



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ECONOMÍA

MENSUALIZACIÓN DEL
PRODUCTO INTERNO BRUTO REAL.
ESTIMACIÓN Y PRONÓSTICOS, 2003-2008

T E S I N A

QUE PARA OBTENER EL GRADO DE
LICENCIADO EN ECONOMÍA
PRESENTA

ANGEL FERNANDO PINEDA SOLIS

DIRECTOR

FRANCISCO DAVID PAYNO Y SÁNCHEZ



CIUDAD UNIVERSITARIA, MÉXICO D.F. 2008



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Quiero agradecer:

A ti Alejandra, por compartir tu vida conmigo, por estar junto a mí en todo momento, por qué sin tu apoyo y consejos, además de tu comprensión, no podría salir adelante en mis proyectos, ni en mis sueños. Te amo.

A Héctor, Jorge y Luis, por su comprensión al no poder darles todo el tiempo que quisieran, por su paciencia y amor hacia mí. Son lo mejor, los quiero mucho.

A mis padres, por su amor incondicional, confianza y fe en mí, por su ejemplo y por enseñarme que se puede lograr siempre lo que uno se propone. Gracias por estar siempre para mí, los quiero mucho.

A la familia Barragán Domínguez, por su apoyo y comprensión al haberme adoptado como parte de su familia. Los quiero mucho.

También de manera muy especial a todas aquellas personas que me ayudaron, de una u otra forma, en la culminación de este proyecto, por su apoyo, fe, consejos, opinión y sobre todo con su tiempo. Gracias: Francisco Payno, Fernando García, Ernesto García, Daniel Vargas, José Terán y a todos aquellos que por mi falta de memoria, más no de reconocimiento, no pude listar.

Índice General

1. Justificación y delimitación	7
2. Objetivos	9
2.1. General	9
2.2. Particulares	9
3. Marco conceptual	10
3.1. Agregados Macroeconómicos	10
3.2. Números Índice	15
3.3. Modelos econométricos de series de tiempo	21
3.3.1. Promedio Móvil y Autorregresivo	24
3.3.2. Promedio Móvil Autorregresivo Integrado	28
3.3.3. Promedio Móvil Autorregresivo Integrado Estacional	42
3.3.4. Elección del Modelo	47
3.4. Producto Interno Bruto, serie trimestral	49
3.5. Indicador Global de la Actividad Económica, serie mensual	54
4. Mensualización del PIB a precios constantes	58
4.1. Ajuste empírico	58
4.2. Ajuste econométrico y proyecciones	62
5. Conclusiones y recomendaciones	72
Bibliografía	76

ANEXOS

A. IGAE mensual, enero 2003 – marzo 2008	81
B. PIB real trimestral, 1 ^{er} Trimestre 2003 – 1 ^{er} Trimestre 2008	81
C. IGAE trimestral, 1 ^{er} Trimestre 2003 – 1 ^{er} Trimestre 2008	81
D. PIB real mensual, enero 2003 – marzo 2008	82
E. Comparación del PIB real trimestral publicado con el agregado a partir del mensual, 1 ^{er} Trimestre 2003 – 1 ^{er} Trimestre 2008	82
F. Comparación del índice del PIB real mensual con el IGAE mensual ,enero 2003 – marzo 2008	83
G. Comparación del Índice del PIB real trimestral publicado con el agregado a partir del Índice mensual, 1 ^{er} Trimestre 2003 – 1 ^{er} Trimestre 2008	85
H. Comparación del índice del PIB real trimestral con el IGAE trimestral, 1 ^{er} Trimestre 2003 – 1 ^{er} Trimestre 2008	86
I. Gráfica de la serie del PIB real mensual, enero 2003 – diciembre 2007	87
J. Resultado de la prueba aumentada de Dickey-Fuller en primera diferencia para la serie del PIB real mensual.	88
K. Correlograma en primera diferencia de la serie del PIB real mensual	89

L.	Resultados de la Prueba de Theil del modelo SARIMA seleccionado	90
M.	Especificaciones y resultados del modelo SARIMA seleccionado	91
N.	Raíces de los procesos AR y MA del modelo SARIMA seleccionado	92
O.	Correlograma de los residuales del modelo SARIMA seleccionado	94
P.	PIB real mensual estimado con el modelo SARIMA seleccionado, marzo 2006 – diciembre 2008	95
Q.	Comparación de las estimaciones del PIB real con el modelo SARIMA seleccionado, con los datos publicados del primer trimestre de 2008	95
R.	Comparación de las estimaciones del PIB real con el modelo SARIMA seleccionado, con los datos publicados, enero 2006 – diciembre 2008	96
S.	Resultados de la Prueba de Theil del modelo ARIMAX seleccionado	97
T.	Especificaciones y resultados del modelo ARIMAX seleccionado	98
U.	Raíces de los procesos AR y MA del modelo ARIMAX seleccionado	99
V.	Correlograma de los residuales del modelo ARIMAX seleccionado	101

W. PIB real mensual estimado con el modelo ARIMAX seleccionado, octubre 2006 – diciembre 2008	102
X. Comparación de las estimaciones del PIB real con el modelo ARIMAX seleccionado, con los datos publicados del primer trimestre de 2008	102
Y. Comparación de las estimaciones del PIB real con el modelo ARIMAX seleccionado, con los datos publicados, enero 2006 – diciembre 2008	103
Z. Comparación de las estimaciones del PIB real con el modelo ARIMAX seleccionado, con los datos publicados, enero 2003 – diciembre 2008	104

1. Justificación y delimitación

A la hora de plantear el estudio de algún fenómeno económico, uno de los aspectos a tener en cuenta es la periodicidad de la información que podemos utilizar, para realizar estimaciones confiables.

La presencia de datos con distintas periodicidades suele provocar que el modelo resultante tenga la menor periodicidad de las posibles, lo cual conlleva a pérdida de información.

Para que esto no ocurra es necesario que los valores de menor frecuencia sean estimados para mayores frecuencias. Uno de los casos más habituales es el paso de datos anuales a trimestrales y de estos últimos a mensuales.

El mundo tan cambiante en que vivimos, provoca la necesidad de tener que tomar decisiones con gran frecuencia y rapidez. Para ello, es necesario que se disponga de la información necesaria, fiable y lo más actual posible.

Sin embargo, los principales agregados económicos normalmente son anuales, mientras que los indicadores económicos se presentan con mayor frecuencia, ya sea trimestral o mensual.

Así pues, tenemos que actualmente, en México, no existe un cálculo que mida el Producto Interno Bruto (PIB) real de forma mensual, los indicadores que se tienen para el mismo son de periodicidad trimestral y anual.

El Instituto Nacional de Estadística, Geografía e Informática (INEGI), encargado de calcular y publicar los indicadores anteriores, elabora una desagregación mensual del PIB real¹, de uso interno, como insumo para el Sistema de Indicadores Compuestos Coincidente y Adelantado.

Para la mensualización preliminar de la serie del PIB real, se utiliza como variable auxiliar, el Indicador Global de la Actividad Económica (IGAE), que se calcula y publica mensualmente; sin embargo, al agregar trimestralmente la serie preliminar, presenta diferencias con el PIB real trimestral, mismas que obedecen a un efecto estacional.

El presente trabajo pretende desarrollar un procedimiento que elimine las diferencias descritas, permitiendo que haya una representación confiable en las estimaciones, de tal forma que la agregación trimestral del PIB real mensual, sea idéntica al PIB real trimestral, para que una vez definida la serie de tiempo, se puedan hacer predicciones mensuales y trimestrales más precisas sobre el comportamiento del PIB real, para el corto y mediano plazo, considerando efectos estacionales.

¹ Véase: Guerrero, Víctor M. (2004). Nota sobre la estimación del PIB mensual de México. México: INEGI-ITAM.

2. Objetivos

2.1. General

Desarrollar un procedimiento que permita generar una serie de tiempo mensual para el PIB real, del mes de enero de 2003 al mes de diciembre de 2007, para realizar predicciones de corto y mediano plazo.

2.2. Particulares

- Comprobar que la agregación trimestral del PIB real mensual sea idéntica al PIB real trimestral que se publica.
- Establecer un modelo econométrico apropiado para ajustar las observaciones que pudieran presentar diferencias estacionales.
- Realizar los pronósticos mensuales del PIB real para el año 2008.
- Comparar los datos estimados para 2008, con los que se publiquen del primer trimestre del mismo año.

3. Marco Conceptual

3.1. Agregados Macroeconómicos

El interés principal de esta sección, es tener una idea general de cómo se miden y se relacionan la producción, la demanda, el consumo privado, el gasto de gobierno, la inversión, las exportaciones, las importaciones, entre otros.

Para poder hacerlo, debemos de considerar a la contabilidad nacional como la compilación y medición sistemática de estas variables y de otros conceptos macroeconómicos, en cuadros o cuentas, según criterios y normas que los relacionan en un sistema coherente.

Estas variables son resultado de las relaciones económicas –transacciones– que efectúan las diferentes entidades económicas: familias, empresas, gobierno y el resto del mundo.

Aunque el estudio de la contabilidad nacional es relativamente reciente, finales de la primera mitad del siglo XX, los precursores pertenecen a los siglos XVII y XVIII.

El inglés William Petty, midió por primera vez el ingreso y el consumo de Inglaterra, además sugirió la medición sistemática de la actividad económica en su libro "*Political Arithmetick*" (1690).

El irlandés Richard Cantillón en su libro *“Ensayo sobre la Naturaleza del Comercio en General”* (1755) analizó los flujos de ingreso entre los principales sectores de la economía.

Posteriormente, el francés François Quesnay en su *“Tableau Économique”* (1758) hizo una presentación más refinada de los flujos de ingreso entre sectores, que ayudó a la medición de la actividad económica de Francia, es decir, a la construcción de la primera contabilidad económica nacional.

La aparición de la contabilidad nacional, como técnica de medición y conjunto de relaciones que describen la estructura básica del sistema económico, coincide con el desarrollo de la Macroeconomía Moderna, después de la publicación de la *“Teoría General de la Ocupación, el Interés y el Dinero”*, en 1936, de John Maynard Keynes.

El economista americano Simon Kuznets (1941) y los economistas ingleses Richard Stone y James Meade (1944), fueron los pioneros del desarrollo de la contabilidad nacional, cuya estandarización posterior le correspondió a Naciones Unidas (1953). Richard Stone fue el que participó en la preparación de las recomendaciones formuladas en 1947 y 1953 por Naciones Unidas para la construcción de las cuentas nacionales.

Una vez en este contexto, conviene señalar que *“los agregados y relaciones entre las variables macroeconómicas que desarrollaremos, corresponden a identidades contables, válidas en sí mismas por*

*definición*². Para exponer los principales equilibrios y agregados del sistema, iniciaremos con la identidad de oferta y demanda total.

La oferta total es igual a la suma del valor bruto de la producción de bienes y servicios a precios de mercado, más las importaciones de bienes y servicios, en el total de la economía. La oferta encuentra su balance con la demanda total, que se compone de la demanda intermedia y la demanda final.

De tal forma tenemos:

$$OT = DT \text{ sustituyendo}$$

$$VBP + M = DI + DF \text{ en donde}$$

$$DI = CI \text{ y } DF = CP + CG + FBKF + \Delta E + X \text{ entonces tenemos:}$$

$$VBP + M = CI + CP + CG + FBKF + \Delta E + X \quad (1)$$

donde:

OT = Oferta total

DT = Demanda total

VBP = Valor bruto de la producción de bienes y servicios

M = Importaciones de bienes y servicios

² Bibliografía 14.

DI = Demanda intermedia

CI = Consumo intermedio

DF = Demanda final

CP = Consumo privado

CG = Consumo de gobierno

$FBKF$ = Formación bruta de capital fijo

ΔE = Variación de existencias

X = Exportaciones de bienes y servicios

De la identidad (1) se deduce la oferta y demanda global. La oferta global se corresponde con la oferta agregada, la cuál se encuentra representada por el PIB. La demanda global por tanto, equivale a la demanda agregada.

Por lo que si se parte de la identidad de oferta y demanda total, expresada en la ecuación (1), se puede definir el PIB despejando de la siguiente forma:

$$VBP - CI = CP + CG + FBKF + \Delta E + X - M \quad (1.1)$$

donde:

$$VBP - CI = PIB \quad (2)$$

sustituyendo (2) de (1.1) tenemos:

$$PIB = CP + CG + FBKF + \Delta E + X - M \quad (3)$$

El PIB mide entonces, el valor a precios de mercado de la producción de bienes y servicios finales, generados por los factores de la producción residentes dentro del territorio nacional, en un periodo de tiempo determinado, generalmente asociado a un año calendario, trimestre y/o mes, respectivamente.

Para el total de la economía, existen tres metodologías para obtener el PIB: el método de la producción, el método del gasto y el método del ingreso o pago a los factores de la producción.

El método de la producción consiste en deducir al valor bruto de la producción a precios de mercado, el consumo intermedio para calcular así el valor agregado bruto a precios de mercado, que bajo esta valoración es idéntico al PIB, como se puede observar en la identidad (2).

El segundo método es el del gasto, el cual permite medir la demanda de bienes y servicios de utilización final, esto es, por destino final de la producción. La valoración de sus componentes se hace a precios de mercado, es decir, al precio que efectivamente pagan los compradores para hacerse cargo de un bien o servicio. Este método se formula de acuerdo con la identidad (3).

Por último, el tercer método, el del ingreso, consiste en calcular y sumar los componentes del valor agregado: las remuneraciones, los impuestos netos de subsidios sobre la producción y el excedente bruto de operación, que incluye el ingreso de los trabajadores por cuenta propia, los intereses, las regalías, las utilidades, las retribuciones a los empresarios y el consumo de capital fijo, entre otros conceptos. La fórmula para el cálculo del PIB por este método, se expresa como sigue:

$$PIB = R_s + EBO + TNS \quad (4)$$

donde:

R_s = Remuneraciones de asalariados

EBO = Excedente bruto de operación

TNS = Impuestos netos de subsidios sobre la producción

Una vez que se han definido las identidades anteriores, se considera que no es necesario definir alguna otra, dado que el desarrollo del presente trabajo esta relacionado sólo con la oferta y demanda, total y agregada, como origen para la determinación del PIB.

3.2. Números índice

Como su nombre lo dice, los números índice son medidas que indican en cuánto han variado ciertas cosas o bien como se comparan entre sí. Así

pues, se puede decir que un número índice es un valor relativo, que se puede expresar como un cociente o un porcentaje, que mide un periodo dado contra un periodo base determinado. Esta relación mide por lo tanto, la dinámica de un indicador de un periodo en comparación con otro.

De manera formal, *"un número índice puede definirse como una medida estadística que nos proporciona la variación relativa de una magnitud, ya sea simple o compleja, a lo largo del tiempo o el espacio"*³.

Los números índice se pueden clasificar, tomando en cuenta su naturaleza, en simples o compuestos, y además de acuerdo con la importancia relativa de cada componente, es decir, ponderados o sin ponderar.

Los números índices simples son los que estudian la dinámica en el tiempo de una magnitud que sólo tiene un componente, así pues tenemos:

$$I'_X = \left(\frac{X_t}{X_0} \right) 100$$

donde:

I'_X = Número índice en el periodo t de la magnitud X

X_t = Valor de la magnitud X en el periodo t

³ Bibliografía 21.

X_0 = Valor de la magnitud X en el periodo base 0

Por lo general, este tipo de índices se emplean principalmente para estudios de producción y ventas de los productos elaborados.

