Financial Cross-hedging

Managing OpEx Risk Through Hedging Strategies in Commodity Markets

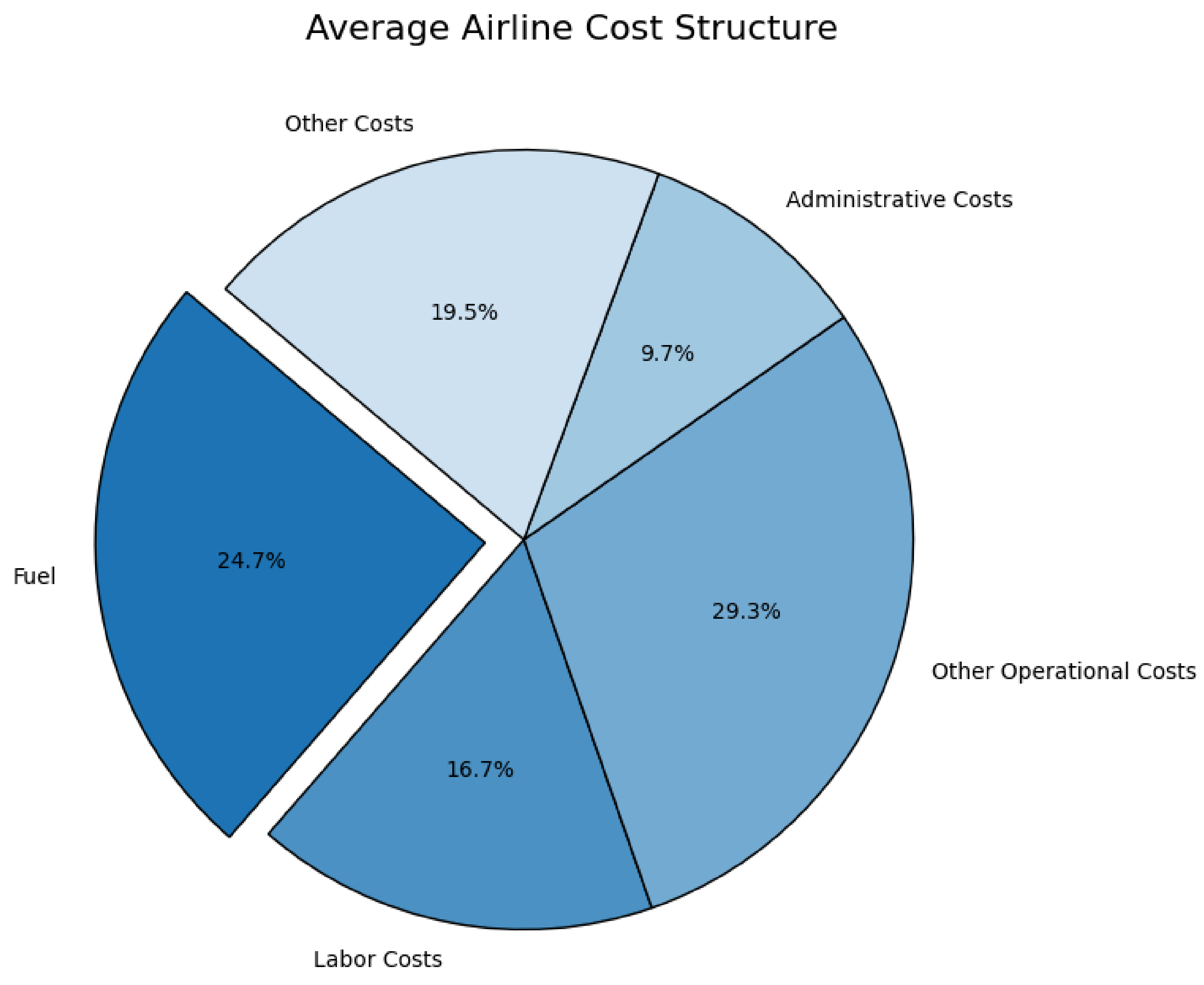
Riccardo Perego
Giovanni Piva
Emilio Zheng

Overview

- Many companies rely on specific assets for their operations, such as airlines depending on jet fuel, and fluctuations in the prices of these assets directly impact their operational expenses and overall profitability.
- Due to the volatility of these asset prices, companies may face significant financial uncertainty, leading to unpredictable costs and fluctuations in their earnings.
- To mitigate this risk, some companies engage in hedging by using financial instruments like futures contracts, allowing them to lock in asset prices; while this does not change the expected cost, it significantly reduces variance and stabilizes financial outcomes.

