International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2022, 12(6), 237-247.

Reference Price for the Mexican Crude Oil Mix Export Price: An Alternative Estimation for the Budget and Fiscal Responsibility Law

Alberto Gallegos-David[1]*, Arturo Lorenzo Valdés[2], Bárbara Trejo Becerril[3]

[1]Instituto Tecnológico Autónomo de México, Río Hondo 1, Col. Progreso Tizapán, Álvaro Obregón, Ciudad de México, C.P. 01080, Mexico, [2]Universidad Iberoamericana Puebla, Boulevard del Niño Poblano 2901, Col. Reserva Territorial Atlixcáyotl, San Andrés Cholula, Puebla, C.P. 72810, Mexico, [3]Tecnológico de Monterrey Campus Ciudad de México, Calle del Puente 222, Col. Ex ejido de Huipulco, Tlalpan, Ciudad de México, C.P. 14380, Mexico *Email: alberto.gallegos@itam.mx

Received: 03 May 2022 **Accepted:** 19 August 2022 **DOI:** https://doi.org/10.32479/ijeep.13176

ABSTRACT

This paper aims to perform an alternative methodology the Ministry of Finance and Public Credit (SHCP) applies to estimate the annual Mexican Crude Oil Mix Export Price (MXM), a crucial element of the General Economic Policy Criteria in the Economic Package. We first identify the MXM and the West Texas Intermediate (WTI) relation, computing tail conditional dependence between both series. Subsequently, we use a market risk analysis approach that considers some methodologies to estimate the value at risk (VaR), including an ARIMA-TGARCH model for the innovations of the MXM's price to forecast its behavior using data daily data from January 03rd, 1996, to December 30th, 2021. Once we identify the VaR and the ARIMA-TGARCH components, we aim to design an alternative method to estimate the annual average MXM's price.

Keywords: Oil Prices, Copulas, VaR, TGARCH
JEL Classifications: H27, C51, G15, G22

1. INTRODUCTION

The budgeting process for fiscal authorities in emerging and industrialized countries that depend on the exports of non-renewable commodities relies heavily on how these commodities' price estimations affect short and long-run terms for budgeting purposes. Moreover, terms of trade often play a crucial role in identifying the likely effects of price shocks for food, metal, or energy commodities on the budgeting processes. Therefore, it is significatively important if export revenues for net commodity-exporter countries represent a significant contribution -in the not always well-diversified tax base- among the sources of revenues.

Like in many other net commodity-exporter countries, each year in Mexico, the budgeting process starts with a detailed review of the leading structural indicators stated in the General Economic Policy Criteria of the last Economic Package handed over to Congress. Among the structural indicators, estimating the annual MXM price is crucial to determine oil revenues for the Federal Budget Law that, along with the Federal Expenditures Budget and the General Economic Policy Criteria, is an essential component of the Economic Package of each fiscal year.

Figure 1 shows the evolution of the public revenues in Mexico since 2006. Accordingly, oil revenues have gradually been decreasing, reflecting the exhaustion of the existing oil fields and the limited amounts of public and private investment spent to explore and extract oil from some of the proved oil fields, Mexican authorities have reported. In 2021, oil revenues represented 8.4% of the total public budgetary revenues[1].

This gradual and significant decrease in the contribution of the oil revenues since 2008 explains by the decline in the oil production and export platforms, as Figure 2 shows[2].

In a context of increased price volatility observed during the last economic and financial crises, we apply an approach to market risk management to examine the relationship between the MXM and WTI crude oil prices. This study aims to identify the stochastic process that rules MXM's price statistical behavior, the short and long terms drifters, the parameters of an ARIMA general-time series class of model, and the TGARCH processes that can be involved[3].

Modeling the statistical behavior of commodities prices, specifically crude oil, for predictive purposes usually falls into two workhorse kinds of models. First, those models based on the weak hypothesis of the market's efficiency where crude oil prices follow a random walk without drifters, as Hamilton (2009) pointed out. And the second stream of models, where empirical deviations from the risk prime associated with the expected realization of crude oil prices, determine its statistical behavior. For example, modeling time series analysis to identify ARIMA components, GARCH, and TGARCH are among the most commonly applied.

This paper falls into the second stream of research. Here, we conduct a time series analysis and tail conditional dependence between the distribution of the MXM and the WTI prices as a dependent measure. Furthermore, we apply a market risk analysis approach and some methodologies to estimate the value at risk (VaR) in high volatility periods.

The remaining of this paper organizes as follows. In Section 2, we survey the specialized literature on this subject. Then, in Section 3, we propose the methodology. In section 4, we describe the data and results. Finally, section 5 presents the main conclusions and suggested lines for future research.

2. LITERATURE REVIEW

We choose the Value at Risk (VaR) among the several methods to estimate market risks[4]. The VaR is a statistical measure to estimate potential losses in asset portfolios. According to the capital requirements and risk measures implemented with the Basel Agreements, the VaR methodology is relatively new[5]. However, among several investors, the consensus points out that VaR's origin

1 According to the Ministry of Finance and Public Credit (SHCP), in 2008 oil revenues contributed with 44.2% out of the total revenues, a global maximum.
2 The observed global maximum for the production and the export platforms is 3.4 and 1.9 daily million barrels, respectively, achieved in 2004, according to Mexican Oils Co. (PEMEX).
3 These episodes are the 2008-2009 financial and economic crisis and the 2020-2021 sanitary and economic crisis.
4 Review Angelidis et al. (2004) and Grajales and Pérez (2010) for a survey of the different methodologies that have been proposed to measure market risks.
5 Basel I, II and III agreements represent the institutional efforts to improve international banking regulations under the Basel Committee of the Bank of International Settlements since 1974.

Figure 1: Evolution of oil revenues and their contribution to the public revenues