Los números índice compuestos sin ponderar, son poco utilizados en el ámbito económico, debido a que aunque poseen varios componentes, su evolución en el tiempo se estudia asignando el mismo peso relativo a cada uno de sus componentes, aplicándole así promedios simples a la suma de los valores relativos. De forma matemática tenemos:

Sea una magnitud compleja agregada de N componentes $(1, 2, \dots, N)$ e I_{i0}^t los números índice simples de cada componente i en el periodo t

$$I_0^t = \left(\frac{1}{N} \sum_{i=1}^N I_{i0}^t \right) 100 = \left(\frac{1}{N} \sum_{i=1}^N \frac{X_{it}}{X_{i0}} \right) 100$$

donde:

I_0^t = Número índice total en el periodo t de la magnitud agregada

I_{i0}^t = Número índice simple del componente i en el periodo t

X_{it} = Valor del componente i en el periodo t

X_{i0} = Valor del componente i en el periodo base 0

Para el caso de los números índice compuestos ponderados, estos también tienen varios componentes, pero a diferencia del índice anterior, se asigna un peso diferente o coeficiente de ponderación determinado a cada componente, de tal forma que mide la evolución de las magnitudes de manera más eficiente. Formalmente:

Sea una magnitud compleja agregada de N componentes $(1, 2, \dots, N)$, I_{i0}^t los números índice simples de cada componente i en el periodo t y (w_1, w_2, \dots, w_N) los coeficientes de ponderación de los componentes.

$$I_0^t = \left(\frac{\sum_{i=1}^N I_{i0}^t w_i}{\sum_{i=1}^N w_i} \right) 100 = \left(\frac{\sum_{i=1}^N \frac{X_{it}}{X_{i0}} w_i}{\sum_{i=1}^N w_i} \right) 100$$

donde:

I_0^t = Número índice total en el periodo t de la magnitud agregada

I_{i0}^t = Número índice simple del componente i en el periodo t

X_{it} = Valor del componente i en el periodo t

X_{i0} = Valor del componente i en el periodo base 0

w_i = Coeficiente de ponderación del componente i

Este tipo de índices son los que realmente se emplean en el análisis de la dinámica de fenómenos complejos de naturaleza económica.

Para efectos de la presente investigación utilizaremos solo los índices compuestos y para tal efecto sólo estudiaremos los índices denominados de Laspeyres⁴ y de Paasche⁵.

Índice de Laspeyres o método del año base.

a) Índice de precios⁶:

$$P_L = \left(\frac{\sum_{i=1}^N \frac{p_{it}}{p_{i0}} w_i}{\sum_{i=1}^N w_i} \right) 100 \text{ donde: } w_i = p_{i0} q_{i0} \text{ entonces tenemos:}$$

$$P_L = \left(\frac{\sum_{i=1}^N \frac{p_{it}}{p_{i0}} p_{i0} q_{i0}}{\sum_{i=1}^N p_{i0} q_{i0}} \right) 100 = \left(\frac{\sum_{i=1}^N p_{it} q_{i0}}{\sum_{i=1}^N p_{i0} q_{i0}} \right) 100$$

b) Índice de volumen físico:

$$Q_L = \left(\frac{\sum_{i=1}^N \frac{q_{it}}{q_{i0}} w_i}{\sum_{i=1}^N w_i} \right) 100 \text{ donde: } w_i = q_{i0} p_{i0} \text{ entonces tenemos:}$$

⁴ Llamado así por su desarrollador: Ernest Laspeyres, economista y estadístico alemán.

⁵ Llamado así por su desarrollador: Hermann Paasche, economista alemán.

⁶ Para los índices de precios y volumen físico, se considera p para precios y q para cantidades del producto i , ya sea en la base 0 o en algún punto del tiempo t .

$$Q_L = \left(\frac{\sum_{i=1}^N \frac{q_{it}}{q_{i0}} q_{i0} p_{i0}}{\sum_{i=1}^N q_{i0} p_{i0}} \right) 100 = \left(\frac{\sum_{i=1}^N q_{it} p_{i0}}{\sum_{i=1}^N q_{i0} p_{i0}} \right) 100$$

Índice de Paasche o método del año corriente.

c) Índice de precios:

$$P_P = \left(\frac{\sum_{i=1}^N \frac{p_{it}}{p_{i0}} w_i}{\sum_{i=1}^N w_i} \right) 100 \text{ donde: } w_i = p_{i0} q_{it} \text{ entonces tenemos:}$$

$$P_P = \left(\frac{\sum_{i=1}^N \frac{p_{it}}{p_{i0}} p_{i0} q_{it}}{\sum_{i=1}^N p_{i0} q_{it}} \right) 100 = \left(\frac{\sum_{i=1}^N p_{it} q_{it}}{\sum_{i=1}^N p_{i0} q_{it}} \right) 100$$

d) Índice de volumen físico:

$$Q_P = \left(\frac{\sum_{i=1}^N \frac{q_{it}}{q_{i0}} w_i}{\sum_{i=1}^N w_i} \right) 100 \text{ donde: } w_i = q_{i0} p_{it} \text{ entonces tenemos:}$$

$$Q_P = \left(\frac{\sum_{i=1}^N \frac{q_{it}}{q_{i0}} q_{i0} p_{it}}{\sum_{i=1}^N q_{i0} p_{it}} \right) 100 = \left(\frac{\sum_{i=1}^N q_{it} p_{it}}{\sum_{i=1}^N q_{i0} p_{it}} \right) 100$$

3.3. Modelos econométricos de series de tiempo

Un aspecto importante de la ciencia económica es realizar pronósticos con cierto grado de precisión. En la toma de decisiones, al menos implícitamente, se debe tener alguna idea del futuro.

Se pueden realizar pronósticos a partir de información cualitativa o cuantitativa, a este último tipo pertenecen las series de tiempo –también denominada histórica o cronológica– y se les puede definir como: *“un conjunto de observaciones sobre valores que toma una variable cuantitativa, correspondiente a un fenómeno económico, ordenados cronológicamente en diferentes momentos del tiempo, así pues, un modelo de series de tiempo es un método complejo de extrapolación de datos”*⁷.

Los datos son de la forma (y_t, t) donde:

y_t = Variable endógena o dependiente

t = Variable exógena o independiente

De tal forma que sólo hay una variable a estudiar que es y_t .

En el análisis de regresión simple se tienen dos variables, se explica una variable a partir de la otra. Para este caso y como se dijo, sólo hay una variable, por lo que se explica una variable a partir de su pasado histórico.

⁷ Bibliografía 21.

El análisis de series de tiempo es el procedimiento mediante el cual se identifican y se aíslan los factores relacionados con el tiempo que influyen en los valores observados en las series de tiempo. Una vez identificados, estos factores pueden contribuir a la interpretación de valores históricos de series de tiempo y a pronosticar valores futuros de las mismas. El método clásico para el análisis de series de tiempo identifica cuatro de estas influencias o componentes:

- La tendencia (T).- Es un componente de la serie temporal (Y) que refleja su evolución a largo plazo. Puede ser de naturaleza estacionaria o constante, de naturaleza lineal, de naturaleza parabólica, de naturaleza exponencial, entre otras.
- Las variaciones cíclicas (C).- Es un componente de la serie que recoge oscilaciones periódicas de amplitud superior a un año. Estas oscilaciones periódicas no son regulares y se presentan en los fenómenos económicos cuando ocurren de forma alternativa etapas de prosperidad o de depresión.
- Las variaciones estacionales (S).- Es un componente de la serie que recoge oscilaciones que se producen alrededor de la tendencia, de forma repetitiva y que se consuman en el término de un año y se repiten anualmente. Estas variaciones suelen identificarse con base en datos mensuales o trimestrales.

- Las variaciones aleatorias o irregulares (R).- Es un componente de la serie que recoge movimientos provocados por factores imprevisibles (desastres naturales, cambios súbitos en el nivel de precios, colapsos en bolsas de valores, etc). También reciben el nombre de variaciones irregulares, residuales o erráticas.

De acuerdo con la forma en la que se suponga como están relacionados los cuatro componentes, se pueden formular dos tipos de modelos de series de tiempo, para separar estos componentes y así formular predicciones con respecto a la misma serie (Y). Los modelos de las series de tiempo usualmente son aditivos de la forma

$$Y = T + S + C + R$$

o multiplicativos de la forma

$$Y = T \times S \times C \times R$$

Para un modelo aditivo se supone que los cuatro componentes son independientes entre sí, mientras que para el multiplicativo se encuentran relacionados entre sí.

En este contexto, continuaremos analizando diferentes tipos de procesos estocásticos, útiles para modelar las series de tiempo: proceso auto-regresivo (**AR**), proceso de promedio móvil (**MA**), proceso de promedio móvil auto-regresivo (**ARMA**) y proceso de promedio móvil auto-regresivo integrado (**ARIMA**).

3.3.1. Promedio Móvil y Autorregresivo

Modelos de Promedio Móvil (MA)

En el proceso de promedio móvil de primer orden, cada observación y_t es generada por un promedio ponderado de perturbaciones aleatorias u_t que se remontan a un periodo. Este proceso se denota como **MA(1)** y está definido por:

$$y_t = \beta_0 + u_t - \theta_1 u_{t-1}$$

Para un proceso de promedio móvil de segundo orden, en donde y_t es ahora generada por un promedio ponderado de perturbaciones aleatorias u_t que se remontan a dos periodos, que se denota como **MA(2)** y está definido por:

$$y_t = \beta_0 + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2}$$

Y por último el proceso de promedio móvil de orden general q , en el que y_t es generada por un promedio ponderado de perturbaciones aleatorias u_t que se remontan q periodos. Denotamos este proceso como **MA(q)** y escribimos su ecuación como:

$$y_t = \beta_0 + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - \dots - \theta_q u_{t-q}$$

Donde los parámetros $\theta_1, \dots, \theta_q$ pueden ser positivos o negativos.⁸

En el modelo de promedio móvil se supone que las perturbaciones aleatorias están distribuidas en forma independiente a lo largo del tiempo; es decir, son generadas por un proceso de ruido blanco⁹. Dicho proceso puede que no ocurra frecuentemente; sin embargo, las sumas ponderadas de estos procesos, pueden ser representativos de procesos que no son ruido blanco.

En el modelo de promedio móvil se deben cumplir con los siguientes supuestos:

Sea y_t una variable estacionaria donde:

$$1) E[y_t] = \mu < \infty$$

$$2) \text{Var}[y_t] = \sigma^2 < \infty$$

$$3) \text{Cov}[y_t, y_{t-s}] = \gamma_s < \infty$$

⁸ Siguiendo las convenciones, ponemos un signo de menos frente a $\theta_1, \dots, \theta_q$. En algunos libros de texto, el modelo **MA(q)** se escribe como:

$$y_t = \beta_0 + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q}$$

⁹ El término de ruido blanco se refiere a perturbaciones aleatorias que suponen tener media igual a cero, varianza constante y no estar autocorrelacionadas.

De manera general, con base en los supuestos anteriores, se dice que un proceso estocástico es estacionario, si se mantiene fijo en el tiempo¹⁰; es decir, si su media y su varianza son constantes en el tiempo y si el valor de la covarianza entre dos periodos depende solamente del rezago entre estos dos periodos de tiempo y no del tiempo en el cual se ha calculado la covarianza¹¹. Donde tanto la media, la varianza y la covarianza son finitas.

Modelos Autorregresivos (AR)

Al igual que en los modelos **MA**, los modelos **AR** son procesos estocásticos débilmente estacionarios y las perturbaciones aleatorias también son generadas por un proceso de ruido blanco, por lo tanto los modelos **AR** están sujetos a los mismos supuestos que los modelos **MA**.

En el proceso autorregresivo de primer orden, la observación actual y_t es generada por un promedio ponderado de observaciones pasadas que se remontan a un periodo, junto con una perturbación aleatoria u_t en el periodo actual. Este proceso se denota como **AR(1)** y está definido por:

$$y_t = \beta_0 + \phi_1 y_{t-1} + u_t$$

Para un proceso autorregresivo de segundo orden, la observación actual y_t es generada por un promedio ponderado de observaciones pasadas

¹⁰ Esto permite modelar el proceso a través de una ecuación con coeficientes fijos que pueden estimarse a partir de datos pasados.

¹¹ En el estudio de las series de tiempo, un proceso estocástico como éste se conoce como proceso estocástico débilmente estacionario. Para efectos del presente trabajo y en la mayor parte de las situaciones prácticas, este tipo de estacionalidad será suficiente.

que se remontan a dos periodos, junto con una perturbación aleatoria u_t en el periodo actual. Este proceso se denota como **AR(2)** y está definido por:

$$y_t = \beta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t$$

De tal forma que para un proceso autorregresivo de orden general, este se denota como **AR(p)** y establece que la observación actual y_t es generada por un promedio ponderado de observaciones pasadas que se remontan a p periodos, junto con una perturbación aleatoria u_t en el periodo actual, definiéndose por:

$$y_t = \beta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t$$

Modelos de Promedio Móvil Autorregresivo (ARMA)

Debido a que es muy probable que y_t tenga características de **AR** y de **MA**, analizaremos los modelos mixtos que los combinan, los cuales se denominan modelos de promedio móvil autorregresivos de orden (p, q) .

Estos procesos se denotan como **ARMA(p,q)** y se definen como:

$$y_t = \beta_0 + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q}$$

Para facilitar la notación, podemos utilizar el operador de rezago L . Se define como: $L^z y_t = y_{t-z}$ para todo z . Por lo tanto, $Ly_t = y_{t-1}$,

$L^2 y_t = y_{t-2}$, $L^{-1} y_t = y_{t+1}$, etcétera. Con esta notación, el proceso **ARMA(p,q)** puede especificarse en su forma reducida; aplicando el operador de rezago tenemos:

$$y_t - \phi_1 y_{t-1} - \dots - \phi_p y_{t-p} = \beta_0 + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q}$$

$$y_t - \phi_1 L y_t - \dots - \phi_p L^p y_t = \beta_0 + u_t - \theta_1 L u_t - \dots - \theta_q L^q u_t$$

$$(1 - \phi_1 L - \dots - \phi_p L^p) y_t = \beta_0 + (1 - \theta_1 L - \dots - \theta_q L^q) u_t$$

donde

$$1 - \phi_1 L - \dots - \phi_p L^p = \phi(L) \text{ y } 1 - \theta_1 L - \dots - \theta_q L^q = \theta(L)$$

Por lo que

$$\phi(L) y_t = \beta_0 + \theta(L) u_t$$

Para la estacionariedad del componente **AR**, se requiere que las raíces de $\phi(L) = 0$ queden fuera del círculo unitario, y para la invertibilidad del componente **MA**, que las raíces de $\theta(L) = 0$ estén fuera del mismo círculo.

3.3.2. Promedio Móvil Autorregresivo Integrado

Box-Jenkins presentaron los procesos estocásticos **ARMA** y **ARIMA** para la construcción de modelos de series de tiempo. Como se vio anteriormente, en la especificación de los procesos **ARMA** se hace explícita la

dependencia de una variable respecto a su evolución histórica a través de la inclusión, tanto de los niveles pasados de la propia variable **AR** como de los shocks ocurridos en la serie a lo largo del periodo de estimación **MA**.

La diferencia de los modelos **ARIMA** con los **ARMA** es que se aplica el operador diferencia para eliminar una posible tendencia polinomial de orden d . En la práctica d toma valores de 0, 1 ó a lo más 2.

Un modelo **ARIMA (p,q,d)** indica que está formado de un polinomio autorregresivo de orden p , de una diferencia de orden d y de un polinomio de promedio móvil de orden q . Así tenemos:

$$\phi \nabla^d y_t = \beta_0 + \theta u_t, \quad \forall d \geq 1$$

donde

$$\nabla^d = \text{Operador diferencia}$$

De tal forma que se parte de los siguientes supuestos:

y_t es una variable no estacionaria, pero existe una transformación que si lo es $\nabla^d y_t$, donde

$$1) E[\nabla^d y_t] = \mu < \infty$$

$$2) \text{Var}[\nabla^d y_t] = \sigma^2 < \infty$$

$$3) \text{Cov}[\nabla^d y_t \nabla^d y_{t-s}] = \gamma_s < \infty$$

y el operador diferencia se define como:

$$\nabla y_t = y_t - y_{t-1}$$

$$\nabla^2 y_t = \nabla(\nabla y_t) = \nabla y_t - \nabla y_{t-1}$$

⋮

$$\nabla^d y_t = \nabla(\nabla^{d-1} y_t) = \nabla^{d-1} y_t - \nabla^{d-1} y_{t-1}$$

y la relación entre operadores rezago y diferencia se define como:

$$\nabla y_t = y_t - y_{t-1} = y_t - L y_t = (1 - L) y_t$$

$$\nabla^2 y_t = \nabla y_t - \nabla y_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$$

$$= (1 - L) y_t - (1 - L) y_{t-1}$$

$$= (1 - L)(y_t - y_{t-1})$$

$$= (1 - L)(1 - L) y_t$$

$$= (1 - L)^2 y_t$$

⋮

$$\nabla^d y_t = (1-L)^d y_t$$

Entonces se puede representar el modelo **ARIMA(p,q,d)** como lo hicimos anteriormente con operador diferencia

$$\phi \nabla^d y_t = \beta_o + \theta u_t$$

ó con operador de rezago

$$\phi(1-L)^d y_t = \beta_o + \theta u_t$$

Por otra parte, Box-Jenkins, definen cuatro procesos en el modelado de series de tiempo para poder realizar pronósticos, los cuales se describen a continuación:

a) Identificación

La identificación tiene como objetivo principal determinar tentativamente el orden de los polinomios autorregresivos y de promedios móviles, así como el número de veces que deberá aplicarse el operador diferencia para cancelar la no-estacionariedad homogénea. La propiedad de estacionariedad significa que el proceso está en un estado de equilibrio estadístico¹².

