https://economix.fr/

# Market Efficiency and Optimal Hedging Strategy for the US Ethanol Market

Document de Travail
Working Paper
2018-6

Emmanuel Hache
Anthony Paris

UMR 7235

EconomiX - UMR7235
Université Paris Nanterre
Bâtiment G - Maurice Allais, 200, Avenue de la République
92001 Nanterre cedex

Email : secretariat@economix.fr

# Market Efficiency and Optimal Hedging Strategy for the US Ethanol Market

HACHE Emmanuel* † ‡ PARIS Anthony† § ¶

November 2017

**Abstract**

The aim of this paper is to study the ethanol price dynamics in the US market and find the optimal hedging strategy. To this end, we first attempt to identify the long-term relationship between ethanol spot prices and the prices of futures contracts on the Chicago Board of Trade (CBOT). Then, we model the short-term dynamics between these two prices using a Markov-switching vector error correction model (Ms-VECM). Finally, accounting for the variance dynamics using a Gjr-MGarch error structure, we compute a time-varying hedge ratio and determine the optimal hedging strategy in the US ethanol market.

*JEL Classification*: Q41, Q42, G15, C41

*Keywords*: Ethanol prices, Futures markets, Markov-switching regime models, Hedge ratio

*Corresponding author. Tel.: +33 1 47 52 67 49; fax: +33 1 47 52 70 66. E-mail address: emmanuel.hache@ifpen.fr.

†IFP Energies nouvelles, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison, France.

‡The French Institute for International and Strategic Affairs (IRIS), 2 bis, rue Mercoeur, 75011 Paris, France.

§EconomiX-CNRS, University of Paris Nanterre, 200 avenue de la République, F92000 Nanterre France.

¶University of Paris 8 (LED), 2 rue de la Liberté, 93526 Saint-Denis, France.

# 1 Introduction

Ethanol is derived from various agricultural products (cassava, corn, hemp, sugar beet or sugarcane) and has been increasingly added to gasoline blends for several reasons: (i) it helps to reduce greenhouse gases emissions (GHG) in the transportation sector, (ii) if produced with agricultural feedstock, ethanol can be seen as a renewable energy, and (iii) from a technical point of view, the use of ethanol helps to boost octane numbers and leads to an improvement in thermal engine efficiency. All these factors have contributed to the development of ethanol's use worldwide. Such recent evolution calls for a detailed investigation of the ethanol market to fully understand its dynamics. Specifically, the aim of this paper is to study the ethanol price dynamics and determine the optimal hedging strategy on the US market.

Ethanol policy is a story that has many chapters in the past 40 years in the US. Ethanol inclusion in US gasoline blends began in 1908 when the Model-T Ford could be customized to run on gasoline or alcohol. It was not until the late Seventies, however, that the meaningful inclusion of ethanol came about. The first government involvement for ethanol was the Energy Tax Act of 1978 (a tax exemption for adding ethanol to the gasoline blend) in the wake of geopolitical concerns in the oil market with the 2$^{nd}$ world oil shock. The Surface Transportation Assistance Act of 1982 and the Tax Reform Act of 1984 gave an impetus for ethanol inclusion despite a decrease of the tax exemption during the 1992-2000 period with the Omnibus Budget Recollection Act. The Renewable Fuel Standard (RFS) program, created by the Energy Policy Act of 2005 and extended by the Energy Independence and Security Act of 2007, has led to a new expansion of the US ethanol market. Ethanol production and consumption have since been multiplied by four between 2005 and 2016, increasing approximately from 300 to 1,200 million gallons.

Since 2009 the US has become a net exporter in the ethanol market. According to the US Census Bureau, the Department of Commerce, and the Department of Agriculture, the US exported 836 million gallons of ethanol in 2015 (5.7% of total US ethanol production) and imported 93 million gallons of fuel ethanol (less than 1% of US ethanol consumption). Canada (30% of US exports), Brazil (14%), Philippines (9%), China (8%), and India (6%) are the top destinations of US ethanol in 2015. Brazil also remains the main supplier for the US with 73% of the imported ethanol volume in 2015. This export-import structure within the ethanol market with Brazil can be easily explained by the RFS and California Low Carbon Fuel Standard (LCFS) targets put in place for the reduction of GHG emissions that impose more stringent requirements. As mentioned by the Energy Information Administration,[1] life cycle analysis (LCA) studies demonstrate that ethanol from sugarcane has

[1] https://www.eia.gov/todayinenergy/detail.php?id=25312

a better scoring in terms of GHG emissions than products based on corn feedstock. It contributes to the substitution of corn-ethanol production from the countryside with imported sugarcane-ethanol from Brazil. The ethanol market structure is already driven by (i) the inclusion policy of different countries, (ii) energy prices and especially the evolution of the crude oil price, and (iii) the regulatory framework. But recent changes prove that the production process (ethanol is derived from different agricultural products) could also impact the international market structure and ethanol price dynamics. Ethanol prices registered several ups and downs since 2008, with prices ranging from $1.47 per gallon to more than $4 per gallon following the volatility observed in energy and agricultural prices.

Due to the increase of ethanol production and consumption in the US in the first part of the last decade, futures contracts on corn-based ethanol were launched on March 2005 on the Chicago Board of Trade (CBOT).[2] Derivatives markets allow commercial players to reduce their price risk exposure with various hedging strategies and different tools (futures contracts, options, etc.). These tools protect against adverse price movements in order to reduce the risk of loss in the business. In this context, the optimal hedging strategy is to minimize the variance of the hedge portfolio containing spot and futures contracts. A variety of questions has been asked regarding the derivatives strategy, which is related to traders' behavior, speculation, price volatility, etc.

The motivations of this paper are threefold. Firstly, considering the role the ethanol market could play in the transportation sector for its own energy transition, we study the long-term relationship between ethanol spot prices and the prices of futures contracts on the CBOT; allowing us to investigate the weak form of the efficient market hypothesis.[3] To the best of our knowledge, this is the first article focusing on ethanol in this research field. Secondly, we have a methodological motivation and contribution. Indeed, we compute a wide range of time-varying hedge ratios[4] with different econometric models to look for the optimal hedging strategy for ethanol commercial players. We consider adjustments to long-term equilibrium and regime shifts governed by a Markov chain (as Alizadeh et al. (2008)) and short-run dynamics between spot and futures price changes (as in Salvador and Arago

[2]CME Group is the world's leading and most diverse derivatives marketplace, made up of four markets, CME, CBOT, NYMEX, and COMEX. Each market offers a wide range of global benchmarks across major asset classes.

[3]An efficient market is characterized by prices that reflect all available information. The weak form of market efficiency considers only historical price or return series in the information set (Fama, 1970).

[4]One of them is the hedge ratio which is initially defined as the estimated coefficient between spot and futures price changes based on Ordinary Least Squares (OLS) estimation (Ederington, 1979) i.e., the ratio of the unconditional spot and futures price changes covariance over the unconditional variance of the futures price changes. It provides the number of futures contracts to buy or sell for one unit of the underlying asset (in the case of this article: ethanol) to minimize the variance of the hedged portfolio returns.

(2014)). In addition, we extend the work of Salvador and Arago (2014) by allowing short-run dynamics between prices to be state-dependent on price volatility. Hamilton (1989) proposes the Markov-switching model while Krolzig (1999) extends this specification to the vector autoregressive model. By including structural breaks in the variance equation, we take into account the high volatility persistence (Lamoureux and Lastrapes, 1990). With structural breaks in the short-run dynamics, we allow for time-varying behavior in the adjustment to the equilibrium and the short-run dynamic processes. We then include an informational link between mean and volatility processes across each market state (Alizadeh et al., 2008). Finally, relying on the Gjr framework (Glosten et al., 1993), we introduce asymmetric behavior to the variance process (see also Brooks et al. (2002)) to take into account different responses to new information according to the past shocks sign. Therefore, we estimate a Markov-switching vector error correction model with a Gjr-MGarch error structure (Ms-VECM-Gjr-MGarch). To overcome Johansen (1988)'s approach drawbacks,[5] we use Nielsen (2010)'s nonparametric cointegration approach to analyze its ability to improve hedging strategy. As Nielsen (2010)'s nonparametric cointegration procedure does not require model specification, we assume non-linear dynamics in short-run and variance equations. Thirdly, we check the performance of a cross-hedging strategy[6] with the gasoline futures market. Indeed, Franken and Parcell (2003) highlight its efficiency while Dahlgran (2009) concludes there is a lower performance with this market compared to the ethanol futures market.

The rest of the paper is organized as follows. Section 2 briefly reviews the literature on storable commodity market efficiency and hedging-ratio estimation. In section 3 we present data and the Markov-switching vector error correction model (Ms-VECM-Gjr-MGarch). Section 4 presents empirical results on the efficient market hypothesis and the optimal hedging strategy. The main conclusions are summarized in the final section.

## 2 A brief overview of literature

Following the works of Kaldor (1939), Working (1948), Brennan (1958) and Telser (1958), spot and futures prices of a storable commodity should be equal. The difference between these prices is explained by the cost of storage and the interest rate as,

$$F_t^T = S_t exp[(r_t + \bar{s})(T - t)] \tag{1}$$

[5]In particular, Johansen's procedure could lead to an estimation bias due to the restrictions imposed on the short-run dynamics which are supposed to be linear.

