Journal of Kyzylorda Scholarly Review

P-ISSN:3006-0265 | E-ISSN:3007-4681. Published by the Open University of Kyzylorda

https://bulletin.ouk.kz/index.php/bulletin/index

Application of GME Tsallis with a Kink Regression Model for Intelligent Forecasting of Crude Oil Prices via the Nonlinear Relationship with Gold

Saad Obaid Jameel Al Masoodi
Bucharest Academy of Economic Studies, Bucharest, Romania
sjameel@uowasit.edu.iq

Abstract. This study stems from an important research need to understand the mechanisms through which safe-haven assets influence fluctuations in energy markets, particularly at a time when economic transformations and geopolitical changes are accelerating, This study aims to reveal the nature of the relationship between the price of crude oil (Brent) on the one hand and the price of an ounce of gold on the other, by employing a kink regression model with an unknown threshold, as developed by Hansen (2017). This is done using the generalized maximum entropy method based on the Tsallis scale, as it is a smart estimation measure; Tarkhamtham, Yamaka, and Sriboonchitta (2018) have demonstrated that it shows clear superiority when the tails are non-normal. A practical application was conducted on real monthly data covering the period from February 2015 to December 2025, i.e., 131 months These real-world observations were obtained from official sources: oil prices were obtained from the U.S. Energy Information Administration, and gold prices were obtained from the official websites of the World Bank. After conducting an in-depth analysis of the data, the results showed that there was a significant kink point when the gold price reached 2,254$ per ounce, and that the relationship between the gold and oil prices: it was a strong positive relationship (β_1^-=0.696) and then shifted (changed) to a negative relationship (β_1^+=−0.441) before and after the kink point. The Wald test value reached 16.04 at the 1% level according to the table established by Hansen (2017), and the mean absolute percentage error (MAPE) was 4.48% in the forecast within the sample, and that this finding confirms the usefulness of the proposed framework for identifying nonlinear structural relationships and for providing reliable forecast values for oil prices.

Keywords: GME Tsallis, kink regression model, Brent crude oil price, gold price, Wald test, in-sample prediction, nonlinear inflection points.

Introduction

Crude oil prices hold particular and pivotal importance in the global economic system, as their fluctuations have a significant impact on most productive sectors as well as financial sectors. Moreover, a precise explanation of these fluctuations is no longer limited to supply, demand, and traditional factors; rather, it is necessary and essential to examine a vast network of complex relationships with monetary and mechanisms. Foremost among these variables is gold, which plays the traditional role of money in times of uncertainty. Gold has had a long-standing relationship with oil throughout history; this relationship is well-documented and is viewed by investors through what is known as the gold-to-oil ratio. In this context, Drzazga-Szczęśniak et al. (2023) in this context is that financial time series are embedded in internal structures that can be uncovered using concepts from information theory, and that major events such as economic and financial crises and geopolitical conflicts have direct effects on the entropy of these series before their effects manifest in prices. Along the same lines, Pele and Mazurencu -Marinescu-Pele (2019) have explained that the use of entropy measures helps to accurately reveal the dynamics of financial markets and that they are significantly superior to what traditional models offer; this view supports the importance of resorting to statistical intelligence tools to understand oil price behavior. Traditional analysis of the relationship between oil and gold prices is often linked to an implicit assumption that the relationship is linear and remains stable across different price levels; however, this assumption rarely holds up in practice, as market behavior varies significantly when a certain price threshold is crossed. This means that when the price of gold rises above record levels often due to significant economic concerns it may affect oil demand differently than is observed during stable periods. This requires us to employ nonlinear regression models capable of handling and adapting to such structural changes [1], [2].

In this section, we highlight the importance of using the kink regression model introduced by Hansen (2017) as a more flexible analytical tool, as it combines economic realism with statistical rigor. The model allows the function to change its slope abruptly at an unknown point referred to as a kink while maintaining the continuity of the function at that point a feature that suits many economic phenomena in which sensitivity changes without the relationship level jumping. Hansen demonstrated that the resulting estimates possess significant convergence properties and introduced the Wald test to detect the significance of the kink point [3].