¹² Quiere decir que su comportamiento estadístico no cambia a través del tiempo.

Para volver estacionaria la serie es preciso estabilizar la varianza y estabilizar el nivel (media). Se debe diferenciar a la serie original y_t las veces que sea necesario para lograr la estacionariedad. Luego se identifica el proceso **ARMA**. Para lograr lo anterior, se apoyará con las gráficas de autocorrelación y la función de autocorrelación parcial.

b) Estimación

Una vez identificado un modelo de series de tiempo, se deben encontrar estimadores eficientes para los parámetros del modelo. Identificándose el orden del polinomio autorregresivo y de promedios móviles p y q , así como el grado de diferenciación, se dice que el modelo:

$$\phi(L)\nabla^d y_t = \beta_0 + \theta(L)\mu_t$$

esto es:

$$(1 - \phi_1 L - \dots - \phi_p L^p)\nabla^d y_t = \beta_0 + (1 - \theta_1 L - \dots - \theta_q L^q)\mu_t$$

Resulta adecuado para representar la serie.

c) Verificación del modelo

Esta etapa consiste en verificar que los supuestos básicos del modelo se cumplan. Esto es, que los errores tengan distribución

normal con media cero, varianza constante e independiente¹³, de no ser así, se procede a repetir las etapas anteriores, hasta que se cumplan estos supuestos. Aunque el análisis principal se centra en los residuos, no se deben descuidar los análisis de la bondad de ajuste del modelo estimado¹⁴ y el análisis de los parámetros del modelo¹⁵.

d) Pronóstico

Una vez identificado el proceso **ARIMA** que genera la serie temporal de interés, estimados los parámetros del modelo **ARIMA** correspondiente y haber pasado la etapa de verificación, se utiliza el modelo para realizar pronósticos, con el menor error de predicción posible.

Para verificar que la serie de tiempo presenta estacionariedad, se utiliza una prueba de raíz unitaria conocida convencionalmente como prueba τ (tau) o Dickey-Fuller. En su forma más simple, se estima la siguiente regresión:

$$Y_t = \rho Y_{t-1} + u_t$$

y se divide el cociente estimado ρ por su error estándar para calcular el estadístico τ de Dickey-Fuller y se consultan las tablas para ver si la hipótesis

¹³ Características del proceso de ruido blanco.

¹⁴ Error estándar de los residuos, estadístico Criterio Bayesiano de Schwartz, entre otros.

¹⁵ Condición de invertibilidad y condición de estacionariedad.

nula $\rho = 1$ es rechazada. Sin embargo esas tablas no son completamente adecuadas y han sido ampliadas por Mackinnon a través de simulaciones de Montecarlo.

Si el valor absoluto calculado del estadístico τ excede los valores absolutos críticos de Mackinnon, entonces no se rechaza la hipótesis de que la serie de tiempo dada es estacionaria. Si por el contrario, este es menor que el valor crítico, la serie de tiempo es no estacionaria.

Por razones teóricas y prácticas, la prueba Dickey-Fuller se aplica a regresiones efectuadas en las siguientes formas:

$$\Delta Y_t = \delta Y_{t-1} + u_t$$

$$\Delta Y_t = \beta_1 + \delta Y_{t-1} + u_t$$

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + u_t$$

donde t es la variable de tiempo o tendencia. En cada caso, la hipótesis nula de que $\delta = 0$ donde $\delta = (\rho - 1)$, es decir, que hay una raíz unitaria¹⁶. La diferencia entre estas regresiones es la inclusión de la constante y el término de tendencia.

¹⁶ Si una serie de tiempo tiene raíz unitaria se conoce como una caminata aleatoria, por lo que se identifica como no estacionaria.

En caso de que la serie resulte ser no estacionaria¹⁷ con los datos de la serie en niveles, se le aplica la 1ª o 2ª diferencia hasta que el estadístico de la prueba Dickey-Fuller sea mayor en valores absolutos que los valores críticos de Mackinnon al 5% de significancia.

Existe una relación directa entre las diferencias aplicadas a la serie y el nivel de integración. Es decir, si la serie de tiempo satisface la prueba de raíz unitaria aplicando la 1ª diferencia, se tiene una serie estacionaria con orden de integración uno, denotándose como $I(1)$; y si lo hiciera satisfaciendo la 2ª diferencia, el orden de integración sería $I(2)$, en general, si una serie de tiempo debe ser diferenciada d veces, se dice que ésta es integrada de orden d o $I(d)$.

Como se mencionó en la descripción del proceso de identificación de Box-Jenkins, una vez que la serie se identifica como estacionaria y se establece su grado de integración, se debe identificar el proceso **ARMA** mediante el análisis de las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) para determinar los valores p y q asociados a los términos **AR** y **MA**; en contraste, actualmente muchos investigadores basan esta selección en la formulación en un modelo general con un elevado número de términos **AR** y **MA**, el cual es reducido a una especificación más eficiente por medio

¹⁷ Esto es que la media, la varianza y la covarianza no sean constantes a lo largo del tiempo.

de los estadísticos t^{18} , el Criterio de Información Akaike (AIC), el Criterio Bayesiano de Schwartz (BIC) conocido también como Criterio Bayesiano de Información y el Coeficiente de Desigualdad de Theil.

Como se indicó, se deben identificar la ACF, la PCAF y los correlogramas resultantes, que son simplemente los gráficos de ACF y PCAF frente a la longitud de rezago.

La ACF al rezago k , denotada por ρ_k , se define como:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \text{ donde } \gamma_k \text{ es la covarianza al rezago } k \text{ y } \gamma_0 \text{ la varianza.}$$

Si se gráfica γ_k frente a k , la gráfica obtenida se conoce como correlograma poblacional. Puesto que en la práctica, sólo se tiene la realización de un proceso estocástico, solamente se puede calcular la función de autocorrelación muestral $\hat{\rho}_k$, teniendo entonces:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \text{ donde } \hat{\gamma}_k \text{ es la covarianza muestral del rezago } k \text{ y } \hat{\gamma}_0 \text{ la varianza muestral.}$$

La gráfica de $\hat{\rho}_k$ frente a k se conoce como correlograma muestral.

¹⁸ Los valores de este estadístico permiten discriminar los términos que no deben considerarse como candidatos, desechando todos aquellos que se encuentren por debajo de 1.96, en valor absoluto, o por encima de 0.05 en probabilidad.

El concepto de autocorrelación mide la correlación entre observaciones que están separadas k periodos de tiempo manteniendo constantes las correlaciones en los rezagos intermedios, es decir, la autocorrelación parcial es la correlación entre Y_t y Y_{t-1} después de eliminar el efecto de las Y intermedias.

Para poder reducir a una especificación más eficiente, cuando un modelo **AR**, **MA** o **ARMA** se ajusta a una serie de tiempo determinada, es aconsejable verificar que el modelo en realidad brinde una descripción adecuada de los datos, para lo cual se emplean con frecuencia el AIC o el BIC.

Consideremos que p es el número total de parámetros estimados, entonces:

$$AIC(p) = n \log \hat{\sigma}_p^2 + 2p \quad 19$$

$$BIC(p) = n \log \hat{\sigma}_p^2 + p \log n \quad 20$$

donde n es el tamaño de la muestra y $\hat{\sigma}_p^2$ es el cociente de la suma de cuadrados residual entre $(n - p)$.

¹⁹ Este criterio determina el número de rezagos a incluir, penalizando más fuerte la adición de variables que la \bar{R}^2 reduciendo el número de grados de libertad.

²⁰ Al igual que el criterio de información de Akaike, este criterio determina la longitud del rezago y penaliza aún más fuerte la adición de variables que la \bar{R}^2 reduciendo el número de grados de libertad y ofreciendo un error cuadrático medio más bajo.

Así pues, si se consideran varios modelos, se elegirá aquel que tenga el criterio de información más bajo, ya sea de Akaike o de Schwartz, considerando que estos criterios de bondad de ajuste son similares a los de tipo \bar{R}^2 o $\hat{\sigma}^2$.

Una de las principales pruebas para verificar el modelo, es el Coeficiente de Desigualdad de Theil, el cual se define como:

$$U = \frac{\sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t^s - Y_t^a)^2}}{\sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t^s)^2 + \frac{1}{T} \sum_{t=1}^T (Y_t^a)^2}}$$

Donde el numerador de U es exactamente la raíz cuadrática media del error de pronóstico, pero la escala del denominador es tal que U siempre tendrá valores entre cero y uno. Si $U = 0$, $Y_t^s = Y_t^a$ para toda t , entonces hay un ajuste perfecto; pero si U es igual a uno, el desempeño que pronostica del modelo es tan malo como podría ser. Así pues U mide la raíz cuadrática media del error en términos relativos. Un valor menor a 0.05 es razonablemente aceptable para U .

U debe ser descompuesto en tres proporciones que en suma deben ser igual a uno:

1. Proporción de sesgo U^M , la cual es un indicador de error sistemático, dado que mide la extensión en la que los valores promedio de la serie simulada y real se desvían entre sí. El valor esperado para esta proporción es cercano a cero, por lo que un valor por encima de 0.1 ó 0.2 significaría la presencia de un sesgo sistemático, lo que representaría la revisión del modelo.
2. Proporción de varianza U^S , indica la capacidad del modelo para replicar el grado de variabilidad en la variable de interés. Si esta proporción es grande, significa que la serie ha fluctuado en forma considerable, mientras la serie simulada muestra poca fluctuación, o viceversa; esto indicaría de nuevo que el modelo debería revisarse.
3. Proporción de covarianza U^C , mide el error que no es sistemático, de tal forma que representa el error restante después de que se han explicado las desviaciones de los valores promedio. Al no ser razonable esperar que los pronósticos se correlacionen a la perfección con los resultados reales, este componente del error es menos inquietante que los otros dos.

Así pues, estas proporciones se utilizan como medio de descomposición del error de simulación en sus fuentes características.

Análisis de Intervención

Existen los modelos **ARIMA** con variables de intervención, en los cuales las series económicas son afectadas por fenómenos externos, tales como cambios tecnológicos, huelgas, cambios en medidas de política o económicas, cambios en la legislación o escala de algún impuesto, cambios metodológicos en la medición de las estadísticas, etc. Estos fenómenos son llamados intervenciones ya que interfieren en el comportamiento original de la serie, por lo tanto se debe evaluar su efecto e incorporarlo al modelo **ARIMA** a través de variables artificiales binarias (análisis de intervención).

Se recurre a variables que explican la presencia de fenómenos exógenos en la serie de tiempo. Se incorporan como variables dummies en la forma de impulsos y escalones que se utilizan para representar cambios temporales o permanentes en el nivel de las series debidos a eventos especiales. La no incorporación de variables artificiales conduce a sesgos en las estimaciones de los parámetros, a elevar el error estándar residual y en ocasiones a errores en la especificación del modelo **ARIMA**.

La mayoría de veces, a priori no se conocen los fenómenos exógenos que afectan la serie de tiempo y más bien se utiliza una primera aproximación del modelo **ARIMA** para determinar la presencia de valores anómalos que son posteriormente incorporados al modelo.

A continuación se describen las principales variables de intervención:

Variables Impulso: Recoge el efecto de fenómenos que intervienen en la serie en un único momento T_0 . Esto se traduce en una variable que contiene un uno en T_0 y ceros en el resto. Afecta el componente irregular de la serie.

Variable escalón: Recoge el efecto de un cambio en el nivel en la serie, es decir, que contienen ceros hasta el momento T_0 y unos en adelante. Afecta el componente tendencia de la serie.

Variable tendencia o rampa: Estas contienen ceros en un tramo de la serie hasta un momento T_0 , a partir del cual empieza a crecer en forma ascendente. Afecta la tendencia de la serie.

Efecto calendario: Este efecto se refiere al hecho de que cabe esperar un mayor nivel de actividad en aquellos meses con mayor número de días laborales, por lo cual hay que tomar en cuenta no solo el número de días de cada mes, sino también su diferente composición porcentual en términos de lunes, martes, etc., en cada mes.

Efecto de la semana santa o pascua (Easter effect): Con este efecto se intenta representar la influencia de la festividad móvil de semana santa ejerce sobre la actividad económica en los meses de marzo y abril.

Días de comercio (Trading-Days): Consiste en el ciclo semanal que se presenta cuando los días de la semana tienen un nivel de actividad distinto, unido a la distinta longitud de los meses; de tal modo que por ejemplo, un mes en particular podría tener un nivel de ventas superior a otro, debido únicamente a que posee un mayor número de días.

Las intervenciones se incluyen en la parte determinística de la serie y el modelo **ARIMA** en la parte estocástica.

3.3.3. Promedio Móvil Autorregresivo Integrado Estacional

Hasta el momento solamente se ha analizado la parte regular del modelo **ARIMA**. Sin embargo, muchas series económicas presentan un elevado componente estacional, por lo que esta parte estacional también tendrá un modelo **ARIMA**. En este caso también se sigue el proceso descrito antes, con la diferencia de que los rezagos que se deben analizar son el rezago 12, 24 y 36 en el caso de series mensuales y los rezagos 3, 6, 9 y 12 para las series trimestrales. La notación del modelo **ARIMA** con parte estacional es la siguiente:

ARIMA (p,d,q) X (P,D,Q)s

donde:

p: Número de parámetros autorregresivos.

d: Número de diferenciaciones para que la serie sea estacionaria.

q: Número de parámetros de medias móviles.

P: Número de parámetros autorregresivos en la parte estacional.

D: Número de diferenciaciones para que la serie sea estacionaria, en la parte estacional.

Q: Número de parámetros de medias móviles en la parte estacional.

s: Periodicidad de serie ($s=12$ serie mensual, $s=4$ serie trimestral) trimestrales.

Existen dos extensiones de los modelos **ARIMA** para incluir estacionalidad:

1. Estacionalidad endógena, modelo **SARIMA**.
2. Estacionalidad exógena, modelo **ARIMAX**

Modelos SARIMA

Los modelos **SARIMA** se utilizan cuando existen procesos **AR** y **MA** con estacionalidad. La utilización de modelos **SARIMA** modifican las condiciones de convergencia. En este tipo de modelos la estacionalidad es medida con diferencias y procesos estacionales.

Cuando se está modelando una serie de tiempo, como primer paso debemos observar la gráfica de nuestra variable, a primera vista podemos notar cierta tendencia, pero a fin de verificarlo puntualmente se aplica la prueba de Dickey-Fuller.

Una vez que se verifica la estacionariedad de la serie, se deben identificar la ACF, la PCAF y los correlogramas resultantes. Esta identificación es de suma importancia, ya que permite analizar cómo se comporta la serie y determinar la regularidad en que se presenten los retardos. Así pues, con el correlograma ACF se establecen los procesos **MA** y con el de PACF los procesos **AR**.

Un correlograma donde existen datos, con cierta regularidad, que se ubican fuera de las bandas Bartlett establecidas, constituye un indicador del factor estacional. Los correlogramas se analizan de acuerdo a la regularidad en que se presenten los retardos, por ejemplo en múltiplos de 3 y 12, etc. lo que permite determinar si este fenómeno es trimestral o anual.

La utilización de un modelo **SARIMA** garantiza capturar la estacionalidad y sus efectos en la serie que se esté analizando.

Se debe advertir que la inclusión de un factor estacional podría alterar la modelación de la serie que se tenía anteriormente de acuerdo a los procesos **MA** y **AR** que se hubieran incluido. Por lo que es necesario verificar nuevamente que los coeficientes estimados en la ecuación sean estadísticamente significativos y que se cumplan las condiciones de

invertibilidad y estacionariedad, verificando que las raíces sean menores a 1 en módulo.

Una vez que satisfacemos las condiciones necesarias podemos señalar el orden del modelo **SARIMA**.

Si incluimos en la ecuación un AR(1), MA(1), MA(6), SMA(12)²¹

Se tendrá un **SARIMA** (12, 1, 1, 6)

↓	↓	↓	↓
SMA	AR	I	MA
↓	↓	↓	↓
s	(p,	d,	q)

Una vez que se ha modelado la serie de tiempo, es necesario analizar el correlograma de los residuales para verificar que todas las autocorrelaciones sean mayores a 0.05 en probabilidad. De esta forma se estará garantizando que no faltan procesos estocásticos a ser incluidos en el modelo de pronóstico.

