



# EL COLEGIO DE MÉXICO

## CENTRO DE ESTUDIOS ECONÓMICOS

### **MAESTRÍA EN ECONOMÍA**

TRABAJO DE INVESTIGACIÓN PARA OBTENER EL GRADO DE MAESTRO EN  
ECONOMÍA

*ESTIMACIÓN DE UN MODELO DE DEMANDA DE  
GASOLINAS PARA LA CIUDAD DE MÉXICO*

*JUAN CARLOS RODRÍGUEZ DÍAZ*

**PROMOCIÓN 1979 - 1981**

**ASESOR: DR. GERARDO ESQUIVEL HERNÁNDEZ**

2007



---

A Sandra L. Barojas Beltrán y Juan María Rodríguez Barojas

---

## **Agradecimientos**

Quiero hacer patente mi agradecimiento al Dr. Gerardo Esquivel Hernández, profesor del Centro de Estudios Económicos del Colegio de México, por su valiosa orientación para la terminación de esta tesis, así como el tiempo que generosamente me ofreció. No está de más mencionar que sus comentarios, sugerencias y aportes bibliográficos enriquecieron el contenido de este trabajo.

Asimismo, quiero expresar un reconocimiento al Dr. Miguel Gerardo Breceda Lapeyre, Coordinador del Programa de Energía de la Universidad Autónoma de la Ciudad de México, por su interés, ayuda y apoyo para la culminación de este documento.

---

## Resumen

En este documento se modela, en primer término, una función de demanda de gasolinas para la Ciudad de México en el periodo 1993 – 2005, mediante el método de mínimos cuadrados ordinarios (MCO), posteriormente y para el mismo periodo, se construye un modelo ARIMA utilizando la metodología de Box – Jenkins y adicionalmente, para establecer la existencia (o ausencia) de una relación de largo plazo entre las variables implicadas, se aplica el Análisis de Cointegración.

Los resultados del MCO muestran que las elasticidades ingreso de corto y largo plazo, en la Ciudad de México, son mayores que los resultados obtenidos en otros estudios para todo el país. La elasticidad precio de corto plazo coincide con el promedio de las elasticidades para todo el país y la zona fronteriza norte y la elasticidad precio de largo plazo resulta ligeramente menor que la estimada para estas zonas. Por su parte, las elasticidades ingreso y precio de largo plazo obtenidas mediante el análisis de cointegración resultaron superiores a las del MCO.

Por último, se demuestra que la serie “ventas de gasolina en la Ciudad de México” para el periodo enero de 1993 – diciembre de 2005, presenta un quiebre estructural al final de 2003 e inicio de 2004, utilizando la prueba de raíces unitarias y cambio estructural desarrollada por Perron y aplicando también la técnica del análisis de intervención para un modelo ARIMA.

La importancia del presente trabajo radica en: primero, existen pocas estimaciones de una función de demanda de gasolinas para la Ciudad de México, segundo, resulta interesante conocer el valor de las elasticidades ingreso y precio para esta región y como comparan con las elasticidades calculadas, en otros estudios, para todo el país y, tercero, confrontar los resultados obtenidos mediante mínimos cuadrados ordinarios con los conseguidos mediante técnicas de cointegración.

---

# INDICE

	Página
Introducción.....	9
Antecedentes.....	9
Problemática.....	9
Objetivo.....	10
Hipótesis.....	11
Metodología.....	11
Sección 1. Marco de referencia.....	13
Sección 2. Marco teórico.....	17
a. Modelo de demanda de gasolinas.....	17
b. Presentación de los datos.....	19
Sección 3. Revisión de la literatura.....	22
a. Demanda de gasolinas automotrices en México.....	22
b. Demanda de gasolinas automotrices en otros países.....	27
c. Demanda de gasolina estimadas con técnicas de cointegración.....	28

---

<b>Sección 4. Estimación de un modelo estructural de la demanda de gasolina.....</b>	<b>30</b>
a. Estimación del modelo.....	30
b. Cálculo de elasticidades.....	35
c. Comparación con otros resultados.....	36
<b>Sección 5. Cambio estructural.....</b>	<b>38</b>
a. Prueba de cambio estructural utilizando el modelo estimado.....	38
b. Raíces unitarias y cambio estructural.....	44
<b>Sección 6. Modelo de series de tiempo de la demanda de gasolinas y análisis de intervención.....</b>	<b>49</b>
a. Modelo de series de tiempo para la demanda de gasolinas.....	49
b. Análisis de intervención.....	71
<b>Sección 7. Estimación de un modelo con técnicas de cointegración.....</b>	<b>77</b>
a. Orden de integración de las series.....	77
b. Cointegración de las series y modelo de corrección de errores.....	82

---

<b>Sección 8. Conclusiones y Recomendaciones.....</b>	<b>87</b>
<b>Bibliografía.....</b>	<b>91</b>
<b>Anexo</b>	
<b>1. Pruebas econométricas.....</b>	<b>94</b>
<b>2. Modelos de series de tiempo.....</b>	<b>101</b>
<b>3. La conducta del consumidor.....</b>	<b>109</b>
<b>4. Elasticidad precio y elasticidad ingreso.....</b>	<b>113</b>
<b>5. Funciones de demanda de energía.....</b>	<b>114</b>

---

**CUADROS**

Cuadro 1.1. Población total del DF, municipios conurbados y ZMVM,1940 – 2020.....	13
Cuadro 3.1. Elasticidad ingreso y elasticidad precio de la demanda de gasolina en todo el país, zonas metropolitanas, frontera norte y resto del país.....	23
Cuadro 3.2. Elasticidad ingreso y elasticidad precio en zonas fronterizas.....	23
Cuadro 3.3. Elasticidad ingreso y elasticidad precio de la demanda de gasolinas bajo el programa Hoy no circula.....	24
Cuadro 3.4. Elasticidad ingreso y elasticidad precio de la demanda de gasolinas Nova.....	25
Cuadro 3.5. Elasticidades de corto y largo plazo.....	26
Cuadro 3.6. Comparación de estimaciones de la elasticidad precio de la demanda de gasolina.....	26
Cuadro 3.7. Elasticidad ingreso y precio en Canadá.....	27
Cuadro 3.8. Elasticidad ingreso y precio en Estados Unidos.....	28
Cuadro 3.9. Elasticidades de la demanda de gasolinas estimadas con técnicas de cointegración.....	29
Cuadro 4.1. Parámetros y estadísticos de la estimación del Modelo 1.....	31
Cuadro 4.2. Batería de pruebas del Modelo 1.....	31
Cuadro 4.3. Parámetros y estadísticos de la estimación del Modelo 2.....	32
Cuadro 4.4. Batería de pruebas del Modelo 2.....	32
Cuadro 4.5. Parámetros y estadísticos de la estimación del Modelo 3.....	33
Cuadro 4.6. Batería de pruebas del Modelo 3.....	33
Cuadro 4.7. Parámetros y estadísticos de la estimación del Modelo 4.....	34
Cuadro 4.8. Batería de pruebas del Modelo 4.....	34
Cuadro 4.9. Resumen de los parámetros y batería de pruebas de los modelos....	35
Cuadro 4.10. Elasticidades de corto y largo plazo del modelo estimado.....	36
Cuadro 4.11. Comparación de estimaciones de las elasticidades ingreso y precio de la gasolina con revisiones de la literatura.....	36
Cuadro 5.1. Estimación 1 & 2 del logaritmo de las ventas de gasolina en la Ciudad de México.....	40

---

Cuadro 5.2. Histograma y estadísticos de la serie diferencias.....	41
Cuadro 5.3. Estadístico de Q de Ljung y Box de la serie diferencia.....	43
Cuadro 5.4. Parámetros de la regresión.....	48
Cuadro 6.1. Volumen de ventas de gasolina para la Ciudad de México.....	52
Cuadro 6.2. Determinación de $\hat{\sigma}^2$ para la serie ventas de gasolina.....	52
Cuadro 6.3. Identificación de la serie mediante la FAC y la FACP.....	54
Cuadro 6.4. Valor de la FACP muestral para la serie diferenciada de las ventas de gasolina.....	56
Cuadro 6.5. Valores de la FAC calculada para la serie diferenciada de las ventas de gasolina.....	57
Cuadro 6.6 Valores de la FAC muestral para la serie diferenciada de las ventas de gasolina.....	57
Cuadro 6.7 Resultados de la estimación del modelo ARI (1,0,1) de las ventas de gasolina.....	59
Cuadro 6.8 Resultados de la estimación del modelo ARI (2,0,1) (12,0,0) de las ventas de gasolina.....	59
Cuadro 6.9 Resultados de la estimación del modelo ARIMA (1,1,1) de las ventas de gasolina.....	60
Cuadro 6.10 Resultados de la estimación del modelo ARIMA (1,1,1) (12,0,12) de las ventas de gasolina.....	61
Cuadro 6.11 Estadísticas descriptivas e histograma del residuo del modelo ARIMA (1,0,1) de las ventas de gasolina.....	63
Cuadro 6.12 Estadístico Q de Ljung y Box de los residuos del modelo ARI (1,0,1).....	64
Cuadro 6.13 Estadísticas descriptivas e histograma del residuo del modelo ARI (2,0,1) (12,0,0) de las ventas de gasolina.....	65
Cuadro 6.14 Estadístico Q de Ljung y Box de los residuos del modelo ARI (2,0,1) (12,0,0).....	66
Cuadro 6.15 Estadísticas descriptivas e histograma del residuo del modelo ARIMA (1,1,1) de las ventas de gasolina.....	67

---

Cuadro 6.16 Estadístico Q de Ljung y Box de los residuos del modelo ARIMA (1,1,1).....	68
Cuadro 6.17 Estadísticas descriptivas e histograma del residuo del modelo ARIMA (1,1,1) (12,0,12) de las ventas de gasolina.....	69
Cuadro 6.18 Estadístico Q de Ljung y Box de los residuos del modelo ARIMA (1,1,1) (12,0,12).....	70
Cuadro 6.19 Resumen de pruebas de los modelos ARIMA .....	70
Cuadro 6.20. Parámetros del modelo estimado.....	76
Cuadro 7.1 Correlograma del logaritmo de las ventas de gasolina.....	79
Cuadro 7.2 Correlograma del logaritmo del PIB.....	79
Cuadro 7.3 Correlograma del logaritmo del precio de la gasolina.....	80
Cuadro 7.4 Prueba de raíces unitarias para el logaritmo de las variables.....	81
Cuadro 7.5 Prueba de raíces unitarias para la primera diferencia del logaritmo de las variables.....	81
Cuadro 7.6 Resultados de la regresión de lvg contra c, lpib & lpg.....	82
Cuadro 7.7 Prueba de Dickey – Fuller para el error de la regresión.....	82
Cuadro 7.8 Prueba de Johansen para lvg, lpib, lpg, .....	83
Cuadro 7.9 Vector de cointegración para lvg, lpib, lpg.....	84
Cuadro 7.10 Modelo de corrección de errores para lvg, lpib, lpg.....	85

## GRÁFICAS

Gráfica 1.1. Pirámide de población por grupos de edad y sexo en el DF, 2000 y 2030.....	14
Gráfica 1.2. Variación del PIB nacional y del DF, 1994 – 2003.....	15
Gráfica 2.1. Evolución del tipo de gasolina en la Ciudad de México.....	19
Gráfica 2.2. Precio de la gasolina en la Ciudad de México.....	21
Gráfica 2.3. Consumo de gasolina en la Ciudad de México.....	21
Gráfica 4.1. Tasas de crecimiento de la demanda de gasolinas y PIB en la Ciudad de México.....	37
Gráfica 5.1. Prueba de CUSUM.....	38
Gráfica 5.2. Log de ventas de gasolina reales y estimaciones 1 & 2.....	39
Gráfica 5.3. Serie diferencia.....	42
Gráfica 5.4. Autocorrelación y autocorrelación parcial de la serie diferencia.....	43
Gráfica 5.5. Variable dummy de un salto.....	46
Gráfica 5.6. Variable dummy de cambio de nivel.....	47
Gráfica 6.1. Función de autocorrelación muestral de las ventas de gasolina.....	53
Gráfica 6.2. Función de autocorrelación muestral de la primera diferencia de las ventas de gasolina.....	54
Gráfica 6.3. FAC y FACP de la primera diferencia del logaritmo de las ventas de gasolina.....	55
Gráfica 6.4. Comportamiento del residuo del modelo ARI(1,0,1) de las ventas de gasolina.....	64
Gráfica 6.5. Comportamiento del residuo del modelo ARI (2,0,1) (12,0,0) de las ventas de gasolina.....	66
Gráfica 6.6. Comportamiento del residuo del modelo ARIMA (1,1,1) de las ventas de gasolina.....	68
Gráfica 6.7. Comportamiento del residuo del modelo ARIMA (1,1,1) (12,0,12) de las ventas de gasolina.....	69
Gráfica 6.8. Gráfica de la función de impulso $\varepsilon_{it}$ .....	74
Gráfica 7.1. Evolución del logaritmo de las ventas de gasolina, PIB y precio de la gasolina de la Ciudad de México.....	78
Gráfica 7.2. Gráfica del error de la regresión.....	83

## **Introducción**

### **Antecedentes**

La edad de piedra no acabó por falta de piedras, y la era de los combustibles fósiles tampoco terminará por el agotamiento del petróleo, el gas natural y el carbón<sup>1</sup>; sin embargo esto no excluye que se “controle” su consumo, así como el de sus derivados, entre los que se encuentra la gasolina.

Desde un punto de vista tecnológico, los recursos energéticos se pueden clasificar en tres clases. La primera comprende aquellos que son baratos pero limitados, como es el caso del petróleo y el gas; la segunda clase abarca los abundantes pero menos atractivos y son aquellos que se usan cuando los de la primera clase comienzan a escasear, ejemplo de estos son: carbón, petróleo y gas cuya ubicación genera altos costos de extracción y uranio de alto grado utilizado en los reactores nucleares, entre otros y la tercera clase son los recursos superabundantes que en un futuro próximo podrían brindar cantidades ilimitadas de energía: fusión, fisión, energía solar, viento, por citar los más conocidos.

La gasolina es un petrolífero derivado de los procesos de refinación del petróleo, el cual, de acuerdo a la clasificación anterior, es un recurso energético escaso, y dado que no existen en el País palancas directas para controlar la tasa de consumo de este producto, resulta esencial conocer el comportamiento de su demanda para poder establecer políticas que permitan regularla a través de instrumentos tales como tasas de interés, impuestos, control de precios, subsidios y restricciones ambientales, entre otras.

### **Problemática**

La mayoría de los estudios realizados a la fecha obtienen funciones de demanda de gasolinas en el ámbito nacional o de la Zona Metropolitana del Valle de México

---

<sup>1</sup> José Santamaría, Director de Word Watch, España, noviembre 2004

y pocos son los que tratan el caso específico de la demanda en la Ciudad de México.

Por lo anterior, resulta interesante modelar una función de demanda de gasolinas para el caso de la Ciudad de México, la cual puede utilizarse para establecer políticas que permitan regular el consumo de este producto en base a impactos en su demanda por cambio de precio o bien el tamaño de las inversiones necesarias para elevar la producción nacional de este bien, que permitan satisfacer la demanda ante los crecimientos observados en la economía nacional y como alternativa a su importación.

### **Objetivo**

Tomando como referencia la teoría económica, en especial la teoría de la conducta del consumidor, el objetivo de este trabajo es el de estimar una función de demanda de gasolinas. El propósito principal es el de estimar confiablemente el sentido y grado de conexión entre el consumo de gasolina, el ingreso y el precio del energético, primero mediante el método de mínimos cuadrados ordinarios, posteriormente aplicar el análisis de series de tiempo para construir un modelo ARIMA y por último utilizar técnicas de cointegración para estar seguros de que el modelo establecido mediante mínimos cuadrados no representa una “relación espuria”. Se elaborará y evaluará un modelo econométrico que permita estimar el valor y signo de la elasticidad ingreso y de la elasticidad precio, para poder establecer políticas de control de demanda en base a estas variables.

Otro objetivo presente en este trabajo, es el de probar si existe cambio estructural en las ventas de gasolina para la Ciudad de México, en el periodo analizado, recurriendo a la metodología de raíces unitarias desarrollada por Perron<sup>1</sup> y al análisis de intervención en series de tiempo.

---

<sup>1</sup> Perron, P. “The Great Crash, the Oil Price Shock and the Unit Root Hypothesis”, *Econometrica*, 57, 1989

## Hipótesis

A partir de la función de demanda de las gasolinas, la primera hipótesis se refiere al comportamiento de las elasticidades, ingreso y precio, de acuerdo a la teoría de la demanda del consumidor y la segunda al quiebre estructural de la serie.

### Hipótesis 1

- a) La elasticidad de la demanda es positiva y elástica con respecto al ingreso de los consumidores, esto implica un valor en una vecindad del uno.
- b) La elasticidad de la demanda es negativa e inelástica con respecto al precio del combustible, por lo que el valor de ésta es menor a uno.

### Hipótesis 2

Existe un cambio estructural en las ventas de gasolinas para la Ciudad de México.

## Metodología

La estimación de la función de demanda se efectuará mediante el método de mínimos cuadrados, ya que como establece Galindo<sup>2</sup>, “...*bajo ciertas condiciones, el modelo estimado por mínimos cuadrados ordinarios reúne las condiciones para convertirse en una aproximación adecuada del proceso generador de información...*”, arguyendo a favor del uso exhaustivo de pruebas de diagnóstico para conocer las propiedades estadísticas de los modelos econométricos y seleccionar la mejor opción posible.

Las pruebas de diagnóstico que se utilizarán son las siguientes:

---

<sup>2</sup> Galindo P., Luís Miguel, “ La econometría aplicada moderna: Los mínimos cuadrados ordinarios y las pruebas de diagnostico”, Cuaderno de Trabajo de Economía Aplicada, N° 17, Maestría en Ciencias Económicas (UACPyP, CCH, UNAM), 1995

- i. Pruebas de bondad del ajuste utilizando el coeficiente de determinación múltiple ( $R^2$ ), suma de los cuadrados de los errores (SSE) y error estándar de la regresión (ESR)
- ii. Prueba de significancia estadística  $t$  de los parámetros
- iii. Prueba de normalidad de Jarque – Bera para los residuales
- iv. Prueba de Durbin Watson para detectar autocorrelación
- v. Prueba del multiplicador de Lagrange para detectar autocorrelación si se incluyen simultáneamente variables endógenas y exógenas
- vi. Prueba de White para la heteroscedasticidad suponiendo que es función de las variables independientes de la ecuación inicial
- vii. Prueba de Arch para detectar heteroscedasticidad

Para la construcción del modelo ARIMA se utiliza el método de trabajo de Box – Jenkins, el cual es constructivo, esto es, no se trata de decir si se puede construir un modelo ARIMA para una serie de tiempo determinada  $X_t$  sino de mostrar como se establece esta representación lineal (Ludlow<sup>3</sup>), la idea es efectuar las etapas: identificación, estimación, verificación y uso del modelo que muestren explícitamente que la serie original  $X_t$  se puede modelar por un ARIMA (p,d,q).

Para establecer un modelo estructural mediante análisis de cointegración de la demanda de gasolinas se aplicó la metodología sugerida por Holden y Perman<sup>4</sup> en donde primero se establece el orden de integración de las series a través de su correlograma y el estadístico de Ljung – Box y se comprueba aplicando la prueba de Dickey – Fuller Aumentada. Posteriormente, para establecer si las series están o no cointegradas, se siguen dos procedimientos el de Engle – Granger y el de Johansen y por último se estima el modelo de corrección de errores.

---

<sup>3</sup> Ludlow W, Jorge, “Modelos , Pronósticos y Volatilidad de las Series de Tiempo Generadas en la Bolsa Mexicana de Valores”, Biblioteca de Ciencias Sociales y Humanidades, Serie Economía, UAM Azcapotzálco, México, 1997.

<sup>4</sup> Holden, D. y Perman, R., “Unit Roots and Cointegration for the Economist”, Cointegration for the Applied Economist editado por B. Bhaskara Rao, Martin Press NY, 1995

## Sección 1. Marco de referencia<sup>5</sup>

La Constitución Política de los Estados Unidos Mexicanos señala que la Ciudad de México es el Distrito Federal (D.F.), la cual geográficamente se localiza en el suroeste de la Cuenca de México. Está conformada por 16 delegaciones, y en el año 2000 tenía una población de 8.6 millones de habitantes. Su jurisdicción territorial tiene una extensión de 60,203 hectáreas de tejido urbano y una “zona rural” aún mayor que abarca 88,442 hectáreas. El Distrito Federal o Ciudad de México no constituye una forma urbana delimitada sino que es sólo una parte de la ciudad real, es decir, de la Zona Metropolitana del Valle de México (ZMVM).

Cuadro 1.1. Población total del DF, municipios conurbados y ZMVM, 1940 – 2020 (millones de habitantes)

	1940	1960	1980	2000	2010	2020
<b>D.F.</b>	1.6	5.2	13.2	8.6	8.8	9.0
<b>Municipios conurbados</b>	-	0.5	0.2	9.8	11.7	13.3
<b>ZMVM</b>	-	5.7	13.0	18.4	20.5	22.3

Fuente: La Ciudad de México Hoy, Bases para un diagnóstico. Fideicomiso de estudios estratégicos sobre la Ciudad de México, México, 2000

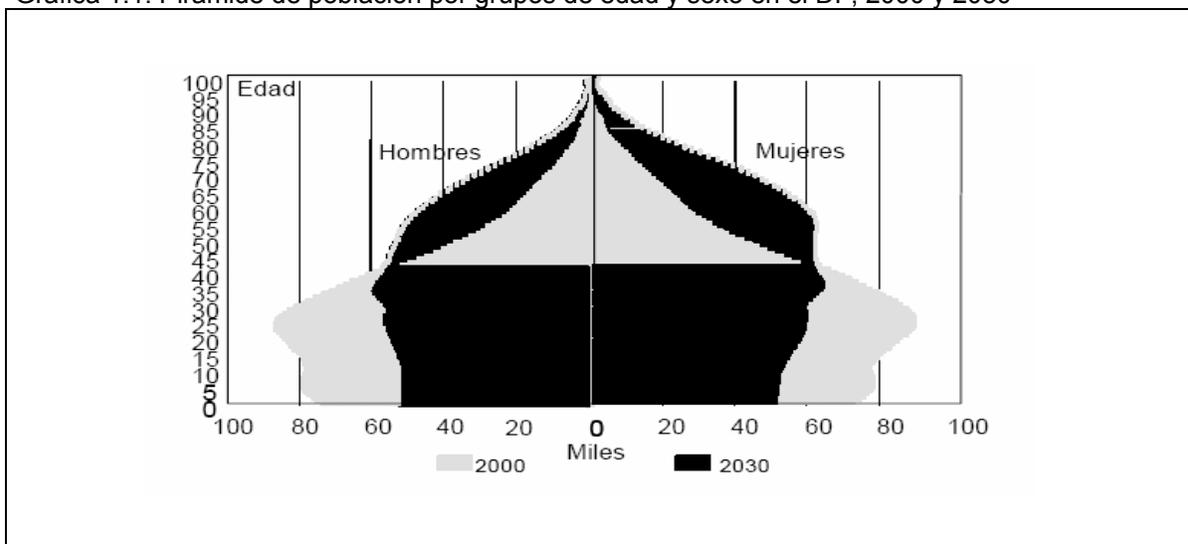
Por lo que respecta a la densidad demográfica, éste es sólo un indicador auxiliar para el análisis de la concentración poblacional, ya que en el D.F. las densidades de población cambian durante el día, dada la movilidad de los habitantes de los municipios conurbados que acceden al D.F. con fines de trabajo, estudio, atención a la salud o diversos servicios, cambiando notablemente la concentración de personas. De hecho, la población de la Ciudad de México es un mosaico de

<sup>5</sup> Casi la totalidad de la información de esta Sección se obtuvo del documento: *Hacia la Agenda XXI de la Ciudad de México, Gobierno del Distrito Federal, noviembre, 2004*

grupos sociales de muy diversas características. En el mismo territorio y contexto biogeofísico coexisten descendientes de los pobladores originarios de la cuenca y personas recién llegadas a la capital del país en busca de mejores perspectivas de vida, que provienen de otros estados de la República.

La estructura demográfica de la ciudad ha sufrido modificaciones en las últimas décadas por la disminución de la inmigración, la reducción paulatina de las tasas de mortalidad y como consecuencia del incremento de la esperanza de vida al nacer. En 1998 la población menor de 15 años representaba el 30.4 por ciento de la población y la de más de 65 años sólo el 4.2 por ciento; para el año 2000 el primer grupo representaba el 28.5 por ciento y el segundo el 4.6 por ciento. Se calcula que para el 2010, estos segmentos de la población representarán el 23.4 y 6.4 por ciento, respectivamente; la edad promedio de la población pasará de 27 años a 32 años.

Gráfica 1.1. Pirámide de población por grupos de edad y sexo en el DF, 2000 y 2030

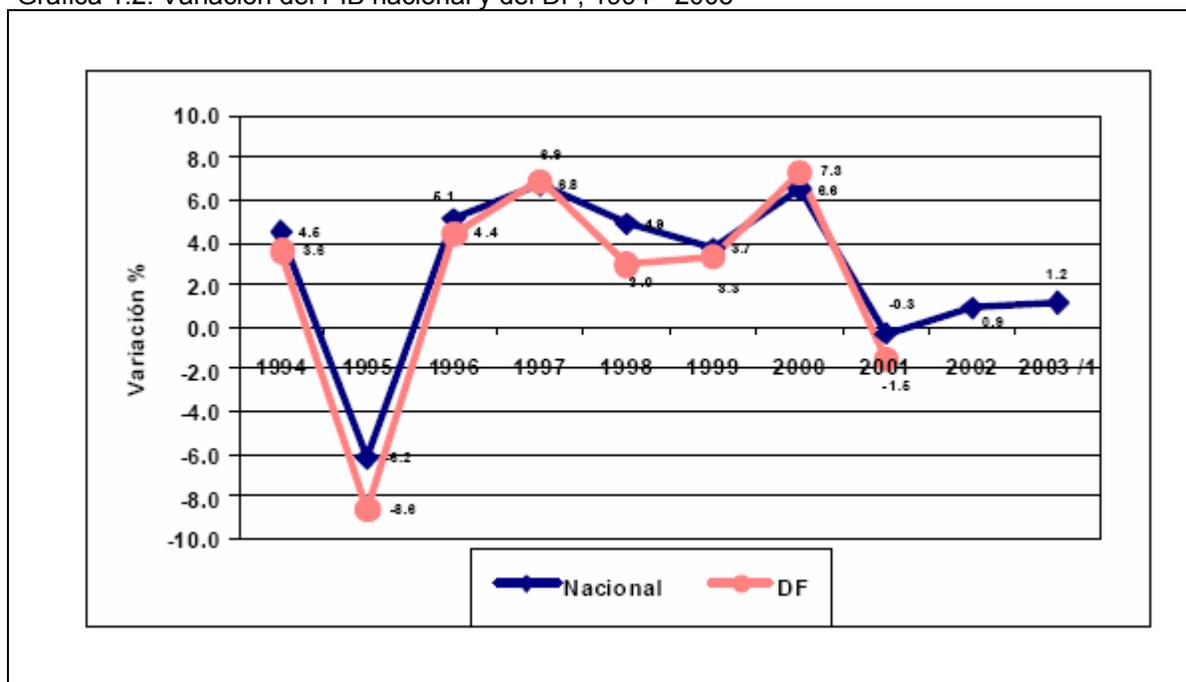


Fuente: Hacia la Agenda XXI de la Ciudad de México, Gobierno del Distrito Federal, noviembre, 2004, pag. 18

En la Ciudad de México se efectúa una gran parte de la actividad económica del país, su contribución al Producto Interno Bruto (PIB) supera el 20 por ciento y representa más de la mitad del PIB de la región central del país . Tres quintas

partes de la actividad bancaria tiene lugar en esta localidad y en ella se realiza el 75 por ciento del ahorro financiero nacional. Aquí también se concentra el mayor número de establecimientos económicos, prácticamente una quinta parte de los activos fijos y cerca de la quinta parte de la población ocupada del país.

