https://doi.org/10.29201/peipn.v17i34.86
PANORAMA ECONÓMICO, vol. XVII, núm. 34, julio-diciembre, 2021, pp. 153-184

Relationships among gasoline prices of eight brands at the north of Mexico City

Edgar I. Faustino-Cruz* Francisco Ortiz-Arango**[1]

(*Recibido: diciembre 2020/Aceptado: abril 2021*)

Abstract

This document uses a vector error correction model to obtain the decomposition into permanent and transitory components of gasoline prices offered by eight brands in the north of Mexico City, through an impulse response analysis and variance decomposition, from July 1, 2018, to june 17, 2020. The main findings are that are multiple influences and interdependencies among the eight brands' prices analyzed. In the short term, three patterns in pricing are identified: (*a*) prices initially explained by themselves but rapidly influenced by the rest of prices, (*b*) prices explained throughout the cycle mainly by their disturbances, and (*c*) prices that depend strongly on the rest of prices. In the long term, these patterns consequently determine that there are three cointegration vectors between all prices. The results found in the analyzed period suggest that it is perhaps still early to expect that there will be an equilibrium price vector derived from a competitive market in Mexico.

Keywords: gasoline prices, econometric modeling, vector error correction model, variance decomposition, Mexico City.
JEL classification: C51, D41.

* Doctorando en la Universidad Panamericana, México, City and Comisión Reguladora de Energía México. <edgarivanfc@gmail.com>.
* Profesor-investigador en la ECEE, Universidad Panamericana, México, City, México. <fortizar@up.edu.mx>.
[1] Correspondence autor. ORCID: https://orcid.org/0000-0003-2707-4459.

Relaciones entre los precios de la gasolina de ocho marcas en el norte de la Ciudad de México

Resumen

Este documento utiliza un modelo de corrección de error vectorial para obtener la descomposición en componentes permanentes y transitorios de los precios de la gasolina que ofrecen ocho marcas en el norte de la Ciudad de México, a través de un análisis de impulso-respuesta y descomposición de la varianza, a partir de julio 1 de 2018 al 17 de junio de 2020. Los principales hallazgos son que existen múltiples influencias e interdependencias entre los precios de las ocho marcas analizadas. En el corto plazo, se identifican tres patrones en la fijación de precios: *a*) precios inicialmente explicados por sí mismos pero rápidamente influenciados por el resto de precios, *b*) precios explicados a lo largo del ciclo principalmente por sus perturbaciones, y *c*) precios que dependen fuertemente del resto de precios. A largo plazo, estos patrones determinan en consecuencia que existan tres vectores de cointegración entre todos los precios. Los resultados encontrados en el período analizado sugieren que quizás aún sea prematuro para esperar que exista un vector de precios de equilibrio derivado de un mercado competitivo en México.

Palabras clave: precios de la gasolina, modelado econométrico, modelo de corrección de errores vectoriales, descomposición de la varianza, Ciudad de México.
Clasificación JEL: C51, D41.

1. Introduction

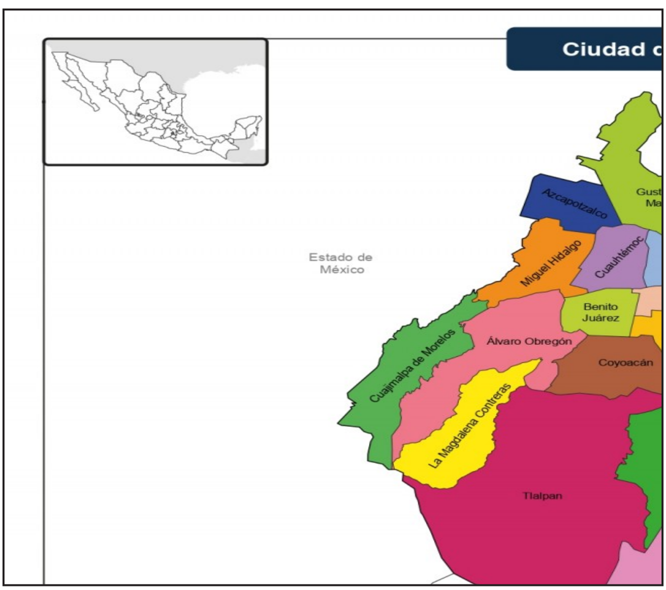
One of the main objectives of the 2013 energy reform in Mexico was to promote an environment of greater competition in the gasoline market. In this reform process, the Energy Regulatory Commission (CRE) finalized the schedule's implementation for making the gasoline and diesel markets more flexible on November 30, 2017. The 11, 774[2] service stations (gas stations) established at that moment in the national territory began to sell gasoline at free-market prices.

Into this context, this work aims to examine the short and long-term relationships among the leading competitors' gasoline prices to determine the

[2] https://www.gob.mx/cre/articulos/estrategia-de-flexibilizacion-de-los-mercados-de-gasolinas-y-diesel. Accessed June 3, 2020.

degree of determination of prices via competition or whether these are determined unilaterally, and we tried to verify the existence of a single price law associated with a competitive gasoline market. To this end, the volatility transmission mechanism and the interactive relationships between gasoline prices are analyzed using a Vector Error Correction (VEC) model. As previously mentioned, the different competitors face the same logistical conditions like road infrastructure and supply terminals. Consumers can quickly drive to choose among other brands. It would be expected that any series diverges from the rest, and there are relationships of cointegration.

For developing this analysis, we focus on four mayors of Mexico City: Gustavo A. Madero, Miguel Hidalgo, Cuauhtémoc, and Venustiano Carranza, this is due to the precision and reliability in collecting data and that it can be considered as a relevant market, the latter defined by two elements: 1) geographic delimitation and 2) by the similarity of logistics costs; regarding the first, it is considered that a gas station located close to others faces greater competition than an isolated gas station since consumers do not face costs (or very few) to transport from one gas station to another in search of a better price; regarding the second element, it can be considered that all gas stations obtain their supply from the same storage terminals, so the differences in their prices cannot be due to logistics costs.



Source: National Institute of Statistics and Geography (INEGI).

Figure 1
Political division in mayors of Mexico City

A relevant contribution of this work is that there is no precedent in the Mexican case of cointegration to analyze the competition level in gasoline commercialization. It allows identifying the existence of a single price law associated with a competitive gasoline market in the region under study. Although many studies are dedicated to analyzing fuel demand, these focus on analyzing it through integration techniques at the international level. In the bibliographic review carried out, in no case is the level of competition in the commercialization of fuels analyzed either internationally or in Mexico, much less trying to identify the existence of a single price law associated with a competitive gasoline market. The most relevant conclusions of this work are that there are multiple influences among the eight brands' prices analyzed. The gasoline market in these regions is segmented, and distributors exercise a certain level of market power.

Some previous works related to this kind of study are, for example, (Bentzen, 1994), who finds a stable and positive long-term relationship between the demand for gasoline and its economic determinants in Denmark. Another case is (Cheung & Thomson, 2004), between 1980 and 1999, found that gasoline demand was relatively inelastic in the face of price changes, both in the short and long term. The long-term income elasticity was 0.97, which implies that gasoline consumption's future growth rate will approximate that economy's growth rate.

Further, in the period 1970-1989 (Eltony & Mutairi, 1995), found that gasoline demand is inelastic concerning the price in the short and long term. While it is elastic for a long time, the gasoline demand is inelastic concerning the price with respect to income in the short term. This suggests that gasoline demand response is greater to income changes in the long term than in the short term. Additionally, gasoline consumption is adjusted to its long-term level, with approximately 52% of the first year's adjustment.

For the period 1978-2005 (Akinboade, Ziramba, & Kumo, 2008), found that the demand for gasoline in South Africa was inelastic in terms of prices and income. In Fiji's case (Rao & Rao, 2009), using five-time series techniques, they find that the gasoline demand is inelastic both in prices and income.

The antecedent of this type of analysis in the case of Mexico is presented in the work of (Reyes, Escalante, & Matas, 2010), who for the period 1960-2008 find that the estimates of price and income elasticities of long and long and short term short term were: -0.285, -0.041, 1.004 and 0.721, respectively, this implies that the demand for fuel is sensitive to the trajectory of income and is inelastic to prices. While (Ferrer & Escalante, 2014), for the period 1980-2012, they find a short-term income elasticity of 0.49 and a short-term price elasticity of -0.12. That is, demand is sensitive to income inelastic to price.

In the study of (Ibarra Salazar & Sotres Cervantes, 2008), they estimate the price elasticity of gasoline demand for the border area. For the rest of Mexico, for this they use a data panel that combines monthly time series, from january 1997 to december 2003, with a cross-section of the Mexican states, finding that the estimated price elasticity for demand –both for the border region and for the non-border region– was negative. For the non-border region, the elasticity's numerical value varies between -0.15 and -1.06 (average of -0.67). In contrast, for the northern border region, it varies between -0.67 and -1.57 (average of -1.18); these differences indicate, among other things, that the competition faced by gas stations on the northern border means that, in the face of price changes, the gasoline demand is more sensitive in this small region than in the interior of the country.