Modelos ARIMAX

En los modelos **ARIMAX** se incluyen variables exógenas. La utilización de este tipo de modelos modifica el punto de equilibrio. En este caso la estacionalidad es medida con variables dummies.

²¹ El término **SMA**(12) representa la inclusión de un factor estacional de tipo mensual en la variable **MA**.

Para modelar una serie de tiempo utilizando un **ARIMAX** se procede de la misma manera que con los modelos **SARIMA**. La diferencia con un modelo **ARIMAX** es que se incluyen variables exógenas llamadas dummies, en lugar de factores de estacionales, para corregir y capturar el efecto de la estacionalidad.

Una vez que se ha determinado el orden de integración se procede a observar el correlograma tanto en la función de autocorrelación simple como en la función de autocorelación parcial y, de esta forma, se observan los procesos que rebasan las bandas de Bartlett a fin de determinar la periodicidad de la estacionalidad.

Una vez identificada la regularidad de los procesos tanto en la columna de los **MA** como en la de los **AR**, se crean las variables dummies relacionando cada variable con el número de procesos que se incluirán en el modelo de estimación.

Se pueden incluir las variables dummies necesarias, hasta que nuestro modelo cumpla con las condiciones de estacionariedad e invertibilidad, verificando que las raíces sean menores a 1 en módulo y que los coeficientes de la ecuación sean en su totalidad estadísticamente significativos bajo los estadísticos t .

Una vez que se satisfacen las condiciones necesarias, es posible determinar el orden del proceso estocástico **ARIMAX**. Pensemos en un modelo de

pronóstico cuya ecuación de regresión está representada por los siguientes procesos:

AR(2), MA(1), MA(3), @SEAS(1), @SEAS(2), @SEAS(3), @SEAS(4)²² y nuestra serie tiene un orden de integración 1; es decir, con la primera diferencia la serie se convierte en estacionaria.

Así entonces, el orden del modelo ARIMAX es

(4,	2,	1,	3)
↓	↓	↓	↓
D	AR	I	MA
↓	↓	↓	↓
s	(p,	d,	q)

Una vez logrado lo anterior es necesario verificar en el correlograma de los residuales que todas las probabilidades sean mayores a .05, de lo contrario, no se estarán considerando todos los procesos estocásticos tanto **AR** como de **MA**, para la estimación de los pronósticos.

3.3.4. Elección del modelo

Uno de los criterios de selección del modelo, puede ser en función de su consistencia, por lo que se analiza:

²² Los términos @SEAS(*n*) representan la inclusión de dummies que pueden ir de 11 a 12 dependiendo de la inclusión o no de constante, ya que si se desea hacer la estimación incluyendo todas las variables dicotómicas (dummies), no debe incluirse el término constante a fin de cuidar el problema de la dependencia lineal, si se desea incluir el intercepto forzosamente se tendrá que incluir "*n-1*" variables binarias para garantizar la independencia lineal.

1. Cuando el modelo verdadero (es decir, el proceso generador de datos) está entre los modelos considerados, la probabilidad de seleccionarlo tiende a uno cuando aumenta el tamaño de la muestra.
2. Cuando el modelo verdadero no está entre los considerados, y es imposible seleccionarlo, la probabilidad de seleccionar la mejor aproximación al proceso generador de datos tiende a uno cuando aumenta el tamaño de la muestra.

Naturalmente que es deseable la consistencia. Si el proceso generador de datos está entre los considerados, se espera que a medida que se incrementa el tamaño de la muestra acabaremos seleccionándolo. El AIC aunque penaliza los grados de libertad, es inconsistente; aun cuando aumente el tamaño de muestra, el AIC selecciona modelos que son demasiado grandes. Sin embargo, el BIC, que penaliza con máxima intensidad los grados de libertad, si es consistente.

Otro criterio de selección es la llamada eficiencia asintótica el cual elige, a medida que aumenta el tamaño de la muestra, una serie de modelos cuyas varianzas de error de pronóstico a un paso tienden a lo que se obtendría con el modelo verdadero, con parámetros conocidos, con una rapidez cuando menos igual a la de cualquier otro criterio de selección de modelo. El AIC, aunque es inconsistente, sí es asintóticamente eficiente, mientras que el BIC no lo es.

En la práctica de los pronósticos, se registran y examinan aquéllos obtenidos con el AIC y EL BIC. Con mucha frecuencia es el mismo modelo el que se selecciona con ambos. Pero cuando no es así, y a pesar de la propiedad teórica de eficiencia asintótica del AIC, muchos autores recomiendan usar el modelo más parsimonioso que selecciona el BIC, en igualdad de circunstancias. Esto se apega al principio de parsimonia ²³ y a los resultados de estudios en los que se comparan la eficiencia de pronóstico fuera de la muestra, de modelos seleccionados con distintos criterios.

3.4. Producto Interno Bruto, serie trimestral

El PIB se define como el valor a precios de mercado de la producción de bienes y servicios finales, generados por los factores de la producción residentes dentro del territorio nacional, en un periodo de tiempo determinado, generalmente asociado a un año calendario, trimestre y/o mes, respectivamente.

En la medición trimestral del PIB los criterios metodológicos que se siguen son similares a los que se emplean en los cálculos anuales de las Cuentas

²³ Principio atribuido al fraile franciscano inglés del siglo XIV Guillermo de Ockham, que indica que las explicaciones nunca deben multiplicar las causas sin necesidad. Así pues cuando dos explicaciones se ofrecen para un fenómeno, la explicación completa más simple es preferible.

de Bienes y Servicios²⁴, cuando la información básica disponible en el corto plazo así lo permite.

Se aplican los criterios de clasificación de las actividades económicas del Sistema de Clasificación Industrial de América del Norte (SCIAN) 2002, que sirven para ordenar también la información estadística básica que cuenta con oportunidad mensual. Para el proceso de agregación de las clases a los subsectores de actividad económica se procede a obtener subtotales por subsector con las 668 clases medidas de un total de 737, con las que se obtienen valores representativos de los subsectores de actividad que se alinean a los valores anuales, infiriéndose de esta forma la evolución probable de las clases que, por no contar con información de corte trimestral y por su mínima importancia relativa, no son directamente medidas. Así, la cobertura alcanzada para el PIB trimestral es de 93.7 por ciento.

Los datos se expresan en términos reales o precios constantes de 2003, y en términos nominales o precios corrientes; y corresponden al total de la economía, así como para cada uno de los sectores que la componen:

- Agricultura, ganadería, aprovechamiento forestal, pesca y caza.
- Minería; Electricidad, agua y suministro de gas por ductos al consumidor final.

²⁴ Véase: Instituto Nacional de Estadística, Geografía e Informática (2008). *Sistema de Cuentas Nacionales de México. Cuentas de Bienes y Servicios 2003-2006 Base 2003. Tomo I.* México: INEGI.

- Construcción.
- Industrias manufactureras.
- Comercio.
- Transportes, correos y almacenamiento.
- Información en medios masivos.
- Servicios financieros y de seguros.
- Servicios inmobiliarios y de alquiler de bienes muebles e intangibles.
- Servicios profesionales, científicos y técnicos.
- Dirección de corporativos y empresas.
- Servicios de apoyo a los negocios y manejo de desechos y servicios de remediación.
- Servicios educativos.
- Servicios de salud y de asistencia social.
- Servicios de esparcimiento, culturales, deportivos, y otros servicios recreativos.
- Servicios de alojamiento temporal y de preparación de alimentos y bebidas.

- Otros servicios excepto actividades del Gobierno.
- Actividades del Gobierno.

Se incorpora una industria ficticia de ajuste, referida a los Servicios de Intermediación Financiera Medidos Indirectamente, que se incluye dada la imposibilidad de medir, por actividad económica, el gasto referido al pago de intereses realmente por las industrias en el registro del consumo intermedio.

Adicionalmente, se agrega la información correspondiente a las Actividades Primarias, Secundarias y Terciarias, para su presentación. La información se complementa con los Índices de Precios Implícitos, los cuales surgen de relacionar los valores corrientes con los constantes del PIB para cada trimestre.

El cálculo a precios constantes parte de la elaboración de índices mensuales y trimestrales de volumen físico de la producción, de formulación Laspeyres, que tiene su base fija en el año de 2003, para cada una de las clases de actividad preseleccionadas, con los cuales se extrapolan los respectivos valores agregados en cada trimestre.

Cabe señalar que en la Agricultura los niveles registrados difieren de los obtenidos mediante el cálculo anual, debido a que en este último se cuantifica la producción del "año agrícola", mientras que trimestralmente

se mide el valor agregado en cada uno de los trimestres comprendidos en un año calendario.

Respecto al PIB Nominal, o a precios corrientes, éste se elabora a partir de los cálculos a precios constantes. Para tal efecto se lleva a cabo la evaluación y selección de un conjunto de indicadores de precios y salarios que tienen como fuentes principales al Banco de México, PEMEX, Comisión Federal de Electricidad, y la Encuesta Industrial Mensual, entre otros.

La elección de indicadores adecuados toma en cuenta la cobertura conceptual y metodológica de las actividades, así como el tipo de bienes y servicios producidos en ellas, para cada una de las clases de actividad económica SCIAN seleccionadas.

Los índices de valor y de precios seleccionados, con base 2003, permiten hacer una medición del PIB a precios corrientes, que se agrega de manera similar al método de los cálculos a precios constantes.

La información se complementa con los Índices de Precios Implícitos, los que resultan de relacionar los datos en valores nominales o corrientes de cada año, con los medidos a precios constantes de 2003, por tal razón los índices de precios implícitos quedan expresados con base 100 en el año de 2003.

No obstante, para evitar los problemas derivados del uso de metodologías y/o fuentes de información diferentes, las estimaciones de corto plazo se

articulan a los niveles anuales tan pronto como se producen los resultados definitivos.

Los resultados de PIB a precios constantes de cada trimestre se presentan 45 días después de terminado el periodo en cuestión. Para ello se hace necesario que una parte de los datos utilizados en el cálculo, particularmente los correspondientes al último mes del periodo, sean preliminares.

Los datos de PIB a precios corrientes de cada trimestre se divulgan una semana después de terminado el cálculo a precios constantes. Cuando se conforma el cuarto trimestre de cada año, además de obtenerse los datos preliminares del año en estudio, también se procede a la articulación de los resultados anuales de los años inmediatos anteriores, pudiendo ser revisados los dos últimos años.

Los datos de cada trimestre, tanto corrientes como constantes, se presentan en millones de pesos en términos anualizados, es decir multiplicados por cuatro, con objeto de expresar el nivel que alcanzaría la economía del país o cualquier sector económico, si en el resto del año se mantuvieran las condiciones observadas en el trimestre en estudio.

3.5. Indicador Global de la Actividad Económica, serie mensual

Con el fin de proporcionar información que permita un conocimiento mucho más oportuno sobre la evolución de la actividad económica del

país, el INEGI calcula el Indicador Global de la Actividad Económica (IGAE) de periodicidad mensual.

Este indicador muestra la evolución de la actividad económica del país, a través de un índice de volumen físico que se elabora, al igual que el PIB, utilizando el esquema conceptual y metodológico del SCN, así como la clasificación por actividades económicas del SCIAN y las fuentes de información que cuentan con una oportunidad mensual.

Sus fuentes de información son: encuestas sectoriales del INEGI, instituciones públicas y privadas, organismos públicos, cámaras y asociaciones, así como empresas públicas y privadas.

La elaboración de este indicador emplea datos estadísticos provenientes de actividades de los diferentes sectores de actividad, los cuales se listan a continuación:

- Actividades Primarias
 - Agricultura
 - Ganadería
 - Aprovechamiento forestal
 - Pesca y caza
- Actividades Secundarias o Industriales

- Minería
- Electricidad
- Agua y suministro de gas por ductos al consumidor final
- Construcción
- Industrias manufactureras
- Actividades Terciarias o de Servicios
 - Comercio
 - Transportes, correos y almacenamiento
 - Información en medios masivos
 - Servicios financieros y de seguros
 - Servicios inmobiliarios y de alquiler de bienes muebles e intangibles
 - Servicios profesionales, científicos y técnicos
 - Servicios de apoyo a los negocios
 - Servicios educativos; Servicios de salud
 - Servicios de esparcimiento, culturales, deportivos, y otros servicios recreativos

- Servicios de alojamiento temporal y de preparación de alimentos y bebidas
- Otros servicios excepto actividades del Gobierno
- Actividades del Gobierno

Al igual que en el PIB trimestral se incorpora una industria ficticia de ajuste, referida a los Servicios de Intermediación Financiera Medidos Indirectamente.

Con estas características, los resultados del IGAE presentan una alta correlación con los del PIB trimestral, por lo cual se constituye en un excelente indicador del comportamiento del mismo.

Es importante destacar que la información básica que incorpora el IGAE contiene datos muy preliminares y está sujeta a revisión por parte de las empresas y organismos públicos y privados; adicionalmente, no incluye la totalidad de las actividades como lo hace el PIB trimestral. Por ello, los resultados del IGAE pueden diferir con los del PIB trimestral y debe considerársele como un indicador de la tendencia o dirección de la actividad económica del país en el corto plazo.

Los resultados del IGAE que se presentan, se expresan en índices de volumen físico base 2003=100, la serie inicia en enero de 2003, y están disponibles 53 días después de concluido el mes de referencia.

4. Mensualización del PIB a precios constantes

Uno de los problemas recurrentes en el campo de la estadística y la econometría es la falta de información disponible por parte de las fuentes primarias, más aún, el nivel de desagregación en la periodicidad de las variables relevantes, es de suma importancia para evitar posibles errores de estimación obteniendo un mayor número de observaciones en la muestra representativa.

Es por ello que en este capítulo, se describe el desarrollo empírico que se siguió para generar una serie de tiempo mensual para el PIB real con base 2003, que va de enero de 2003 a diciembre de 2007, para después describir el proceso de selección del modelo econométrico adecuado con el que se pudo ajustar dicha serie de tiempo y realizar los pronósticos para el año 2008.

4.1. Ajuste empírico.

Para la estimación de la serie del PIB real mensual, se cuenta con la información de la serie del IGAE mensual, del mes de enero de 2003 al mes de marzo de 2008 (**ANEXO A**); así como con la serie del PIB real trimestral, que comprende primer trimestre de 2003 al primer trimestre de 2008 (**ANEXO B**).

A partir de la serie del IGAE mensual se calcula una de corte trimestral (**ANEXO C**), como un promedio simple de la agregación de las observaciones mensuales:

$$IGAE_j = \frac{\sum_{i=1}^n IGAE_{ij}}{3}, \text{ en donde:}$$

$IGAE_j$ = IGAE del trimestre j

$IGAE_{ij}$ = IGAE del mes i correspondiente al trimestre j

Una vez definida la serie del IGAE trimestral, se calcula una serie del PIB real mensual, usando como ponderador del PIB real trimestral la proporción que representa el IGAE mensual con relación del trimestral respectivo:

$$PIB_{ij}^R = PIB_j^{RP} \left(\frac{IGAE_{ij}}{IGAE_j} \right), \text{ en donde:}$$

PIB_{ij}^R = PIB real del mes i correspondiente al trimestre j

PIB_j^{RP} = PIB real publicado del trimestre j

$IGAE_{ij}$ = IGAE del mes i correspondiente al trimestre j

$IGAE_j$ = IGAE del trimestre j

La serie estimada para el PIB real mensual se presenta en el **ANEXO D**. Con esta serie es posible calcular el PIB trimestral como un promedio simple con los meses respectivos, de tal forma que es posible comprobar que

mediante esta agregación, se obtiene una serie trimestral idéntica a la publicada (**ANEXO E**), formalmente tenemos:

$$PIB_j^R = \frac{\sum_{i=1}^n PIB_{ij}^R}{3}, \text{ en donde:}$$

$$PIB_j^R = \text{PIB real del trimestre } j$$

$$PIB_{ij}^R = \text{PIB real del mes } i \text{ correspondiente al trimestre } j$$

comprobándose que el $PIB_j^R = PIB_j^{RP}$.

