Economic Analysis and Policy 78 (2023) 648–660

Contents lists available at ScienceDirect

Economic Analysis and Policy

journal homepage: www.elsevier.com/locate/eap

Analyses of topical policy issues

US biofuel market persistence and mean reversion properties[☆]

Miguel A. Martin-Valmayor [a,b,*], Luis A. Gil-Alana [c,a], Asís Pardo Martín [d]

[a] Universidad Francisco de Vitoria, Madrid, Spain
[b] Universidad Complutense de Madrid, Madrid, Spain
[c] University of Navarra, Pamplona, Spain
[d] IESE Business School, Universidad de Navarra, Spain

ARTICLE INFO

Article history:
Received 22 March 2023
Received in revised form 8 April 2023
Accepted 9 April 2023
Available online 11 April 2023

JEL classification:
C22
G15

Keywords:
Bioethanol
Corn
Fractional integration
Long memory
Time series analysis

ABSTRACT

This paper investigates the market persistence and mean reversion properties for corn, bioethanol and gasoline prices in the US biofuel industry, evaluating long memory effects with fractional integration techniques from January 1982 to May 2022 with USDA data. Empirical results show evidence of no mean reversion properties for the prices in the three series though some support of it is found when the differences of bioethanol and gasoline are taken with respect to corn. Thus, external shocks in the original series are expected to remain persistent and would require additional policy measures to recover the original trend. Furthermore, the impact of Covid on the time series has been analyzed by comparing the scenarios pre and post pandemic, finding evidence of no major changes in the integration orders in all the series under analysis.

1. Introduction

The use of biofuels as a vehicle fuel has significantly reduced the amount of greenhouse gas emissions in recent times (Bouri et al., 2021). Further points to consider regarding biofuel usage is the decrease in crude oil prices and the level of dependency large biofuel producer countries have on it. However, the possibility of price increases in biofuel feedstocks such as sugar, corn or soybeans is a potential impediment for the rest of the economy. In this paper we check mean reversion and market persistence properties of bioethanol and gasoline, and their impact on the feedstocks in the US, the largest biofuel producer.

By way of a brief introduction, the most common biofuels are bioethanol, synthesized from carbohydrates (normally coarse grain or sugarcane); and biodiesel, usually generated from fats and oils, typically vegetable oil such as oil from soybeans. Biodiesel is the dominant biofuel produced in Europe and Asia Pacific, while bioethanol is the main one in North America and South and Central America. The BP Statistical Review of World Energy (2021) identified that global biofuel production fell by 6% worldwide in 2020. As for the previous 10 years, it had been undergoing a 6% compound average growth and the industry is questioning if production levels had reached saturation point. With regards to the biofuel and gasoline relationship, Barros et al. (2014) investigated this market in Brazil (2000–2012) with two interesting

☆ Luis A. Gil-Alana and Miguel A. Martin-Valmayor gratefully acknowledges financial support from the MICIN-AEI-FEDER PID2020-113691RB-I00 project from 'Ministerio de Ciencia e Innovación' (MICIN), Spain, 'Agencia Estatal de Investigación' (AEI) Spain and 'Fondo Europeo de Desarrollo Regional' (FEDER).

* Correspondence to: Universidad Francisco de Vitoria, Facultad de Derecho, Empresa y Gobierno, E-28223 Pozuelo de Alarcon - Madrid, Spain.
E-mail addresses: miguel.martin@ufv.es, mamart02@ucm.es (M.A. Martin-Valmayor).

https://doi.org/10.1016/j.eap.2023.04.008
0313-5926/

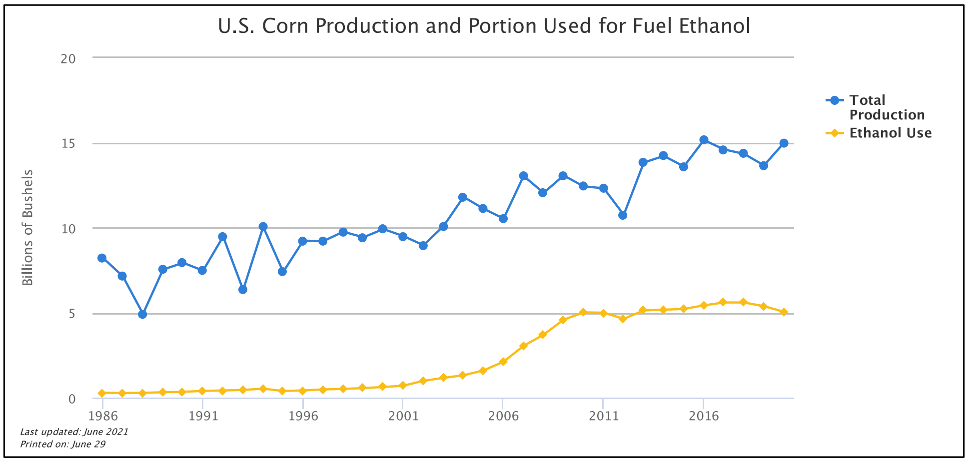

Fig. 1. U.S. Corn Production and Portion Used for Fuel Ethanol.
Source: Taken from AFDC, US department of energy https://afdc.energy.gov/data/.

findings: First, that ethanol/gasoline prices tend to grow steadily over time while consumption does not; and second, they found that individually ethanol and gasoline prices were time series with integration orders smaller than 1, while the joint ratio ethanol/gasoline had orders of integration clearly above 1. Thus, shocks for individual series were expected to disappear in the long run, but the joint relationship leads to shocks having a permanent effect that might affect other parts of the economy.

In terms of the global size of the biofuel market, the United States is today the largest producer of biofuel, as it is for crude oil. In 2020, total production was about 1653 thousand barrels of oil equivalent per day (beq/day), of which the US production represents about 602 thousand beq/day (Sönnichsen, 2022), and only 3.65% of the total crude oil produced in the US (BP, 2021). Brazil and Indonesia ranked second and third, with production in thousand beq/day of 395 and 126 respectively. Before 2007, the US was not the global leader in either market, but after this financial crisis, the US followed an energy strategy of developing local technologies (such as fracking) to extract more crude oil and to develop internal oil alternatives to reduce the external dependency on crude oil fuel. In fact, the US bioethanol production increased from 91 million gallons in 2005 to reach between 1500 and 1800 million gallons per year after 2016. However, from this year onwards the production tends to remain stable with a production capacity amounting to about 17 billion gallons per year, with over 200 bio-refineries across the US (Sönnichsen, 2022).