This work is made up of the following sections: The second section presents the methodology and data used, and the construction of the empirical model is also described. In the third section, the econometric tests are carried out, the description and interpretation of the results obtained from the empirical analysis are also carried out. Finally, the conclusions of the study are presented.

2. Methodology and justification

Traditional standard regression techniques, such as ordinary least squares (OLS), require the variables to be stationary (a variable is stationary if its mean and all its autocovariance are finite and do not change over time). However, in practice, a lot of economic time series seems to be "first difference stationary" (as in our case), meaning that the time series level is not stationary, but its first difference is.[3] If non-stationarity is ignored, relationships could be established when, in reality, they do not exist, which in turn can lead to false conclusions (Granger & Newbold, 1974). Next, we will see how to analyze non-stationary series and obtain necessary information from them.

According to (Wooldridge, 2010) if $\{y_t: t = 0, 1, \ldots\}$ and $\{x_t: t = 0, 1, \ldots\}$ are two I (1) processes, then, in general, $y_t - Bx_t$ is a process I(1) for any number B. However, it is possible that for $B \neq 0$, and $y_t - Bx_t$ is an I(0) process, which means that it has a constant mean, constant variance, and the autocorrelations that depend only on the elapsed period between any two variables in the series and is not asymptotically correlated.

If such B exists, y and x are said to be cointegrated, and B is called a cointegration parameter. In this sense, cointegration exposes the presence of a long-term equilibrium towards which the system converges. The differences (or error term) in the cointegration equation are interpreted as the disequilibrium error for each particular point of time (Engle, 1987).

[3] In section 2.1 it is shown that the series used in this analysis are stationary in first difference.

When multiple time series are analyzed, the natural extension of the autoregressive model is vector autoregression (VAR), in which a vector of variables is modeled as dependent on its own lags and the lags of the other variables in the vector. (Sims, 1980) introduced the VAR model in the economic field and promoted its widespread application in the economic system's dynamic analysis.

Since the gasoline price series are non-stationary, we are inclined to use vector autoregressive models (VAR) and cointegration over other statistical techniques such as OLS. This is so given that, on the one hand, VAR models, together with vector autoregressive models (VEC), are used to model time series in multivariate contexts where there are dynamic dependencies between different series. Thus, VAR models constitute a direct extension when one wants to capture the dynamic dependencies that may exist between these series.

Similarly, since Engle and Granger's (1987) appearance, cointegration analysis has been widely used because it allows analyzing if the transmission of related events in the short term produces a common in the long term trends.

2.1. VEC model

To analyze the non-stationary series, we can use two versions of the cointegration models:
the vector autoregressive and the vector error correction model (VECM). The advantage of the error correction model is that it includes not only differentiated variations but also levels. Hence, it is advantageous to adopt this version of the autoregressive model since it provides both short-term and long-term parameters and allows to make future predictions by studying the analysis of the impulse-response functions and the decomposition of the variance. According to (Johansen S., 1988), the parameters of a VECM can be written as:

$$\Delta_{yt} = \alpha\beta' y_{t-1} + \sum_{i=1}^{p-1} \tau_i \Delta_{yt-i} + v + dt + \varepsilon t$$

where:

y_t is a $K \times 1$ vector of endogenous variables
α is a matrix of parameters $K \times r$
β is a matrix of parameters $K \times r$
$\tau_i, \ldots, \tau_{p-1}$ are matrices of parameters $K \times K$
v is a vector of parameters $K \times 1$,
δ is a $K \times 1$ vector of trend coefficients,
t is a time trend.[4]

[4] According to the graph of the data, we assume that the data do not have quadratic trends, and this implies that they are restricted to the cointegration equations to be stationary around a constant mean.

Data:

The gas stations used for this study are located in the northern area of Mexico City, in the municipalities of Azcapotzalco, Gustavo A. Madero, Miguel Hidalgo, Cuauhtémoc, and Venustiano Carranza, everyone identified from their Web pages,[5] such as is shown in table 1:

Table 1
Number of service stations by municipality and brand[6]

Brand / # of stations	**Azcapotzalco**	**Gustavo A. Madero**	**Miguel Hidalgo**	**Cuauhtémoc**	**Venustiano Carranza**	**Total**
Hidrosina	0	0	1	1	1	3
Petro Seven	1***	1	0	0	0	2
G500	0	1	0	1	1	3
Shell	0	1	1	1	0	3
BP	1	0	0	1	0	2
Total México	2	0	1	0	0	3
Repsol	0	1	0	0	0	1
Pemex	2	2	3	2	2	11
Total amount	6	6	6	6	4	28

Source: own elaboration from the Web pages of the different companies.
*** It is considered a station located in the state of Mexico in the limits with the Azcapotzalco mayor's office (200 meters of distance in a straight line on the highway).

For this research, the final daily consumer prices are used, reported by the stations for the period from July 1, 2018, to June 16, 2020; care was taken to ensure that all gas stations maintained the same ownership since November 2017, the date on which gas stations were allowed to set their prices freely. However, the first months of 2018 are excluded because they show too much volatility and alter the econometric tests results; this is offset by the fact that we include essential data referring to the Covid scenario.

In the case of brands with more than one station, the average of their stations is used; the sample's smallness is because it was ensured that the gas stations in the sample kept the same commercial brand during the period analyzed. The prices were obtained from CRE. Graph 1 shows the dynamics

[5] Date of consultation: june 11, 2020.
[6] The gas stations included in the sample and the source are shown in the Annex II.

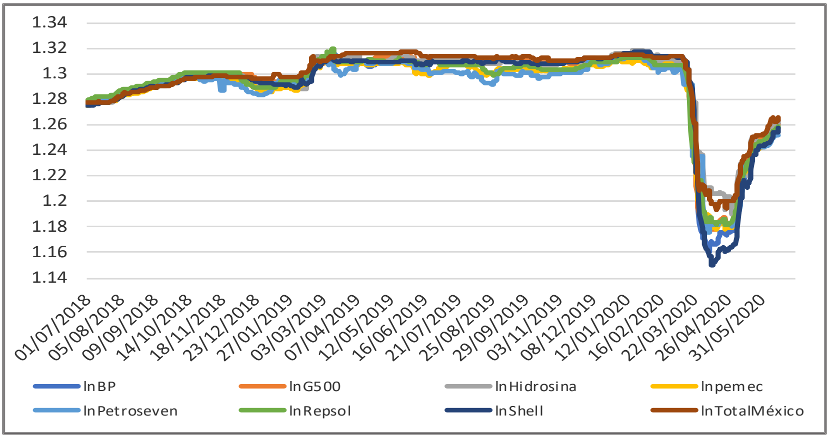
of prices, where the same trend can be seen in all series during the analyzed period, and we present the basic statistics in table 2.

Table 2
Basic statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
BP	187	19.63533	1.448209	14.49	20.73
G500	187	19.6804	1.360467	15.07332	20.74
Hidrosina	187	19.70965	1.188687	15.49	20.84333
Pemex	187	19.48214	1.294839	15.05909	20.55591
Petroseven	187	19.461118	1.27045	14.99	20.595
Repsol	187	19.60845	1.30114	15.19	20.89
Shell	187	19.59152	1.510204	14.13667	20.77667
Total Mexico	187	19.8032	1.24776	15.62333	20.74

Source: own elaboration

Graph 1
The logarithm of the prices reported by the stations



Source: own elaboration based on data from the Energy Regulatory Commission.

Development of econometric tests and analysis of the results

Econometric tests

The unit root test of ADF (Augmented Dickey-Fuller) was applied to test the stationarity of each of the series of logarithms of consumer prices. The test results are shown below:

The test results in table 3 show that the series is the first-order stationary.[7] To estimate the VEC model, the next step is to determine the lag intervals. Thus, the lag length criteria and the AR roots graph were adopted, according to (Khim & Liew, 2004) and (Nielsen 2001). to determine the lag intervals, as shown below:

Table 3
ADF unit root tests

Variable	ADF Statistic	P-value
D1logBP	-14.817	0.0000
D1logG500	-13.724	0.0000
d1logHidrosina	- 21.845	0.0000
d1logPemex	- 14.368	0.0000
d1logPetroSeven	- 23.812	0.0000
d1logRepsol	- 24.321	0.0000
d1logShell	- 17.174	0.0000
d1logTotalMxico	- 14.695	0.0000

Source: own elaboration based on data from the Energy Regulatory Commission.

According to table 4, both the Final Prediction Error (FPE) test and the Akaike Information Criterion (AIC) recommend 19 lags. In contrast, the Hannan-Quinn (HQIC) recommends 15 lags, and Schwarz Information Criteria (BIC) suggest 7 lag. In this case, the number of lags selected is 19 because it is also suggested by the likelihood ratio (LR) test.