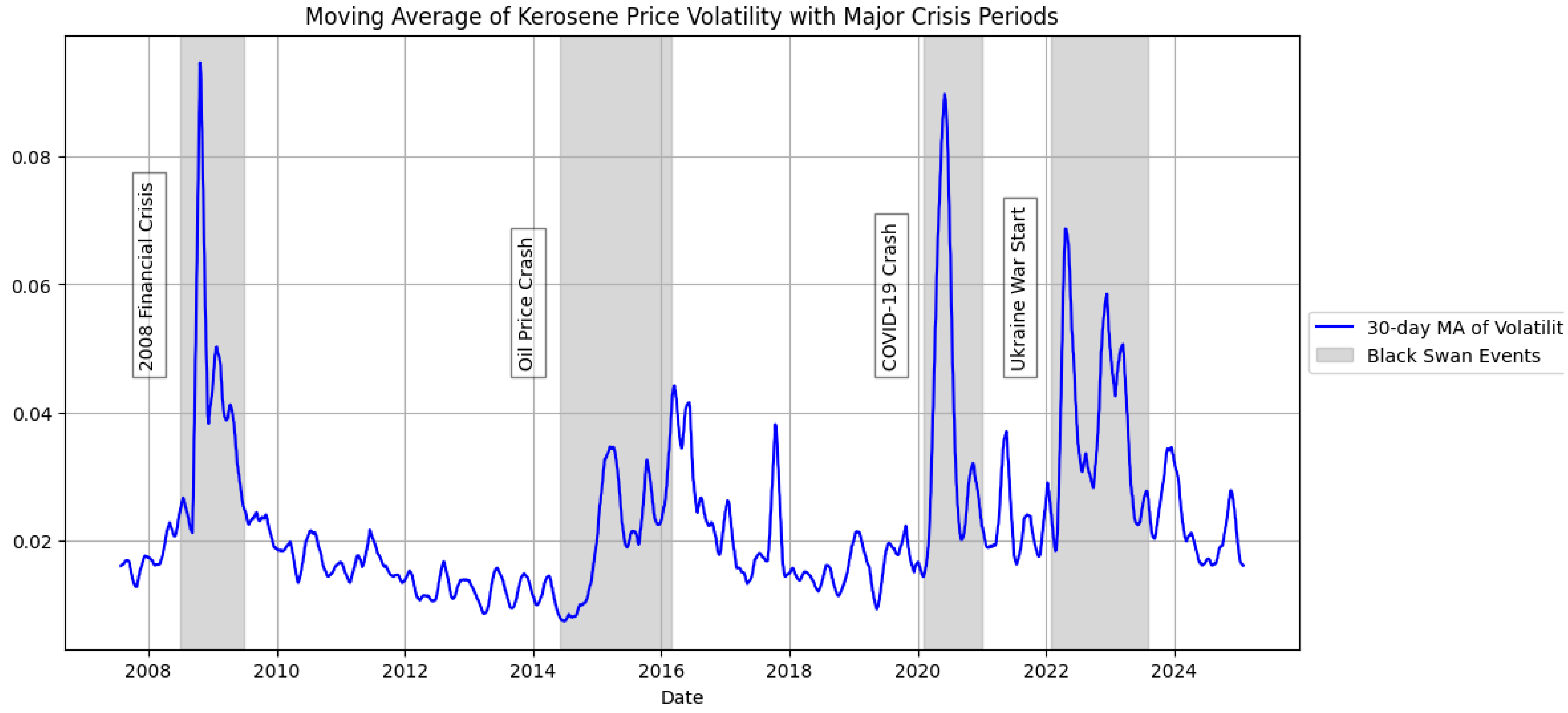
The Airline Industry

- The airline industry sees fuel (kerosene) as a key cost driver, making up about 24.7% of total operating expenses on average.



The Airline Industry

- This is a key concern as fuel prices are highly volatile, posing a significant threat to the stability of company profits.



Problem

- This situation highlights the need for the previously mentioned risk management strategy: hedging, where airlines can use financial instruments to lock in kerosene prices and reduce exposure to market fluctuations.
- But, there is a problem...
- In financial markets, there is no direct futures contract for kerosene, making it impossible for airlines to hedge their positions in a conventional way...

Solution

- When no direct futures contract exists for an asset, companies can hedge using a correlated asset, a strategy known as **cross-hedging**.
- Kerosene is a derivative of crude oil, meaning its price closely follows oil price movements, making oil futures a potential hedging instrument.
- Since crude oil futures are widely traded, we will be showing how airlines can use them as a proxy hedge to reduce exposure to kerosene price volatility

Cross-hedging

- Cross-hedging involves a time series regression analysis to determine the relationship between the price of the asset being hedged (kerosene) and the price of a correlated financial instrument (such as oil futures).
- There are multiple oil-related financial products available in the market, including crude oil futures, heating oil futures, and gas oil futures, each with varying degrees of correlation to kerosene prices.
- To optimize the hedging strategy, we will analyze these different financial products and select the one that provides the strongest and most stable correlation with kerosene prices.

Project Phases

- Data Collection & Merging
- Exploratory Data Analysis (EDA)
- Correlation Analysis and choice of Independent Variable
- First Regression Model (Time Series based)
- PCA for the second Model
- Second Regression Model: multi-asset approach
- Model Comparison
- Hedge Ratio calculation

Data Sources

Table 1
Data sources

	Trading exchanges	Data sources
U.S. Gulf Coast Kerosene - type Jet Fuel Spot Price FOB	-	U.S. Energy Information Administration
Brent Crude Oil	NYMEX	Yahoo Finance
West Texas Intermediate Crude Oil	NYMEX	Yahoo Finance
Heating Oil	NYMEX	Yahoo Finance
RBOB Gasoline	NYMEX	Yahoo Finance

Data Processing Steps

1 - Download all separate data sources.

2 - Join the datasets by date

3 - Remove non-useful columns, NaNs, and negative numbers

4 - Calculate log returns

Note on Log Returns

- We analyze returns instead of prices because prices are highly correlated over time, while daily returns show less dependency.
- Using returns provides a clearer and more reliable measure of price movements for statistical modeling.
- Log returns are preferred in finance as they are time-additive and exhibit more stable variance, improving analysis accuracy.

Normal Returns

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}}$$

Log Returns

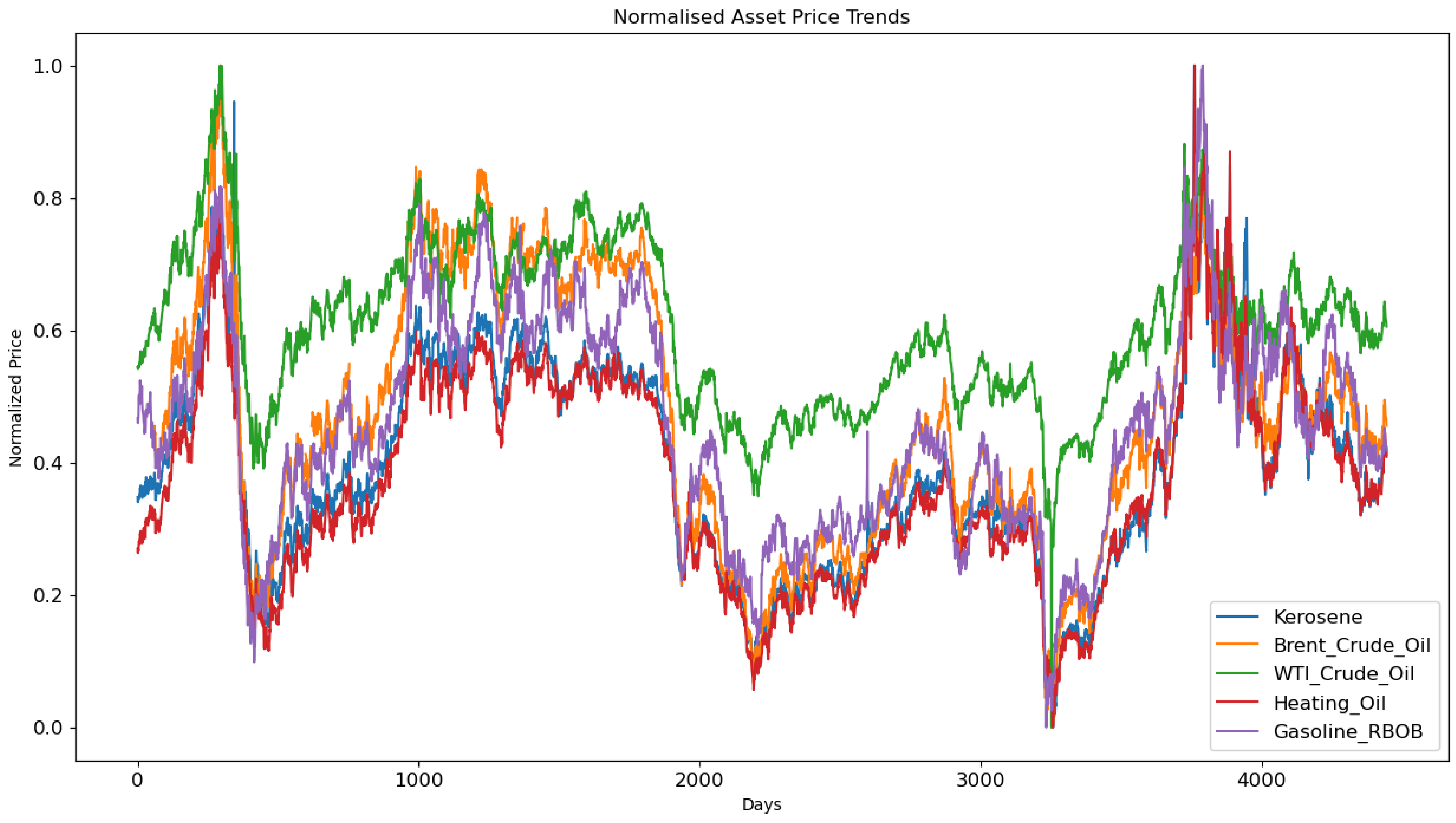
$$z_t = \log\left(\frac{p_t}{p_{t-1}}\right)$$

