Investigation of the relationship between price of futures contracts of energy commodities and their energy generation contribution to US energy grid

Luca Renz & Kasper Aleksander Kuznik

December 2024

Contents

1. Introduction

Energy generation and its relationship to commodity markets have significant implications for financial stability and market behavior. This paper examines the influence of U.S. energy generation—categorized by sources such as natural gas, crude oil, and coal—on the corresponding commodities futures prices. By employing time series analysis, we explore how shifts in energy generation impact the pricing dynamics of these key commodities. While the focus of this documentation is on the primary findings and visualizations, supplementary charts and detailed analyses are included in the appendix.

2. Literature Review

TBD

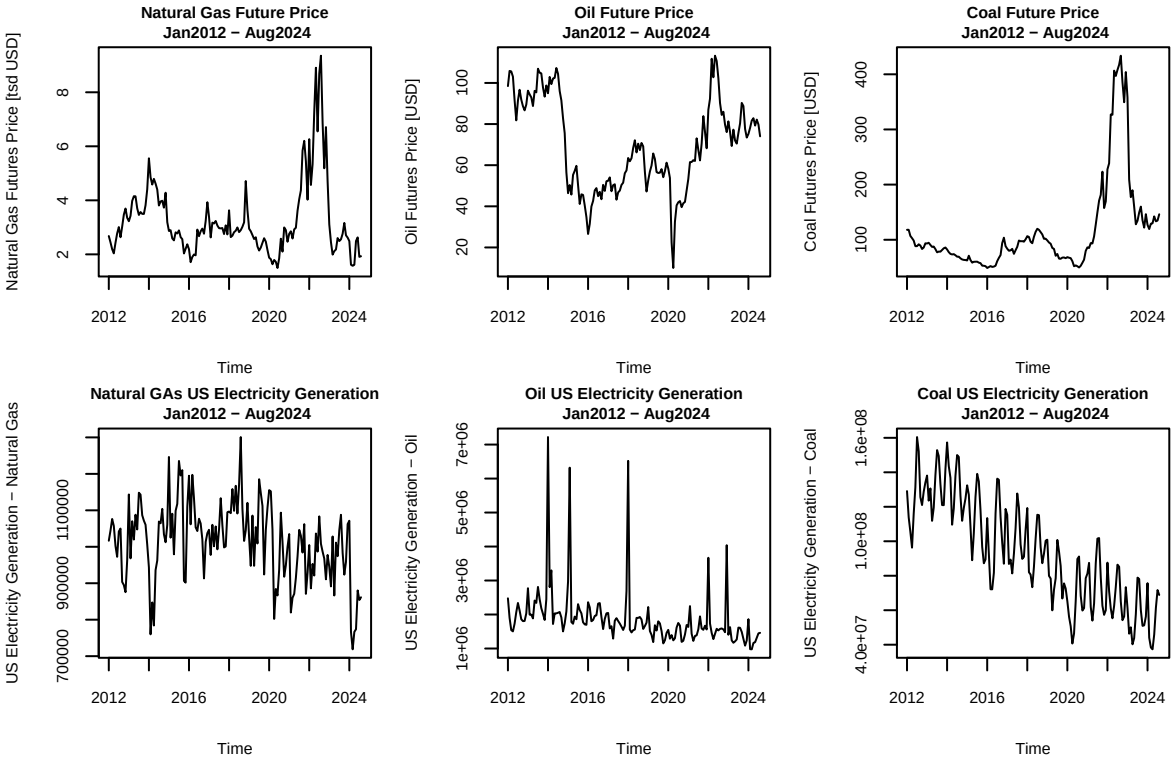
3. Methods

To investigate the causal relationship between U.S. energy generation by source and commodities futures prices for Natural Gas, Crude Oil, and Coal, the first step is to prepare the data to ensure comparability. After cleansing and preprocessing the data, the "Augmented Dickey-Fuller Test" is applied to determine whether the time series are stationary. If any time series are found to be non-stationary, they are transformed using the diff(log()) method until the "Dickey-Fuller Test" confirms stationarity. Once the series are stationary, they are used to construct vector autoregressive models (VAR models) to analyze the dynamic relationships. Additionally, Granger causality tests are performed to evaluate the direction and strength of causality between the variables.