[7] The prefix d1 indicates that the series is in the first difference

Table 4
Lag length tests

Sample: 7/20/2018 - 6/17/2020 Number of obs = 699								
Lag	**LL**	**LR**	**df**	**p**	**FPE**	**AIC**	**HQIC**	**SBIC**
0	22171.7				4.E-38	-63.4155	-63.3953	-63.3634
1	28521.2	12699.00	64	0.000	6.E-46	-81.3997	-81.2185	-80.9310
2	29055.5	1068.60	64	0.000	2.E-46	-82.7453	-82.4031	-81.8601
3	29505.0	899.01	64	0.000	5.E-47	-83.8483	-83.3451	-82.5466
4	29937.6	865.23	64	0.000	2.E-47	-84.9030	-84.2388	-83.1847
5	30230.8	586.29	64	0.000	1.E-47	-85.5587	-84.7334	-83.4238
6	30480.6	499.76	64	0.000	6.E-48	-86.0905	-85.1042	-83.5391
7	30695.5	429.82	64	0.000	4.E-48	-86.5223	-85.3749	-83.5543*
8	30890.6	390.13	64	0.000	3.E-48	-86.8973	-85.5889	-83.5127
9	31113.8	446.46	64	0.000	2.E-48	-87.3529	-85.8834	-83.5518
10	31297.8	367.89	64	0.000	1.E-48	-87.6961	-86.0656	-83.4784
11	31429.3	263.10	64	0.000	9.E-49	-87.8894	-86.0978	-83.2551
12	31548.5	238.26	64	0.000	8.E-49	-88.0471	-86.0945	-82.9963
13	31683.9	270.81	64	0.000	7.E-49	-88.2514	-86.1378	-82.7840
14	31820.3	272.92	64	0.000	5.E-49	-88.4587	-86.1841	-82.5748
15	31968.7	296.85	64	0.000	4.E-49	-88.7003	-86.2646*	-82.3998
16	32059.5	181.54	64	0.000	4.E-49	-88.7769	-86.1802	-82.0598
17	32182.6	246.18	64	0.000	3.E-49	-88.9459	-86.1882	-81.8123
18	32272.3	179.45	64	0.000	3.E-49	-89.0196	-86.1008	-81.4694
19	32388.0	231.3*	64	0.000	2.8e-49*	-89.1673*	-86.0875	-81.2006

Source: own elaboration.

3. Cointegration test

The cointegration test consists of selecting the appropriate form of the cointegration test and the order of lag. The cointegration relationship between variables in the VAR model is generally tested with the Johansen method (Johansen S., 1988).

The trace statistic suggests the possible presence of 4 cointegration relationships. In contrast, the statistic corresponding to the maximum auto value (max statistic) has a critical value very close to 5% in the case of the null hypothesis for a number no more significant than three cointegration relationships. Since the eight variables are cointegrated, a vector error correction model (VECM) is the appropriate model to apply.

Table 5
Cointegration Test

Johansen tests fer cointegration Trend: constant Number of obs = 699 Sample: 7/20/2018 - 6/17/2020 Lags = 19					
maximum rank	**parms**	**LL**	**eigenvalue**	**trace value**	**5% critical value**
0	1160	32280.411.	.	215.153.	156.00
1	1175	32307.911.	0.07567	160.154.	124.24
2	1188	32332.332.	0.06749	111.311.	094.15
3	1199	32352.175.	0.05519	71.625.	068.52
4	1208	32368.475.	0.04557	39.0244*	047.21
5	1215	32379.119.	0.02999	17.738.	029.68
6	1220	32383.932.	0.01368	8.112.	015.41
7	1223	32387.091.	0.00900	1.793.	003.76
8	1224	32387.988.	0.00256		
–	–	–	–	–	–
0	1160	32280.411.		54.999.	051.42
1	1175	32307.911.	0.07567	48.842.	045.28
2	1188	32332.332.	0.06749	39.686.	039.37
3	1199	32352.175.	0.05519	32.601.	033.46
4	1208	32368.475.	0.04557	21.287.	027.07
5	1215	32379.119.	0.02999	9.626.	020.97
6	1220	32383.932.	0.01368	6.319.	014.07
7	1223	32387.091.	0.00900	1.793.	003.76
8	1224	32387.988.	0.00256		

Source: own elaboration

VECM estimation and analysis

Table 6
Results and test of the VECM estimation.

```
Sample:  7/20/2018 - 6/17/2020                        No. of obs      =        699
                                                       AIC             =   -89.1573
Log likelihood =  32368.48                             HQIC            = -86.11773
Det(Sigma_ml)  =  8.29e-51                             SBIC            = -81.29467
```

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_lnBP	149	.000705	0.9194	6252.689	0.0000
D_lnG500	149	.000684	0.8720	3732.849	0.0000
D_lnHidrosina	149	.001362	0.7375	1539.259	0.0000
D_lnpemec	149	.000756	0.8372	2817.459	0.0000
D_lnPetroseven	149	.001452	0.8182	2465.818	0.0000
D_lnRepsol	149	.001186	0.7819	1964.099	0.0000
D_lnShell	149	.000757	0.9049	5216.154	0.0000
D_lnTotalMxico	149	.000656	0.8703	3678.749	0.0000

Source: own elaboration.

According to the cointegration tests (table 5), based on the relationships presented and in order to maintain the greatest possible simplicity, a cointegrated model with 4 long-term relationships was adjusted using the STATA 12 software. The VECM (table 6) has taken the first difference of the logarithms of the variables, which are represented as D_lnBP, D_lnG500, D_lnPetroseven, D_lnHidrosina, D_lnRepsol, D_lnShell, D_lnTotalMxico, D_lnpemex. Furthermore, the *R* squared value of the variables are good enough to justify its causality, and p values close to zero also indicate significance.

The first part of the VECM (annex I) contains the estimates of the short-term parameters and their standard errors, statistics, and confidence intervals; the coefficients of L. ce1, L. ce2, L. ce3, and L.ce4 are the parameters of the fit matrix α for this model. All the coefficients are significant at the 5% level, except those of BP, G500, Hidrosina, Petroseven, Repsol, Shell, and Total in the fourth cointegration equation, and Hidrosina, Pemex, Petroseven, and Total in the second. Using the previous notation, the following was estimated:

$$\hat{\alpha} = \begin{bmatrix} -.111 & -.059 & .114 & .038 & -.182 & -.017 & .042 & -.024 \\ .045 & .061 & -.036 & .030 & .039 & .021 & .060 & .036 \\ .036 & -.037 & -.076 & -.009 & -.024 & .001 & .026 & -.005 \\ -.019 & -.031 & .014 & -.078 & .056 & .003 & .013 & .037 \end{bmatrix}$$

$$\hat{B} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0.46 & -0.69 & -0.66 & -0.10 \\ 0 & 1 & 0 & 0 & 0.50 & -0.87 & -0.52 & -0.07 \\ 0 & 0 & 1 & 0 & -0.37 & 0.59 & -0.54 & -0.53 \\ 0 & 0 & 0 & 1 & -0.12 & -0.47 & -0.17 & -.018 \end{bmatrix}$$

$$\hat{v} = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0]$$

Table 7
Cointegration parameters

Johansen normalization restrictions imposed

beta	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_ce1						
lnBP	1	.	.	.	.	.
lnG500	-2.13e-17	.	.	.	.	.
lnHidrosina	4.34e-18	.	.	.	.	.
lnpemec	0	(omitted)				
lnPetroseven	.4698102	.0924485	5.08	0.000	.2886145	.651006
lnRepsol	-.6900985	.1216109	-5.67	0.000	-.9284516	-.4517455
lnShell	-.6696464	.070639	-9.48	0.000	-.8080963	-.5311966
lnTotalMxico	-.1006719	.0539324	-1.87	0.062	-.2063774	.0050337
_cons	-.0109377	.	.	.	.	.
_ce2						
lnBP	-2.08e-17	.	.	.	.	.
lnG500	1	.	.	.	.	.
lnHidrosina	-3.47e-18	.	.	.	.	.
lnpemec	2.78e-17	.	.	.	.	.
lnPetroseven	.5051339	.1059461	4.77	0.000	.2974834	.7127844
lnRepsol	-.8761395	.1393662	-6.29	0.000	-1.149292	-.6029867
lnShell	-.5255644	.0809523	-6.49	0.000	-.6842281	-.3669008
lnTotalMxico	-.0784803	.0618066	-1.27	0.204	-.199619	.0426584
_cons	-.0322456	.	.	.	.	.
_ce3						
lnBP	1.11e-16	.	.	.	.	.
lnG500	-5.55e-17	.	.	.	.	.
lnHidrosina	1	.	.	.	.	.
lnpemec	-2.22e-16	.	.	.	.	.
lnPetroseven	-.3720254	.231046	-1.61	0.107	-.8248673	.0808164
lnRepsol	.5927128	.3039283	1.95	0.051	-.0029757	1.188401
lnShell	-.5473304	.1765399	-3.10	0.002	-.8933423	-.2013186
lnTotalMxico	-.5381019	.1347871	-3.99	0.000	-.8022797	-.2739241
_cons	-.1761359	.	.	.	.	.
_ce4						
lnBP	-5.55e-17	.	.	.	.	.
lnG500	-2.78e-17	.	.	.	.	.
lnHidrosina	-2.78e-17	.	.	.	.	.
lnpemec	1	.	.	.	.	.
lnPetroseven	-.1279084	.1083442	-1.18	0.238	-.3402592	.0844423
lnRepsol	-.4791934	.1425208	-3.36	0.001	-.7585291	-.1998576
lnShell	-.1737006	.0827847	-2.10	0.036	-.3359557	-.0114455
lnTotalMxico	-.1888016	.0632056	-2.99	0.003	-.3126823	-.0649209
_cons	-.0361424	.	.	.	.	.

