The Missing Tariff: Why Didn't the Trump Administration Tax Oil? A Systemic Perspective from Wall Street to Main Street

Nicole Yang, Grace Yu, Xiaolan Deng, Karolina Ochman

Europe Women's Datathon Spring 2025

Citadel | Citadel Securities & Correlation One

Non-Technical Executive Summary

This study investigates the complex and far-reaching consequences of oil price shocks in the United States, using the Trump administration's 2025 "missing tariff" on crude oil as a lens. Drawing on supply decomposition, time series analysis, machine learning, and scenario simulations, we trace how oil price changes ripple through energy supply, financial markets, transportation safety, and household welfare.

We show that, despite significant gains in energy independence through the shale revolution, the U.S. remains vulnerable to global oil price volatility. Financial markets respond with rapid and synchronized swings, while the impact on road safety and household costs unfolds more gradually and unevenly. Empirical results reveal that higher gasoline prices are linked with lagged reductions in traffic fatalities, echoing behavioral adaptation in travel patterns. For families, oil price increases are sometimes absorbed by the market or delayed by regulation, but can still break through during crises, affecting daily life in unpredictable ways.

Simulation of a hypothetical oil tariff demonstrates how even a single policy change could set off wide-ranging, if short-lived, shocks in markets and society alike. Ultimately, our findings underscore the persistent influence of oil—no longer just a story of dependency or independence, but of interconnected risks and adjustments that span from Wall Street to Main Street.

Contents

Non-Technical Executive Summary 1

1 Introduction 3

2 Measuring U.S. Oil Import Dependency 4
- 2.1 Results and Interpretation 6
- 2.2 Data Sources and Processing 7

3 Financial Market Transmission Mechanisms 8
- 3.1 Rolling and Lagged Correlation Analysis between Gasoline Prices and Energy Stocks 8
  - 3.1.1 Market Dynamics in 2022: War in Ukraine 10
  - 3.1.2 Conclusion 11
- 3.2 VAR and Impulse Response Analysis 11
  - 3.2.1 Methodology 11
  - 3.2.2 Conclusion 13
  - 3.2.3 Bonus: Event Study—Simulating a 10% Oil Tariff by the Trump Administration 13
- 3.3 Data Sources and Processing 14

4 Road Safety 15
- 4.1 Data Sources and Processing 17

5 Household Welfare 18
- 5.1 Import Price of Crude Oil and CPI of Fuel Oils and Other Fuels 18
- 5.2 CPI for Household Energy, Fuel Oils, and Other Fuels 19
- 5.3 Import Price of Crude Oil and Vehicle Miles Traveled 21

6 Conclusion 22

Acknowledgements 23

1 Introduction

In April 2025, the Trump administration imposed sweeping retaliatory tariffs on hundreds of imported goods, fundamentally reshaping U.S. trade relationships and sparking volatility across global markets. Yet, in a striking move, crude oil—arguably the most strategically vital commodity in the American economy—was conspicuously left off the tariff list. [1] This "missing tariff" raises a provocative question: What if oil had also been targeted? What ripple effects might such a policy have unleashed throughout the U.S. economy and society?

Over the past two decades, the U.S. energy landscape has been transformed by a complex interplay of global economic forces, domestic policy reforms, and technological advancements. Key geopolitical events such as the COVID-19 pandemic and the Russian invasion of Ukraine have triggered sharp volatility in global oil markets, which in turn has disrupted U.S. fuel prices and supply chains. Meanwhile, regulatory initiatives, environmental priorities, and shifting consumer behavior have shaped both energy production and demand, altering the structure of gasoline supply, household energy costs, and America's position in the global energy trade. While the country has made substantial strides toward greater energy independence and diversification, oil prices remain a central channel through which shocks are transmitted across financial, social, and economic domains.

This report uses the "missing tariff" as a lens to systematically investigate how oil price shocks propagate and amplify across multiple dimensions of the U.S. system:

- First, we reconstruct and analyze the structure of U.S. gasoline supply and oil dependency, providing quantitative evidence of how changes in global and domestic conditions shape the nation's energy security.

- Second, we examine how oil price fluctuations are rapidly transmitted to financial markets and sectoral stocks, and use simulated policy scenarios—such as a hypothetical oil tariff—to assess the magnitude of potential market disruption.

- Third, we explore the lagged but significant effects of oil prices on transportation behavior and public safety, employing both macroeconomic and machine learning approaches to reveal the complex relationship between energy costs and highway fatality rates.

- Finally, we investigate how oil price dynamics affect household energy expenditures and consumer behavior, highlighting threshold effects, price pass-through, and patterns of adaptation.

By integrating empirical statistics, time series modeling, machine learning, and event study analysis, this research reveals the interconnected mechanisms by which oil price shocks reverberate across markets, behaviors, and social outcomes. Our findings provide data-driven insight and policy recommendations for managing risk and promoting economic and social resilience in the face of energy price volatility.

2 Measuring U.S. Oil Import Dependency

To quantify the United States' structural dependence on imported oil for gasoline consumption, we reconstructed a year-by-year breakdown of gasoline supply using publicly available data from the U.S. Energy Information Administration (EIA). We disaggregated the supply into three key components: gasoline produced from domestically refined crude oil, gasoline refined from imported crude, and direct imports of finished gasoline. By accounting for net exports, we derived the effective domestic supply used for U.S. consumption.

A key indicator, the Dependency Ratio, is defined as the proportion of gasoline originating from imported crude and direct gasoline imports, relative to the total net supply:

$$\text{Dependency Ratio} = \frac{G_{\text{imp.crude}} + G_{\text{import}}}{G_{\text{total}} - G_{\text{export}}}$$

In contrast, the Self-sufficiency Score measures the share of gasoline derived from U.S.-produced crude oil:

$$\text{Self-sufficiency Score} = \frac{G_{\text{dom.crude}}}{G_{\text{total}} - G_{\text{export}}}$$

To compute these values, we assumed a crude-to-gasoline conversion efficiency of 46.6%, denoted by θ. This conversion efficiency is based on the average yield of finished motor gasoline from crude oil input at U.S. refineries, as reported by the U.S. Energy Information Administration [4]. The gasoline refined from domestic and imported crude is calculated as:

$$G_{\text{imp.crude}} = C_{\text{imp}} \cdot \theta$$

$$\theta = 0.466$$

Here, G_{import} represents direct gasoline imports, G_{export} is gasoline exports, and G_{total} is the total supply before exports. C_{dom} and C_{imp} refer to the volume of crude oil input from domestic and imported sources, respectively.