[6]Cross-hedging occurs when the asset underlying the contract is different than the asset whose price is being hedged (Hull, 2005).

and with a log-transformation,

$$f_t^T = s_t + (r_t + \bar{s})(T - t) \tag{2}$$

Here, $F_t^T$ (resp. $f_t^T$) is the price (resp. log-price) of futures contract at the time $t$ for a maturity $T$. $S_t$ (resp. $s_t$) is the spot price (resp. log-price) at the same date. $r_t$ and $\bar{s}$ refer, respectively, to the risk-free interest rate and the cost of carry, this latter is supposed to be constant. According to the aforementioned works, the difference between spot and futures prices is instantaneously compensated by arbitrageurs.

This hypothesis has been relaxed by Garbade and Silber (1983). They mention that arbitrageurs operate in the markets if the spread between these prices is large enough to enlarge their profits according to the transaction and information costs. Therefore, the unit relationship between spot and futures prices is only valid in the long term.The spot and futures markets are thus efficient if prices are cointegrated as in Chowdhury (1991) or Lai and Lai (1991). In addition, Garbade and Silber (1983) show that futures markets integrate new information faster than in the underlying spot market, leading to a causality from futures to spot prices. It helps the price discovery process registered in commodities markets which leads to informational efficiency for physical and financial markets.

Figuerola-Ferretti and Gonzalo (2010) extent this model by integrating the convenience yield, i.e., the premium attributed by agents for physically holding the commodity instead of holding a futures contract. It depends on various market characteristics in the spot market (weather conditions, geopolitical unrest, transaction costs, etc.).[7] With a constant free-risk interest rate, one-period futures contract and the approximation of the convenience yield, $y_t$, used by these authors, as

$$y_t = \gamma_1 s_t - \gamma_2 f_t \tag{3}$$

equation (2) becomes

$$f_t = \frac{1-\gamma_1}{1-\gamma_2} s_t + \frac{\bar{r} + \bar{s}}{1-\gamma_2} \tag{4}$$

Their theoretical framework allows a long-term relationship, i.e., a cointegrating relationship, with a non-unit coefficient between spot and futures prices. In addition, they mention that the coefficient value depends on the spot market condition. The parameter is greater (resp. smaller) than unity if the spot market is in contango (resp. backwardation).

Literature about the estimation of an optimal hedge ratio has been developed since the seminal work of Ederington (1979) who proposes using the estimated coefficient between changes in spot and futures prices with an ordinary least square

[7]See Routledge et al. (2000) or Heaney (2002) for more details on the convenience yield.

estimator (OLS). However, this hedge ratio is unsatisfactory for many markets (Cecchetti et al., 1988; Myers and Thompson, 1989). Baillie and Myers (1991) and Kroner and Sultan (1993) state that the hedge ratio should be time-varying based on the time-varying distribution of many asset prices. They propose computing this dynamic optimal hedge ratio ($\delta_t$) for each period by taking into account all past information ($\Omega_{t-1}$) such as

$$\delta_t|\Omega_{t-1} = \frac{\sigma_{t-1}(\Delta F_{t-1}, \Delta S_{t-1})}{\sigma^2_{t-1}(\Delta F_{t-1})} \tag{5}$$

Many studies estimate those conditional covariance ($\sigma_{t-1}$) and variance ($\sigma^2_{t-1}$) with the multivariate Garch model proposed by Engle and Kroner (1995) as, for instance, Kroner and Sultan (1993), Garcia et al. (1995) or Kavussanos and Nomikos (2000) and conclude there has been an improvement of the hedging strategy with the dynamic hedge ratio compared to the constant formulation. The improvement degree depends on the market and the futures maturity studied (Lien and Tse, 2002).

The estimation of the dynamic hedge ratio should integrate the possible existence of a cointegrating relationship between spot and futures prices. Kroner and Sultan (1993), Ghosh (1993), Chou et al. (1996) or Lien (1996) highlighted an underestimated hedge ratio if this characteristic is not accounted for. In addition, Brooks et al. (2002) show the improvements of the hedge ratio effectiveness with the integration of the asymmetric volatility response against positive and negative shocks, i.e., the leverage effect. Furthermore, the conditional mean (Sarno and Valente, 2000) and variance (Lamoureux and Lastrapes, 1990) estimations can be biased if regime shifts exist. Thus, the hedge ratio effectiveness can be improved by integrating regime shifts in the estimation. Lee and Yoder (2007a,b) include regime shifts in the variance process and show an improvement – but not always significant – of the hedge ratio effectiveness. Alizadeh et al. (2008) extent this model by integrating regime shifts in variance and conditional mean processes and highlight a significant effectiveness improvement for most of the markets studied. Finally, Salvador and Arago (2014) propose incorporating (i) the regime shifts, the cointegrating relationship and the leverage effect in the same model in order to estimate an optimal dynamic hedge ratio, as well as (ii) the short-run dynamics between spot and futures price changes.

The literature concerning hedging strategies on energy markets is well developed with, for instance, Lien and Yang (2008) for heating and crude oil markets, Alizadeh et al. (2008) on crude oil, unleaded gasoline and heating oil markets, Hanly (2017) with WTI and Brent crude oils, natural gas, unleaded gasoline, heating oil and gasoil. However, the literature on hedging strategies on ethanol market is very scarce. Franken and Parcell (2003) highlight the cross-hedging efficiency between ethanol spot price and unleaded gasoline futures markets. However, while they correct the estimation for autocorrelation and heteroscedasticity, they do not incorporate the error correction term, regime switching and time-varying variance process. Finally, Dahlgran (2009) compares direct hedging for ethanol commercial agents with cross-hedging strategy with unleaded and Reformulated Gasoline Blendstock for Oxygen Blending (RBOB gasoline) futures markets. He demonstrates that the direct hedging strategy outperforms cross-hedging for a four-week, and more, hedge horizon.

# 3 Data and methodology

As stressed above, our article deals with the relationship between the spot prices and the futures prices of ethanol. As transaction volumes have risen, in particular for the shortest maturities, we focus on the relationship between the spot prices and the prices for the two-month futures contracts. The data studied are relative to the ethanol in the North American market: the spot price for ethanol (Argus Ethanol USGC barge/rail fob Houston), the futures prices of ethanol on the CBOT, and the transaction volumes and open interest for the same market, (weekly market business reports of the Commodity Futures Trading Commission [CFTC]). Apart from the spot price of ethanol, these pieces of information are all in the public domain. The data cover the period from July 2008 to December 2016, corresponding to 468 weekly observations. The prices are expressed in US dollars per gallon and are log-transformed.

Table 1 presents some descriptive statistics and tests results. Unit root tests confirm the stationarity of spot and futures prices series in their first-difference.[8] In addition, the Ljung and Box (1978) and ARCH tests confirm the presence of autocorrelation in most cases and heteroscedasticity, respectively. These characteristics justify the choice of a specification with autoregressive terms and heteroscedastic errors.

We apply the Johansen (1988)'s test to check the existence of a long-term relationship with unit cointegrating vectors and to estimate the conditional mean with a Markov switching vector error correction model (Ms-VECM) within a bivariate framework. The inclusion of a multivariate generalized autoregressive conditional heteroscedasticity (MGarch) error structure allows us to compute the dynamic hedge ratio. By including a long-term equilibrium, we eliminate the bias in the hedge ratio

[8] In view of the conflicting results for the spot log-price series, we apply the Perron (1990)'s unit root test which confirms its non-stationarity with a break in mean on March 12 2014. We choose this test in view of series characteristics, i.e., the absence of trend and a potential break in the mean. We present results with innovational-outlier model for break date determination. Results with additional-outlier model are similar.

estimation mentioned by Kroner and Sultan (1993) and Ghosh (1993). In addition, the nonlinear specification avoids estimation bias due to the existence of multiple regimes in the mean (Sarno and Valente, 2000) and variance (Lamoureux and Lastrapes, 1990) equations. Furthermore, the dynamic hedge ratio computed with this specification outperforms OLS hedge ratio in many energy markets (Alizadeh et al., 2008). Finally, we take into account the leverage effect within the Gjr framework.

Table 1: Summary statistics and unit root tests

| Variables | Log | | First-log differences | |
|---|---|---|---|---|
| | Spot | Futures | Spot | Futures |
| Mean | / | / | 0.000 | 0.000 |
| Std. errors | / | / | 0.050 | 0.040 |
| Skewness | / | / | 0.047 | -0.283 |
| Kurtosis | / | / | 6.039 | 4.288 |
| ADF | 0.047* | 0.297 | 0.001* | 0.001* |
| PP | 0.099* | 0.306 | 0.001* | 0.001* |
| KPSS | 0.010 | 0.010 | 0.100* | 0.100* |
| Perron | -1.148 | -1.229 | / | / |
| | -3.8 | -3.8 | / | / |
| Q(6) | 0.001 | 0.001 | 0.001 | 0.681 |
| $Q^2(6)$ | 0.001 | 0.001 | 0.001 | 0.001 |

Note: This table reports descriptive statistics and the p-value of the unit root tests applied, i.e., Augmented Dickey and Fuller (1981)'s test (ADF), Phillips and Perron (1988)'s test (PP) and Kwiatkowski et al. (1992)'s test (KPSS). The Perron's line refers to the Perron (1990) test with the test's statistic and the critical value at a 5% significance level in the first and second line, respectively. The critical value comes from Perron and Vogelsang (1992). The null hypothesis of unit root with break is rejected if the test statistics is greater than the critical value. The star mentions the stationarity of the variable at a 10% significance level. Q(6) and $Q^2(6)$ are the p-value of the Ljung and Box (1978)'s test and ARCH test (Engle, 1982) for $6^{th}$ order autocorrelation, respectively.