3.1 Data collection and preprocessing

The historical futures prices for Natural Gas, Crude Oil (WTI), and Newcastle Coal were sourced from Barchart.com in the form of CSV files. Data on U.S. electricity generation, categorized by energy source, was obtained from the U.S. Energy Information Administration as XLSX file. This method required manual extraction and preparation. The availability of data varies: Crude Oil futures pricing data begins in 1983, Natural Gas futures data in 1990, Coal futures data in 2009, and U.S. energy generation data in 2001. All datasets are reported on a monthly basis and are available up to August 2024. Despite these differences in starting dates, the data was aligned to ensure consistency for the analysis period.

3.2 Data Preparation and Visualization



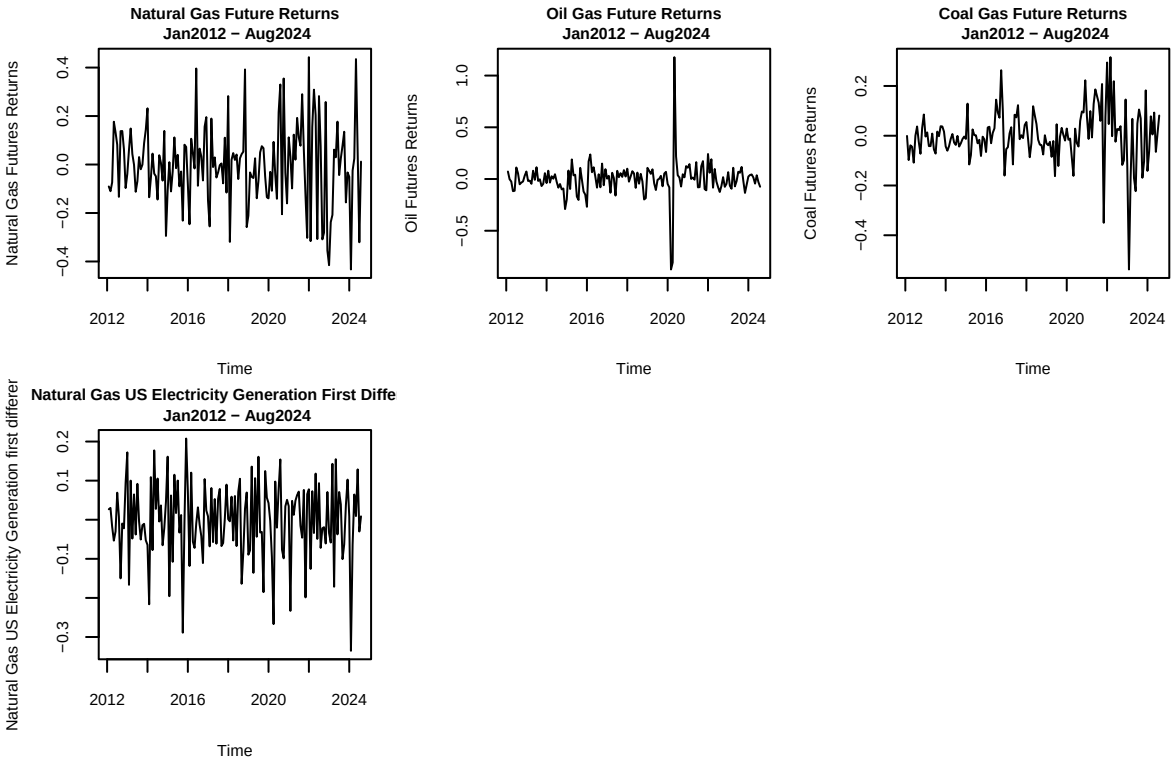
3.3 Time Series Sationarity

To develop the VAR models, it is essential to first ensure that all time series are stationary. A stationary time series is defined as one with consistent statistical properties over time, such as mean, variance, and autocorrelation. To verify stationarity, the Augmented Dickey-Fuller test is applied to each of the investigated time series. The test results, including p-values, are used to assess the stationarity of the time series data, such as futures prices of Natural Gas, Crude Oil, Coal, as well as U.S. energy generation by the energy source.