Summary Statistics

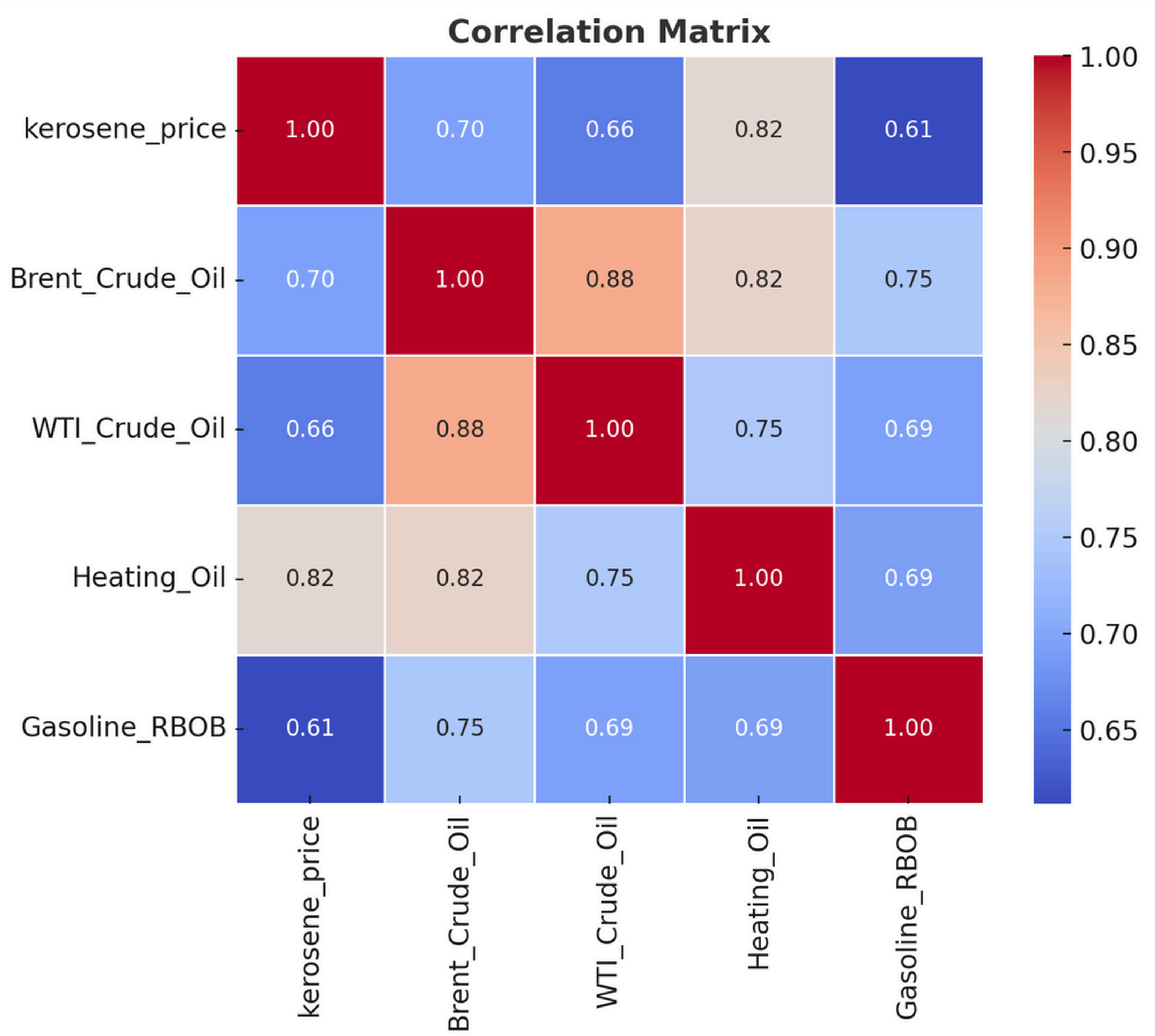
	Kerosene	Brent_Crude_Oil	WTI_Crude_Oil	Heating_Oil	Gasoline_RBOB
count	4290.0	4290.0	4290.0	4290.0	4290.0
mean	0.0001	0.0	0.0001	0.0001	0.0
std	0.0267	0.0233	0.027	0.0225	0.0266
min	-0.2775	-0.2758	-0.2822	-0.2475	-0.3854
25%	-0.012	-0.0103	-0.0124	-0.0108	-0.012
50%	0.0	0.0006	0.0008	0.0008	0.001
75%	0.0124	0.0114	0.0128	0.0116	0.0135
max	0.3264	0.1908	0.3196	0.1236	0.224