However, estimating the parameters of this model using traditional least squares methods may lose its efficiency when data are sparse or when errors involve a non-normal distribution. This is where the role of the Generalized Maximum Entropy method, developed by Golan, Judge, and Miller (1996), which was subsequently used in several higher-order measures, including the Tsallis measure. Tarkhamtham, Yamaka, and Sriboonchitta (2018) demonstrated that the GME Tsallis estimator clearly outperforms the least squares and maximum likelihood estimators in cases where the sample is small and the normal distribution assumption is not met. This was further reinforced by Tarkhamtham and Yamaka (2019) in a subsequent study, which focused on high-order entropy estimators [4], [5], [6].

Based on the foregoing, this study seeks to answer the following questions: First, does the relationship between the price of crude oil (Brent) and the price of gold contain a kink point that is statistically significant and economically meaningful? What is the nature of the change in this relationship before and after the kink point? How accurate is the GME Tsallis -Kink model in explaining oil price behavior within the sample under study? What are the key contributions of this study across three integrated dimensions? The first dimension is that it represents the first application of the kink model with the GME Tsallis estimator to the relationship between oil and gold, as previous studies were limited to applications in the fields of finance and economic growth. The second aspect involves the use of real-world data obtained from official sources, spanning a long period marked by major events such as the COVID-19 pandemic and the Russia-Ukraine war, The third aspect involves providing a predictive assessment within the sample that demonstrates the model's ability to capture changes in oil prices with high accuracy.

Methods

This study relies on two complementary pillars that integrate with one another, and this constitutes the analytical framework of the study. Among the pillars emphasized are statistical flexibility and inferential intelligence in the case of the kink regression model, specifically when the threshold value is unknown. One of the first researchers to introduce this model was Hansen in 2017. To fully understand this model, it is important to note that it belongs to a larger class of threshold models, whose basic concepts were clarified and established by Tong (1990) in the context of nonlinear autoregression. Furthermore, Chan and Tsay (1998) made a significant and important contribution by developing the continuous threshold model, which differs entirely from previously known threshold models (classical models) because it assumes that the function is continuous at the kink point. In light of this assumption, Hansen developed the mathematical model to allow for a break in the slope at an unknown point while maintaining the continuity condition. The general form of the model for a single independent variable with a kink point is as follows [7], [8]:

$$y_t = \beta_0 + \beta_1(x_t - r)^- + \beta_1(x_t - r)^+ + \varepsilon_t \qquad (1)$$

where β_0 represents the constant term, and β_1^- and β_1^+ represent the slopes of the function before and after the inflection point r, respectively. The negative and positive functions are defined as follows:

$$(x_t - r)^- = \min(x_t - r, 0), \qquad (x_t - \gamma)^+ = \max(x_t - r, 0) \qquad (2)$$

Hansen (2017) demonstrated that the estimate of the kink point converges to its true value at a rate of $n^{-(1/3)}$, an unusual rate that highlights the non-differentiable nature of the refractive index at the kink point. However, the estimates of the slope coefficients converge at the standard rate of $n^{-(1/2)}$, allowing for standard statistical inference. To test for the presence of a kink point, the Wald statistic is calculated using the following formula:

$$T_n = n \cdot \frac{\text{SSE}_R - \text{SSE}_U}{\text{SSE}_U} \qquad (3)$$

where (SSE_R) represents the sum of squares of the residuals under the linear model constrained by the null hypothesis of no refraction, and (SSE_U) represents the sum of squares of the residuals under the unconstrained refractive model. It is known that the asymptotic distribution of this statistic under the null hypothesis is not the usual chi-square distribution but rather a non-standard distribution based on a Gaussian process. Hansen (2017) provided critical value tables based on Monte Carlo simulations.