It should be emphasized that the Johansen (1988) cointegration test requires assumptions regarding the short-run dynamics that must follow a linear process. Using Johansen (1988)'s procedure with a non-linear short-run specification may lead to bias in both cointegration test results and long-term estimations, generating in turn a bias on the short-run and conditional variance estimations. To overcome these major drawbacks, we rely on Nielsen (2010)'s nonparametric variance ratio testing approach as this methodology does not require assumptions in the short-run specification.[9] The nonparametric variance ratio trace statistic is defined by

$$\Lambda_{n,r}(d_1) = T^{2d_1} \sum_{j=1}^{n-r} \lambda_j \tag{6}$$

where $\lambda_j$, $j = 1, ..., n$, are the eigenvalues, listed by increasing order, of the observed $(n \times T)$ time series matrix, $r$ is the cointegration rank tested and $d_1$ is a summation parameter fixed to 0.1.[10] The eigenvalues of the price series matrix are given by the solutions of

$$|\lambda B_T - A_T| = 0 \tag{7}$$

[9]For more details on the testing procedure, see Nielsen (2010).

[10]As mentioned by Nielsen (2010), the choice of $d_1 = 0.1$ maximizes the power of the test.

with

$$\begin{aligned} A_T &= \sum_{t=1}^{T} Z_t Z_t' \\ B_T &= \sum_{t=1}^{T} \tilde{Z}_t \tilde{Z}_t' \end{aligned} \tag{8}$$

where $\tilde{Z}_t$ is the fractional difference of $Z_t$ truncated by $d_1$. $Z_t$ is our time series matrix after demeaning. The null hypothesis is the presence of $r-1$ cointegration relationships. A test statistic that is greater than the critical value leads to the rejection of the null hypothesis in favor of the alternative, i.e., the existence of $r$ cointegration relationships. In addition, the estimated cointegration coefficients are provided by the eigenvectors associated with eigenvalues and converge to their real values. Therefore, by using both Johansen (1988) and Nielsen (2010) cointegration approach, we can analyze the effect of the long-term estimation bias on the hedge ratio efficiency.

The Ms-VECM with Gjr-MGarch[11] error structure can be expressed by

$$\begin{aligned} \Delta X_t &= c + \Gamma_{st} \Delta X_{t-1} + \Pi_{st} X_{t-1} + \epsilon_{t,st} \\ \epsilon_{t,st} &= \begin{pmatrix} \epsilon_{s,t,st} \\ \epsilon_{f,t,st} \end{pmatrix} |\Omega_{t-1} \backsim IN(0, H_{t,st}) \end{aligned} \tag{9}$$

where $\Delta X_t = (\Delta s_t, \Delta f_t)'$ (resp. $X_{t-1} = (s_{t-1}, f_{t-1})'$) is the vector of log-returns (resp. log-price) and c is a vector of constant. $\Gamma_{st}$ and $\Pi_{st}$ are coefficient matrices related to short- and long-term dynamics, respectively.[12] These $(2 \times 2)$ matrices depend on the regime $st$, $st = 1, 2$. $\epsilon_{t,st}$ is a regime-dependent Gaussian white noise vector. With our multivariate Garch error structure, the error covariance matrix, $H_{t,st}$, is time- and regime-dependent.

As mentioned by Alizadeh et al. (2008), two steps are necessary to estimate this model. Firstly, we check the existence of a cointegrating relationship between spot and futures prices. Considering a linear process, we apply the Johansen (1988)'s test. The $\lambda_{max}$ and $\lambda_{trace}$ statistics allow us to check the rank of the matrix $\Pi$. Under the alternative hypothesis, there is at least one cointegrating relationship. If the rank of the long-term adjustment is non-null, $\Pi$ can be decomposed such as $\Pi = \alpha\beta'$. The vectors $\alpha$ and $\beta$ are $(2 \times 1)$ coefficient vectors referring to the error correction coefficients, i.e., characterizing the adjustment process to the long-term equilibrium, and the long-term coefficients, describing the long-term equilibrium, respectively. In addition, we apply the likelihood ratio test from Johansen (1995) to

[11] We estimate a wide range of specifications but only detail the more complex model.

[12] We integrate only one lag in the short-run dynamics according to the information criterion BIC from Schwarz (1978) during the Johansen cointegration procedure.

check the existence of unitary long-term coefficients between spot and futures prices. The non-reject of the null hypothesis of unit coefficient will favor the Garbade and Silber (1983) model against that proposed by Figuerola-Ferretti and Gonzalo (2010).

Secondly, we introduce regime shifts depending on an unobserved state variable $st$. The latter can takes two values, $st = 1, 2$, corresponding to two different regimes. This variable follows a first order Markov process with the transition probability matrix,

$$P = \begin{pmatrix} P_{11} & P_{21} \\ P_{12} & P_{22} \end{pmatrix} = \begin{pmatrix} 1 - P_{12} & P_{21} \\ P_{12} & 1 - P_{21} \end{pmatrix} \tag{10}$$

where $P_{12}$ (resp. $P_{21}$) is the probability that the system will shift from state 1 (resp. 2) to state 2 (resp. 1). $P_{11}$ (resp. $P_{22}$) is the probability that the system will stay in regime 1 (resp. regime 2). We obviously have $P_{11} + P_{12} = 1$ and $P_{21} + P_{22} = 1$.

All the coefficients depend on the regime $st$ except for the long-term coefficients, $\beta$. Indeed, variables with a nonlinear cointegrating relationship do not admit an error correction model (Gonzalo and Pitarakis, 2006). In the presence of a cointegrating relationship, the $\Pi_{st}$ matrix is decomposed as $\Pi_{st} = \alpha_{st}\beta'$.

The conditional covariance matrix of error terms, $H_{t,st}$, is regime-dependent, time-varying, and follows a multivariate Garch specification with a Baba et al. (1987) framework, i.e., BEKK, as

$$H_{t,st} = C'_{st}C_{st} + A'_{st}\epsilon_{t-1}\epsilon'_{t-1}A_{st} + B'_{st}H_{t-1}B_{st} + D'_{st}\eta_{t-1}\eta'_{t-1}D_{st} \tag{11}$$

with $\epsilon_{t-1}$ and $H_{t-1}$ being the vector of mean equation residuals and the global covariance matrix for the past period, respectively. $\eta_{t-1}$ is negative past shocks, i.e., $\eta_{t-1} = min(\epsilon_{t-1}, 0)$. $C_{st}$ is a $(2 \times 2)$ lower triangular matrix containing regime-dependent coefficients. $A_{st}$, $B_{st}$ and $D_{st}$ are $(2 \times 2)$ diagonal matrices of coefficients measuring the past shock effects on the conditional covariance matrix, their persistence and the additional effect of a past negative shock, respectively. However, the conditional covariance matrix depends on the sequence of all previous regimes through $H_{t-1}$. With this path-dependence problem, the estimation by the maximum likelihood method is numerically infeasible. To overcome this problem, we follow the formulations of Gray (1996) and Lee and Yoder (2007b) concerning the conditional variances, $h_{ss}$ and $h_{ff}$, and the conditional covariance, $h_{sf}$, respectively, as

$$h_{ss,t} = \pi_{1,t}(r^2_{s,1,t} + h_{ss,1,t}) + (1 - \pi_{1,t})(r^2_{s,2,t} + h_{ss,2,t}) - [\pi_{1,t}r_{s,1,t} + (1 - \pi_{1,t})r_{s,2,t}]^2 \tag{12}$$

$$h_{ff,t} = \pi_{1,t}(r^2_{f,1,t} + h_{ff,1,t}) + (1 - \pi_{1,t})(r^2_{f,2,t} + h_{ff,2,t}) - [\pi_{1,t}r_{f,1,t} + (1 - \pi_{1,t})r_{f,2,t}]^2 \tag{13}$$

$$\begin{aligned} h_{sf,t} = {} & \pi_{1,t}(r_{s,1,t}r_{f,1,t} + h_{sf,1,t}) + (1 - \pi_{1,t})(r_{s,2,t}r_{f,2,t} + h_{sf,2,t}) \\ & -[\pi_{1,t}r_{s,1,t} + (1 - \pi_{1,t})r_{s,2,t}][\pi_{1,t}r_{f,1,t} + (1 - \pi_{1,t})r_{f,2,t}] \end{aligned} \tag{14}$$