Data Exploration

Data Exploration

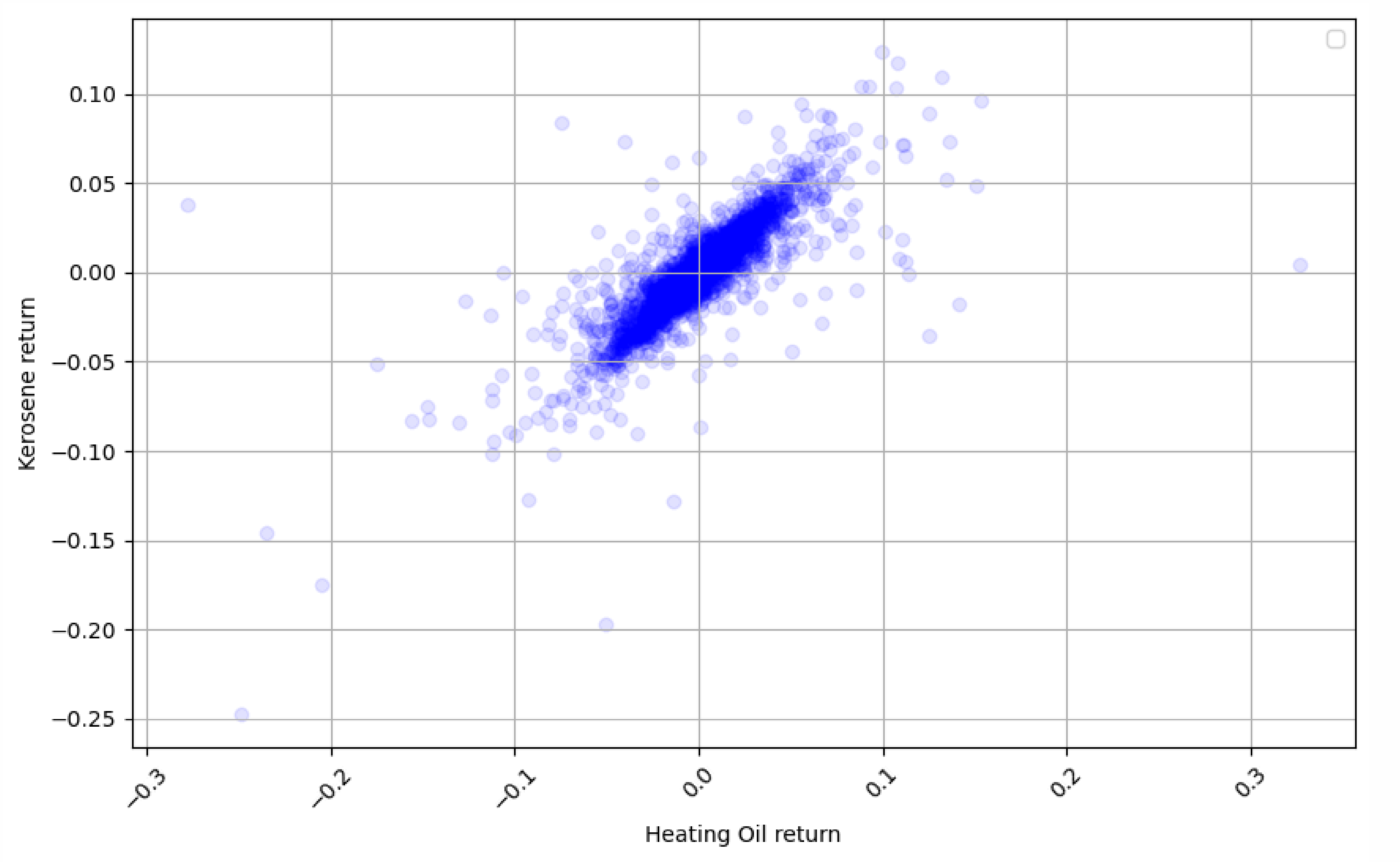


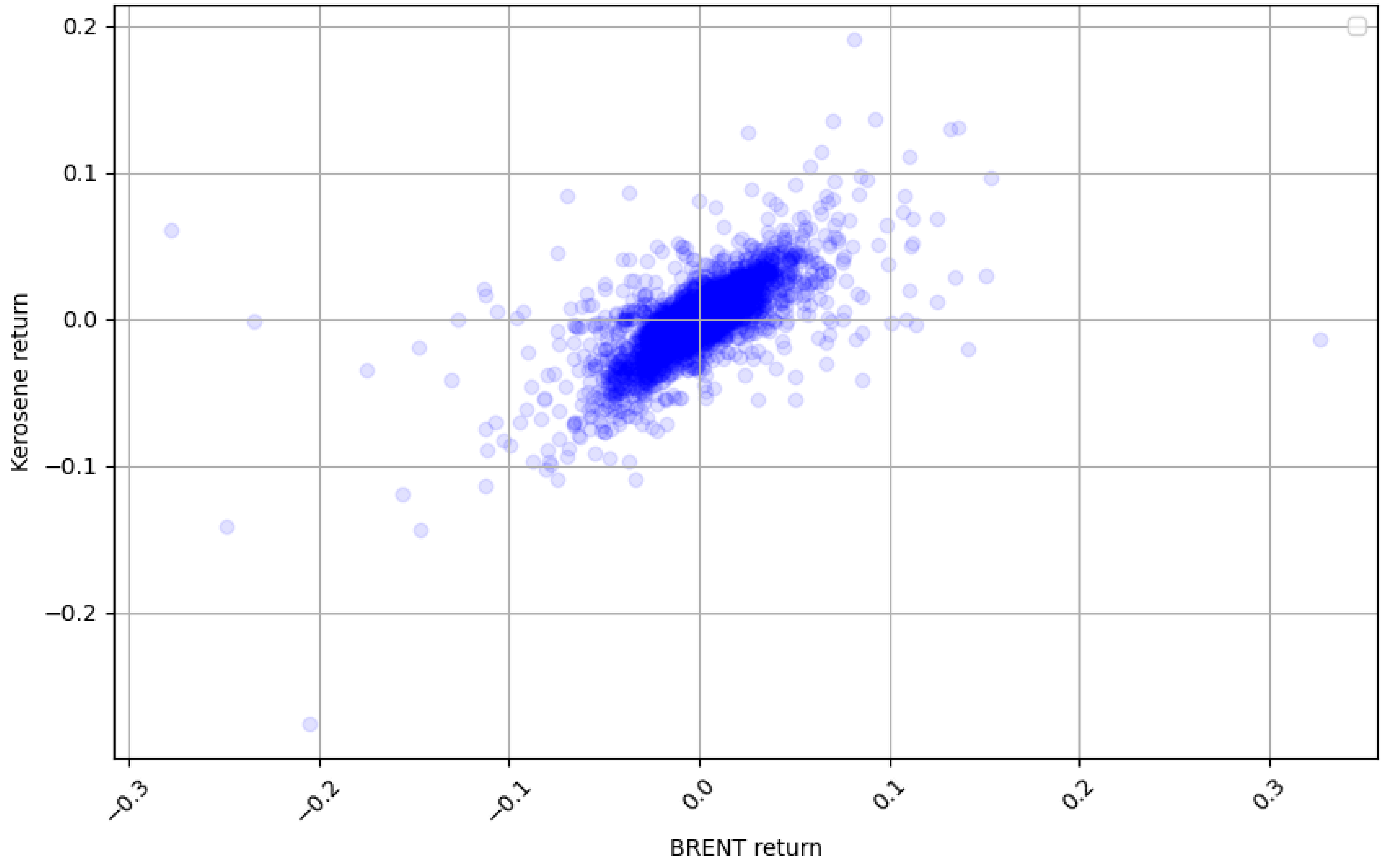
Data Exploration - Correlation Analysis



- As an independent variable to be used in the first model, we choose the oil most correlated with jet fuel.
- Heating Oil is the commodity most related to Kerosene (**0.817**).

Data Exploration - Correlation Analysis





Time Series Analysis