Source: own elaboration.

The estimation table (table 7) contains the cointegration vectors' estimated parameters for this model, together with their standard errors, statistics, and confidence intervals. It can be seen that all coefficients other than zero and one are statistically significant; the equations of cointegration can be expressed as follows:

$$lnBP = -0.46lnPetroseven + 0.69lnRepsol + 0.66InShell + 0.10InTotalMxico + .010$$

$$lnG500 = -0.50lnPetroseven + 0.87\,lnRepsol + 0.52InShell + 0.07InTotalMxico + 0.32$$

$$lnHidrosina = 0.37lnPetroseven\ - 0.59\,lnRepsol + 0.54lnShell + 0.53\,lnTotalMxico + .032$$

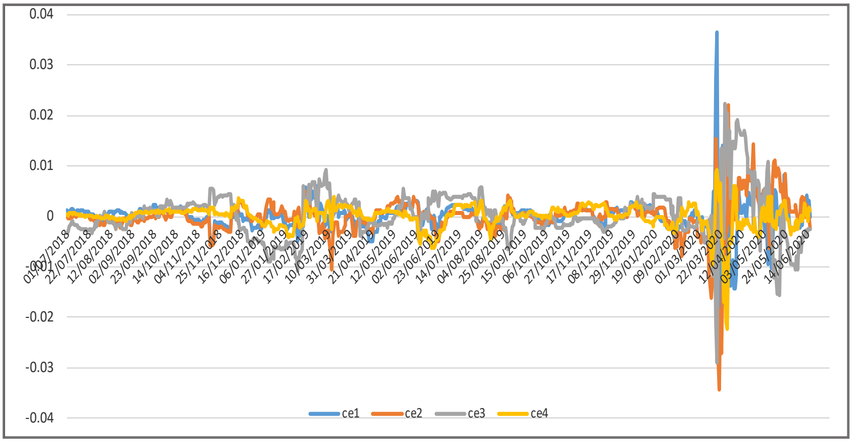
$$lnPemex = 0.12lnPetroseven + 0.47lnRepsol + 0.17InShell + 0.18InTotalMxico + .03$$

The three previous equations show that BP and G500 have the same sign for all the explanatory variables. It is almost the same for Hidrosina and Pemex, except for Repsol's sign in the Hidrosina equation.

Diagnosis of VECM

Next, to verify that the VECM is correctly specified, a set of diagnostic tests are performed.

Graph 3
Cointegration relationships



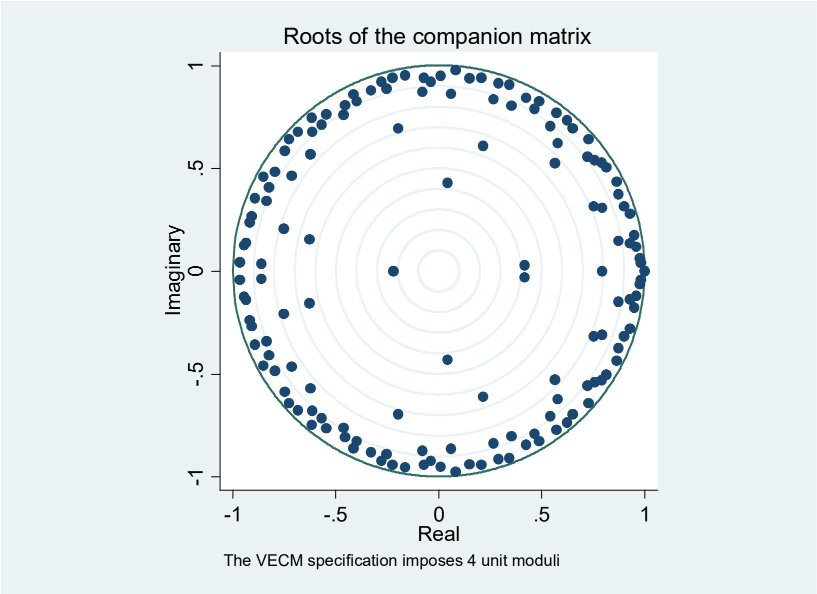
Source: own elaboration.

The zero mean lines represent a stable and long-term equilibrium relationship between the variables of each cointegration relationship (ce).

4. Stability test

If a VECM has *K* endogenous variables and *r* cointegrating vectors, there will be *K* - *r* unit modules in the complementary matrix; for the estimated model. It can be seen (graph 4) that 4-unit root modules are effectively imposed, the root of the other results of the residual stability test is less than 1, so the VECM model satisfies the stability condition.

Graph 4
Unit root test



Source: own elaboration.

Simultaneously, the serial correlation test (table 8) shows no serial correlation in the residuals at lag 18 (again, remember that the software automatically has taken the first difference of the logarithms of the variables, so we have lost one lag), according to the previous diagnoses; in general, the VECM model has good effects.

Table 8
Serial correlation test

Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	149.6380	64	0.00000
2	146.6746	64	0.00000
3	117.9828	64	0.00005
4	140.4341	64	0.00000
5	101.7981	64	0.00186
6	158.1344	64	0.00000
7	156.3072	64	0.00000
8	111.2937	64	0.00023
9	120.8812	64	0.00002
10	109.7630	64	0.00033
11	122.5160	64	0.00002
12	109.9605	64	0.00031
13	112.2318	64	0.00018
14	116.3093	64	0.00007
15	104.6957	64	0.00100
16	108.1592	64	0.00047
17	147.5436	64	0.00000
18	67.3307	64	0.36386
19	88.4906	64	0.02304

H0: no autocorrelation at lag order

Source: own elaboration.

Finally, we test for normality; as we can see in table 9, only BP is usually distributed.[8] The graph of the Standardized values of residuals (figure 2) is highly peaked and moderately skewed, but they are quasi-normal in general terms. Although the ideal would be to observe normality in the whole model, we can go ahead because of the following. First, this is so because the purpose of the model estimation is to examine the relationships between the variables and any long-term relationships between the series (and not forecasting).

Second, although Johansen indeed derived in 1988 Maximum likelihood estimation under the assumption of a normal likelihood (i.e., normal errors), later derives the large-sample distribution of his estimators under much broader moment conditions, thus (Johansen S., 2009) himself wrote: "*the asymptotic results available from the Gaussian analysis need not hold. Methods for checking vector autoregressive models, including a test for normality of residuals.*

[8] With a 91.5% of confidence.

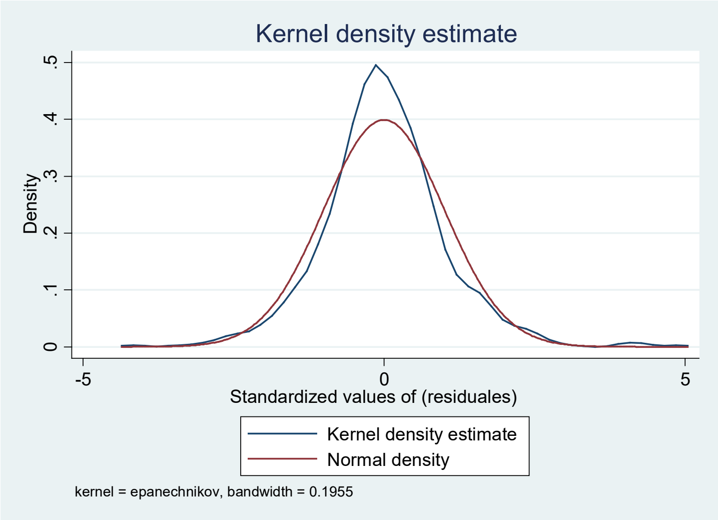
[...]Thus, the limit results hold for i.i.d. errors with finite variance, and not just for Gaussian errors". This means that no passing a test for normality has no implications on the validity of either tests or estimators in VECMs.

