficit público y desempeño económico en los noventa: el caso colombiano*, Borradores de Economía, Banco de la República., N 261, pag 61 - 111.
- [4] Azevedo, J.V.E., Koopman, S.J., Rua, A. (2006), *Tracking the business cycle of the Euro area: A multivariate model based band-pass filter*, Journal of Business & Economic Statistics, 24(3), pag 278 - 290.
- [5] Banerji, A. (1999), *The lead profile and others non-parametrics tools to evaluate survey series as leading indicators. In use of Survey Data for Industry*, Research and Economic Policy, pag 89 - 98.
- [6] Baxter, M. y King R.G. (1995), *Measuring Business Cycle Aproximate Band-Pass Filters for Economics Time Series*, National Bureau of Economic Research, Working paper N 5022.
- [7] Bujosa M. García Ferrer A., De Juan A. (2013), *Predicting Recessions with Factor Linear Dynamic Harmonic Regressions.*, Journal of Forecasting, 32(6), pag 481 - 499.
- [8] Bujosa M. García Ferrer A. (2000), *Forecasting OECD industrial turning points using unobserved component models with business survey data.*, International Journal of Forecasting, 16, 207 - 227.
- [9] Bry, G. y Boschan, E. I. (1971), *Programmed Selection of Cyclical Turning Points in Cyclical Analysis of Time Series: Selected Procedures and Computer Programs.*, Eds. Bry, G. and Boschan, C., NBER, pag 7 - 63.
- [10] Caporello, G. y Maravall, A. (2003), *TSW Program.*, Banco de España, Madrid.
- [11] Charry L. F., López M., Ramírez J. M., Tenjo, F. (2007), *Acelerador financiero y ciclos económicos en Colombia: un ejercicio exploratorio.*, Borradores de Economía, Banco de la República, N 451, Agosto.

-
- [12] Crone, Theodore M. y Clayton-Matthews A. (2005), *Consistent Economic Indexes for the 50 States.*, Review of Economics and Statistics, 87, pag 593 - 603.
- [13] Departamento Administrativo Nacional de Estadística - DANE (2016), *Metodología general indicador de seguimiento a la economía - ISE*, Dirección de Síntesis y Cuentas Nacionales, Agosto 2016.
- [14] Fisher, R. A. (1935), *The design of experiments*, Edinburgh, Oliver & Boyd.
- [15] Guerrero V.M. (2013), *Capacidad predictiva de los índices cíclicos compuestos para los puntos de giro de la economía mexicana.*, EconomíA Mexicana. Nueva Época, Centro de Investigación y Docencia Económicas, A.C. México, 22(1), pag 47 - 99.
- [16] Granger C.W.J. y Newbold P. (1986), *Forecasting economic time series.*, Harcourt Brace Jovanovich, New York.
- [17] Granger C., Hernández Y., Ramos J., Toro J., Zárate H., (2018), *La postura fiscal en Colombia a partir de los ajustes a las tarifas impositivas.*, Borradores de Economía, 1038, Banco de la República, 2018.
- [18] Harding, D. y Pagan, A. (1999), *Dissecting the Cycle.*, Melbourne Institute Working paper, No 13.
- [19] Hodrick, R.J. y Prescott, E.C. (1986), *Postwar US business cycles: an empirical investigation.*, Journal of Money, Credit, and Banking, 29(1):1 - 16.
- [20] Iranzo P. D. (2007), *Análisis de outliers: un caso a estudio I*, Tesis Doctoral, Universidad de Valencia, Facultad de Economía, España, Mayo.
- [21] Leamer, E.E. (2009), *Macroeconomic Patterns and Stories*. Springer: Berlin.
- [22] Martínez G. (2008), *Anatomía de los ciclos económicos en Colombia 1970-2007.*, Borradores de Economía, 496, Banco de la República, Marzo.
- [23] Martínez W., Nieto F.H. and Poncela P. (2016), *Choosing a dynamic common factor as a coincident index.*, Statistics and Probability Letters, pag 89 - 98.
- [24] Melo L. F. and Nieto F. H. (2001), *About a coincident index for the state of the economy.*, Borradores de economía, Banco de la República, November 19.
- [25] Montes E., Ojeda J. y Rubio O. (2003), *Deuda externa, inversión y crecimiento en Colombia, 1970 y 2002*, Borradores de economía, 272, Banco de la República, Diciembre de 2003.
- [26] Moses, L.E. (1952), *A Two-Sample Test.*, Psychometrika, N 17, pag 239 - 247.
- [27] Nieto F. H. (2004), *A coincident index for thwe state fo the economy.*, Internat. Statist. Rev. 72., pag 355 - 376.
- [28] Nieto, F. H. (1998), *Ex-post and Ex-Ante prediction of unobserved economic time series: a case study.*, Journal of Forecasting, N 17, pag 35 - 58.
- [29] Nieto, F. H., Peña, D. and Saboya, D. (2016), *Common seasonality in multivariate time series.*, Statistica Sinica, pag 1389 - 1410.