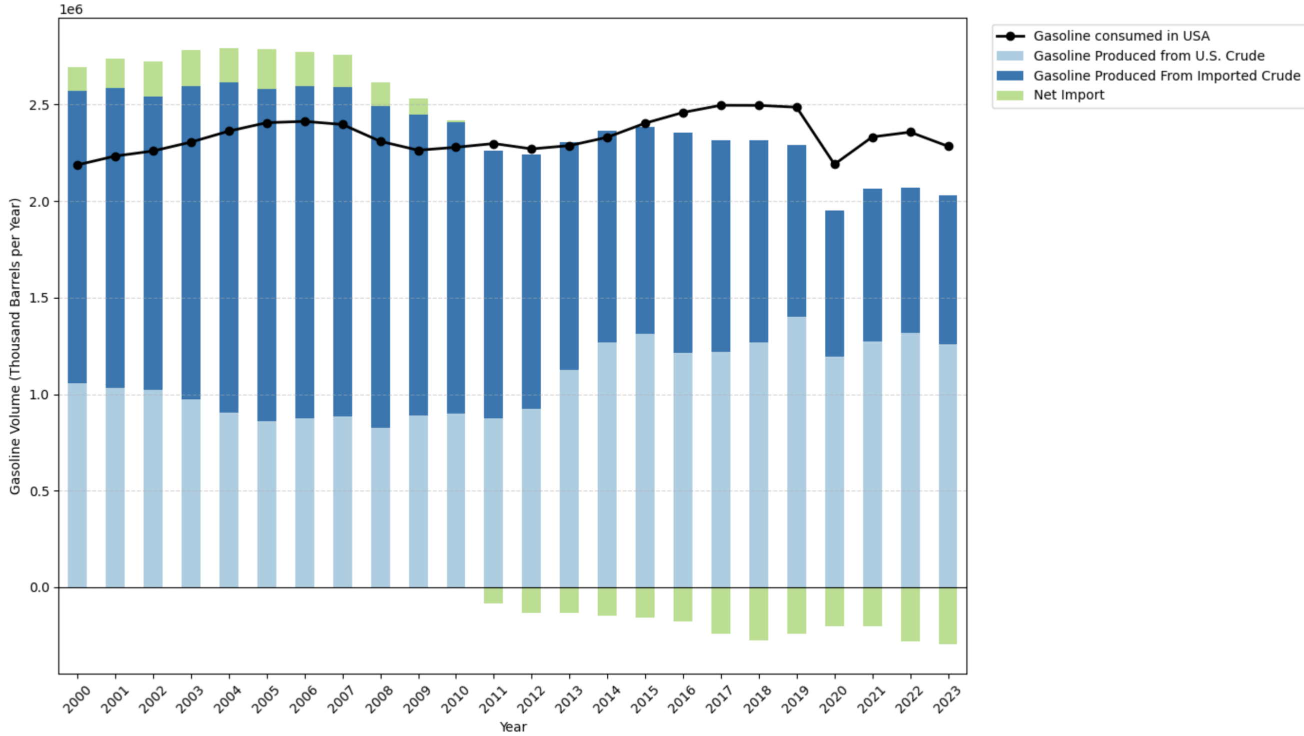

Figure 1: **Decomposition of U.S. Gasoline Supply and Consumption (2000–2023).** This figure visualizes the structural sources of U.S. gasoline supply, decomposed into gasoline refined from domestic crude (light blue), gasoline refined from imported crude (blue), and direct imports of finished gasoline (green). Net exports are subtracted to reflect domestic availability. The black line represents the EIA-reported total gasoline consumption, scaled to be comparable with the supply structure.

A visible decline in dependency on imported crude is observed after 2010, reflecting the impact of the shale oil boom. However, imports, including crude and finished gasoline, remained a significant component of U.S. gasoline supply, particularly in the early 2000s. Domestic gasoline production experienced a steady increase, rising from just above 1 billion barrels in 2000 to about 1.25 billion barrels in 2023. In contrast, gasoline imports declined markedly over the same period, falling from around 155 million barrels annually in 2000 to negligible levels by the early 2010s. This dramatic shift was driven by rising domestic production and refining capacity, as well as the 2011 policy change that allowed broader gasoline exports. Following this regulatory shift, the U.S. emerged as a significant gasoline exporter, with volumes surging to approximately 298 million barrels annually by 2023. This transition marked a pivotal turn in U.S. energy dependence, from a heavy reliance on foreign gasoline to becoming a dominant net exporter. Meanwhile, the overall volume of gasoline in the U.S. market declined, reflecting reduced domestic demand. This was shaped by advances in fuel efficiency driven by federal standards, changes in driving behavior such as increased urbanization and remote work, and the growing adoption of hybrid and electric vehicles. Additionally, the expanded use of ethanol blends like E10 and E15 further decreased the demand for petroleum-based gasoline. Environmental regulations and climate policies also supported fuel diversification and emissions reduction efforts. Collectively, these technological, behavioral, and regulatory shifts contributed to a structural decrease in gasoline consumption, even as the U.S. strengthened its energy independence and global export position.

2.1 Results and Interpretation

Year	Dependency Ratio	Self-sufficiency Score
2000	0.608	0.392
2001	0.623	0.377
2002	0.624	0.376
2003	0.650	0.350
2004	0.676	0.324
2005	0.692	0.308
2006	0.684	0.316
2007	0.679	0.321
2008	0.684	0.316
2009	0.647	0.353
2010	0.640	0.376
2011	0.639	0.411
2012	0.618	0.445
2013	0.543	0.517
2014	0.501	0.563
2015	0.491	0.578
2016	0.530	0.549
2017	0.527	0.571
2018	0.516	0.594
2019	0.460	0.644
2020	0.465	0.643
2021	0.463	0.644
2022	0.458	0.673
2023	0.477	0.662

Table 1: Dependency Ratio and Self-sufficiency Score (2000–2023)

As shown in Table 1, the U.S. gasoline supply was significantly dependent on imports in the early 2000s, with the Dependency Ratio exceeding 60% and the Self-sufficiency Score consistently below 40%. For reference, many developed countries such as Japan and South Korea maintain energy dependency rates above 80%, while European countries generally fall within the 50–70% range [2]. In contrast, U.S. energy policy has long emphasized strategic autonomy.

From around 2013 onward, the U.S. began reducing its reliance on imported crude and finished gasoline. This transition is strongly associated with the *shale revolution*, which refers to the large-scale adoption of horizontal drilling and hydraulic fracturing technologies that significantly boosted domestic crude oil production [3]. By 2020–2023, the Self-sufficiency Score exceeded 64%, while the Dependency Ratio fell below 50%, suggesting a major structural shift in supply resilience. Nonetheless, even in recent years, imported components still constitute a notable portion of the U.S. gasoline supply.

2.2 Data Sources and Processing

All analyses in this section are based on the official dataset `all_commodities.csv` provided by the organizers. Using this file, we extracted annual values for each component of the gasoline supply chain and computed the Dependency Ratio and Self-sufficiency Score according to the formulas given in the methodology. Basic data cleaning and calculation were performed in Python to ensure accuracy and consistency with the definitions used in the competition.

3 Financial Market Transmission Mechanisms

In summary, this section uses both correlation-based and time series methods to quantify the rapid and uneven ways in which oil price changes affect financial assets. These findings lay the groundwork for evaluating the broader market consequences of energy policy shocks.

3.1 Rolling and Lagged Correlation Analysis between Gasoline Prices and Energy Stocks

To investigate the evolution and complexity of financial linkages between gasoline prices and major energy stocks, we begin with a 90-day rolling correlation analysis over the period 2000–2024. This method captures the time-varying relationship between the two markets and allows us to identify episodes of elevated or diminished comovement under different market regimes.

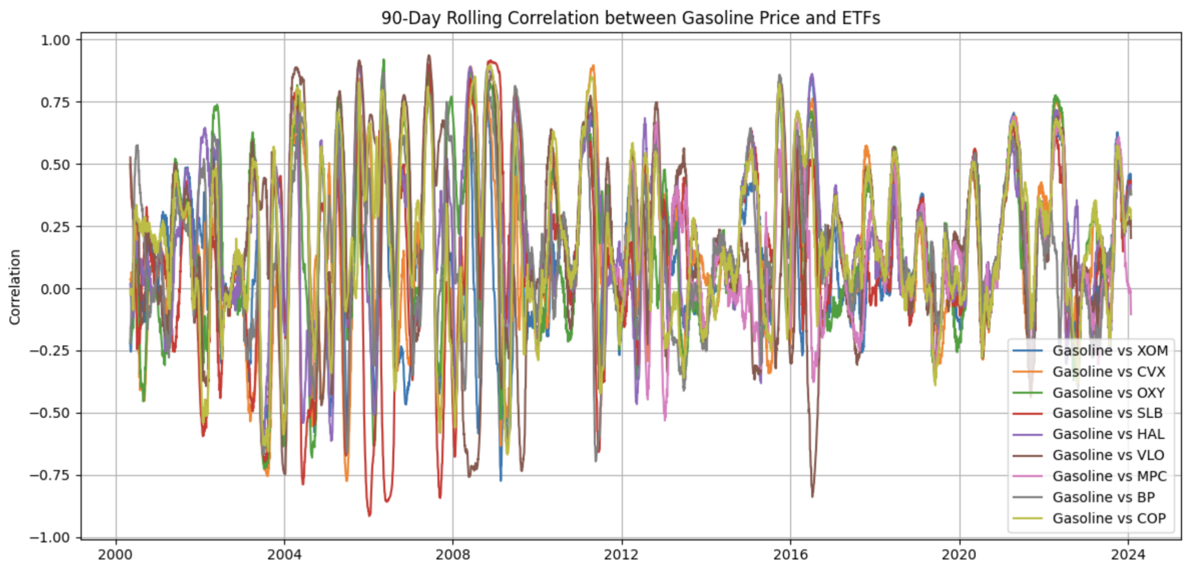

Figure 2: 90-Day Rolling Correlation between Gasoline Price and Major Energy Stocks (2000–2024)

As shown in Figure 2, the correlations between gasoline prices and energy stocks are highly dynamic, fluctuating between positive and negative values throughout the sample period. Episodes of increased correlation frequently coincide with periods of market turbulence or major external shocks, such as the 2008 financial crisis or oil market disruptions. This pronounced variability underscores the limitations of static correlation measures, emphasizing the importance of accounting for time-varying risk transmission channels in financial market analysis. Furthermore, the diversity observed across different assets suggests that the sensitivity to oil prices is both asset-specific and temporally unstable.