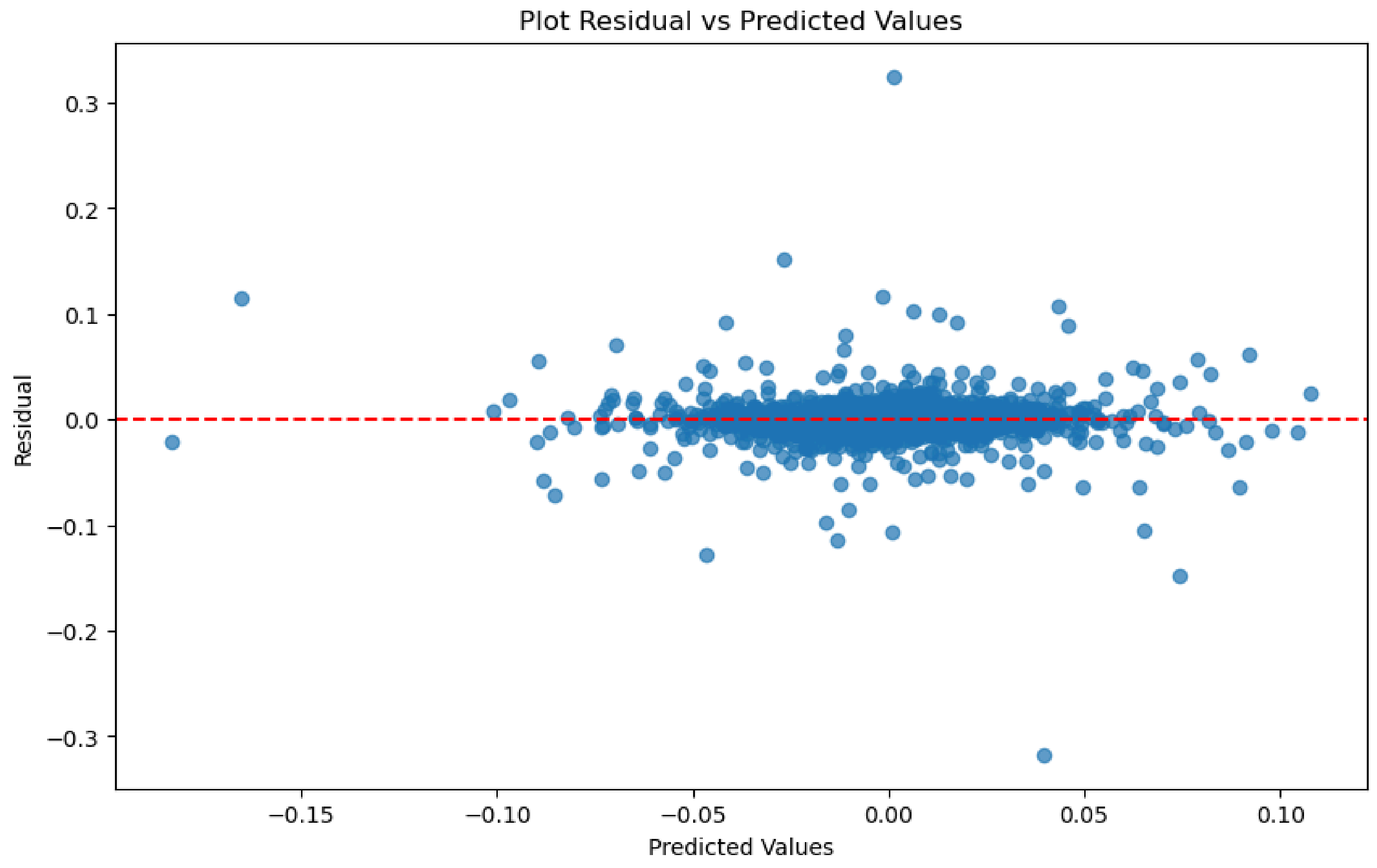
- A time series is a sequence of data observed at regular time intervals (in our case, daily kerosene and oil prices).
- Each observation **depends** on past values.
  Therefore, the order of the data is crucial, so we cannot use the 'classical' methods to randomly split the dataset.
- Variations are random and unpredictable (*stochasticity*).
- But when can we use OLS in Time Series?