-
- [30] Lam, C. and Yao, Q. (2012), *Factor modeling for high-dimensional time series: inference for the number of factors.*, The Annals of Statistics, 4, pag 694 - 726.
- [31] Peña D., Poncela P. (2006), *Nonstationary dynamic factor analysis.*, Journal of Statistical planning and Inference, 136(4), pag 1237 - 1257.
- [32] Stock, J.H., Watson, M.W. (1991), *A probability model of the coincident economic indicators.*, Lahiri, K., Moore, G.H. (Eds.), Leading Economic Indicators: New Approaches and Forecasting Records, 4, Cambridge University Press, New York, pag 63 - 85.

Apéndice 1: Soporte formal a la prueba de permutación

Con el fin de darle un soporte formal a la prueba de permutación que fundamenta la construcción del índice y a partir de él investigar las propiedades estadísticas de la prueba de permutación, consideramos el siguiente marco teórico para ello:

Sean Y y Z los tiempos de ocurrencia (por ejemplo, meses, trimestres, etc.) de los puntos de giro de $\{C_t\}$ e $\{I_t\}$, respectivamente. Ellos son variables aleatorias sobre un espacio de probabilidad, digamos (Ω, \mathcal{A}, P) , con valores en \mathbb{Z}^+ , el conjunto de los números enteros positivos. Considérese el vector aleatorio (Y, Z) el cual está definido sobre (Ω, \mathcal{A}, P) y toma valores en $\mathbb{Z}^+ \times \mathbb{Z}^+$, el producto cartesiano de \mathbb{Z}^+ consigo mismo. Así, dado $\omega \in \Omega$, $(Y, Z)(\omega) = (Y(\omega), Z(\omega)) = (t, s)$, donde $(t, s) \in \mathbb{Z}^+ \times \mathbb{Z}^+$. Ahora, se define la variable aleatoria $D = Y - Z$ sobre el espacio muestral de (Y, Z) con valores en \mathbb{Z} . Acá, $D(t, s) = t - s$ para todo $(t, s) \in \mathcal{R}_{(Y, Z)}$, el recorrido o espacio muestral de (Y, Z) . Es claro que la distribución de D , representada por la probabilidad inducida $P_{(Y, Z)}$ de (Y, Z) , no es fácilmente conocida, a menos que se considere supuestos adicionales. Ni tampoco las marginales de Y y Z .

Sea m la mediana de D . El conjunto de hipótesis propuesto en la Subsección 2.2. puede ser re-escrito como $H : m = 0$ v.s. $K : m \neq 0$. Como una estadística de prueba, se propone a $S = \sum_{i=1}^Q D_i$, donde D_i denota la diferencia D en el par i , $i = 1, \dots, Q$. Bajo la hipótesis nula, las diferencias de pares están alrededor de cero en una forma simétrica aproximada, entonces su suma es cercana a cero. En consecuencia, valores de S lejanos de cero señalarán un rechazo de H .