To further explore the heterogeneity in timing and strength of the oil-stock relationship, we conduct a lagged correlation analysis. For each stock, we calculate the maximum correlation coefficient with gasoline prices across a range of lags, recording the lag at which this maximum occurs. This approach enables us to distinguish between synchronous, lagged, and leading responses.

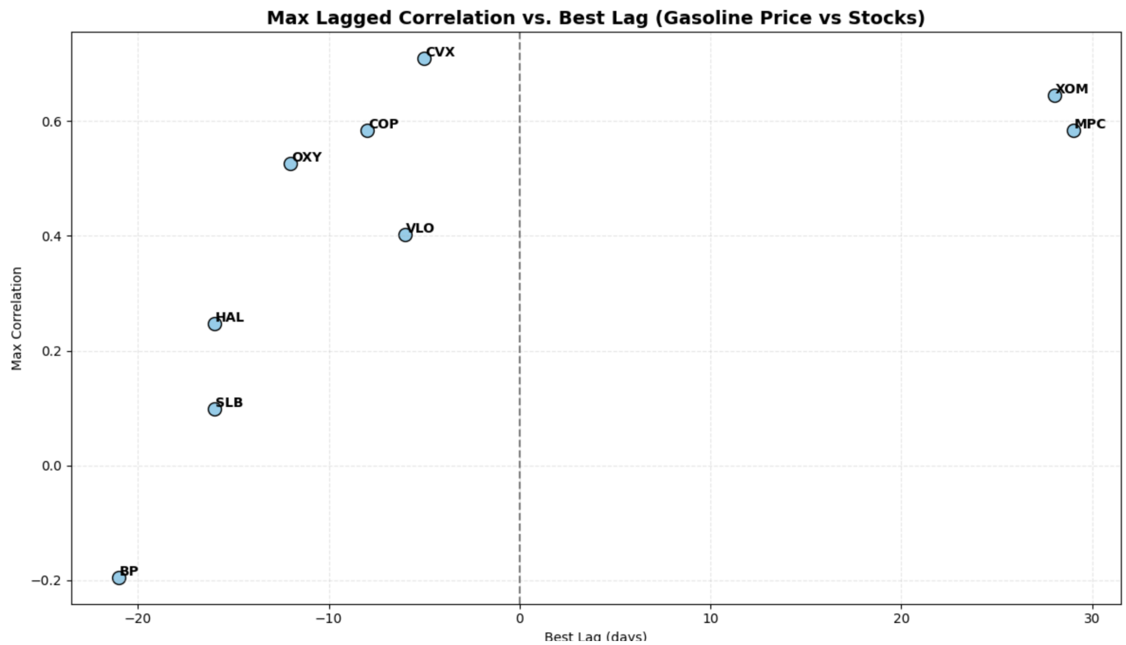

Figure 3: Max Lagged Correlation vs. Best Lag (Gasoline Price vs. Stocks)

The resulting scatter plot (Figure 3) and summary statistics (Table 2) provide detailed insights into the timing and magnitude of stock responses to gasoline price movements. Most stocks exhibit their highest correlation with gasoline prices at a positive lag, indicating that stock returns tend to react after changes in the energy market. For example, XOM and MPC are classified as 'Lead' responders with best lags of 28 and 29 days, respectively, while CVX displays both the highest maximum correlation (0.71) and a relatively synchronous response to oil price changes. The magnitude of the maximum correlation also varies considerably across assets, with some stocks (e.g., VLO, SLB) displaying particularly strong sensitivity, while others (e.g., BP, MPC) are less affected. The SyncScore further quantifies the degree of synchrony for each asset, highlighting substantial heterogeneity in the speed of adjustment across the sector.

Ticker	MaxCorrelation	BestLag	LagCategory	SyncScore
CVX	0.710	-5	Synchronous	0.118
COP	0.583	-8	Lag	0.065
VLO	0.403	-6	Lag	0.058
OXY	0.527	-12	Lag	0.041
XOM	0.645	28	Lead	0.022
MPC	0.584	29	Lead	0.019
HAL	0.247	-16	Lag	0.015
SLB	0.099	-16	Lag	0.006
BP	-0.196	-21	Lag	-0.009

Table 2: Summary of Lagged Correlation Analysis: Maximum Correlation, Best Lag, and Synchronization Score by Ticker

3.1.1 Market Dynamics in 2022: War in Ukraine

In 2022, international oil prices experienced their most significant volatility since the 2008 financial crisis, largely driven by geopolitical shocks such as the outbreak of the Russia-Ukraine war. To better understand the dynamic linkages between the energy and financial markets under extreme conditions, we focus specifically on 2022 and analyze the 30-day rolling correlations between gasoline prices and major energy stocks (see Figure 4).

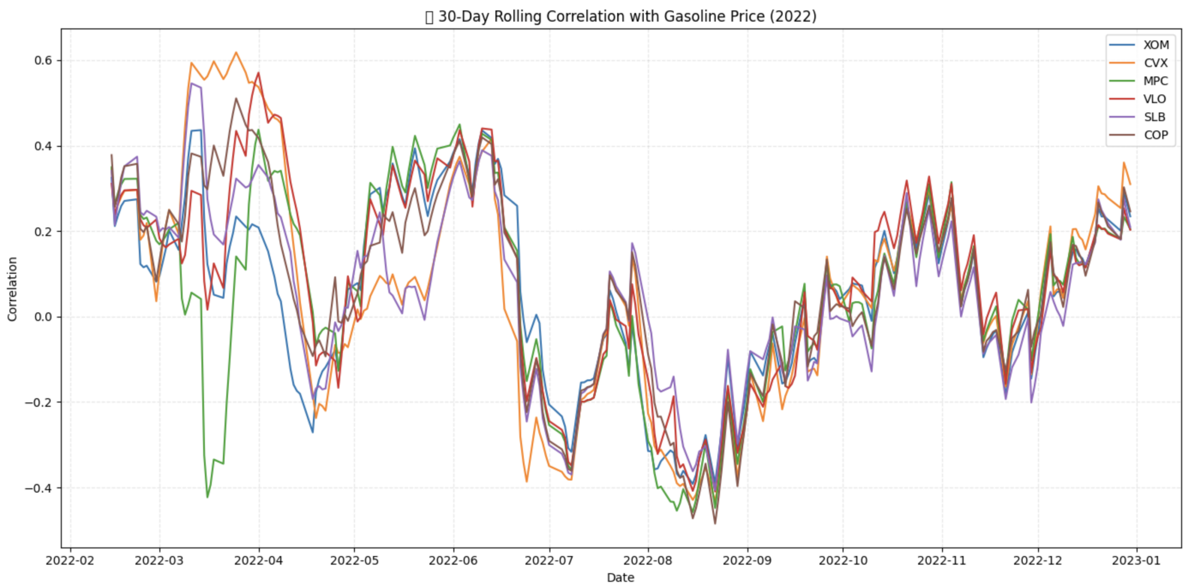

Figure 4: 30-Day Rolling Correlation between Gasoline Price and Selected Energy Stocks (2022)

As shown in the figure, during the sharp surge in oil prices from February to March 2022, rolling correlations between energy stocks and gasoline prices increased dramatically, with several stocks exhibiting coefficients above 0.6. This pattern indicates that the sensitivity of energy stocks to oil price fluctuations intensifies significantly during periods of external shocks, resulting in heightened market interconnectedness.

In the second half of 2022, as oil price volatility moderated, the overall correlations between energy stocks and gasoline prices declined, typically falling to the 0.2–0.3 range. At the same time, heterogeneity among individual stocks became more pronounced. For example, leading firms such as XOM and CVX saw their correlations decrease more rapidly, while others, such as MPC and COP, maintained relatively high levels of comovement. These observations further confirm the heterogeneity in asset sensitivity and the temporal instability of these relationships.

Overall, the 2022 case study clearly demonstrates that: (1) extreme market events can sharply amplify the impact and speed of oil price transmission to financial assets; (2) correlations are highly dynamic, switching between crisis and normal periods rather than remaining stable; and (3) there are substantial differences in sensitivity across assets. This underscores the necessity of employing dynamic, asset-specific frameworks for risk monitoring and policy evaluation, rather than relying solely on static correlation measures.

3.1.2 Conclusion

Taken together, the results from both the long-term analysis and the 2022 case study highlight the time-varying and asset-specific nature of financial linkages between energy markets and sector stocks. These findings provide a strong rationale for further investigation using formal time series models. In the following section, we employ vector autoregression (VAR) and impulse response analysis to rigorously quantify the direction and magnitude of oil price shocks on financial assets.