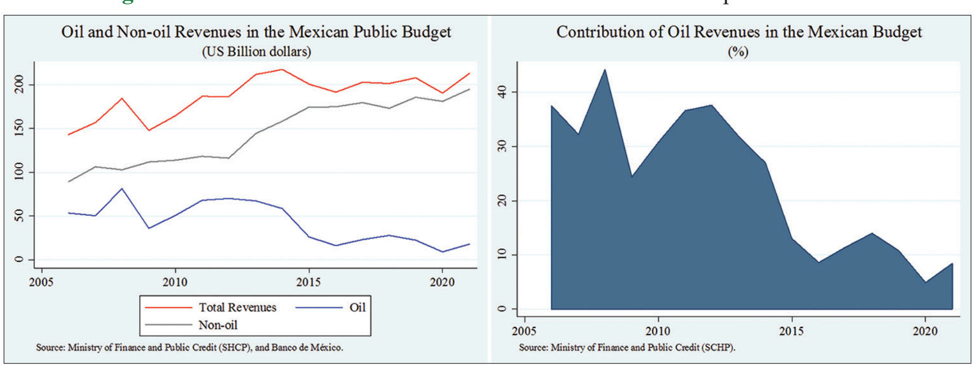
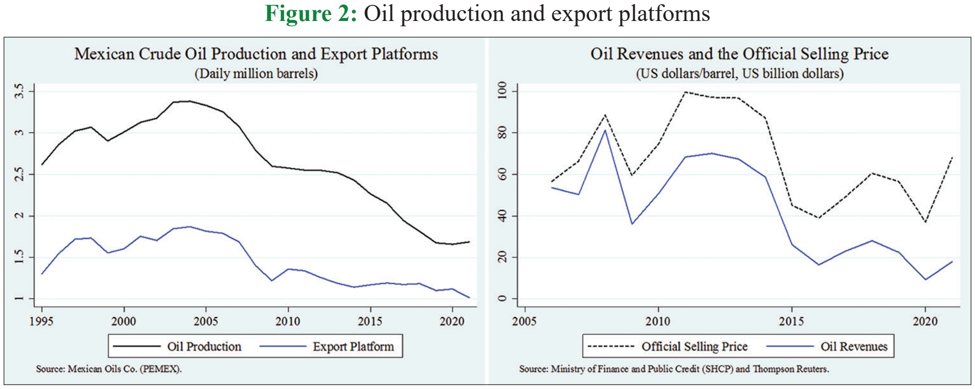


Figure 2: Oil production and export platforms



is the technical document RiskMetrics that JP Morgan released in the mid-90s[6].

In applying the market risk measure approach through the VaR and TGARCH estimation, we propose to minimize the risk of loss of oil revenues due to an over or underestimation of the annual average oil price for the MXM in a context of high volatility observed in energy commodities markets during the most recent economic and financial crisis. The aim is to calculate a price estimation that reflects the lesser loss of the oil revenues.

Grajales and Pérez (2010) point out that considering volatility ensures efficiency estimations from the previously known procedures applying the VaR methodology.

On the other hand, Wets and Rios (2015) proposed a new methodology considering standard short and long-run terms to identify time series drifters for each period, some mean reversibility in the stochastic process, and some specific information about the state of the commodity's market for copper.

Bautista and Mora (2019) pointed out that determining the risk in periods of high volatility is critical because it directly impacts the provision of reserves necessary to face potential losses. Subsequently, they calculated the VaR for the prices of Brent, WTI, and MXM using different distributions for the innovation process. Finally, they found that the VaR-stable model is more robust and accurate.

According to Schwartz (1997), the stochastic behavior characterizes by a mean reversion process, the standard for several commodity prices like crude oil. Ramírez et al. (2019) adapted a stochastic difference-equation to the series of Mexican oil prices. They found that a persistent reversion to the long-term mean shocks produced on real prices does not involve permanent changes.

Likewise, Dafas (2004) characterized the stochastic behavior of crude oil prices as mean-reverting and Markov-switching jump-diffusion processes. Dafas (2004) applied the expectation-maximization algorithm and the Hamilton filter to estimate the model's parameters.

According to Dafas (2004), jumps in the spot price of crude oil are well captured through a Poison process assuming a constant rate of events. With these assumptions, Dafas (2004) pointed out that the probability distribution of the logarithm of the oil spot price has fatter tails and a thinner body than the normal distribution.

On the other hand, Kilian (2009) and Kilian et al. (2009) set up a VAR model to identify the determinants of oil prices. They assume a combination between global aggregate demand and precautionary demand shocks that indirectly reflect disturbances from the supply side that explain the drivers of oil prices globally[7].

6 At the end of the 70s and early 80s several companies started to design risk measure systems. RiskMetrics was designed by JP Morgan in the mid-90s.

7 By precautionary-demand, Kilian (2009) refers to oil demand explained by the uncertainty about expected shortfalls of supply relative to expected demand, i.e., inventories of crude oil by precautionary purposes.

Unlike Kilian et al. (2009) and Ramirez et al. (2019), in this paper, we do not pursue the identification of the sources of variation of the MXM's price based on its supply and demand conditions nor some of the deviations proposed in Hamilton (2009). Instead, we aim to determine through a time series analysis the kind of stochastic process that generates the MXM's price in a context of high volatility and the relation with WTI by calculating tail conditional dependence[8].

In so doing, we carry out the following steps, (i) we measure the relationship between MXM and WTI-with which the MXM's price is highly dependent-, by calculating dependency measures using copulas, explicitly conditional probabilities, (ii) we estimate an ARIMA-TGARCH model for MXM price returns and volatility, (iii) we estimate the price of oil with specific risk management criteria (VaR) and compare it with the Mexican government's proposal.

In this context, in a Monte Carlo simulation, Boutouria and Abid (2010) identified that through a time-continuous stochastic process involving the variance, convenience yield, and the interest rate, the stochastic volatility is one of the specific factors determining copper prices.

Zhang and Zhang (2015) pointed out that since 2009, WTI and Brent crude oil benchmark prices have been experiencing some abnormal spreads between them and increased volatility. Zhang and Zhang (2015) proposed a Markov-based switching model with dynamic autoregressive coefficients to identify WTI and Brent crude oil prices volatility regimes, suggesting a *sui generis* classification from "sharply downward" to "sharply upward" and "relatively stable" regimes.

3. METHODOLOGY

Commodities price statistical behavior is comparable with asset-portfolio returns described by probability density functions with fat tails, skewness, and high volatility, reflecting extreme events that normal distributions do not replicate. The MXM's price density function reflects both features-fat tails and skewness-and high volatility, confirming that the assumptions of normal distributions do not fit these types of time series.

In Mexico and several emerging market commodity-exporter economies, the Ministry of Finance and Public Credit estimates the annual average price of the MXM for budgetary purposes[9]. Calculations involve the following variables:

$$P_{MXM_s} = \theta\left\{\alpha\frac{\sum_{t=-120}^{s-1} p_t^{MXM}}{120} + (1-\alpha)\frac{\sum_{j=s+40}^{m} p_j^{WTI}}{m-(s+40)}\right\} + \\ +(1-\theta)\delta\gamma\frac{\sum_{k=s+3}^{l} p_k^{WTI}}{l-(s+3)} \quad (1)$$

8 Explicitly, the aim is to determine the order of the general ARIMA-TGARCH model that explains the short and long terms evolution of the MXM's price for predictive purposes.

