LUISS Guido Carli University

Degree Program in Management and Computer Science

Course of: Computational Finance

Machine Learning and Policy Shocks: Predicting Stock Market Reactions to the 2025 U.S. Steel and Aluminum Tariffs

Supervisor: Prof. Nicola Borri

Candidate: Santiago Cuesta

ID Number: 289581

Academic Year: 2024/2025

Contents

Chapter 1

Introduction

1.1 Motivation

The usage of machine learning techniques and artificial intelligence (AI) in financial decision-making has enhanced the accuracy of predictions for projecting stock prices leading to more accurate, data driven decisions than ever before. For instance, a study published in *Expert Systems with Applications* discusses various machine learning methods applied to stock market prediction, highlighting their effectiveness in capturing complex patterns in financial data.[1]

On the other hand, traditional financial models, such as the Capital Asset Pricing Model (CAPM) and the Fama-French three-factor model, have long been utilized to estimate expected returns based on dependent variables like market risk premiums and others. Studies have shown that these statistical models often exhibit relatively low R^2 values, suggesting they explain only a small proportion of the variation in stock returns.[2]

To address this limitation, recent studies have adopted more flexible approaches, such as machine learning-inspired workflows that incorporate panel regression models trained on historical data. By combining the structure and interpretability of panel data methods with the predictive mindset of training and testing splits, this approach seeks to balance explanatory power and forecasting performance. While it maintains interpretability through fixed effects and interaction terms—such as those capturing event-day impacts—it also leverages predictive validation to ensure robustness. Ultimately, model selection remains context-dependent: the "best" model depends not on universal superiority but on alignment with the research objective, whether interpretability, predictive power, or causal inference.

[1] Li, J., Liang, Y., Wang, J., & Zhou, X. (2022). A survey of multi-view clustering: Current techniques, emerging challenges and future directions

[2] Tang, Z. (2024). Risk-return analysis of equity portfolios: Comparison between CAPM and Fama-French three-factor model

Motivated by the adage 'history repeats itself,' this analysis aims to examine whether this notion holds true in the context of financial markets. Empirical studies have provided evidence supporting this concept. For instance, research published in the *Journal of Financial Markets* examined the validity of stock index models over different time periods and found patterns suggesting that, in the context of share trading, "history repeats itself".[3] In the case of this study, the impact of a hyper specific event will be examined, US government's steel and aluminum tariffs in 2018, to predict the stock prices of companies that suffered heavy losses or profits due to the implementation of these since the same tariffs have been announced once again in 2025.

1.2 Objective

The primary objective of this thesis is to develop a predictive model integrating machine learning techniques to forecast the impact of the newly implemented 2025 U.S. steel and aluminum tariffs on stock returns. By using historical stock market responses from the 2018 tariffs enacted under President Donald Trump, this research seeks to identify complex market patterns through the usage of advanced machine learning techniques.

The integration of these methods enables a deeper analysis of stock market behavior, a potential increase in variance explanation (accuracy of our predictions), and provides useful information for investors, policymakers, and businesses affected by these tariffs. The model's output will highlight the practical forecasting potential of machine learning–enhanced methods, underscoring their ability to predict market responses to policy events with improved accuracy.

[3] Michailidis, G., & de Andrade e Silva, M. (2004). The effect of beta estimation on the performance of the CAPM: An empirical study.

Chapter 2

Historical Context: The 2018 U.S. Steel and Aluminum Tariffs

During his first term in 2018, American President Donald Trump decided to impose tariffs on all steel and aluminum imports using national security concerns as a justification. These tariffs were announced in March and the order was signed on the 8th of said month. The precise fees of the tariffs were 25% on steel imports and 10% on aluminum. They were effective fifteen days after the order, on March 23rd, but initially excluded key allies like the European Union, Canada, and Mexico. Two months later, on May 29th, President Trump announced via Twitter (now X) that the tariffs were being extended to the European Union, Canada, and Mexico effective the next day.

The rationale behind the tariff implementation was to protect American industries. During this period, domestic steel and aluminum production were declining, while there was a surge in steel and aluminum imports. Under Section 232, the President may act if imports are found to threaten national security, and an investigation led by the Department of Commerce classified the increase in imports a national security threat.[4] There had been a significant decline in domestic aluminum and steel production, with steel industry employment falling 35% over the past two decades and the aluminum industry losing 65% of its workforce in just five years.[5] Commerce Department officials argued that overcapacity and unfair trade practices were compromising the U.S. metal industries.