El PIB real mensual puede representarse mediante un índice de quantum tipo Laspeyres, debido a que se consideran fijos los precios relativos en la base y el único cambio que se observa es el que se realiza en las cantidades, de tal forma que se considera de la siguiente manera para su cálculo:

$$IPIB_i^R = \left(\frac{PIB_i^R}{PIB_0^{2003}} \right) 100 \text{ donde}$$

$$IPIB_i^R = \text{Índice de Volumen Físico del PIB real mensual}$$

$$PIB_i^R = \text{PIB real mensual}$$

$$PIB_0^{2003} = \text{PIB anual del año base 2003}$$

recordemos que lo anterior se sustenta en que el PIB real trimestral es resultado de la extrapolación a través de índices tipo Laspeyres, donde

$$PIB_j^{RP} = \frac{Q_t^L}{Q_0^L} = \frac{\left(\frac{\sum_{i=1}^N q_{it} P_{i0}}{\sum_{i=1}^N q_{i0} P_{i0}} \right) 100}{\left(\frac{\sum_{i=1}^N q_{i0} P_{i0}}{\sum_{i=1}^N q_{i0} P_{i0}} \right) 100} = \left(\frac{\sum_{i=1}^N q_{it} P_{i0} \sum_{i=1}^N q_{i0} P_{i0}}{\sum_{i=1}^N q_{i0} P_{i0} \sum_{i=1}^N q_{i0} P_{i0}} \right)$$

donde al simplificar tenemos

$$PIB_j^{RP} = \left(\frac{\sum_{i=1}^N q_{it} P_{i0}}{\sum_{i=1}^N q_{i0} P_{i0}} \right) 100$$

así pues, también podemos decir que

$$IPIB_i^R = \left(\frac{PIB_i^R}{PIB_0^{2003}} \right) 100 = \left(\frac{\sum_{i=1}^N q_{it} P_{i0}}{\sum_{i=1}^N q_{i0} P_{i0}} \right) 100$$

En el **ANEXO F** se presenta la serie del índice del PIB real mensual comparado con el IGAE mensual respectivo, donde también se comparan sus crecimientos anuales.

Si calculamos un promedio simple para la serie del Índice del PIB real mensual correspondiente a sus trimestres respectivos, podemos obtener una serie del índice del PIB real trimestral, que puede ser comparada con la publicada (**ANEXO G**). Para demostrar nuevamente la igualdad entre la agregación trimestral de los cálculos de la serie del PIB real mensual con la trimestral publicada, matemáticamente tenemos:

$$IPIB_j^R = \frac{\sum_{i=1}^n IPIB_{ij}^R}{3}, \text{ en donde:}$$

$IPIB_j^R$ = Índice del PIB real del trimestre j

$IPIB_{ij}^R$ = Índice del PIB real del mes i correspondiente al trimestre j

comprobándose que $IPIB_j^R = IPIB_j^{RP}$ al igual que $PIB_j^R = PIB_j^{RP}$.

En el **ANEXO H** se presenta la serie del índice del PIB real trimestral comparada con la del IGAE trimestral, comparando además sus crecimientos anuales.

4.2. Ajuste econométrico y proyecciones

Una vez que se estimó la serie para el PIB real mensual, consideraremos como universo para nuestro estudio econométrico, las observaciones comprendidas en el periodo de enero de 2003 a diciembre de 2007, con el objeto de poder comparar la eficiencia de los resultados pronosticados

para los meses del primer trimestre de 2008. Es importante mencionar que para todos los procesos del modelado y sus pruebas, se utilizará el software econométrico EViews en su versión 5.0.

De acuerdo con la metodología propuesta, incluiremos modelos candidatos, tanto **SARIMA** como **ARIMAX**, con el objeto de elegir aquel que presente la mayor consistencia y eficiencia, con base en los resultados de las pruebas respectivas.

La serie para el periodo propuesto presenta estacionalidad y tendencia, pero no estacionariedad, tal como se puede observar en la Gráfica del **ANEXO I**. Aplicando la prueba de Dickey-Fuller sobre los niveles de la serie, se obtiene el estadístico -0.143815, el cual es menor en términos absolutos que el valor crítico de Mackinnon -2.915522 al 5%, lo que comprueba la no estacionariedad de la serie.

Con base en la evidencia anterior, se debe aplicar la primera diferencia para inducir la estacionariedad y cumplir con los requisitos de los procesos de Box-Jenkins.

Verificando nuevamente con la prueba de Dickey-Fuller sobre la primera diferencia de la serie se obtiene el estadístico -8.579411, el cual es mayor en términos absolutos que el valor crítico de Mackinnon -2.923780 al 5%, por lo que se comprueba la estacionariedad de la serie en primera diferencia.

El resultado de esta prueba se presenta completo en el **ANEXO J**, en el cuál también se muestran que todos coeficientes son significativos al

presentar valores estadísticos t mayores a 1.96, así como un coeficiente de determinación ajustado (R^2) de 0.921755 y un valor para el BIC de 27.15830.

La serie del PIB real mensual es $I(1)$, es decir su orden de integración es uno, de acuerdo con el estadístico de prueba de Dickey-Fuller.

Una vez satisfecha la condición de estacionariedad para la serie, se construye el correlograma (**ANEXO K**) de la misma en primera diferencia, con el fin de analizar la ACF y PACF, para identificar el proceso **ARMA** y determinar los valores asociados a los términos **AR** y **MA**.

A través del análisis del correlograma se identifican los siguientes candidatos **AR** y **MA**: AR(1), AR(4), AR(6), AR(10), AR(11), AR(16), AR(22), AR(23), AR(24), MA(1), MA(5), MA(6), MA(7), MA(12) y MA(13).

Partiendo de la identificación anterior, se procede a establecer modelos candidatos del tipo **SARIMA** y **ARIMAX**, dado que se ha podido observar un fuerte componente estacional.

Los modelos **SARIMA** candidatos que pasaron las pruebas a) significancia individual para cada uno de sus coeficientes; b) condiciones de estacionariedad e invertibilidad, y; c) probabilidades del correlograma de los residuales; son los siguientes:

Modelo	Procesos	R ²	R ² ajustado	AIC	BIC
1	AR(1), AR(11), MA(8), MA(11), SMA(12)	0.731554	0.706582	26.82108	27.01600
2	AR(1), AR(11), MA(11), SMA(12)	0.716971	0.697673	26.83231	26.98825
3	C, MA(1), MA(12), SMA(12)	0.760839	0.747794	26.54693	26.68778
4	C, MA(1), MA(12), SMA(12)	0.798626	0.783710	26.40886	26.58492
5	AR(1), AR(5), MA(15), SAR(12)	0.916124	0.909503	25.58692	25.75241

Para caso de los **ARIMAX**, satisfaciendo las mismas pruebas tenemos:

Modelo	Procesos	R ²	R ² ajustado	AIC	BIC
1	@SEAS(1), @SEAS(2), @SEAS(3), @SEAS(5), @SEAS(9), @SEAS(10), AR(23), MA(2), MA(6)	0.887253	0.853847	26.30896	26.70484
2	@SEAS(1), @SEAS(2), @SEAS(3), @SEAS(4), @SEAS(5), @SEAS(10), AR(1), MA(1), MA(12)	0.866577	0.844794	26.14392	26.46364
3	@SEAS(1), @SEAS(2), @SEAS(3), @SEAS(4), @SEAS(5), @SEAS(9), @SEAS(10), AR(1), MA(6), MA(1)	0.903592	0.885516	25.85346	26.20871
4	@SEAS(1), @SEAS(2), @SEAS(3), @SEAS(5), @SEAS(9), @SEAS(10), AR(5), AR(23), MA(1), MA(2), MA(5), MA(6), MA(8), MA(13), MA(14)	0.965954	0.943257	25.44485	26.10465

Los modelos seleccionados, con base en el criterio de información de Schwartz²⁵, son el número 5 para el método **SARIMA** y el número 4 para el **ARIMAX**. Al final se elegirá como mejor modelo, bajo el mismo criterio, el modelo **SARIMA** antes mencionado, por presentar el BIC más bajo de todos los modelos candidatos; sin embargo, para fines ilustrativos se describirán las características principales del modelado de ambos casos.

Modelo SARIMA seleccionado

El periodo de observaciones considerado es de enero de 2003 a diciembre de 2007 y el mejor periodo de pronóstico es de marzo de 2006 a diciembre de 2008, con base en los resultados de la prueba de Theil (**ANEXO L**), donde:

Concepto	Valor
Coefficiente de desigualdad de Theil	0.008844
Proporción de sesgo	0.037936
Proporción de varianza	0.132995
Proporción de covarianza	0.829069

La especificación del modelo de acuerdo con los procesos incluidos, que satisface las condiciones de Box-Jenkins, es:

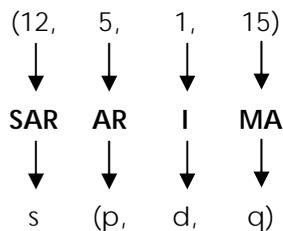
²⁵ Como se mencionó en el capítulo 3, éste es consistente y penaliza con máxima intensidad los grados de libertad, considerando irrelevante que no sea eficientemente asintótico, debido a que la muestra de observaciones no es grande, en cuyo caso se podía preferir el AIC.

$$d(\text{PIB}) = -\phi_1 \text{AR}(1) + \phi_5 \text{AR}(5) + \theta_{15} \text{MA}(15) + \phi_{12}^s \text{AR}(12)$$

$$d(\text{PIB}) = -\phi_1 Y_{t-1} + \phi_5 Y_{t-5} + \theta_{15} U_{t-15} + \phi_{12}^s Y_{t-12}$$

$$d(\text{PIB}) = -0.640701 Y_{t-1} + 0.281914 Y_{t-5} + 0.919819 U_{t-15} + 0.892205 Y_{t-12}$$

Una vez especificado señalamos el orden del modelo como:



Por último, como se vio con anterioridad, se verifican en el modelo las tres condiciones siguientes:

- I. Todos los parámetros son estadísticamente significativos al 95% de confianza. (**ANEXO M**)
- II. Todas las raíces de los procesos **AR** y **MA** son $< \|1\|^{26}$ a fin de garantizar estacionariedad e invertibilidad. (**ANEXO N**)
- III. De acuerdo con el correlograma de los residuales, todas las autocorrelaciones tanto simples como parciales son estadísticamente no significativas ($P > 0.05$) a fin de garantizar que se están

²⁶ Uno en módulo, que incluye a los números complejos.

considerando todos los procesos **AR** y de **MA** en el modelo de pronóstico. (**ANEXO O**)

Una vez verificadas todas las pruebas, se realizan los pronósticos correspondientes (**ANEXO P**), que se comparan contra los ya publicados para 2008, tanto mensuales como trimestrales, con el objeto de observar las desviaciones de la estimación con los datos disponibles. Esta información se presenta en el **ANEXO Q**, en forma gráfica se verifica la comparación de los valores en el **ANEXO R**.

Modelo ARIMAX seleccionado

El periodo de observaciones considerado es de enero de 2003 a diciembre de 2007 y el mejor periodo de pronóstico es de octubre de 2006 a diciembre de 2008, con base en los resultados de la prueba de Theil (**ANEXO S**), donde:

Concepto	Valor
Coefficiente de desigualdad de Theil	0.004556
Proporción de sesgo	0.036312
Proporción de varianza	0.004031
Proporción de covarianza	0.959657

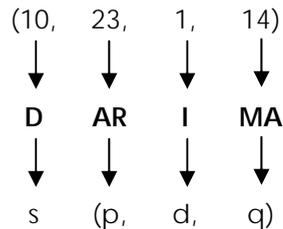
La especificación del modelo de acuerdo con los procesos incluidos, que satisface las condiciones de Box-Jenkins, es:

$$d(PIB) = -\alpha_1 D_1 - \alpha_2 D_2 + \alpha_3 D_3 + \alpha_5 D_5 - \alpha_9 D_9 + \alpha_{10} D_{10} - \phi_5 AR(5) + \\ + \phi_{23} AR(23) + \theta_1 MA(1) - \theta_2 MA(2) - \theta_5 MA(5) + \theta_6 MA(6) - \theta_8 MA(8) + \\ + \theta_{13} MA(13) - \theta_{14} MA(14)$$

$$d(PIB) = -\alpha_1 D_1 - \alpha_2 D_2 + \alpha_3 D_3 + \alpha_5 D_5 - \alpha_9 D_9 + \alpha_{10} D_{10} - \phi_5 Y_{t-5} + \\ + \phi_{23} Y_{t-23} + \theta_1 U_{t-1} - \theta_2 U_{t-2} - \theta_5 U_{t-5} + \theta_6 U_{t-6} - \theta_8 U_{t-8} + \theta_{13} U_{t-13} - \\ - \theta_{14} U_{t-14}$$

$$d(PIB) = 181359.1D_1 - 261097.0D_2 + 417811.9D_3 + 295237.3D_5 - \\ - 243019.9D_9 + 292005.2D_{10} - 0.359960Y_{t-5} + 0.497774Y_{t-23} + \\ + 1.465726U_{t-1} - 0.640722U_{t-2} - 0.529714U_{t-5} + 0.976701U_{t-6} - \\ - 0.196006U_{t-8} + 0.302427U_{t-13} - 0.433351U_{t-14}$$

Una vez especificado señalamos el orden del modelo como:



Por último, como se vio con anterioridad, se verifican en el modelo las tres condiciones siguientes:

- I. Todos los parámetros son estadísticamente significativos al 95% de confianza. **(ANEXO T)**

- II. Todas las raíces de los procesos **AR** y **MA** son $< \|1\|$ a fin de garantizar estacionariedad e invertibilidad. **(ANEXO U)**

III. De acuerdo con el correlograma de los residuales, casi todas las autocorrelaciones, tanto simples como parciales son estadísticamente no significativas ($P > 0.05$) a fin de garantizar que se están considerando todos los procesos **AR** y de **MA** en el modelo de pronóstico (**ANEXO V**). Nótese que la primera autocorrelación correspondiente a esta serie de residuales, no satisface el valor sugerido asociado a la probabilidad (valor $p > 0.05$), reportando un valor $p = 0.022$; no obstante, los pronósticos que se obtienen con este modelo predictivo son muy satisfactorios, por lo que recurriendo al principio de parsimonia, considerando la objetividad del modelo, se considera oportuno incluirlo con el propósito de poder realizar comparaciones ilustrativas respecto de los resultados derivados del modelo **SARIMA** seleccionado.

Una vez más, verificadas todas las pruebas, se realizan los pronósticos correspondientes (**ANEXO W**), que se comparan contra los ya publicados para 2008, tanto mensuales como trimestrales, con el objeto de observar las desviaciones de la estimación con los datos disponibles. Esta información se presenta en el **ANEXO X**, en forma gráfica se ilustra la comparación de los valores en el **ANEXO Y**.

Elección del modelo

Después de analizar los dos modelos de pronóstico, en cuanto a sus estadísticos de prueba, como ya se había mencionado, se puede determinar que el modelo **SARIMA** aunque tiene un R^2 ajustado más bajo que el del modelo **ARIMAX**, 91% vs. 94% respectivamente, tiene también un valor más bajo en el Criterio de información de Schwartz, 25.75241 vs. 26.10465 en el mismo orden, por lo tanto se infiere que tiene un error cuadrático medio más bajo, lo que permite generar pronósticos más precisos para el corto y el mediano plazo, tal como se demostró, motivo por el que se elige el modelo **SARIMA**. En el **ANEXO Z** se muestra una comparación gráfica de los pronósticos de ambos modelos.

5. Conclusiones y recomendaciones.

Una vez seleccionado el modelo **SARIMA** y comparado contra los pronósticos que ofrece el **ARIMAX**, podemos observar su comportamiento y concluir que el modelo elegido ofrece un escenario más conservador a lo largo del año 2008. Obteniendo una expectativa de crecimiento para 2008 de 2.8 puntos porcentuales con el modelo adoptado, contra 4.0 por ciento del de referencia.

Este resultado del modelo designado es congruente con las proyecciones de crecimiento anual para el periodo 2007-2008, que de manera oficial, realiza tanto el BANXICO que considera será entre 2.4 y 2.9 por ciento, como la Secretaría de Hacienda y Crédito Público (SHCP), que lo estima en 2.8 por ciento.

La desviación que ofrece el pronóstico trimestral del modelo **SARIMA** contra los datos publicados, es menor y por lo tanto más acertada que para el **ARIMAX**, aunque éste último presenta una menor variación para los meses de enero y marzo, pero que es poco significativa.

Es importante resaltar que la diferencia porcentual entre los datos publicados y los estimados para el mes de marzo, es la que mayor desviación presenta, cerca de 3.5 por ciento. Esta situación puede ser explicada a partir del número de observaciones que contiene la serie original, considerando las características particulares de este mes a lo largo de la misma.

La principal de las características que contribuyeron a esta diferencia, se refieren al efecto del número de días efectivamente trabajados, debido a la semana santa y a la celebración del natalicio de Benito Juárez, lo que ocasionó un puente en la mayoría de los casos, del lunes 17 al viernes 21 marzo, así como la conjunción de dos fines de semana, uno al principio y otro al final de este mismo periodo, teniendo así, alrededor de cinco a nueve días inhábiles, situación única durante la serie, por lo que no es posible capturar todo su efecto en los pronósticos.