9 See Articles 31st and 15th of the Budget and Fiscal Responsibility Law (Ley Federal de Presupuesto y Responsabilidad Hacendaria) and its bylaw (Reglamento de la Ley Federal de Presupuesto y Responsabilidad Hacendaria). The resulting price enters as an annual average price for the next fiscal year as of the day of its approval for the Low Chamber of the Congress.

with:

$$\theta = \frac{1}{2} = \alpha$$

where:

- s stands for the calculation time, generally at the end of August of each year
- t runs from the observed monthly price 10 years before the calculation time
- j considers at least 3 years as of the time of calculation s
- k runs from December of the year of calculation to November of the estimated fiscal year
- δ is the observed price difference in percentage between the WTI and the MXM during the estimation period, usually between April and August
- γ=0.84 is a conservative factor to underestimate oil revenues for budgetary purposes.

This study uses pure risk management criteria to contrast them with the SHCP's methodology. Likewise, we use different methods to calculate the VaR, describing the stochastic process and the relation between the MXM and WTI returns.

3.1. TGARCH Modeling

We use daily prices to build the series of the oil rate returns under the assumption that returns on m days follow a compound interest process[10]. The compound interest distributions are estimations that use time-series models of the ARCH-GARCH family[11].

Here we show the TGARCH model used to estimate the marginal density distributions of the innovations associated with the oil rate returns. This model extends the traditional GARCH model proposed by Zakoian (1994)[12]. The TGARCH model has been recognized among the best to describe the statistical behavior of asset returns in developing economies. Moreover, the TGARCH model can capture some features that characterize many financial and economic series. Worth mentioning is the existence of non-constant volatilities, skewed and leptokurtic, volatility clustering, distributions, and leverage effects.

From a modeling perspective, the main feature of the TGARCH model is that it allows the volatility of the return series on period t, r_t, to depend on the "news" arriving at the market (i.e., the lagged innovation u_{t-1}). We describe such volatility with the following specification of the conditional variance of the innovations, σ_t^2 :

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \gamma u_{t-1}^2 I\left(u_{t-1} < 0\right) + \beta \sigma_{t-1}^2 \qquad (2)$$

where innovations u_t are, by assumption, distributed as a normal distribution. The parameters α_0, α_1, β, and γ are non-negative by assumption, and I defines as an indicator function:

$$I = \begin{cases} 1\, if\ u_{t-1} < 0 \\ 0\, if\ u_{t-1} \geq 0 \end{cases}$$

The specification of the conditional variance given by the expression (2) allows us to analyze the effects of qualitative news on the current volatility of the return series: (i) good news, $u_{t-1} > 0$, have an effect equal to α_1 on σ_t^2 , and (ii) bad news, $u_{t-1} < 0$, have an effect equal to $\alpha_1+\gamma$. Thus, when $\gamma \neq 0$, bad news has measurable effects on the volatility of these series. Mainly, when bad news occurs and $\gamma > 0$, these series show the "leverage effect" (i.e., the volatility caused by bad news is more significant than the one caused by the good news). Therefore, γ could be considered a measure of the sensitivity to bad news prevailing in the market.

We use the AR (1)-TGARCH(1,1) model with a normal distribution to estimate the marginal distributions of the oil rate returns. This model has a three-equation system structure. The first expression is the conditional mean of the series of returns, r_t, during the period t. The second one is the condition that defines an ARCH process. The third one is the specification of the conditional variance. The structure that defines the TGARCH models estimated is:

$$r_t = \phi_0 + \phi_1 r_{t-1} + u_t\ u_t = \sigma_t \varepsilon_t\ \sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \gamma u_{t-1}^2 I\left(u_{t-1} < 0\right) + \beta \sigma_{t-1}^2 \qquad (3)$$

We use some complementary tests to validate the estimation procedure. Specifically, we use ADF and KPSS tests to assess the order of integration of the log's series of the oil prices. We use both tests due to their complementarity and to avoid spurious estimations[13]. In addition, we use ARCH-LM tests of the type proposed by Engle (1982) to examine the convenience of using models of the ARCH-GARCH family for modeling and analyzing the series of returns. Furthermore, we use Ljung-Box tests of the type proposed by Ljung and Box (1978) to assess potential misspecification problems.

3.2. Copula Modeling

Copula modeling allows us to describe multivariate distribution functions through their marginal distribution functions and copula's dependence function. Nelsen (1999) presented both the theoretical and practical characteristics. The aim is to isolate the dependence structure from the structure of the marginal distributions.

A bivariate copula $C(u_1, u_2)$ is a cumulative distribution function (CDF) with uniform marginal distribution functions on the unit interval. Sklar's theorem (Sklar, 1959) states that if $F_j(x_j)$ is the CDF of two univariate continuous random variables X_j for any $j = 1,2$, then $C(F_1(x_1), F_2(x_2))$ is a bivariate CDF for $X(X_1, X_2)$ with marginal distributions F_j. Conversely, if F is a continuous bivariate CDF with univariate marginals F_1, F_2, then there is a unique bivariate copula C such that $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$.

10 We define the *m-return* for an asset during the day t , r_t , as the change in logs of the price on m days of such asset, P_t. Therefore $r_t = InP_t - InP_{t-m}$.

11 The ARCH-GARCH family includes more than a hundred time-series models. Particularly, the ARCH and GARCH acronyms stand for Auto Regressive Conditional Heteroscedasticity and Generalised Auto Regressive Conditional Heteroscedasticity. These time-series models have their origins in Engle (1982) and Bollerslev (1986).

12 The TGARCH acronym stands for Threshold Generalized Autoregressive Conditional Heteroscedasticity model.

13 The ADF and the KPSS tests have complementary null hypotheses. The null hypothesis of the ADF test is that the data generating process contains a unit root. The null hypothesis of the KPSS test is that the data generating process is stationary. The joint use of both tests allows us to guarantee the estimation of robust results regarding the order of integration of the series.

The properties of copulas allow us to study dependencies more easily in financial markets. Among these properties is that copulas are invariant to monotone transformations of random variables. Secondly, as widely used by Kendall's tau, there is a direct relationship between copulas parameters and concordance measures (Kendall, 1938). Third, they provide an asymptotic dependence treatment in the tails of the distributions.

Kendall's tau (τ) is a measure of concordance between two random variables. Two points (x_1, x_2) and (y_1, y_2) are said to be concordant if (x_1–y_1) (x_2–y_2) >0, and discordant if (x_1–y_1) (x_2–y_2) >0. Likewise, two random vectors (X_1, X_2) and (Y_1, Y_2) are concordant if the probability $P[(x_1-y_1)(x_2-y_2)>0]$ is greater than $P[(x_1-y_1)(x_2-y_2)<0]$; that is X_1 and X_2 tend to increase together. They are discordant if the opposite happens. Kendall's τ measures differences in probability:

$$\tau(X_1,X_2)=P\left[(X_1-Y_1)(X_2-Y_2)>0\right] -P\left[(X_1-Y_1)(X_2-Y_2)<0\right] \tag{4}$$

Kendall's τ is related to copulas through the following equation:

$$\tau(X_1,X_2)=4\iint C(u_1,u_2)\,dC(u_1,u_2)-1 \tag{5}$$

Additionally, an alternative dependence measure defined by copulas is the asymptotic tail dependence, which measures the performance of random variables during extreme events. This paper uses a parameter to estimate the probability that an extreme increase (decrease) in MXM returns occurred if we observe an extreme increase (decrease) in WTI returns.