The second pillar is the general maximum entropy method according to the Tsallis scale, as this method derives from the principle established by Jaynes (1957) as an extension of the information theory founded by Shannon (1948). Golan, Judge, and Miller (1996) generalized this principle to serve as a standard tool in econometrics for addressing ill-defined problems and data limitations. Subsequently, Tsallis (1988) introduced an entropy measure that overcomes the limitations of Shannon's classical logarithmic formula. The Tsallis entropy measure is expressed by the following formula [3], [4], [9], [10]:

$$H_q(\mathrm{p}) = \frac{1}{q-1}\left[1 - \sum_{i=1}^{K} p_i^q\right] \qquad (4)$$

where (p) is the probability vector and (q) is the order parameter; as the value of (q) approaches one, the Tsallis metric transitions to the classical Shannon metric. Tarkhamtham, Yamaka, and Sriboonchitta (2018) applied this measure to estimate the kink regression model using Monte Carlo simulations and demonstrated that the estimate using the Tsallis GME clearly outperformed the estimates obtained via the least squares method and the maximum likelihood method when the errors were not normally distributed. A subsequent study by Tarkhamtham and Yamaka (2019) demonstrated that using higher orders of the coefficient (α) yielded higher accuracy according to the MAE criterion. Based on these results, the fourth-order coefficient (α=4) was adopted in this study [5].

Data and the Empirical Model

This study is based on actual monthly data spanning a period of 131 months, beginning in February 2015 and ending in December 2025. This period was selected because it includes significant economic and geopolitical events, including the 2016 oil price crash, followed by the COVID-19 pandemic (2020–2021), during which the price of a barrel of oil reached a historic low of 18.38$ in April 2020, and the outbreak of the Russian-Ukrainian war (2022), which drove the price to a peak of 122. 71$ per barrel in June of the same year. On the other hand, this period coincided with a historic surge in the value of gold, as the price per ounce rose from 1,075$ in December 2015 to approximately 4,309$ in December 2025.

Crude oil (Brent) price data was obtained from the U.S. Energy Information Administration (EIA) (2025), and gold price data was obtained from World Bank commodity reports (2025). The dependent variable in the model represents the logarithm of the price of a barrel of crude oil in U.S. dollars, and the independent variable represents the logarithm of the price of an ounce of gold in U.S. dollars. Logarithmic transformations were adopted to mitigate large fluctuations in the time series and to facilitate the interpretation of the coefficients as elasticities, and the applied model equation takes the following form:

$$\ln(\mathrm{Oil}_t) = \beta_0 + \beta_1(\ln(\mathrm{Gold}_t) - r)^- + \beta_1(\ln(\mathrm{Gold}_t) - r)^+ + \varepsilon_t \qquad (5)$$

Descriptive statistics for the data revealed two distinct patterns for the variables: the average price of oil per barrel during the study period was 66. 59$, with a standard deviation of 18.60$, indicating significant volatility. In contrast, the average price of gold per ounce was 1,774$, with a standard deviation of 666$, reflecting a strong upward trend during the study period. and the value of the linear correlation coefficient between the two logarithmic series reached 0.369, which is a moderate value indicating a relationship not fully captured by linear models; this clearly confirms the actual need to adopt a nonlinear model capable of detecting latent inflection points, As for the convergence property of the inflection point estimator, it takes the following form, as demonstrated by Hansen (2017):

$$n^{1/3}(\hat{r} - r_0) \xrightarrow{d} \xi \qquad (6)$$

The augmented Dickey-Fuller test was applied to examine the stationarity of the series, and it was found that both series are first-order integrated (I(1)), a common property in financial time series, as noted by Drzazga-Szczęśniak et al. (2023). This does not pose an obstacle to applying the kink model at the logarithmic level, as the model aims to detect a structural relationship between levels rather than a short-term relationship between differences [2].

Results and Discussion

The application of the proposed methodological framework to the study data revealed highly significant results that combine statistical significance with economic relevance, as the Wald statistic ($T_n = 16.04$) exceeded the critical value at the 1% significance level (9.21 according to Hansen, 2017) by a clear and significant margin. This indicates the presence of a fundamental breakpoint in the relationship between oil and gold. Our confidence in this result is supported by the findings of Tarkhamtham and Yamaka (2019) that the GME Tsallis estimator retains its efficiency in detecting true breakpoints even in the absence of a normal distribution assumption.