3.2 VAR and Impulse Response Analysis

To rigorously investigate the dynamic transmission mechanisms between oil prices and energy stocks, we estimate VAR (vector autoregression) models using weekly returns for gasoline prices and a set of major energy sector stocks (e.g., CVX, XOM, VLO). This modeling framework, widely used in the literature (e.g., Kilian & Park, 2009 [6]), allows us to examine not only the direction but also the magnitude and persistence of shock transmission between markets.

3.2.1 Methodology

We construct a VAR model including gasoline price returns and returns for selected energy stocks, with the optimal lag length determined using the Akaike Information Criterion (AIC). Granger causality tests are applied to assess whether oil prices can be considered leading indicators for stock returns, and vice versa. We then compute impulse response functions (IRFs) to trace the response of each variable to a one-standard-deviation shock in another, providing an intuitive visualization of dynamic effects.

The impulse response analysis presented in Figure 5 provides a detailed view of how major energy stocks respond dynamically to shocks in gasoline prices. The results demonstrate that following a positive shock to oil prices, energy sector stock returns exhibit an immediate and statistically significant increase. For most stocks, the peak response occurs within the first one or two periods, after which the effect diminishes rapidly and returns to baseline within approximately a week. This pattern highlights the "short and sharp" nature of price transmission in the energy-financial nexus, where the market rapidly incorporates new information before the impact fades.

Importantly, the magnitude and persistence of the response vary substantially across assets, underscoring pronounced heterogeneity in risk exposure. For instance, XOM displays the strongest initial reaction, with a peak impulse response around 0.15, whereas VLO exhibits a more muted and delayed effect, possibly reflecting differences in operational leverage and market positioning. The confidence intervals for the initial response periods are generally well above zero for most stocks, indicating that these effects are both statistically and economically significant in the short term. However, as the number of periods increases, the confidence intervals quickly widen and encompass zero, confirming that the market impact of oil price shocks is transitory.

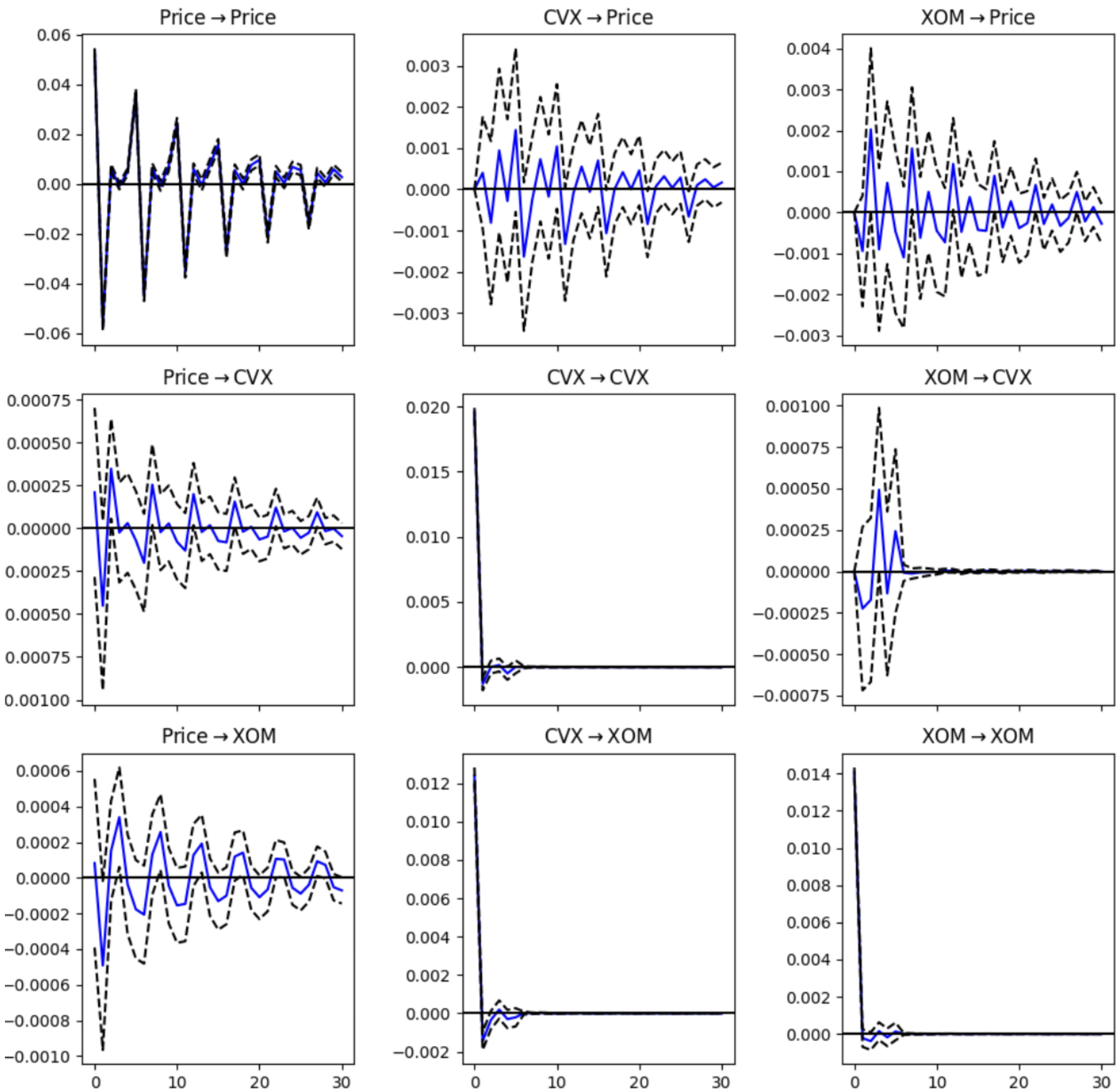

Figure 5: Impulse Response Functions for Oil Prices and Energy Stocks (selected pairs)

Further analysis reveals that the transmission of risk is predominantly unidirectional: shocks to energy stock returns have little to no discernible effect on subsequent oil prices, as evidenced by impulse responses that are consistently indistinguishable from zero. This finding corroborates the view that oil prices serve as a key driver of risk transmission to the financial sector, rather than vice versa.

In addition, the impulse response results and synchronization measures indicate that the leading energy firms—such as XOM, CVX, and other large-cap stocks—exhibit exceptionally high correlation in their responses to oil price shocks. This pronounced synchronization suggests that the sector as a whole prices in oil-related news almost instantaneously, magnifying both the speed and scale of market reactions. As a result, any significant policy intervention or external shock would be rapidly transmitted across all major players, amplifying systemic risk and sector-wide volatility.

3.2.2 Conclusion

Our results demonstrate that financial markets—especially the energy sector—are highly sensitive and synchronised in their response to oil price changes. Any significant increase in oil prices, such as from a new tariff, would be rapidly transmitted across major firms and ETFs, leading to sharp and widespread volatility. Although these market reactions tend to be short-lived, their intensity can trigger substantial financial disruption, increase uncertainty, and amplify systemic risk throughout the sector. This underscores the need for cautious policy design, as abrupt oil price shocks can have pronounced and immediate negative effects on both investors and the broader economy.

3.2.3 Bonus: Event Study—Simulating a 10% Oil Tariff by the Trump Administration

To provide a forward-looking perspective on potential policy impacts, we conduct a counterfactual event study to simulate the effect of a hypothetical 10% retaliatory tariff on oil by the Trump administration. This scenario is inspired by the series of tariffs imposed on other product categories during the same period, but which never extended to oil.

Figure 6 displays the cumulative abnormal returns (CAR) of major energy stocks (CVX, XOM, VLO) and the S&P 500 index (SPY) surrounding the simulated tariff announcement date (marked by the red dashed line). The CARs are computed relative to a market model benchmark, isolating the reaction attributable to the policy shock.