Ahora, para analizar formalmente las permutaciones de las diferencias signadas como se indicó en la Subsección 2.2, considérese el siguiente esquema:

El mecanismo probabilístico que genera los signos $+, -$ puede ser representado por el espacio de probabilidad $(\Omega_g = \{+, -\}, \mathbb{F} = \mathcal{P}(\Omega_g), P_g)$, donde $\mathcal{P}(\Omega_g)$ es el conjunto de todos los subconjuntos de Ω_g . Bajo equiprobabilidad, $P_g(\{+\}) = 1/2$. Ahora, considérese el espacio de medida producto $(\Omega_g^Q, \mathcal{F}, \bar{P})$, donde $\Omega_g^Q = \times_{i=1}^Q \Omega_g$ es el producto cartesiano de Ω_g consigo mismo Q veces, \mathcal{F} es la σ -álgebra generada por el producto cartesiano de \mathbb{F} consigo mismo Q veces y \bar{P} es la medida producto. Nótese que $\bar{P}(\{\delta_1, \dots, \delta_Q\}) = \prod_{i=1}^Q P_g(\{\delta_i\}) = 1/2^Q$, con $\delta_i \in \{+, -\}$, $i = 1, \dots, Q$. Fijan-

do $|d_1|, \dots, |d_Q|$, se define el elemento aleatorio A de $(\Omega_g^Q, \mathcal{F}, \bar{P})$ en el espacio medible $(\mathbb{Z}^Q, \mathcal{P}(\mathbb{Z}^Q))$, donde $\mathbb{Z}^Q = \times_{i=1}^q \mathbb{Z}$ y, como antes, \mathcal{P} indica el conjunto de partes, tal que $A(\delta_1, \dots, \delta_Q) = (\delta_1 |d_1|, \dots, \delta_Q |d_Q|)$ con $\delta_i \in \{+, -\}$. Nótese que esta imagen por A es una sucesión signada de las diferencias observadas correspondiente a la permutación $(\delta_1, \dots, \delta_Q)$. Sea \bar{P}_A la medida inducida por A , entonces $\bar{P}_A(A(\delta_1, \dots, \delta_Q)) = 1/2^Q$, para todo $(\delta_1, \dots, \delta_Q) \in \Omega^Q$. Esta medida de probabilidad será considerada la distribución de permutación.

Sea s_j el valor observado de la estadística S en la j -ésima sucesión signada en el espacio muestral de A , $j = 1, \dots, Q$. Entonces $s_0 = s_{j_0}$ para algún $j_0 = 1, \dots, Q$. Además, sea

$$p_0 = \frac{\#\{s_j : |s_j| \geq |s_{j_0}|, j = 1, \dots, Q\}}{2^Q},$$

donde $\#$ denota el cardinal del conjunto en referencia. Este número es el p -valor de la sucesión signada observada bajo la distribución de permutación \bar{P}_A .

Observación. S es un elemento aleatorio de $(\mathbb{Z}^Q, \mathcal{P}(\mathbb{Z}^Q))$ en el espacio medible $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}))$. Sea $\bar{P}_{A,S}$ su medida de probabilidad inducida. Esta no es necesariamente equiprobable pero si $B = \{y \in \mathbb{Z} : |y| \geq |s_{j_0}|\}$, entonces $\bar{P}_{A,S}(B) = \bar{P}_A(S^{-1}(B)) = p_0$, donde $S^{-1}(B)$ es la imagen inversa de B por S .

Al nivel de significancia α , $0 < \alpha < 1$, la regla de decisión es rechazar la hipótesis nula y mantener la alternativa si $p_0 \leq \alpha$. Un rechazo de la hipótesis nula significa que s_{j_0} es un valor extremo de S bajo la distribución $P_{A,S}$ o que la sucesión observada d_1, \dots, d_Q es un resultado extremo bajo la distribución P_A .