As for the inflection point itself, the research yielded a logarithmic value of 7.7206, corresponding to a gold price of 2,254$ per ounce. This value carries profound economic significance, as it coincides with the phase in which gold surpassed its record levels in 2024, thus heralding the beginning of a new economic era characterized by inflationary fears and escalating geopolitical issues. It should be noted that this point marks a turning point between two systems, each of which exhibits completely different behavior regarding gold prices, and this is clearly reflected in the nature of the relationship with oil prices, as shown in the following table:

Table 1. Parameters of the kink regression model estimated using the Tsallis GME method (α=4)

Parameter	Estimate	Standard error	t-value
β_0	4.4129	0.0502	87.871***
β_1^-(Gold)	0.6956	0.1143	6.085***
β_1^+(Gold)	−0.4407	0.2119	−2.080**

$\hat{r}$ = $2254.29/oz,*   *$R^2$ = 0.2304,   *Wald T_n = 16.039,*   *n = 131*

*** Significant at the 1% level, ** Significant at the 5% level

The results in Table (1) reveal a fundamental discovery regarding the nature of the relationship between gold and oil, namely that this relationship reverses direction when the price crosses the estimated kink point . In the region before the kink point, when the price of gold is below 2,254$ per ounce, a strong positive relationship emerges between the two variables with an elasticity of 0.696 and significance at the 1% level . This behavior is consistent with the standard economic interpretation, whereby the two variables move together under normal conditions under the influence of common monetary factors, including inflation and the strength of the dollar, which is in line with the view of Drzazga-Szczęśniak et al. (2023), who confirmed that financial assets share a response to general shocks in conditions of relative stability.

What is notable is the region beyond the kink point, where gold exceeds the 2,254$ threshold; that is, here the relationship reverses to a negative one with an elasticity of (0.441-) and a significance level of 5%. This reversal requires a gradual economic explanation, as the rise in gold prices after the kink point does not occur in a vacuum but is typically accompanied by a flight to safe-haven assets a trend that, in turn, arises only amid expectations of an imminent economic slowdown. and as this expected slowdown is projected to affect actual oil demand, prices begin to decline. This explanation aligns with the observations made by Pele and Mazurencu-Marinescu-Pele (2019) that high asset price regimes often carry informational signals that precede actual developments in other markets.

In-sample forecasting

After estimating the parameters of the refractive model (which contains a kink) using the GME Tsallis method, the study proceeded to evaluate its predictive ability within the sample under study. This evaluation focuses on measuring the model's accuracy in generating estimated oil prices based on observed gold prices over the entire period from February 2015 to December 2025, as this type of evaluation is of great importance in the statistical literature because it reveals the model's consistency with the data as well as its ability to explain the underlying dynamics, as noted by Tarkhamtham, Yamaka, and Sriboonchitta (2018) in their evaluation of the performance of Generalized Maximum Entropy (GME) estimators.

For the evaluation, three integrated metrics were used, the statistical context of which is based on their application in the financial forecasting literature. The first of these metrics is the Mean Absolute Percentage Error (MAPE), the second is the root mean square error (RMSE), and the third is the Theil U statistic, which takes a value close to zero when the forecast is accurate. The MAPE formula is as follows:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| * 100\% \qquad (7)$$

The Theil U statistic is calculated using the following formula:

$$U = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n} y_t^2} + \sqrt{\frac{1}{n}\sum_{t=1}^{n} \hat{y}_t^2}} \qquad (8)$$

The application of these metrics yields results that indicate a high degree of accuracy for the proposed model, as shown in the following table:

Journal of Kyzylorda Scholarly Review
P-ISSN:3006-0265 | E-ISSN:3007-4681
https://bulletin.ouk.kz/index.php/bulletin/index.php

Table 2. In-sample prediction accuracy metrics for the GME Tsallis -Kink model

Measure of accuracy	Value	Interpretation according to Lewis (1982)
MAPE	4.48%	Excellent prognosis (<10%)
RMSE (lograthmic)	0.2643	Low
Theil U	0.0317	Near zero (good match)
R^2	0.2304	A plausible explanation for the discrepancy

The results presented in Table (2) reveal that the proposed model achieves excellent predictive accuracy according to Lewis's (1982) classification, as the absolute error did not exceed 4.48% This value indicates the model's ability to effectively capture oil price dynamics, especially considering the extreme volatility in oil prices during the study period, which was marked by major geopolitical and economic events. Furthermore, the Theil U statistic, which reached 0. 032 confirms the quality of the fit between the predicted and actual values, as values close to zero in this statistic indicate excellent predictive performance, as explained by Lewis (1982) in his establishment of the fundamentals of financial forecasting.