The work of Dutta (2019) focuses specifically on the modelization of US ethanol pricing, claiming that this issue has not received great attention in the biofuel literature. They proposed a corn price implied volatility (CIV) index, with a GARCH approach to explain the variation in the ethanol price from corn grain, this being its main feedstock. Sometime before, Abbott (2013) investigated the role of biofuels in determining rises in corn and other agricultural commodity prices, concluding that the contribution of biofuels to high corn prices remained debatable. In fact, the USDA[1] reported a very stable corn price at $3/bushel in the 2015–2019 period, but after year 2020 and the COVID pandemic, it rose sharply to $6/bushel questioning again if there was a capacity problem and if biofuel corn feeding could cause problems for the US corn market. After the rise of US corn production to produce ethanol (2005 to 2010), the distribution of US biofuel-corn usage (see, Fig. 1) appears to remain constant at 5 bill. bushels, with a more volatile non-biofuel corn at around 10 bill. bushels and this large and it is unclear if this persistent new demand for corn has definitively changed the US corn and biofuel market price dynamics.

The objective of this paper is to complement these previous studies by determining the impact of the biofuel industry on corn and gasoline pricing in the US, this being the largest current biofuel producer. The proposed approach of the paper is to study the market persistence and mean reversion properties for the corn, bioethanol and gasoline prices, evaluating the long memory effects with fractional integration techniques. This methodology is very appropriate for the analysis of persistence and to evaluate the nature of the shocks, being more flexible and general than the traditional methods that focus exclusively on integer degrees of differentiation.

The rest of the paper is structured as follows: Section 2 includes a short literature review on biofuel, persistence, and mean reversion properties; Section 3 explains the methodology; Section 4 describes the datasheet and sources, while Section 5 displays the main results. Section 6 develops the main conclusions of the paper and outlines further lines of research.

1 https://www.ers.usda.gov/topics/farm-economy/bioenergy/data/.

2. Literature review

The analysis of the bioethanol price persistence literature is a very specific topic that has not been particularly popular in recent times. For the specific case of the US, prior studies to Serra et al. (2011), were mainly theoretical studies with some empirical simulations. This study was the first to use empirical analysis such as VECM (vector error correction model) to assess the price relationships within the US ethanol industry with monthly prices (1990 to 2008), finding evidence of long-run relationships among the prices analyzed and strong links between energy and food prices. More focused on food sustainability, To and Grafton (2015) estimated autoregressive models of global food prices in terms of US oil prices, GDP per capita and biofuel production, finding evidence of a significant effect of biofuel production on US food prices. Chiu et al. (2016) explored the relationship between the prices of crude oil, corn, and ethanol (1986–2015) using a VAR, VECM and autoregressive distributed lag to explore the connections between crude oil, corn, and ethanol markets in the context of the US, finding evidence that corn prices were driven by ethanol prices, however corn prices did not influence ethanol prices until 2005 and structural breaks were endogenously determined. Pal and Mitra (2017) analyzed US diesel and soybean prices (2004–2014) with quantile autoregressive distributed lag models with monthly samples, finding evidence of strong links between diesel and soybean prices over the long run. Morris et al. (2017) studied the US retail gasoline pricing, in terms of crude oil and corn-ethanol with a polynomial distributed lag to a price transmission model, finding evidence of asymmetry between wholesale and retail gasoline prices but lack of asymmetry between ethanol and gasoline prices. Saghaian et al. (2018) analyzed if corn-ethanol prices could lead to volatility spillovers between food and energy prices using a BEKK-multivariate-GARCH approaches with daily, weekly, and monthly futures prices (2007–2015), finding also evidence of asymmetric volatility transmission between corn and ethanol prices.

With regards to other markets, Kristoufek et al. (2014) analyzed the biofuel price transmission with price cross-elasticities combining generalized least squares estimation for EU industry and Germany markets, with evidence that both ethanol and biodiesel prices were responsive to their production factors (ethanol to corn, and biodiesel to German diesel). For the Brazilian case, Du and Carriquiry (2013) analyzed biofuel and sugar dynamics in the context of Brazilian flex-fuel vehicles (FFVs), with evidence that price dynamics were largely determined by market factors and prices exhibited strong mean-reversion, in line with the previously mentioned work of Barros et al. (2014) that concluded that when considered individually ethanol pricing was mean reverting. Finally, David et al. (2018) studied the fractional dynamic behavior of the ethanol prices on the Brazilian spot price market, with the application of fractional integration tools (ARIMA and ARFIMA models, and the Hurst and Lyapunov exponents), concluding that ethanol prices series were anti-persistent, and with price shocks expected to dissipate in the long run.

Finally, regarding recent specific biofuel feedstocks studies, Etienne et al. (2016) proposed an investigation concerning the price and volatility transmission between natural gas, fertilizers, and corn markets (1994–2014), using multivariate GARCH models with evidence of a mild linkage between prices and volatility between corn markets and natural gas in the long term (1994–2014) although no linkage in the most recent period (2006–2014) was observed. Musunuru (2016) used GARCH models with its variations EGARCH, TGARCH and APARCH to examine the presence of volatility persistence in soybean futures data and concluded that soybean return series exhibit volatility characteristics typical of a financial time series. Oláh et al. (2017) investigated the impact of biofuel production on the increased volatility of food prices, finding evidence of correlations between cereal, sugar and vegetable oil price indexes and crude oil prices from 2003 to 2016, however their findings show that the main driver for food price fluctuation was mainly the oil price shock. Barros et al. (2018) studied the dynamics of sugarcane production in the context of biofuel production in Brazil, with evidence that the total production is highly persistent, with an integration order smaller than 1 but close to it, with shocks expecting to have a permanent nature and requiring policy measures to recover the level from exogenous shocks.

3. Data description

To obtain ethanol pricing data, we follow the traditional approach of Serra et al. (2011), following the U.S. ethanol spot price taken from USDA. Thus, we select monthly data (1982–2022) from Corn, Ethanol and Gasoline from Feed Grains Data (Yearbook Tables) USDA, ERS Feed Grains Database page https://www.ers.usda.gov/data-products/feed-grains-database/. In particular Corn prices ($/bushel) were chosen as the ones received by farmers; Ethanol ($/gallon) and gasoline ($/gallon) were rack prices for wholesale truckload transport FOB (free on board). Fig. 2 summarizes this data comparing the price of the Corn and the Ethanol commodities, and Table 1 shows the main statistical descriptors of these values over the whole period examined.