Time series	p-value
Natural Gas Futures Price	???
Crude Oil Futures Price	???
Coal Futures Prices	???
US Electricity Generation Natural Gas	???
US Electricity Generation Crude Oil	???
US Electricity Generation Coal	???

Table 1: Augmented Diickey-Fuller test results for investigated time series

Time series with a P-value greater than 0.05, such as Natural Gas U.S. energy generation and futures prices of Natural Gas, Crude Oil, and Coal, are classified as non-stationary. To address this, stationarity is introduced by applying diff(log()) transformation.



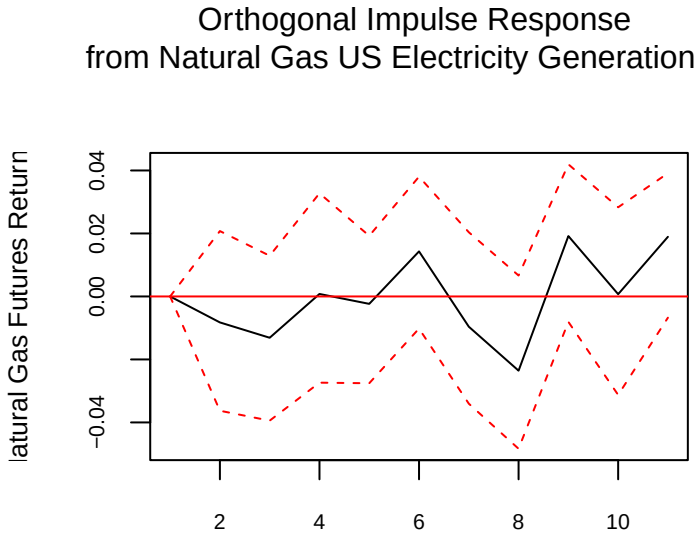
Time series	p-value
Natural Gas Futures Price	???
Crude Oil Futures Price	???
Coal Futures Prices	???
US Electricity Generation Natural Gas	???

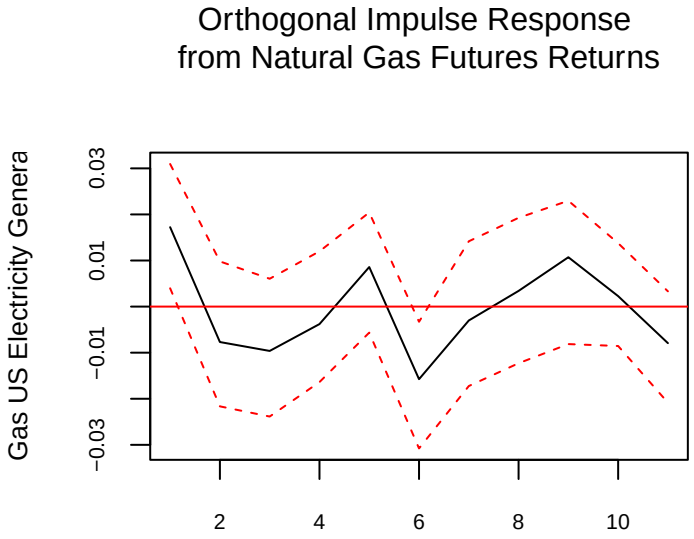
Table 2: Augmented Diickey-Fuller test results for investigated time series after first differences

4. VAR Model and Granger Causality

Once the time series are made stationary by applying first differences, the next step is to fit a Vector Autoregressive model to the stationary data. The VAR model captures the linear relationships between a variable and its own past values, as well as the past values of other endogenous variables... TBD