Table 9
Normality test

Jarque-Bera test

Equation	chi2	df	Prob > chi2
D_lnBP	8.292	2	0.01583
D_lnG500	308.059	2	0.00000
D_lnHidrosina	176.608	2	0.00000
D_lnpemec	111.236	2	0.00000
D_lnPetroseven	1284.567	2	0.00000
D_lnRepsol	999.051	2	0.00000
D_lnShell	73.266	2	0.00000
D_lnTotalMxico	326.643	2	0.00000
ALL	3287.722	16	0.00000

Source: own elaboration.



Source: own elaboration.

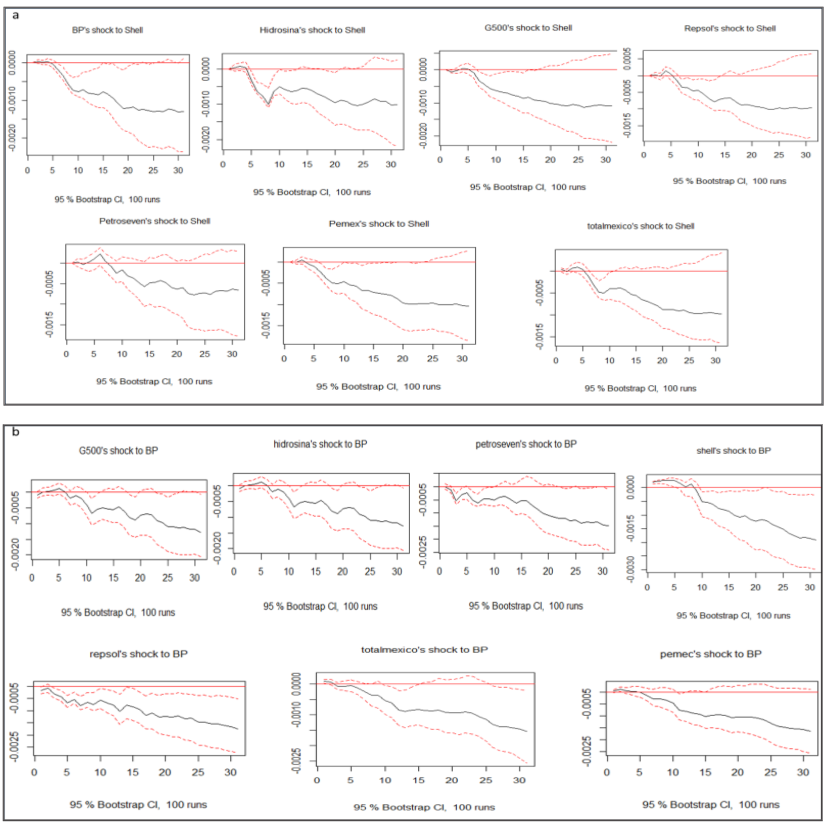
Figure 2
Standardized values of residuals vs. normal distribution

5. Impulse response function

With a model that is now acceptably well specified, we can use the impulse-response function. The impulse-response functions agglomerate the system's response to unanticipated shocks in the variables of the vector components. Thus, an alteration in a variable's behavior will directly affect it and be transmitted to the rest through the model's dynamic structure. Additional analysis is performed through the impulse response function based on the VECM, obtaining the results for 30 days.

It is essential to mention that variables modeled in a cointegration VECM are not reversible to the mean. The unit modules in the matrix imply that the effects of some shocks will not disappear with time. So when the impact of an impulse does not disappear over time, the result is permanent.

Graphs 5
a and *b*. Impulse response functions

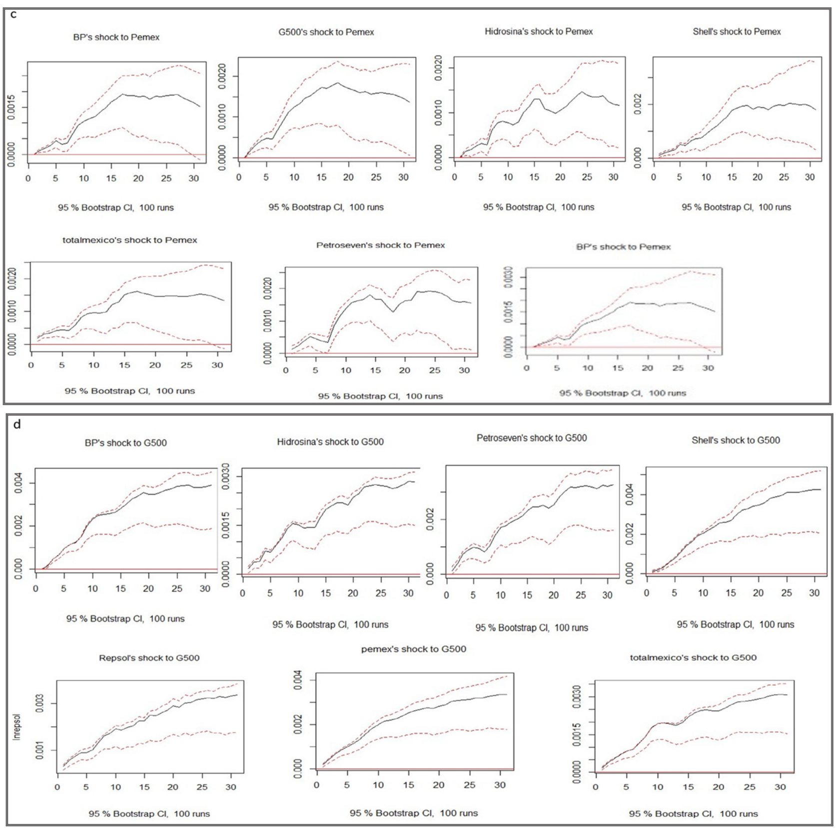


Source: own elaboration.

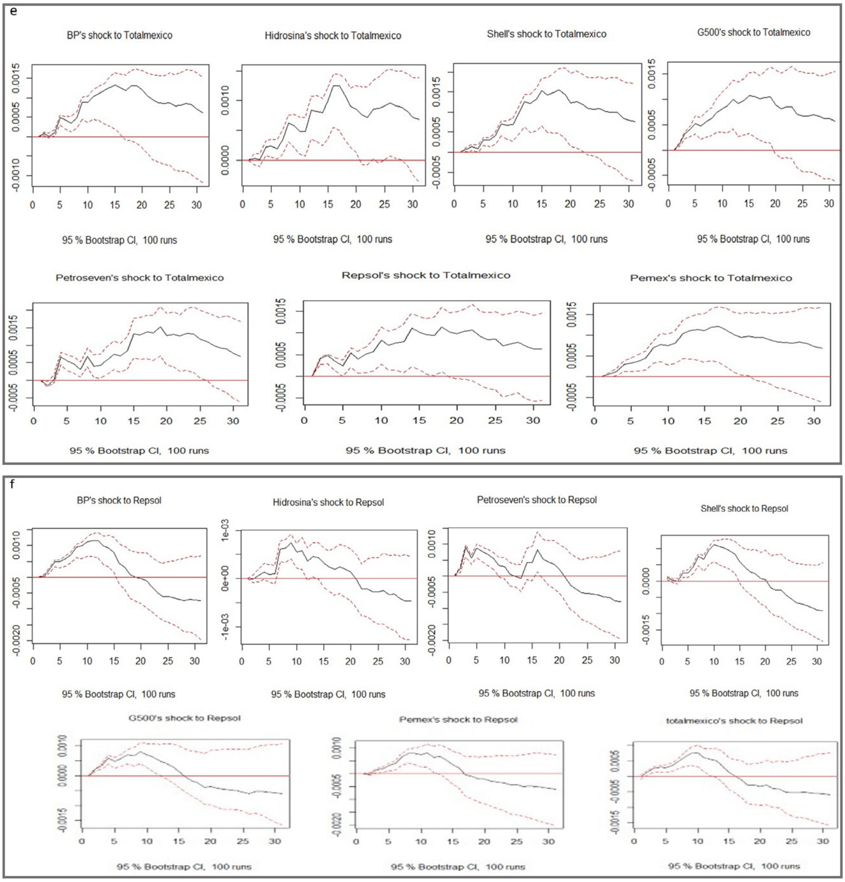
As shown in graphs 5 (*a*) and 5 (*b*), the charts indicate that orthogonal shock to the average price of Shell has a permanent effect on the price of the rest of the brands; this shock, together with those of BP are the only ones that cause the rest of prices to fall rapidly during the first days.

On the other hand, in graph 5(*c*), 5(*d*), 5(*e*), and 5(*f*), the charts indicate that an orthogonal shock to the average price of Pemex, G500, TotalMéxico, and Repsol, respectively has a permanent effect on the price of the rest of the brands, these shocks causes the rest of prices to rise during the first days. The case of Repsol stands out, where the impact is of shorter duration and the tendency to sap around 15 to 20 days.