La primera conclusión a la que podemos llegar, es que los modelos **ARIMAX** tienen un mejor resultado de pronósticos en el muy corto plazo, ya que en la medida en que se comienza a pronosticar en el tiempo y la muestra de observaciones en la serie es pequeña, no se captura con eficacia los efectos de los procesos que incluye, es decir, es de memoria corta.

Por otro lado, podemos ver que con los modelos **SARIMA**, es posible generar pronósticos de corto y mediano plazo, con resultados más confiables, ya que capturan mejor los efectos de los procesos incluidos, aún con series con pocas observaciones.

También es importante resaltar, que gracias al procedimiento de mensualización que se propone, es posible contar con una muestra mayor de observaciones, pasando de 20 que se capturan con periodicidad trimestral, a 60 que se obtienen mediante la mensualización. Lo que

permite una mayor bondad de ajuste en los pronósticos y estimaciones realizadas.

Este procedimiento demostró ajustarse perfectamente a los valores tanto trimestrales como anuales, permitiendo capturar la estacionalidad que se presenta en cada uno de los meses del año, como por ejemplo el caso del mes de marzo mencionado con anterioridad.

Queda pues, sentada la base con la que se puede mensualizar, de manera precisa, el PIB real a partir de datos trimestrales, por lo que será necesario alimentar esta base con los datos del IGAE mensual y el PIB real trimestral, conforme se pongan a disponibilidad por parte del INEGI.

Es importante remarcar, que conforme se vaya actualizando y ampliando la base del PIB real mensual, será también necesario establecer nuevos modelos candidatos para poder estimar los pronósticos más confiables para periodos venideros.

Recordemos que como mencionamos, conforme se aumentan las observaciones de la base, pueden resultar más eficaces en algún momento, los modelos **ARIMAX** en el corto plazo, ya que son de memoria corta, y los **SARIMA** en caso de pronósticos para periodos un poco más extensos. Por lo que será necesario volver a modelar bajo ambos procedimientos, generando los candidatos respectivos, para definir cual resultará más confiable de acuerdo con el plazo en el cual se quiere pronosticar.

Esta última actividad es ardua, pero necesaria, ya que como se vio a lo largo del desarrollo del presente trabajo, se trata de procesos iterativos, que es preciso establecer mediante procesos de prueba y error hasta encontrar el modelo con las características adecuadas y que además cumplan con los requisitos establecidos.

Como se explica con anterioridad, una de las características más importantes de esta aportación, es que se establece una metodología sistemática que permite explicar los valores futuros del PIB real, a partir de los datos contenidos en la serie de tiempo.

Queda pues, como antecedente, para trabajos posteriores en los que se decida incluir la parte estructural, definiendo la causalidad de las variaciones del PIB real, con la construcción de modelos combinados de regresión y series de tiempo.

Por último, sólo resta concluir que se cumplió con la expectativa que se planteó en los objetivos de este trabajo y la satisfacción de poder aportar un método que puede servir para incrementar la periodicidad de variables que aún se encuentran sólo para periodicidades menores.

Bibliografía

1. Abel, Andrew B.; Bernanke, Ben S. (2004). *Macroeconomía*. España: Pearson Educación.
2. Almagro Vázquez, Francisco (2004). *El Sistema de Cuentas Nacionales y sus Aplicaciones*. México: Instituto Politécnico Nacional.
3. Canavos, George C. (1996). *Probabilidad y Estadística Aplicaciones y métodos*. México: McGraw-Hill.
4. Cochrane, John H. (1997). "*Time Series for Macroeconomics and Finance*". Estados Unidos: Universidad de Chicago. Edición Electrónica.
Texto completo en www-gsb.uchicago.edu/fac/john.cochrane/research/Papers/timeser1.pdf
5. Diebold, Francis (2001). *Elementos de Pronósticos*. México: International Thomson Editores.
6. Dornbusch, Rudiger; Fischer, Stanley; y, Startz Richard (2002). *Macroeconomía*. España: McGraw-Hill.
7. Guerrero, Víctor M. (1991). *Análisis Estadístico de Series de Tiempo Económicas*. México: UAM Iztapalapa.
8. Guerrero, Víctor M. (2003) "*Monthly disaggregation of a quarterly time series and forecasts of its unobservable monthly values*". *Journal of Official Statistics* 19, 215-235.

9. Gujarati, Damodar N. (1999). *Econometría*. Colombia: McGraw-Hill.
10. Hernández Sampieri, Roberto; Fernández Collado, Carlos; y, Baptista Lucio, Pilar (2002). *Metodología de la Investigación*. México: McGraw-Hill.
11. Instituto Nacional de Estadística, Geografía e Informática (2008). *Estadísticas Económicas. Producto Interno Bruto Trimestral Base 2003. Metodología*. México: INEGI.
12. Instituto Nacional de Estadística, Geografía e Informática (2008). *Sistema de Cuentas Nacionales. Oferta y Demanda Global a Precios Corrientes Base 2003. Metodología*. México: INEGI.
13. Instituto Nacional de Estadística, Geografía e Informática (2008). *Sistema de Cuentas Nacionales. Oferta y Utilización Trimestral a Precios Constantes de 2003. Metodología*. México: INEGI.
14. Jiménez, Felix (1999). *Agregados Macroeconómicos y Contabilidad Nacional*. Edición electrónica.
Texto completo en <http://www.pucp.edu.pe/departamento/economia/images/documentos/DDD168.pdf>
15. Kazmier, Leonard J. (1999). *Estadística Aplicada a la Administración y a la Economía*. México: McGraw-Hill.
16. Kohler, Heinz (2005). *Estadística para Negocios y Economía*. México: CECOSA.

17. Ludlow Wiechevs, Jorge (1999). *Econometría, Modelos y Pronósticos*. México: UAM.
18. Massad, Carlos; Pattillo, Guillermo (2000). *Macroeconomía en un Mundo Interdependiente*. Chile: McGraw-Hill.
19. Naciones Unidas (2003). *System of National Accounts 1993, Handbook*. Estados Unidos: ONU.
20. Pindyck, Robert S.; Rubinfeld, Daniel L. (2001). *Econometría: Modelos y Pronósticos*. México: McGraw-Hill.
21. Ruiz Muñoz, David; Sánchez Sánchez; Ana María (2006). *Apuntes de Estadística*. Edición electrónica.
Texto completo en www.eumed.net/libros/2006/rms/.
22. Samuelson, Paul A.; Nordhaus, William D.; Dieck, Lourdes; y, Salazar, José de Jesús (1997). *Macroeconomía con aplicaciones a México*. México: McGraw-Hill.
23. Sánchez Fernández, Jesús (2004). *Introducción a la Estadística Empresarial*. Edición Electrónica.
Texto completo en www.eumed.net/cursecon/libreria/index.htm
24. Sérúzier, Michel (2003); *Medir la Economía de los Países según el Sistema de Cuentas Nacionales*. Colombia: CEPAL en coedición con Alfaomega Colombiana.

25. Wooldridge, Jeffrey M. (2002); *Introducción a la Econometría: Un enfoque moderno*. México: Thomson Learning.

ANEXOS

Nota: en este apartado se utiliza la notación ND para datos no disponibles y NA cuando no aplica la presentación de algún dato.

A. IGAE mensual, enero 2003 – marzo 2008

Mes	Año					
	2003	2004	2005	2006	2007	2008
Enero	97.3	98.9	102.4	109.2	112.0	112.0
Febrero	94.8	98.0	101.1	105.5	107.7	107.7
Marzo	100.1	105.4	104.5	113.1	115.8	115.8
Abril	97.9	101.2	106.7	108.4	111.3	ND
Mayo	100.9	104.1	107.9	115.6	118.4	ND
Junio	100.7	105.7	108.0	115.4	118.4	ND
Julio	101.6	104.4	106.4	112.8	118.1	ND
Agosto	99.5	104.9	109.7	115.2	118.8	ND
Septiembre	98.1	103.1	107.2	111.3	114.4	ND
Octubre	103.1	105.7	109.9	116.1	121.9	ND
Noviembre	101.8	108.9	112.5	116.4	121.1	ND
Diciembre	104.2	108.5	112.5	114.9	119.3	ND

B. PIB real trimestral, 1er Trimestre 2003 – 1er Trimestre 2008

Trimestre	Año					
	2003	2004	2005	2006	2007	2008
Primer	7,367,746	7,617,682	7,767,347	8,233,098	8,438,636	8,660,282
Segundo	7,540,240	7,819,639	8,108,631	8,521,543	8,743,224	ND
Tercer	7,535,231	7,866,297	8,129,515	8,528,532	8,820,493	ND
Cuarto	7,779,997	8,127,261	8,409,225	8,721,858	9,090,534	ND

C. IGAE trimestral, 1er Trimestre 2003 – 1er Trimestre 2008

Trimestre	Año					
	2003	2004	2005	2006	2007	2008
Primer	97.4	100.8	102.7	109.3	111.8	114.8
Segundo	99.8	103.7	107.5	113.1	116.0	ND
Tercer	99.7	104.1	107.8	113.1	117.1	ND
Cuarto	103.0	107.7	111.6	115.8	120.8	ND

D. PIB real mensual, enero 2003 – marzo 2008

Mes	Año					
	2003	2004	2005	2006	2007	2008
Enero	97.3	98.9	102.4	109.2	112.0	112.0
Febrero	94.8	98.0	101.1	105.5	107.7	107.7
Marzo	100.1	105.4	104.5	113.1	115.8	115.8
Abril	97.9	101.2	106.7	108.4	111.3	ND
Mayo	100.9	104.1	107.9	115.6	118.4	ND
Junio	100.7	105.7	108.0	115.4	118.4	ND
Julio	101.6	104.4	106.4	112.8	118.1	ND
Agosto	99.5	104.9	109.7	115.2	118.8	ND
Septiembre	98.1	103.1	107.2	111.3	114.4	ND
Octubre	103.1	105.7	109.9	116.1	121.9	ND
Noviembre	101.8	108.9	112.5	116.4	121.1	ND
Diciembre	104.2	108.5	112.5	114.9	119.3	ND

E. Comparación del PIB real trimestral publicado con el agregado a partir del mensual, 1er Trimestre 2003 – 1er Trimestre 2008

Periodo		PIB		
Año	Trimestre	Publicado	Estimado	Diferencia
2003	1	7,367,746	7,367,746	0
	2	7,540,240	7,540,240	0
	3	7,535,231	7,535,231	0
	4	7,779,997	7,779,997	0
2004	1	7,617,682	7,617,682	0
	2	7,819,639	7,819,639	0
	3	7,866,297	7,866,297	0
	4	8,127,261	8,127,261	0
2005	1	7,767,347	7,767,347	0
	2	8,108,631	8,108,631	0
	3	8,129,515	8,129,515	0
	4	8,409,225	8,409,225	0

Periodo		PIB		
Año	Trimestre	Publicado	Estimado	Diferencia
2006	1	8,233,098	8,233,098	0
	2	8,521,543	8,521,543	0
	3	8,528,532	8,528,532	0
	4	8,721,858	8,721,858	0
2007	1	8,438,636	8,438,636	0
	2	8,743,224	8,743,224	0
	3	8,820,493	8,820,493	0
	4	9,090,534	9,090,534	0
2008	1	8,660,282	8,660,282	0
	2	ND	ND	ND
	3	ND	ND	ND
	4	ND	ND	ND

F. Comparación del índice del PIB real mensual con el IGAE mensual, enero 2003 – marzo 2008

Periodo		Indicador		Crecimiento Anual	
Año	Mes	PIB real	IGAE	PIB real	IGAE
2003	Enero	97.4	97.3	NA	NA
	Febrero	94.9	94.8	NA	NA
	Marzo	100.2	100.1	NA	NA
	Abril	97.9	97.9	NA	NA
	Mayo	100.9	100.9	NA	NA
	Junio	100.7	100.7	NA	NA
	Julio	101.6	101.6	NA	NA
	Agosto	99.5	99.5	NA	NA
	Septiembre	98.1	98.1	NA	NA
	Octubre	103.0	103.1	NA	NA
	Noviembre	101.7	101.8	NA	NA
	Diciembre	104.1	104.2	NA	NA
2004	Enero	99.0	98.9	1.58	1.64
	Febrero	98.1	98.0	3.31	3.38
	Marzo	105.5	105.4	5.23	5.29
	Abril	101.0	101.2	3.24	3.37
	Mayo	103.9	104.1	3.04	3.17
	Junio	105.5	105.7	4.83	4.97

Periodo		Indicador		Crecimiento Anual	
Año	Mes	PIB real	IGAE	PIB real	IGAE
2004	Julio	104.4	104.4	2.74	2.76
	Agosto	104.9	104.9	5.41	5.43
	Septiembre	103.1	103.1	5.08	5.10
	Octubre	105.6	105.7	2.46	2.52
	Noviembre	108.8	108.9	6.91	6.97
	Diciembre	108.4	108.5	4.06	4.13
2005	Enero	102.5	102.4	3.62	3.54
	Febrero	101.2	101.1	3.24	3.16
	Marzo	104.6	104.5	-0.78	-0.85
	Abril	106.5	106.7	5.40	5.43
	Mayo	107.7	107.9	3.62	3.65
	Junio	107.8	108.0	2.14	2.18
	Julio	106.2	106.4	1.77	1.92
	Agosto	109.5	109.7	4.43	4.58
	Septiembre	107.0	107.2	3.83	3.98
	Octubre	109.6	109.9	3.79	3.97
	Noviembre	112.2	112.5	3.12	3.31
	Diciembre	112.2	112.5	3.50	3.69
2006	Enero	108.9	109.2	6.21	6.64
	Febrero	105.2	105.5	3.93	4.35
	Marzo	112.8	113.1	7.79	8.23
	Abril	108.1	108.4	1.48	1.59
	Mayo	115.2	115.6	7.02	7.14
	Junio	115.0	115.4	6.73	6.85
	Julio	112.6	112.8	5.97	6.02
	Agosto	115.0	115.2	4.97	5.01
	Septiembre	111.1	111.3	3.78	3.82
	Octubre	115.7	116.1	5.63	5.64
	Noviembre	116.0	116.4	3.45	3.47
	Diciembre	114.5	114.9	2.12	2.13
2007	Enero	111.9	112.0	2.71	2.56
	Febrero	107.6	107.7	2.23	2.09
	Marzo	115.6	115.8	2.53	2.39
	Abril	111.0	111.3	2.71	2.68
	Mayo	118.1	118.4	2.46	2.42
	Junio	118.1	118.4	2.64	2.60
	Julio	117.7	118.1	4.58	4.70
	Agosto	118.4	118.8	3.01	3.13
	Septiembre	114.0	114.4	2.67	2.79
	Octubre	121.4	121.9	4.93	5.00
	Noviembre	120.6	121.1	3.98	4.04

Periodo		Indicador		Crecimiento Anual	
Año	Mes	PIB real	IGAE	PIB real	IGAE
2007	Diciembre	118.9	119.3	3.77	3.83
2008	Enero	116.0	116.2	3.75	3.75
	Febrero	114.0	114.1	5.95	5.94
	Marzo	113.9	114.0	-1.55	-1.55
	Abril	ND	ND	ND	ND
	Mayo	ND	ND	ND	ND
	Junio	ND	ND	ND	ND
	Julio	ND	ND	ND	ND
	Agosto	ND	ND	ND	ND
	Septiembre	ND	ND	ND	ND
	Octubre	ND	ND	ND	ND
	Noviembre	ND	ND	ND	ND
	Diciembre	ND	ND	ND	ND

G. Comparación del Índice del PIB real trimestral publicado con el agregado a partir del Índice mensual, 1^{er} Trimestre 2003 – 1^{er} Trimestre 2008

Periodo		PIB		
Año	Trimestre	Publicado	Estimado	Diferencia
2003	1	97.5	97.5	0
	2	99.8	99.8	0
	3	99.7	99.7	0
	4	103.0	103.0	0
2004	1	100.8	100.8	0
	2	103.5	103.5	0
	3	104.1	104.1	0
	4	107.6	107.6	0
2005	1	102.8	102.8	0
	2	107.3	107.3	0
	3	107.6	107.6	0
	4	111.3	111.3	0
2006	1	109.0	109.0	0
	2	112.8	112.8	0
	3	112.9	112.9	0
	4	115.4	115.4	0