In equations (12), (13) and (14), $\pi_{st,t}$ is the probability of being in the state $st$ at the time $t$. $h_{ss,st,t}$ (resp. $h_{ff,st,t}$) is the regime-dependent variance concerning the spot (resp. futures) price at the time $t$ and is contained in $H_{t,st}$. Similarly, $h_{sf,st,t}$ is the state-dependent covariance at the time $t$ and is an element of the same matrix. $r_{s,st,t}$ (resp. $r_{f,st,t}$) is the regime-dependent conditional mean of the spot (resp. futures) price equation at the time $t$. These latter are calculated from the following equations:

$$\epsilon_{s,t} = \Delta s_t - [\pi_{1,t} r_{s,1,t} + (1 - \pi_{1,t}) r_{s,2,t}] \tag{15}$$

$$\epsilon_{f,t} = \Delta f_t - [\pi_{1,t} r_{f,1,t} + (1 - \pi_{1,t}) r_{f,2,t}] \tag{16}$$

This Ms-VEC model is estimated by maximizing of the likelihood function. Each state-dependent error follows a N-dimensional normal distribution with zero mean and $H_{t,st}$ covariance matrix. The global density function is a mixture of these distributions weighted by the probability of being in each regime:

$$\begin{aligned} f(X_t, \theta) = & \frac{\pi_{1,t}}{2\pi} |H_{t,1}|^{-\frac{1}{2}} exp(-\frac{1}{2} \epsilon'_{t,1} H_{t,1}^{-1} \epsilon_{t,1}) \\ & + \frac{\pi_{2,t}}{2\pi} |H_{t,2}|^{-\frac{1}{2}} exp(-\frac{1}{2} \epsilon'_{t,2} H_{t,2}^{-1} \epsilon_{t,2}) \end{aligned} \tag{17}$$

$$L(\theta) = \sum_{t=1}^{T} log f(X_t, \theta) \tag{18}$$

with $\theta$ denoting the parameter vector. The log-likelihood function (equation (18)) is maximized using the expectation-maximisation algorithm proposed by Dempser et al. (1977) under constraints like $\pi_{1,t} + \pi_{2,t} = 1$, $\pi_{1,t} \geqslant 0$ and $\pi_{2,t} \leqslant 1$.

With our specification, we can compute the dynamic hedge ratio as

$$\delta_t | \Omega_{t-1} = \frac{h_{sf,t-1}}{h_{ff,t-1}} \tag{19}$$

where $h_{sf,t-1}$ et $h_{ff,t-1}$ are defined in equations (14) and (13), respectively.

In order to analyze the hedging strategies' performance of each specification[13] we compute hedged portfolios each week and their returns variance over the samples chosen as

$$VAR(\Delta s_t - \delta_t \Delta f_t) \tag{20}$$

[13]We estimate 22 specifications including 8 linear and 14 nonlinear models. Specifications vary about inclusion, or not, of error correction and autoregressive terms in mean equation, asymmetry in variance equation, as well as parameters allowed to switch. In addition, we use an OLS model and a naive model, i.e., with a unit hedging ratio.

In addition, as in Kroner and Sultan (1993) or Alizadeh et al. (2008) among others, we compute the hedger's utility function as

$$E_{t-1}U(\Delta s_t - \delta_t \Delta f_t) = E_{t-1}(\Delta s_t - \delta_t \Delta f_t) - k \times VAR_{t-1}(\Delta s_t - \delta_t \Delta f_t) \tag{21}$$

where $k$ is the degree of risk aversion. This utility function represents economic benefits from the hedging strategy. Another way to consider this benefit is the value-at-risk (VaR) exposure and is computed as

$$VaR = W_0[E(\Delta s_t - \delta_t \Delta f_t) + Z_\alpha \sqrt{VAR(\Delta s_t - \delta_t \Delta f_t)}] \tag{22}$$

where $W_0$ is the initial value of the portfolio and $Z_\alpha$ is the normal distribution quantile.

# 4 Empirical results

Futures contracts on corn-based ethanol were launched on floor-based trading in March 2005. The CBOT launched the ethanol contract on the electronic platform in 2006 contributing to an increase in liquidity within the market. In 2007 options contracts were also launched in the market. For the first time, the volume reached 1,000 contracts in July 2006 and it really took off after 2009 with a sharp increase in the spot prices. During previous decades, and especially in the initial phase of construction of the ethanol futures market, the main objective was to attract and concentrate the liquidity required for commercial traders to achieve hedging activities. Nevertheless, the rise in transaction volumes has been accompanied by a concentration of traders' liquidity on the shortest maturity contracts exchanged on commodity markets. This factor has been observed and studied, for example on the WTI market in the US (Hache and Lantz, 2013).[14] For ethanol futures prices, we observed a decrease in transaction volumes between 2008 and 2016 as contract terms grew longer (Figure 1), and a virtual absence of liquidity for long-term contracts (compared to short-term maturity). In fact, the inadequate information available at any given moment $t$ on contracts whose maturity period is greater than one year does not give traders the incentives to trade in the market. As a consequence, the liquidity for distant contracts at a maturity greater than five months decreases sharply. Moreover the maturity greater than two months registered a sharp decline in transaction volumes after 2012.

On the one hand, by studying available data from 2008 to 2016, we observed a marked rise in transaction volumes for each maturity. Measured in batches of 29,000 gallons (a standard financial contract for ethanol on the CBOT), these transactions

[14] See also the literature review in Lautier (2005).

have risen, for two-month term contracts, from around 78,864 in 2008 to 404,133 in 2016, i.e., multiplied by a factor of 5 (Figure 2). On the other hand, the share of non-commercial players increased from around 15% before 2008 to over 35% on average since 2014 (Figure 3). However, both the increase in the volume of transactions on financial trading floors and the growing share of non-commercial players should be kept in perspective. As mentioned previously, during the previous three decades and especially in the initial phase of construction of the commodities markets, the main objective of the different derivatives marketplaces was to attract and concentrate the liquidity required for commercial traders to achieve hedging activities. In October 1974, the NYMEX launched the first energy contracts for industrial fuel oil. Simon (1984) explains the failure of this first attempt by the under-development of the financial markets and because of the very specific contract specifications (the delivery point of the futures contracts was Rotterdam which held no appeal for the American commercial players). A contract for heating oil in the NYMEX was also launched in 1978 and was abandoned because of inadequate liquidity's volume. During the 1980s in the context of deregulation put in place by the Reagan administration, the NYMEX decided a simultaneous launch of energy contracts: gasoline (1981), crude oil (1983), and heating oil (1990). In Europe the International Petroleum Exchange (IPE) launched its first fuel oil contract in 1981. Since then, financial markets have registered an increase in transactions volume and in the share of non-commercial players in the exchange markets. In the petroleum sector, competition between the two main exchanges (i.e., the NYMEX in New York and the Intercontinental exchange [ICE] in London) led to a strong deregulation process. In the US, for example, the introduction at the end of December 2000 of the law modernizing commodities markets, the Commodity Futures Modernization Act (CFMA), triggered market instability in the crude oil market (Medlock and Jaffe, 2009; Hache and Lantz, 2013).

Figure 1: Open interest by contracts maturity

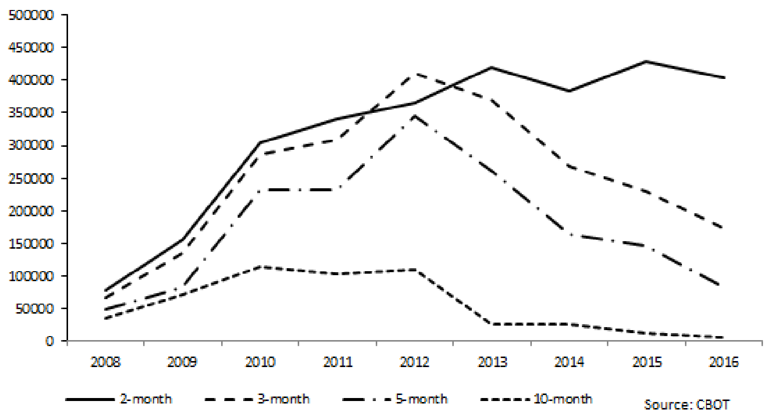


Figure 2: Position (number of contracts) by actors

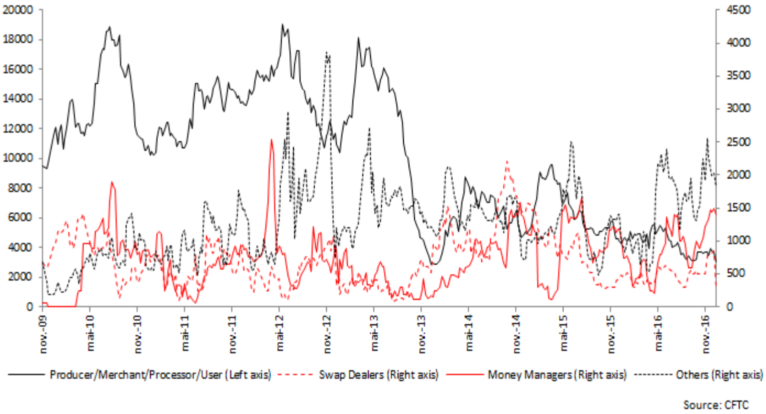
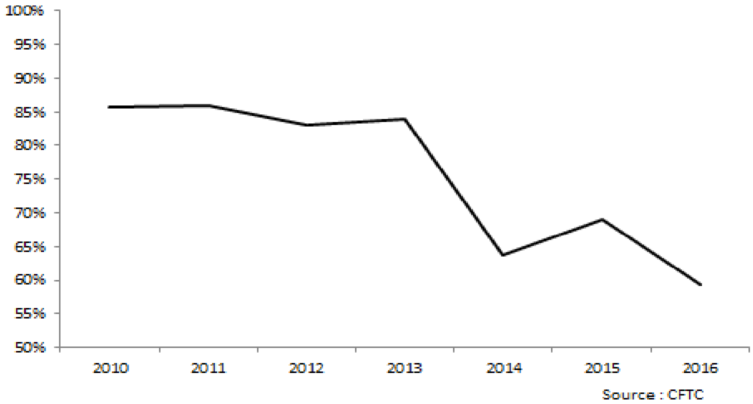