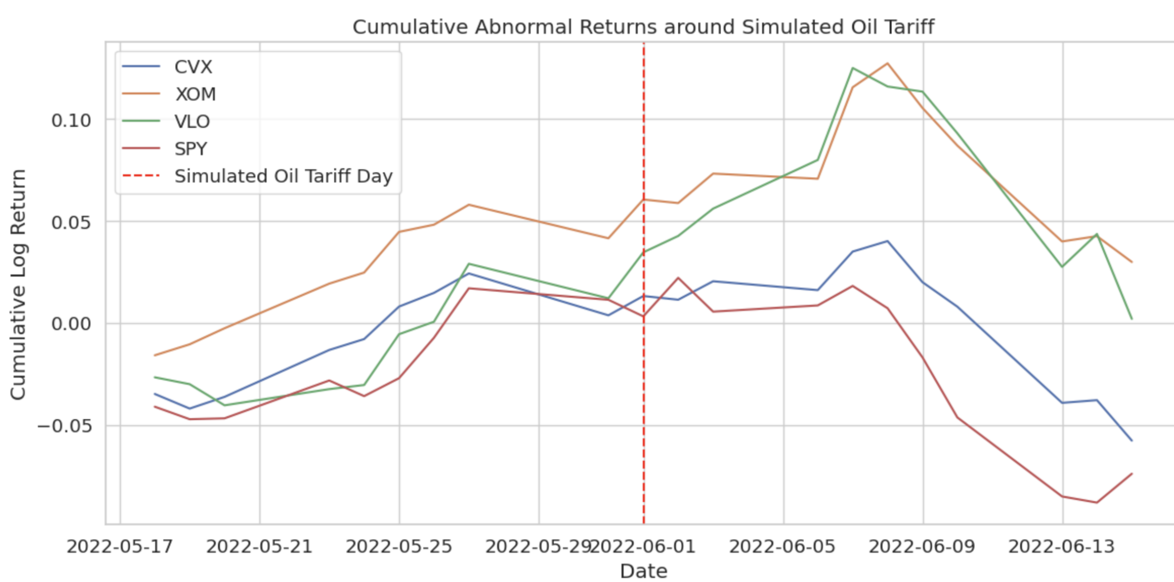

Figure 6: Cumulative Abnormal Returns around Simulated Oil Tariff Announcement

As shown in the figure, all tracked energy stocks, as well as the market benchmark, exhibit pronounced volatility immediately following the simulated tariff event. Notably, energy stocks such as XOM and VLO experience a sharp rise in cumulative abnormal returns in the days after the simulated announcement, reaching their peaks roughly a week after the event before reversing course and declining rapidly. In contrast, SPY—representing the broader market—also shows a temporary positive response, but quickly diverges from energy stocks and enters a sustained decline, suggesting a market-wide reassessment of risk and growth prospects.

This simulated scenario illustrates several key points: First, the financial markets would likely respond with immediate and pronounced swings in valuation if a major oil tariff were enacted, consistent with the sensitivity documented in earlier sections. Second, the rebound and subsequent reversal in energy stock returns underscores the “short and sharp” adjustment dynamics previously identified in the VAR analysis. Finally, the divergence between sector stocks and the broader market index highlights the systemic and sector-specific risks such a policy shock could introduce.

These results reinforce the broader argument that—even in the absence of historical precedent—major policy interventions in the oil market could trigger swift and substantial volatility, affecting not only energy firms but also the stability of the wider financial system.

3.3 Data Sources and Processing

We used weekly gasoline price data and stock price data (`weekly_gasoline_prices.csv`, `all_stocks_and_etfs.csv`) provided by the organizers. All time series were first checked for missing or inconsistent values, then converted to a uniform date format and merged based on weekly timestamps. We filtered for the major energy sector tickers analyzed in this section. Data cleaning and feature engineering—including log-returns computation for stock prices—were conducted in Python. The cleaned dataset was then used to calculate rolling and lagged correlations, as well as to fit VAR models and generate impulse response functions, using the `pandas`, `numpy`, and `statsmodels` packages.

4 Road Safety

Understanding the broader societal impacts of energy price shocks has become increasingly important in light of the ongoing volatility in global oil markets. A substantial body of literature, notably Burke et al. (2015) [7], has established that higher gasoline prices are associated with reductions in U.S. traffic fatality rates, though with considerable temporal lag, reflecting gradual behavioral adjustments by households and commuters.

Building on these findings, we conduct a detailed empirical analysis using quarterly U.S. highway fatality data, gasoline prices, and relevant macroeconomic indicators. Our approach integrates lagged correlation analysis with state-of-the-art predictive modeling, aiming to both quantify the delay and magnitude of oil price effects on road safety and evaluate the predictive value of these relationships.

Figure 7 plots the lagged Pearson correlation coefficients between gasoline prices and highway fatalities across a range of lags (0–24 months), along with corresponding P-values. The results reveal that the positive association between gasoline prices and fatality rates strengthens at longer lags, with correlation coefficients rising to approximately 0.25 and P-values dropping below 0.01 at lags of 12 months and beyond. This delayed effect supports the hypothesis articulated in the literature [7], that changes in fuel prices impact driving and safety outcomes not instantaneously but through adjustments in travel patterns, vehicle miles traveled, and perhaps substitution toward safer vehicles over time.

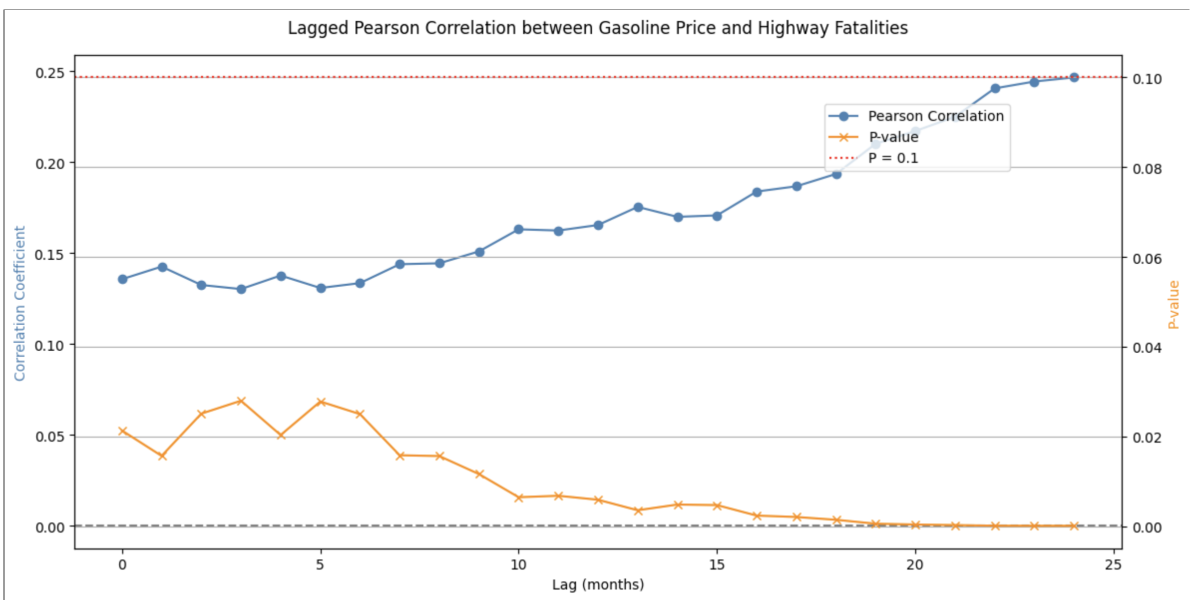

Figure 7: Lagged Pearson Correlation between Gasoline Price and Highway Fatalities

To further evaluate the predictive value of macroeconomic indicators for highway fatalities, we train a suite of machine learning models—including linear regression, ridge regression, random forest, gradient boosting, and support vector regression (SVR)—using quarterly data. Figure 8 compares the actual and predicted highway fatality counts for each model. While all models broadly capture the overall direction and cyclical movements of fatalities, they systematically underestimate peak values and sharp fluctuations, suggesting the existence of important behavioral or contextual variables not captured by macro-level predictors alone.