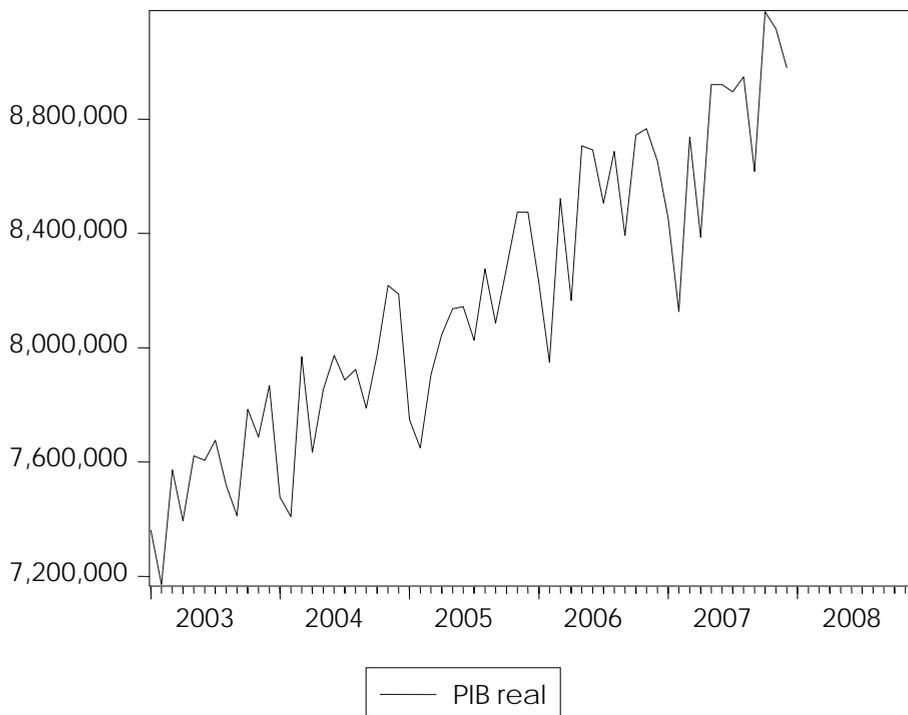
Periodo		PIB		
Año	Trimestre	Publicado	Estimado	Diferencia
2007	1	111.7	111.7	0
	2	115.7	115.7	0
	3	116.7	116.7	0
	4	120.3	120.3	0
2008	1	114.6	114.6	0
	2	ND	ND	ND
	3	ND	ND	ND
	4	ND	ND	ND

H. Comparación del índice del PIB real trimestral con el IGAE trimestral, 1^{er} Trimestre 2003 – 1^{er} Trimestre 2008

Periodo		Indicador		Crecimiento Anual	
Año	Trimestre	PIB real	IGAE	PIB real	IGAE
2003	1	97.5	97.4	NA	NA
	2	99.8	99.8	NA	NA
	3	99.7	99.7	NA	NA
	4	103.0	103.0	NA	NA
2004	1	100.8	100.8	3.39	3.46
	2	103.5	103.7	3.71	3.84
	3	104.1	104.1	4.39	4.41
	4	107.6	107.7	4.46	4.53
2005	1	102.8	102.7	1.96	1.89
	2	107.3	107.5	3.70	3.73
	3	107.6	107.8	3.35	3.49
	4	111.3	111.6	3.47	3.65
2006	1	109.0	109.3	6.00	6.43
	2	112.8	113.1	5.09	5.21
	3	112.9	113.1	4.91	4.95
	4	115.4	115.8	3.72	3.73
2007	1	111.7	111.8	2.50	2.35
	2	115.7	116.0	2.60	2.56
	3	116.7	117.1	3.42	3.54
	4	120.3	120.8	4.23	4.29

Periodo		Indicador		Crecimiento Anual	
Año	Trimestre	PIB real	IGAE	PIB real	IGAE
2008	1	114.6	114.8	2.63	2.62
	2	ND	ND	ND	ND
	3	ND	ND	ND	ND
	4	ND	ND	ND	ND

I. Gráfica de la serie del PIB real mensual, enero 2003 – diciembre 2007



J. Resultado de la prueba aumentada de Dickey-Fuller en primera diferencia para la serie del PIB real mensual

Null Hypothesis: D(DATOS2) has a unit root

Exogenous: Constant

Lag Length: 10 (Automatic based on SIC, MAXLAG=10)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-8.579411	0.0000
Test critical values:		
1% level	-3.574446	
5% level	-2.923780	
10% level	-2.599925	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(DATOS2,2)

Method: Least Squares

Date: 07/11/08 Time: 21:52

Sample (adjusted): 2004M01 2007M12

Included observations: 48 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(DATOS2(-1))	-10.18865	1.187570	-8.579411	0.0000
D(DATOS2(-1),2)	8.289107	1.133726	7.311386	0.0000
D(DATOS2(-2),2)	7.501690	1.073480	6.988198	0.0000
D(DATOS2(-3),2)	6.717012	0.989950	6.785205	0.0000
D(DATOS2(-4),2)	5.836334	0.895544	6.517083	0.0000
D(DATOS2(-5),2)	5.182675	0.768793	6.741316	0.0000
D(DATOS2(-6),2)	4.378990	0.650232	6.734507	0.0000
D(DATOS2(-7),2)	3.548669	0.498215	7.122763	0.0000
D(DATOS2(-8),2)	2.689251	0.365781	7.352084	0.0000
D(DATOS2(-9),2)	1.811989	0.236546	7.660195	0.0000
D(DATOS2(-10),2)	0.822454	0.127354	6.458016	0.0000
C	257329.8	36071.76	7.133828	0.0000
R-squared	0.940068	Mean dependent var	-6598.233	
Adjusted R-squared	0.921755	S.D. dependent var	486069.8	
S.E. of regression	135964.7	Akaike info criterion	26.69050	
Sum squared resid	6.66E+11	Schwarz criterion	27.15830	
Log likelihood	-628.5719	F-statistic	51.33452	
Durbin-Watson stat	2.014583	Prob(F-statistic)	0.000000	

K. Correlograma en primera diferencia de la serie del PIB real mensual

Date: 07/11/08 Time: 21:49
 Sample: 2003M01 2008M12
 Included observations: 59

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.486	-0.486	14.654	0.000
		2	0.075	-0.211	15.007	0.001
		3	-0.074	-0.182	15.360	0.002
		4	-0.238	-0.495	19.071	0.001
		5	0.413	0.007	30.452	0.000
		6	-0.347	-0.302	38.638	0.000
		7	0.356	0.074	47.419	0.000
		8	-0.199	-0.035	50.213	0.000
		9	-0.042	-0.001	50.339	0.000
		10	-0.035	-0.347	50.429	0.000
		11	-0.217	-0.536	53.954	0.000
		12	0.581	-0.055	79.770	0.000
		13	-0.259	0.085	85.029	0.000
		14	0.064	-0.008	85.355	0.000
		15	-0.071	0.151	85.770	0.000
		16	-0.132	0.272	87.229	0.000
		17	0.166	-0.030	89.580	0.000
		18	-0.063	0.247	89.924	0.000
		19	0.116	0.123	91.139	0.000
		20	-0.048	0.144	91.355	0.000
		21	-0.070	0.137	91.824	0.000
		22	-0.032	0.490	91.921	0.000
		23	-0.071	0.638	92.428	0.000
		24	0.216	1.481	97.226	0.000

L. Resultados de la Prueba de Theil del modelo SARIMA seleccionado

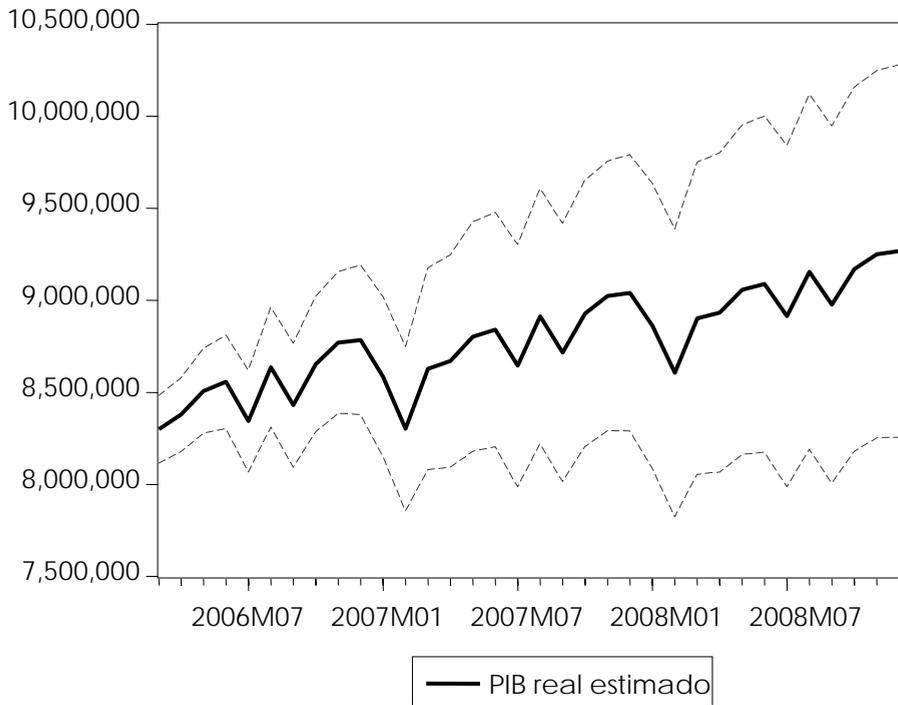
Forecast: PIB real estimado

Actual: PIB real

Forecast sample: 2006M03 2008M12

Included observations: 22

Root Mean Squared Error	153460.6
Mean Absolute Error	133349.1
Mean Absolute Percentage Error	1.543722
Theil Inequality Coefficient	0.008844
Bias Proportion	0.037936
Variance Proportion	0.132995
Covariance Proportion	0.829069



M. Especificaciones y resultados del modelo SARIMA seleccionado

Dependent Variable: D(DATOS2)
 Method: Least Squares
 Date: 07/12/08 Time: 14:53
 Sample (adjusted): 2004M07 2007M12
 Included observations: 42 after adjustments
 Convergence achieved after 10 iterations
 Backcast: 2003M04 2004M06

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.640701	0.109835	-5.833285	0.0000
AR(5)	0.281914	0.107379	2.625420	0.0124
SAR(12)	0.892205	0.078691	11.33814	0.0000
MA(15)	-0.919819	0.028842	-31.89139	0.0000
R-squared	0.916124	Mean dependent var		23979.01
Adjusted R-squared	0.909503	S.D. dependent var		276658.8
S.E. of regression	83226.68	Akaike info criterion		25.58692
Sum squared resid	2.63E+11	Schwarz criterion		25.75241
Log likelihood	-533.3252	Durbin-Watson stat		2.128029
Inverted AR Roots	.99	.86+.50i	.86-.50i	.68
	.50+.86i	.50-.86i	.13+.71i	.13-.71i
	.00-.99i	-.00+.99i	-.50-.86i	-.50+.86i
	-.79+.41i	-.79-.41i	-.86-.50i	-.86+.50i
	-.99			
Inverted MA Roots	.99	.91+.40i	.91-.40i	.67-.74i
	.67+.74i	.31-.95i	.31+.95i	-.10-.99i
	-.10+.99i	-.50-.86i	-.50+.86i	-.80+.58i
	-.80-.58i	-.97-.21i	-.97+.21i	

N. Raíces de los procesos AR y MA del modelo SARIMA seleccionado

Inverse Roots of AR/MA Polynomial(s)
 Specification: D(DATOS2) AR(1) AR(5) MA(15)
 SAR(12)
 Date: 07/13/08 Time: 09:29
 Sample: 2003M01 2008M12
 Included observations: 42

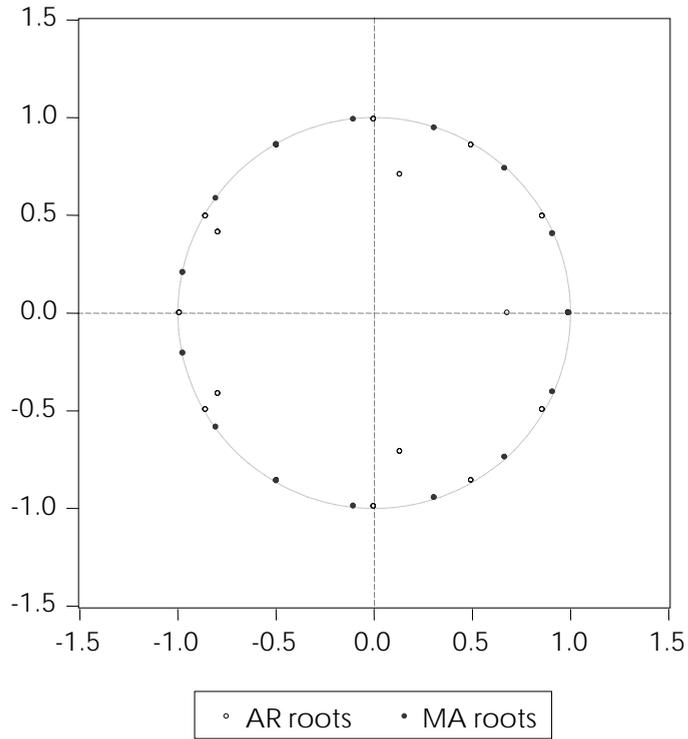
AR Root(s)	Modulus	Cycle
0.857833 ± 0.495270i	0.990540	12.00000
0.990540	0.990540	
-0.495270 ± 0.857833i	0.990540	3.000000
0.495270 ± 0.857833i	0.990540	6.000000
-3.26e-16 ± 0.990540i	0.990540	4.000000
-0.990540	0.990540	
-0.857833 ± 0.495270i	0.990540	2.400000
-0.792891 ± 0.413634i	0.894299	2.361431
0.132666 ± 0.707788i	0.720114	4.534931
0.679749	0.679749	

No root lies outside the unit circle.
 ARMA model is stationary.

MA Root(s)	Modulus	Cycle
-0.972713 ± 0.206756i	0.994444	2.142857
-0.497222 ± 0.861213i	0.994444	3.000000
0.665413 ± 0.739016i	0.994444	7.500000
0.307300 ± 0.945772i	0.994444	5.000000
-0.804522 ± 0.584519i	0.994444	2.500000
0.908469 ± 0.404477i	0.994444	15.00000
-0.103948 ± 0.988996i	0.994444	3.750000
0.994444	0.994444	

No root lies outside the unit circle.
 ARMA model is invertible.