Figure 3: Commercial positions




In order to analyze the ability of Garbade and Silber (1983) and Figuerola-Ferretti and Gonzalo (2010) to explain the ethanol market, we apply Johansen(1988)'s cointegration tests. The results in Table 2 confirm the presence of a long-term relationship between spot and futures ethanol prices regardless of the cointegration test used. The Likelihood Ratio test does not reject the null hypothesis of unit coefficient at a 10% significant level. Thus, the Garbade and Silber (1983)'s theory is a valid explanation of the long-term relationship between spot and futures prices in the ethanol market. Finally, the long-term causality tests conclude in favor of a price discovery process from futures to spot prices, at a 10% significant level. These findings are in line with the informational efficiency of the US ethanol market.

We estimate the Ms-VEC model with two states applied to both the mean and the variance equations. These two states refer to low and high volatility regimes. Table 3 presents results with the Nielsen's cointegration specification.[15] In each

[15] Table 3 presents results for the best models in terms of explanatory power and hedging strategy.

state, only futures prices adjusts to equilibrium ($\alpha_{f,st}$). This result highlights the minor role of futures prices in the discovery process in the short term. Note that the adjustment process is faster during the low volatility regime ($st = 2$), compared to the high volatility regime ($st = 1$). Concerning the short-run dynamics ($\gamma_{ij,st}$), these two markets (spot and futures) seem to be disconnected during normal periods and only past changes of futures prices have a significant impact on spot prices for the high volatility state ($\gamma_{sf,1}$). This last result highlights the fact that futures market can help in understanding the ethanol price dynamics during periods of instability. Furthermore, the relationship is regime-dependent, confirming the ability of our Markov-switching specification to describe it. Figure 4 presents the probability of being in the regime of high volatility.[16] Two main periods of high volatility are observed in 2008-2009 and 2013-2014. Market volatility during these two periods could be explained by the low liquidity during the first one (Figure 1) and by few positions taken by commercial agents for the second period (Figure 2).

Table 2: Cointegration and causality tests

$$\beta_s s_t + f_t + \beta_0 = u_t$$

| Lags | $H_0$ | P-value | | Cointegration vector | LR test | |
|---|---|---|---|---|---|---|
| | | $\lambda_{max}$ test | $\lambda_{trace}$ test | ($\beta_s$ 1 $\beta_0$) | $H_0 : \beta_s$=-1 | $H_0 : \beta_0$=0 |
| 1 | r=0 | 0.001 | 0.001 | (-1.044 1 0.109) | 0.078 | 0.001 |
| - | $H_0$ | Test stat | Critical Value | Cointegration vector | - | - |
| - | r=0 | 3.78 | 3.57 | (-1.010 1 -) | - | - |

| Causality test | P-value |
|---|---|
| Spot to futures prices | 0.867 |
| Futures to spot prices | 0.087 |

Note: The two first lines present the Johansen (1988)'s test results. The lags column mentions the number of lags in the VEC Model. Lag length choice is based on Schwarz (1978)'s Information Criterion. The two P-value columns refer to the P-value of two tests mentioned. P-value inferior to 0.05 leads to the null hypothesis reject of zero cointegrating relationship against one. Cointegration vector column mentions coefficients estimated with $\hat{\beta}_s$ normalized to unity. The LR test checks the existence of a one-to-one relationship between spot and futures prices. We mention the P-value of the test. The next two lines present the Nielsen (2010) test results with the test statistic and the critical value associated at a 5% significance level. The chosen specification is constant and without trend. The null hypothesis is rejected when the test statistic is superior to the critical value. Note that constant is not estimated with this procedure. The causality test refers to the Toda and Yamamoto (1995) test whose null hypothesis is the absence of long-term causality.

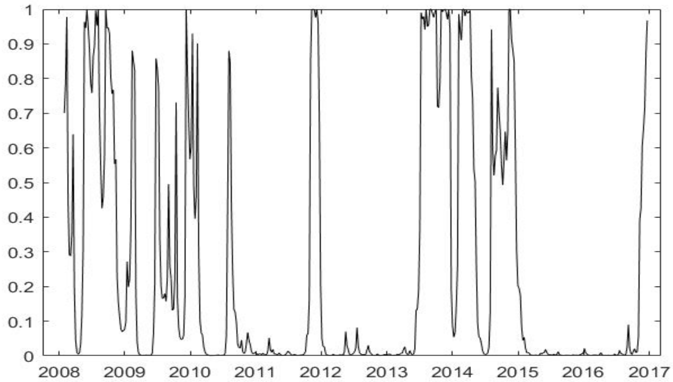
Turning to the conditional variance equation, as expected, we note a high persistence degree ($a^2_{ii,st} + b^2_{ii,st}$ for $st = 1, 2$) during volatile periods. This feature is common to oil and gasoline markets (Fong and See, 2002; Alizadeh et al., 2008). In addition, our specification captures well the leverage effect especially during volatile periods with high and significant coefficients $d_{ii,st}$. Losses in players' portfolio, i.e., negative shocks, have a greater impact on future volatilities than gains, i.e., positive shocks. Finally, the probability of switching from high to low variance states ($P_{12}$) is

We interpret results only for the more explanatory model. Results concerning the 22 specifications are available upon request.

[16]We represent the smoothed probability which provides the best estimation of the states at each time using full-sample information. See Krolzig (1997) for further details on its calculation as well as on other existing probabilities. This figure concerns the Ms-VECM$^N$-Gjr-MGarch. The figure for the VECM$^J$-Ms-MGarch is similar and available upon request to the authors.

greater compared to the probability of switching from low to high variance regimes ($P_{21}$). This result indicates a shorter duration for high volatility regimes and is confirmed by the average expected state duration calculation proposed by Hamilton (1989).[17] This latter result is interesting as it supports the idea of a certain efficiency of the financial market in the short run through the arbitrage process realized by the different players. These durations are nine and twenty weeks for high and low volatility regimes, respectively.

Figure 4: Smoothed probabilities of being in a high volatility state




Our different model specifications allow us to compute the dynamic hedge ratios. We also compute the naive ($\delta = 1$) and OLS hedge ratios of Ederington (1979). We provide information about a non-hedged strategy for comparison purpose. In addition, we compute cross-hedge ratios with gasoline futures markets estimating from our different specifications. These latter will allow us to compare hedging with the ethanol futures market and cross-hedging with the gasoline futures market.[18] The gasoline market could be used by ethanol commercial agents for risk hedging (Franken and Parcell, 2003). Table 4 provides variance, utility and value-at-risk for main specifications and each market,[19] as well as the variance improvement of the best strategy compared to each other. During the final period of our sample, i.e., Panel A, the optimal specification is a VAR-Gjr-MGarch. The lack of high volatility (Figure 4) during this period explains this result, as well as a possible lack of a cointegration relationship. In addition, all cross-hedging strategies underperform both direct hedging strategies and the situation without a hedging strategy.

[17]The average expected duration of state 1 (resp. 2) can be calculated by $(P_{12})^{-1}$ (resp. $(P_{21})^{-1}$).

[18]New York Harbor Reformulated RBOB Regular Gasoline Contract. More details on http://www.cmegroup.com/trading/energy/refined-products/rbob-gasoline.html

[19]A table with all the specifications is available upon request to the authors.