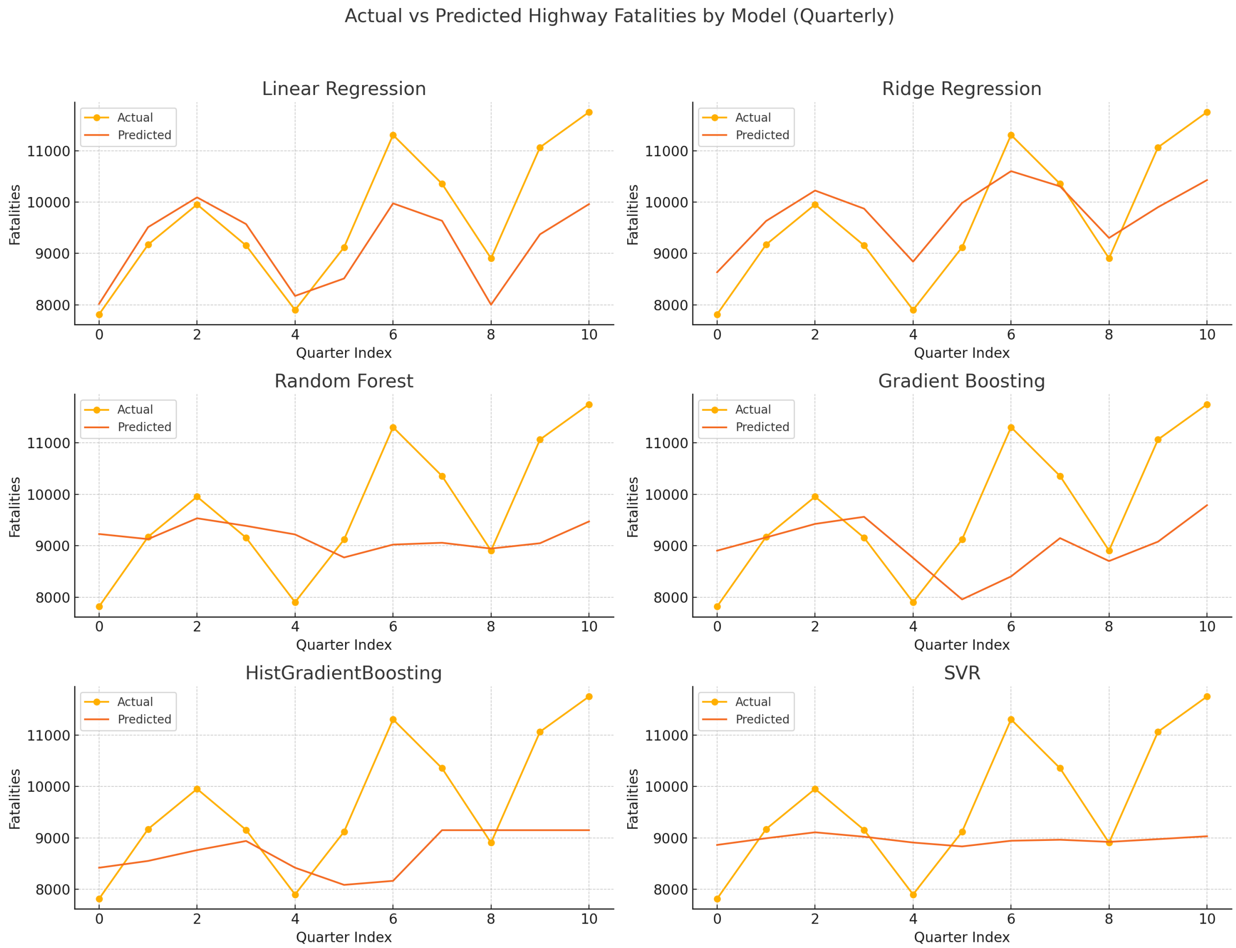

Figure 8: Actual vs. Predicted Highway Fatalities by Model (Quarterly)

The comparative predictive performance of these models is summarized in Table 3. Notably, ridge regression achieves the lowest mean absolute error (MAE) and root mean squared error (RMSE), as well as the highest explanatory power ($R^2 = 0.61$), substantially outperforming both ordinary linear regression and advanced ensemble methods. The relatively weak performance of random forest, gradient boosting, and SVR—each yielding negative R^2 values—highlights the limitations of flexible algorithms when data are sparse, noisy, and highly collinear, as is often the case with quarterly traffic safety data.

Table 3: Quarterly Model Prediction Performance: Highway Fatality Forecasting

Model	**MAE**	**RMSE**	$\mathbf{R}^2$
Ridge Regression	≈ 701	≈ 790	0.61
Linear Regression	≈ 765	≈ 952	0.43
Random Forest	≈ 1063	≈ 1357	-0.16
Gradient Boosting	≈ 1121	≈ 1397	-0.23
SVR	≈ 1098	≈ 1421	-0.27

The superior performance of ridge regression likely reflects its ability to mitigate overfitting and address multicollinearity through regularization, making it particularly well-suited for small-sample, high-noise settings where explanatory variables are strongly correlated. This finding aligns with established results in the statistical learning literature [9] and suggests that in the context of macro-social data, regularized linear models may offer more robust and interpretable forecasts than more complex, non-parametric approaches.

Overall, our findings corroborate the established literature on the delayed impact of oil prices on traffic fatalities [7, 8] and add new evidence that predictive modeling based solely on macro-level features remains limited. To improve forecasting and inform policy, future work should incorporate behavioral, regulatory, and high-frequency data, moving beyond aggregate indicators to capture the multifaceted mechanisms linking energy costs and road safety.

4.1 Data Sources and Processing

This section utilizes quarterly data on U.S. highway traffic fatalities, gasoline prices, and macroeconomic indicators, as provided in `monthly_transportation_statistics.csv` and `weekly_gasoline_prices.csv`. Monthly data were aggregated to a quarterly frequency to match the fatality data. All series were aligned on date, checked for missing values, and outliers were visually inspected and, where appropriate, removed or imputed. Predictor variables, including lagged gasoline prices and unemployment rates, were constructed using `pandas`. The resulting panel was used for lagged correlation analysis and to train predictive models in Python (`scikit-learn`), with model performance evaluated via cross-validation.

5 Household Welfare

This section explores how oil price fluctuations are transmitted to U.S. households through a variety of economic channels. By examining the relationships between crude oil import prices, consumer energy price indices, and transportation demand, we assess both the magnitude and timing of price pass-through, as well as the adaptive behaviors that shape household vulnerability to energy shocks. Together, these analyses provide new insight into the complex and sometimes asymmetric ways that global oil market volatility impacts American consumers.

5.1 Import Price of Crude Oil and CPI of Fuel Oils and Other Fuels

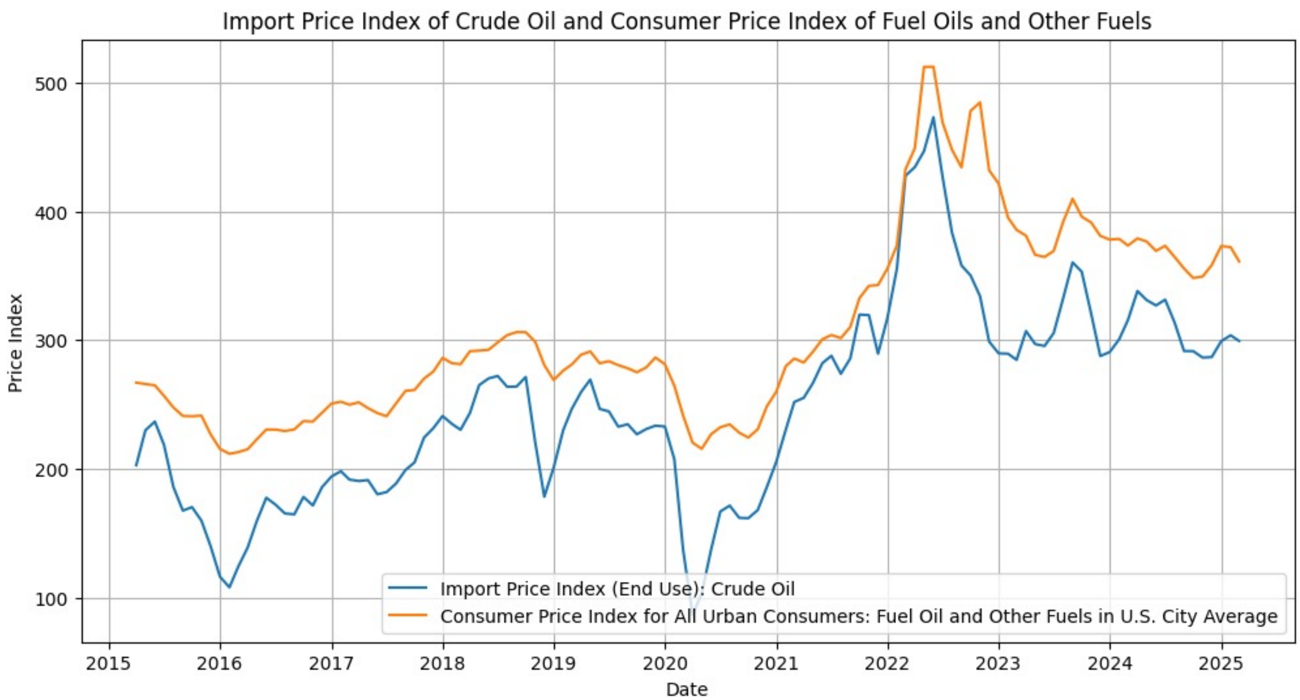

Figure 9: Import Price Index of Crude Oil and Consumer Price Index of Fuel Oils and Other Fuels.