Moreover, as global markets quote these commodities, in order to evaluate the frequency impact weekly prices have been taken from the Bloomberg Commodities database for Corn commodity (C1 COMB Comdty 1972–2022); Ethanol (CUA1 COMB Comdty 2005–2022) and Gasoline (XB1 COMB Comdty 2005–2022). In comparison with previous USDA data, these indexes did not use the same units; however, data has been useful for analytical purposes in terms of persistence and mean reversion properties. Table 1 summarizes the main statistics of both sources of monthly (USDA) and weekly (Bloomberg) data with evidence of a smaller volatility coefficient (standard deviation/average) on the weekly data starting after 2005.

For the USDA datasheet, a market correlation coefficient between ethanol and gasoline measured as $\frac{Covariance\,(Ethanol, Gasoline)}{Desvstd\,(ethanol) Desvstd\,(gasoline)}$ has been taken for the whole period, with a value of 0.793. This value, positive and less than

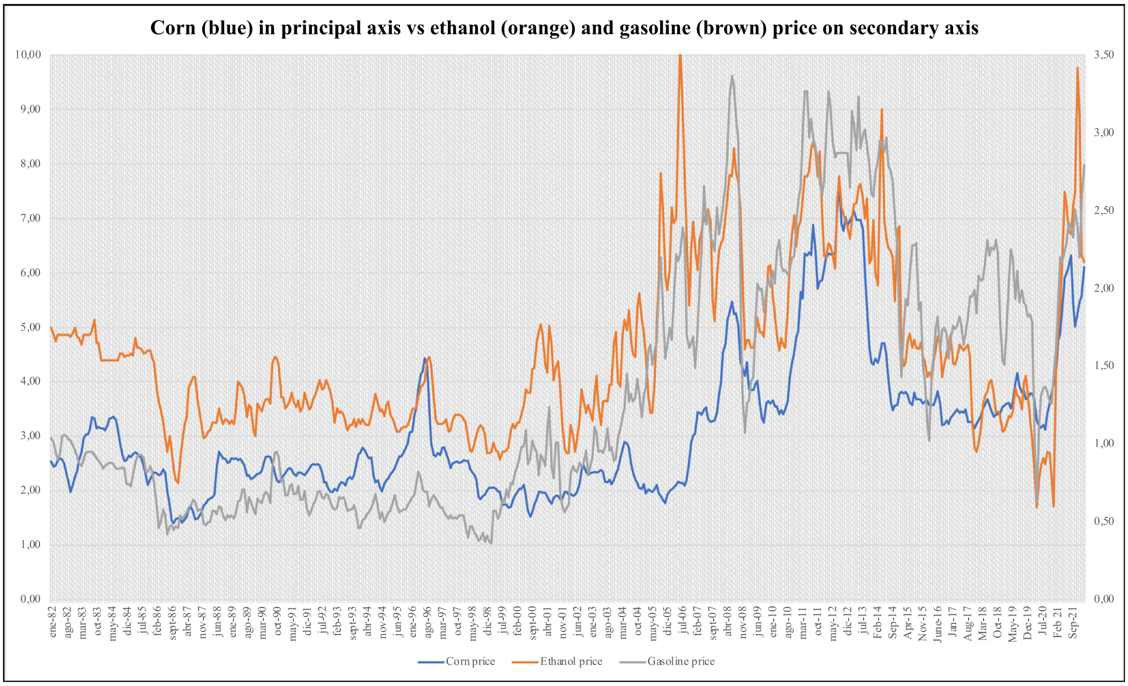

Fig. 2. Corn and Ethanol/Gasoline price comparison. Monthly data taken from USDA.

Table 1
Main statistic descriptors of the datasheet.

	Monthly data (USDA) 1982–2022			Weekly data (Bloomberg) (2005–2022)		
	Corn ($/Bushel)	Ethanol ($/Gall)	Gasoline ($/Gall)	Corn (CUCOMB)	Ethanol (CUA1 COMB)	Gasoline (XB1 COMB)
MIN	1,40	0,59	0,36	189,75	0,84	57,37
MAX	7,63	3,58	3,37	824,50	3,45	401,58
AVERAGE	3,11	1,58	1,34	448,40	1,90	209,15
STDEV	1,29	0,54	0,81	143,98	0,48	62,79
STDEV/AVERAGE	0,415	0,339	0,607	0,32	0,25	0,30

1 could mean that Ethanol prices should rise when Gasoline tends to rise, but not as fast. However, by measuring a time series beta with monthly samples and different time spans, it can be seen that this positive contribution is not constant. Fig. 3a shows this monthly beta for different time spans (1 yr, 2 yr and 3 yr) for monthly data and where positive values are mostly typical, however there is evidence of negative periods, especially in recent times (2017–2019) where Gasoline price tends to rise but the Ethanol prices drop. Fig. 3b uses a higher frequency with weekly data, with evidence of more fluctuations especially with smaller window spans.

4. Methodology

The methodology used in the paper is based on the concept of long memory that implies that the infinite sum of the autocovariances of a stationary process is infinite. That is, defining $\gamma_u = E\left[(x(t) - Ex(t))(x(t+u) - Ex(t))\right]$, then, x(t) displays the property of long memory if:

$$\sum_{u=-\infty}^{\infty} \gamma_u = \infty. \tag{1}$$

Within the long memory class of models, fractional integration is a special category, that is characterized because the number of differences required in a series to render it stationary I(0) may be a fractional value. In other words, x(t) is integrated of order d, and denoted as I(d) if:

$$(1-B)^d x_t = u_t, \qquad t = 1, 2, \ldots, \tag{2}$$

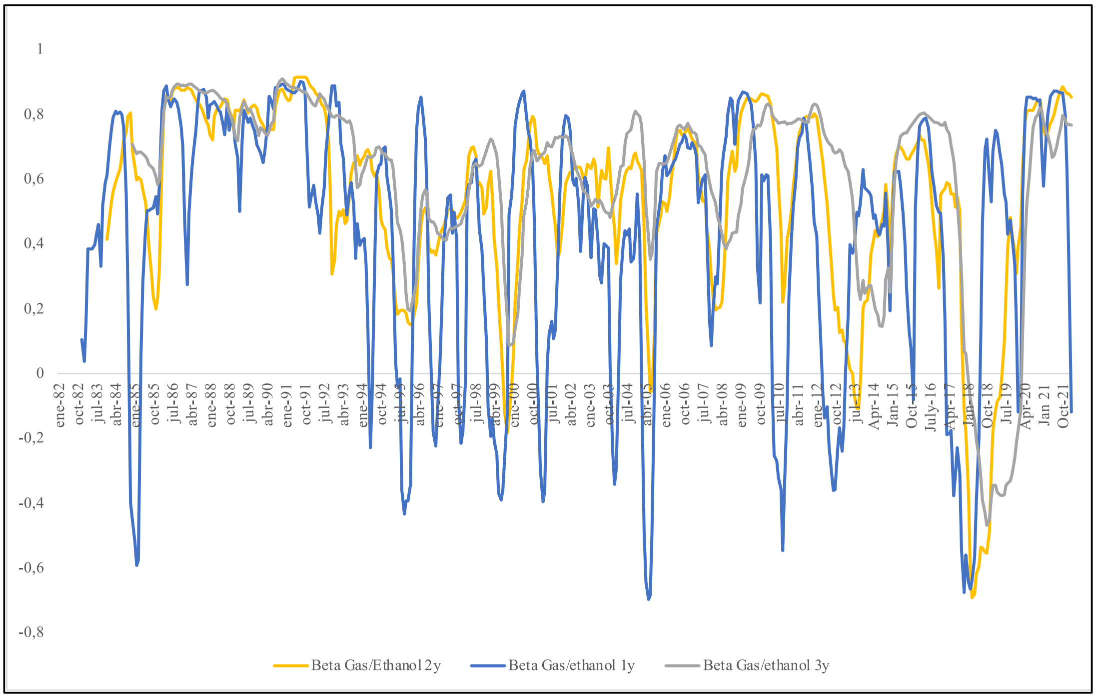

Fig. 3a. Correlation coefficient $\left(\frac{Covar(eth;gas)}{Desvest_{eth}Desvest_{gasoline}}\right)$ between Ethanol and Gasoline with different time spans with monthly samples (USDA).

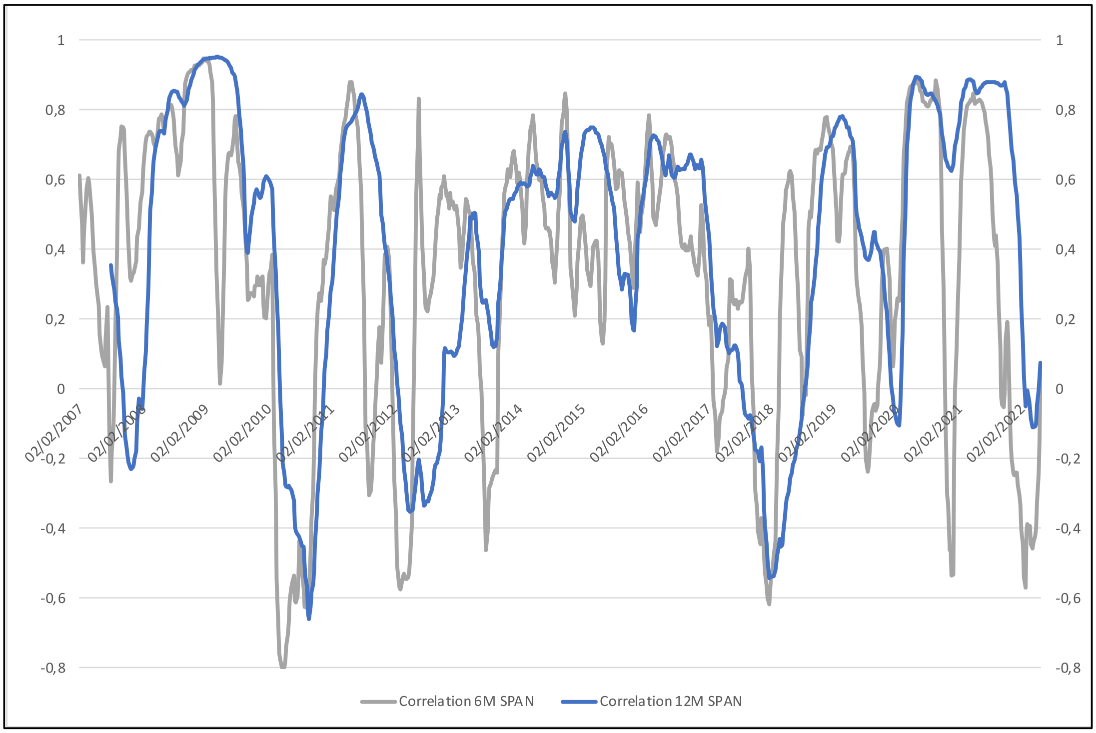

Fig. 3b. Correlation coefficient $\left(\frac{Covar(eth;gas)}{Desvest_{eth}Desvest_{gasoline}}\right)$ between Ethanol and Gasoline with different time spans with weekly samples (Bloomberg).

Table 2
Estimates of d with monthly data.

(i) Original data			
Series (monthly)	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	1.21 (1.13, 1.30)	**1.30 (1.21, 1.40)**	1.30 (1.21, 1.40)
Ethanol ($/Gall)	1.01 (0.94, 1.10)	**1.00 (0.92, 1.11)**	1.00 (0.92, 1.11)
Gasoline ($/Gall)	1.07 (0.97, 1.18)	**1.11 (1.00, 1.23)**	1.11 (1.00, 1.23)
(ii) Logged data			
Series (monthly)	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	1.19 (1.11, 1.29)	**1.37 (1.27, 1.49)**	1.37 (1.27, 1.49)
Ethanol ($/Gall)	1.06 (0.96, 1.18)	**1.06 (0.96, 1.18)**	1.06 (0.96, 1.18)
Gasoline ($/Gall)	0.98 (0.90, 1.09)	**0.98 (0.90, 1.09)**	0.98 (0.90, 1.09)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

where B is the backshift operator and u_t is I(0), which is defined as a process where the infinite sum of its autocovariances is finite, i.e.,

$$\sum_{u=-\infty}^{\infty} \gamma_u < \infty. \tag{3}$$

The long memory feature is made explicit by mean of the Binomial expansion of the polynomial above in Eq. (2), noting that, for all real d,

$$(1-B)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j B^j = 1 - dB + \frac{d(d-1)}{2} B^2 - \cdots \tag{4}$$

and thus, the equation given in (2) can be expressed as

$$\mathrm{x}_t = \mathrm{d}\,\mathrm{x}_{t-1} - \frac{\mathrm{d}(\mathrm{d}-1)}{2}\mathrm{x}_{t-2} + \cdots + \mathrm{u}_t. \tag{5}$$

In this context, if d is a non-integer value, x_t will be a function of all its past history, and higher the value of d is, the higher the degree of dependence is between the data. The estimation of the differencing parameter is based on a frequency domain version of the Whittle function as proposed in Dahlhaus (1989) and as in Robinson (1994). In fact, we use the latter approach which is testing procedure based on the Lagrange Multiplier (LM) principle and that has numerous advantages in relation with other methods. In particular, it allows us consider values of the differencing parameter, d, which are not constrained to the stationary region as is the case with other all the other approaches. Moreover, it has a limit standard N(0,1) distribution, which is unaffected by the inclusions of deterministic terms in the model. Finally, it is also the most efficient method, in the Pitman sense, against local departures from the null. (See, Gil-Alana and Robinson, 1994, for the version of Robinson's (1994) tests used in this application).