Graphs 5
c and *d*. Impulse response functions



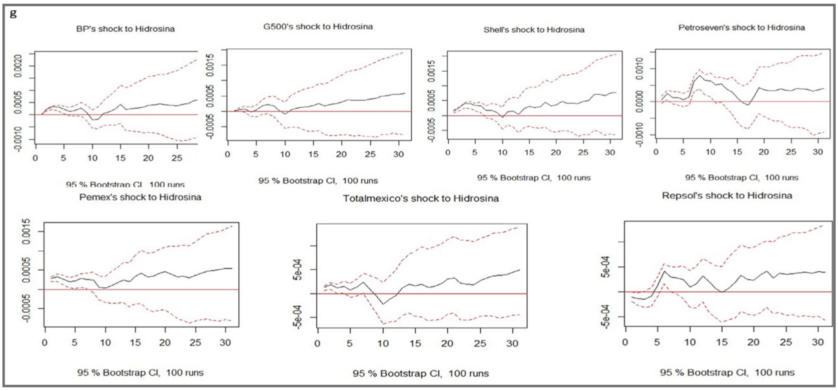
Graphs 5 e and f. Impulse response functions



Source: own elaboration

In graphs 5 (g), the charts indicate that an orthogonal shock to the average prices of Hidrosina has a particular effect on the price of the rest of the brands. These shocks cause the rest of the prices to rise slightly during the first days; unlike the following case (Petroseven), here, the effect seems to have a reversible impact on the mean in the long term.

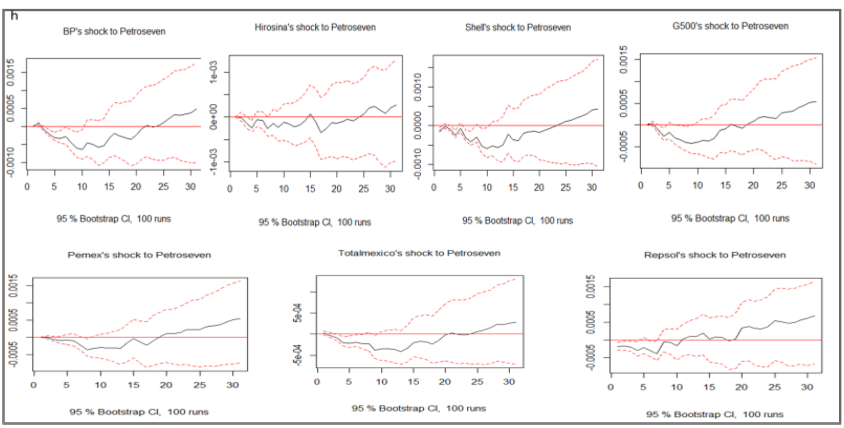
Graphs 5 g
Impulse response functions



Source: own elaboration.

Finally, in graphs 5(h), it is possible to see that an orthogonal shock to the average price of PetroSeven has a notable impact on the price of the rest of the brands. In general terms, all brands slightly reduce their prices in the first days. However, This effect changes direction around days 7 to 12.

Graphs 5 h
Impulse response functions

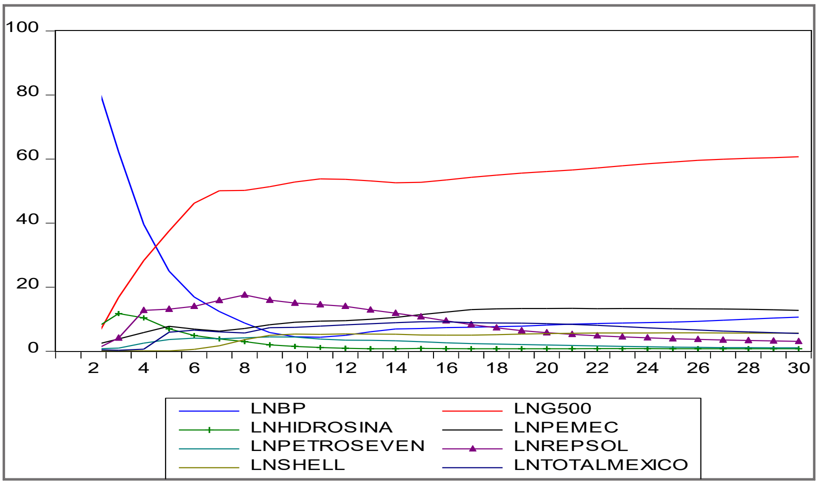


Source: own elaboration.

6. Variance decomposition

The decomposition of the variance refers to the decomposition of the mean square error in each variable's contributions. The variance decomposition can be applied to analyze the influence of each variable's innovation on other variables, which shows relative effects; if the contributions of its own disturbances explain a significant proportion of a variable's variance, it would be relatively more exogenous than others. Using the econometric software Eviews 10, the results of the variance decomposition for the first 30 days are obtained, as shown below:

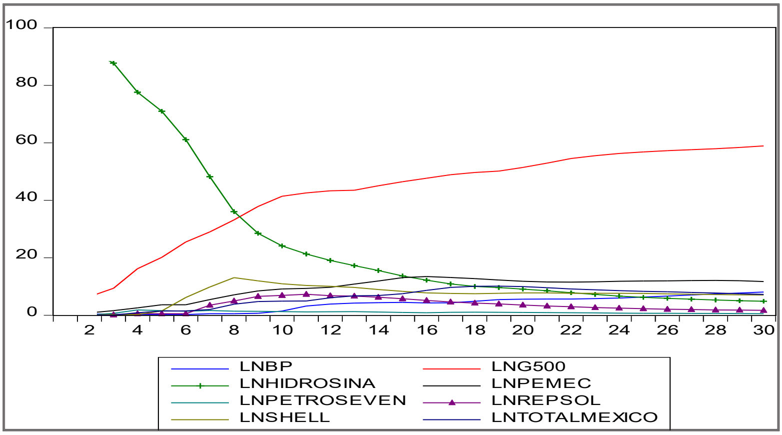
Graphs 6
Variance decomposition of BP.



Source: own elaboration

In graphs 6, it can be seen that in the first periods, the variance of BP strongly depends on itself but decreases rapidly; around day 6, the variance of G500 influences more than the same variance of BP, maintaining a certain dominance from that moment. (50%) in the dynamics of this variable. Pemex becomes the second relevant variable (10%) of variance from period 16.

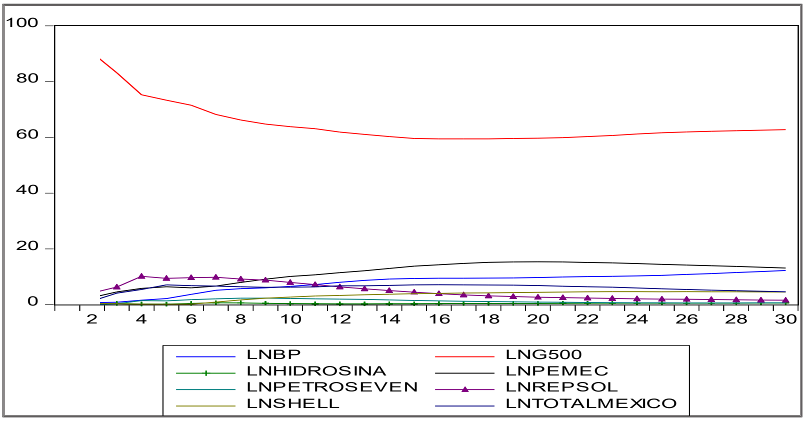
Graphs 7
Variance decomposition of hidrosina



Source: own elaboration.

In graphs 7, it can be seen that in the first periods, the variance of hidrosina strongly depends on itself but decreases rapidly; around day 9, the variance of G500 influences more than the same variance of BP, maintaining a certain dominance from that moment (50%) in the dynamics of this variable. Pemex becomes the second relevant variable (10%) of variance from period 16.

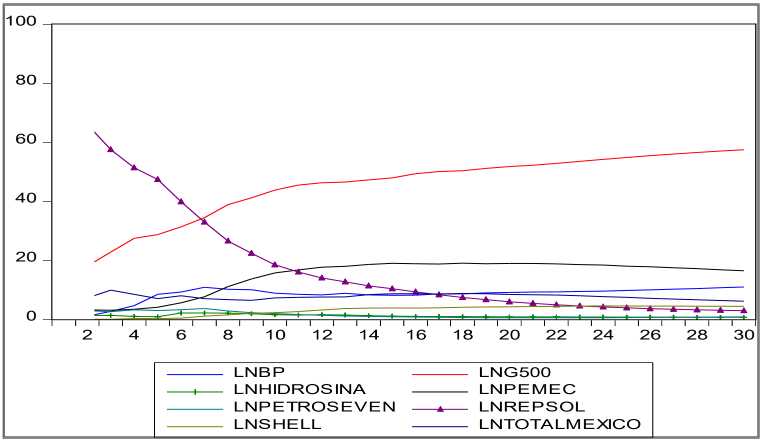
Graphs 8
Variance decomposition of G500



Source: own elaboration.