Aside from the sharp divergence in 2020, the CPI for fuel oil and other fuels generally moves in tandem with U.S. import oil prices, as both are driven by similar supply and demand dynamics in global energy markets. However, despite the dramatic drop in import oil prices in 2020, when U.S. crude briefly traded at negative prices due to oversupply and storage issues, the CPI for fuel oil and other fuels for all urban consumers did not fall nearly as sharply. This difference is largely because the CPI reflects retail prices paid by consumers, not raw commodity prices. The CPI incorporates various additional costs such as refining, transportation, distribution, and retail markups, all of which do not fluctuate as rapidly or dramatically as crude oil prices. Moreover, price transmission from the wholesale to retail level tends to lag, and prices often exhibit downward rigidity. Also companies may be slower to reduce prices, especially amid uncertainty about future supply and demand. Additionally, the pandemic introduced logistical disruptions and reduced refinery operations, which limited the extent to which lower crude costs could be passed on to consumers. As a result, while the underlying cost of oil was extremely low, the final prices faced by households remained relatively more stable.

5.2 CPI for Household Energy, Fuel Oils, and Other Fuels

This scatter plot illustrates the relationship between the U.S. Import Price Index for Crude Oil and the Consumer Price Index (CPI) for Household Energy. Overall, the plot reveals a threshold-like relationship. At lower and moderate levels of crude oil prices (below approximately 300), the household energy CPI remains relatively stable, clustered between 180 and 220. This suggests that under typical market conditions, household energy costs are somewhat insulated from fluctuations in crude oil import prices—likely due to regulatory factors, fixed utility contracts, or reliance on other energy sources like electricity and natural gas. However, once crude oil prices rise beyond a certain threshold (above 300), a noticeable shift occurs: the household energy CPI increases more steeply, forming a scattered upper band between 220 and 270+. This pattern reflects the delayed but significant pass-through effect of oil price shocks on household energy expenses, particularly during periods like 2022–2023 when global oil prices surged.

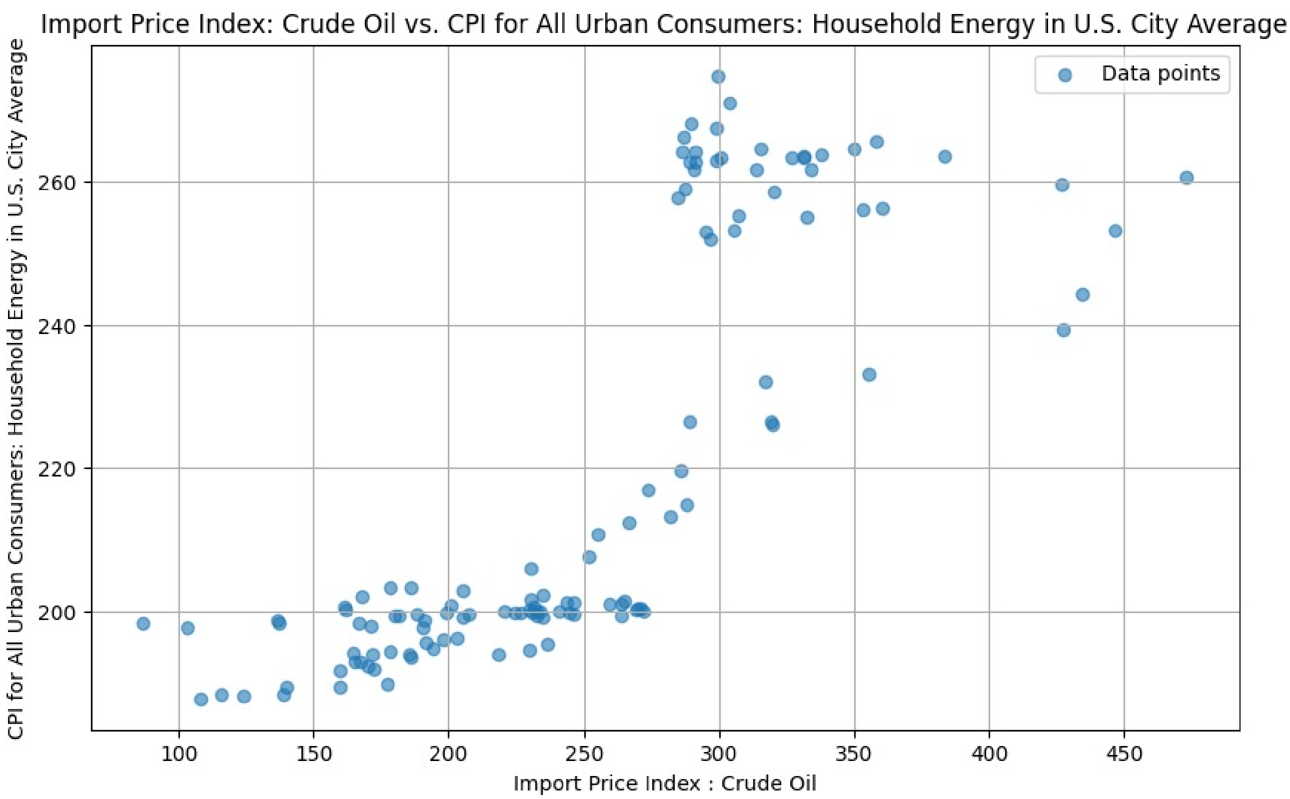

Figure 10: Import Price Index of Crude Oil and Consumer Price Index for All Urban Consumers: Household Energy in U.S. City Average.

When analyzing the graph showing the Consumer Price Index (CPI) for all urban consumers in the U.S. on household energy and on fuel oil and other fuels, a few key economic and geopolitical events help explain the trends. One major decline in the fuel-related CPI occurs in 2020, which corresponds with the onset of the COVID-19 pandemic and the oil price war between Saudi Arabia and Russia. During this time, global demand for fuels plummeted as lockdowns and travel restrictions were implemented worldwide. At the same time, Saudi Arabia and Russia ramped up oil production in a price war that flooded the market with supply. This led to an unprecedented situation where oil futures briefly turned negative in the U.S., meaning suppliers were effectively paying buyers to take oil off their hands due to lack of storage capacity, an extreme case of over-supply. The CPI for fuel oil dropped sharply in response. In contrast, by 2022, the economic dynamics had shifted dramatically. The Russian invasion of Ukraine led to widespread sanctions on Russian exports, including oil and natural gas, reducing the global energy supply and driving prices up. This was particularly impactful in Europe, which relied heavily on Russian energy, but the shockwaves were felt globally, including in the U.S. As a result, the CPI for fuel oil and other fuels spiked. However, the CPI for general household energy did not increase as steeply. This more moderate rise can likely be attributed to the diverse energy mix used in U.S. households. Natural gas, electricity from renewables, and nuclear power helped cushion the impact, preventing a uniform surge across all household energy components.

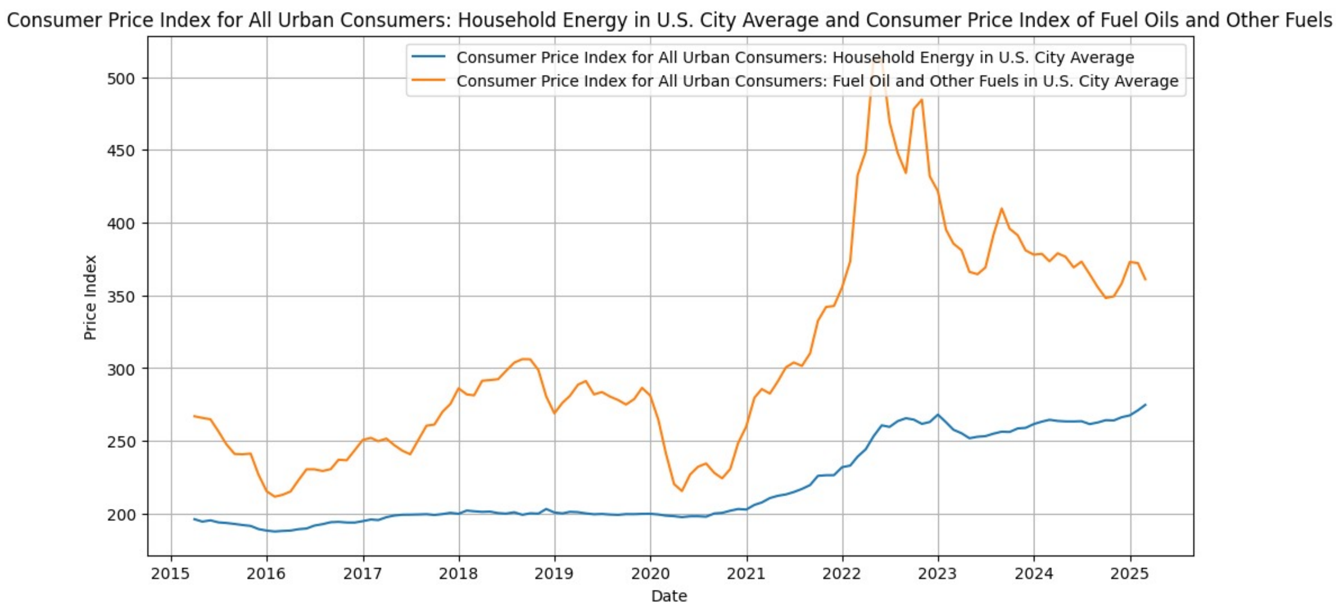

Figure 11: Consumer Price Index for All Consumers: Household Energy in U.S. Average and Consumer Price Index of Fuel Oils and Other Fuels.