The lower τ^L and upper τ^U asymptotic tail dependence coefficients definitions are:

$$\begin{aligned}\tau^L &= \lim_{\alpha\to 0^+} P\left(X_2<F_2^{-1}(\alpha)\mid X_1<F_1^{-1}(\alpha)\right) = \lim_{\alpha\to 0^+}\frac{C(\alpha,\alpha)}{\alpha},\\ \tau^U &= \lim_{\alpha\to 1^-} P\left(X_2>F_2^{-1}(\alpha)\mid X_1>F_1^{-1}(\alpha)\right) = \lim_{\alpha\to 1^-}\frac{1-2\alpha+C(\alpha,\alpha)}{1-\alpha}\end{aligned} \tag{6}$$

According to equation (6), there is independence in the tail if the value is zero and dependence if the value is between zero and one. Furthermore, if the value is one, there is a perfect dependence.

In doing so, we use Clayton and Gumbel copulas. The Clayton copula shows lower tail dependence, while the Gumbel copula shows upper tail dependence. The bivariate Clayton copula is given by:

$$C_\theta^{CL}(u_1,u_2)=\left(u_1^{-\theta}+u_2^{-\theta}-1\right)^{-\frac{1}{\theta}} \tag{7}$$

where: $\theta\in[-1,\infty)\setminus\{0\}$. In this case $\tau^L=2^{-\frac{1}{\theta}}$. The bivariate Gumbel copula is given by:

$$C_\theta^{GU}(u_1,u_2)=\exp\left[-\left((-\ln u_1)^\theta+(-\ln u_2)^\theta\right)^{\frac{1}{\theta}}\right] \tag{8}$$

where: $\theta\in[-1,\infty)$. In this case $\tau^U=2-2^{-\frac{1}{\theta}}$.

Following Patton (2006), we estimated the conditional tail dependence measures. The lower τ^L and upper τ^U tail coefficients are supposed to be time-dependent. The evolving dynamic is as follows:

$$\begin{aligned}\tau^L &= \Lambda\left(\lambda_{0L}+\lambda_{1L}\tau_{t-1}^L+\lambda_{2L}\left|u_{1t-1}-u_{2t-1}\right|\right),\\ \tau^U &= \Lambda\left(\lambda_{0U}+\lambda_{1U}\tau_{t-1}^U+\lambda_{2U}\left|u_{1t-1}-u_{2t-1}\right|\right)\end{aligned} \tag{9}$$

Where: Λ is the logistic transformation used to keep the values between zero and one. To describe the marginal behavior, we use the closing MXM and WTI oil prices at time t, i.e., P_t. Continuous returns calculations followed equation (10):

$$r_t=\ln P_t-\ln P_{t-1} \tag{10}$$

We also examined the oil returns volatility using a TGARCH model. For the conditional mean, we used an AR(1)-TGARCH as in equation (3) modeling.

3.3. The Value at Risk *(VaR)*

The Value at Risk (*VaR*) is a statistical measure to estimate potential losses in asset portfolios. The *VaR* measurements have two main features: firstly, it represents the amount of the maximum loss in an asset portfolio with a given likelihood, and secondly, *VaR* estimations consider correlations among different risk factors. Both features of the *VaR* computations allow investors to assign capital to various assets. To express the *VaR* calculations, we use the following analytical way:

Be R the set of all historical risk market factors[14] that have effects on a given financial asset portfolio, where R is the matrix $(n+1)\times m$, with $(n+1)$ representing the number of periods and m the number of risk factors, i.e.,

$$R=\left\{\overline{r}_0,\overline{r}_1,\ldots\overline{r}_n\right\} \tag{11}$$

Every component in (11) represents the value all factor risks take in a given date, with m as the number of factors for each component $\overline{r}_i$, where m is its dimension, i.e., $\left|\overline{r}_i\right|=m$. Therefore, $\overline{r}_i^T=\left\{r_i^1,r_i^2,\ldots r_i^{m-1},r_i^m\right\}$, where m is, as already mentioned, the number of risk factors that affect the value the hypothetical portfolio has at the time i, where T stands for the transposed vector. Consequently, the value of the risk factors for the present scenario is $\overline{r}_0^T=\left\{r_0^1,r_0^2,\ldots r_0^{m-1},r_0^m\right\}$.

For the data in R, we design a set of scenarios $\overline{S}_1,\overline{S}_2,\ldots\overline{S}_n$ where every $\overline{S}_j\ \ \forall j=1,2,\ldots,n$, identifies the risk factor scenario

14 In practice, it refers to the database of the market factor risks.

explicitly given the observations in $\bar{r}_j \wedge \bar{r}_{j+1}$. Where $\bar{S}_j$ is the forecast of the risk factors in a given period, known as the "*holding period*," given the risk factors at the $j \wedge j+1$ dates.

3.4. Scenarios

Be the $j-th$ scenario $\bar{S}_j$:

$$\bar{S}_j = g\left(\bar{r}_0, f\left(\bar{r}_j, \bar{r}_{j+1}\right)\right) \forall j = 1,2,\ldots,n \quad (12)$$

In (12) $f\left(\bar{r}_j, \bar{r}_{j+1}\right) = \left(\frac{r_j^1}{r_{j+1}^1}, \frac{r_j^2}{r_{j+1}^2}, \cdots, \frac{r_j^m}{r_{j+1}^m}\right)$ is the quotient among risk factors, whereas g (.)is the value of the risk factors in the $j-th$ scenario given the returns among the data from $j \wedge j+1$ and the present value of given risk factors, i.e.,

$$\bar{S}_j = g\left(\bar{r}_0, f\left(\bar{r}_j, \bar{r}_{j+1}\right)\right) = \left(r_0^1 \cdot \frac{r_j^1}{r_{j+1}^1}, r_0^2 \cdot \frac{r_j^2}{r_{j+1}^2}, \cdots, r_0^m \cdot \frac{r_j^m}{r_{j+1}^m}\right) \quad (13)$$

This general methodology considers the risk factors database; however, taking the specific case of a single risk factor, (13) can be re-expressed in the following way:

$$\bar{S}_j = g\left(r_0, f\left(r_j, r_{j+1}\right)\right) = r_0 f\left(r_j, r_{j+1}\right)$$
$$= r_0\left[1 + \frac{r_{j+1}}{r_j - 1}\right] = r_0\left[1 + \underset{j}{\rho}_{j+1}\right] \quad (14)$$

where: $\underset{j}{\rho}_{j+1}$ represents the percentage change of the observed factor risk from the date j to j+1, according to eq. (15)[15]:

$$_j\rho_{j+1} = \frac{r_{j+1} - r_j}{r_j} = \frac{r_{j+1}}{r_j} - 1 \quad (15)$$

Equations (13), (eqn: jscen_314), and (15) specifically show how the multiplicative scenarios of a historical simulation are applied to obtain the *VaR*.