5. Empirical results and discussion

Table 2 refers to the estimates of the differencing parameter d, for the original data (panel i) and the logged monthly values (panel ii) respectively, in a model given by the following equation

$$y_t = \alpha + \beta t + x_t, \quad (1-B)^d\ x_t = u_t, \quad u_t = \rho\, u_{t-12} + \varepsilon_t. \tag{6}$$

where α and β refer to unknown coefficients corresponding to an intercept and a linear time trend; x_t is an I(d) process and u_t follows a seasonal (monthly) AR process.

We display the values of d (and the 95% confidence intervals) for the three well-known cases of (i) no terms, i.e., $\alpha = \beta = 0$ in (6); (ii) with an intercept ($\beta = 0$), and (iii) with an intercept and a linear time trend, and we mark in bold in the tables the selected models according to these deterministic terms.

Table 3
Estimates of d with monthly data in the differenced series.

Series (in logs)	No terms	With an intercept	With an intercept and a linear trend
Ethanol – Corn	0.66 (0.56, 0.77)	**0.56** **(0.47, 0.70)**	0.56 (0.47, 0.70)
Gasoline – Corn	0.73 (0.66, 0.83)	**0.65** **(0.57, 0.77)**	0.65 (0.57, 0.77)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

Table 4
Estimates of d with monthly data (ending at December 2019).

(i) Original data			
Series	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	1.18 (1.10, 1.27)	**1.29** **(1.20, 1.39)**	1.29 (1.20, 1.39)
Ethanol ($/Gall)	0.97 (0.88, 1.09)	**0.98** **(0.88, 1.11)**	0.98 (0.88, 1.11)
Gasoline ($/Gall)	1.05 (0.95, 1.17)	**1.08** **(0.97, 1.20)**	1.08 (0.97, 1.20)
(ii) Logged data			
Series	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	1.17 (1.07, 1.27)	**1.36** **(1.26, 1.48)**	1.36 (1.26, 1.48)
Ethanol ($/Gall)	1.09 (0.98, 1.22)	**1.09** **(0.98, 1.23)**	1.09 (0.98, 1.23)
Gasoline ($/Gall)	0.98 (0.89, 1.09)	**0.98** **(0.89, 1.09)**	0.98 (0.89, 1.09)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

Starting with the original data we observe that the time trend coefficient is insignificant in the three series, the intercept being sufficient to describe the deterministic part of the model. We see that for two of the series (Ethanol and Gasoline) the unit root null hypothesis cannot be rejected, while this hypothesis is rejected in favor of $d > 1$ for Corn. Looking at the logged transformed data (in the lower panel) though there are some quantitative differences, qualitatively the results are very similar, and the unit root null cannot be rejected for Ethanol and Gasoline, being rejected in favor of higher degrees of integration for corn.

In Table 3, we look at the log-series and focus on the differences between Gasoline and Ethanol with respect to Corn. In both cases we observe a substantial reduction in the degree of integration, the estimated values of d being 0.56 (for Ethanol-Corn difference) and 0.65 for Gasoline-Corn. More importantly, in the two series we observe that the highest value in the intervals is smaller than 1 supporting thus the hypothesis of convergence and mean reversion in the differences.

In order to check if the Covid-pandemic has had any effect on the degree of persistence in the data, we also make the computation with data ending at December 2019. The results for the unlogged and logged data are reported in Table 4. We observe that broadly speaking, the same conclusions hold, finding evidence of unit roots for both ethanol and gasoline but higher orders of integration for corn. Thus, in any single case we support the hypothesis of mean reversion even with the pre-Covid data. Similarly, for the differenced series, (Table 5) the values are very similar to those obtained when using the whole sample size, providing support of mean reversion in the differenced series.

In comparison with other previous studies as Barros et al. (2014) that measured the integration coefficient for the Brazilian ethanol in times of financial crisis, here a similar behavior can be seen in the pre and post shock results. In that paper, time series come from 1983 with yearly samples and results ending in the financial crisis revealed a temporal increase in the Ethanol integration parameter evolving from 1.06 (2007) and 1.11 (2010) to 0.97 (2013) and 0.93 (2016). In our study for the US-based ethanol with monthly samples, pre-covid value is 0.98 (ending in 2019), and post-covid is 1.00 (ending February 2022, before the Ukrainian war). Thus, both studies measuring the ethanol pricing but produced in different countries reveal both long memory and almost no evidence of mean reversion. The smaller sampling frequency of Barros et al. (2014) might explain the greater changes in the integration factor parameter. On the other hand, looking at Serra et al. (2011), conclusions regarding Ethanol and Corn also appear to show a similar behavior. In the case of Ethanol, weak mean reversion properties are in line with the ethanol market equilibrium pointed by Serra et al. (2011) between

Table 5
Estimates of d with monthly data in the differenced series (with data ending at December 2019).

Series (in logs)	No terms	With an intercept	With an intercept and a linear trend
Ethanol – Corn	0.66 (0.56, 0.75)	**0.57** **(0.48, 0.70)**	0.57 (0.48, 0.70)
Gasoline – Corn	0.79 (0.69, 0.90)	**0.70** **(0.60, 0.82)**	0.70 (0.60, 0.82)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

Table 6
Estimates of d with weekly data and white noise errors.

(i) Original data			
Series (weekly)	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	1.03 (0.98, 1.08)	**1.03** **(0.98, 1.09)**	1.03 (0.98, 1.09)
Ethanol ($/Gall)	0.99 (0.94, 1.05)	**1.00** **(0.95, 1.07)**	1.00 (0.95, 1.07)
Gasoline ($/Gall)	1.05 (1.01, 1.10)	**1.04** **(0.99, 1.09)**	1.04 (0.99, 1.09)
(ii) Logged data			
Series	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	1.00 (0.95, 1.05)	**1.02** **(0.98, 1.08)**	1.02 (0.98, 1.08)
Ethanol ($/Gall)	1.00 (0.95, 1.07)	**1.02** **(0.96, 1.08)**	1.02 (0.96, 1.08)
Gasoline ($/Gall)	1.01 (0.96, 1.06)	**1.06** **(1.01, 1.11)**	1.06 (1.01, 1.11)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

ethanol market oil producers and the crude oil industry. In the case of Corn, where there is no mean reversion behavior (d = 1.29), the corn price responds to crude oil price increases, but only when the ethanol market is far from its equilibrium leading this shock to a permanent behavior.