Table 3: Estimation results

| | Ms-VECM$^N$-Gjr-MGarch | | | | VECM$^J$-Ms-MGarch | | | |
|---|---|---|---|---|---|---|---|---|
| $\beta_s$ | -1.011 | (-) | | | -1.044 | (-) | | |
| $\beta_f$ | 1 | (-) | | | 1 | (-) | | |
| $\beta_0$ | - | (-) | | | 0.109 | (-) | | |
| | $st=1$ | | $st=2$ | | | | | |
| $c_{s,st}$ | -0.005 | (0.417) | 0.002 | (0.429) | 0.001 | (0.791) | | |
| $c_{f,st}$ | -0.011 | (0.049) | 0.006 | (0.017) | 0.003 | (0.203) | | |
| $\alpha_{s,st}$ | -0.001 | (0.960) | -0.001 | (0.977) | -0.001 | (0.923) | | |
| $\alpha_{f,st}$ | -0.083 | (0.048) | -0.119 | (0.001) | -0.076 | (0.001) | | |
| $\gamma_{ss,st}$ | 0.037 | (0.801) | -0.145 | (0.390) | -0.114 | (0.140) | | |
| $\gamma_{sf,st}$ | 0.317 | (0.029) | 0.230 | (0.214) | 0.260 | (0.002) | | |
| $\gamma_{fs,st}$ | 0.065 | (0.552) | 0.072 | (0.651) | 0.051 | (0.445) | | |
| $\gamma_{ff,st}$ | 0.025 | (0.822) | -0.071 | (0.686) | -0.027 | (0.733) | | |
| | | | | | $st=1$ | | $st=2$ | |
| $c_{11,st}$ | 0.034 | (0.001) | 0.011 | (0.016) | 0.030 | (0.001) | 0.013 | (0.001) |
| $c_{21,st}$ | 0.037 | (0.001) | 0.029 | (0.001) | 0.034 | (0.001) | 0.029 | (0.001) |
| $c_{22,st}$ | 0.043 | (0.001) | 0.029 | (0.001) | 0.049 | (0.001) | 0.030 | (0.001) |
| $a_{11,st}$ | 0.626 | (0.001) | 0.292 | (0.037) | 0.815 | (0.001) | 0.376 | (0.001) |
| $a_{22,st}$ | 0.309 | (0.060) | 0.189 | (0.144) | 0.536 | (0.001) | 0.256 | (0.027) |
| $b_{11,st}$ | 0.001 | (0.971) | 0.214 | (0.266) | 0.001 | (0.887) | 0.149 | (0.458) |
| $b_{22,st}$ | 0.427 | (0.021) | 0.001 | (0.945) | 0.391 | (0.006) | 0.001 | (0.869) |
| $d_{11,st}$ | 0.657 | (0.025) | 0.320 | (0.053) | - | - | - | - |
| $d_{22,st}$ | 0.571 | (0.014) | 0.442 | (0.005) | - | - | - | - |
| $P_{11}$ | 0.890 | (0.001) | | | 0.880 | (0.001) | | |
| $P_{12}$ | 0.110 | (0.009) | | | 0.120 | (0.001) | | |
| $P_{21}$ | 0.050 | (0.001) | | | 0.056 | (0.001) | | |
| $P_{22}$ | 0.950 | (0.001) | | | 0.944 | (0.001) | | |
| LogL | $1.951\times10^3$ | | | | $1.936\times10^3$ | | | |
| | Spot | Futures | | | Spot | Futures | | |
| JB | 0.001 | 0.001 | | | 0.001 | 0.001 | | |
| Q(6) | 0.160 | 0.665 | | | 0.004 | 0.834 | | |
| $Q^2(6)$ | 0.001 | 0.001 | | | 0.001 | 0.001 | | |

Note: J and N refer to Johansen (1988) and Nielsen (2010)'s cointegration estimation, respectively. For each parameter, we mention the estimated coefficients and the P-value of the Student test in bracket. The coefficient is significant at the 10%, 5% or 1% if P-value is less than 0.10, 0.05 or 0.01, respectively. LogL, JB, Q(6) and $Q^2(6)$ are the log-likelihood, the Jarque and Bera (1980) test for normality, the Ljung and Box (1978) test for autocorrelation and the ARCH test (Engle, 1982) for heterosckedasticity, respectively.

Table 5 displays results of the main hedging strategies for two other panels,[20] i.e., the first half of 2010 and 2012. The VECM-Ms-MGarch with Johansen's cointegration provides the best strategy for both periods. This result confirms the suitability of Markov-switching and Johansen's cointegration specifications for the hedging strategy on the ethanol market. Note that coefficients of this model are consistent with the previous specification presented (Table 3). Hedgers can decrease \$1,268 and \$2,453 of their average weekly value-at-risk with an initial portfolio value of \$1,000,000 compared to the simple OLS specification. These weekly decreases correspond to \$9,144 and \$17,689 annualized decreases, that is to say, only 0.09% and 1.77% of the initial portfolio value. Finally, cross-hedging strategies outperform the non-hedged situation for each period with the OLS and Ms-VECM$^N$-Gjr-MGarch for panels B and C, respectively. This last result highlights the ability of the Nielsen procedure to provide a good hedging strategy.

Table 6 present results for the out-of-sample simulation concerning the period from December 28, 2016, to June 21, 2017, i.e., 25 observations. Concerning non-

[20] A table with all the specifications is available upon request to the authors.

linear specifications, we estimate the model at each point of time to forecast states' probabilities as well as state-dependent conditional mean and variance-covariance matrix. We then compute the prediction of the hedge ratio after recomposition of the global variance-covariance matrix coming from equations (13) and (14). The optimal hedging strategy is the linear multivariate Garch specification with a variance improvement of 80.5% and 22.9% compared to a no-hedged situation and the OLS-based hedge ratio, respectively. However, this strategy does not significantly outperform most of the specifications studied especially the naive strategy consisting of a unit hedge ratio. In addition, cross-hedging with the gasoline market is not efficient compared to the direct hedging strategy. This result is also valid compared to the no-hedged situation for most of the strategies. Furthermore, nonlinear specifications do not seem efficient for hedging in the ethanol market. This result could be explained by the difficulty in well-forecasting the states' probability or by the absence of high volatility periods. Finally, the Johansen (1988) cointegration procedure outperforms the nonparametric approach of Nielsen (2010) for 10 strategies against eight with ethanol markets but out-performs for eight against 10 strategies with cross-hedging in gasoline futures market.[21]

Table 4: In-sample hedging simulation

| | Ethanol spot and futures | | | | Ethanol spot and gasoline futures | | | |
|---|---|---|---|---|---|---|---|---|
| | Var. | V. Impr. | Util. | VaR | Var. | V. Impr. | Util. | VaR |
| Panel A | | | | | | | | |
| No Hedged | 12.82 | 56.5% | -5.127 | 59,073 | **12.82** | **56.5%** | **-5.127** | **59,073** |
| Naive | 5.650 | 1.44% | -2.260 | 39,219 | 36.67 | 84.8% | -14.67 | 99,912 |
| OLS | 6.053 | 8.02% | -2.423 | 40,596 | 18.16 | 69.3% | -7.266 | 70,322 |
| MGarch | 5.640 | 1.28% | -2.256 | 39,187 | 18.96 | 70.6% | -7.583 | 71,841 |
| VAR-Gjr-MGarch | **5.568** | **-** | **-2.227** | **38,933** | 18.18 | 69.3% | -7.271 | 70,348 |
| VECM$^J$-MGarch | 5.592 | 0.43% | -2.237 | 39,019 | 18.92 | 70.5% | -7.569 | 71,774 |
| VECM$^N$-Gjr-MGarch | 5.611 | 0.76% | -2.244 | 39,084 | 18.24 | 69.4% | -7.295 | 70,465 |
| Ms-MGarch | 6.054 | 8.03% | -2.422 | 40,599 | 18.13 | 69.2% | -7.250 | 70,247 |
| VAR-Ms-Gjr-MGarch | 6.467 | 13.9% | -2.587 | 41,960 | 18.36 | 69.6% | -7.346 | 70,708 |
| VECM$^J$-Ms-Gjr-MGarch | 6.596 | 15.5% | -2.638 | 42,376 | 18.17 | 69.3% | -7.267 | 70,327 |
| VECM$^N$-Ms-MGarch | 6.620 | 15.8% | -2.648 | 42,453 | 18.00 | 69.0% | -7.201 | 70,007 |
| Ms-VAR-MGarch | 5.626 | 1.02% | -2.250 | 39,135 | 18.06 | 69.2% | -7.222 | 70,112 |
| Ms-VECM$^J$-Gjr-MGarch | 5.598 | 0.53% | -2.239 | 39,038 | 18.37 | 69.7% | -7.347 | 70,712 |
| Ms-VECM$^N$-MGarch | 5.855 | 4.91% | -2.342 | 39,927 | 18.36 | 69.7% | -7.345 | 70,705 |

Note: Panel A refers to 5/25/16-12/21/16. Variance (Var.) and Utility (Util.) are presented in $10^{-4}$ and $10^{-3}$, respectively. Variance improvement (V. Impr.) measures the incremental variance reduction of the best strategy versus the other strategies with the formula: $[Var(Strategy_i) - Var(Best)]/Var(Strategy_i)$. VaR is in US dollars for an initial investment of \$1 million and $k = 4$. J and N refer to Johansen (1988) and Nielsen (2010)'s cointegration estimation, respectively.

[21] These results come from the comparison between Johansen (1988) and Nielsen (2010)'s approach for each specification including those not presented in Table 4, 5 and 6.