Graphs 8 shows that in the first periods, the variance of G500 strongly depends on itself and decreases slowly, maintaining a robust self-regressive behavior since, after 30 days, most of its variance is still self-explanatory (60%).

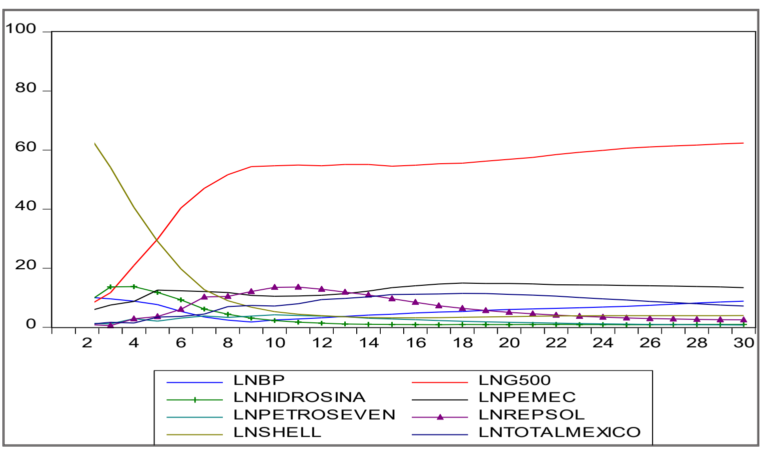
Graphs 9
Repsol variance decomposition



Source: own elaboration.

In graphs 9, it can be seen that in the first periods, the variance of Repsol is determined mainly by itself (around 80% for the first period) but gradually decreases. Around day 7, the variance of G500 influences more than the same variance of Repsol, maintaining from that moment on a certain dominance (50%) in this variable's dynamics.

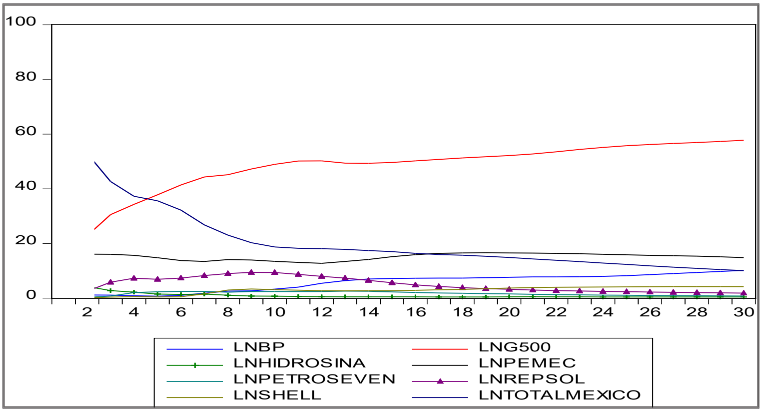
Graphs 10
Petroseven variance decomposition



Source: own elaboration

In graphs 10, it can be seen that in the first periods, the variance of PetroSeven depends strongly but decreases rapidly; around day 26, the variance of G500 influences more than the same variance of BP, maintaining a certain dominance from that moment in the dynamics of this variable. Pemex becomes the second relevant variable (20%) of variance from period 16.

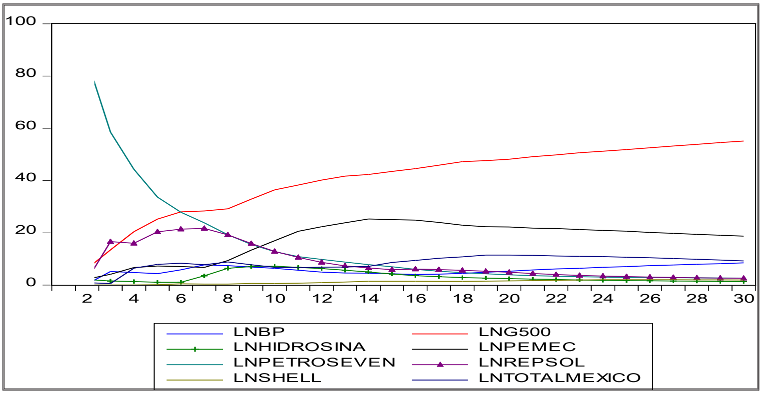
Graphs 11
Shell variance decomposition



Source: own elaboration.

In graphs 11, it can be seen that in the first periods, the variance of Shell is not determined solely by itself (around 75% for the first period) and gradually decreases. Around day 5, the variance of G500 influences more than the same variance of Shell, maintaining from that moment a certain dominance (60%) in the dynamics of this variable.

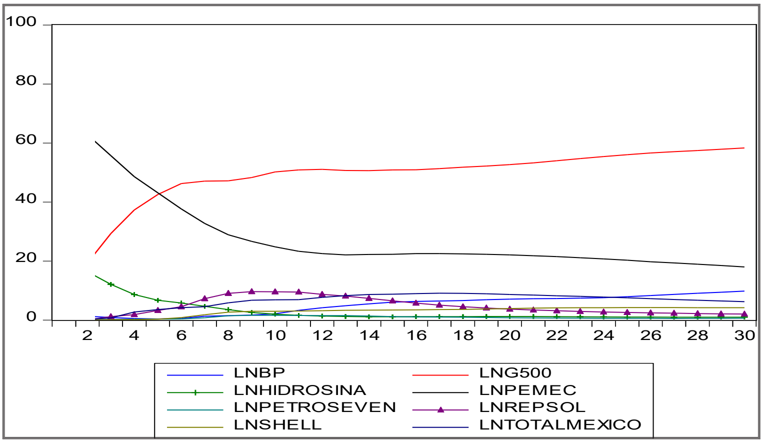
Graphs 12
Variance decomposition of total México



Source: own elaboration.

In graphs 12, it can be seen that in the first periods, the variance of total is not determined solely by itself (around 75% for the first period) and gradually decreases. Around day 5, the variance of G500 influences more than the same variance of total, maintaining from that moment a certain dominance (60%) in the dynamics of this variable.

Graphs 13
Variance decomposition of Pemex



Source: own elaboration.

In graphs 13, it can be seen that the variance of Pemex is not determined solely by itself (around 70% for the first period) and gradually decreases. Around day 7, the variance of G500 influences more than the same variance of Pemex, maintaining from that moment a certain dominance (50%) in the dynamics of said variable.

4. Conclusions

This document established a correlation model of the different gasoline prices in the northern region of Mexico City. It examined the causal relationships between eight different commercial brands. The exercise carried out attempts to capture the relationships between the cycles due to the decision-making its dynamization. The evidence of cointegration of at least four vectors between prices, contrary to a single vector of equilibrium prices, would be expected in a competitive market. Besides, these cointegration vectors acted to move without deviating too much from their selected long-term equilibria. At least four vectors can be explained by the patterns found by analyzing the impulse response function and the variance decomposition.

In this sense, according to the analysis of the impulse response function, four patterns were found: *i*) shocks (G500, Pemex, Repsol, and Total México) that would cause positive effects on the other prices, which would rise rapidly during the first days; *ii*) shocks (Shell and BP), which would cause a permanent negative impact on the price of the rest of the brands, which fell rapidly during the first days; *iii*) transitory shocks (Hidrosina) since its impact on other prices is of lesser magnitude and seems to be reversible to the mean (red line) in the long run, and *iv*) shocks (Petroseven) that would cause positive effects on the other prices, but during the first days the effect is negative.

Likewise, the variance decomposition analysis yielded exciting elements on the relative (short-term) dependence of each variable on the rest of the variables. According to the graphs (8), in the first days, G500 is explained by the contributions of their own disturbances. This variable would be relatively more exogenous than the others in the initial determination of their prices. This variable maintains a strong self-regressive behavior since, after 30 days, most of their variances continue to be self-explanatory. Graphs 6, 7, and 10 show that, even though the variances of BP, Hidrosina, and Petroseven are mainly determined by themselves in the first periods, but they are affected relatively quickly by the rest of the prices. On the other hand, in graphs 9, 11, 12, and 13, a third pattern is shown in the prices of Repsol, Shell, Total, and Pemex, consisting in that these prices are not initially explained only by the contributions of their own disturbances, and that they are quickly affected and explained by the rest of other prices.

Finally, the correlations between the different prices are definitely complex; thus, for example, the impact of G500 and Pemex prices have the most significant short-term impact (variance decomposition) on the rest of the prices, and both are the only ones whose shocks cause permanent positive effects (impulse-response) on the rest of prices. Regardless of the approach is too local (north of Mexico City), the results found indicate that it is perhaps still early to expect that in Mexico to have a unique equilibrium price vector derived from a competitive market and instead, there seems to be a certain presence of segmented markets in which each brand could be exercising specific market power.