To corroborate this hypothesis, we also conducted the analysis on a weekly basis, and the results are reported across Tables 6–11. Starting with a model where the residuals are uncorrelated (in Table 6) we see that all the estimates of d are around 1.00, supporting thus the unit root null hypothesis in all except one single case (Gasoline, logs). If autocorrelation is permitted throughout the model of Bloomfield (1973), (Table 7), the results are similar though now Gasoline (original) also displays an integration order significantly higher than 1.

Table 8 focuses on the differences and the values are now higher than with the monthly data, and the evidence of mean reversion is weak with the estimates of the differencing parameter being close to 1.

Finally, we repeat the computation using data ending at December 2019. (Tables 9–11) and the same qualitative results hold as in previous tables. Evidence of d = 1 in most of the individual series (with the exception of gasoline in the two cases (original and logged) with autocorrelation), and values of d slightly below 1 for the differenced series, finding thus very weak evidence of mean reversion.

As a robustness method, and in relation with the results reported so far, we also conducted the analysis with other fractionally integrated approaches, including the maximum likelihood estimation in the time domain of Sowell (1992) along with various semiparametric methods (Geweke and Porter-Hudak, 1986; Shimotsu and Phillips, 2005) and the results, though not reported, were almost identical to those reported in the paper. In addition, we also investigated the potential presence of nonlinear trends in the data, and for this purpose, we replaced the first equality in (6) by the Chebyshev polynomials in time, such that the new model becomes

$$y_t = \sum_{i=0}^{m} \theta_i P_{iT}(t) + x_t, \quad (1-L)^d x_t = u_t, \quad t = 1, 2, \ldots, \tag{7}$$

Table 7
Estimates of d with weekly data and autocorrelated errors.

(i) Original data			
Series (weekly)	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	0.98 (0.91, 1.06)	**0.98 (0.90, 1.05)**	0.98 (0.90, 1.05)
Ethanol ($/Gall)	0.92 (0.85, 1.03)	**0.92 (0.82, 1.04)**	0.91 (0.82, 1.04)
Gasoline ($/Gall)	1.11 (1.04, 1.21)	**1.13 (1.05, 1.24)**	1.13 (1.05, 1.24)
(ii) Logged data			
Series	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	0.97 (0.91, 1.05)	**1.02 (0.94, 1.10)**	1.02 (0.95, 1.10)
Ethanol ($/Gall)	0.99 (0.86, 1.04)	**0.94 (0.86, 1.08)**	0.94 (0.86, 1.08)
Gasoline ($/Gall)	1.00 (0.94, 1.08)	**1.10 (1.00, 1.20)**	1.10 (1.00, 1.20)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

Table 8
Estimates of d with weekly data in the differenced series.

(i) White noise errors			
Series (in logs)	No terms	With an intercept	With an intercept and a linear trend
Ethanol – Corn	1.00 (0.96, 1.05)	**0.93 (0.88, 0.99)**	0.93 (0.88, 0.99)
Gasoline – Corn	0.99 (0.94, 1.04)	**0.96 (0.90, 1.00)**	0.96 (0.90, 1.00)
(ii) Autocorrelated errors			
Series (in logs)	No terms	With an intercept	With an intercept and a linear trend
Ethanol – Corn	1.02 (0.95, 1.11)	**0.87 (0.79, 0.99)**	0.88 (0.80, 0.99)
Gasoline – Corn	0.99 (0.92, 1.09)	**0.93 (0.82, 1.01)**	0.93 (0.82, 1.01)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

where m refers to the orthogonal Chebyshev polynomials order in time, expressed as:

$$P_{0,T}(t) = 1, \tag{8}$$

$$P_{i,T}(t) = \sqrt{2}\cos(i\pi(t-0.5)/T), \quad t = 1, 2, \ldots, T; \quad i = 1, 2, \ldots \tag{9}$$

The results, based on Cuestas and Gil-Alana (2016) and reported in the Appendix, clearly show no evidence of non-linear trends in any of the series under investigation.

6. Conclusions

In this paper, we have studied the market persistence and mean reversion properties for the corn, bioethanol, and gasoline prices, evaluating the long memory effects with fractional integration techniques for the case of the US market, currently the largest bioethanol and crude oil producer. To this end, we have employed monthly data from January 1982 to February 2022 from USDA and weekly data from October 2005 to May 2022; and used fractional integration techniques for the analysis. If mean reversion properties apply, the time-series under study is forecasted to recover to its mean value with no additional policies. Thus, investors can profit this property and expect corrections to this value when volatility applies to the time series.

Empirical results show evidence of no mean reversion for the prices of the three series (ethanol, gasoline and corn prices) though some evidence is found in the differences of ethanol and gasoline with respect to corn. In particular, the

Table 9
Estimates of d with weekly data and white noise errors (ending 2019).

(i) Original data			
Series (weekly)	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	1.03 (0.97, 1.09)	**1.03 (0.97, 1.09)**	1.03 (0.97, 1.09)
Ethanol ($/Gall)	0.98 (0.92, 1.05)	**1.02 (0.95, 1.09)**	1.02 (0.95, 1.09)
Gasoline ($/Gall)	1.03 (0.98, 1.09)	**1.04 (1.00, 1.10)**	1.04 (1.00, 1.10)
(ii) Logged data			
Series	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	1.00 (0.95, 1.06)	**1.02 (0.97, 1.08)**	1.02 (0.97, 1.08)
Ethanol ($/Gall)	0.99 (0.92, 1.06)	**1.03 (0.96, 1.10)**	1.03 (0.96, 1.10)
Gasoline ($/Gall)	1.00 (0.95, 1.06)	**1.04 (0.99, 1.09)**	1.04 (0.99, 1.09)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

Table 10
Estimates of d with weekly data and autocorrelated errors (ending 2019).