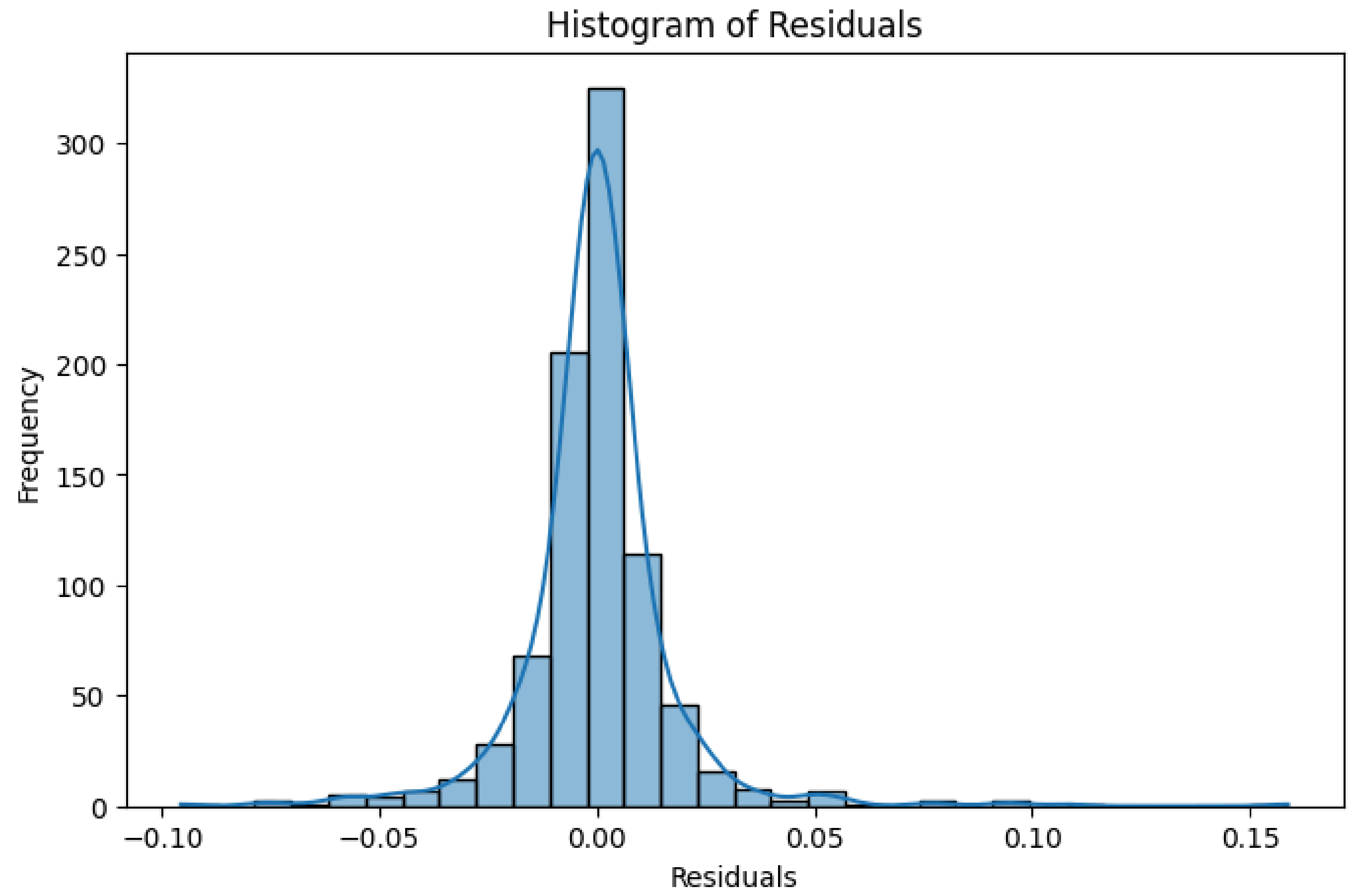
Assumptions

1. Linear In parameters
2. No perfect collinearity
3. Strict exogeneity $E[\epsilon_t|\mathbf{X}] = 0$
4. Homoskedasticity $\mathrm{Var}(\epsilon_t|\mathbf{X}) = \sigma^2$
5. No serial correlation $\mathrm{corr}(\epsilon_t, \epsilon_s|\mathbf{X}) = 0 \text{ for all } t \neq s$
6. Normality and independence of the error terms

Checking Assumptions

- *Breush-Pagan* test:
  H0: Homoskedasticity
  H1: Heteroskedasticity
  p-value: **0.73** (> 0.05) (fail to reject)
- *Durbin-Watson* test statistic: **2.28**
  - Value close to 2 -> no evidence of serial correlation in the residuals.
- *Shapiro test:*
  H0: Sample normally distributed
  H1: Sample not normally distributed
  p-value: 0.00 (< 0.05) (reject)
  Not too concerning due to **CLT**