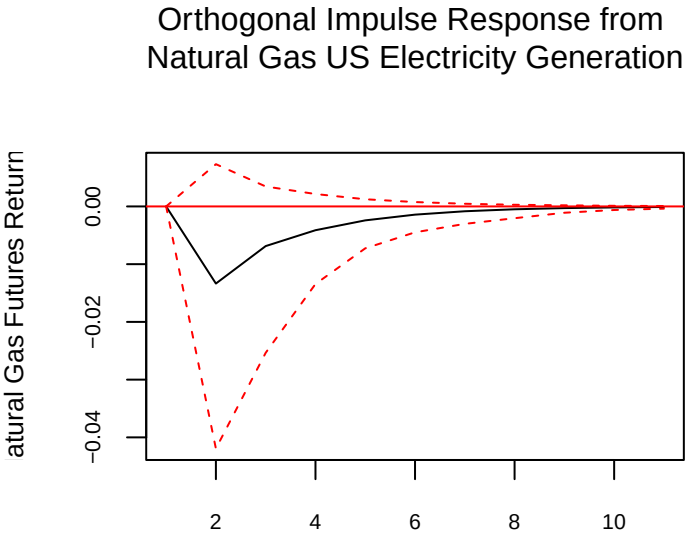
4.1a Natural Gas Futures Returns vs Natural Gas US Electricity Generation Difference

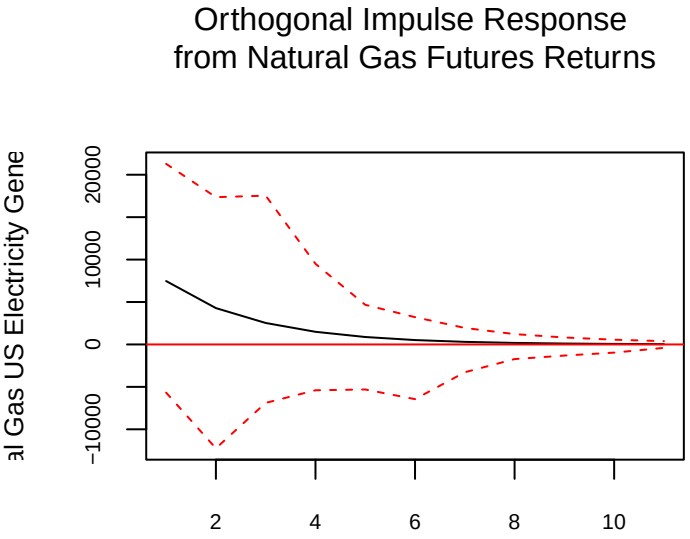




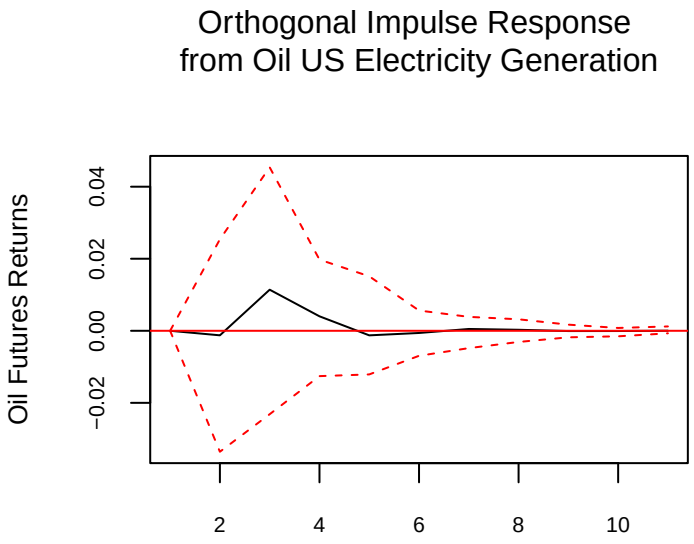
4.1b Natural Gas Futures Returns vs Natural Gas US Electricity Generation