3.5. Loss and Profit Forecasts

When the scenarios are defined, we obtain the theoretical value in each of the n scenarios, which constitutes the corresponding forecast to the "*holding period*." Let $V_j = V\left(\bar{S}_j\right)$ be the theoretical value of the instrument in the $j-th$ scenario, therefore, $\bar{V} = \left(V\left(\bar{S}_1\right), V\left(\bar{S}_2\right), \ldots, V\left(\bar{S}_n\right),\right)$ represents the series of theoretical values of the instrument i in each of the n scenarios.

In this context, the function V(.)[16] vary depending on the investment portfolio under analysis, given the aggregated valuation of the portfolio components. Therefore, given that $V_0 = V\left(\bar{r}_0\right)$ identifies the base value of the portfolio, or the known "*mark to market*," we calculate loss and gains in each scenario like $PnL_j = V_j - V_0 \forall j = 1,2,\ldots n$; where the difference represents the change in the portfolio value concerning the base value under the $j-th$ scenario[17].

In doing so, the loss and profit series *PnL* obtained after reassessing each of the $j-th$ scenarios are:

$$PnL = \left(PnL_1, PnL_2, \ldots, PnL_n\right) \quad (16)$$

Finally, in equation (16), we arrange elements of *PnL* in an ascending way:

$$PnL = \left(PnL_{j:1}, PnL_{j:2}, \ldots, PnL_{j:n}\right) \quad (17)$$

In equation (17) the time series is already ordered, therefore:

$$1\ PnL_{j:1} \leq PnL_{j:2} \leq \cdots \leq PnL_{j:n} \quad (18)$$

In equation (18), j represents that any *PnL* from the original time series could be in the first position, any *PnL* in the second position, and so forth. In contrast, the index numbers represent an ascending order, i.e., a statistic order.

3.6. *VaR* by Historical Simulation

The estimation of the *VaR* of an asset portfolio by Historical Simulation with a given α confidence level $\left(VaR_{HS}^{\alpha}\right)$, leads us to the identification of the element inside k or $PnL_{j:k}$ in *PnL*, which corresponds to the level of the required confidence α:

$$\begin{aligned} k &= n\left(1-\alpha\right) \\ VaR_{HS}^{\alpha} &= PnL_{j:k} \end{aligned} \quad (19)$$

In doing so, eq. (eqn: var_119) allows us to obtain the *VaR* estimation through the Historical Simulation methodology. The *VaR* analysis identifies the maximum expected loss for a given time-setting and confidence level.

$$VaR_{\alpha} = \inf\left\{PnL_{\{K\}} : P\left(L > PnL_{\{K\}}\right) \leq 1-\alpha\right\} \quad (20)$$

In eq. (20), VaR^{α} represents the Value at Risk at the α percent, $PnL_{\{K\}}$ denotes the *PnL* in the k scenario of the ordered series *PnL*, and $P\left(L > PnL_{\{K\}}\right) \leq 1-\alpha$ shows that the likelihood of observing a loss or gain is more significant than that estimated with the *VaR* for a given asset portfolio is α.

3.7. *VaR* Normal Delta (*VaR−δN*) and *VaR*-TGARCH

Among the parametric models, the estimation of the $VaR-\delta N$ assumes that the density function of the asset portfolio returns is known, with parameters not previously identified. Notably, this methodology assumes normality in the distribution of the daily returns with given first and second moments $R_p \sim N(\mu, \sigma^2)$. In this context, we obtained *VaR* estimations that reflect the change in the asset portfolio value from one day to another so that the loss or gain, like in equation (21):

15 This is the period that corresponds to the "holding period" that generally is 1 day.

16 The V(.) function implies the underlying valuation models.

17 This change is known as the "Profit and Loss" or *PnL*.

$$VaR = V_0 \times R_p^* \tag{21}$$

where: V_0 is the initial value of the asset portfolio, and R_p^* represents the daily return with an α confidence level. Given the density function $R_p \sim N(\mu,\sigma^2)$, therefore $x = \frac{R_p - \mu}{\sigma} \sim N(0,1)$.

Then, for a given level of confidence α:

$$P\left(x \le \frac{R_p^* - \mu}{\sigma}\right) = 1 - \alpha \tag{22}$$

If Φ is the density function $N(0,1)$, then:

$$\Phi^{-1}(1-\alpha) = \frac{R_p^* - \mu}{\sigma} \tag{23}$$

where: Φ^{-1} is the inverse function of the normal standard distribution. Assuming that the expected return is zero, we have:

$$R_p^* = \sigma \times \Phi^{-1}(1-\alpha) \tag{24}$$

Therefore, the *VaR* estimation for a one-day time setting is:

$$VaR = V_0 \times \sigma \times \Phi^{-1}(1-\alpha) \tag{25}$$

Furthermore, if we assume that daily returns are independent and identically distributed, we have that the $VaR\text{-}\delta N$ for a given time setting of t days is:

$$VaR - \delta N = V_0 \times \sigma \times \Phi^{-1}(1-\alpha) \times \sqrt{t}. \tag{26}$$

Finally, the calculation of *VaR*-TGARCH risk measure for the oil rate of return is:

$$VaR_\alpha^t = \mu_{t+1} + \sigma_{t+1} q_\alpha(Z) \tag{27}$$

where: Z is a random variable with a normal distribution function and $q_\alpha(Z)$ is the *α–quantil* of Z.

4. DATA AND RESULTS

For this paper, we retrieved data based on the series of the daily closing Mexican Crude Oil Mix Export prices (MXM) from the Central Bank of Mexico (Banco de México)[18] and the West New York Mercantile Exchange Contracts (NYMEX) from the US Energy Information Administration for the period January 3rd, 1996 to December 30th, 2021.

Table 1 presents descriptive statistics of daily returns. The last two columns are Pearson's correlation and Kendall's Tau, respectively, which measure the dependence between the MXM and WTI prices.

We confirmed the non-normality of MXM returns except for 2005, 2006, and 2010 by the Jarque-Bera statistic based on kurtosis and skewness. Furthermore, the Pearson's correlation coefficient is a measure of linear dependence: (i) if this value is equal to 1 (a positive one), there is a perfect positive linear dependence, (ii) if it is equal to-1 (a negative one), there is a perfect negative linear dependence. Therefore, we confirmed the positive linear correlation between the MXM and WTI returns except for 2014. On the other hand, Kendall's tau determines the positive dependence except for 2014, like the Pearson's correlation.