Gráfica 1.2. Variación del PIB nacional y del DF, 1994 - 2003



Fuente: Sistema de Cuentas Nacionales de México. PIB por Entidad Federativa, 2004 - 2003, INEGI

Por lo que se refiere al consumo de energía, en la Ciudad de México se gasta alrededor del 10 por ciento de la energía utilizada en todo el territorio nacional. El mayor consumidor de energía es el sector transporte (50 por ciento del total). Este consumo se constituye principalmente de hidrocarburos (98.8 por ciento), siendo las gasolinas el energético de mayor utilización (73.8 por ciento).

Dentro de la oferta de transporte en la Ciudad de México, como lo mencionan Hadda y Bazán<sup>6</sup> es interesante mencionar que casi el 60 por ciento de los viajes lo ofrece el servicio público concesionado (combis, microbuses y autobuses) y los

<sup>6</sup> Hadda C., Gabriel y Bazán N., Gerardo, *Energía y Contaminación del Aire en la Zona Metropolitana del Valle de México*, Contextos Culturales Editores, 2000.

automóviles privados contribuyen con el 19 por ciento de los viajes, pero significan el 97 por ciento del parque vehicular estimado en 3.1 millones para el año 2000.

Por lo que respecta al parque vehicular, los mismos autores establecen que éste ha crecido de manera constante con tasas del orden de 10 por ciento anual, y el índice de motorización (cantidad de vehículos por cada mil habitantes) ha ido creciendo a través de los años: en 1940 era de 27, en 1960 llegó a 51, en 1980 a 127 y en 1990 se incrementó a 168, esto es, un automóvil por cada 6 habitantes.

Según proyecciones del Instituto Nacional de Estadística, Geografía e Informática (INEGI) y el Departamento de Energía de Estados Unidos, en la Ciudad de México, para 2012, habrá 218 vehículos por cada mil habitantes, un automóvil por cada 4 habitantes, lo que generará una demanda creciente de gasolinas.

## Sección 2. Marco teórico

### a. Especificación del modelo de demanda de gasolina.

En el corto plazo la demanda de energía del consumidor está básicamente determinada por su acervo de bienes de capital que usan energía, aunque algunos autores, como lo mencionan Houthakker, Verleger y Sheehan<sup>2</sup>, suponen que este es el único factor que tiene influencia en la demanda. De acuerdo con Houthakker y Taylor<sup>3</sup>, este supuesto es inconsistente con la teoría de la demanda aceptada y en su lugar se debe emplear un supuesto menos restrictivo basado en un modelo de ajuste en el flujo de la demanda.

En tales modelos el acervo de bienes de capital usuarios de energía se supone fijo en el corto plazo, y su utilización se establece como una función de la influencias de una economía normal. En ese caso, se supone que hay una demanda deseada de energía por los consumidores  $q_t^*$ , al tiempo  $t$ . Esta demanda es función del ingreso y del precio

$$q_t^* = f(y_t, p_t) \quad (2.1)$$

Por simplicidad se supone que  $f(\quad)$  es log-lineal

$$q_t^* = \alpha y_t^\beta p_t^\gamma \quad (2.2)$$

Además, se supone un ajuste muy sencillo en el modelo

<sup>2</sup> Houthakker, H.S, Verleger, P.K., Sheehen, D.P., "Dynamic Demand Analices for Gasolina and Residential Electricity", American Journal of Agricultural Economics 56: 412 – 418. 1974.

<sup>3</sup> Houthakker, H.S. y Taylor L.D., "Consumer Demand in the United States, Analyses and Projections", Harvard University Press, Cambridge, 1970.

$$q_t / q_{t-1} = (q_t^* / q_{t-1})^\theta \quad (2.3)$$

en donde  $0 < \theta < 1$ . En tal caso la ecuación a estimarse es:

$$\ln q_t = \theta \ln \alpha + \theta\beta \ln y_t + \theta\gamma \ln p_t + (1 - \theta) \ln q_{t-1} \quad (2.4)$$

Debe notarse que este modelo supone la una misma tasa de utilización de energía para todos los tipos de bienes de capital.

Por lo anterior, el modelo a estimar para la demanda de gasolina establece que ésta depende del ingreso del consumidor y de su precio:

$$\log G_t = \alpha_0 + \alpha_1 \log Y_t + \alpha_2 \log p_t \quad (2.5)$$

Como la demanda de gasolina no es directamente observable se supone que el consumo de gasolina es proporcional a su demanda y que está en función de las mismas variables.

Cabe señalar, que en el análisis de la demanda de gasolina debe distinguirse entre el corto y el largo plazo, ya que el tiempo que los individuos tardan en modificar su conducta se justifica porque ello implica un coste, como lo puntualiza Espino Bravo<sup>4</sup>. Por lo que cabría esperar que las elasticidades de largo plazo fuesen más altas, en términos absolutos, que las de corto plazo.

Por lo que se refiere al signo y valor de los parámetros, se espera que la gasolina sea un bien ordinario ( $\alpha_1 > 0$ ) y necesario ( $|\alpha_2| < 1$ ).

---

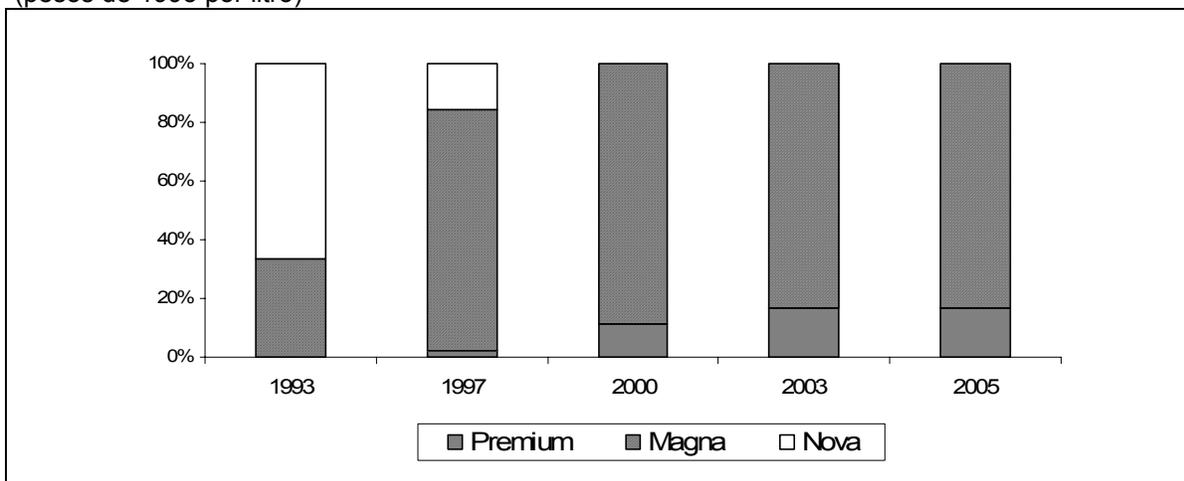
<sup>4</sup> Espino Bravo, J.M., "Estimación de la elasticidad de la demanda de gasolina en México, 1993 – 2003", Documents de Recerca del Programa de Doctorado en Economía Aplicada, Universitat Autònoma de Barcelona, Octubre 2005.

## b. Presentación de los datos

El modelo se estima para un conjunto de datos que van de enero de 1993 a diciembre de 2005.

En este periodo coexistieron tres tipos de gasolina: *Nova* (con plomo), que dejó de proveerse en 1997, *Pemex Magna* (sin plomo) y *Pemex Premium* (de alto octanaje la cual comenzó a surtirse en 1997). A principios de los noventa Pemex elaboraba dos tipos de gasolinas automotrices: *Nova* y *Magna Sin*. La primera de octanaje menor y con plomo; la segunda introducida en 1990, sin plomo y con octanaje equivalente a la gasolina regular sin plomo (*Unleaded Regular*) del mercado norteamericano. Las gasolinas se vendían con dos especificaciones: normal y oxigenada, esta última destinada a zonas ambientales críticas, como las ciudades de México, Guadalajara y Monterrey.

Gráfica 2.1. Evolución del tipo de gasolina en la Ciudad de México (pesos de 1993 por litro)



Fuente: BDR, Base de datos de Pemex Refinación

El precio que se utilizó es el de la gasolina *Pemex Magna*, debido a que fue el único tipo de gasolina que prevaleció durante todo el periodo además que representa más del 80 por ciento del consumo de gasolina y se evitan los problemas de endogeneidad que traería consigo ponderar por consumo. Es

---

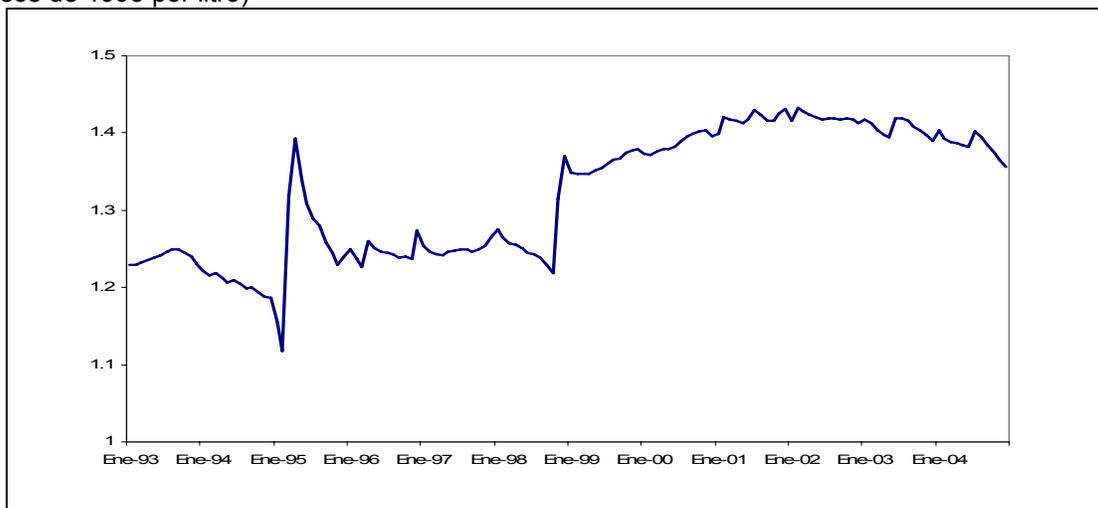
pertinente señalar que las diferencias entre los tres precios son mínimas y, generalmente, su tasa de variación es la misma.

El precio de la gasolina *Pemex Magna* se obtuvo del Banco de Información Económica que el Instituto Nacional de Estadística, Geografía e Informática (INEGI) tiene en su página electrónica ([www.inegi.gob.mx](http://www.inegi.gob.mx)) y fue deflactado por el Índice de Precios al Consumidor para el Área Metropolitana de la Ciudad de México (base 1993 = 100) publicado en la página electrónica del Banco de México ([www.banxico.mx](http://www.banxico.mx)).

Petróleos Mexicanos (Pemex) es el único proveedor de gasolina en todo el país. El precio lo determina el gobierno federal a principio de cada año en base al precio que las gasolina de características similares tienen en la Costa Norteamericana del Golfo más el Impuesto Especial sobre Producción y Servicios (IEPS) y el Impuesto al Valor Agregado (IVA). Se aplica también una tasa de incremento mensual del precio de la gasolina para todo el año, a fin de que el incremento acumulado resulte igual a la meta de inflación determinada por Banco de México. Por lo anterior, los cambios en el precio real de la gasolina se deben a choques que desvían la inflación observada de la inflación objetivo.

La fluctuaciones en el precio internacional de la gasolina son amortiguadas por el gobierno federal a través del IEPS. Como ejemplo basta decir que durante los últimos 10 años el IEPS tuvo una tasa promedio de 40 por ciento, mientras que en 2005 ante los altos precios internacionales de la gasolina este impuesto fue, en promedio, tan sólo del 5 por ciento. En los dos últimos meses de 2005, el precio de referencia fue superior al nacional, por lo que el IEPS fue negativo. En suma la determinación del precio de las gasolina no toma en cuenta criterios de eficiencia. En consecuencia, el precio real de la gasolina mostró una marcada estabilidad durante todo el periodo de análisis.

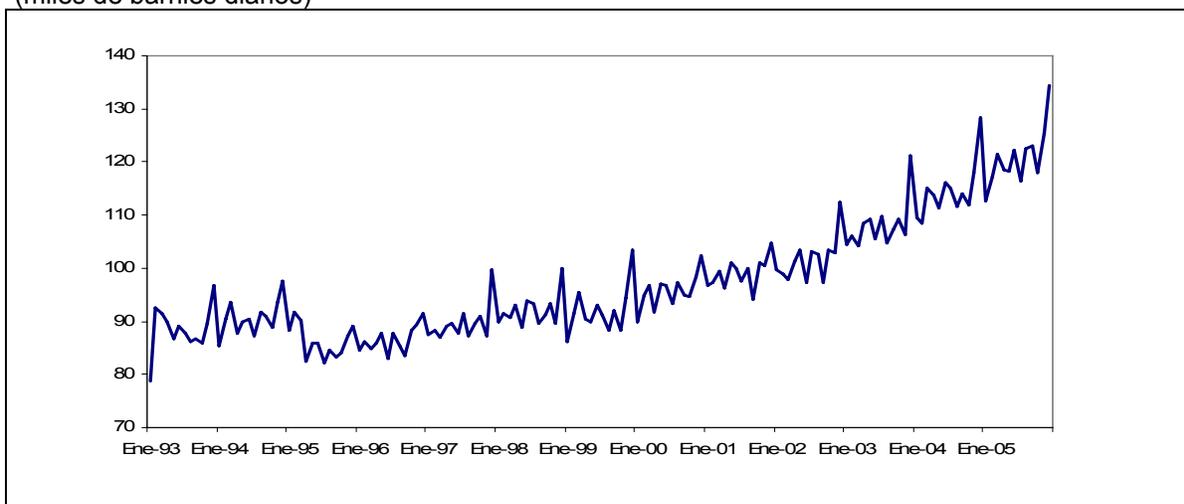
Gráfica 2.2. Precio de la gasolina en la Ciudad de México  
(pesos de 1993 por litro)



Fuente: Banco de Información Económica, INEGI

El consumo de gasolina se aproximó con las ventas internas de Pemex en base a los reporte de las Terminales de Almacenamiento y Distribución de Añil, Barranca del Muerto, San Juanico y Azcapotzálco, que son las que surten a la gasolineras ubicadas en la Ciudad de México. Estos reportes se obtuvieron de la Base de Datos de Refinación (BDR) que es una base electrónica de información perteneciente a Pemex Refinación.

Gráfica 2.3. Consumo de gasolina en la Ciudad de México  
(miles de barriles diarios)



Fuente: BDR, Base de datos de Pemex Refinación

---

### Sección 3. Revisión de la literatura

En esta sección se hace un recuento sobre los resultados de varios estudios sobre la demanda de gasolinas en México y en Estados Unidos y Canadá, destacando el tipo de modelo usado para la estimación y el periodo considerado.

#### a. Demanda de gasolinas automotrices en México

El primer estudio que se analiza es el realizado por Luis Miguel Galindo y Enrique Salinas<sup>5</sup>, en donde los autores utilizaron para la estimación un modelo de vectores autorregresivos (VAR), el cual supone una determinación simultánea del conjunto de las variables endógenas que se hacen función de los valores rezagados. La estimación del VAR se realizó con datos mensuales de julio de 1977 a junio de 1995 para el consumo de gasolinas de la ZMVM. Con base en los resultados obtenidos, los autores mencionan que, *“...las estimaciones realizadas para la demanda de gasolinas en México sugieren una elasticidad ingreso de la demanda de gasolinas que oscila entre 0.8 y 1.2 y una elasticidad precio de entre -0.2 y -0.8...”*, señalando además que, *“...Estos resultados indican, más allá de los resultados puntuales, que la demanda de gasolinas en México es sensible a movimientos en el ingreso y en el precio.”*

Por su parte, Petróleos Mexicanos<sup>6</sup>, a través de Dirección Corporativa de Finanzas, estimó la demanda de gasolinas automotrices a escala nacional y para la frontera norte, utilizando dos modelos combinados. El primero de ellos estima la demanda global de gasolinas automotrices a partir de la ecuación base del laboratorio de Cavendish y en una segunda etapa formulan un modelo de selección binaria tipo logit, el cual determina la proporción del consumo de gasolinas Premium o Magna en función de los diferenciales de precios, de sus

---

<sup>5</sup> Galindo, L.M. y Salinas, E., “La demanda de gasolinas y los instrumentos económicos en México”, Gaceta Ecológica, n° 41 (INE-SEMARNAP), México, 1996

<sup>6</sup> Petróleos Mexicanos, “Modelos de demanda interna de productos petrolíferos y gas natural, México, 2000

rezagos y de una variable tendencial simple. El periodo que consideran es de enero de 1995 a diciembre de 2000.

Cuadro 3.1. Elasticidad ingreso y elasticidad precio de la demanda de gasolina en todo el país, zonas metropolitanas, frontera norte y resto del país.

	elasticidad			
	ingreso		precio	
	corto plazo	largo plazo	corto plazo	largo plazo
Todo el país		0.647		-0.154
Zonas Metropolitanas	0.251		-0.112	
Frontera Norte	0.083		-0.083	
Resto del país	0.054		-0.37	

Fuente: Petróleos Mexicanos

Valores diferentes, sobre todo para la elasticidad ingreso al establecido por Galindo, pueden encontrarse en el trabajo realizado por Haro López e Ibarrola Pérez<sup>7</sup>, cuyo objetivo es analizar y determinar la sensibilidad de la demanda de gasolina comercializada en la zona fronteriza y estatal del norte de México. Los autores estimaron un modelo precio relativo de la demanda con dos muestras que abarcan distintos periodos, enero de 1995 a diciembre de 1998 y enero de 1995 a julio de 1999. Es importante destacar que en estas regiones existe un bien sustituto de la gasolina nacional.

Cuadro 3.2. Elasticidad ingreso y elasticidad precio en zonas fronterizas

	elasticidad	
	ingreso	precio
zonas fronterizas	0.580	-0.415
estados fronterizos	0.402	-0.312

Fuente: Haro e Ibarrola

<sup>7</sup> Haro López, R.A. y Ibarrola Pérez, J.L., "Cálculo de la elasticidad precio de la demanda de gasolina en la zona fronteriza norte de México", Gaceta Económica, año 6, num. 11., otoño 2000

En un documento publicado en 1997 Eskeland y Feyzioglu<sup>8</sup> estimaron una función de demanda de gasolina para la Ciudad de México en el periodo 1983-1992, para medir los efectos del programa *Hoy no circula* en el consumo de este petrolífero. Para ello supusieron que el consumo agregado de gasolina en la Ciudad de México depende del precio de la gasolina y del ingreso familiar.

Cuadro 3.3. Elasticidad ingreso y elasticidad precio de la demanda de gasolinas bajo el programa *Hoy no circula*

elasticidad	Sin regulación	Bajo regulación
precio	-0.17	-0.05
ingreso	0.06	0.24

Nota: regulación = aplicación del programa Hoy no circula

Fuente: Eskeland y Feyzioglu

Cabe hacer notar que los autores utilizaron las llamadas telefónicas internacionales salidas desde la Ciudad de México como una aproximación del ingreso, por lo que el valor de la elasticidad ingreso difiere de los resultados presentados en los otros documentos. Además de que la muestra podría ser obsoleta. Ello debido a que los factores institucionales que afectan la demanda de gasolina en México se han modificado en los últimos años, como resultado de la introducción de programas de inspección de automóviles en las principales ciudades del país; la desaparición de la gasolina con plomo en 1997; la instrumentación del programa de racionamiento del automóvil en la Ciudad de México en 1989 y la caída en el precio real de la gasolina durante 1988, como consecuencia de su utilización como ancla nominal durante el programa de estabilización heterodoxo. Más aún, el periodo muestral corresponde a un periodo de crisis y depresión económica en México, en el cual, el crecimiento económico estuvo por debajo de su potencial de largo plazo. En ese sentido, la reacción de los individuos ante los cambios de precio podría ser distinta cuando la economía

<sup>8</sup> Eskeland, G.S. y Feyzioglu, T., "Rationing Can Backfire: The "Day without a Car" in México City", *The World Bank Economic Review*, Volume 11, number 3, September 1997.

está en crecimiento. En suma, las elasticidades precio y renta de la gasolina podrían haber cambiado.

En otro estudio, publicado en 1980 Pindyck<sup>9</sup> utilizó datos combinados de Brasil y México para estimar un modelo dinámico de forma log-lineal, a partir del cual obtuvo elasticidades precio de algunos combustibles. Para el caso de las gasolinas, obtiene una estimación para la elasticidad precio de corto plazo que va de  $-0.051$  a  $-0.137$  y la de largo plazo resulta entre  $-1.13$  y  $-1.94$ .

Utilizando un modelo que considera velocidades de ajuste diferentes ante cambios en precios e ingresos, modelo de elasticidad constante con ajuste parcial, De Alba y Samaniego<sup>10</sup> obtuvieron las siguientes elasticidades, para el periodo 1977 - 1984.

Cuadro 3.4. Elasticidad ingreso y elasticidad precio de la demanda de gasolinas Nova

	elasticidad precio		elasticidad ingreso	
	corto plazo	largo plazo	corto plazo	largo plazo
Nova	-0.11	-0.24	0.58	1.24

Fuente: Alba y Samaniego

En una versión reducida de un trabajo de investigación presentado por Juan Manuel Espino Bravo<sup>11</sup>, el autor estima las elasticidades precio y renta de la demanda de gasolina en México para 1993 – 2003, aplicando el Método Generalizado en Primeras Diferencias y especificando la demanda de gasolina por un modelo de ajuste parcial.

<sup>9</sup> Pindyck, R.S., "The Structure of World Energy Demand", MIT Press, 1980.

<sup>10</sup> De Alba, E. y Samaniego, R., "Estimación de la demanda de gasolinas y diesel y el impacto de sus precios sobre los ingresos del Sector Público", Documento de Trabajo No. 1985-VIII, Centro de Estudios Económicos, El Colegio de México, 1985.

<sup>11</sup> Espino Bravo, J.M., "Estimación de la elasticidad de la demanda de gasolina en México, 1993-2003", Documents de Recerca del Programa de Doctorado en Economía Aplicada, Universitat Autònoma de Barcelona, Octubre 2005.

En los resultados presentados la elasticidad ingreso resulta menor que las de otros estudios y esto puede atribuirse a que una parte del efecto ingreso está expresado en la elasticidad de la variable “automóviles por adulto”.

Cuadro 3.5. Elasticidades de corto y largo plazo

variable	corto plazo	largo plazo
precio	-0.693	-1.172
ingreso	0.287	0.485
automóviles por adulto	-0.404	-0.683

Fuente: Espino Bravo, J.M.

Se muestra también que la elasticidad precio es mayor que la mediana registrada en las principales recopilaciones sobre el tema (Cuadro 3.6). El autor asegura que lo anterior es consistente con la hipótesis de que la elasticidad precio en los países en desarrollo es más alta que la de los países desarrollados.

Cuadro 3.6. Comparación de estimaciones de la elasticidad precio de la demanda de gasolina

Autor	Elasticidad precio	
	corto plazo	largo plazo
Drollas (1984)	0.27	0.80
Dahl y Sterner (1991)	0.17	0.80
Espey (1998)	0.23	0.43
Graham y Glaister (2002)	0.30	0.70
Espino Bravo (2005)	0.69	1.17

Fuente: Espino Bravo, J.M.

## b. Demanda de gasolinas automotrices en otros países

En mayo de 2000, Christopher J. Nicol<sup>12</sup> publicó un escrito cuyo objetivo era medir el grado de respuesta de la demanda de gasolina al cambio en su precio derivado de la aplicación de impuestos sobre este petrolífero. Para ello, estimaron un sistema completo de ecuaciones de demanda incorporando varias características que se suponía eran importantes para determinar el comportamiento del consumidor, usando datos del gasto familiar en Canadá para los años 1969, 1974, 1978, 1982, 1984, 1986, 1990 y 1992 y del gasto del consumidor en Estados Unidos para el periodo 1980 - 1992.

En este documento establece que la demanda de gasolina es inelástica a su precio e inelástica al ingreso, excepto para un tipo de familia de Canadá. Además menciona que la demanda de gasolina, generalmente, responde en mayor medida a los cambios de precios e ingreso en Canadá, pero esto no es totalmente cierto para todos los tipos de familias. También indica que, mientras que se observan diferencias regionales en las elasticidades en Canadá y Estados Unidos, el tamaño de la familia y el estatus en la tenencia de la casa tiene mayores impactos sobre la diferencia en elasticidades entre las familias.

Cuadro 3.7. Elasticidad ingreso y precio en Canadá

	casa hipotecada sin hijos	casa hipotecada un hijo	casa hipotecada dos hijos	casa rentada sin hijos	casa rentada un hijo	casa rentada dos hijos
ingreso	0.523	1.296	0.443	0.673	0.922	0.680
precio	-0.466	-0.185	-0.580	-0.670	-0.103	-0.853

Fuente: Nicol, J.C.

<sup>12</sup> Nicol, J.C., "Elasticities of demand for gasoline in Canada and the United States", Discussion paper # 84, Department of Economics, University of Regina, mayo 2000, Canada.

Cuadro 3.8. Elasticidad ingreso y precio en Estados Unidos

	casa hipotecada sin hijos	casa hipotecada un hijo	casa hipotecada dos hijos	casa rentada sin hijos	casa rentada un hijo	casa rentada dos hijos
ingreso	0.285	0.621	0.559	0.750	0.837	0.941
precio	-0.162	-0.339	-0.028	-0.026	-0.598	-0.125

Fuente: Nicol, J.C.

### c. Demanda de gasolina estimadas con técnicas de cointegración.

En la Memoria de Investigación de Espino Bravo<sup>13</sup>, se menciona que como las series de los modelos de demanda de gasolinas son no estacionarias, se aplican técnicas de cointegración para estimar las elasticidades de estos modelos, sobre todo a partir del estudio de Bentzen<sup>14</sup> en donde explica el consumo de gasolina per capita por su precio, el acervo de vehículos per capita y la creciente eficiencia en el uso del combustible representada por una tendencia temporal.

Después de Bentzen, algunos de los modelos estimados con técnicas de cointegración incorporan como variables explicativas sólo el precio y la renta en términos reales (i.e. Eltony y Al-Mutairi<sup>15</sup>, Cheung y Thomson<sup>16</sup>). En el siguiente cuadro, el cual se obtuvo del trabajo de Espino Bravo, se presentan algunas elasticidades precio y renta de la demanda de gasolina estimadas mediante técnicas de cointegración, para distintos países y periodos.