Figure (1) visually illustrates the model's performance and presents a comparison between actual and estimated oil prices during the study period, highlighting the periods in which major events occurred, such as the COVID-19 and the Russian-Ukrainian war. The figure demonstrates the model's ability to track the general trend of prices, with some deviations observed during exceptional periods a behavior that is expected given the sudden nature of these shocks.

Figure 1. In-sample forecasting performance of the GME Tsallis -Kink model, showing the time series of actual and estimated values

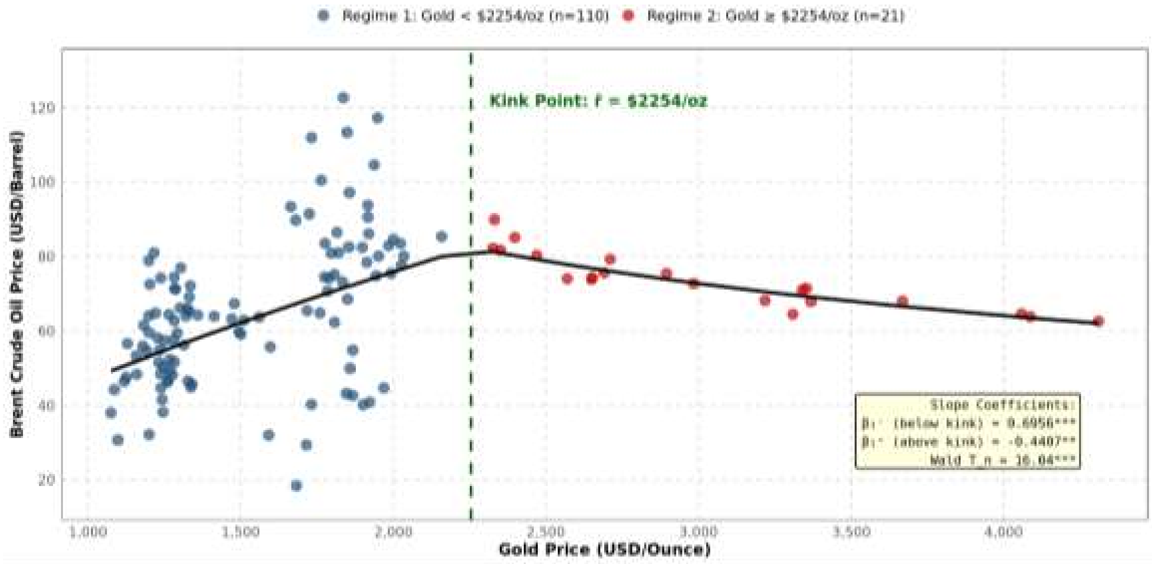

Figure 2. illustrates the nonlinear relationship between oil and gold, with the inflection point estimated using the GME Tsallis -Kink method

Although the model demonstrates excellent predictive performance within the sample, these results are based on an in-sample analysis. Therefore, generalizing them to future periods requires the researcher to periodically re-estimate the parameters, given the dynamic nature of the stock market and the possibility of structural breaks, such as those observed by Drzazga-Szczęśniak et al. (2023) during major events. In this context, the results of this model support the argument put forward by Tak and Pele (2025) that intelligent forecasting frameworks require continuous updating of their data in order to maintain their effectiveness in volatile financial environments [11], [12], [13], [14], [15].