(i) Original data			
Series (weekly)	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	0.97 (0.89, 1.06)	**0.97 (0.88, 1.05)**	0.97 (0.88, 1.05)
Ethanol ($/Gall)	0.89 (0.81, 0.98)	**0.90 (0.80, 1.02)**	0.90 (0.80, 1.02)
Gasoline ($/Gall)	1.12 (1.01, 1.19)	**1.14 (1.05, 1.24)**	1.14 (1.05, 1.24)
(ii) Logged data			
Series	No terms	With an intercept	With an intercept and a linear trend
Corn ($/Bushel)	1.00 (0.92, 1.08)	**1.02 (0.94, 1.10)**	1.01 (0.94, 1.10)
Ethanol ($/Gall)	0.87 (0.77, 0.99)	**0.91 (0.80, 1.02)**	0.91 (0.81, 1.02)
Gasoline ($/Gall)	0.99 (0.92, 1.10)	**1.14 (1.04, 1.23)**	1.14 (1.04, 1.23)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

value 1 is in all cases within the confidence interval showing this lack of mean reversion properties. Thus, it appears that external shocks in the individual series are expected to remain persistent and additional policy measures would be required to recover the original trends, in line with other studies such as Barros et al. (2014) for the Brazilian industry.

Furthermore, the possible Covid impact in the time series has been analyzed by comparing the results ending in 2019 with those ending in 2022, finding evidence of no major changes in the integration orders in the series. Following other studies with yearly data, a temporal increase would be expected in this integration factor; however, this issue does not hold in our empirical results with weekly and monthly samples and can be associated with the smaller sampling frequency used.

This article can be extended in various directions. First, from a methodological viewpoint, other structures incorporating, for example, structural breaks can be taken into account. In addition, non-linear models based on Fourier function in times (Gil-Alana and Yaya, 2021) or neural networks (Yaya et al., 2021) can also be considered. Nevertheless, the results reported in the Appendix seems to reject the hypothesis of non-linear structures. On the other hand, issues such as the Russia-Ukraine war and its effects on the series examined is something that will be investigated in future papers.

Table 11
Estimates of d with weekly data in the differenced series (ending 2019).

(i) White noise errors			
Series (in logs)	No terms	With an intercept	With an intercept and a linear trend
Ethanol – Corn	1.01 (0.96, 1.07)	**0.95** **(0.89, 0.99)**	0.95 (0.89, 1.00)
Gasoline – Corn	0.99 (0.94, 1.05)	**0.96** **(0.90, 1.00)**	0.96 (0.90, 1.00)
(ii) Autocorrelated errors			
Series (in logs)	No terms	With an intercept	With an intercept and a linear trend
Ethanol – Corn	1.04 (0.97, 1.13)	**0.86** **(0.75, 0.97)**	0.87 (0.75, 0.97)
Gasoline – Corn	0.98 (0.91, 1.08)	**0.98** **(0.85, 1.10)**	0.98 (0.85, 1.10)

The values reported are the estimated values of the order of integration. In parenthesis, they are 95% confidence bands, and in bold we mark the selected specifications according to the deterministic terms.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and material statement

The authors declare that all data supporting the findings of this study is available within the article and its supplementary information files. In particular, the calculation datasets generated during and/or analyzed during the current study are available from the corresponding author on request. The results/data/figures in this manuscript have not been published elsewhere, nor are they under consideration by another publisher, and there are not hyperlinks to publicly archived datasets analyzed or generated during the study except for the public information collected.

Funding

Professors Gil-Alana and Martín-Valmayor notifies that for the development of this research, it has been received funds from the MCIN-AEI-FEDER (Government of Spain) under Grant Agreement No PID2020-113691RB-I00 and gratefully acknowledges this financial support from project from 'Ministerio de Ciencia e Innovación' (MCIN), 'Agencia Estatal de Investigación' (AEI) Spain and 'Fondo Europeo de Desarrollo Regional' (FEDER). All authors declare that there is not any competing financial and/or non-financial interests in relation to the work described

Declarations

Ethics approval and consent to participate

I declare on behalf of all authors that the manuscript has not been submitted to more than one publication for simultaneous consideration, and that this work is original and has not been published elsewhere in any form or language (partially or in full). Our results are clear, honest, and without fabrication, falsification or inappropriate data manipulation, to discipline-specific rules for acquiring, selecting and processing data. No data, text, or theories by others are presented as if they were the author's own ('plagiarism'). Proper acknowledgments to other works are given (including material that is closely copied), summarized and/or paraphrased), quotation marks (to indicate words taken from another source) are used for verbatim copying of material, and permissions secured for material that is copyrighted.

I also declare that research articles, relevant literature and non-research articles are cited appropriately. We have avoided untrue statements about entities (who can be an individual person or a company) or descriptions of their behavior or actions that could potentially be seen as personal attacks or allegations about that entity. This research has no threats to public health or national security (e.g. dual use of research). If necessary, we are prepared to send relevant documentation or data in order to verify the validity of the results presented and sensitive information in the form of confidential or proprietary data is excluded.

Authors declare all of the above to make sure to respect third parties' rights such as copyright and/or moral rights.

Consent for publication

I declare that all authors consent for are responsible for correctness of the statements provided in the manuscript, as defined by Springer.

Appendix

(a) Monthly data					
Original data					
Corn ($/Bushel)	1.30 (1.21, 1.40)	0.560 (0.04)	0.984 (0.14)	−0.110 (−0.04)	0.538 (0.37)
Ethanol ($/Gall)	0.99 (0.88, 1.11)	1.280 (0.70)	−0.434 (−0.39)	0.101 (0.18)	0.240 (0.65)
Gasoline ($/Gall)	1.10 (0.99, 1.23)	1.226 (0.42)	−0.526 (−0.29)	0.063 (0.07)	0.334 (0.65)
Logged data					
Corn ($/Bushel)	1.37 (1.27, 1.49)	0.349 (0.14)	0.300 (0.11)	−0.024 (−0.02)	0.142 (0.27)
Ethanol ($/Gall)	1.05 (0.95, 1.18)	0.212 (0.10)	−0.354 (−028)	0.087 (0.14)	0.217 (0.57)
Gasoline ($/Gall)	0.97 (0.87, 1.08)	0.267 (0.25)	−0.500 (−0.80)	0.094 (0.29)	0.243 (1.11)
(b) Weekly data					
Original data					
Corn ($/Bushel)	1.02 (0.97, 1.08)	389.919 (1.06)	0.536 (0.02)	−36.494 (−0.33)	−96.587 (−1.34)
Ethanol ($/Gall)	1.00 (0.94, 1.07)	2.661 (1.59)	0.205 (0.22)	0.032 (0.07)	−0.227 (−0.76)
Gasoline ($/Gall)	1.04 (0.99, 1.08)	202.834 (0.97)	24.466 (0.19)	−19.502 (−0.32)	−25.192 (−0.63)
Logged data					
Corn ($/Bushel)	1.02 (0.97, 1.07)	**5.751** **(7.79)**	−0.014 (−0.03)	−0.087 (−0.40)	−0.209 (−1.44)
Ethanol ($/Gall)	1.02 (0.96, 1.08)	0.973 (1.20)	0.119 (0.24)	−0.074 (−0.03)	−0.115 (−0.72)
Gasoline ($/Gall)	1.06 (1.01, 1.12)	**5.260** **(4.18)**	0.139 (0.17)	−0.094 (−0.26)	−0.111 (−0.47)

References

Abbott, P., 2013. Biofuels, Binding Constraints and Agricultural Commodity Price Volatility. In: NBER Working Paper Series, Cambridge National Bureau of Economic Research, Inc.