<sup>13</sup> Espino Bravo, J.M., "Eficacia del impuesto a la gasolina para reducir las emisiones de los automóviles. Estimación de la demanda de gasolinas en México, 1993 – 2003", Memoria de Investigación, Programa de Doctorado en Economía Aplicada, Universidad Autónoma de Barcelona, septiembre de 2005.

<sup>14</sup> Bentzen, J., "An empirical analysis of gasoline demand in Denmark using cointegration techniques". *Energy Economics* 16 (2), 39-143, 1994.

<sup>15</sup> Eltony, M.; Al-Mutairi, N "Demand for gasoline in Kuwait: An empirical analysis using cointegration techniques". *Energy Economics* 17, p. 249-253, 1995

<sup>16</sup> Cheung, Kui-Yin; Thomson, Elspeth, "The demand for gasoline in China: A cointegration analysis". *Journal of Applied Statistics* 31 (5), 533-544, 2004

Cuadro 3.9. Elasticidades de la demanda de gasolinas estimadas con técnicas de cointegración

Autores	País	Periodo	Elasticidad-precio		Elasticidad-ingreso	
			Corto plazo	Largo plazo	Corto plazo	Largo plazo
Bentzen	Dinamarca	1948-91	-0.320	-0.414	0.890	1.044
Eltony y Al-Mutairi	Kuwait	1970-89	-0.372	-0.463	0.472	0.919
Ramanathan	India	1972-94	-0.209	-0.319	1.178	2.682
Dahl y Kurtubi	Indonesia	1970-95	-0.036	-0.631	0.190	1.289
Alves y De Losso	Brasil	1974-99	-0.092	-0.465	0.122	0.122
Cheung y Thomson	China	1980-99	-0.194	-0.560	1.636	0.970

Fuente: Espino Bravo, J.M.

Lo interesante que puede observarse en este cuadro es que las elasticidades precio de largo plazo estimadas con técnicas de cointegración resultaron menores que los promedios registrados con otro tipo de modelos (con datos de sección cruzada agregados, con datos de panel, con micro datos, entre otros). Ello podría ser consecuencia del reconocimiento explícito de la naturaleza no estacionaria de las series

---

## Sección 4. Estimación de un modelo estructural de la demanda de gasolina

### a. Estimación del modelo

La estimación del modelo de demanda de gasolinas se realizará a partir de la ecuación 2.5 mencionada en la Sección 2, la cual establece que la demanda de gasolina depende del ingreso del consumidor y de su precio. Utilizando el método de mínimos cuadrados ordinarios y el uso exhaustivo de pruebas de diagnóstico, así como retrasos tanto en la variable endógena como en las exógenas, se determinarán los parámetros de la ecuación que representa una adecuada aproximación del proceso generador de datos (DGP, por sus siglas en inglés). Ya que como lo menciona Galindo<sup>17</sup>, los datos observados se consideran una posible realización de este DGP, el cual se compone de una secuencia de vectores aleatorios que se suponen como independientes e idénticamente distribuidos, por lo que el principal objetivo de la econometría moderna es especificar y estimar un modelo estadístico que represente una adecuada aproximación del DGP.

#### Modelo 1

El primer modelo estimado considera únicamente al ingreso disponible de los consumidores y al precio de las gasolinas

$$\ln VG_t = \alpha_0 + \alpha_1 \ln PIB_t + \alpha_2 \ln\left(\frac{PG}{INPC}\right)_t \quad (4.1)$$

De la estimación del modelo se obtienen los siguientes resultados para los parámetros:

---

<sup>17</sup> Galindo P., Luís Miguel, "La metodología econométrica moderna: Una versión aplicada", Cuaderno de Trabajo de Economía Aplicada, N° 18, maestría en Ciencias Económicas (UACPyP, CCH, UNAM), 1995

Cuadro 4.1. Parámetros y estadísticos de la estimación del Modelo 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-24.00307	1.428319	-16.80512	0.0000
lpib	1.473358	0.074444	19.79159	0.0000
lpg	-0.672730	0.104658	-6.427904	0.0000
R-squared	0.827716	Mean dependent var		4.571615
Adjusted R-squared	0.825464	S.D. dependent var		0.111207
S.E. of regression	0.046459	Akaike info criterion		-3.281431
Sum squared resid	0.330247	Schwarz criterion		-3.222780
Log likelihood	258.9516	F-statistic		367.5353
Durbin-Watson stat	1.238267	Prob(F-statistic)		0.000000

Y los siguientes valores para la batería de pruebas, en donde se puede observar, dado el valor de “p”, que el modelo presenta problemas de autocorrelación en los términos de error.

Cuadro 4.2. Batería de pruebas del Modelo 1

	Jarque-Bera	LM 2 rezagos	ARCH 1 rezago	White
Valor del estadístico	3.077	18.069	0.531	1.501
Probabilidad	0.215	0.000	0.467	0.216

## Modelo 2

A este modelo se le agregó como variable explicativa la demanda de gasolinas retrasada un periodo, dado el carácter autorregresivo de la demanda de gasolinas, esto es, el consumo del mes actual depende en gran medida del consumo del mes anterior.

$$\ln VG_t = \alpha_0 + \alpha_1 \ln PIB_t + \alpha_2 \ln\left(\frac{PG}{INPC}\right)_t + \alpha_3 \ln VG_{t-1} \quad (4.2)$$

De la estimación del modelo se obtienen los siguientes estimadores para los parámetros:

Cuadro 4.3. Parámetros y estadísticos de la estimación del Modelo 2

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-15.13814	2.100117	-7.208239	0.0000
lpib	0.924718	0.122654	7.539247	0.0000
lpg	-0.450345	0.104680	-4.302090	0.0000
lv(-1)	0.390542	0.073292	5.328544	0.0000
R-squared	0.855537	Mean dependent var	4.572947	
Adjusted R-squared	0.852667	S.D. dependent var	0.110312	
S.E. of regression	0.042342	Akaike info criterion	-3.460604	
Sum squared resid	0.270720	Schwarz criterion	-3.382064	
Log likelihood	272.1968	F-statistic	298.0826	
Durbin-Watson stat	2.138571	Prob(F-statistic)	0.000000	

Los valores para la batería de pruebas muestran que en este modelo los residuales se distribuyen normalmente, no existe correlación entre los términos de error y es homoscedástico.

Cuadro 4.4. Batería de pruebas del Modelo 2

	Jarque-Bera	LM 2 rezagos	ARCH 1 rezago	White
Valor del estadístico	0.373	3.331	0.907	0.583
Probabilidad	0.829	0.048	0.342	0.712

### Modelo 3

Para este modelo se tomaron los precios con un retraso y se conservó a la variable dependiente con un rezago.

$$\ln VG_t = \alpha_0 + \alpha_1 \ln PIB_t + \alpha_2 \ln\left(\frac{PG}{INPC}\right)_{t-1} + \alpha_3 \ln VG_{t-1} \quad (4.3)$$

De la estimación del modelo se obtienen los siguientes estimadores para los parámetros:

Cuadro 4.5. Parámetros y estadísticos de la estimación del Modelo 3

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-15.39728	2.138828	-7.198933	0.0000
lpib	0.940109	0.124896	7.527120	0.0000
lpg(-1)	-0.458334	0.105864	-4.329461	0.0000
lvg(-1)	0.381905	0.074010	5.160163	0.0000
R-squared	0.855738	Mean dependent var	4.572947	
Adjusted R-squared	0.852872	S.D. dependent var	0.110312	
S.E. of regression	0.042313	Akaike info criterion	-3.461996	
Sum squared resid	0.270344	Schwarz criterion	-3.383456	
Log likelihood	272.3047	F-statistic	298.5682	
Durbin-Watson stat	2.079210	Prob(F-statistic)	0.000000	

Al igual que el modelo anterior, en este caso se tiene normalidad en la distribución de los residuales, homoscedasticidad y ausencia de correlación en los términos de error.

Cuadro 4.6. Batería de pruebas del Modelo 3

	Jarque-Bera	LM 2 rezagos	ARCH 1 rezago	White
Valor del estadístico	0.620	2.354	0.528	0.553
Probabilidad	0.733	0.098	0.468	0.735

#### Modelo 4

Para este modelo se retrasó la variable del ingreso y la variable dependiente

$$\ln VG_t = \alpha_0 + \alpha_1 \ln PIB_{t-1} + \alpha_2 \ln\left(\frac{PG}{INPC}\right)_t + \alpha_3 \ln VG_{t-1} \quad (4.4)$$

De la estimación del modelo se obtienen los siguientes estimadores para los parámetros:

Cuadro 4.7. Parámetros y estadísticos de la estimación del Modelo 4

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-14.86506	2.206105	-6.738146	0.0000
lpib(-1)	0.907312	0.128672	7.051330	0.0000
lpg	-0.442321	0.108820	-4.064703	0.0001
lvg(-1)	0.405032	0.075496	5.364913	0.0000
R-squared	0.850413	Mean dependent var		4.572947
Adjusted R-squared	0.847441	S.D. dependent var		0.110312
S.E. of regression	0.043086	Akaike info criterion		-3.425751
Sum squared resid	0.280322	Schwarz criterion		-3.347211
Log likelihood	269.4957	F-statistic		286.1486
Durbin-Watson stat	2.147585	Prob(F-statistic)		0.000000

Las pruebas de diagnóstico para este modelo permiten afirmar que existe autocorrelación en los términos de error debido al valor de “p” obtenido en la prueba del Multiplicador de Lagrange.

Cuadro 4.8. Batería de pruebas del Modelo 4

	Jarque-Bera	LM 2 rezagos	ARCH 1 rezago	White
Valor del estadístico	0.398	3.779	2.560	1.207
Probabilidad	0.819	0.025	0.111	0.308

En resumen y de acuerdo a la siguiente tabla, el modelo que presentó las mejores características de acuerdo a la batería de pruebas realizadas fue el Modelo 2

Cuadro 4.9. Resumen de los parámetros y batería de pruebas de los modelos

	R2	error estand	suma cuad	criterio	criterio	estadístico	probabilidad			
		regresión	residuales	Akaike	Schwartz	D-W	Jarque-Bera	LM	ARCH	White
Modelo 1	0.827	0.0464	0.330	-3.281	-3.222	1.238	0.215	0.000	0.467	0.216
<b>Modelo 2</b>	<b>0.855</b>	<b>0.0423</b>	<b>0.270</b>	<b>-3.460</b>	<b>-3.382</b>	<b>2.138</b>	<b>0.829</b>	<b>0.048</b>	<b>0.342</b>	<b>0.712</b>
Modelo 3	0.855	0.0423	0.270	-3.461	-3.383	2.079	0.733	0.098	0.468	0.735
Modelo 4	0.850	0.043	0.28	-3.425	-3.347	2.147	0.819	0.025	0.111	0.328

Este modelo establece que el consumo de gasolinas, para la Ciudad de México, tiene una relación positiva directa con el ingreso de los consumidores (PIB), negativa con el precio real del combustible y positiva con el consumo de gasolinas de un periodo anterior.

#### b. Cálculo de elasticidades

Tomando el Modelo 2 seleccionado como el óptimo, se tiene que

$$\ln VG = - 15.1381 + 0.924 \ln PIB - 0.450 \ln\left(\frac{PG}{INPC}\right) + 0.390 \ln VG_{t-1} \quad (4.5)$$

Los valores de los parámetros estimados en esta ecuación permiten establecer que la gasolina es un bien ordinario ( $\alpha_1 = 0.924 > 1$ ) y necesario ( $\alpha_2 = 0.450 < 1$ ). Además, la demanda de gasolina exhibe una gran respuesta a los cambios en el ingreso, dado que su elasticidad en el corto es casi la unidad y en el largo plazo es superior a ésta; por su parte el menor valor de la elasticidad precio, en el corto plazo, muestra una respuesta débil de la demanda de gasolinas al cambio en su precio, no así en el largo plazo.

Cuadro 4.10. Elasticidades de corto y largo plazo del modelo estimado

ingreso		precio	
corto plazo	largo plazo	corto plazo	largo plazo
0.924	1.514	-0.450	-0.737

### c. Comparación con otros resultados

El siguiente cuadro presenta un comparativo entre las elasticidades calculadas en el presente trabajo y las comentadas en la Sección 3.

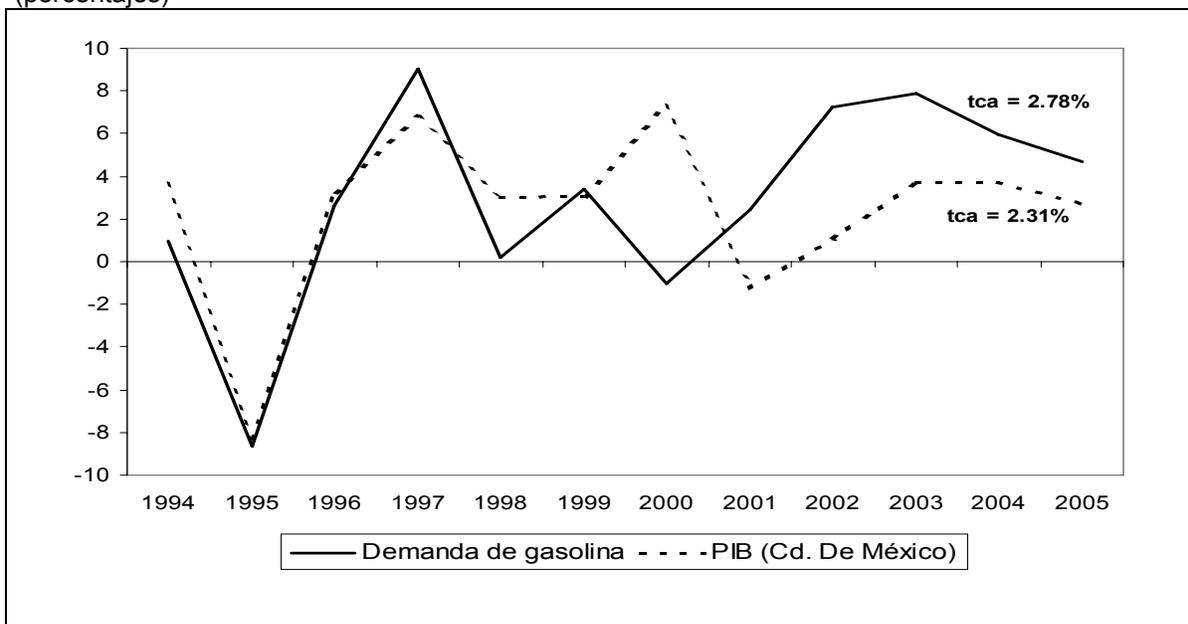
Cuadro 4.11. Comparación de estimaciones de las elasticidades ingreso y precio de la gasolina con revisiones de la literatura

Autor	región	Elasticidad ingreso		Elasticidad precio	
		corto plazo	largo plazo	corto plazo	largo plazo
Galindo y Salinas	todo el país	(0.8 , 1.2)		(-0.2 , -0.8)	
Petróleos Mexicanos (2000)	todo el país		0.647		-0.154
Petróleos Mexicanos (2000)	zonas metropolitanas	0.251		-0.112	
Haro e Ibarrola (2000)	zonas fronterizas	0.580		-0.415	
Eskeland y Feyzioglu (1997)	todo el país	0.630	0.840	-1.040	-1.390
Alba y Samaniego (1985)	todo el país	0.580	1.240	-0.110	-0.240
Espino Bravo (2005)	todo el país	0.287	0.485	-0.693	-1.172
Este documento (2005)	Ciudad de México	0.924	1.514	-0.450	-0.737

Del cual se desprenden los siguientes comentarios:

La elasticidad ingreso, tanto en el corto como en el largo plazo, en el caso de la Ciudad de México es mayor que los valores calculados para el resto del país en otros estudios. Lo anterior indica que la demanda de gasolina en el corto plazo crece a tasas iguales que el ingreso de los consumidores, mientras que en el largo plazo el ajuste en la demanda es mayor que el crecimiento del PIB.

Gráfica 4.1. Tasas de crecimiento de la demanda de gasolinas y PIB en la Ciudad de México (porcentajes)



En el caso del precio de la gasolina, la respuesta de los consumidores en la Ciudad de México, en el largo plazo, es menor a la estimada para el resto del país, mientras que en el corto plazo esta respuesta es similar al promedio nacional.

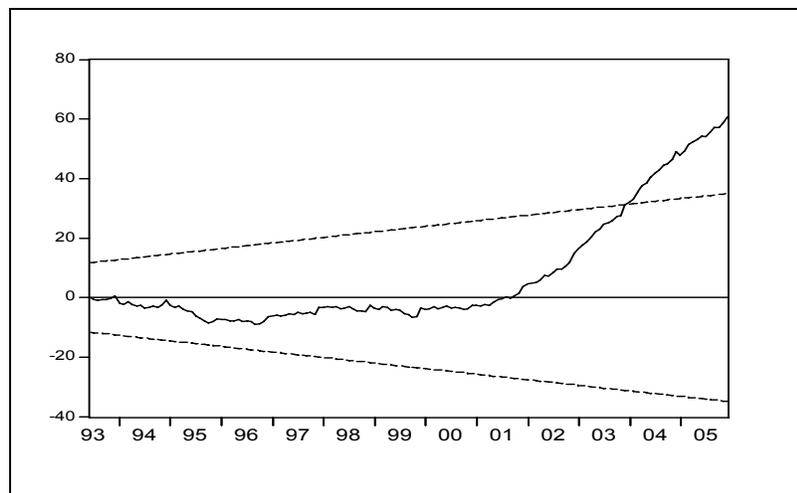
## Sección 5. Cambio estructural

### a. Prueba de cambio estructural utilizando el modelo estimado

La prueba CUSUM permite exhibe inestabilidad en los parámetros si el valor de la suma acumulativa de los residuales recursivos cae fuera del área limitada por dos líneas críticas.

La gráfica 5.1 muestra la prueba CUSUM del modelo óptimo de la sección anterior, en donde claramente se observa que la suma acumulativa cae fuera de las líneas críticas a partir de 2004.

Gráfica 5.1. Prueba de CUSUM



Para analizar si realmente existe un cambio estructural en los parámetros de la ecuación de regresión, a partir de 2004, se realizaron dos estimaciones, la primera para el periodo completo (enero de 1993 a diciembre de 2005), la cual se le denominó “estimación 1” y la segunda para un periodo más breve (enero de 2000 a diciembre de 2005), llamada “estimación 2” en donde se supone se produce el cambio estructural.

Estimación 1: enero de 1993 – diciembre de 2005

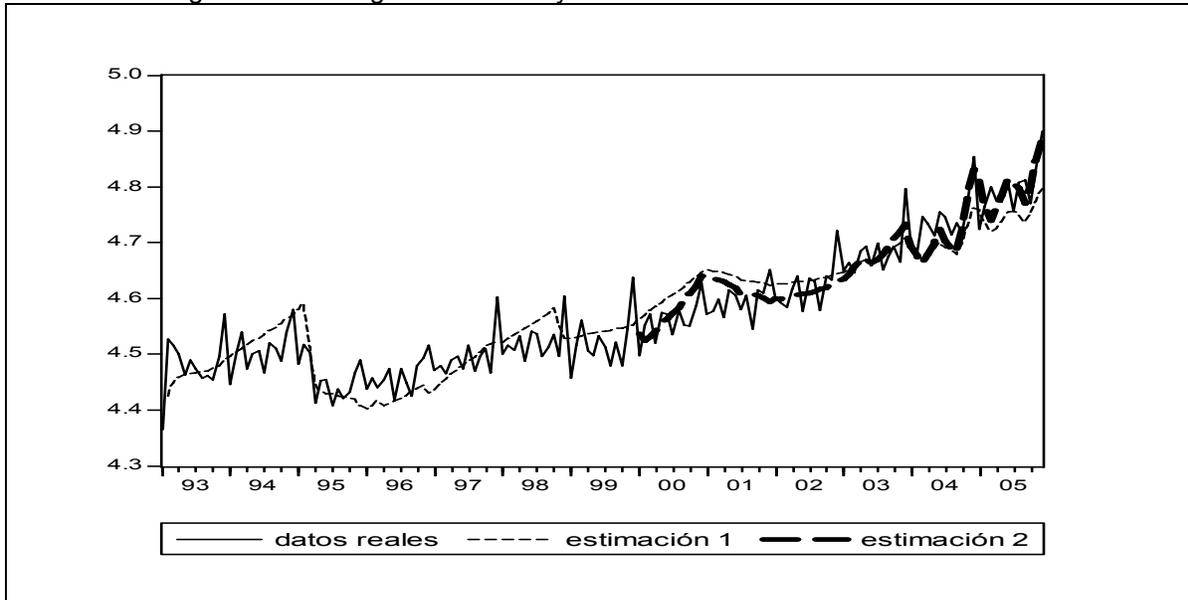
$$\ln VG = - 15.138 + 0.924 \ln \text{PIB} - 0.450 \ln\left(\frac{PG}{INPC}\right) + 0.390 \ln VG_{t-1} \quad (5.1)$$

Estimación 2: enero de 200 – diciembre de 2005

$$\ln VG = - 32.514 + 1.857 \ln \text{PIB} - 0.917 \ln\left(\frac{PG}{INPC}\right) + 0.229 \ln VG_{t-1} \quad (5.2)$$

El modelo lineal general establece que los parámetros del modelo estimado deben ser constantes dentro de la muestra considerada, lo cual a simple vista no se cumple en el modelo óptimo. La gráfica 5.2 muestra los datos reales de las ventas de gasolina para la Ciudad de México y las dos estimaciones realizadas.

Gráfica 5.2. Log de ventas de gasolina reales y estimaciones 1 & 2



Los valores de las estimaciones, así como su diferencia, se muestran en el cuadro siguiente para el periodo enero 2000 – diciembre 2005.

Cuadro 5.1. Estimación 1 &amp; 2 del logaritmo de las ventas de gasolina en la Ciudad de México

	estimación 1	estimación 2	Diferencia
Ene-00	4.561	4.537	-0.02
Feb-00	4.571	4.525	-0.05
Mar-00	4.578	4.531	-0.05
Abr-00	4.586	4.541	-0.05
May-00	4.593	4.552	-0.04
Jun-00	4.600	4.563	-0.04
Jul-00	4.607	4.573	-0.03
Ago-00	4.613	4.582	-0.03
Sep-00	4.621	4.595	-0.03
Oct-00	4.629	4.608	-0.02
Nov-00	4.638	4.623	-0.01
Dic-00	4.648	4.640	-0.01
Ene-01	4.651	4.642	-0.01
Feb-01	4.648	4.635	-0.01
Mar-01	4.648	4.633	-0.01
Abr-01	4.646	4.631	-0.02
May-01	4.643	4.625	-0.02
Jun-01	4.640	4.619	-0.02
Jul-01	4.634	4.609	-0.02
Ago-01	4.630	4.605	-0.03
Sep-01	4.631	4.607	-0.02
Oct-01	4.630	4.605	-0.02
Nov-01	4.628	4.601	-0.03
Dic-01	4.624	4.595	-0.03
Ene-02	4.626	4.600	-0.03
Feb-02	4.626	4.599	-0.03
Mar-02	4.627	4.601	-0.03
Abr-02	4.629	4.605	-0.02
May-02	4.630	4.607	-0.02
Jun-02	4.631	4.608	-0.02
Jul-02	4.632	4.609	-0.02
Ago-02	4.633	4.612	-0.02
Sep-02	4.636	4.616	-0.02
Oct-02	4.638	4.619	-0.02
Nov-02	4.641	4.626	-0.02
Dic-02	4.644	4.630	-0.01
Ene-03	4.647	4.634	-0.01
Feb-03	4.652	4.643	-0.01
Mar-03	4.659	4.657	0.00
Abr-03	4.666	4.667	0.00
May-03	4.670	4.672	0.00
Jun-03	4.667	4.664	0.00
Jul-03	4.670	4.669	0.00
Ago-03	4.675	4.679	0.00
Sep-03	4.682	4.692	0.01
Oct-03	4.689	4.704	0.01
Nov-03	4.699	4.720	0.02
Dic-03	4.707	4.734	0.03
Ene-04	4.689	4.694	0.00
Feb-04	4.678	4.677	0.00
Mar-04	4.669	4.663	-0.01
Abr-04	4.675	4.679	0.00
May-04	4.686	4.700	0.01
Jun-04	4.700	4.724	0.02
Jul-04	4.692	4.704	0.01
Ago-04	4.685	4.691	0.01
Sep-04	4.680	4.682	0.00
Oct-04	4.699	4.724	0.02
Nov-04	4.730	4.779	0.05
Dic-04	4.761	4.832	0.07
Ene-05	4.760	4.816	0.06
Feb-05	4.739	4.773	0.03
Mar-05	4.720	4.739	0.02
Abr-05	4.725	4.758	0.03
May-05	4.739	4.784	0.05
Jun-05	4.755	4.814	0.06
Jul-05	4.757	4.810	0.05
Ago-05	4.751	4.798	0.05
Sep-05	4.737	4.770	0.03
Oct-05	4.750	4.800	0.05
Nov-05	4.775	4.848	0.07
Dic-05	4.797	4.884	0.09

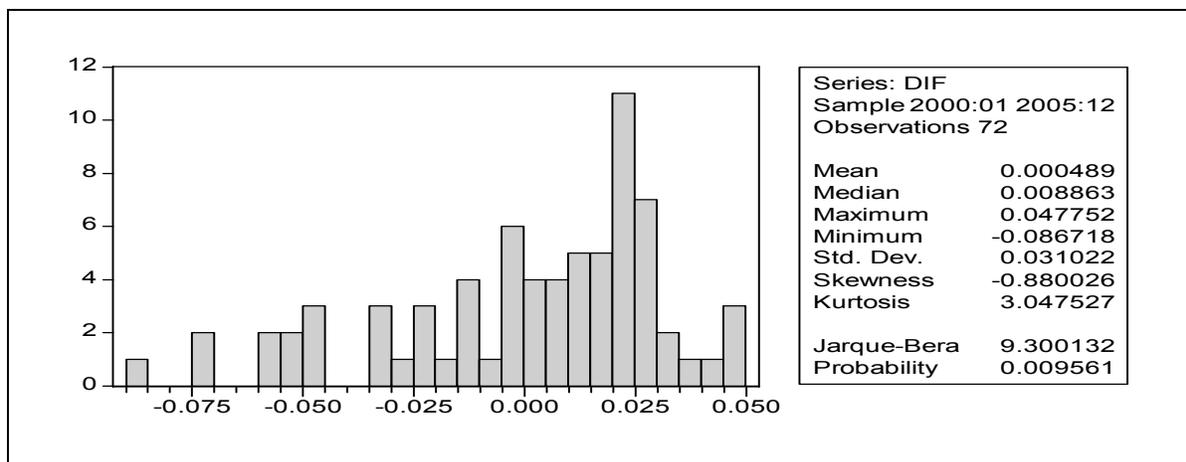
El paso siguiente será comprobar si la diferencia entre ambas estimaciones es sólo ruido blanco, de no ser así se estaría ante la existencia de cambio estructural en los parámetros.