Conclusion

This study aimed to provide an intelligent analytical framework for examining the relationship between crude oil prices (Brent) and gold prices, utilizing statistical intelligence techniques to estimate the parameters of a kink regression model using the Generalized Maximum Entropy method of high order according to the Tsallis scale. This analytical approach revealed a statistically significant kink point when gold prices (at 2,254 $ per ounce), at which point the relationship shifts from a strong positive correlation to a negative one. This mathematical model has demonstrated high predictive capability within the sample, with a reliability of 4.48% according to the MAPE criterion, and that this value gives it the necessary importance and classifies it as one of the models recommended for forecasting, as described by Lewis (1982).
These results carry important practical implications for decision-makers and investors in the energy sector, as they encourage a re-examination of the common implicit

assumption that the relationship between oil and gold remains stable across different price cycles. However, gold's crossing of the 2,254$ threshold should be regarded as an important and potential warning sign of an imminent decline in demand for oil, especially if this coincides with other indicators of an economic slowdown. Furthermore, this result provides empirical evidence for the argument put forward by Tak and Pele (2025) that modern financial time series contain informational signals that can be effectively utilized in forecasting indicators of other markets.

This study also opens the door to subsequent research, including: First, expanding the framework to incorporate additional variables such as digital assets, including Bitcoin, which has begun to play a growing role in the global asset ecosystem, as demonstrated by Pele and Mazurencu-Marinescu-Pele (2019); second, exploring the dynamics of this relationship during specific geopolitical shocks, following the approach taken by Drzazga-Szczęśniak et al. (2023) in their study of the impact of the Russian-Ukrainian war on financial markets; and third: integrating the proposed framework with deep learning techniques and paying due attention to the efforts of Han et al. (2024) to develop hybrid models combining traditional methods with modern artificial intelligence, and that the fundamental contribution of this study remains the establishment of a methodological bridge between the literature on entropy and the literature on nonlinear regression in the field of energy markets, as this is the area we hope will be expanded upon in future studies.

References

[1] D. T. Pele and M. Mazurencu-Marinescu-Pele, "Using high-frequency entropy to forecast Bitcoin's daily value at risk," *Entropy*, vol. 21, no. 2, p. 102, 2019, doi: 10.3390/e21020102.

[2] E. A. Drzazga-Szczęśniak, P. Szczepanik, A. Z. Kaczmarek, and D. Szczęśniak, "Entropy of financial time series due to the shock of war," *Entropy*, vol. 25, no. 5, p. 823, 2023, doi: 10.3390/e25050823.

[3] B. E. Hansen, "Regression kink with an unknown threshold," *Journal of Business & Economic Statistics*, vol. 35, no. 2, pp. 228–240, 2017, doi: 10.1080/07350015.2015.1073595.

[4] A. Golan, G. Judge, and D. Miller, *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. Chichester: John Wiley & Sons, 1996.

[5] P. Tarkhamtham, W. Yamaka, and S. Sriboonchitta, "The generalized maximum Tsallis entropy estimator in kink regression model," *J. Phys. Conf. Ser.*, vol. 1053, p. 012103, 2018, doi: 10.1088/1742-6596/1053/1/012103.

[6] P. Tarkhamtham and W. Yamaka, "High-order generalized maximum entropy estimator in kink regression model," *Thai Journal of Mathematics*, pp. 185–200, 2019.

[7] H. Tong, *Non-linear Time Series: A Dynamical System Approach*. Oxford: Oxford University Press, 1990.

[8] K. S. Chan and R. S. Tsay, "Limiting properties of the least squares estimator of a continuous threshold autoregressive model," *Biometrika*, vol. 85, no. 2, pp. 413–426, 1998.

[9] E. T. Jaynes, "Information theory and statistical mechanics," *Physical Review*, vol. 106, no. 4, pp. 620–630, 1957.

[10] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical*

Journal, vol. 27, no. 3, pp. 379–423, 1948.
[11] C. Tsallis, "Possible generalization of Boltzmann-Gibbs statistics," *J. Stat. Phys.*, vol. 52, no. 1–2, pp. 479–487, 1988.
[12] W. Bank, "Commodity Markets Outlook: Pink Sheet Data," World Bank Group, Washington, DC, 2025. [Online]. Available: https://www.worldbank.org/en/research/commodity-markets
[13] C. D. Lewis, *Industrial and Business Forecasting Methods*. London: Butterworth Scientific, 1982.
[14] H. Han *et al.*, "A hybrid time series forecasting method based on GARCH and CEEMDAN-GCN model," *Journal of Cloud Computing*, vol. 13, no. 1, p. 2, 2024.
[15] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, vol. 50, no. 4, pp. 987–1007, 1982.