Table 5: In-sample hedging simulation with panel B and C

| | Ethanol spot and futures | | | | Ethanol spot and gasoline futures | | | |
|---|---|---|---|---|---|---|---|---|
| | Var. | V. Impr. | Util. | VaR | Var. | V. Impr. | Util. | VaR |
| Panel B | | | | | | | | |
| No Hedged | 9.774 | 48.3% | -3.910 | 51,584 | 9.774 | 48.3% | -3.910 | 51,584 |
| Naive | 5.797 | 12.9% | -2.319 | 39,728 | 17.87 | 71.7% | -7.148 | 69,751 |
| OLS | 5.395 | 6.50% | -2.158 | 38,325 | **7.752** | **34.9%** | **-3.101** | **45,939** |
| Gjr-MGarch | 5.606 | 10.0% | -2.242 | 39,065 | 7.917 | 36.2% | -3.167 | 46,426 |
| VAR-MGarch | 5.612 | 10.1% | -2.245 | 39,088 | 8.072 | 37.5% | -3.229 | 46,879 |
| VECM$^J$-MGarch | 5.617 | 10.2% | -2.247 | 39,105 | 8.033 | 37.2% | -3.213 | 46,764 |
| VECM$^N$-Gjr-MGarch | 5.645 | 10.6% | -2.258 | 39,201 | 8.069 | 37.4% | -3.228 | 46,869 |
| Ms-MGarch | 5.104 | 1.17% | -2.042 | 37,278 | 7.916 | 36.2% | -3.167 | 46,425 |
| VAR-Ms-MGarch | 5.129 | 1.65% | -2.052 | 37,369 | 7.848 | 35.7% | -3.139 | 46,224 |
| VECM$^J$-Ms-MGarch | **5.044** | **-** | **-2.018** | **37,057** | 7.883 | 36.0% | -3.153 | 46,327 |
| VECM$^N$-Ms-MGarch | 5.160 | 2.24% | -2.064 | 37,480 | 7.818 | 35.4% | -3.127 | 46,134 |
| Ms-VAR-Gjr-MGarch | 5.153 | 2.11% | -2.061 | 37,454 | 7.973 | 36.7% | -3.189 | 46,590 |
| Ms-VECM$^J$-MGarch | 5.251 | 3.94% | -2.100 | 37,808 | 7.822 | 35.5% | -3.129 | 46,148 |
| Ms-VECM$^N$-Gjr-MGarch | 5.105 | 1.19% | -2.042 | 37,280 | 8.232 | 38.7% | -3.293 | 47,339 |
| Panel C | | | | | | | | |
| No Hedged | 10.95 | 75.9% | -4.380 | 54,600 | 10.95 | 75.9% | -4.380 | 54,600 |
| Naive | 3.133 | 16.0% | -1.253 | 29,207 | 14.19 | 81.4% | -5.676 | 62,152 |
| OLS | 2.850 | 7.75% | -1.140 | 27,853 | 11.11 | 76.3% | -4.444 | 54,996 |
| Gjr-MGarch | 3.414 | 22.9% | -1.366 | 30,486 | 11.02 | 76.1% | -4.406 | 54,763 |
| VAR-Gjr-MGarch | 3.309 | 20.5% | -1.324 | 30,016 | 11.09 | 76.2% | -4.438 | 54,957 |
| VECM$^J$-MGarch | 3.288 | 20.0% | -1.315 | 29,919 | 10.96 | 76.0% | -4.385 | 54,633 |
| VECM$^N$-MGarch | 2.800 | 6.10% | -1.120 | 27,608 | 10.95 | 75.9% | -4.379 | 54,594 |
| Ms-MGarch | 2.676 | 1.75% | -1.071 | 26,994 | 10.98 | 76.0% | -4.391 | 54,668 |
| VAR-Ms-MGarch | 2.742 | 4.12% | -1.097 | 27,324 | 10.78 | 75.6% | -4.314 | 54,186 |
| VECM$^J$-Ms-MGarch | **2.629** | **-** | **-1.052** | **26,754** | 10.52 | 75.0% | -4.209 | 53,523 |
| VECM$^N$-Ms-Gjr-MGarch | 2.645 | 0.60% | -1.058 | 26,835 | 10.57 | 75.1% | -4.228 | 53,642 |
| Ms-VAR-Gjr-MGarch | 2.762 | 4.81% | -1.105 | 27,421 | 11.53 | 77.2% | -4.461 | 55,104 |
| Ms-VECM$^J$-MGarch | 2.848 | 7.68% | -1.139 | 27,844 | 10.80 | 75.7% | -4.321 | 54,231 |
| Ms-VECM$^N$-Gjr-MGarch | 2.704 | 2.77% | -1.082 | 27,133 | **9.811** | **73.2%** | **-3.924** | **51,681** |

Note: Panel B and C refer to 1/06/10-8/04/10 and 1/09/12-8/01/12, respectively. Variance (Var.) and Utility (Util.) are presented in $10^{-4}$ and $10^{-3}$, respectively. Variance improvement (V. Impr.) measures the incremental variance reduction of the best strategy versus the other strategies with the formula: $[Var(Strategy_i) - Var(Best)]/Var(Strategy_i)$. VaR is in US dollars for an initial investment of \$1 million and $k = 4$. Figures in bold denote the best-performing model for each market. J and N refer to Johansen (1988) and Nielsen (2010)'s cointegration estimation, respectively.

Table 6: Out-sample hedging simulation

| | Ethanol spot and futures | | | | Ethanol spot and gasoline futures | | | |
|---|---|---|---|---|---|---|---|---|
| | Var. | V. Impr. | Util. | VaR | Var. | V. Impr. | Util. | VaR |
| No Hedged | 13.18 | 80.5%$^{***}$ | -5.271 | 59,895 | 13.18 | 80.5%$^{***}$ | -5.271 | 59,895 |
| Naive | 2.803 | 8.35% | -1.120 | 27,622 | 24.35 | 89.5%$^{***}$ | -9.742 | 81,428 |
| OLS | 3.331 | 22.9%$^{**}$ | -1.332 | 30,116 | 13.32 | 80.7%$^{***}$ | -5.328 | 60.221 |
| MGarch | **2.569** | **-** | **-1.027** | **26,445** | 13.54 | 81.0%$^{***}$ | -5.416 | 60,715 |
| VAR-Gjr-MGarch | 2.647 | 2.95% | -1.059 | 26,847 | 13.45 | 80.9%$^{***}$ | -5.381 | 60,519 |
| VECM$^J$-MGarch | 2.679 | 4.11% | -1.072 | 27,009 | 13.56 | 81.1%$^{***}$ | -5.422 | 60,750 |
| VECM$^N$-Gjr-MGarch | 2.706 | 5.06% | -1.082 | 27,140 | 13.58 | 81.1%$^{***}$ | -5.431 | 60,797 |
| Ms-Gjr-MGarch | 3.040 | 15.5% | -1.216 | 28,768 | 13.46 | 80.9%$^{***}$ | -5.384 | 60,533 |
| VAR-Ms-Gjr-MGarch | 3.077 | 16.5% | -1.231 | 28,942 | 13.92 | 81.5%$^{***}$ | -5.570 | 61,569 |
| VECM$^J$-Ms-Gjr-MGarch | 3.080 | 16.6% | -1.232 | 28,957 | 13.48 | 80.9%$^{***}$ | -5.393 | 60,588 |
| VECM$^N$-Ms-Gjr-MGarch | 3.098 | 17.1% | -1.239 | 29,042 | 13.58 | 81.1%$^{***}$ | -5.432 | 60,805 |
| Ms-VAR-MGarch | 3.021 | 15.0% | -1.208 | 28,678 | 13.65 | 81.2%$^{***}$ | -5.458 | 60,952 |
| Ms-VECM$^J$-MGarch | 3.083 | 16.7% | -1.233 | 28,973 | **13.11** | **80.4%$^{***}$** | **-5.243** | **59,738** |
| Ms-VECM$^N$-Gjr-MGarch | 3.098 | 17.1% | -1.239 | 29,042 | 13.21 | 80.6%$^{***}$ | -5.286 | 59,981 |

Note: Variance (Var.) and Utility (Util.) are presented in $10^{-4}$ and $10^{-3}$, respectively. Variance improvement (V. Impr.) measures the incremental variance reduction of the best strategy versus the other strategies with the formula: $[Var(Strategy_i) - Var(Best)]/Var(Strategy_i)$. Stars (*, **, ***) indicate that the best strategy outperforms the competing model at a 10%, 5% and 1% significance level, respectively. The P-values are provided from White (2000)'s reality check using the stationary bootstrap of Politis and Romano (1994). VaR is in US dollars for an initial investment of \$1 million and $k = 4$. J and N refer to Johansen (1988) and Nielsen (2010)'s cointegration estimation, respectively.

## 5 Conclusion

In this paper, we analyze the ethanol prices dynamics in the US from 2008 to 2016. For this purpose, we use a Markov-switching vector error correction model with an asymmetric Garch error structure. This specification allows us to study the short-term, long-term and variance dynamics across different volatility regimes. From the cointegration test we could not reject the hypothesis of a long-term equilibrium relationship between spot and futures prices. Two distinct states (low and high volatility) should be distinguished for the short-term dynamics. We provide several dynamic hedge ratios and we examine their performance through in-sample and out-sample simulations.