Recordemos que a una sucesión de variables aleatorias  $\{a_t\}$  se le conoce como ruido blanco si:

- su media es constante (generalmente se le considera igual a cero).
- su varianza  $\sigma_a^2$  no es función del tiempo.
- las variables aleatorias  $\{a_t\}$  son mutuamente independientes.

Guerrero<sup>10</sup> sugiere construir el cociente  $\frac{\sqrt{N}m(a)}{\sigma_a}$ , en donde N es el número de observaciones y si el valor absoluto de dicho cociente es menor que dos se dirá que no hay evidencia de que la media del proceso de ruido blanco sea distinta de cero.

Cuadro 5.2. Histograma y estadísticos de la serie diferencias



Sustituyendo algunos de los valores del cuadro anterior se tiene que:

<sup>10</sup> Guerrero, V.M., "Análisis Estadístico de Series de Tiempo Económicas", Thomson, México, 2003

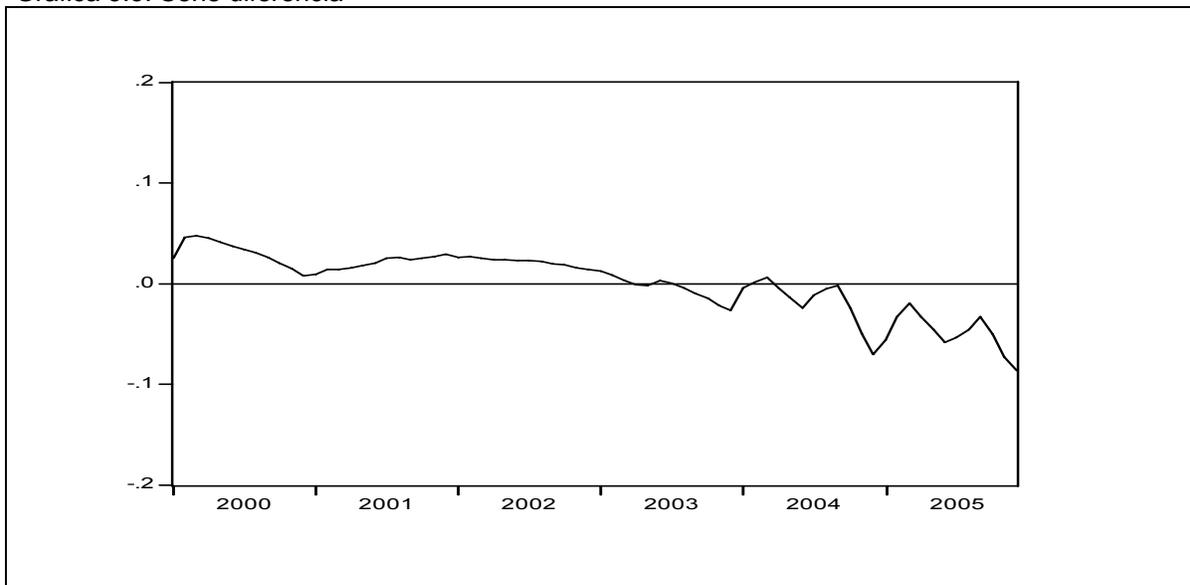
$$\left| (\sqrt{N}) \frac{m(a)}{\sigma_a} \right| = \left| (\sqrt{72})(0.000489) / 0.031022 \right| = 0.133 < 2 \quad (5.3)$$

de donde se puede concluir que no hay evidencia de que la media de las diferencias sea distinta de cero.

La gráfica de la diferencia contra el tiempo permite observar si la varianza es o no constante, aunque esta verificación visual puede parecer muy burda, la idea es que sólo las violaciones muy notorias, del supuesto de desviación estándar constante, son las que realmente llegan a causar problemas.

La gráfica muestra un patrón de varianza no monótona, lo que permite suponer que la desviación estándar no es constante.

Gráfica 5.3. Serie diferencia



El tercer supuesto de no autocorrelación requiere que  $\rho_k(a)=0$  para toda  $k \neq 0$ .



Por lo anterior, se tiene que la diferencia de las estimaciones presenta las siguientes características:

- la media es igual a cero
- la desviación estándar es función del tiempo
- existe autocorrelación

Por lo que se puede concluir que la variable diferencia no se puede considerar como ruido blanco, ya que no cumple con las dos últimas características.

Lo anterior permite establecer que sí se dio un cambio estructural en los parámetros de la ecuación estimada, y que este cambio ocurrió al inicio de 2004.

## b. Raíces unitarias y cambio estructural

Perron<sup>11</sup> desarrollo un procedimiento formal para la prueba de raíces unitarias en presencia de un cambio estructural en el periodo  $t = \tau + 1$ , considerando la hipótesis nula de un salto una sola vez en el nivel de un proceso de raíces unitarias contra la hipótesis alternativa de un cambio en el nivel de la intersección de un proceso estacionario. Formalmente, las hipótesis nula y alternativa se expresan:

$$\begin{aligned} H_0 : y_t &= a_0 + a_1 y_{t-1} + \mu_1 D_p + \varepsilon_t \\ H_1 : y_t &= a_0 + a_2 t + \mu_1 D_L + \varepsilon_t \end{aligned} \quad (5.5)$$

donde  $D_p$  representa una variable dummy de *pulso* tal que  $D_p = 1$  si  $t = \tau + 1$  y cero en otro caso, y  $D_L$  una variable dummy de *nivel* tal que  $D_L = 1$  si  $t > \tau$  y cero en otro caso.

---

<sup>11</sup> Perron, P. , et al

Bajo la hipótesis nula,  $\{y_t\}$  es un proceso de raíces unitarias con un salto único en el nivel de la sucesión en el periodo  $t = \tau + 1$ . Bajo la hipótesis alternativa,  $\{y_t\}$  es de tendencia estacionaria con un salto en la intersección.

Perron sigue estimar la siguiente ecuación

$$y_t = a_0 + a_1 y_{t-1} + a_2 t + \mu_1 D_p + \mu_2 D_L + \sum_{i=1}^k \beta_i \Delta y_{t-i} + \varepsilon_t \quad (5.4)$$

Para probar si la serie en cuestión se modela mejor como un proceso de raíces unitarias con un salto en el nivel de la serie en el periodo  $t = \tau + 1$ , es decir los valores de los periodos anteriores de la serie influyen en el comportamiento actual

$$y_t = a_0 + y_{t-1} + \mu_1 D_p + \varepsilon_t \quad (5.5)$$

o bien como un proceso estacionario con tendencia y un cambio en el nivel de la serie en el periodo  $t = \tau + 1$

$$y_t = a_0 + a_2 t + \mu_2 D_L + \varepsilon_t \quad (5.6)$$

El modelo general de Perron considera la regresión

$$y_t = a_0 + a_1 y_{t-1} + a_2 t + \mu_1 D_p + \mu_2 D_L + \sum_{i=1}^k \beta_i \Delta y_{t-i} + \varepsilon_t \quad (5.7)$$

en donde  $D_p$  representa una variable *dummy* de un salto tal que

$$\begin{aligned} D_p &= 1 & \text{si} & \quad t = \tau + 1 \\ D_p &= 0 & \text{si} & \quad t \neq \tau + 1 \end{aligned} \quad (5.8)$$

y  $D_L$  es una variable *dummy* de cambio de nivel tal que

$$\begin{aligned} D_L &= 1 & \text{si} & \quad t > \tau + 1 \\ D_L &= 0 & \text{si} & \quad t < \tau + 1 \end{aligned} \quad (5.9)$$

En donde, bajo el supuesto de un cambio en una sola vez en la media de un proceso de raíces unitarias, se tiene que

$$a_1 = 1 \quad a_2 = 0 \quad \mu_2 = 0 \quad (5.10)$$

Bajo el supuesto de la hipótesis alternativa de un cambio permanente en el tiempo en la tendencia de un modelo estacionario, se cumple que:

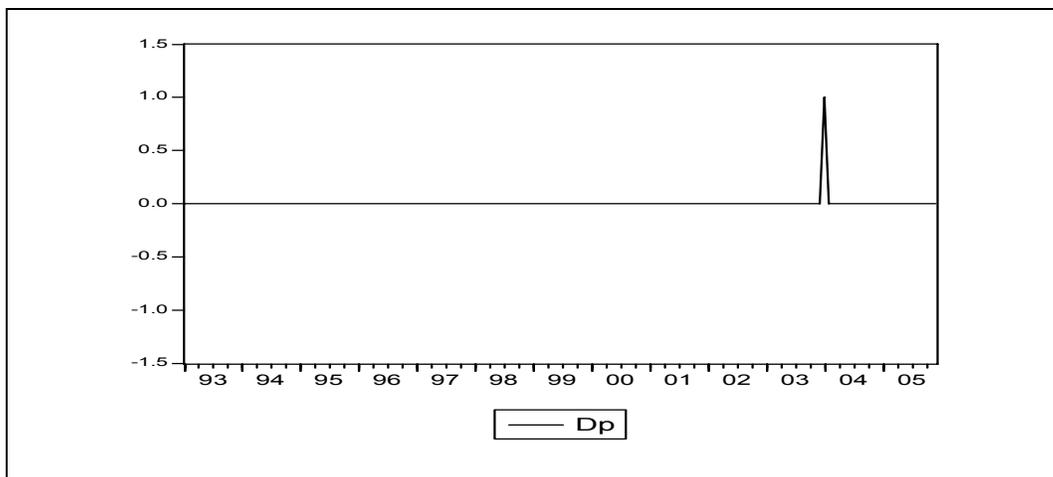
$$a_1 < 1 \quad \mu_1 = 0 \quad (5.11)$$

A continuación se utilizara la metodología sugerida por Perron para corroborar si existen cambios cualitativos en la demanda de gasolinas para la Ciudad de México a partir de enero de 2004 y si estos dependen del comportamiento anterior de la serie u obedecen a un cambio estructural. Por lo que se definen las siguientes variables *dummy*:

Variable *dummy* de un salto

$$\begin{aligned} D_p &= 1 & \text{si } t &= \text{enero de 2004} \\ D_p &= 0 & \text{si } t &\neq \text{enero de 2004} \end{aligned} \quad (5.12)$$

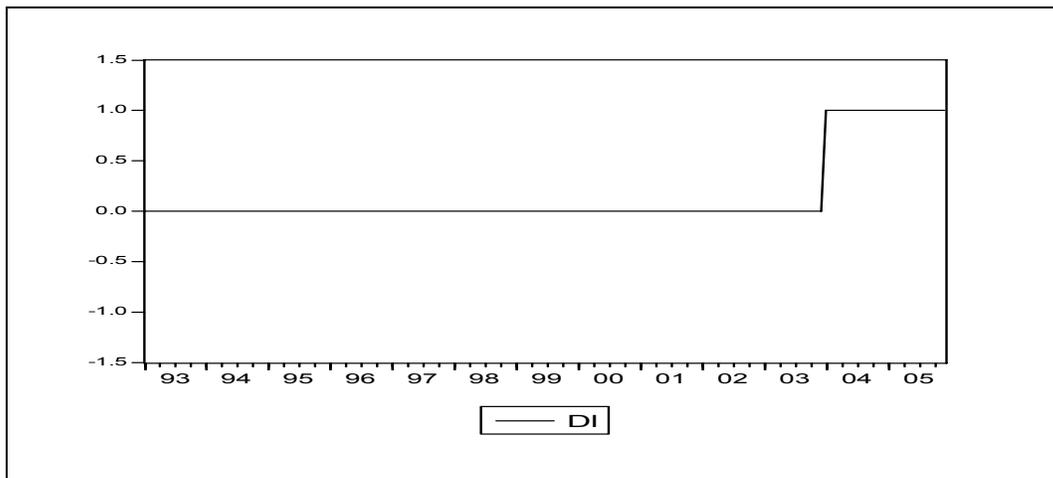
Gráfica 5.5. Variable dummy de un salto



Variable *dummy* de cambio de nivel

$$\begin{aligned} D_L &= 1 & \text{si } t > \text{enero de 2004} \\ D_L &= 0 & \text{si } t < \text{enero de 2004} \end{aligned} \quad (5.13)$$

Gráfica 5.6. Variable dummy de cambio de nivel



La regresión estimada mediante mínimos cuadrados ordinarios para la ecuación de Perron

$$y_t = a_0 + a_1 y_{t-1} + a_2 t + \mu_1 D_P + \mu_2 D_L + \varepsilon_t \quad (5.14)$$

arrojó los siguientes valores para los parámetros

Cuadro 5.4. Parámetros de la regresión

Variable	Coefficiente	Error estándar	t-estadístico	Probabilidad
$a_0$	2.885003	0.343888	8.389373	0.0000
$a_1$	0.349168	0.077627	4.498032	0.0000
$a_2$	0.001018	0.000155	6.545027	0.0000
$\mu_1$	0.077156	0.042842	1.800946	0.0737
$\mu_2$	0.071472	0.014509	4.926093	0.0000
R-cuadrada	0.861823	Media var. dependiente		4.572947
R-cuadrada ajustada	0.858138	Error estándar var. depend.		0.110312
Error estándar regresión	0.041548	Criterio Akaike		-3.492188
Suma cuadrado residuales	0.258941	Criterio Schwarz		-3.394013
Log max verosimilitud	275.6446	F-statistic		233.8904
Durbin-Watson	2.002302	Prob(F-statistic)		0.000000

De los valores mostrados en el Cuadro 5.4 y dado que:

$$a_1 = 0.349 < 1$$

$$\mu_1 = 0 \quad (\text{dado que } p = 0.737 > 0.5)$$

De acuerdo con la metodología de Perron se puede afirmar que  $\log(vg)$  presenta una tendencia estacionaria con un salto a partir de enero de 2004.

## Sección 6. Modelo de series de tiempo de la demanda de gasolinas y análisis de intervención

### a. Modelo de series de tiempo para la demanda de gasolinas

El propósito de este punto es el de desarrollar un modelo ARIMA para la demanda de gasolinas de la Ciudad de México, y posteriormente aplicar el método denominado “Análisis de Intervención” para medir explícitamente la influencia de algún evento anormal o extraño sobre la serie, el cual en este caso es el impacto del mercado ilícito en la demanda de gasolinas.

La estrategia de construcción de modelos para series de tiempo fue desarrollada por Box y Jenkins la cual consta de cuatro etapas fundamentales que son:

- i. *Identificación* de un posible modelo dentro de la clase de modelos ARIMA; es decir, determinación de los valores  $p$ ,  $d$  y  $q$  que especifican el modelo ARIMA apropiado,
- ii. *Estimación* de los parámetros involucrados en el modelo, a través de técnicas de estimación no lineal,
- iii. *Verificación* de que el modelo *proporciona* un ajuste adecuado y de que los supuestos básicos, implícitos en el modelo, se satisfacen; de no cumplirse los supuestos, se determinan las modificaciones necesarias y, de hecho, se repiten las etapas anteriores hasta que la verificación indique resultados aceptables,
- iv. *Uso del modelo* para los fines que el investigador haya tenido en mente al construirlo; dichos fines son por lo general de pronóstico, control, simulación o explicación del fenómeno en estudio.

## i. Identificación

En esta etapa el objetivo principal es el de determinar 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 no homogénea. De manera general podría decirse que la etapa de identificación consiste en determinar, primero, una serie estacionaria en función de la serie original, para la cual se pueda tener una representación ARMA (p,q) y, posteriormente, fijar los valores de  $p$  y  $q$ .

### i.1 Estabilización de la varianza

El procedimiento para elegir una transformación potencia que estabilice la varianza de una serie de tiempo se basa en el artículo de Barlett<sup>11</sup> quien afirma que "...si la variación tiende a cambiar con el nivel medio de las mediciones, la varianza solamente se estabilizara con un cambio apropiado de escala".

Suponiendo que la varianza  $\sigma_t^2$  de la variable aleatoria  $Z_t$  puede expresarse como una función de su media  $\mu_t$ , o sea

$$\sigma_t^2 = f(\mu_t) \quad (6.1)$$

entonces, si  $T()$  es una función cuya primera derivada existe, puede obtenerse la siguiente aproximación lineal a  $T(Z_t)$

$$T(Z_t) = T(\mu_t) + \left( \frac{dT}{dZ_t} \Big|_{Z_t=\mu_t} \right) (Z_t - \mu_t) \quad (6.2)$$

de donde se obtiene que una aproximación lineal a la varianza de  $T(Z)$  es

<sup>11</sup> Barlett, M.S. (1974) "The use of transformations", en *Biométrica* 3, 39

$$Var(T) = \left( \frac{dT}{dZ_t} \Big|_{Z_t=\mu_t} \right)^2 f(\mu_t) \quad (6.3)$$

Así, ya que se desea que  $T(Z_t)$  tenga varianza constante  $C^2$ , se sigue que

$$\frac{dT}{dZ_t} \Big|_{Z_t=\mu_t} = \frac{C}{\sqrt{f(\mu_t)}} \quad (6.4)$$

y por lo tanto

$$T(\nu\mu_t) = \int \frac{C}{\sqrt{f(\mu_t)}} d\mu_t \quad (6.5)$$

Como puede observarse, es necesario conocer la forma funcional de  $f$  para estar en condiciones de utilizar la aproximación anterior y de esta manera determinar la transformación  $T$  que establezca la varianza. Si se restringe el tipo de transformación a la *familia de transformaciones potencia*, la cual se utiliza frecuentemente porque proporciona buenos resultados en la práctica, es factible emplear la relación anterior. Esto es, si la variable  $Z_t$  es positiva y si es razonable suponer que  $\sigma_t^2$  es proporcional a  $\mu_t^{2(1-\lambda)}$  para algún valor de  $\lambda$ ; es decir, si

$$f(\mu_t) \propto \mu_t^{2(1-\lambda)} \quad (6.6)$$

se sigue que

$$T(\mu_t) \propto \begin{cases} \mu_t^\lambda & \text{si } \lambda \neq 0 \\ \log(\mu_t) & \text{si } \lambda = 0 \end{cases} \quad (6.7)$$

Este último resultado sugiere que la función  $T$ , que vuelve aproximadamente constante la varianza de  $Z_t$ , debe ser la transformación potencia

$$T(Z_t) \propto \begin{cases} Z_t^\lambda & \text{si } \lambda \neq 0 \\ \log(Z_t) & \text{si } \lambda = 0 \end{cases} \quad (6.8)$$

Tomando la variable  $Z_t$  como la ventas de gasolina en la Ciudad de México  $\{VG_t\}$ , para el periodo enero de 1993 – diciembre de 2005, se procede a estabilizar su varianza utilizando el método anterior.

Cuadro 6.1. Volumen de ventas de gasolina para la Ciudad de México (miles de barriles diarios)

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
ene	78.7	85.3	88.4	84.6	87.5	90.0	86.2	89.8	96.7	99.7	104.6	109.6	112.6
feb	92.5	90.4	91.6	86.2	88.2	91.4	91.4	94.9	97.2	98.8	106.0	108.4	117.2
mar	91.4	93.6	90.2	84.8	86.9	90.7	95.6	96.7	99.4	97.9	104.2	115.2	121.4
abr	90.0	87.7	82.4	86.0	89.1	93.1	90.5	91.8	96.1	101.2	108.4	113.8	118.6
may	86.8	90.0	85.9	87.7	89.6	88.9	89.8	97.0	101.0	103.5	109.3	111.4	118.3
jun	89.1	90.5	86.0	83.0	87.7	93.8	93.1	96.7	100.1	97.2	105.5	116.1	122.2
jul	87.7	87.1	82.1	87.7	91.4	93.4	91.3	93.2	97.6	103.1	109.8	115.2	116.3
ago	86.2	91.8	84.6	85.5	87.3	89.6	88.2	97.4	100.1	102.6	104.8	111.6	122.5
sep	86.6	90.9	83.2	83.5	89.6	91.3	92.0	94.9	94.1	97.4	107.3	113.9	123.1
oct	86.0	88.9	84.1	88.2	90.9	93.2	88.2	94.7	101.0	103.5	109.3	111.8	117.9
nov	89.6	93.6	87.1	89.5	87.1	89.6	94.3	98.3	100.4	102.8	106.2	117.9	125.4
dic	96.7	97.6	89.1	91.4	99.7	99.9	103.3	102.2	104.8	112.3	121.1	128.3	134.3

Fuente: BIE, [www.inegi.com.mx](http://www.inegi.com.mx)

El cuadro siguiente muestra los cálculos necesarios para determinar la transformación potencia estabilizadora de varianza

Cuadro 6.2. Determinación de  $\lambda$  para la serie ventas de gasolina

	-1	-0.5	0	0.5	1
1993	0.00245	0.011011	0.049493	0.465441	4.377122
1994	0.001327	0.006952	0.036426	0.346748	3.300756
1995	0.001307	0.006873	0.036149	0.335658	3.116747
1996	0.000825	0.004866	0.028716	0.267079	2.484045
1997	0.001546	0.007795	0.039314	0.372127	3.522356
1998	0.001043	0.005804	0.032297	0.309923	2.974047
1999	0.002336	0.010627	0.048336	0.463607	4.446638
2000	0.001134	0.006179	0.033674	0.329289	3.220052
2001	0.000818	0.004837	0.028601	0.284623	2.832391
2002	0.001662	0.008232	0.040771	0.411103	4.145214
2003	0.001789	0.0087	0.042300	0.439660	4.569715
2004	0.002034	0.009579	0.045105	0.482500	5.161468
2005	0.002078	0.009733	0.045587	0.501083	5.507863
Coefficiente de variación	0.350351	0.266846	0.18048	0.206663	0.247292

Fuente: elaboración propia

Del cuadro anterior se observa que el mínimo coeficiente de variación (dentro del grupo de valores de  $\lambda$  elegido) se logra con  $\lambda=0$ , por lo cual una posible transformación para estabilizar la varianza de  $\{VG_t\}$  es

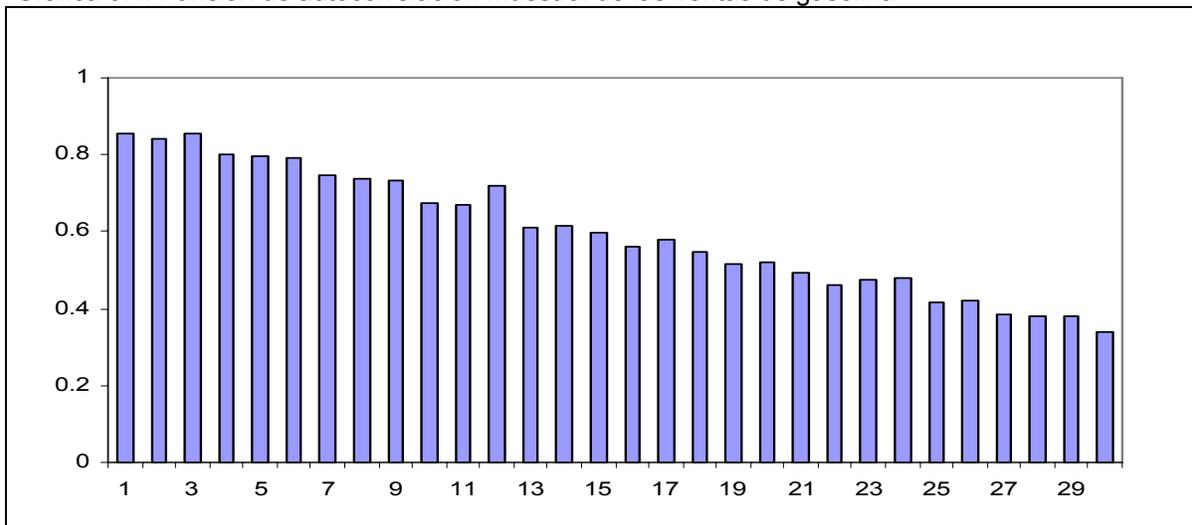
$$T(VG_t) = \log(VG_t) \quad (6.9)$$

## i.2 Estabilización del nivel

Después de haber seleccionado la expresión adecuada de  $T(Z_t)$ , se procede a estabilizar el nivel de la serie mediante la aplicación del operador diferencia un número apropiado de veces. La principal herramienta para determinar el grado de diferenciación correcto es la función de autocorrelación (FAC) muestral, ya que un decaimiento rápido de las autocorrelaciones a cero es indicativo de que la serie es estacionaria, en cuanto a nivel se refiere.

La FAC de la serie  $T(VG_t) = \log(VG_t)$  claramente demuestra que ésta no es estacionaria

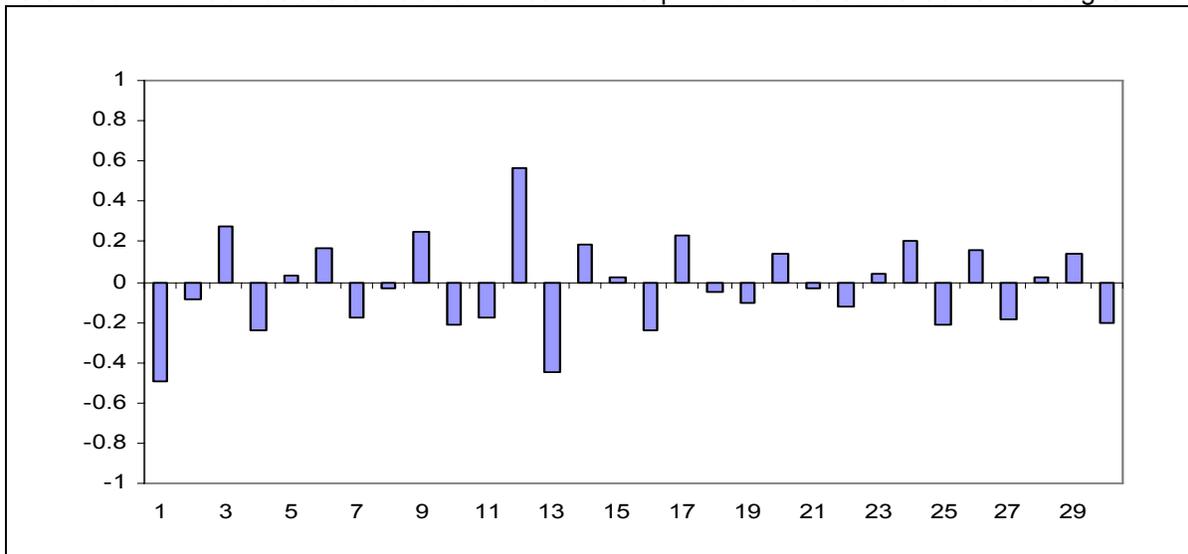
Gráfica 6.1. Función de autocorrelación muestral de las ventas de gasolina



Si aplicamos una vez el operador diferencia a la serie  $T(VG_t)$  podemos observar que esta se vuelve estacionaria, esto es

$$\nabla T(VG_t) = \nabla \log(VG_t) = \log(VG_t) - \log(VG_{t-1}) \quad (6.10)$$

Gráfica 6.2. Función de autocorrelación muestral de la primera diferencia de las ventas de gasolina



### i.3 Utilización de la FAC y FACP para identificar el modelo

De acuerdo al comportamiento de las funciones de autocorrelación (FAC) y autocorrelación parcial (FACP) se puede llevar a cabo la etapa de identificación, conviene recordar que

Cuadro 6.3. Identificación de la serie mediante la FAC y la FACP

Proceso	FAC	FACP
AR(p)	Convergencia a cero, con comportamiento indicado por la ecuación:	Solamente las primeras p autocorrelaciones parciales son distintas de cero
	$\phi(B)\rho_k = 0$ para $k \geq p$	
MA(q)	Solo las primeras q autocorrelaciones son distintas de cero	Sucesión infinita convergente a cero

## ARMA(p,q)

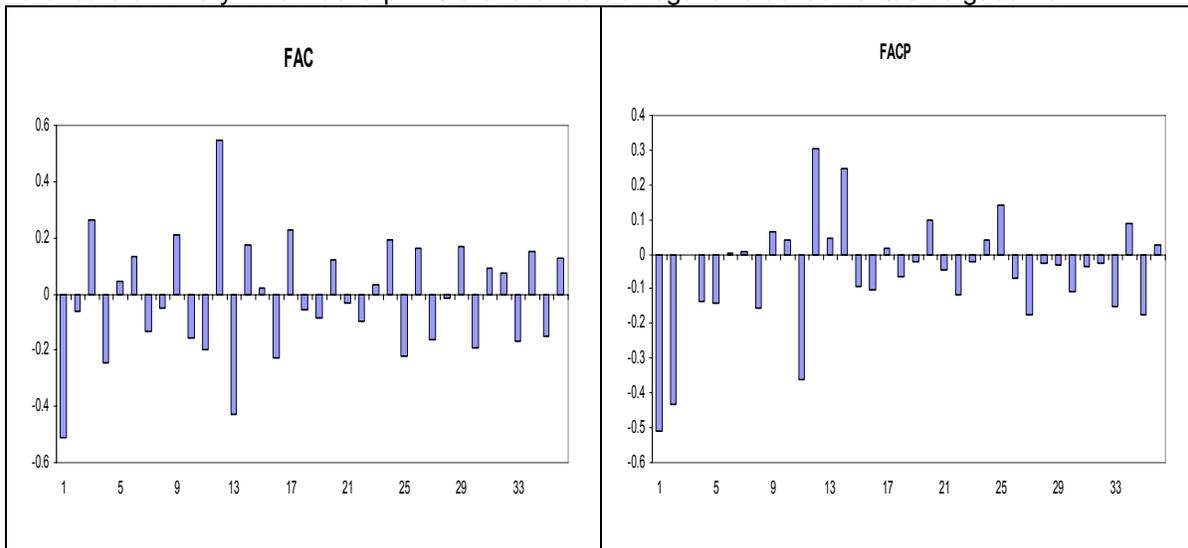
Comportamiento irregular de las Sucesión infinita convergente a cero  
primeras q autocorrelaciones y después  
convergencia a cero de acuerdo con

$$\phi(B)\rho_k = 0 \text{ para } k > 0$$

Fuente: elaboración propia

Por lo que atendiendo al comportamiento de las FAC y FACP de la serie diferenciada, las gráficas permiten proponer un modelo del tipo ARI(1,1).