According to the results obtained, we can summarize the most relevant findings through the following three conclusions:

- There are multiple influences and interdependencies among the prices of the eight brands analyzed.
- The gasoline market in these four mayors considered is segmented, and that the distributors exercise a certain level of market power.
- The results found in the analyzed period suggest that it is perhaps still early to expect that there will be an equilibrium price vector derived from a competitive market in Mexico.

This en lugar de The research does not pretend to be definitive but rather to contribute to the academic and regulatory discussion on the effectiveness of the deregulation of gasoline market prices. Future research could prove or rule out collusion agreements' possible existence and analyze the cost structure (storage and transportation) and location patterns to conclude competitiveness.

References

Akinboade, O.; E. Ziramba & W. Kumo (2008). "The Demand for Gasoline in South Africa: An Empirical Analysis Using Co-integration Techniques. *Energy Economics*, 3222-3229. https://doi.org/10.1016/j.eneco.2008.05.002.

Bentzen, J. (1994). An empirical analysis of gasoline demand in Denmark using cointegration techniques. *Energy Economics,* 139-143. https://doi.org/10.1016/0140-9883(94)90008-6.

Cheung, K.-Y., & E. Thomson (2004). The demand for gasoline in China: a cointegration analysis. *Journal of Applied Statistics,* 533–544. https://doi.org/10.1080/0266476 0410001681837.

Eltony, M., & N. Mutairi (1995). Demand for gasoline in Kuwait: an empirical analysis using cointegration techniques. *Energy Economics,* 249–253. https://doi.org/10.1016/0140-9883(95)00006-g.

Engle, R., & C. Granger (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica*, 251-276. https://doi.org/10.2307/1913236.

Ferrer, C., & S. Escalante (2014). Demanda de gasolina en la zona metropolitana del Valle de México: análisis empírico de la reducción del subsidio. *Revista de Economía del Rosario*, 89-117. https://doi.org/10.12804/rev.econ.rosario.17.01.2014.04.

Granger, C. (1980). Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control* 2, 329-352.

Granger, C., & P. Newbold (1974). Spurious regressions in econometrics. *Journal of Econometrics* 2, 111-120. https://doi.org/10.1016/0304-4076(74)90034-7.

Ibarra Salazar, J., & L. Sotres Cervantes (2008). La demanda de gasolina en México *El efecto en la frontera norte*. *Frontera Norte*, 131-156. https://doi.org/10.2307/j.ctv26d8jg.8.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 231-254. https://doi.org/10.1016/0165-1889(88)90041-3.

___ (2009). Cointegration: Overview and Development. *Handbook of Financial Time Series*, 671-693. https://doi.org/10.1007/978-3-540-71297-8_29.

Khim, V., & S. Liew (2004). Which Lag Length Selection Criteria Should We Employ? *Economics Bulletin*, 1-9.

Liew, V. (2004). Which Lag Selection Criteria Should We Employ? *Economics Bulletin*, 1-8.

Nielsen, B. (2001). Order determination in general vector autoregressions. *Economics Papers* 2001-W10, Economics Group.

Rao, B., & G. Rao (2009). Cointegration and the Demand for Gasoline. *Energy Policy*, 3978-3983. https://doi.org/10.1016/j.enpol.2009.04.046.

Reyes, O.; R.,Escalante, & A. Matas (2010). La demanda de gasolinas en México: Efectos y alternativas ante el cambio climático. *Economía teoría y práctica,* 23-110. https://doi.org/10.24275/etypuam/ne/322010/reyes.

Sims, C. (1980). Macroeconomics and Reality. *Econometrica*, 1-48.

Wooldridge, J. (2010). *Introducción a la econoemtría: un enfoque moderno*. Ciudad de México: Cengage Learning.

Annex I. Table A. Short-term estimation of the VECM Model

	Coef.	Std.Err.	z	P>z	[95%	Conf. Interval]
D_lnBP						
_ce1						
L1.	-0.1117521	0.0320738	-3.48	0	-0.1746156	-0.0488886
_ce2						
L1.	0.0452998	0.0238155	1.9	0.057	-0.0013777	0.0919773
_ce3						
L1.	0.0364468	0.0120673	3.02	0.003	0.0127952	0.0600983
_ce4						
L1.	-0.019737	0.0254746	-0.77	0.438	-0.0696664	0.0301923
D_lnG500						
_ce1						
L1.	-0.0592005	0.0311172	-1.9	0.057	-0.1201892	0.0017882
_ce2						
L1.	0.0618951	0.0231052	2.68	0.007	0.0166097	0.1071805
_ce3						
L1.	0.0379989	0.0117074	3.25	0.001	0.0150527	0.060945
_ce4						
L1.	-0.0311673	0.0247149	-1.26	0.207	-0.0796076	0.017273

Annex I.

	Coef.	Std.Err.	z	P>z	[95%	Conf. Interval]
D_lnHidrosina						
_ce1						
L1.	0.1145927	0.0619821	1.85	0.064	-0.0068899	0.2360753
_ce2						
L1.	-0.0365454	0.046023	-0.79	0.427	-0.1267489	0.053658
_ce3						
L1.	-0.0765967	0.0233199	-3.28	0.001	-0.1223029	-0.0308905
_ce4						
L1.	0.0140954	0.0492293	0.29	0.775	-0.0823923	0.110583
D_lnpemec						
_ce1						
L1.	0.0387671	0.0343864	1.13	0.26	-0.0286291	0.1061632
_ce2						
L1.	0.0302789	0.0255327	1.19	0.236	-0.0197642	0.0803219
_ce3						
L1.	-0.0099578	0.0129374	-0.77	0.441	-0.0353146	0.0153991
_ce4						
L1.	-0.0786554	0.0273114	-2.88	0.004	-0.1321848	-0.025126
D_lnPetroseven						
_ce1						
L1.	-0.1828542	0.0660688	-2.77	0.006	-0.3123467	-0.0533616
_ce2						
L1.	0.0396893	0.0490575	0.81	0.418	-0.0564617	0.1358403
_ce3						
L1.	-0.0242123	0.0248575	-0.97	0.33	-0.0729321	0.0245075
_ce4						
L1.	0.0560364	0.0524752	1.07	0.286	-0.0468131	0.1588858
D_lnRepsol						
_ce1						
L1.	-0.0177115	0.0539826	-0.33	0.743	-0.1235156	0.0880925
_ce2						
L1.	0.2101264	0.0400833	5.24	0	0.1315646	0.2886882
_ce3						
L1.	0.0012731	0.0203102	0.06	0.95	-0.0385342	0.0410805

Conclution. Annex I.

	Coef.	Std.Err.	z	P>z	[95%	Conf. Interval]
_ce4						
L1.	0.0035629	0.0428757	0.08	0.934	-0.080472	0.0875978
D_lnShell						
_ce1						
L1.	0.0423335	0.0344683	1.23	0.219	-0.0252231	0.1098901
_ce2						
L1.	0.0602945	0.0255934	2.36	0.018	0.0101323	0.1104567
_ce3						
L1.	0.0263239	0.0129682	2.03	0.042	0.0009066	0.0517412
_ce4						
L1.	0.0138168	0.0273764	0.5	0.614	-0.0398401	0.0674736
D_lnTotal-Mxico						
_ce1						
L1.	-0.0249394	0.0298662	-0.84	0.404	-0.0834762	0.0335974
_ce2						
L1.	0.0368087	0.0221763	1.66	0.097	-0.0066561	0.0802736
_ce3						
L1.	-0.0005896	0.0112368	-0.05	0.958	-0.0226133	0.0214341
_ce4						
L1.	0.0376855	0.0237213	1.59	0.112	-0.0088074	0.0841783

Annex II. Sample

Numer de Permiso
PL/3201/EXP/ES/2015
PL/3211/EXP/ES/2015
PL/12734/EXP/ES/2015
PL/2724/EXP/ES/2015
PL/3626/EXP/ES/2015
PL/537/EXP/ES/2015
PL/2092/EXP/ES/2015
PL/435/EXP/ES/2015
PL/7540/EXP/ES/2015
PL/843/EXP/ES/2015
PL/8538/EXP/ES/2015
PL/8584/EXP/ES/2015
PL/9102/EXP/ES/2015
PL/9257/EXP/ES/2015
PL/6565/EXP/ES/2015
PL/5601/EXP/ES/2015
PL/10857/EXP/ES/2015
PL/8771/EXP/ES/2015
PL/1059/EXP/ES/2015
PL/6427/EXP/ES/2015
PL/7904/EXP/ES/2015
PL/4614/EXP/ES/2015
PL/5299/EXP/ES/2015
PL/7717/EXP/ES/2015
PL/8770/EXP/ES/2015
PL/1467/EXP/ES/2015
PL/1308/EXP/ES/2015
PL/988/EXP/ES/2015

All prices were collected from the Energy Regulatory Commission website, available at the following link: https://www.cre.gob.mx//ConsultaPrecios/GasolinasyDiesel/GasolinasyDiesel.html.