The ethanol market is characterized by its efficiency and a price-discovery process from futures to spot prices in the long term. The cost-of-carry model from Garbade and Silber (1983) is able to well-explain the long-term relationship. In addition, the ethanol futures market can well-explain the spot prices dynamics during the periods of high volatility. Furthermore, hedging strategies based on ethanol futures contracts always outperform the cross-hedging strategy based on the use of gasoline futures contracts. Markov-switching specification and Johansen (1988)'s cointegration procedure are able to provide an efficient hedging strategy for two-third of the periods analyzed. Then , a simple multivariate Garch model is the best hedging strategy during the first half of 2017 according to the out-of-sample simulation. Finally, while Nielsen (2010)'s nonparametric tool provides a clear explanation power for the price dynamics, it cannot be used as a hedging strategy in the ethanol market.

In order to get a full understanding of the different hedging strategies in the financial ethanol market, this paper could be extended in various ways. The methodology used with RBOB gasoline market could be applied to other commodities futures markets such as crude oil, corn or sugar. More globally compared to mature futures market (crude oil, sugar, etc.), the ethanol market was launched in 2005 and the traders' behaviour could have been influenced by many factors such as a lack of information regarding the physical production, commercial strategy (anti-dumping), fiscal policy in producing countries (taxes in Brazil, in the US, etc. and more global uncertainties regarding energy and environmental policies all around the world, etc.). It could explain the fact that the ethanol futures market has not appealed to the traders since the beginning of the last decade but it could become a key market in the near future with the environmental and regulatory constraints of a 2°C scenario.

## References

Alizadeh, A. H., Nomikos, N. K., and Pouliasis, P. K. (2008). A markov regime switching approach for hedging energy commodities. *Journal of Banking and Finance*, 32(9):1970–1983.

Baba, Y., Engle, R. F., Kraft, N. K., and Kroner, K. F. (1987). Multivariate simultaneous generalized arch. University of California, San Diego, unpublished manuscript.

Baillie, R. T. and Myers, R. J. (1991). Bivariate garch estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics*, 6(2):109–124.

Brennan, M. J. (1958). The supply of storage. *American Economic Review*, 48(1):50–72.

Brooks, C., Henry, O. T., and Persand, G. (2002). The effect of asymmetries on optimal hedge ratios. *Journal of Business*, 75(2):333–352.

Cecchetti, S. G., Cumby, R. E., and Figlewski, S. (1988). Estimation of the optimal futures hedge. *The Review of Economics and Statistics*, 70(4):623–630.

Chou, W. L., Fan, D. K. K., and Lee, C. F. (1996). Hedging with the nikkei index futures: The conventional model versus the error correction model. *The Quarterly Review of Economics and Finance*, 36(4):495–505.

Chowdhury, A. R. (1991). Futures market efficiency: Evidence from cointegration tests. *The Journal of Futures Markets*, 11(5):577–589.

Dahlgran, R. A. (2009). Inventory and transformation hedging effectiveness in corn crushing. *Journal of Agricultural and Resource Economics*, 34(1):154–171.

Dempser, A., Laird, N., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38.

Dickey, D. A. and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4):1057–1072.

Ederington, L. H. (1979). The hedging performance of the new futures markets. *The Journal of Finance*, 34(1):157–170.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variances of the united kingdom inflation. *Econometrica*, 50(4):987–1008.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous generalized arch. *Econometric Theory*, 11(1):122–150.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417.

Figuerola-Ferretti, I. and Gonzalo, J. (2010). Modelling and measuring price discovery in commodity markets. *Journal of Econometrics*, 158(1):95–107.

Fong, W. and See, K. (2002). A markov switching model of the conditional volatility of crude oil prices. *Energy Economics*, 24(1):71–95.

Franken, J. R. and Parcell, J. L. (2003). Cash ethanol cross-hedging opportunities. *Journal of Agricultural and Applied Economics*, 35(3):509–516.

Garbade, K. D. and Silber, W. L. (1983). Price movements and price discovery in futures and cash markets. *The Review of Economics and Statistics*, 65(2):289–297.

Garcia, P., Roh, J.-S., and Leuthold, R. M. (1995). Simultaneously determined, time-varying hedge ratios in the soybean complex. *Applied Economics*, 27(12):1127–1134.

Ghosh, A. (1993). Cointegration and error correction models: Intertemporal causality between index and futures prices. *The Journal of Futures Markets*, 13(2):193–198.

Glosten, L. R., Jagannathan, R., and Runke, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801.

Gonzalo, J. and Pitarakis, J.-Y. (2006). *Threshold Effects in Multivariate Error Correction Models*, volume 1: Econometric Theory of *Palgrave Handbook of Econometrics*. Palgrave MacMillam, New-York.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42(1):27–62.

Hache, E. and Lantz, F. (2013). Speculative trading and oil price dynamic: A study of the wti market. *Energy Economics*, 36:334–340.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384.

Hanly, J. (2017). Managing energy price risk using futures contracts: A comparative analysis. *The Energy Journal*, 38(3):93–112.

Heaney, R. (2002). Approximation for convenience yield in commodity futures pricing. *Journal of Futures Markets*, 22(10):1005–1017.

Hull, J. (2005). *Options, Futures, and Other Derivatives.* Prentice Hall, Upper Saddle River, 6 edition.

Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255–259.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3):231–254.

Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford, Oxford university press edition.

Kaldor, N. (1939). Speculation and economic stability. *The Review of Economic Stability*, 7(1):1–27.

Kavussanos, M. G. and Nomikos, N. K. (2000). Hedging in the futures market. *The Journal of Derivatives*, 8(1):41–58.

Krolzig, H.-M. (1997). *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*. Lecture Notes in Economics and Mathematical Systems. Springer,, Berlin.

Krolzig, H.-M. (1999). Statistical analysis of cointegrated var process with markovian regime shifts. Department of Economics, University of Oxford, unpublished manuscript.

Kroner, K. F. and Sultan, J. (1993). Time-varying distributions and dynamic hedging with foreign currency futures. *Journal of Financial and Quantitative Analysis*, 28(4):535–551.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3):159–178.

Lai, K. S. and Lai, M. (1991). A cointegration test for market efficiency. *The Journal of Futures Markets*, 11(5):567–575.

Lamoureux, C. G. and Lastrapes, W. D. (1990). Persistence in variance, structural change, and the garch model. *Journal of Business and Economic Statistics*, 8(2):225–234.

Lautier, D. (2005). Term structure models of commodity prices: a review. *Journal of Alternative Investments*, 8(1):42–64.

Lee, H.-T. and Yoder, J. (2007a). Optimal hedging with a regime-switching time-varying correlation garch model. *The Journal of Futures Markets*, 27(5):495–516.

Lee, H.-T. and Yoder, J. K. (2007b). A bivariate markov regime switching garch approach to estimate time varying minimum variance hedge ratios. *Applied Economics*, 39(10):1253–1265.

Lien, D. (1996). The effect of the cointegration relationship on futures hedging: A note. *The Journal of Futures Markets*, 16(7):773.

Lien, D. and Tse, Y. K. (2002). Some recent developments in futures hedging. *Journal of Economic Survey*, 16(3):357–396.

Lien, D. and Yang, L. (2008). Asymmetric effect of basis on dynamic futures hedging: Empirical evidence from commodity markets. *Journal of Banking & Finance*, 32(2):187–198.

Ljung, G. M. and Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Medlock, K. B. and Jaffe, A. M. (2009). Who is in the oil futures market and how has it changed. James A. Baker III Institute For Public Policy, Rice University, unpublished manuscript.

Myers, R. J. and Thompson, S. R. (1989). Generalized optimal hedge ratio estimation. *American Journal of Agricultural Economics*, 71(4):858–868.

Nielsen, M. O. (2010). Nonparametric cointegration analysis of fractional systems with unknown integration orders. *Journal of Econometrics*, 155(2):170–187.

Perron, P. (1990). Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics*, 8(2):153–162.

Perron, P. and Vogelsang, T. J. (1992). Testing for a unit root in a time series with a changing mean: Corrections and extensions. *Journal of Business and Economic Statistics*, 10(4):467–470.

Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(4):1303–1313.

Routledge, B. R., Seppi, D. J., and Spatt, C. S. (2000). Equilibrium forward curves for commodities. *The Journal of Finance*, 55(3):1297–1337.

Salvador, E. and Arago, V. (2014). Measuring hedging effectiveness of index futures contracts: Do dynamic models outperform static models? a regime-switching approach. *The Journal of Futures Markets*, 34(4):374–398.

Sarno, L. and Valente, G. (2000). The cost of carry model and regime shifts in stock index futures markets: An empirical investigation. *The Journal of Futures Markets*, 20(7):603–624.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Simon, Y. (1984). Les marchés à terme de produits pétroliers. *L'Actualité Economique*, 60(3):375.

Telser, L. G. (1958). Futures trading and the storage of cotton and wheat. *Journal of Political Economy*, 66(3):233–255.

Toda, H. Y. and Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66(1-2):225–250.

White, H. (2000). A reality check for data snooping. *Econometrica*, 68(6):1097–1126.

Working, H. (1948). Theory of the inverse carrying charge in futures markets. *Journal of Farm Economics*, 30(1):1–28.