Gráfica 6.3. FAC y FACP de la primera diferencia del logaritmo de las ventas de gasolina



Para corroborar lo anterior, se puede utilizar la aproximación sugerida por Quenouille<sup>12</sup> la cual indica que si el proceso es AR(p), las autocorrelaciones parciales muestrales se distribuyen de manera independiente, con:

$$\text{media} \quad E(\hat{\phi}_{ii}) = \phi_{ii} \quad (6.11)$$

$$\text{varianza} \quad \text{Var}(\hat{\phi}_{ii}) = 1 / (N-d) \quad \text{para } i > p \quad (6.12)$$

<sup>12</sup> Quenouille, M.H. (1949) "Approximate test of correlation in time series", Journal of the Royal Statistical Society B-11, 68.

A partir de esta fórmula, se puede establecer que  $\phi_{ii}$  es distinto de cero ( al nivel de significación de aproximadamente 5%) si el valor calculado de  $\hat{\phi}_{ii}$  se encuentra fuera del intervalo definido por

$$\pm 2\sqrt{\text{Var}(\hat{\phi}_{ii})} = \pm \frac{2}{\sqrt{N-d}} \quad \text{para } i > p \quad (6.13)$$

Utilizando el siguiente cuadro que muestra el valor de la FACP para la serie diferenciada

Cuadro 6.4. Valor de la FACP muestral para la serie diferenciada de las ventas de gasolina

FACP muestral											
1	2	3	4	5	6	7	8	9	10	11	12
-0.509	-0.431	-0.002	-0.138	-0.140	0.003	0.009	-0.156	0.064	0.040	-0.361	0.303
13	14	15	16	17	18	19	20	21	22	23	24
0.048	0.248	-0.091	-0.102	0.019	-0.065	-0.019	0.100	-0.045	-0.115	-0.021	0.041

Y tomando  $N=141$  y  $d=1$ ,  $\hat{\phi}_{ii}$  sería significativamente distinta de cero sólo si

$$|\hat{\phi}_{ii}| > 2/\sqrt{N-d} = 0.1690 \quad (6.14)$$

como se puede observar, la autocorrelación estimada para el retraso 3 no satisface esto, por lo que se rechaza la posibilidad del modelo ARI(1,1)

Además, si el proceso fuese en realidad un ARI(1,1), la FAC muestral debería mostrar convergencia a cero de acuerdo con la ecuación  $(1-\phi B)r_k = 0$ , pero esto no sucede ya que, por ejemplo si  $\phi$  fuese 0.5, entonces se debería tener

Cuadro 6.5. Valores de la FAC calculada para la serie diferenciada de las ventas de gasolina

$r_1=0.5$	$r_2=0.25$	$r_3=0.125$	$r_4=0.063$
	$r_5=0.031$	$r_6=0.016$	

y la discrepancia entre estos valores y los del siguiente cuadro es notable

Cuadro 6.6 Valores de la FAC muestral para la serie diferenciada de las ventas de gasolina

FAC muestral											
1	2	3	4	5	6	7	8	9	10	11	12
-0.509	-0.061	0.263	-0.247	0.044	0.135	-0.132	-0.049	0.209	-0.156	-0.199	0.547
13	14	15	16	17	18	19	20	21	22	23	24
-0.427	0.174	0.019	-0.23	0.225	-0.054	-0.086	0.12	-0.034	-0.099	0.03	0.194

Por lo anterior, es quizás razonable pensar en un modelo mixto ARIMA(1,1,1), e incluso si se observa con mayor cuidado las gráficas presentadas en la página 60 de la FAC y FACP muestrales están presentando valores que permiten suponer que el proceso posee estacionalidad.

Entonces los modelos propuestos a estimar y verificar para las ventas de gasolinas en la Ciudad de México serían:

$$\left. \begin{array}{l}
 \text{ARI}(1,0,1) \\
 \text{ARI}(2,0,1) \text{ (12,0,0)} \\
 \text{ARIMA (1,1,1)} \\
 \text{ARIMA (1,1,1) (12,0,12)}
 \end{array} \right\} (6.15)$$

cuya representación sería, respectivamente

$$\begin{aligned}
 (1 - \phi) \nabla T(\text{VG}) &= a_t \\
 (1 - \phi_1 B - \phi_2 B^2)(1 - \Phi B^{12}) \nabla T(\text{VG}) &= a_t \\
 (1 - \phi) \nabla T(\text{VG}) &= (1 - \theta) a_t \\
 (1 - \phi B)(1 - \Phi B^{12}) \nabla T(\text{VG}) &= (1 - \theta B)(1 - \Theta B^{12}) a_t
 \end{aligned}
 \tag{6.16}$$

## ii. Estimación

Se utilizó el EViews para la estimación de los distintos modelos, el cual emplea técnicas de estimación no lineales para los parámetros y permite obtener no sólo estimaciones puntuales para estos, sino también intervalos de confianza. Este método sugerido por Box y Jenkins<sup>13</sup> y basado en el algoritmo de Marquardt, tiene como fundamento un desarrollo en series de Taylor que linealiza a la serie  $a_t$  definida por

$$a_t = W_t - \phi_1 W_{t-1} - \dots - \phi_p W_{t-p} - \theta_0 + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \tag{6.17}$$

en donde  $W_t = \nabla^d T(Z_t)$  y condicionada en que se conocen los valores muestrales de  $W$  y los valores iniciales de los parámetros  $\phi, \theta_0$  y  $\theta$ . Los valores iniciales se corregirán iterativamente con el objetivo final de minimizar  $S(\phi, \theta_0, \theta)$ ; el proceso iterativo termina cuando se logra la “convergencia”, lo cual sucede por ejemplo, si el cambio relativo en cada parámetro no es mayor que un cierto valor pequeño, fijado de antemano.

### Modelo 1. $(1 - \phi) \nabla T(\text{VG}) = a_t$

El resultado de la estimación fue el siguiente

<sup>13</sup> Box, G.E.P. y Jenkins, G. (1970) “Time Series Analysis, Forecasting and Control”, San Francisco, Holden-Day

Cuadro 6.7 Resultados de la estimación del modelo ARI (1,0,1) de las ventas de gasolina

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.002622	0.002297	1.141453	0.2555
AR(1)	-0.494118	0.067962	-7.270485	0.0000
R-squared	0.258030	Mean dependent var		0.002421
Adjusted R-squared	0.253148	S.D. dependent var		0.049276
S.E. of regression	0.042585	Akaike info criterion		-3.461731
Sum squared resid	0.275648	Schwarz criterion		-3.422290
Log likelihood	268.5533	F-statistic		52.85995
Durbin-Watson stat	2.502854	Prob(F-statistic)		0.000000
Inverted AR Roots	-.49			

por lo que el modelo queda de la siguiente manera

$$\nabla \log VG_t = 0.002622 - 0.4941 \nabla \log VG_{t-1} \quad (6.18)$$

**Modelo 2.**  $(1 - \phi_1 \mathbf{B} - \phi_2 \mathbf{B}^2) (1 - \Phi \mathbf{B}^{12}) \nabla \mathbf{T}(\mathbf{VG}) = \mathbf{a}_t$

El resultado de la estimación para este segundo modelo fue el siguiente

Cuadro 6.8 Resultados de la estimación del modelo ARI (2,0,1) (12,0,0) de las ventas de gasolina

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.002889	0.002811	1.027829	0.3058
AR(1)	-0.848686	0.070423	-12.05124	0.0000
AR(2)	-0.573312	0.069697	-8.225775	0.0000
SAR(12)	0.673540	0.061138	11.01672	0.0000
R-squared	0.720858	Mean dependent var		0.002561
Adjusted R-squared	0.714746	S.D. dependent var		0.049351
S.E. of regression	0.026358	Akaike info criterion		-4.406129
Sum squared resid	0.095180	Schwarz criterion		-4.322476
Log likelihood	314.6321	F-statistic		117.9300
Durbin-Watson stat	1.955458	Prob(F-statistic)		0.000000
Inverted AR Roots	.97	.84 -.48i	.84+.48i	.48+.84i
	.48 -.84i	.00+.97i	-.00 -.97i	-.42+.63i
	-.42 -.63i	-.48+.84i	-.48 -.84i	-.84 -.48i
	-.84+.48i	-.97		

por lo que el modelo queda de la siguiente manera

$$\begin{aligned} \nabla \log VG_t = & 0.00289 - 0.8486 \nabla \log VG_{t-1} - 0.5733 \nabla \log VG_{t-2} \\ & - 0.673540 \nabla \log VG_{t-12} + 0.57162 \nabla \log VG_{t-13} + 0.38614 \nabla \log VG_{t-14} \end{aligned} \quad (6.19)$$

**Modelo 3.**  $(1 - \phi) \nabla T(\mathbf{VG}) = (1 - \theta) \mathbf{a}_t$

El resultado para este tercer modelo fue

Cuadro 6.9 Resultados de la estimación del modelo ARIMA (1,1,1) de las ventas de gasolina

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.002216	0.000743	2.980776	0.0034
AR(1)	-0.183385	0.101833	-1.800846	0.0737
MA(1)	-0.711210	0.074949	-9.489273	0.0000
R-squared	0.441623	Mean dependent var		0.002421
Adjusted R-squared	0.434227	S.D. dependent var		0.049276
S.E. of regression	0.037065	Akaike info criterion		-3.733019
Sum squared resid	0.207442	Schwarz criterion		-3.673857
Log likelihood	290.4424	F-statistic		59.71332
Durbin-Watson stat	1.993819	Prob(F-statistic)		0.000000
Inverted AR Roots	-.18			
Inverted MA Roots	.71			

por lo que el modelo queda de la siguiente manera

$$\nabla \log VG_t = 0.002216 - 0.1833 \nabla \log VG_{t-1} + 0.7112 \mathbf{a}_{t-1} \quad (6.20)$$

**Modelo 4.**  $(1 - \phi \mathbf{B})(1 - \Phi \mathbf{B}^{12}) \nabla T(\mathbf{VG}) = (1 - \theta \mathbf{B})(1 - \Theta \mathbf{B}^{12}) \mathbf{a}_t$

Para este último modelo los resultados obtenidos fueron

Cuadro 6.10 Resultados de la estimación del modelo ARIMA (1,1,1) (12,0,12) de las ventas de gasolina

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.027814	0.031292	0.888861	0.3756
AR(1)	-0.312906	0.074819	-4.182175	0.0001
SAR(12)	0.980524	0.023237	42.19730	0.0000
MA(1)	-0.667273	0.073021	-9.138132	0.0000
SMA(12)	-0.903734	0.028853	-31.32191	0.0000
R-squared	0.735135	Mean dependent var		0.002792
Adjusted R-squared	0.727401	S.D. dependent var		0.049252
S.E. of regression	0.025715	Akaike info criterion		-4.448906
Sum squared resid	0.090593	Schwarz criterion		-4.344828
Log likelihood	320.8723	F-statistic		95.06102
Durbin-Watson stat	2.223091	Prob(F-statistic)		0.000000
Inverted AR Roots	1.00	.86+.50i	.86 -.50i	.50+.86i
	.50 -.86i	.00+1.00i	-.00 -1.00i	-.31
	-.50+.86i	-.50 -.86i	-.86 -.50i	-.86+.50i
	-1.00			
Inverted MA Roots	.99	.86+.50i	.86 -.50i	.67
	.50+.86i	.50 -.86i	.00 -.99i	-.00+.99i
	-.50+.86i	-.50 -.86i	-.86+.50i	-.86 -.50i
	-.99			

por lo que la ecuación queda de la siguiente forma

$$\nabla \log VG_t = 0.002781 - 0.3129 \nabla \log VG_{t-1} + 0.9805 \nabla \log VG_{t-12} + \quad (6.21)$$

$$0.3068 \nabla \log VG_{t-13} + 0.6672 a_{t-1} + 0.9037 a_{t-12} + 0.6030 a_{t-13}$$

### iii. Verificación

La etapa de verificación de la metodología de Box-Jenkins tiene su origen en la idea de que todo modelo es erróneo, puesto que los modelos son representaciones simplificadas de la realidad. Por lo que sí hay que elegir entre varios modelos, habrá que seleccionar aquel que presente menos fallas, o poco importantes; por este motivo habrá que poner todos los modelos en tela de juicio para detectar sus fallas (que se miden como violaciones a los supuestos que fundamentan el modelo)

Primero se analizarán los residuos de cada modelo verificando que estos sean ruido blanco mediante la comprobación de los supuestos:

- ✓ Supuesto 1.  $\{a_t\}$  tiene media cero
- ✓ Supuesto 2.  $\{a_t\}$  tiene varianza constante
- ✓ Supuesto 3.  $\{a_t\}$  son mutuamente independientes
- ✓ Supuesto 4.  $\{a_t\}$  tiene una distribución normal

y posteriormente se compararán los siguientes estadísticos:

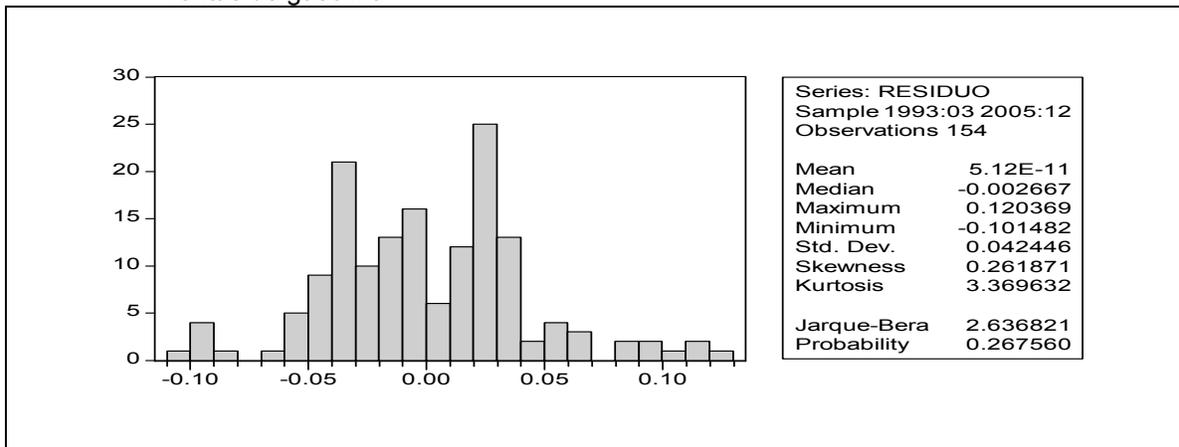
- significancia de parámetros (t), esto es, se desecha la hipótesis de que los coeficientes individuales son iguales a cero.
- $R^2$ , bondad del ajuste.
- estadístico Q de Box y Pierce, el menor valor del estadístico muestra que el conjunto de residuos del modelo es un proceso estocástico de ruido blanco.
- Criterio de Akaike, el menor valor como una prueba de minimización de la varianza de la regresión y de parsimonia, en forma conjunta.
- Criterio de Schwartz, igual que el anterior.
- Durbin-Watson
- error estándar de estimación, el menor valor como una prueba práctica de parsimonia.

### **Análisis de residuos del Modelo 1**

Supuesto 1.

Se obtuvieron los siguientes estadísticos descriptivos para los residuales del primer modelo

Cuadro 6.11 Estadísticas descriptivas e histograma del residuo del modelo ARI (1,0,1) de las ventas de gasolina



Utilizando el cociente sugerido por Guerrero  $\frac{\sqrt{N-d-p} m(a)}{\sigma_a}$  y recordando que si el valor absoluto de dicho cociente es menor que dos se dirá que no hay evidencia de que la media del proceso de ruido blanco sea distinta de cero.

Entonces para este primer modelo el estimador propuesto por Guerrero es:

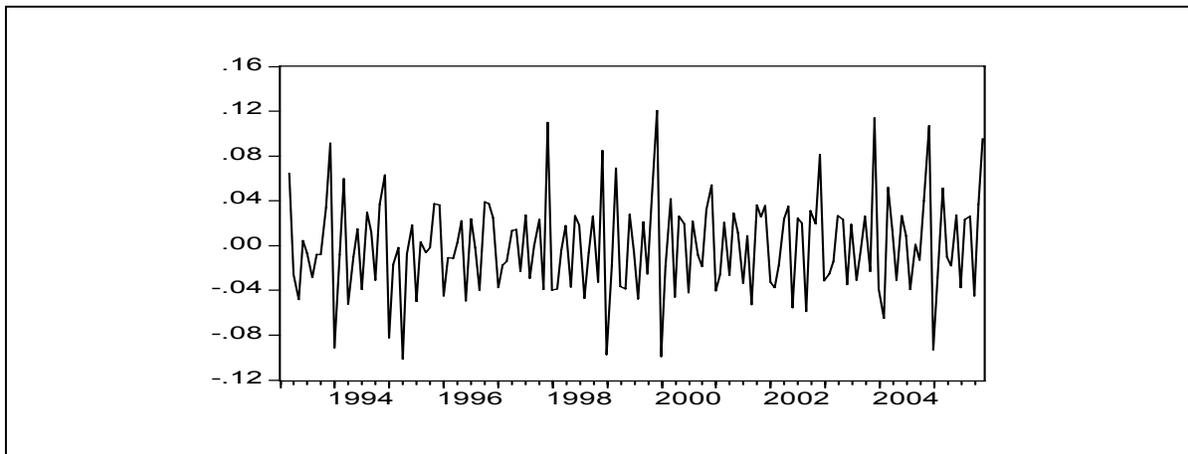
$$\left| \frac{\sqrt{N-d-p} m(a)}{\sigma_a} \right| = \left| \frac{\sqrt{154-1-1} \cdot 0.00}{0.424} \right| = 0.00 < 2 \quad (6.22)$$

como el cociente es menor que 2, se concluye que la media del proceso es cero.

Supuesto 2.

La gráfica muestra, visualmente, que la varianza de  $a_t$  que permanece constante a través del tiempo

Gráfica 6.4. Comportamiento del residuo del modelo ARI(1,0,1) de las ventas de gasolina



Supuesto 3.

Debido a que la independencia implica no autocorrelación se requiere que  $\rho_k(a)=0$  para toda  $k \neq 0$ .

Para probar esta hipótesis se muestra el estadístico Q de Ljung y Box y su probabilidad, para diferentes retrasos.

Cuadro 6.12 Estadístico Q de Ljung y Box de los residuos del modelo ARI (1,0,1)

	1	4	8	12	16	20	24	28	32	36
Q-Stat	11.914	47.505	62.639	158.2	187.38	205.36	236.87	248.05	263.98	276.96
Prob	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

dado que  $p < 0.05$  en todos los casos, se rechaza la hipótesis nula de no autocorrelación.

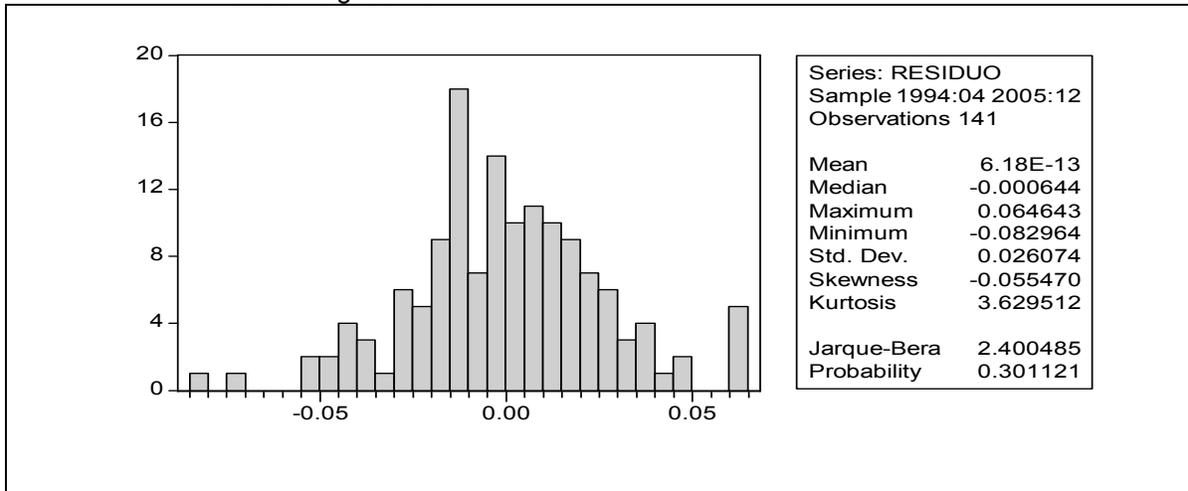
Supuesto 4.

En el Cuadro 6.11 de la página 63 se observa que el estadístico Jarque – Bera tiene un valor de probabilidad  $p = 0.267 > 0.05$ , por lo que se acepta la hipótesis nula de que los errores se distribuyen en forma normal.

## Análisis de residuales del Modelo 2.

Supuesto 1.

Cuadro 6.13 Estadísticas descriptivas e histograma del residuo del modelo ARI (2,0,1) (12,0,0) de las ventas de gasolina



Aplicando el cociente sugerido por Guerrero, se tiene que:

$$\left| \frac{\sqrt{N-d-p} \cdot m(a)}{\sigma_a} \right| = \left| \frac{\sqrt{141-1-1} \cdot (0.00/0.026)}{0.026} \right| = 0.00 < 2$$

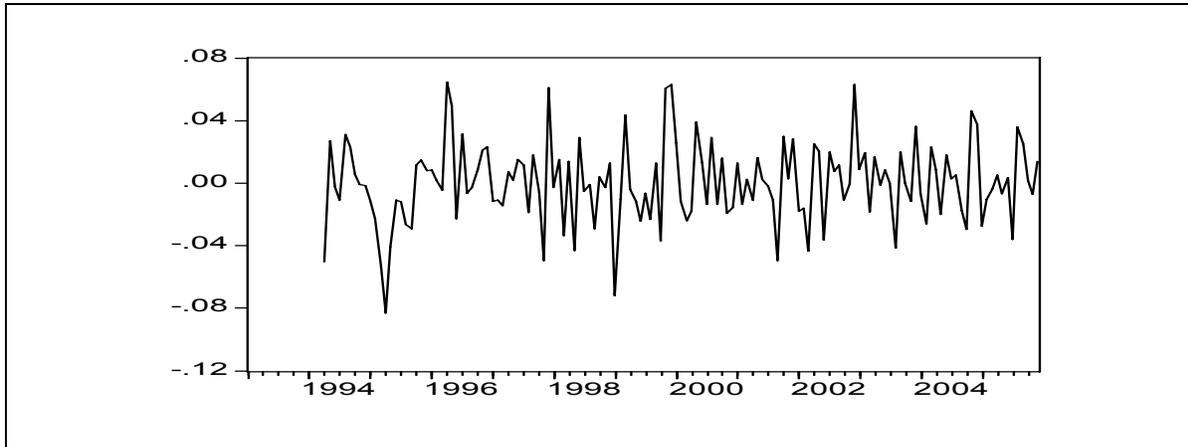
(6.23)

No hay razón para pensar que la media de los residuos sea distinta de cero.

Supuesto 2.

La gráfica de la varianza de  $a_t$  muestra visualmente que permanece constante a través del tiempo

Gráfica 6.5. Comportamiento del residuo del modelo ARI (2,0,1) (12,0,0) de las ventas de gasolina



Supuesto 3.

Cuadro 6.14 Estadístico Q de Ljung y Box de los residuos del modelo ARI (2,0,1) (12,0,0)

	1	4	8	12	16	20	24	28	32	36
Q-Stat	0.0091	0.8145	3.8949	12.535	23.12	25.991	47.719	53.063	56.702	64.472
Prob	0.924	0.936	0.866	0.404	0.111	0.166	0.003	0.003	0.005	0.002

El valor de la probabilidad del estadístico Q para todos los rezagos permite asegurar la no-existencia de autocorrelación en estos residuos.