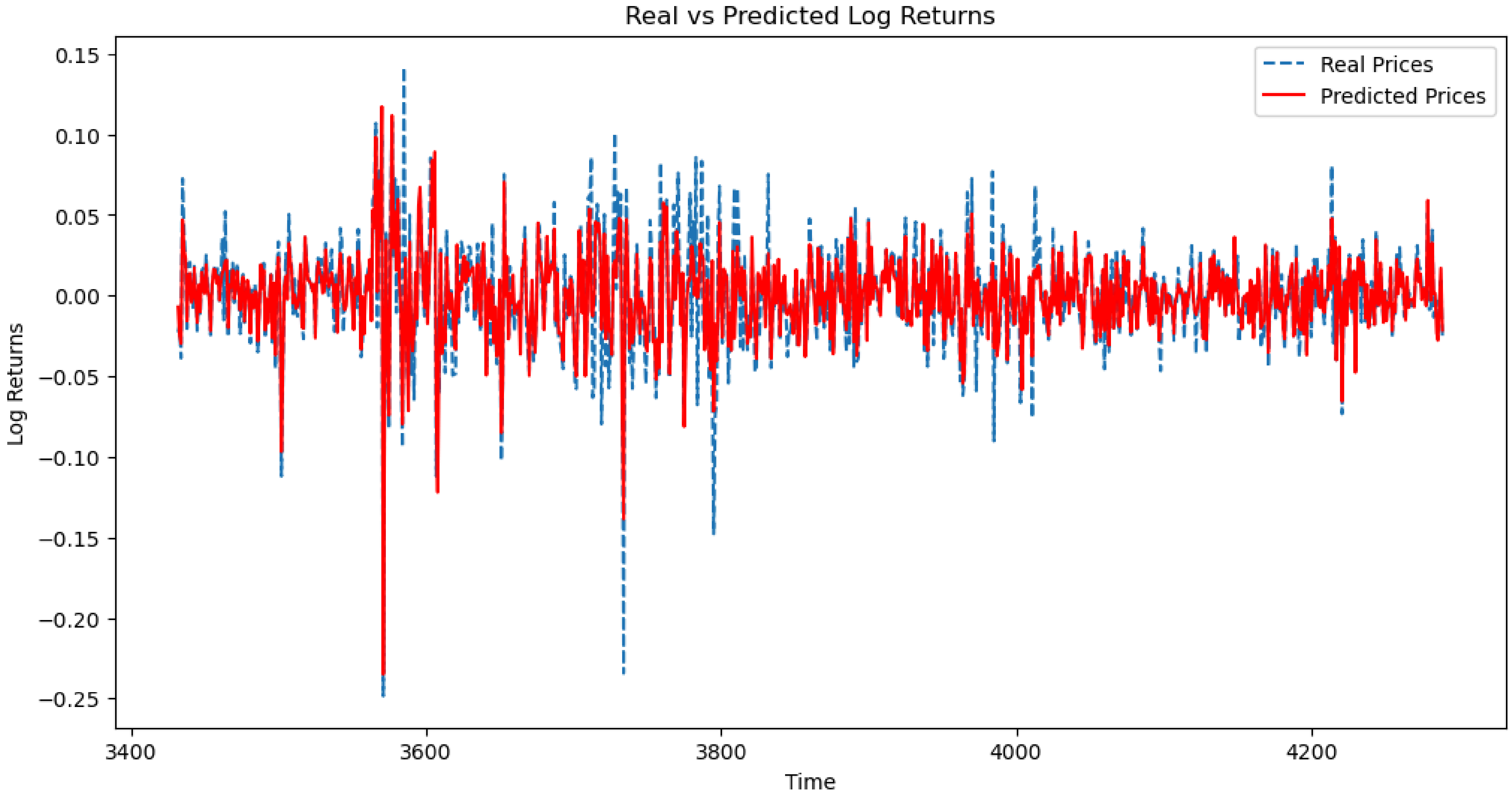
First Model Summary

$$R_K = \beta_0 + \beta_1 \cdot R_{HO} + \epsilon$$

Table 1: First Regression (OLS)

Model	OLS - Least Squares	
Dependent Variable	kerosene	
No. Observations	3432	
R-squared	0.644	
Adjusted R-squared	0.644	
F-statistic	6196	
Prob(F-statistic)	0.000	
AIC	-19120	
BIC	-19110	
Durbin-Watson	2.277	
Condition Number	47.4	
Variables	**Coef**	**Std Err**
Constant	-5.907e-06	0.000
Heating_Oil	0.9492	0.012
Statistics	**t-value**	**P t**
Constant	-0.023	0.981
Heating_Oil	78.712	0.000

First Model Summary



PCA for the Second Model

Heating Oil + BRENT Crude Oil

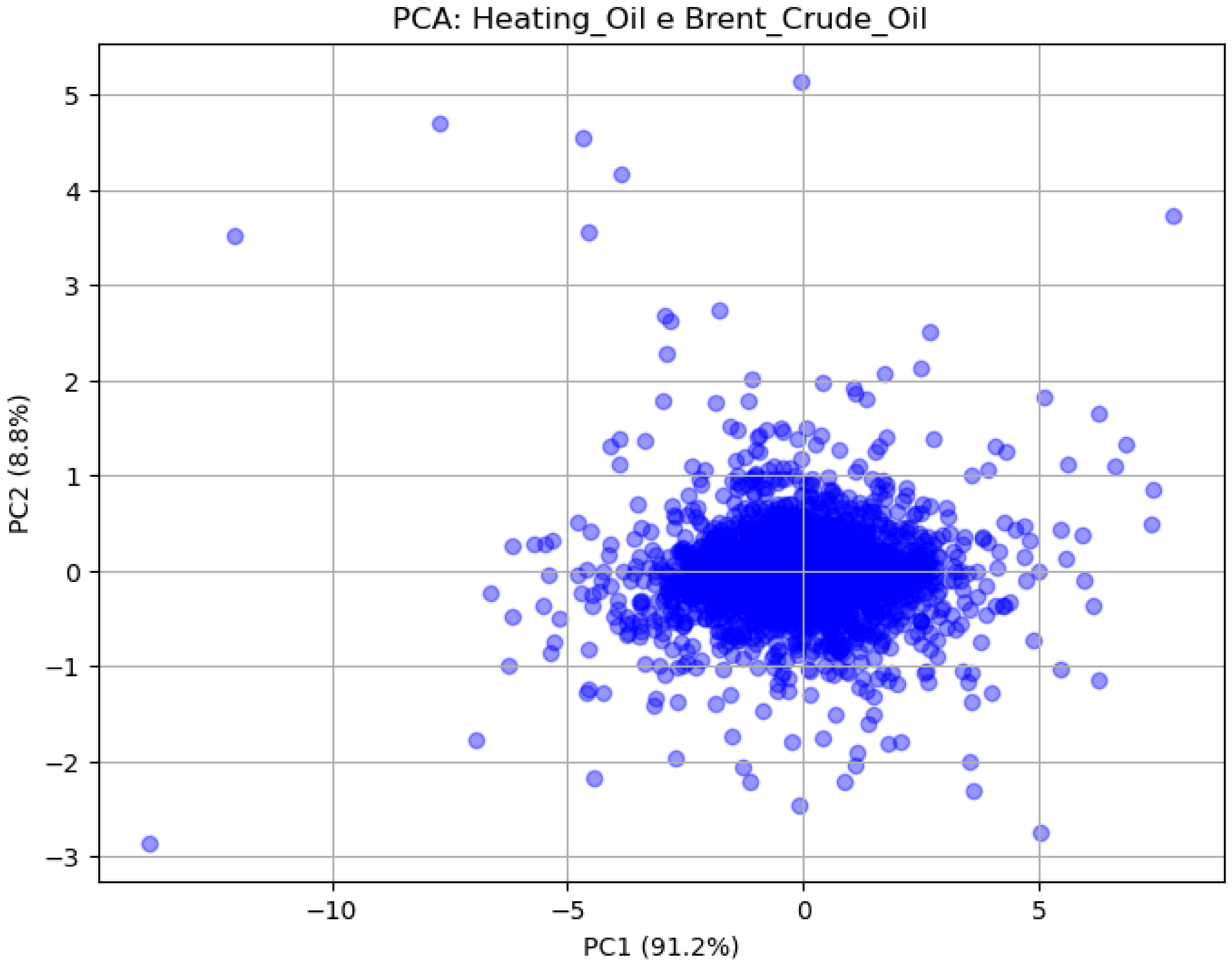
- The resulting graph is without obvious correlation, indicating that the two new components are orthogonal.

- 91.2% of the variance is explained by the first PC. This means that PC1 is the direction along which the data varies the most, and it is the one that summarizes almost all the underlying structure of the data.