Figure 3 shows the graphs of the MXM and WTI daily prices and their daily returns. Once more, the observed evolution is seemingly related, confirming the high dependence measured by Pearson's correlation and Kendall's tau.

4.1. Copula Results

To obtain the time series of tail dependence, we use the results of the Clayton copulas calculations in equation (7), the Gumbel copulas in equation (8), and their own marginal returns behavior as in equation (3). Figure 4 shows these results.

Clayton copulas suggest that reductions in the WTI oil returns likely caused declines observed in the MXM oil returns. Similarly, Gumbel copulas indicate that increases in WTI oil returns likely caused raises in the MXM oil returns.

These effects are very similar in lower probabilities and show autoregressive behavior since the parameters of the equation for both-lower and upper tail dependence-are statistically significant. However, this probability varies in the period under analysis, ranging from 61.99% to 84.6% for lower dependence (Clayton) and 58.56% to 84.27% for upper dependence (Gumbel).

In this context, the estimation of the Clayton and Gumbel copulas indicates a high likelihood that an extreme increase (decrease) in WTI oil prices will result in an extreme increase (decrease) in MXM prices, confirming the high dependence between the MXM and the WTI returns.

4.2. *VaR* Results

Table 2 shows in its second column the average observed daily price of the Mexican Crude Oil Mix Export (MXM), the reference price of the MXM estimated by the SHCP for budgetary purposes (third column), whereas columns fourth to sixth show several reference price proposals based on the calculation of the normal parametric VaR (fourth), historical VaR (fifth) and the 4-month VaR-TGARCH considering a level of $\alpha = 0.1$ and an 8-year window (sixth).

Figure 5 shows the evolution of the observed annual average price of the MMX, SHCP's estimations, and our alternative proposed calculations (VaR Normal, VaR Historical, and VaR-GARCH) comparing their Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE).

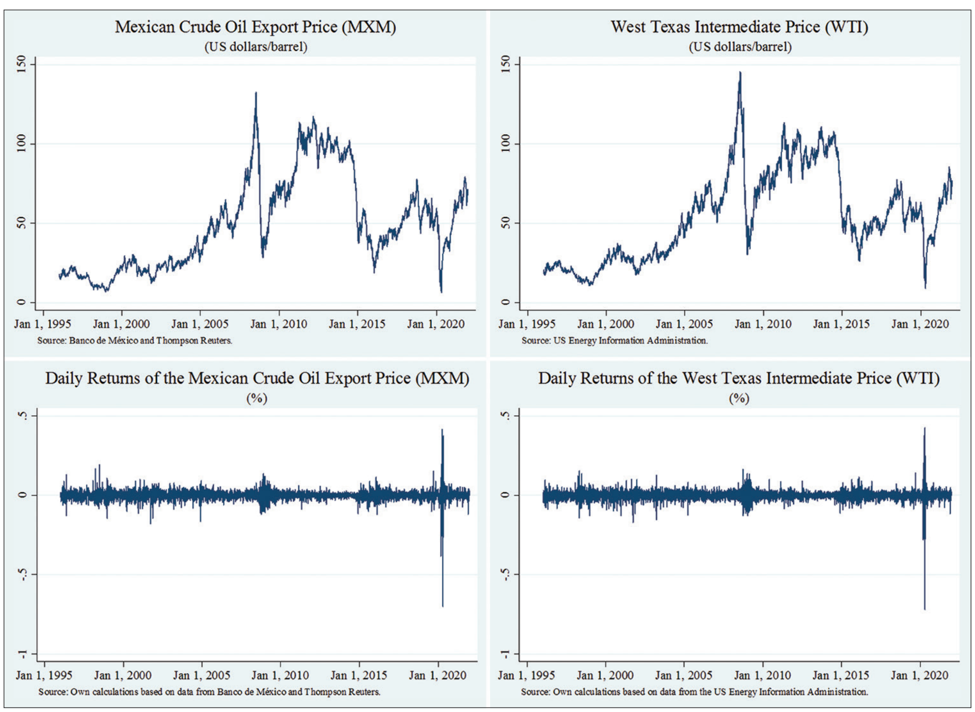
18 Precio del petróleo (2022). Mezcla Mexicana, Dólares por Barril, PMI. (2022).

Table 1: Descriptive statistics for the annualy continuous returns for the MXM and the WTI crude oil prices. Pearson×s correlation and Kendall's tau between the MXM and the WTI crude oil prices

Year	Mean	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	P-value	$\rho_{MXM, WTI}$	$\tau_{MXM, WTI}$
1996	0.0008	0.0241	0.1332	8.9631	384.5048	0.0000	0.6621	0.5938
1997	–0.0021	0.0183	–0.0056	4.5923	27.5748	0.0000	0.7209	0.5561
1998	–0.0016	0.0344	0.9705	9.6303	519.0438	0.0000	0.7480	0.6516
1999	0.0038	0.0225	–0.1916	3.8740	9.9041	0.0071	0.8928	0.7084
2000	–0.0008	0.0267	–0.5736	4.7813	48.6281	0.0000	0.8268	0.6680
2001	–0.0007	0.0322	–0.9140	7.9937	307.5292	0.0000	0.8148	0.6571
2002	0.0022	0.0219	0.1980	4.2960	19.9713	0.0000	0.8301	0.6373
2003	–0.0001	0.0240	–0.4490	5.0711	55.4178	0.0000	0.7882	0.6296
2004	0.0005	0.0254	–1.0720	10.2588	625.3816	0.0000	0.7930	0.6414
2005	0.0019	0.0209	0.2733	3.4625	5.5548	0.0622	0.7050	0.6012
2006	0.0002	0.0175	-0.0450	3.3448	1.3754	0.5027	0.8591	0.6960
2007	0.0020	0.0160	-0.3116	3.7845	10.9186	0.0043	0.8369	0.6552
2008	–0.0034	0.0363	0.0344	5.3594	60.8219	0.0000	0.8789	0.7839
2009	0.0029	0.0288	–0.1477	4.6683	31.2186	0.0000	0.8633	0.7229
2010	0.0005	0.0166	–0.1466	3.4857	3.5011	0.1737	0.9128	0.7840
2011	0.0009	0.0161	–0.7014	6.7850	176.5170	0.0000	0.8704	0.6471
2012	–0.0003	0.0142	–0.1661	5.6555	77.8874	0.0000	0.8434	0.6433
2013	–0.0002	0.0095	0.1343	4.3501	20.6069	0.0000	0.7401	0.6017
2014	–0.0027	0.0133	–2.5738	21.1253	3860.8640	0.0000	0.4890	0.5623
2015	–0.0019	0.0255	0.1857	5.3834	63.2745	0.0000	0.7607	0.5786
2016	0.0020	0.0288	0.0718	4.7210	32.4351	0.0000	0.8239	0.6762
2017	0.0007	0.0141	–0.3138	3.8002	11.2053	0.0037	0.8020	0.6009
2018	–0.0009	0.0166	–0.7648	4.8221	61.5504	0.0000	0.7629	0.5512
2019	0.0009	0.0235	0.5493	11.4157	783.3391	0.0000	0.7696	0.5240
2020	–0.0007	0.0828	–1.7634	28.0016	6959.5588	0.0000	0.8119	0.7315
2021	0.0016	0.0198	–1.1149	9.1528	463.9867	0.0000	0.8233	0.7268