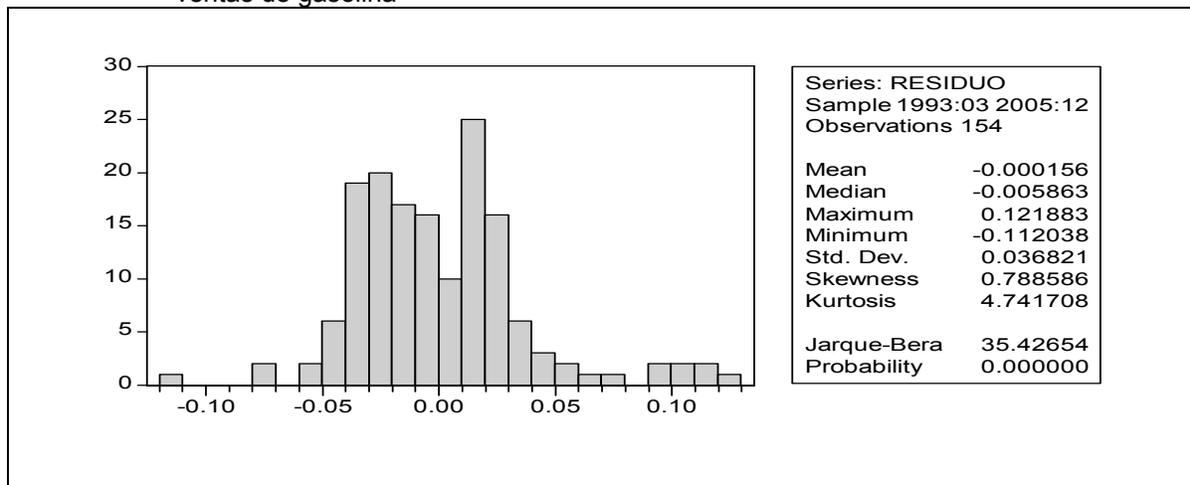
Supuesto 4.

En el Cuadro 6.13 de la página 65 el valor de la probabilidad del estadístico Jarque – Bera es igual a  $0.175 > 0.05$  por lo que se acepta la hipótesis nula de que los errores se distribuyen normalmente.

### Análisis de residuales del Modelo 3.

Supuesto 1.

Cuadro 6.15 Estadísticas descriptivas e histograma del residuo del modelo ARIMA (1,1,1) de las ventas de gasolina



Aplicando el cociente sugerido por Guerrero, se tiene que:

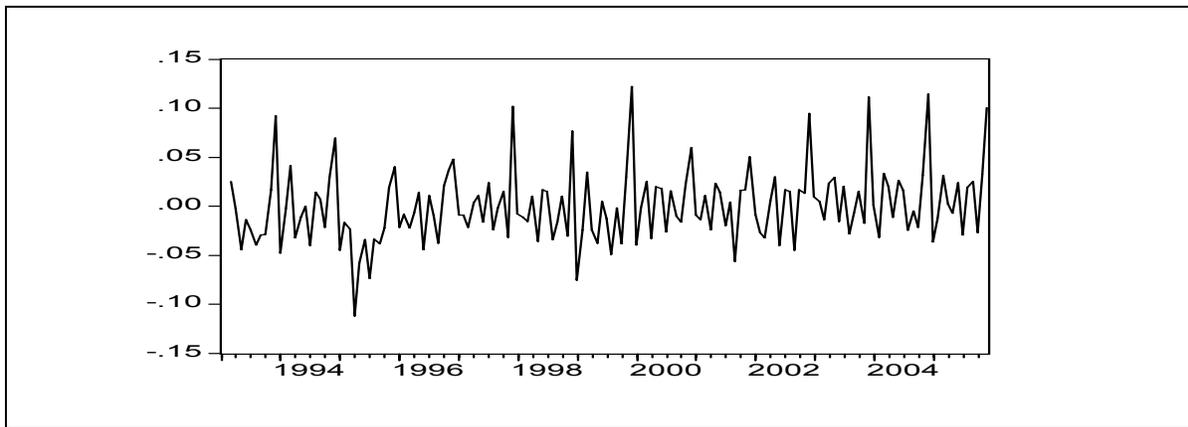
$$\left| \frac{\sqrt{N-d-p} \cdot m(a)}{\sigma_a} \right| = \left| \frac{\sqrt{154-1-1} \cdot (-0.000156)}{0.0368} \right| = 0.052 < 2 \quad (6.24)$$

No hay razón para pensar que la media de los residuos sea distinta de cero.

Supuesto 2.

La gráfica de la varianza de  $a_t$  muestra visualmente que permanece constante a través del tiempo

Gráfica 6.6. Comportamiento del residuo del modelo ARIMA (1,1,1) de las ventas de gasolina



Supuesto 3.

Dado que  $p < 0.05$  en todos los casos (excepto en el primero), se rechaza la hipótesis nula de no autocorrelación

Cuadro 6.16 Estadístico Q de Ljung y Box de los residuos del modelo ARIMA (1,1,1)

	1	4	8	12	16	20	24	28	32	36
Q-Stat	0.082	11.889	19.033	87.772	102.43	109.35	142.84	150.8	158.22	178.04
Prob	0.775	0.018	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.000

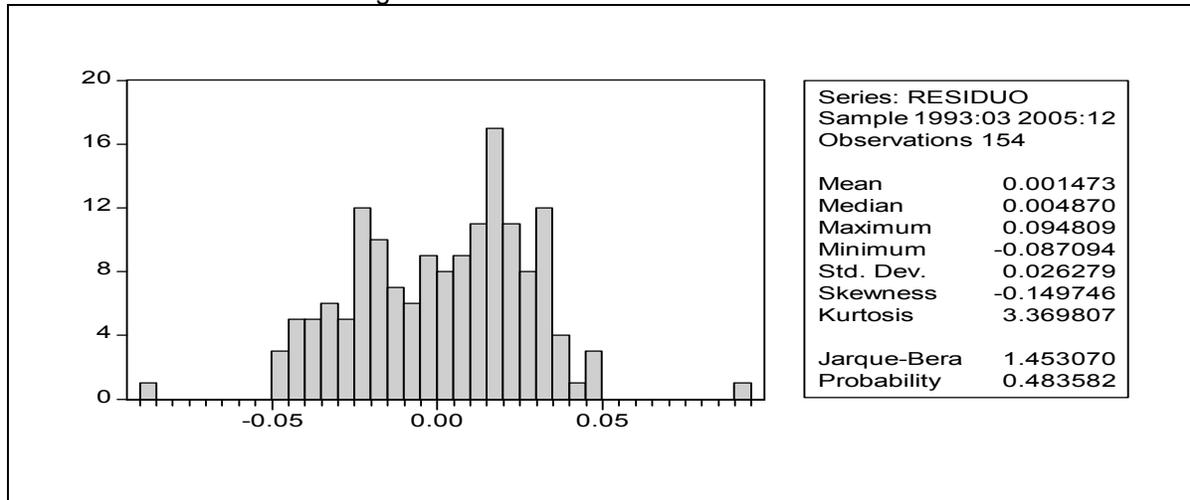
Supuesto 4.

El valor de la probabilidad del estadístico Jarque – Bera en el Cuadro 6.15 de la página 67 ( $p = 0$ ), permite rechazar la hipótesis de que los residuos se distribuyen normalmente.

#### **Análisis de residuales del Modelo 4.**

## Supuesto 1.

Cuadro 6.17 Estadísticas descriptivas e histograma del residuo del modelo ARIMA (1,1,1) (12,0,12) de las ventas de gasolina



Aplicando el cociente sugerido por Guerrero, se tiene que:

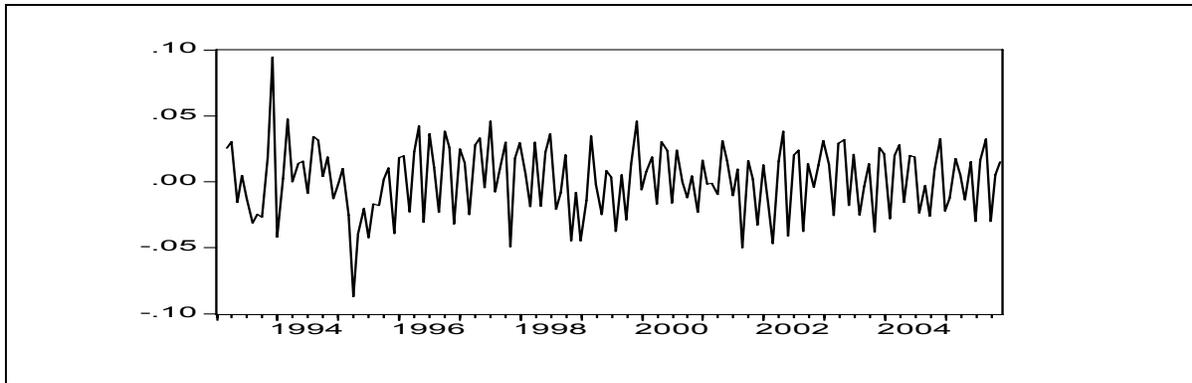
$$\left| \sqrt{N-d-p} \frac{m(a)}{\sigma_a} \right| = \left| \sqrt{154-1-1} \frac{0.001473}{0.02628} \right| = 0.691 < 2 \quad (6.25)$$

No hay razón para pensar que la media de los residuos sea distinta de cero.

## Supuesto 2.

La gráfica de la varianza de  $a_t$  muestra visualmente que permanece constante a través del tiempo

Gráfica 6.7. Comportamiento del residuo del modelo ARIMA (1,1,1) (12,0,12) de las ventas de gasolina



### Supuesto 3.

Cuadro 6.18 Estadístico Q de Ljung y Box de los residuos del modelo ARIMA (1,1,1) (12,0,12)

	1	4	8	12	16	20	24	28	32	36
Q-Stat	2.1345	44.138	75.017	116.54	184.91	216.88	273.61	305.12	340.92	391.67
Prob	0.144	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

en todos los casos, se rechaza la hipótesis nula de no autocorrelación

### Supuesto 4.

El valor de la probabilidad del estadístico Jarque – Bera del Cuadro 6.17 de la página 69, permite no rechazar la hipótesis de que los residuos se distribuyen normalmente.

### Batería de pruebas de los modelos

Cuadro 6.19 Resumen de pruebas de los modelos ARIMA

<b>Modelo</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
	probabilidad de t-estadístico			
constante	0.2555	0.3058	0.0034	0.3756
AR(1)	0.0000	0.0000	0.0737	0.0001
AR(2)		0.0000	0.0000	
SAR(12)		0.0000		0.0000
MA(1)				0.0000
SMA(12)				0.0000
residuales = ruido blanco	NO	SI	NO	NO
R cuadrada	0.2580	0.7208	0.4416	0.7351
QB -P (36)	276.96	64.472	178.04	391.67
Criterio de Akaike	-3.4617	-4.4061	-3.7330	-4.4489
Criterio de Schwartz	-3.4222	-4.3224	-3.6738	-4.3448
Durbin-Watson	2.5028	1.955	1.993	2.223
Error estadístico de estimación	0.0425	0.0263	0.0370	0.0257
Suma de residuales al cuadrado	0.2756	0.0951	0.2074	0.0905

Existen otros supuestos relacionados con el modelo (parsimonioso, admisible y estable) los cuales se prueban con la batería de pruebas del Cuadro 6.19 y se escoge aquel que presenta los menores valores de los estadísticos. De acuerdo con dicho cuadro, el modelo 2 es el más adecuado para representar a las ventas de gasolina de la Ciudad de México.

## **b. Análisis de intervención**

La técnica del análisis de intervención que aquí se presenta está basada en el artículo de Box y Tiao<sup>14</sup> en el cual aparece, prácticamente por vez primera, esta metodología de análisis. Una intervención puede ser interpretada como la ocurrencia de un evento exógeno al comportamiento histórico de la variable en estudio; tal intervención podría ser, como en el caso que nos ocupa, un cambio en las políticas de distribución de petrolíferos por autotank de las terminales de

<sup>14</sup> Box, G.E.P. y Tiao, G.C. (1975) "Intervention analysis with applications to economic and environmental problems", en *Journal of the American Statistical Association* 70,70

almacenamiento y distribución de Petróleos Mexicanos a las gasolineras, lo cual de acuerdo a los resultados obtenidos en las secciones anteriores, esto generó un impacto en las ventas de gasolinas a principios de 2004.

Si se piensa que una serie de tiempo estacionaria  $\{W_t\}$  está constituida por dos partes, una determinista que refleja esencialmente el nivel de la serie ( $\mu_w$ ) y otra estocástica que da origen a las fluctuaciones alrededor de dicho nivel, podrían distinguirse tres tipos de efectos:

- i. los que se dejan sentir como una elevación o caída momentánea del nivel y que desaparecen sin influir en el comportamiento posterior de la serie.
- ii. los que ejercen una influencia sostenida (no momentánea) sobre el nivel de la serie, pero dejan intacta la estructura básica de su parte estocástica.
- iii. los efectos que, independientemente de influir o no sobre la parte determinista, si alteran la estructura de la parte estocástica

Los dos primeros son los que pueden estudiarse mediante el análisis de intervención.

El análisis de intervención puede verse como una extensión de los procesos ARIMA, ya que un modelo para una serie  $\{T(Z_t)\}$  que contenga los efectos de una intervención, se expresa como

$$T(Z_t) = \varepsilon_{I,t} + N_t \quad (6.26)$$

donde  $N_t$  se asocia con un modelo ARIMA, estacionario e invertible, que representa a la parte estocástica de la serie y el cual puede expresarse como

$$\phi(B)\nabla^d N_t = \theta_0 + \theta(B)a_t \quad (6.27)$$

con  $N_t$  medida como desviaciones respecto a su media si  $d=0$ , mientras que  $\varepsilon_{I,t}$  es una función que permite representar los efectos de la intervención.

### b.1 Función de respuesta a impulsos

La función  $\varepsilon_{I,t}$  se especifica en términos de ecuaciones en diferencia y se utiliza la función de pulso definida por

$$P_{I,t} = \begin{cases} 1 & \text{si } t=I \\ 0 & \text{si } t \neq I \end{cases} \quad (6.28)$$

en donde  $I$  denota el momento en el que ocurrió la intervención. Dada esta función, que de hecho define a una serie de tiempo que toma el valor 1 en el momento  $t=I$  y 0 en cualquier otro momento, entonces es razonable suponer que  $\varepsilon_{I,t}$  puede satisfacer una ecuación del tipo

$$(1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r) \varepsilon_{I,t} = (\omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s) P_{I,t} \quad (6.29)$$

cuya solución específica dependerá de la condición  $\varepsilon_{I,t} = 0$  para  $t < 1$ , lo que indica que antes de haber ocurrido la intervención no existen efectos atribuibles a ella.

A partir de este modelo puede postularse todavía un modelo dinámico más general, en el cual se consideren raíces unitarias para el polinomio de retraso que opera sobre  $\varepsilon_{I,t}$ , de esta forma se obtiene

$$\delta(B) \nabla^b \varepsilon_{I,t} = \omega(B) P_{I,t} \quad (6.30)$$

con

$$\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r \quad (6.31)$$

y

$$\omega(B) = \omega_0 - \omega_1 B - \omega_2 B^2 - \dots - \omega_s B^s \quad (6.32)$$

Por consiguiente, el modelo ARIMA mencionado en los párrafos anteriores, para la serie que incluye a la intervención, se convierte en

$$T(Z_t) = \frac{\omega(B)}{\delta(B)\nabla^b} P_{I,t} + \frac{\theta_0 + \theta(B)a_t}{\phi(B)\nabla^d} \quad (6.33)$$

Guerrero sugiere la función

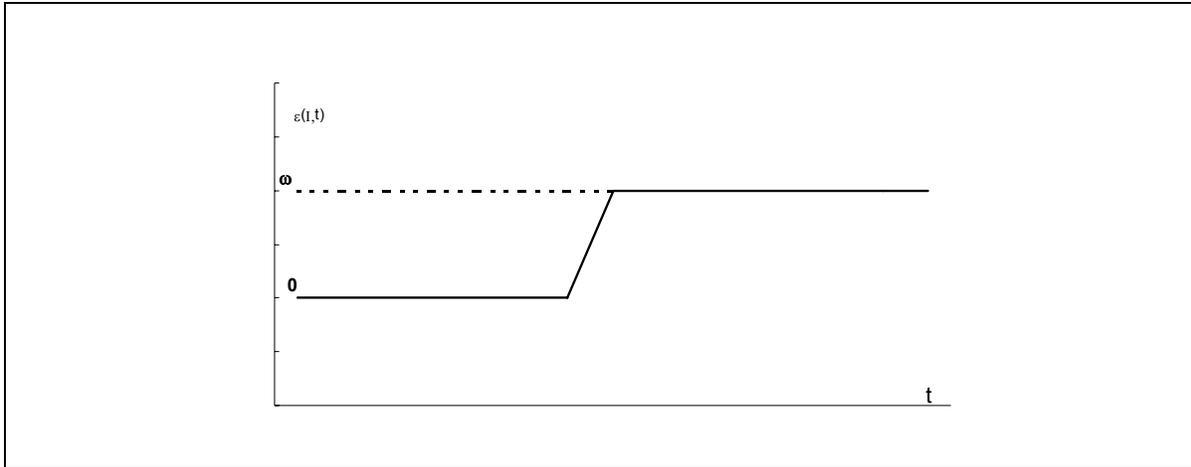
$$\nabla \varepsilon_{I,t} = \omega P_{I,t} \quad (6.34)$$

en donde

$$\varepsilon_{I,t} = \begin{cases} 0 & \text{si } t < I \\ \omega & \text{si } t \geq I \end{cases} \quad (6.35)$$

para el caso de un cambio de nivel en la serie, como el que se ilustra a continuación

Gráfica 6.8. Gráfica de la función de impulso  $\varepsilon_{I,t}$



## b.2 El estadístico C

En muchas ocasiones puede surgir la duda de si una cierta intervención ejerció o no una influencia considerable sobre la serie en estudio; para aclarar esta duda es factible utilizar un estadístico de prueba diseñado por Box y Tiao mediante el cual pueden detectarse cambios en el modelo, atribuibles a la intervención. Este estadístico se expresa como

$$C = \frac{\sum_{h=1}^H e_{I-2+h}^2}{\sigma_a^2} \quad (6.36)$$

en donde  $\sigma_a^2$  es la varianza residual del modelo construido para el periodo previo a la intervención ocurrida en el momento  $t=I$  y  $e_{I-1}, e_I, \dots, e_{I+H-2}$  son los errores del pronóstico un periodo hacia delante.

Por lo que una prueba de que el modelo previamente construido continua siendo válido para periodos posteriores a la intervención, se obtiene al comparar C con valores de tablas de una distribución ji cuadrada con H grados de libertad; si el modelo no es del todo adecuado para representar la serie hasta la observación  $t=1+H-1$ , el valor de C tenderá a ser muy grande. Si eso ocurre, se tendrá

evidencia suficiente para concluir que la intervención ocurrida si afectó de alguna manera a la serie y, por consiguiente, podría procederse a postular un modelo de intervención.

### Caso 1. Cambio en el nivel de la serie a partir de enero de 2004

En el apartado anterior se encontró que el modelo

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \Phi B^{12}) \nabla T(VG) = a_t \quad (6.37)$$

era el más adecuado para representar la parte estocástica de la serie y utilizando la función de impulso sugerida por Guerrero, el modelo para el análisis de intervención sería

$$T(Z_t) = \frac{\omega P_{I,t}}{(1 - \gamma_1 B)} + \frac{a_t}{(1 - \phi_1 B - \phi_2 B^2)(1 - \phi_{12} B^{12}) \nabla} \quad (6.38)$$

en donde

$$P_{I,t} = \begin{cases} 1 & \text{si } t \geq \text{enero 2004} \\ 0 & \text{si } t < \text{enero 2004} \end{cases} \quad (6.39)$$

la estimación del modelo arrojó los siguientes resultados para los parámetros

Cuadro 6.20. Parámetros del modelo estimado

	coeficiente	Significancia (al 5%)
$a_t$	0.002630	0.4175
$\phi_1$	-0.84395	0.0000
$\phi_2$	-0.552107	0.0000
$\phi_{12}$	0.674051	0.0000
$\omega$	-0.013385	0.4145
$\gamma_1$	0.001209	0.9403

en donde se puede observar que los parámetros de la función de intervención no tienen significancia, por lo que en un principio se puede decir que no hay un cambio en el nivel de la serie a partir de enero de 2004.

Utilizando el estadístico C de Box y Tiao, se tiene que su valor es mayor que  $\chi^2(13) = 23.68$

$$C = \frac{\sum_{h=1}^H e_{I-2+h}^2}{\hat{\sigma}_a^2} = \frac{0.082341}{(0.026)^2} = 63.279 \quad (6.40)$$

lo que permite concluir que como C tiende a ser muy grande, hay evidencia suficiente para concluir que la intervención ocurrida si afectó de alguna manera a la serie.

---

## Sección 7. Estimación de un modelo con técnicas de cointegración.

### a. Orden de integración de las series

Como mencionan Holden y Perman<sup>15</sup> la modelación econométrica de las series de tiempo usando una especificación dinámica, la cual incorpora un mecanismo de corrección de error ha resultado ser popular y conveniente y existen un buen número de alternativas para aquellos interesados en estimar relaciones económicas de largo plazo.

Los autores mencionan dos de especial interés: el análisis de vectores - autorregresivos, el cual concede un papel limitado a la teoría en la que se basa el modelo y hace énfasis en la importancia de seleccionar un modelo basado en criterios emanados de los datos y el modelo econométrico estructural, el cual destaca el papel de la teoría económica en el diseño y especificación del modelo econométrico. El análisis de cointegración, afirman Holden y Perman; puede verse como una reconciliación de esas dos alternativas, dado que la existencia de una o más relaciones de cointegración entre un conjunto de variables implica que existen restricciones que conectan los parámetros en el vector de autorregresión.

Por lo anterior, para establecer un modelo estructural de la demanda de gasolinas para la Ciudad de México, se aplica en este capítulo la metodología sugerida por Holden y Perman para cointegrar un conjunto de variables, concentrándose en el uso del procedimiento de Johansen para identificar el número de cointegración de las relaciones del conjunto de variables y estimar los parámetros de largo plazo de estas relaciones.

Denotando como:

$lvg$  = logaritmo natural de las ventas de gasolina de la Ciudad de México

$lipib$  = logaritmo natural del Producto Interno Bruto de la Ciudad de México

---

<sup>15</sup> Holden, D. y Perman, R., "Unit Roots and Cointegration for the Economist, Cointegration for the Applied Economist" editado por B. Bhaskara Rao, Martin Press NY, 1995

$lpg$  = logaritmo natural de el precio de la gasolina en la Ciudad de México

y como:

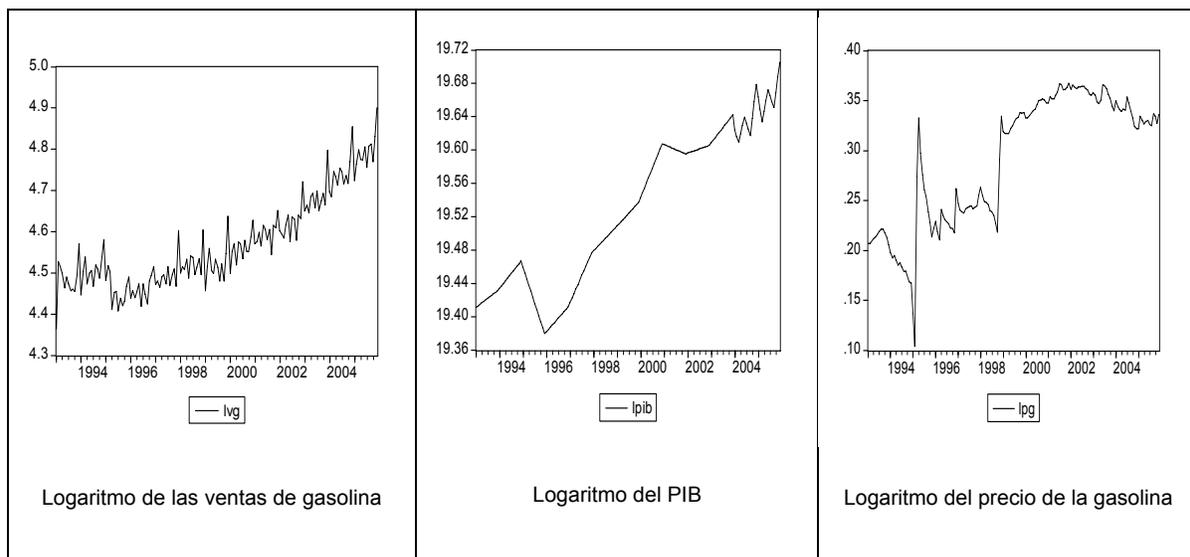
$\Delta lvg$  = primera diferencia de  $lvg$

$\Delta lpib$  = primera diferencia de  $lpib$

$\Delta lpg$  = primera diferencia de  $lpg$

Una revisión visual de las gráficas del logaritmo de las series sugiere una tendencia lineal y dado que presentan una media no constante podría pensarse que son series no estacionarias.

Gráfica 7.1. Evolución del logaritmo de las ventas de gasolina, PIB y precio de la gasolina de la Ciudad de México.



Sin embargo, se debe averiguar si se trata de tendencias estocásticas o simplemente determinísticas, por lo que antes de realizar las pruebas de raíz unitaria es conveniente inspeccionar los correlogramas de las series.

Cuadro 7.1. Correlograma del logaritmo de las ventas de gasolina (lvj)

	AC	PAC	Q-Stat	Prob
1	0.857	0.857	116.71	0.000
2	0.843	0.411	230.53	0.000
3	0.857	0.359	348.83	0.000
4	0.803	-0.021	453.46	0.000
5	0.796	0.059	557.01	0.000
6	0.793	0.073	660.43	0.000
7	0.746	-0.083	752.41	0.000
8	0.737	-0.004	842.97	0.000
9	0.734	0.058	933.43	0.000
10	0.675	-0.142	1010.4	0.000
11	0.671	-0.005	1087.0	0.000
12	0.719	0.335	1175.4	0.000
13	0.612	-0.336	1240.0	0.000
14	0.617	-0.022	1306.1	0.000
15	0.597	-0.120	1368.4	0.000

Cuadro 7.2. Correlograma del logaritmo del PIB (lpib)

	AC	PAC	Q-Stat	Prob
1	0.980	0.980	152.76	0.000
2	0.960	-0.012	300.32	0.000
3	0.942	0.039	443.32	0.000
4	0.928	0.094	583.02	0.000
5	0.913	-0.028	719.18	0.000
6	0.896	-0.058	851.15	0.000
7	0.875	-0.101	977.83	0.000
8	0.854	-0.012	1099.4	0.000
9	0.836	0.019	1216.4	0.000
10	0.820	0.068	1330.0	0.000
11	0.804	-0.030	1439.8	0.000
12	0.785	-0.047	1545.3	0.000
13	0.762	-0.087	1645.4	0.000
14	0.742	0.021	1740.9	0.000
15	0.724	0.033	1832.6	0.000

Cuadro 7.3. Correlograma del logaritmo del precio de la gasolina (lpg)

	AC	PAC	Q-Stat	Prob
1	0.954	0.954	144.84	0.000
2	0.900	-0.120	274.51	0.000
3	0.870	0.260	396.56	0.000
4	0.848	0.008	513.23	0.000
5	0.828	0.085	625.27	0.000
6	0.812	0.046	733.65	0.000
7	0.799	0.059	839.29	0.000
8	0.790	0.066	943.23	0.000
9	0.782	0.037	1045.8	0.000
10	0.771	-0.003	1146.0	0.000
11	0.758	0.015	1243.6	0.000
12	0.749	0.043	1339.5	0.000
13	0.733	-0.076	1432.2	0.000
14	0.708	-0.080	1519.4	0.000
15	0.690	0.049	1602.5	0.000

Como se aprecia en los cuadros anteriores, las correlaciones caen muy lentamente en el tiempo (el estadígrafo de Ljung – Box es siempre altamente significativo). Por otra parte, en todas las series, el primer coeficiente de autocorrelación parcial es cercano a 1, siendo los coeficientes restantes cercanos a cero. Esto es indicativo de la potencial presencia de una raíz unitaria.