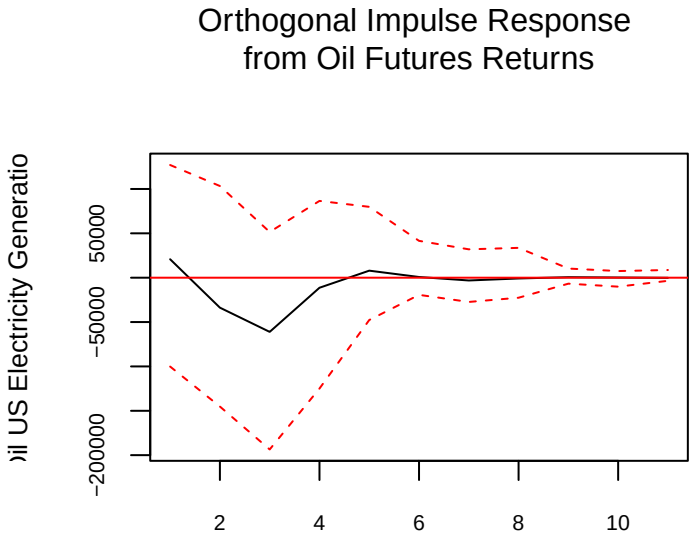
Natural Gas Data is not stationary but give p-value 0.06 checking the results out of curiosity



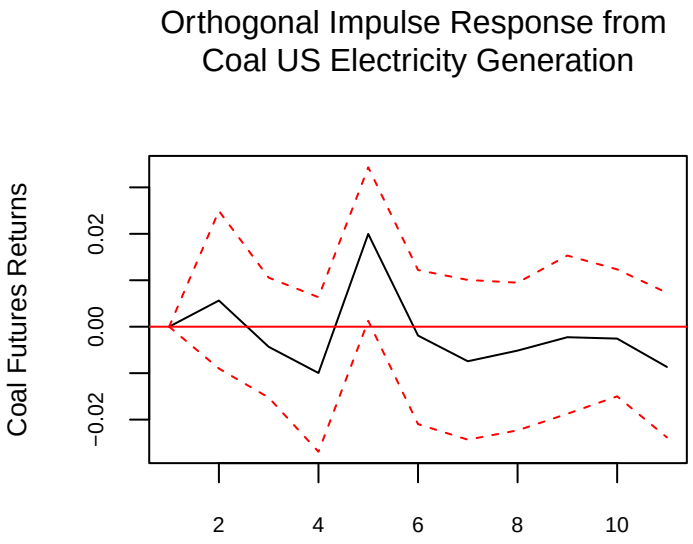


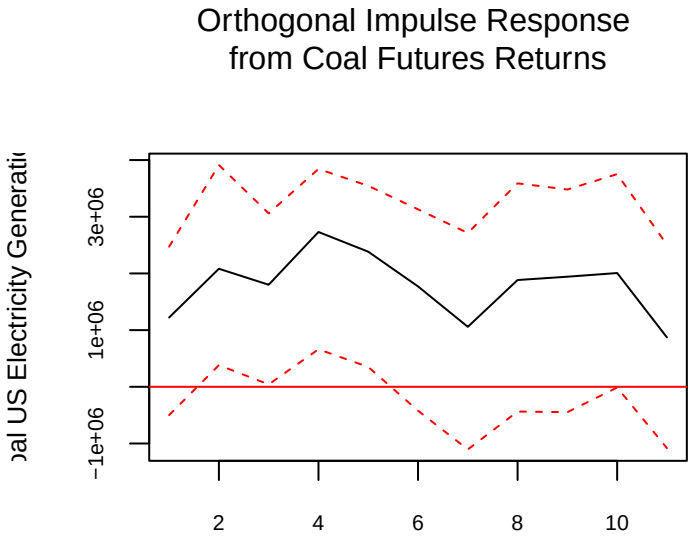
4.2 Oil Futures Returns vs Oil US Electricity Generation





4.3 Coal Futures Returns vs Coal US Electricity Generation





5. Results & Conclusion

TBD

6. Appendix

TBD