Own calculations based on data from Banco de México and the US Energy Information Administration (EIA)

Figure 3: Time series plots in price levels and returns of MXM (left) and WTI (right)



Worth mentioning is that the higher (lower) the RMSE, the better (worse) the forecast for oil revenues[19]. Figure 5 also shows that 2009 and 2015 are atypical years for all the proposed alternative estimations, including the SHCP's calculations[20]. Additionally, as of 2012, VaR-Normal seems to be a better estimation than the SHCP's.

19 SCHP's crude oil rice mix estimations are always underestimated for budgetary purposes.

20 In 2008-2009 we experienced the economic and financial crisis with epicenter in the US economy. In 2015, volatility and uncertainty on the normalization of the US monetary policy reflected the atony in the expansion of the global economic activity.

Figure 4: Conditional tail dependences: Clayton (left) and Gumbel (right) copulas

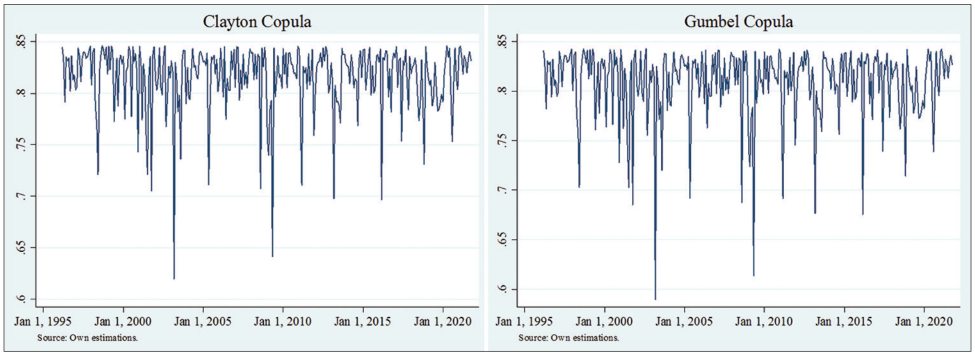


Figure 5: MAE and RMSE on *VaR* alternative estimations on MXM and WTI prices

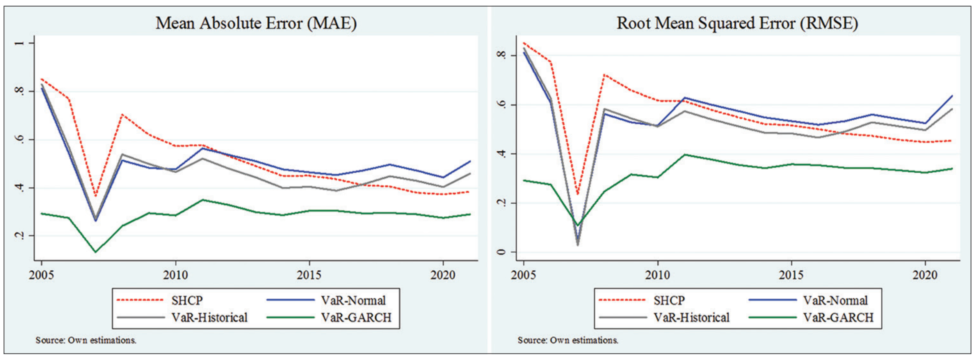


Table 2: Results of normal, historical *VaR* and *VaR*-TGARCH with *a*=0.1, *m*=87 (4 months) and 8-year window for the Mexican crude oil price (MXM)

Year	Average	Reference	VaR-δN	VaR_{HS}	VaR-TGARCH
2005	42.56	23.00	23.48	23.26	32.93
2006	53.24	31.50	41.65	40.61	71.78
2007	61.33	42.50	47.02	46.26	53.10
2008	85.44	46.60	51.42	50.62	67.71
2009	57.37	80.30	89.62	87.31	116.74
2010	72.11	53.90	49.88	55.66	94.06
2011	100.98	63.00	48.48	54.26	57.98
2012	102.11	84.90	75.94	86.27	86.58
2013	98.80	84.90	75.69	86.10	93.20
2014	87.55	81.00	74.94	86.61	105.75
2015	44.21	82.00	67.64	80.96	87.31
2016	35.87	50.00	27.12	29.47	27.56
2017	46.40	42.00	27.43	26.92	39.85
2018	61.78	46.00	33.86	32.86	46.77
2019	56.03	55.00	49.56	47.75	70.14
2020	35.87	49.00	36.27	35.26	38.25
2021	64.72	42.00	25.12	27.49	138.97

Own calculations based on data from Banco de México and thompson reuters

Table 3: Results of normal, historical *VaR* and *VaR*-TGARCH with *a*=0.1, *m*=87 (4 months) and 8-year window for the Mexican crude oil price (MXM)

Year	Average	Reference	VaR-δN	VaR_{HS}	VaR-TGARCH
2005	42.56	23.00	23.60	23.39	32.87
2006	53.24	31.50	42.25	41.27	70.47
2007	61.33	42.50	47.82	46.93	53.59
2008	85.44	46.60	51.43	50.54	65.67
2009	57.37	80.30	90.40	88.07	117.99
2010	72.11	53.90	50.95	56.39	93.41
2011	100.98	63.00	49.94	55.56	58.81
2012	102.11	84.90	77.20	86.92	86.66
2013	98.80	84.90	76.87	86.51	93.56
2014	87.55	81.00	75.26	86.04	103.98
2015	44.21	82.00	67.41	79.89	85.36
2016	35.87	50.00	27.85	30.26	27.88
2017	46.40	42.00	28.15	27.37	40.36
2018	61.78	46.00	34.48	33.46	47.37
2019	56.03	55.00	49.76	47.95	69.70
2020	35.87	49.00	36.48	35.44	37.95
2021	64.72	42.00	26.14	27.94	126.63