El próximo paso consiste en llevar a cabo la prueba de raíces unitarias para cada serie, a fin de verificar rigurosamente las conjeturas anteriores. Los resultados obtenidos, a partir de la prueba de Dickey – Fuller Aumentada, indican que en todos los casos no se rechaza  $H_0$  (presencia de una raíz unitaria) al 95 por ciento de confianza. El siguiente cuadro muestra como el valor del estadístico ADF resulta menor que el valor crítico por lo que no se rechaza la presencia de una raíz unitaria en el logaritmo de las ventas de gasolina. PIB y precio de la gasolina.

Cuadro 7.4. Prueba de raíces unitarias para el logaritmo de las variables

variable	ADF	valor crítico (5%)	H <sub>0</sub>
lv <sub>g</sub>	1.8134 2 rezagos	-1.9429	no se rechaza
lpib	1.8295 7 rezagos	-1.9429	no se rechaza
lp <sub>g</sub>	0.3557 2 rezagos	-1.9429	no se rechaza

A partir de estos resultados se puede establecer que las series no son  $I(0)$ , por lo que se debe demostrar ahora que las series necesitan ser diferenciadas sólo una vez para lograr que sean estacionarias.

Al aplicar la prueba de Dickey – Fuller Aumentada a la primera diferencia de las series lv<sub>g</sub> y lpib y lp<sub>g</sub> se obtiene los siguientes resultados

Cuadro 7.5. Prueba de raíces unitarias para la primera diferencia del logaritmo de las variables

variable	ADF	valor crítico (5%)	H <sub>0</sub>
dlv <sub>g</sub>	-2.3749 12 rezagos	-1.9430	se rechaza
dlpib	-2.5868 3 rezagos	-1.9429	se rechaza
dlp <sub>g</sub>	-11.6143 1 rezago	-1.9429	se rechaza

Del análisis anterior se puede concluir que todas las series son  $I(1)$ . La pregunta que surge ahora es si todas las series están o no cointegradas, es decir, si existe o no una combinación lineal de las variables que sea  $I(0)$ . Para contrastar esta hipótesis, se seguirán dos procedimientos el de Engle – Granger y el de Johansen.

## b. Cointegración de las series y modelo de corrección de errores

La prueba de Engle – Granger exige que se escoja una variable “dependiente” en la relación de cointegración. Aunque los resultados de esta prueba no son inmunes a tal elección, para este caso se tiene que la variable dependiente son las ventas de gasolina. Para estimar la posible relación de cointegración se corrió una regresión considerando a lvg como variable dependiente en función de una constante, lpiib & lpg.

Cuadro 7.6. Resultados de la regresión de lvg contra c, lpiib & lpg.

Dependent Variable: LVG Method: Least Squares				
Sample: 1993:01 2005:12 Included observations: 156				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-24.00307	1.428319	-16.80512	0.0000
LPIB	1.473358	0.074444	19.79159	0.0000
LPG	-0.672730	0.104658	-6.427904	0.0000
R-squared	0.827716	Mean dependent var	4.571615	
Adjusted R-squared	0.825464	S.D. dependent var	0.111207	
S.E. of regression	0.046459	Akaike info criterion	-3.281431	
Sum squared resid	0.330247	Schwarz criterion	-3.222780	
Log likelihood	258.9516	F-statistic	367.5353	
Durbin-Watson stat	1.238267	Prob(F-statistic)	0.000000	

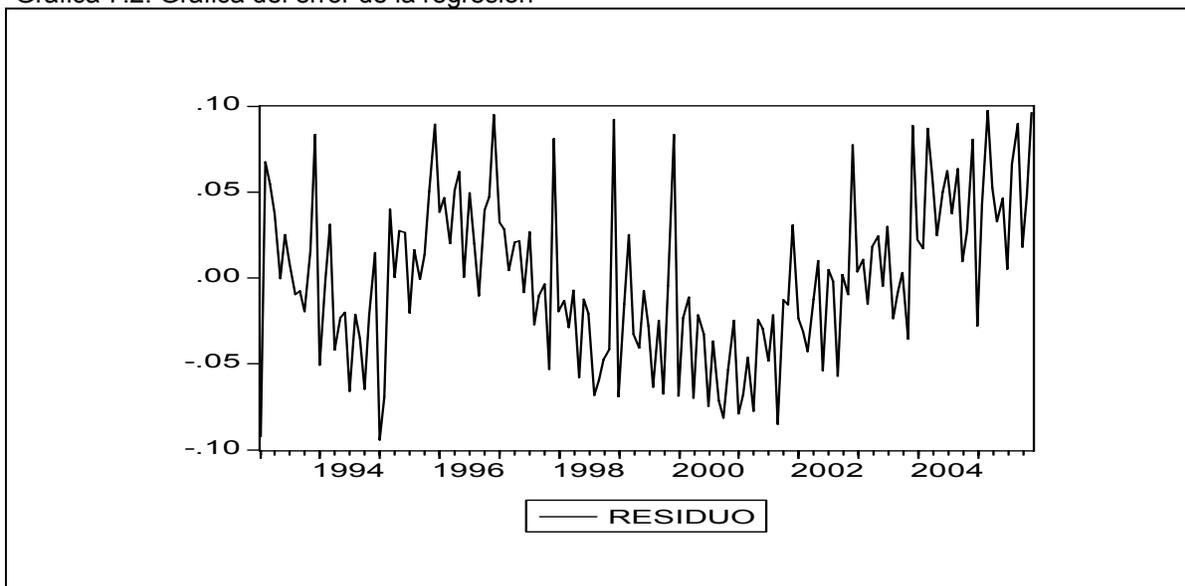
Con los resultados obtenidos se realizó la prueba de Dickey – Fuller para el error de la regresión, cuyos valores se muestran en el siguiente cuadro.

Cuadro 7.7. Prueba de Dickey – Fuller para el error de la regresión

variable	ADF	valor crítico (5%)	H <sub>0</sub>
Error de la regresión	-5.00321 1 rezago	-2.8802	se rechaza

El valor obtenido para el estadístico Dickey – Fuller Aumentado permite rechazar  $H_0$ , por lo que se puede concluir que las variables si están cointegradas, esto es que el error es  $I(0)$  y tal como lo muestra la siguiente gráfica, el residuo muestra una dinámica estacionaria. Por lo tanto, en base a la prueba anterior  $lv_g$  &  $lpib$ ,  $lpg$  están cointegradas.

Gráfica 7.2. Gráfica del error de la regresión



Para complementar la prueba anterior se llevó a cabo la de Johansen, la cual arrojó los siguientes resultados

Cuadro 7.8. Prueba de Johansen para  $lv_g$ ,  $lpib$ ,  $lpg$

Unrestricted Cointegration Rank Test				
Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	5 Percent Critical Value	1 Percent Critical Value
None **	0.211398	50.60797	34.91	41.07
At most 1	0.056103	14.03392	19.96	24.60
At most 2	0.032840	5.142272	9.24	12.97

(\*\*) denotes rejection of the hypothesis at the 5%(1%) level  
Trace test indicates 1 cointegrating equation(s) at both 5% and 1% levels

En el Cuadro 7.8 se puede observar que existe al menos una relación de cointegración para las variables consideradas y de acuerdo a los coeficientes normalizados del Cuadro 7.9.

Cuadro 7.9. Vector de cointegración para lvg, lpib, lpg

1 Cointegrating Equation(s):		Log likelihood	1282.703
Normalized cointegrating coefficients (std.err. in parentheses)			
LVG	LPIB	LPG	C
1.000000	-1.901070 (0.14664)	1.367764 (0.20434)	32.15210 (2.81270)

la relación de cointegración estaría dada por:

$$lvg = -32.15 + 1.90 lpib - 1.36 lpg \quad (4.1)$$

en donde se observa que la elasticidad ingreso de la demanda para el largo plazo es 1.90 mientras que la elasticidad precio de largo plazo es  $-1.36$ .

El modelo de corrección de errores que se muestra en el Cuadro 7.10, describe la dinámica de corto plazo de cada variable del sistema.

El coeficiente asociado a  $u_{t-1}$  en la ecuación de lvg indica que alrededor de 0.08 de la discrepancia entre el logaritmo de las ventas de gasolina actuales y de largo plazo se elimina cada mes. En el caso del PIB, la discrepancia entre el logaritmo de las ventas de gasolina actuales y de largo plazo conduciría a un aumento de 0.012 en el PIB presente. Por el contrario, para el precio de la gasolina la discrepancia entre las ventas actuales de gasolina y las de largo plazo disminuirían en 0.112 dicho precio.

Cuadro 7.10. Modelo de corrección de errores para lvg, lpib, lpg

Cointegrating Eq:		CointEq1		
LVG(-1)		1.000000		
LPIB(-1)		-2.057839 (0.23071) [-8.91957]		
LPG(-1)		1.618867 (0.32448) [ 4.98918]		
C		35.14140		
Error Correction:	D(LVG)	D(LPIB)	D(LPG)	
CointEq1	-0.083604 (0.06585) [-1.26962]	0.012408 (0.00948) [ 1.30925]	-0.112277 (0.03072) [-3.65441]	
D(LVG(-1))	-0.774852 (0.08494) [-9.12210]	-0.016023 (0.01222) [-1.31070]	0.074010 (0.03963) [ 1.86745]	
D(LVG(-2))	-0.477217 (0.07246) [-6.58631]	0.000239 (0.01043) [ 0.02292]	0.000830 (0.03381) [ 0.02456]	
D(LPIB(-1))	1.126710 (0.56363) [ 1.99901]	0.697792 (0.08112) [ 8.60218]	-0.172681 (0.26298) [-0.65664]	
D(LPIB(-2))	-0.148928 (0.59338) [-0.25098]	-0.250639 (0.08540) [-2.93491]	-0.244462 (0.27685) [-0.88300]	
D(LPG(-1))	-0.240786 (0.16657) [-1.44551]	-0.013777 (0.02397) [-0.57470]	0.243473 (0.07772) [ 3.13274]	
D(LPG(-2))	-0.191824 (0.17196) [-1.11548]	-0.018657 (0.02475) [-0.75386]	-0.254746 (0.08023) [-3.17505]	
C	0.003895 (0.00305) [ 1.27652]	0.001129 (0.00044) [ 2.57219]	0.001425 (0.00142) [ 1.00122]	
R-cuadrada	0.505287	0.347641	0.214215	
r-cuadrada ajustada	0.481404	0.316148	0.176280	
suma cuadr resid	0.183690	0.003805	0.039987	
error std ecuación	0.035593	0.005122	0.016606	
estadístico -F	21.15701	11.03860	5.646966	

El análisis de cointegración realizado en primer lugar mediante el procedimiento de Engle – Granger y posteriormente por el método de Johansen permiten evidenciar la existencia de una relación específica de largo plazo entre  $lvg$ ,  $lpib$  &  $lpg$  .

Los signos obtenidos en el vector de cointegración resultan acordes con la teoría económica subyacente; por lo que respecta a sus valores, ambas elasticidades de largo plazo presentan valores ligeramente superiores a los obtenidos tanto en este estudio como en los otros analizados.

---

## Sección 8. Conclusiones y Recomendaciones

### Estimación de un modelo estructural de la demanda de gasolina mediante el método de mínimos cuadrados

El modelo estimado en este documento mediante mínimos cuadrados es congruente con la teoría económica subyacente, la cual establece que la demanda de cualquier energético depende de una variable de ingreso general, en este caso el PIB del DF, y de los precios relativos de las gasolinas y es sensible a los cambios en el ingreso y al movimiento de precios relativos<sup>16</sup>.

Las elasticidades ingreso y precio de la demanda de este modelo permiten establecer que el consumo de gasolinas en la Ciudad de México responde a los movimientos tanto del PIB como del precio de las gasolinas, y si este último se utiliza para controlar el consumo, se debe considerar también que los movimientos en el ingreso de los consumidores se traducirán en cambios en el monto del consumo.

El valor de la elasticidad precio, sobre todo en el corto plazo (-0.450), muestra una respuesta débil de la demanda de gasolinas al cambio en este rubro, por lo que si se deseara inhibir el consumo de gasolinas vía precios, el aumento en estos debería ser considerable, no así en el largo plazo (-0.737).

Por su parte, la elasticidad ingreso, corto y largo plazo (0.924 y 1.514), cuyo valor exhibe a la gasolina como un bien elástico respecto del ingreso, indica que la recuperación del crecimiento económico implica un aumento en el consumo de gasolinas. En este sentido, el control del consumo de gasolinas en un escenario de crecimiento y baja elasticidad precio implica una modificación del precio de las gasolinas en cifras superiores al 50 por ciento. Esta política podría derivarse en altos costos sociales, además de convertirse en un detonador de la inflación.

---

<sup>16</sup> Galindo, L.M. y Salinas, E., *op. cit.*

### **Quiebre estructural**

La serie construida como la diferencia entre las dos estimaciones realizadas con el modelo óptimo de demanda de gasolinas resultó no ser ruido blanco y observando su comportamiento a través del tiempo, se infiere un posible cambio estructural en el volumen de ventas de gasolina en la Ciudad de México, hacia finales de 2003. Aplicando la metodología de Perron se corroboró la existencia de este cambio estructural en las ventas de gasolinas para la Ciudad de México.

### **Modelo de series de tiempo y análisis de Intervención**

Se escogió la especificación ARIMA (2,0,1)(12,0,0) porque de acuerdo a las pruebas de comprobación, revisión de residuos y análisis de la variable de ruido blanco, propuestas por Box y Jenkins, ésta se ajusta bien a la variabilidad de la serie original. A esto se agrega el importantísimo hecho de que la estructura de especificación del modelo seleccionado es la más sencilla.

El modelo seleccionado muestra el fuerte carácter autorregresivo de la serie modelada, el comportamiento del valor actual de las ventas depende de hasta dos rezagos y de la estacionalidad anual de dichas ventas, presente en la serie.

Este mismo modelo sirvió para aplicar la metodología del Análisis de Intervención, con la cual se comprobó que las ventas de gasolina presentan un quiebre estructural a finales de 2003, acorde con los resultados obtenidos con la técnica de Perrón.

### **Estimación de un modelo con técnicas de cointegración**

El modelo de largo plazo para la demanda de gasolinas obtenido mediante técnicas de cointegración muestra elasticidades ingreso (1.90) y precio (-1.36) mayores a las obtenidas en el modelo de corto plazo.

Esto permite concluir que si se presentan periodos prolongados de crecimiento económico en la Ciudad de México, se tendrían demandas muy altas de gasolina difíciles de cubrir con las actuales cuotas de producción de las refinerías nacionales.

En el caso de los precios, como ya se hizo mención en los párrafos anteriores, la respuesta de la demanda de gasolinas al cambio en estos es elástica, por lo que una política sostenida de aumentos en el precio de este combustible permitiría inhibir su consumo.

### **Recomendaciones**

La elasticidad de la demanda para la Ciudad de México, de acuerdo a los valores obtenidos en este trabajo, es más sensible a las variaciones en el ingreso que el resto del país; por lo que cualquier recuperación del crecimiento económico, en el corto y largo plazo, implicaría un aumento en el consumo de gasolinas por encima del promedio nacional.

Este escenario plantearía la necesidad de una mayor producción de gasolinas, que ante la falta de inversiones para la construcción de refinerías en el país, tendría que compensarse con importaciones crecientes cada vez más caras ante la escalada de precios observada en los últimos dos años.

Por lo anterior, habría que preguntarse sobre una política energética adecuada que permita mitigar los efectos negativos en las finanzas públicas derivados del mayor consumo de gasolinas, la no creciente producción nacional y el consiguiente aumento en los volúmenes importados.

Con respecto a los precios de las gasolinas en la Ciudad de México, los resultados de este estudio permiten compartir la conclusión de otros estudios similares realizados para todo el país de que las demandas por gasolina son inelásticas con

respecto al precio. Esta inelasticidad es probable que esté relacionada con la serie de satisfactores sociológicos intangibles adicionales a la movilidad y al confort que proporciona el automóvil.

De hecho, el valor de la elasticidad precio en el corto plazo obtenido en este trabajo, sugiere optar por una política de modernización y mejoramiento del transporte público para desincentivar el consumo de gasolinas que la sola aplicación de un aumento de su precio. Sólo en un escenario de largo plazo, una política de aumento en el precio resultaría relevante dado el valor ligeramente cercano a uno de dicha elasticidad.

---

## Bibliografía

Barceinas P.,F., Bouchain G., R.,Mendoza, M.A., Ruíz G., L.A., *Tópicos en Economía Matemática y Econometría*, Biblioteca de Ciencias Sociales y Humanidades, Serie Economía, 1998.

Bentzen, J., *An empirical analysis of gasoline demand in Denmark using cointegration techniques*. Energy Economics 16 (2), 39-143, 1994.

Box,G.E.P. y Jenkins, G. (1970), *Time Series Análisis, Forecasting and Control*, San Francisco,Holden-Day

Box, G.E.P. y Tiao, G.C. (1975), *Intervention análisis with applications to economic and environmental problems*, Journal of the American Statistical Association 70,70

De Alba, E. y Samaniego, R., *Estimación de la demanda de gasolinas y diesel y el impacto de sus precios sobre los ingresos del Sector Público*, Documento de Trabajo No. 1985-VIII, Centro de Estudios Económicos, El Colegio de México, 1985.

Enders, W., *Applied Econometric Time Series*, John Wiley & Sons, Inc., 1995.

Eskeland,G.S. y Feyzioglu, T., *Rationing Can Backfire: The “Day without a Car” in México City*, The World Bank Economic Review, Volume 11, number 3, September 1997.

Espino Bravo, J.M., *Estimación de la elasticidad de la demanda de gasolina en México, 1993-2003*, Documents de Recerca del Programa de Doctorado en Economía Aplicada, Universitat Autònoma de Barcelona, Octubre 2005.

Galindo P., Luis Miguel, *La econometría aplicada moderna: Los mínimos cuadrados ordinarios y las pruebas de diagnóstico*, Cuaderno de Trabajo de Economía Aplicada, N° 17, Maestría en Ciencias Económicas (UACPyP, CCH, UNAM), 1995

Galindo P., Luis Miguel, *La metodología econométrica moderna: una versión aplicada*, Cuaderno de Trabajo de Economía Aplicada, N° 18, Maestría en Ciencias Económicas (UACPyP, CCH, UNAM), 1995.

Galindo, L.M. y Salinas, E., *La demanda de gasolinas y los instrumentos económicos en México*, Gaceta Ecológica, n° 41 (INE-SEMARNAP), México, 1996

Guerrero G., Víctor Manuel, *Análisis Estadístico de Series de Tiempo Económicas*, Thomson, 2003

Hadda Cella, Gabriel y Bazán Navarrete, Gerardo, *Energía y Contaminación del Aire en la Zona Metropolitana del Valle de México*, Contextos Culturales Editores, 2000.

Hamilton, J.D., *Time Series Analysis*, Princenton University Press, 1994.

Haro López, R.A. y Ibarrola Pérez, J.L., *Cálculo de la elasticidad precio de la demanda de gasolina en la zona fronteriza norte de México*, Gaceta Económica, año 6, núm 11., otoño 2000

Henderson, James M. y Quandt, Richard E., *Teoría Microeconómica: una aproximación matemática*, Editorial Ariel, 1973.

Herstein, I.N., *Álgebra moderna*, Editorial Trillas, México, 1974

Holden, D. y Perman, R., *Unit Roots and Cointegration for the Economist, Cointegration for the Applied Economist* editado por B. Bhaskara Rao, Martin Press NY, 1995

Ludlow W, Jorge, *Modelos , Pronósticos y Volatilidad de las Series de Tiempo Generadas en la Bolsa Mexicana de Valores*, Biblioteca de Ciencias Sociales y Humanidades, Serie Economía, UAM Azcapotzálco, México, 1997.

Nordhaus, William D., *The Efficient Use of Energy Resources*, Cowles Foundation for Research in Economics at Yale University, Monograph 26,1979.

Petróleos Mexicanos, *Modelos de demanda interna de productos petrolíferos y gas natural*, noviembre de 2000.

Perron, P., *The Great Crash, the Oil Price Shock and the Unit Root Hypothesis*, *Econometrica*, 57, 1989

Pindyck, R.S., *The Structure of World Energy Demand*, MIT Press, 1980.

PNUMA, *GEO Ciudad de México: una visión del sistema urbano ambiental*, 2004.

Propuesta del Gobierno del Distrito Federal, *Hacia la Agenda XXI de la Ciudad de México*, Gobierno del Distrito Federal, noviembre, 2004

Quenouille, M.H. (1949), *Approximate test of correlation in time series*, *Journal of the Royal Statistical Society* B-11,68.

Varian, H.R., *Análisis Microeconómico*, Antoni Bosch editor, España, 1992.

## ANEXO

### 1. Pruebas econométricas

#### 1.1 Grado de bondad del ajuste utilizando el coeficiente de determinación múltiple ( $R^2$ )

El coeficiente de determinación reporta el grado o la bondad del ajuste de los valores simulados por el modelo econométrico con respecto a los valores reales.

Este estadístico se construye atendiendo a la diferencia entre el valor real y el esperado:

$$(y_t - y_t^m) = (y_t^e - y_t^m) + u_t$$

donde  $y_t^e$  es el valor proyectado por el modelo econométrico en cuestión y  $y_t^m$  es el valor medio de la serie. La ecuación anterior indica que la diferencia entre el valor real y la media de las series es igual al valor esperado menos la media más un término aleatorio. Sumando todos los datos considerados al cuadrado se obtiene:

$$\sum (y - y^m)^2 = \sum (y^e - y^m)^2 + \sum u_t^2$$

dividiendo por la variación total  $\sum (y - y^m)$  y reordenando

$$R^2 = \sum \left[ \frac{(y^e - y^m)}{(y - y^m)} \right]^2 = \left[ 1 - \sum \frac{u_t^2}{(y - y^m)^2} \right]$$

Esta ecuación indica que el coeficiente de determinación está entre cero y uno y representa la parte explicada por el modelo.

## 1.2 Prueba de significancia estadística $t$ de los parámetros

La prueba más importante de significancia estadística es el estadístico  $t$ . Si expresamos la regresión en forma matricial como:

$$Y = XB + U$$

en donde  $B$  es la matriz de coeficientes y  $U$  representa un vector columna que incluye al término de error.

Como los estimadores son insesgados

$$E[B] = B^v + (X'X)^{-1}X'U = B$$

de donde se obtiene que:

$$B - B^v = (X'X)^{-1}X'U$$

en donde  $B^v$  el valor verdadero de los coeficientes a estimar

La matriz de covarianza de los estimadores se puede representar como:

$$E[B - E(B)] [B - E(B)]' = E[(B - B^v)(B - B^v)']$$

donde:

$$\begin{aligned} \text{COV } B &= E[(X'X)^{-1}X'UU'X(X'X)^{-1}] = (X'X)^{-1}X'E(UU')X(X'X)^{-1} \\ &= \sigma^2 (X'X)^{-1}X'IX(X'X)^{-1} = \sigma^2 (X'X)^{-1} \end{aligned}$$

Así, bajo el supuesto de no autocorrelación, homoscedasticidad y normalidad en los errores puede obtenerse que:

$$[\beta - \beta^v] - N\left[0, \frac{\sigma^2}{\sum (x - x^m)^2}\right]$$

Normalizando la ecuación anterior se obtiene que

$$\frac{[\beta - \beta^v]}{\frac{\sigma^2}{(\sum (x - x^m)^2)^{1/2}}}$$

De donde puede desprenderse que

$$\left[\frac{\beta - \beta^v}{se(\beta)}\right] \sim t_{n-k}$$

donde k es el número de variables explicativas. Normalmente, las pruebas de significancia se realizan considerando que el  $\beta^v = 0$ . De este modo, y atendiendo a que el valor en tablas de una  $t_{25} = 2$ , entonces el coeficiente estimado será significativo si es dos veces el error estándar.

### 1.3 Prueba de normalidad de Jarque – Bera para los residuales

La prueba de normalidad de Jarque-Bera analiza si los residuales tiene una distribución normal. Esta prueba se basa en el análisis de la curtosis y que la dispersión o distribución de las series estadísticas sea simétrica (*skewness*). Estas características se definen de la siguiente forma:

Simetría (*skewness* (SK))

$$SK = \frac{\left(\frac{1}{n-1}\right)\sum (x - x^m)^3}{\left(\frac{1}{n-1}\right)\sum (x - x^m)^2}$$

Curtosis(KC)

$$KC = \left[ \frac{\left(\frac{1}{n-1}\right)\sum (x - x^m)^4}{\left(\frac{1}{n-1}\right)\sum (x - x^m)^2} \right]^{-3}$$

donde  $n$  representa el número de datos de la muestra,  $x$  representa cada uno de los datos de la muestra y  $x^m$  la media correspondiente. La prueba de normalidad se define entonces como:

$$\chi^2 (2) = (n-k)/6 [SK^2 + (1/4)KC^2]$$

$\chi^2 (2)$  es una chi-cuadrada con dos grados de libertad y  $k$  es el número de variables consideradas. La hipótesis nula es que los errores se distribuyen normalmente. Cuando la serie se distribuye normalmente  $SK$  y  $KC$  son cero conjuntamente. La distribución normal en los errores es particularmente importante porque esto favorece la potencia de otras pruebas.

#### 1.4 Prueba de Durbin Watson para detectar autocorrelación

La autocorrelación se define como la existencia de correlación entre los errores ( $u_t$ ) con sus valores pasados. En términos formales la autocorrelación implica que:

$$E(u_t u_{t-k}) \neq 0$$

en donde la ecuación anterior representa la covarianza del término de error.

La principal causa de la autocorrelación serial en los errores de la ecuación estimada pueden resumirse en los siguientes aspectos:

- La omisión de variables relevantes en la ecuación estimada
- Transformaciones en las ecuaciones o ajustes estacionales
- La presencia de rezagos en el proceso de ajuste que no fueron considerados en la ecuación inicial

La función de autocorrelación se define como:

$$r(\rho) = \frac{\sum u_t - u_{t-k}}{\sum u_t^2}$$

donde los intervalos de confianza están dados por:

$$(-1/n) \pm (2/n^{-1/2})$$

n representa el número de observaciones. Los valores fuera de estas bandas indican la presencia de autocorrelación.

La prueba de Durbin-Watson se define como la razón de la suma del cuadrado de la primera diferencia de los residuales con respecto a la suma del cuadrado de los residuales:

$$d = \frac{\sum (u_t - u_{t-1})^2}{\sum u_t^2}$$

La hipótesis nula ( $H_0$ ) es que no existe autocorrelación.

Una aproximación para muestras grandes a esta prueba puede obtenerse utilizando la siguiente ecuación:

$$d = 2(1 - \rho)$$

donde  $u_t = \rho u_{t-1} + v_t$ . De la ecuación anterior se desprende que cuando  $d$  difiere substancialmente de dos entonces existe la posibilidad de autocorrelación serial. Dicha ecuación indica que si la autocorrelación es cero ( $\rho = 0$ ) entonces  $d = 2$ . Por el contrario si existe autocorrelación positiva ( $0 < \rho < 1$ ) entonces  $0 < d < 2$  y si existe una autocorrelación negativa ( $-1 < \rho < 0$ ) entonces  $2 < d < 4$ .