Barros, CP., Gil-Alana, LA., Wanke, P., 2014. Ethanol consumption in Brazil: empirical facts based on persistence, seasonality and breaks. Biomass Bioenergy 63, 313–320. http://dx.doi.org/10.1016/j.biombioe.2014.02.012.

Barros, C., Gil-Alana, LA., Wanke, P., 2018. Ethanol production in brazil: empirical evidence based on persistence. Int. J. Energy Sect. Manage. 12 (4), 566–580. http://dx.doi.org/10.1108/IJESM-05-2017-0004.

Bloomfield, P., 1973. An exponential model in the spectrum of a scalar time series. Biometrika 60, 217–226.

Bouri, E., Dutta, A., Saeed, T., 2021. Forecasting ethanol price volatility under structural breaks. Biofuels, Bioprod. Biorefin. 15 (1), 250–256. http://dx.doi.org/10.1002/bbb.2158.

2021. BP statistical review of world energy. https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energy-economics/statistical-review/bp-stats-review-2021-full-report.pdf, Last viewed June 28th, 2022.

Chiu, F-P., Hsu, C-S., Ho, A., Chen, C-C., 2016. Modeling the price relationships between crude oil, energy crops and biofuels. Energy 109, 845–857. http://dx.doi.org/10.1016/j.energy.2016.05.016.

Cuestas, J.C., Gil-Alana, L.A., 2016. Testing for long memory in the presence of non-linear deterministic trends with Chebyshev polynomials. Stud. Nonlinear Dyn. Econom. 20 (1), 37–56.

Dahlhaus, R., 1989. Efficient parameter estimation for self-similar process. Ann. Statist. 17, 1749–1766.

David, SA., Quintino, DD., Inacio, CMC., Machado, JAT., 2018. Fractional dynamic behavior in ethanol prices series. J. Comput. Appl. Math. 339, 85–93. http://dx.doi.org/10.1016/j.cam.2018.01.007.

Du, X., Carriquiry, MA., 2013. Flex-fuel vehicle adoption and dynamics of ethanol prices: lessons from Brazil. Energy Policy 59, 507–512. http://dx.doi.org/10.1016/j.enpol.2013.04.008.

Dutta, A., 2019. Forecasting ethanol market volatility: new evidence from the corn implied volatility index: forecasting ethanol market volatility. Biofuels, Bioprod. Biorefin. 13 (1), 48–54. http://dx.doi.org/10.1002/bbb.1931.

Etienne, XL., Trujillo-Barrera, A., Wiggins, S., 2016. Price and volatility transmissions between natural gas, fertilizer, and corn markets. Agric. Financ. Rev. 76 (1), 151–171. http://dx.doi.org/10.1108/AFR-10-2015-0044.

Geweke, J., Porter-Hudak, S., 1986. The estimation and application of long memory time series models. J. Time Series Anal. 4 (4), 221–238. http://dx.doi.org/10.1111/j.1467-9892.1983.tb00371.x.

Gil-Alana, L.A., Robinson, P.M., 1994. Testing of unit roots and other nonstationary hypotheses in macroeconomic time series. J. Econometrics 80 (2), 241–268.

Gil-Alana, L.A., Yaya, O.S., 2021. Testing fractional unit roots with non-linear smooth break approximations using Fourier functions. J. Appl. Stat. 48 (13–15), 2542–2559.

Kristoufek, L., Janda, K., Zilberman, D., 2014. Price transmission between biofuels, fuels, and food commodities. Biofuels, Bioprod. Biorefin. 8 (3), 362–373. http://dx.doi.org/10.1002/bbb.1464.

Morris, DH., Parrott, SD., Mehlhorn, J., Tewari, R., Pruitt, R., 2017. The price transmission relationship between ethanol, wholesale gasoline, and blended retail gasoline. Int. J. Food Agric. Econ. 5 (2), 1–8.

Musunuru, N., 2016. Examining volatility persistence and news asymmetry in soybeans futures returns. Atl. Econ. J. 44 (4), 487–500. http://dx.doi.org/10.1007/s11293-016-9517-3.

Oláh, J., Lengyel, P., Balogh, P., Harangi-Rákos, M., Popp, J., 2017. The role of biofuels in food commodity prices volatility and land use. J. Compet. 9 (4), 81–93.

Pal, D., Mitra, SK., 2017. Diesel and soybean price relationship in the USA: evidence from a quantile autoregressive distributed lag model. Empir. Econ. 52 (4), 1609–1626. http://dx.doi.org/10.1007/s00181-016-1114-4.

Robinson, P.M., 1994. Efficient tests of nonstationary hypotheses. J. Amer. Statist. Assoc. 89, 1420–1437.

Saghaian, S., Nemati, M., Walters, C., Chen, B., 2018. Asymmetric price volatility transmission between U.S. biofuel, corn, and oil markets. J. Agric. Resour. Econ. 43 (1), 46–60.

Serra, T., Zilberman, D., José, M.Gil., Goodwin, BK., 2011. Nonlinearities in the U.S. corn-ethanol-oil-gasoline price system. Agric. Econ. 42 (1), 35–45. http://dx.doi.org/10.1111/j.1574-0862.2010.00464.x.

Shimotsu, K., Phillips, P.C.B., 2005. Exact local whittle estimation of fractional integration. Ann. Statist. 33 (4), 1890–1933. http://dx.doi.org/10.1214/009053605000000309.

Sönnichsen, N., 2022. Biofuel industry in the US. Statistics and facts. Stat. Database.

Sowell, F., 1992. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. J. Econometrics 53 (1–3), 165–188.

To, H., Grafton, RQ., 2015. Oil prices, biofuels production and food security: past trends and future challenges. Food Secur. 7 (2), 323–336. http://dx.doi.org/10.1007/s12571-015-0438-9.

Yaya, O.S., Ogbonna, A.E., Furuoka, F., Gil-Alana, L.A., 2021. A new unit root test for unemployment hysteresis based on the autoregressive neural network. Oxf. Bull. Econ. Stat. 83 (4), 960–981.