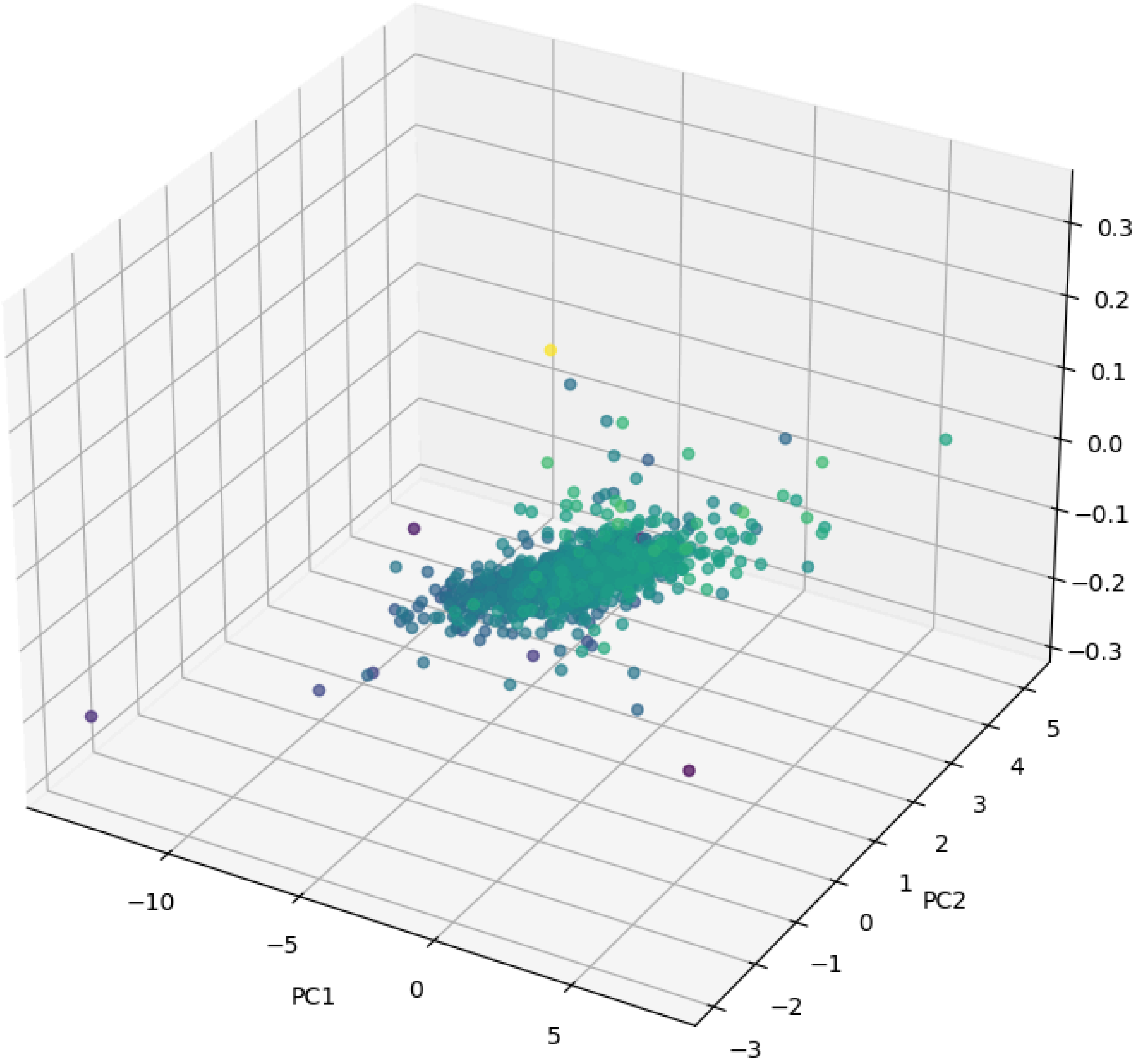
PCA for the Second Model

$$\mathrm{PC1} = \phi_{11} \cdot R_{HO} + \phi_{12} \cdot R_{BR}$$

$$\mathrm{PC2} = \phi_{21} \cdot R_{HO} + \phi_{22} \cdot R_{BR}$$



PCA for the Second Model



Second Model summary

$$R_K = \beta_1 \cdot PC1 + \beta_2 \cdot PC2 + \epsilon$$

Table 1: Second Regression OLS (uncentred)

Model	OLS - Least Squares (Uncentered)	
Dependent Variable	kerosene	
No. Observations	3432	
R-squared	0.668	
Adjusted R-squared	0.668	
F-statistic	3173	
Prob(F-statistic)	0.000	
AIC	-19180	
BIC	-19170	
Log-Likelihood	9590.9	
Df Model	2	
Df Residuals	3430	
Durbin-Watson	2.262	
Condition Number	3.74	
Variables	**Coef**	**Std Err**
PC1	0.0153	0.000
PC2	-0.0101	0.001
Statistics	**t-value**	**P**
PC1	79.596	0.000
PC2	-14.235	0.000

Comparison of the two models

- The second model, which uses two principal components (PC1 and PC2) as independent variables, has a slightly higher R^2 than the first model (**0.67 > 0.64**), which included only Heating Oil. This suggests that adding the second principal component, and thus more hedging instruments, provides a marginal improvement in explaining kerosene price variability.
- Both models have a low standard error, with the second model showing slightly greater precision and reduced uncertainty in the estimates. Additionally, the second model demonstrates stronger statistical significance of the coefficients with a slightly higher t-statistic and lower p-value.

Comparison of the two models

- The β's parameters represent the optimal **Hedge Ratios** associated with the variables, indicating the weight with which each commodity should be used in the hedging strategy to minimize kerosene price risk. They therefore measure the impact on kerosene price variability and determine the most effective linear combination to reduce residual variance.

- (*In the second model, it will obviously be necessary to carry out a transformation of the PCs to interpret the Hedge Ratios in terms of the original commodities*)

Hedge Effectiveness

$$\text{Portfolio Hedged} = Y_{\text{train}} - (\beta_1 \cdot PC1 + \beta_2 \cdot PC2)$$

$$HE = 1 - \frac{\sigma^2_{\text{hedged}}}{\sigma^2_{\text{unhedged}}}$$

$$\sigma^2_{\text{unhedged}} = \text{Var}(Y_{\text{train}})$$

$$\sigma^2_{\text{hedged}} = \text{Var}(\text{Portfolio Hedged})$$

Hedge Effectiveness

- Our second hedging model manages to reduce the risk by 67% compared to the unhedged portfolio, which means that 67% of the variance in the price of kerosene was eliminated by using the two main components.

- An effectiveness of 67% is considered fair, since about one third of the variance remains uncovered. However, we have shown how a composite hedging approach (Heating Oil + BRENT) is better than a single-contract strategy.

Hedge Effectiveness

Table 1: Hedging Effectiveness (HE) comparison

Strategia di Hedging	**Hedging Effectiveness**
Unhedged portfolio	0.00
Hedging Heating Oil (single-contract approach)	0.64
Hedging PCA (composite appr.)	0.67

- Our strategy is a first step towards a Mimicking Portfolio!

Thank you