Inverse Roots of AR/MA Polynomial(s)



O. Correlograma de los residuales del modelo SARIMA seleccionado

Date: 07/13/08 Time: 09:54

Sample: 2004M07 2007M12

Included observations: 42

Q-statistic probabilities adjusted for 4 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.076	-0.076	0.2621	
		2 0.002	-0.004	0.2623	
		3 0.160	0.161	1.4801	
		4 0.118	0.147	2.1536	
		5 -0.143	-0.129	3.1807	0.075
		6 -0.019	-0.077	3.1992	0.202
		7 -0.270	-0.340	7.0394	0.071
		8 -0.160	-0.229	8.4254	0.077
		9 -0.030	-0.028	8.4765	0.132
		10 -0.109	-0.002	9.1620	0.165
		11 -0.212	-0.094	11.852	0.106
		12 0.139	0.108	13.040	0.110
		13 0.112	0.134	13.841	0.128
		14 -0.132	-0.169	14.996	0.132
		15 -0.106	-0.366	15.763	0.150
		16 0.192	-0.107	18.381	0.105
		17 -0.053	-0.110	18.587	0.136
		18 0.065	0.118	18.914	0.168
		19 0.018	0.183	18.941	0.216
		20 -0.038	0.096	19.066	0.265

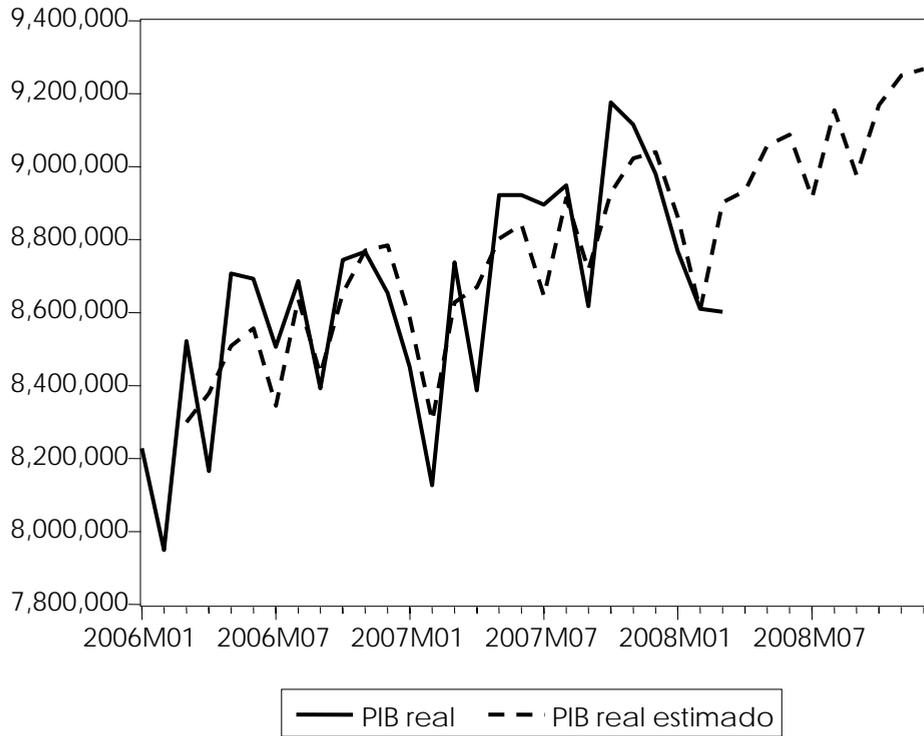
P. PIB real mensual estimado con el modelo SARIMA seleccionado, marzo 2006 – diciembre 2008

Mes	Año		
	2006	2007	2008
Enero	NA	8,584,672	8,861,277
Febrero	NA	8,302,560	8,606,411
Marzo	8,299,317	8,627,964	8,902,443
Abril	8,378,718	8,670,214	8,933,723
Mayo	8,507,991	8,802,385	9,057,037
Junio	8,556,813	8,840,853	9,087,691
Julio	8,344,292	8,645,094	8,914,491
Agosto	8,636,850	8,913,748	9,154,860
Septiembre	8,430,722	8,716,889	8,976,980
Octubre	8,653,138	8,928,372	9,168,622
Noviembre	8,770,447	9,023,244	9,250,340
Diciembre	8,784,348	9,040,187	9,267,743

Q. Comparación de las estimaciones del PIB real con el modelo SARIMA seleccionado, con los datos publicados del primer trimestre de 2008

Mes	PIB real			Crecimiento anual %		
	Publicado	Estimado	Dif %	Publicado	Estimado	Dif
Enero	8,768,441	8,861,277	1.06	3.75	4.85	1.10
Febrero	8,609,975	8,606,411	-0.04	5.95	5.90	-0.04
Marzo	8,602,429	8,902,443	3.49	-1.55	1.88	3.43
1er Trimestre	8,660,282	8,790,043	1.50	2.63	4.16	1.54

R. Comparación de las estimaciones del PIB real con el modelo SARIMA seleccionado, con los datos publicados, enero 2006 – diciembre 2008



S. Resultados de la Prueba de Theil del modelo ARIMAX seleccionado

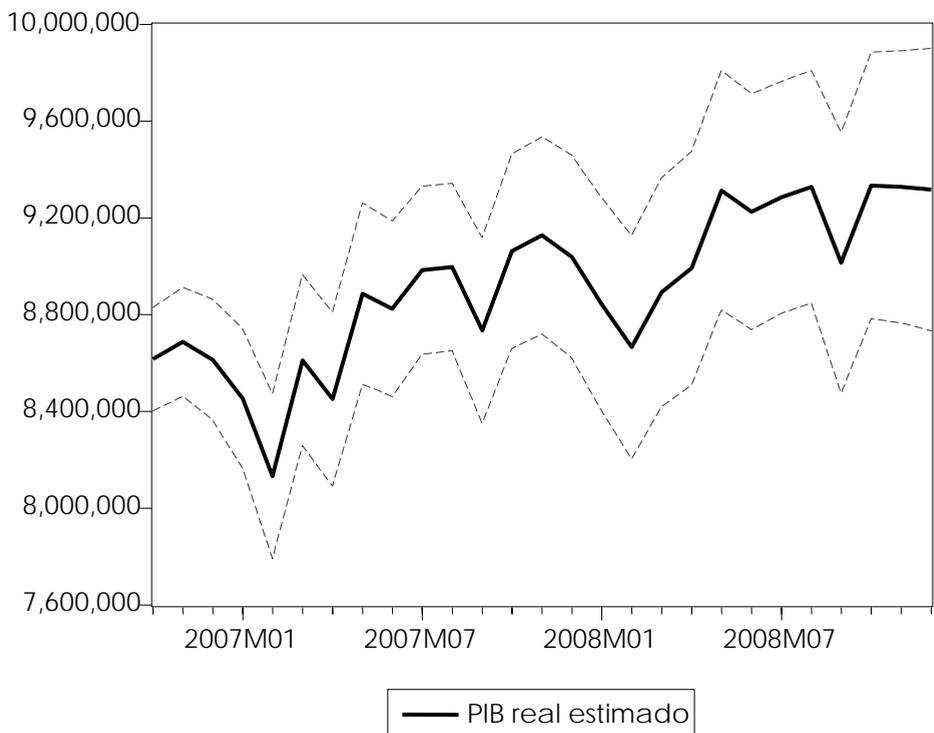
Forecast: PIB real estimado

Actual: PIB real

Forecast sample: 2006M10 2008M12

Included observations: 15

Root Mean Squared Error	79810.04
Mean Absolute Error	67730.57
Mean Absolute Percentage Error	0.769031
Theil Inequality Coefficient	0.004556
Bias Proportion	0.036312
Variance Proportion	0.004031
Covariance Proportion	0.959657



T. Especificaciones y resultados del modelo ARIMAX seleccionado

Dependent Variable: D(DATOS2)
 Method: Least Squares
 Date: 07/12/08 Time: 12:53
 Sample (adjusted): 2005M01 2007M12
 Included observations: 36 after adjustments
 Convergence achieved after 259 iterations
 Backcast: 2003M11 2004M12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
@SEAS(1)	-181359.1	54402.65	-3.333644	0.0032
@SEAS(2)	-261097.0	70821.92	-3.686669	0.0014
@SEAS(3)	417811.9	64677.93	6.459883	0.0000
@SEAS(5)	295237.3	67411.37	4.379636	0.0003
@SEAS(9)	-243019.9	63252.20	-3.842079	0.0009
@SEAS(10)	292005.2	67930.83	4.298567	0.0003
AR(5)	-0.359960	0.174096	-2.067599	0.0512
AR(23)	0.497774	0.124432	4.000379	0.0006
MA(1)	-1.465726	0.066856	-21.92365	0.0000
MA(2)	0.640722	0.058644	10.92553	0.0000
MA(5)	0.529714	0.063770	8.306588	0.0000
MA(6)	-0.976701	0.062109	-15.72550	0.0000
MA(8)	0.196006	0.044811	4.374094	0.0003
MA(13)	-0.302427	0.070435	-4.293703	0.0003
MA(14)	0.433351	0.071360	6.072761	0.0000
R-squared	0.965954	Mean dependent var	22013.95	
Adjusted R-squared	0.943257	S.D. dependent var	293882.1	
S.E. of regression	70004.76	Akaike info criterion	25.44485	
Sum squared resid	1.03E+11	Schwarz criterion	26.10465	
Log likelihood	-443.0073	Durbin-Watson stat	2.458878	
Inverted AR Roots	.95	.92-.28i	.92+.28i	.84-.52i
	.84+.52i	.68+.70i	.68-.70i	.45+.85i
	.45-.85i	.19+.94i	.19-.94i	-.08+.97i
	-.08-.97i	-.33-.93i	-.33+.93i	-.54+.80i
	-.54-.80i	-.74+.61i	-.74-.61i	-.89+.37i
	-.89-.37i	-.98+.12i	-.98-.12i	
Inverted MA Roots	1.00+.07i	1.00-.07i	.77+.63i	.77-.63i
	.70-.71i	.70+.71i	.07-.86i	.07+.86i
	-.29-.95i	-.29+.95i	-.61+.58i	-.61-.58i
	-.90+.18i	-.90-.18i		

U. Raíces de los procesos AR y MA del modelo ARIMAX seleccionado

Inverse Roots of AR/MA Polynomial(s)
 Specification: D(DATOS2) @SEAS(1) @SEAS(2)
 @SEAS(3) @SEAS(5) @SEAS(9)
 @SEAS(10) AR(5) AR(23) MA(1) MA(2)
 MA(5) MA(6) MA(8) MA(13) MA(14)
 Date: 07/13/08 Time: 10:37
 Sample: 2003M01 2008M12
 Included observations: 36

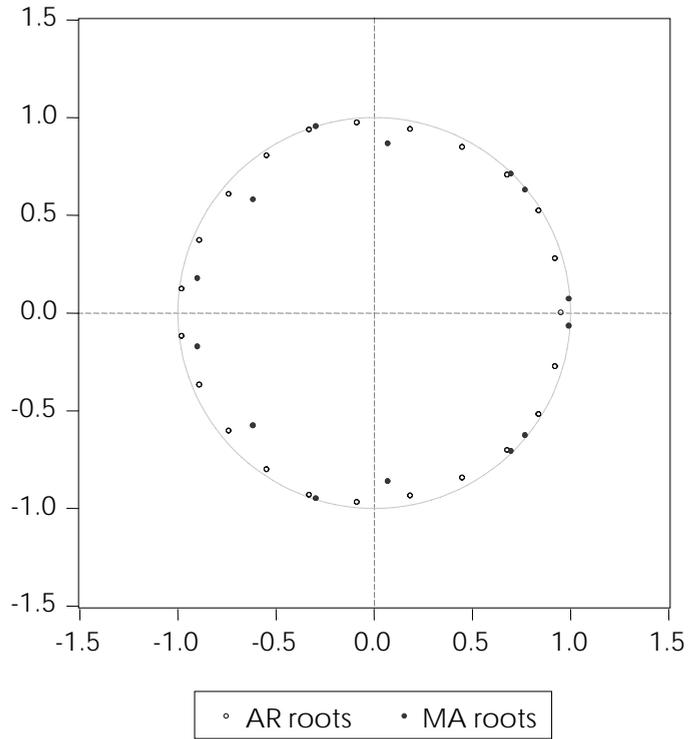
AR Root(s)	Modulus	Cycle
-0.328601 ± 0.933959i	0.990080	3.291167
0.839472 ± 0.520996i	0.988003	11.31196
-0.976827 ± 0.120516i	0.984233	2.081326
0.680550 ± 0.704341i	0.979410	7.828777
-0.084330 ± 0.970565i	0.974222	3.790837
-0.544420 ± 0.801876i	0.969226	2.899161
0.924503 ± 0.275847i	0.964779	21.66884
-0.886921 ± 0.370163i	0.961066	2.287945
0.450421 ± 0.845697i	0.958166	5.810192
0.187233 ± 0.937586i	0.956098	4.573938
-0.738306 ± 0.605530i	0.954862	2.559678
0.954451	0.954451	

No root lies outside the unit circle.
 ARMA model is stationary.

MA Root(s)	Modulus	Cycle
0.995146 ± 0.068413i	0.997495	91.54052
-0.294316 ± 0.952368i	0.996808	3.359055
0.700195 ± 0.707993i	0.995755	7.943989
0.771489 ± 0.627489i	0.994453	9.201718
-0.897469 ± 0.175463i	0.914460	2.130963
0.072398 ± 0.864050i	0.867078	4.224834
-0.614581 ± 0.577322i	0.843214	2.631762

No root lies outside the unit circle.
 ARMA model is invertible.

Inverse Roots of AR/MA Polynomial(s)



V. Correlograma de los residuales del modelo ARIMAX seleccionado

Date: 07/14/08 Time: 10:24

Sample: 2005M01 2007M12

Included observations: 36

Q-statistic probabilities adjusted for 9 ARMA term(s)

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.238	-0.238	2.2053	
		2	-0.086	-0.151	2.5006	
		3	0.158	0.107	3.5290	
		4	0.070	0.137	3.7360	
		5	0.022	0.114	3.7568	
		6	-0.030	0.002	3.7988	
		7	0.072	0.044	4.0410	
		8	-0.109	-0.129	4.6238	
		9	-0.003	-0.074	4.6244	
		10	-0.112	-0.193	5.2842	0.022
		11	0.018	-0.048	5.3019	0.071
		12	-0.027	-0.031	5.3442	0.148
		13	-0.090	-0.033	5.8301	0.212
		14	0.074	0.089	6.1664	0.290
		15	0.028	0.128	6.2165	0.399
		16	-0.199	-0.152	8.9276	0.258

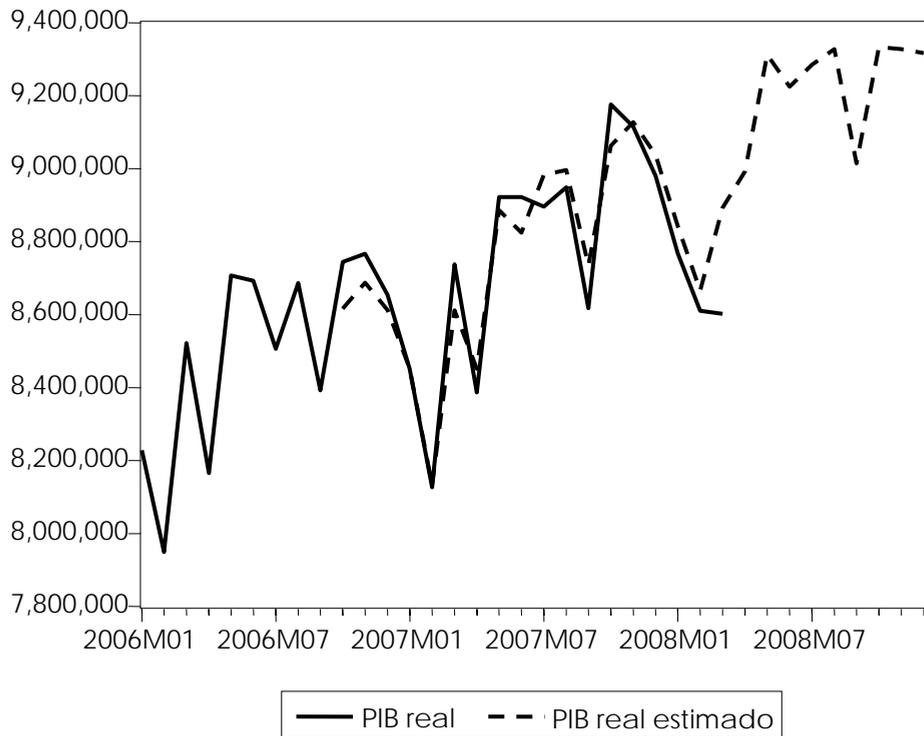
W. PIB real mensual estimado con el modelo ARIMAX seleccionado, octubre 2006 – diciembre 2008

Mes	Año		
	2006	2007	2008
Enero	NA	8,452,596	8,843,045
Febrero	NA	8,132,468	8,665,774
Marzo	NA	8,611,428	8,893,100
Abril	NA	8,451,355	8,992,568
Mayo	NA	8,886,321	9,312,529
Junio	NA	8,824,418	9,224,635
Julio	NA	8,983,111	9,284,546
Agosto	NA	8,996,257	9,327,694
Septiembre	NA	8,734,483	9,014,668
Octubre	8,616,097	9,062,103	9,333,247
Noviembre	8,687,360	9,127,460	9,327,725
Diciembre	8,612,708	9,037,944	9,316,736

X. Comparación de las estimaciones del PIB real con el modelo ARIMAX seleccionado, con los datos publicados del primer trimestre de 2008

Mes	PIB real			Crecimiento anual %		
	Publicado	Estimado	Dif %	Publicado	Estimado	Dif
Enero	8,768,441	8,843,045	0.85	3.75	4.64	0.88
Febrero	8,609,975	8,665,774	0.65	5.95	6.63	0.69
Marzo	8,602,429	8,893,100	3.38	-1.55	1.78	3.33
1er Trimestre	8,660,282	8,800,639	1.62	2.63	4.29	1.66

Y. Comparación de las estimaciones del PIB real con el modelo ARIMAX seleccionado, con los datos publicados, enero 2006 – diciembre 2008



Z. Comparación de las estimaciones del PIB real con el modelo ARIMAX y SARIMA seleccionados, con los datos publicados, enero 2006 - diciembre 2008