5.3 Import Price of Crude Oil and Vehicle Miles Traveled

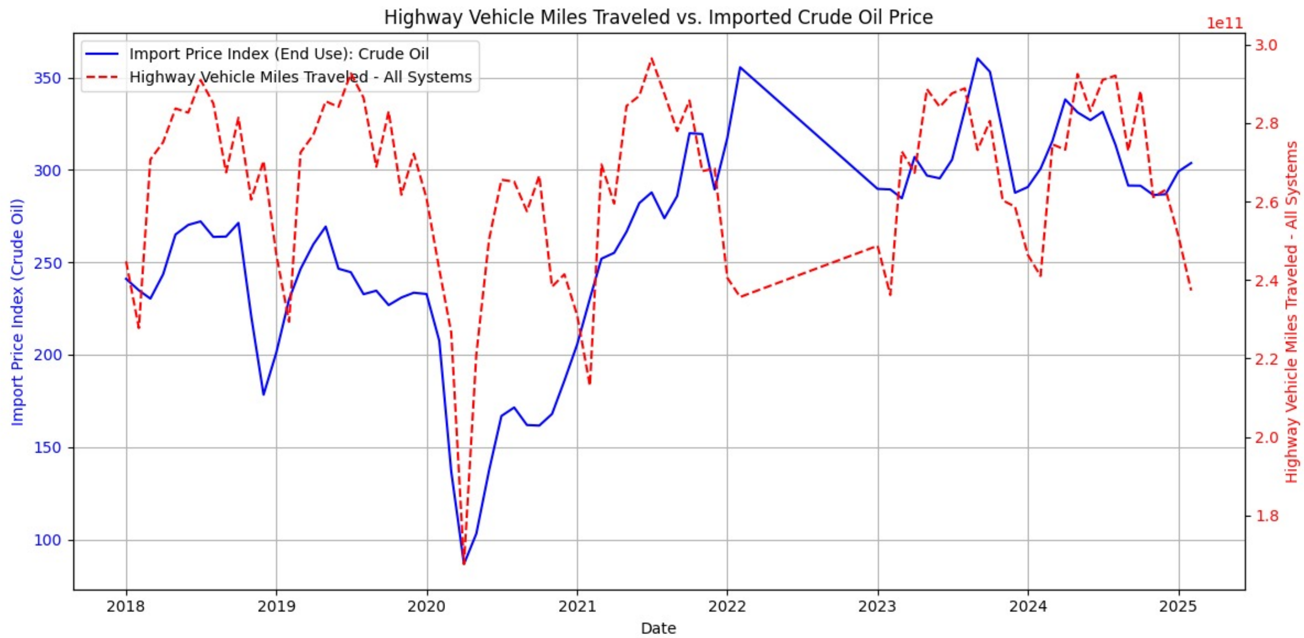

Figure 12: Highway Vehicle Miles Traveled and Imported Crude Oil Price.

Over the long term, VMT appears relatively stable, indicating that demand for highway transportation is inelastic or at least resilient to moderate economic fluctuations. However, a notable deviation occurs during 2022-2023, when crude oil import prices spiked sharply. This period coincides with a drop in VMT, suggesting that high and rising oil prices may have temporarily suppressed transportation demand, likely due to increased fuel costs affecting households and businesses. Interestingly, in the post-2023 period, VMT began to recover even though oil prices remained elevated - though no longer rising. This decoupling could imply adaptation behaviors such as improved fuel efficiency or cost absorption by consumers. Overall, while short-term shocks in oil prices can affect highway transportation volumes, the underlying demand remains structurally stable.

6 Conclusion

Over the past two decades, the United States has dramatically reduced its dependence on imported oil, driven by the shale revolution and policy initiatives, successfully transitioning into a net energy exporter and structurally enhancing its energy security. Nevertheless, oil remains a globally traded commodity and a critical input for the U.S. market, leaving persistent external risk exposures within supply chains and the consumer sector. Oil imports and global price volatility continue to pose potential threats to supply chains, industries, and households.

Oil price shocks are rapidly and powerfully transmitted through financial markets, triggering highly synchronized and short-lived but intense reactions across major energy stocks and related ETFs. These dynamics are further amplified in periods of geopolitical turmoil or policy interventions, as simulated in our counterfactual tariff scenario. At the societal level, rising oil prices exert lagged effects on transportation patterns and public safety, as reflected in the cyclical adjustment of traffic fatality rates—a relationship that is statistically significant but challenging to predict, as evidenced by the limited performance of machine learning models. For households, oil price changes are passed through to final consumer costs via complex, asymmetric mechanisms. The extent of pass-through is shaped by market stickiness, the diversity of energy sources, and threshold effects, with pronounced amplification during crisis periods.

The "missing tariff" is not a coincidence: taxing oil would amplify financial risk and transmit shocks throughout households and society, leading to potentially uncontrollable systemic disruptions—a dynamic demonstrated in our simulation. Enhanced U.S. energy security does not imply immunity to oil price risks. Policy design must recognize the dynamic nature of transmission mechanisms, the amplification effects of highly synchronized markets, and the delayed impacts on society and households. The limitations of available data and models remind us that single macro variables cannot fully explain or predict the breadth of social outcomes.

Future Work. Our analysis also highlights avenues for future research. To improve the explanatory power and practical relevance of predictive models, future work should incorporate higher-frequency, disaggregated, and behavioral data, as well as structural modeling of the mechanisms that link energy markets to financial and social outcomes. Additionally, more granular policy simulation and scenario analysis could help identify pathways to mitigate systemic risk and enhance societal resilience in the face of energy price volatility.

Acknowledgements

We gratefully acknowledge the organizers of the Women's Datathon 2025—Citadel, Citadel Securities, and Correlation One—for their generous support, inspiring mentorship, and for providing us with a unique platform to apply our analytical skills to real-world data challenges. This experience has greatly broadened our horizons and deepened our interest in data science.

References

[1] Energy Connects. (2025). Oil, natural gas and refined product imports exempted from new US tariffs. https://www.energyconnects.com/news/oil/2025/april/oil-natural-gas-and-refined-product-imports-exempted-from-new-us-tariffs/?utm_source=chatgpt.com (Accessed: 2025-05-18)

[2] International Energy Agency. (2023). *World Energy Balances 2023*. IEA. https://www.iea.org/reports/world-energy-balances-overview

[3] U.S. Energy Information Administration. (2020). *The Shale Revolution*. Today in Energy, EIA. https://www.eia.gov/todayinenergy/detail.php?id=42915

[4] U.S. Energy Information Administration. (2024). Refinery yield of finished motor gasoline. https://www.eia.gov/dnav/pet/pet_pnp_pct_dc_nus_pct_m.htm (Accessed: 2025-05-18)

[5] Ruhm, C. J. (2015). Provisional mortality data—United States, 2014: Gasoline prices and traffic fatalities. *American Journal of Public Health*, 105(8), 1370–1371. https://doi.org/10.2105/AJPH.2015.302579

[6] Kilian, L., & Park, C. (2009). The impact of oil price shocks on the U.S. stock market. *International Economic Review*, 50(4), 1267–1287. https://www.jstor.org/stable/40587575

[7] Burke, M., Heft-Neal, S., & Burney, J. (2015). Higher US gasoline prices and the geographic distribution of traffic fatalities. *American Journal of Public Health*, 105(8), 1670–1676. https://doi.org/10.2105/AJPH.2015.302670

[8] Grabowski, D. C., & Morrisey, M. A. (2004). Gasoline prices and motor vehicle fatalities. *Journal of Policy Analysis and Management*, 23(3), 575–593. https://doi.org/10.1002/pam.20023

[9] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). Springer. https://web.stanford.edu/~hastie/ElemStatLearn/

[10] Federal Reserve Bank of St. Louis. (2024). Interest Rate on 10-Year Treasury Constant Maturity (IR10000). FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/IR10000

[11] Federal Reserve Bank of St. Louis. (2024). Consumer Price Index for All Urban Consumers: Energy (CUSR0000SEHE). FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/CUSR0000SEHE