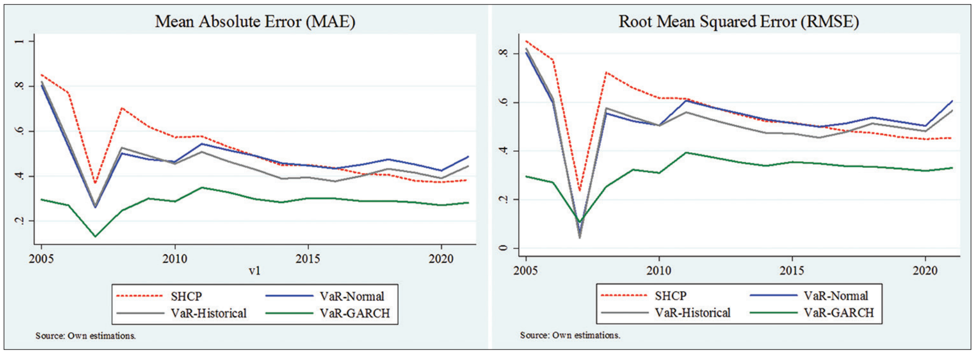
Own calculations based on data from Banco de México and thompson reuters

Table 3 shows nearly the same results considering an adjustment that includes the future prices of WTI -as a proxy for the MXM- but considering the price differential between the MXM and the WTI crude oil prices. We also consider the average between the four future prices. Yet again, the results are approximately identical.

Figure 6 shows the evolution of the observed annual average price of the MMX, SHCP's estimations, and our alternative proposed calculations (VaR Normal, VaR Historical, and VaR-GARCH) comparing their Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE).

As with the previous not-adjusted case, the higher (lower) the RMSE, the better (worse) the forecast is for oil revenues purposes. Figure 6 shows that 2009 and 2015 are atypical years for all the proposed alternative estimations, including the SHCP's calculations. Additionally, as of 2012, VaR Normal seems to be marginally a better estimation than the SHCP's. Remarkably, in the case of 2009, none of the VaR calculations presents better results.

Figure 6: MAE and RMSE on VaR alternative estimations on MXM and WTI prices



5. CONCLUSIONS

This paper presents empirical evidence of tail dependency between the Mexican crude oil price and the international reference West Texas Intermediate. We analyzed daily closing prices for the Mexican Export and WTI oil prices from January 3rd, 1996, to December 30th, 2021.

Empirical results suggest that: (1) each of the series of oil returns can be adequately described with the proposed AR(1)-TGARCH model; (2) a leverage effect exists in MMX and WTI oil returns: volatility increases when returns fall; (3) there are a linear dependence between the MMX and WTI (as indicated by Pearson correlation) returns, and a high degree of concordance (as shown by Kendall's tau); and (4) there is a high degree of conditional dependence in the lower tail and a significant degree of conditional dependence on the upper (right) tail.

This last point leads us to conclude that there is a strong and stable probability of an increase (decrease) in MMX returns following an increase (decrease) in WTI returns varying in time.

Regarding the calculation of the VaR, it is shown as a better option to determine the estimation of the annual Mexican oil blend price, which is an essential component of the Economic Package of each fiscal year. In all cases, the result for previous years is better using VaR.

In this paper, comparisons between using only the MXM or incorporating WTI futures considering a spread were similar because the dependency between MMX and the WTI crude oil prices is extremely high.

Finally, for possible future studies, it is necessary to extend this study to different commodity sectors examining further realized volatility and, consequently, some other copula models accordingly.

REFERENCES

Angelidis, T., Benos, A., Degiannakis, S.A. (2004), The use of GARCH models in VaR estimation. Statistical Methodology, 1(1), 105-128.

Bautista, R.S., Mora, J.A.N. (2020), Valor en riesgo en el sector petrolero: Un análisis de la eficiencia en la medición del riesgo de la distribución α-estable versus las distribuciones t-student generalizada asimétrica y normal. Contaduría y Administración, 65(2), 173.

Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

Boutouria, S., Abid, F. (2010), Modeling Copper Prices. Available from: https://www.ssrn1831309

Dafas, P.A. (2004), Estimating the Parameters of a Mean-reverting Markov-Switching Jump-diffusion Model for Crude Oil Spot Prices. School of Economics Mathematics and Statistics. London: University of London.

Engle, R.F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50(4), 987-1007.

Grajales, C., Pérez, F. (2010), Valor en riesgo para un portafolio con opciones financieras. Revista Ingeniarías Universidad de Medellín, 9(17), 105-118.

Hamilton, J.D. (2009), Understanding oil prices. The Energy Journal, 30(2), 179-206.

Kendall, M.G. (1938), A new measure of rank correlation. Biometrika, 30(1), 81-93.

Kilian, L. (2009), Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. American Economic Review, 99(3), 1053-1069.

Kilian, L., Rebucci, A., Spatafora, N. (2009), Oil shocks and external balances. Journal of International Economics, 77(2), 181-194.

Ley Federal de Presupuesto y Responsabilidad Hacendaria. (2021). Available from: https://www.diputados.gob.mx/LeyesBiblio/pdf/LFPRH.pdf [Last accessed on 2022 Nov 02].

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

Nelsen, R.B. (1999), An Introduction to Copulas, Lecture Notes in Statistics. New York: Springer New York.

New York Mercantile Exchange. Futures Prices. (2022), New York: New York Mercantile Exchange. Available from: https://www.eia.gov/dnav/pet/pet_pri_fut_s1_d.htm [Last accessed on 2022 Jul 02].

Patton, A.J. (2006), Modelling asymmetric exchange rate dependence. International Economic Review, 47(2), 527-556.

Precio Del Petróleo. Mezcla Mexicana, Dólares Por Barril, PMI. (2022), Newtown Square: Project Management Institute. Available from: https://www.banxico.org.mx/apps/gc/precios-spot-del-petroleo-gra.html [Last accessed on 2022 Jul 02].

Reglamento de la Ley Federal de Presupuesto y Responsabilidad Hacendaria. (2015). Available from: http://www.diputados.gob.mx/LeyesBiblio/regley/Reg_LFPRH_131120.pdf [Last accessed on 2022 Nov 02].

Sánchez, J.C.R., Aranda, F.C., Llanos, A.C. (2019), Reversión a la media en las series de precios reales del petróleo en México. Contaduría y Administración, 65(4), 202.

Schwartz, E.S. (1997), The stochastic behavior of commodity prices: Implications for valuation and hedging. The Journal of Finance 52(3), 923-973.

Sklar, A. (1959), Fonctions de Répartition à n Dimensions et Leurs Marges. Publications de l'Institut Statistique de l'Université de Paris, 8, 229-231

Wets, R.J.B., Rios, I. (2015), Modeling and estimating commodity prices: Copper prices. Mathematics and Financial Economics, 9, 247-270.

Zakoian, J.M. (1994), Threshold heteroskedastic models. Journal of Economic Dynamics and Control, 18(5), 931-955.

Zhang, Y.J., Zhang, L. (2015), Interpreting the crude oil price movements: Evidence from the Markov regime switching model. Applied Energy, 143, 96-109.