La Durbin-Watson es válida sólo cuando las variables incluidas en la ecuación son exógenas, pierde potencia cuando se incluyen los valores rezagados de la variable dependiente en la ecuación de la regresión.

### 1.5 Prueba del multiplicador de Lagrange para detectar autocorrelación si se incluyen simultáneamente variables endógenas y exógenas

Esta prueba analiza conjuntamente a la función de autocorrelación  $\nu(\rho)$  con la ventaja de que sigue siendo válida si se incluyen simultáneamente variables endógenas y exógenas. La prueba se define como:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_k y_{t-k} + \alpha_{k+1} x_{it} + \dots + \alpha_{k+1+s} x_{st} + u_t$$

asumiendo que los errores son autorregresivos de orden  $p$  entonces se estima la siguiente regresión:

$$u_t = \phi_1 u_{t-1} + \phi_2 u_{t-2} + \dots + \phi_p u_{t-p} + v_t$$

en donde  $H_0: \phi_1 = \dots = \phi_p = 0$

El estadístico se distribuye como  $\chi^2_{(p)} = nR^2$ . La hipótesis nula es que establece la independencia serial.

## 1.6 Prueba de White para detectar heteroscedasticidad

Esta prueba asume que la heteroscedasticidad es función de las variables independientes de la ecuación inicial. De este modo se utiliza la siguiente regresión:

$$u_t^2 = \beta_0 + \beta_1 x_{1t} + \dots + \beta_k x_{kt} + \gamma_{k+1} x_{it}^2 + \dots + \gamma_{kt} x_{kt}^2 + \mu x_{it} x_{kt} + \dots + \mu x_{kt} x_{1t} + v_t$$

Esta prueba se utiliza también para detectar problemas de especificación y se distribuye como chi-cuadrada con el número de grados de libertad dados por el número de variables incluidas en la regresión auxiliar excluyendo la constante. La hipótesis nula es que no existe heteroscedasticidad. En esta prueba resulta también útil analizar la R2 y las t en la regresión auxiliar.

## 1.7 Prueba de Arch para detectar heteroscedasticidad

Esta prueba se basa en la estimación de una regresión que incluye los valores rezagados al cuadrado de los residuales de la ecuación original:

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_p u_{t-p}^2$$

La hipótesis nula es que no existe heteroscedasticidad. Esta hipótesis se rechaza si los coeficientes de la ecuación son estadísticamente significativos. La prueba se distribuye como una chi-cuadrada con p grados de libertad. Es importante hacer notar que p incluye solamente los valores rezagados de  $u_t$ .

## 2. Modelos de series de tiempo.

### 2.1 Operador de retraso

El operador de retraso que se denota por la letra B (del inglés *backward*) se define mediante la relación

$$BZ_t = Z_{t-1} \quad \text{para toda } t$$

por la aplicación sucesiva del operador B se tiene

$$B^2 Z_t = B(BZ_t) = Z_{t-2}$$

$$B^3 Z_t = B(B^2 Z_t) = Z_{t-3}$$

.....

$$B^k Z_t = B(B^{k-1} Z_t) = Z_{t-k}$$

así que, en general, a la expresión que se llega es

$$B^k Z_t = Z_{t-k} \quad \text{para } k=0,1,2,\dots,t$$

### 2.2 Operador diferencia

El operador diferencia, el cual se denota por  $\nabla$ , se utiliza para expresar relaciones del tipo  $Y_t = Z_t - Z_{t-1}$ , esto es

$$\nabla Z_t = Z_t - Z_{t-1} \quad \text{para toda } t$$

La relación que liga a  $\nabla$  con B es

$$\nabla = 1 - B \text{ o sea } \nabla Z_t = (1-B)Z_t$$

La expresión general para  $\nabla^k$  es

$$\nabla^k Z_t = \sum_{j=0}^k \frac{k!}{j!(k-j)!} (-1)^j Z_{t-j} \quad \text{para } k=0,1,2,\dots,t$$

A través del empleo del teorema del binomio, se tiene que

$$\nabla^k Z_t = (1-B)^k Z_t$$

### 2.3 Procesos estacionarios

Los modelos para los procesos estocásticos se basan en la idea de que una serie de tiempo cuyos valores sucesivos pueden ser altamente dependientes, puede considerarse generada a partir de una serie de choques aleatorios independientes  $\{a_t\}$ . Estos choques aleatorios se supone que son realizaciones independientes de una variable aleatoria cuya media es constante (generalmente se le considera igual a cero) y cuya varianza es  $\sigma^2$ , a esta sucesión de variables aleatorias  $\{a_t\}$  se le conoce como proceso de ruido blanco.

Así el proceso  $\{Z_t\}$  se puede expresar en función de  $\{a_t\}$  mediante la relación lineal

$$\begin{aligned} Z_t &= \mu + a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots \\ &= \mu + \psi(B)a_t \end{aligned}$$

en donde  $\mu$  es un parámetro que determina el nivel (no necesariamente la media) del proceso y  $\psi(B)$  es un polinomio de retraso

$$\Psi(B) = 1 - \psi_1 B - \psi_2 B^2 - \dots$$

que convierte al proceso  $\{a_t\}$  en el proceso  $\{Z_t\}$ .

El concepto de estacionariedad es de suma importancia para el análisis de series de tiempo. Para caracterizar completamente un proceso estocástico es necesario conocer la función de densidad conjunta de sus variables aleatorias, sin embargo en la realidad esto no puede lograrse con una serie de tiempo, por lo que si se obtienen los primeros momentos de las variables aleatorias, estos resumen en buena medida su distribución.

Suponiendo que la media de  $Z_t$  es  $\mu_t$

$$E(Z_t) = \mu_t$$

Entonces, de acuerdo con la expresión que define a  $Z_t$  se tiene

$$\mu_t = \mu + E(a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots)$$

con esta expresión, podría obtenerse la esperanza de la suma en el paréntesis mediante la suma de las esperanzas de cada una de los sumandos, pero esto no es válido necesariamente, a menos que se tenga

$$\psi_0 + \sum_{i=1}^{\infty} |\psi_i| < \infty \quad (1)$$

donde  $\psi_0 = 1$ ; esto es, para que la esperanza de la suma exista se requiere que el valor absoluto de la serie de ponderaciones converja, si esto ocurre se tendría entonces que

$$E(Z_t) = \mu$$

ya que  $E(a_t) = 0$  para toda  $t$ . Por consiguiente, si la ecuación anterior se cumple, la media del proceso no depende del tiempo, lo cual implica en particular que aún cuando en cierto periodo el proceso se aleje de la media, éste siempre regresará a una vecindad de la misma.

La varianza del proceso se obtiene a partir de su definición, o sea, si  $\gamma_0$  denota la varianza de  $Z_t$

$$\begin{aligned} \gamma_0 &= E[(Z_t - \mu)^2] \\ &= E[(a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots)^2] \\ &= E[(a_t^2 + \psi_1^2 a_{t-1}^2 + \psi_2^2 a_{t-2}^2 + \dots) - E(\text{productos cruzados de } \psi_i a_{t-i} \text{ y } \psi_j a_{t-j})] \\ &= \sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2 \end{aligned}$$

expresión que tendrá sentido solamente cuando  $\sum_{i=0}^{\infty} \psi_i^2$  converja, esto ocurre si se cumple (1). La razón por la que desaparece la esperanza de los productos cruzados, es por suponer que  $\{a_t\}$  son variables aleatorias independientes con media cero, entonces se sigue que

$$E(\psi_i a_{t-i} \cdot \psi_j a_{t-j}) = \psi_i E(a_{t-i}) \cdot \psi_j E(a_{t-j}) = 0 \quad \text{para toda } i \neq j$$

Finalmente al considerar una serie de tiempo como un proceso estocástico se debe de tener en consideración a todas las variables  $Z_1, Z_2, \dots, Z_t, \dots, Z_N$ , de tal forma que resulta necesario estudiar la covarianza entre  $Z_t$  y  $Z_{t+k}$ . Dicha covarianza se denota por  $\gamma_k = \text{Cov}(Z_t, Z_{t+k})$ , para  $k = \pm 1, \pm 2, \dots$  por

$$\gamma_k = E[(Z_t - \mu)(Z_{t+k} - \mu)]$$

$$\begin{aligned}
&= E[(a_t - \psi_1 a_{t-1} - \psi_2 a_{t-2} - \dots)(a_{t+k} - \psi_1 a_{t+k-1} - \psi_2 a_{t+k-2} - \dots)] \\
&= E(-\psi_k a_t^2 + \psi_1 \psi_{k+1} a_{t-1}^2 + \psi_2 \psi_{k+2} a_{t-2}^2 + \dots) - \\
&\quad E(\text{productos cruzados de } \psi_i a_{t-i} \text{ y } \psi_j a_{t-j}) \\
&= \sigma_a^2 (-\psi_k + \psi_1 \psi_{k+1} + \psi_2 \psi_{k+2} + \dots) \\
&= \sigma_a^2 \left( \sum_{i=1}^{\infty} \psi_i \psi_{k+i} - \psi_k \right)
\end{aligned}$$

de nuevo, tomar la esperanza para obtener esta última expresión sólo se justifica si  $\sum_{i=1}^{\infty} \psi_i \psi_{k+i}$  converge, y esto es cierto cuando se cumple (1).

## 2.4 Modelos autorregresivos (AR)

El término autorregresivo que se le da al proceso representado por la ecuación

$$\phi(B)\tilde{Z}_t = a_t \quad \text{con } \tilde{Z}_t = Z_t - \mu$$

en donde se supone que  $\{a_t\}$  es un proceso de ruido blanco y se refiere al hecho de que también puede expresarse como

$$Z_t = (1 - \phi_1 - \phi_2 - \dots - \phi_p)\mu + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t$$

la cual es básicamente una ecuación de regresión lineal, con la característica especial de que el valor de la variable dependiente  $Z$  en el periodo  $t$  depende, no de los valores de un cierto conjunto de variables independientes, como sucede en el modelo de regresión, sino de sus propios valores, observados en periodos anteriores a  $t$  y ponderados de acuerdo con los coeficientes autorregresivos  $\phi_1 \dots \phi_p$ .

Un proceso se considerara autorregresivo si se puede describir mediante la ecuación

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t$$

en donde  $\tilde{Z}_t = Z_t - \mu$ , y este será estacionario si y sólo si las raíces de la ecuación característica

$$1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0$$

se encuentran fuera del circulo unitario.

## 2.5 Modelos de promedios móviles (MA)

La idea básica de estos modelos se basa en representar un proceso estocástico  $\{Z_t\}$ , cuyos valores pueden ser dependientes unos de otros, como una suma infinita ponderada de choques aleatorios independientes  $\{a_t\}$ , o sea

$$\tilde{Z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t = \theta(B) a_t$$

en donde  $\{\tilde{Z}_t\}$  representa las desviaciones de  $\{Z_t\}$  respecto a su nivel medio  $\mu$  y  $\theta_1, \theta_2, \dots, \theta_q$  son las ponderaciones (parámetros de promedios móviles) asociados con los choques aleatorios en los periodos  $t-1, t-2, \dots, t-q$ , respectivamente.

El término de promedios móviles parece sugerir que el modelo se obtiene como un promedio de los choques aleatorios que intervienen, pero no es así, puesto que los parámetros no tienen que ser necesariamente positivos, ni su suma debe ser la unidad, como requeriría un promedio.

En general, un proceso estocástico se dirá que sigue un esquema de promedios móviles de orden  $q \geq 1$  si se le puede representar mediante

$$\tilde{Z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

con  $\tilde{Z}_t = Z_t - \mu$ , en donde  $\mu$  es el nivel del proceso,  $\theta_1, \theta_2, \dots, \theta_q$  son los parámetros de los promedios móviles y  $\{a_t\}$  es un proceso de ruido blanco con media cero y varianza constante  $\sigma_a^2$ . Se puede demostrar que todo proceso MA es estacionario, y en particular se observa en las formulas siguientes que ni la media, ni la varianza, ni las covarianzas del proceso MA(q), dependen del tiempo:

$$E(\tilde{Z}_t) = E(Z_t) - \mu = 0$$

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2$$

$$\gamma_k = (-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \sigma_a^2 \quad \text{si } k=1, \dots, q$$

$$\gamma_k = 0 \quad \text{si } k \geq q+1$$

para que esta expresión de  $\gamma_k$  sea válida en general, se define  $\theta_0 = \theta_{q+1} = \theta_{q+2} = \dots = 0$ . De aquí se puede obtener la función de autocorrelación que muestra que el proceso MA(q) tiene una memoria limitada a q periodos.

$$\rho_k \begin{cases} = \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & \text{si } k = 1, \dots, q \\ = 0 & \text{si } k \geq q+1 \end{cases}$$

Una generalización de los modelos AR y MA consiste en combinar ambas clases de modelos para obtener lo que se conoce como modelos autorregresivos y de promedios móviles ARMA (p,q) que se representan mediante

$$\phi(B)\tilde{Z}_t = \theta(B)a_t$$

en donde  $\phi(B)$  y  $\theta(B)$  son polinomios de retraso de orden  $p$  y  $q$  respectivamente,  $\{a_t\}$  es un proceso de ruido blanco y  $\tilde{Z}_t$  es la serie de desviaciones de la variable  $Z_t$  respecto a su nivel  $\mu$ .

Esta generalización surge del hecho de que las series de tiempo que se observan en la práctica, muchas veces presentan características tanto de procesos AR como de procesos MA. Además, considerando que el principio de parsimonia sugiere construir modelos que incluyan el menor número de parámetros, esto puede lograrse si intervienen parámetros tanto autorregresivos como de promedios móviles.

## 2.6 Modelos autorregresivos e integrados de promedios móviles (ARIMA)

Estos modelos pueden ser vistos como una generalización de los modelos ARMA ya que un cierto tipo de no estacionariedad mostrado por algunas series de tiempo, puede representarse mediante la simple toma sucesiva de diferencias de la serie original. Esto permite gran flexibilidad de representación a los modelos ARMA, puesto que en realidad lo que se hace al aplicar el operador diferencia  $\nabla^d$  es eliminar una posible tendencia polinomial de orden  $d$ . Visto desde otro ángulo, si el proceso original  $\{\tilde{Z}_t\}$  adolece de no estacionariedad causada por una tendencia polinomial no determinista, es posible construir el proceso estacionario  $\{W_t\}$ , en donde

$$W_t = \nabla^d \tilde{Z}_t \quad \text{para toda } t$$

para esta nueva serie ya podría ser posible obtener un modelo ARMA:  $\phi(B)\tilde{Z}_t = \theta(B)a_t$ , lo cual sería equivalente a considerar el modelo ARIMA

$$\phi(B)\nabla^d \tilde{Z}_t = \theta(B)a_t, \quad d \geq 1$$

para  $\{\tilde{Z}_t\}$ , en donde  $\{a_t\}$  es un proceso de ruido blanco.

### 3. Conducta del consumidor

Consideremos el caso de un consumidor que se encuentra ante varias cestas de consumo pertenecientes a un conjunto  $X$ , que es su conjunto de consumo.  $X$  es el ortante no negativo de  $R^k$  (el conjunto de todas las  $k$ -plas de números reales).  $X$  es un conjunto convexo y cerrado.

Suponemos que el consumidor tiene unas determinadas preferencias respecto a las cestas de consumo de  $X$ . Cuando se escribe  $x \succeq y$  se dice que “el consumidor piensa que la cesta  $x$  es, al menos, tan buena como la de  $y$ ”. Esta relación binaria (una entre dos elementos de  $X$ ), es una *relación de equivalencia*<sup>1</sup> ya que cumple con las siguientes propiedades:

- i. **Simetria.**  $\forall x, y \in X$ , o bien  $x \succeq y$ , o bien  $y \succeq x$ , o ambos.
- ii. **Reflexiva.**  $\forall x \in X$ ,  $x \succeq x$
- iii. **Transitiva.**  $\forall x, y, z \in X$ , si  $x \succeq y$  e  $y \succeq z$ , entonces  $x \succeq z$

Dada una relación  $\succeq$  que describe una “preferencia débil”, definimos una relación  $\succ$  que describe una “preferencia estricta”.  $x \succ y$  quiere decir que “ $x$  prefiere estrictamente a  $y$ ”. Esta relación es también de equivalencia. Se puede de igual forma definir el concepto de indiferencia representado por  $\sim$  diciendo que  $x \sim y \Leftrightarrow x \succeq y \ \& \ y \succeq x$ .

Otro supuesto sobre la relación de equivalencia definida es:

<sup>1</sup> Herstein, I.N., “Álgebra moderna”, Editorial Trillas, México, 1974

iv. **Continuidad.**  $\forall y \in X$ , los conjuntos  $\{x : x \succeq y\}$  y  $\{x : x \preceq y\}$  son conjuntos cerrados. Por lo tanto  $\{x : x \succ y\}$  y  $\{x : x \prec y\}$  son conjuntos abiertos.

En el análisis económico, suele resultar útil resumir la conducta de un consumidor por medio de una función de utilidad; es decir, una función  $u : X \rightarrow R$  tal que  $x \succ y \Leftrightarrow u(x) > u(y)$ . Las funciones de utilidad suelen ser muy útiles para describir las preferencias, pero no debe dárseles una interpretación psicológica. Su única característica importante es su carácter ordinal. Si  $u(x)$  representa unas determinadas preferencias  $\succeq$  y  $f : R \rightarrow R$  es una función monótona,  $f(u(x))$  representará exactamente las mismas preferencias, ya que  $f(u(x)) \geq f(u(y)) \Leftrightarrow u(x) \geq u(y)$ .

También suelen ser útiles otros supuestos sobre la preferencia

v. **Monotonicidad débil.** Si  $x \geq y$ ,  $\Rightarrow x \succeq y$

vi. **Monotonicidad fuerte.** Si  $x \geq y$  &  $x \neq y$ ,  $\Rightarrow x \succ y$

Dos supuestos más que suelen utilizarse para garantizar que las funciones de demanda de consumo se comporten correctamente son:

vii. **Convexidad.** Dados  $x, y, z \in X$  tal que  $x \succeq y$  &  $y \succeq z \Rightarrow tx + (1-t)y \succeq z \forall t$  tal que  $0 \leq z \leq 1$ .

viii. **Convexidad estricta.** Dados  $x \neq y, z \in X$  si  $x \succeq y$  &  $y \succeq z \Rightarrow tx + (1-t)y \succeq z \forall t$  tal que  $0 < t < 1$ .

### Teorema (existencia de una función de utilidad)

Teorema. Supóngase que las preferencias son completas, reflexivas, transitivas, continuas y monótonas en sentido fuerte, entonces existe una función de utilidad continua  $u : R_+^k \rightarrow R$  que representa esas preferencias.

Demostración. Sea  $\mathbf{e}$  el vector de  $R_+^k$  formado únicamente por unos. En ese caso, dado cualquier vector  $x$ , sea  $u(x)$  un número tal que  $x \sim u(x)\mathbf{e}$ . Se tiene que demostrar que existe ese número y que es único.

Sea  $B = \{t \in R : t\mathbf{e} \succeq x\}$  y  $W = \{t \in R : x \succeq t\mathbf{e}\}$ . En ese caso, la monotonía fuerte implica que  $B$  no es un conjunto vacío;  $W$  tampoco lo es, ya que tiene al menos un elemento, 0. La continuidad implica que los dos conjuntos son cerrados. Dado que la línea real está conectada, existe algún  $t_x$  tal que  $t_x\mathbf{e} \sim x$ . Se tiene que demostrar que esta función de utilidad representa, de hecho, las preferencias subyacentes. Sea

$$u(x) = t_x \quad \text{donde} \quad t_x\mathbf{e} \sim x$$

$$u(y) = t_y \quad \text{donde} \quad t_y\mathbf{e} \sim y$$

en ese caso, si  $t_x < t_y$ , la monotonía fuerte demuestra que  $t_x\mathbf{e} \prec t_y\mathbf{e}$  y la transitividad demuestra que

$$x \sim t_x\mathbf{e} \prec t_y\mathbf{e} \sim y$$

del mismo modo, si  $x \succ y$ , entonces  $t_x\mathbf{e} \succ t_y\mathbf{e}$ , por lo que  $t_x$  debe ser mayor que  $t_y$ .

Partiendo de la hipótesis de que un consumidor racional siempre elige del conjunto de opciones asequibles la cesta por la que muestra una mayor preferencia se puede analizar la conducta del consumidor. En el problema básico de maximización de las preferencias, el conjunto de opciones asequibles es

simplemente el conjunto de todas las cestas que satisfacen la restricción presupuestaria del consumidor. Sea  $m$  la cantidad fija de dinero de que dispone éste y  $\mathbf{p} = (p_1, \dots, p_k)$  el vector de los precios de los bienes,  $x_1, x_2, \dots, x_k$ . El conjunto de cestas asequibles, el conjunto presupuestario del consumidor, viene dado por

$$B = \{x \in X : px \leq m\}$$

El problema de maximización de las preferencias puede expresarse de la siguiente manera:

$$\begin{aligned} &\max u(x) \\ &\text{sujeta a } px \leq m \\ &x \text{ pertenece a } X \end{aligned}$$

Una cesta  $\mathbf{x}$ , maximizadora de la utilidad, debe cumplir con la restricción presupuestaria con igualdad, por lo que el problema del consumidor se puede expresar de la forma

$$\begin{aligned} v(\mathbf{p}, m) &= \max u(\mathbf{x}) \\ &\text{tal que } \mathbf{p}\mathbf{x} = m \end{aligned}$$

La función  $v(\mathbf{p}, m)$  que nos da la utilidad máxima alcanzable a los precios y la renta dados se denomina **función indirecta de utilidad**. El valor de  $\mathbf{x}$  que resuelve este problema es la **cesta demandada** del consumidor.

La función que relaciona  $\mathbf{p}$  y  $m$  con la cesta demandada se denomina **función de demanda** del consumidor y se representa por  $\mathbf{x}(\mathbf{p}, m)$ . Es necesario suponer que hay una única cesta que maximiza la utilidad, la convexidad estricta de las preferencias garantiza esta conducta. Esta función de demanda del consumidor es homogénea de grado 0 en  $(\mathbf{p}, m)$ . Las funciones de demanda expresadas en

función de los precios y del ingreso son observables y se les denomina **funciones de demanda marshallianas**.

La conducta maximizadora puede caracterizarse por medio del cálculo, en la medida en que la función de utilidad sea diferenciable. El lagrangiano del problema de maximización de la utilidad puede expresarse de la forma siguiente:

$$\mathcal{L} = u(x) - \lambda (\mathbf{p}\mathbf{x} - m)$$

donde  $\lambda$  es el multiplicador de Lagrange. Diferenciando el lagrangiano con respecto a  $x_i$  se obtienen las condiciones de primer orden

$$\frac{\partial u(x)}{\partial x_i} - \lambda p_i = 0 \quad \text{siendo } i=1, \dots, k$$

Para interpretar estas condiciones, se puede dividir la condición de primer orden  $i$ -ésima por la  $j$ -ésima a fin de eliminar el multiplicador de Lagrange. De esa manera se tiene que

$$\frac{\frac{\partial u(x^*)}{\partial x_i}}{\frac{\partial u(x^*)}{\partial x_j}} = \frac{p_i}{p_j} \quad \text{siendo } i, j = 1, \dots, k$$

El cociente del primer miembro es la relación marginal de sustitución entre el bien  $i$  y el  $j$  y el del segundo se denomina relación económica de sustitución entre los bienes  $i$  y  $j$ . La maximización implica que estas dos relaciones de sustitución son iguales.

#### 4. Elasticidad precio y elasticidad ingreso

La elasticidad directa de la demanda del bien  $x_i$  ( $\mathcal{E}_i$ ), o elasticidad de la demanda con respecto a su propio precio, se define como el cociente entre la variación proporcional de  $x_i$  y la variación proporcional de su precio  $p_i$ , permaneciendo los demás precios y el ingreso constantes

$$\mathcal{E}_i = \frac{\partial(\log x_i)}{\partial(\log p_i)} = \frac{p_i}{x_i} \frac{\partial x_i}{\partial p_i}$$

La elasticidad de la demanda con respecto al ingreso ( $\eta_i$ ), o elasticidad ingreso de la demanda, se define como el cociente entre la variación proporcional de la cantidad del bien  $x_i$  y la variación proporcional de la renta, permaneciendo todos los demás precios constantes

$$\eta_i = \frac{\partial(\log x_i)}{\partial(\log m)} = \frac{p_i}{x_i} \frac{\partial \varphi(p_1, \dots, p_n, m)}{\partial m}$$

#### 5. Funciones de demanda de energía

El objetivo para la estimación de las relaciones de demanda de energía está determinado por la valuación social de los productos energéticos, observada a través de las actividades del mercado. La valuación social se realiza mediante funciones de preferencia o utilidad individuales, las cuales se agregan en una función de preferencia general, que se supone bien comportada, esto es, expresa una relación completa, reflexiva, transitiva, continua, monótona y convexa. La función de preferencia es maximizada a lo largo del tiempo, de modo que los agentes económicos o consumidores de energía obtienen las canastas que maximizan su utilidad dentro de un conjunto de bienes.

Suponiendo que los consumidores enfrentan dos bienes, los servicio de energía (E) y los bienes no energéticos (X), una aproximación a la función general de preferencias puede obtenerse mediante: (1) especificando su forma en términos de una función paramétrica y estimando los parámetros de dicha función mediante técnicas econométricas; y (2) en otro enfoque, se puede determinar como dependiente de los rezagos y errores aleatorios de la serie de tiempo misma (enfoque de series de tiempo)

Para los estudios explicativos de la demanda de energía, se asume generalmente, que la función de preferencia es separable en el tiempo, y tiene la forma:

$$C_t = X_t + \sum_{i=1}^n A_i E_{it}^{-\alpha_i} (X_t + \sum_{i=1}^n \omega_i E_{it})^{\beta_i}$$

en donde:

$C_t$  = índice de consumo al tiempo t

$X_t$  = índice de consumo de bienes no energéticos al tiempo t

$A_i$  ,  $\alpha_i$  ,  $\beta_i$  = parámetros de la función de preferencias

$\omega_i$  = ponderaciones en el periodo base de la energía en función de otros bienes

$E_i$  = servicios energéticos diferenciados

De esta manera, la parte que aparece entre paréntesis en la ecuación anterior es el ingreso total Y, el cual esta sujeto a la restricción presupuestal agregada:

$$\sum_{i,t} p_{it} E_{it} + X_t = Y_t$$

en donde

$Y_t$  = ingreso total al tiempo t

$p_{it}$  = precio relativo del energético  $i$  al tiempo  $t$

Al maximizar la función de preferencias sujeta a esta restricción presupuestal, se tiene la ecuación de la demanda de cada energético que depende de una variable de ingreso general y de los precios relativos de los combustibles.

$$E_{it} = k_i p_{it}^{-\gamma_i} Y_t^{\delta_i} \quad i=1, \dots, N$$

en donde:

$$k_i = (\alpha_i A_i)^{\frac{1}{(1-\alpha_i)}}$$

$$\gamma_i = \frac{1}{(1-\alpha_i)}$$

$$\delta_i = \beta_i \gamma_i$$

Esta ecuación es llamada la “ecuación base” del laboratorio de Cavendish (Cambridge, Inglaterra) que es la que retuvo la Conferencia Mundial de la Energía en 1979, elaborada por un grupo multinacional de expertos en energía.