Even though China accounted for only a small share of American steel imports, its subsidized excess production was seen as a major factor in driving down global prices. The administration aimed to revitalize U.S. manufacturing by bolstering American metal producers. He stated that the U.S. was being "ripped off" by trading partners on steel. In short, the tariffs were introduced to "stand up" to what the administration called

[4] U.S. Government Accountability Office (2020). Steel and aluminum tariffs: Commerce should improve its exclusion request process and economic impact reviews

[5] Durante, A. (2022). How Section 232 tariffs harmed the economy

unfair steel and aluminum trade to protect U.S. jobs and industrial capacity considered important for economic and defense security.

2.1 Economic and Trade Impact

The economic impact of the tariffs was mixed. On one hand, they provided relief to U.S. metal producers in the form of higher prices and a spike in utilization. A U.S. government report showed improvements in domestic steel output and employment, suggesting the tariffs achieved some immediate objectives for those industries.[4]

On the other hand, many independent studies concluded that the broader impact on the American economy was negative. Local manufacturers that relied on steel and aluminum faced higher costs, which led to increased consumer prices. The Federal Reserve found that by mid-2019, U.S. manufacturing employment was 0.6% lower than it would have been without the metal tariffs, as downstream factories cut jobs due to costlier inputs.[1] One analysis equated this to about 75,000 fewer U.S. manufacturing jobs attributable to the 2018 steel and aluminum tariffs, not counting losses from foreign retaliation. In effect, any jobs saved in steel mills came at the expense of jobs in factories that use steel—according to the Peterson Institute, this translated to a $650,000 cost to consumers for every steel job saved.[5]

This policy had severe repercussions not only domestically but also on the global trade market. One of the most significant repercussions for the U.S., was the weakened relationship with key allies like Mexico, Canada and the European Union. These countries were trading on a large scale with America leaving them completely vulnerable to the tariffs. As a countermeasure, they imposed retaliatory tariffs on American exports, not limited to only steel but ranging from farm produce to motorcycles and alcohol. This led to a broader international trade war, creating uncertainty and heavily affecting businesses and consumers. The tariffs themselves accounted for approximately 4% of the total value of U.S. imports.[6] On the other hand, the indirect effects of said policy were widespread. Exporters were losing foreign sales due to the retaliative measures by key allies and supply chains were disrupted as firms sought to avoid the import taxes. By the final quarter of 2019, U.S. trade data revealed that prices for imported steel had grown substantially, while there was a simultaneous decline in import volumes and American export growth.[5] As evidenced by the previous studies, the tariffs were not all good for the United States considering that they led to the creation of trade barriers, increased costs, tension with key allies without significantly reducing the trade gap. As a matter of fact, the trade deficit in goods reached an all-time high in the following years despite these measures.

[6] Kelly, K. (2020, October 14) Trump steel tariffs bring job losses to swing state Michigan

2.2 Impact on Businesses and Industries

Because of these tariffs, metal-using industries took a financial hit. For example, major us automakers, north American steel exporting companies (Canadian and Mexican) among many others. General Motors and Ford each reported about 1 billion USD in additional steel costs in 2018 due to the tariffs.[6] Ford's CFO stated that the rising metal prices were a major obstacle to profitability and both Ford and GM warned that the tariffs were eroding their profit margins.1 In 2020, industry analysts pointed out that all three Detroit automakers (Ford, GM, Stellantis) closed at least one US plant after the tariffs, partly due to the increased costs and global retaliation impacting exports.

Downstream in the supply chain, smaller businesses that were manufacturing or dependent on metal struggled to handle the sudden increase of costs. For example, a Michigan auto parts supplier lost several contracts worth millions of dollars because competitors in Canada and Mexico could source steel in a cheaper manner. Some U.S. manufacturers even considered moving production overseas to escape the inflated domestic input costs, an ironic outcome for a policy intended to reshore industry.[7]

As mentioned before, these tariffs did bring an initial benefit to the US Steel Corporation by bringing higher market prices, increased orders and a slight increase in employment within steel mills. However, these gains proved fragile because by 2019 the prices fell, and the tariff 'benefits' were null. Focusing on the aluminum side, the tariffs encouraged a few restarts of smelters, but the effect was limited – primary aluminum production rose modestly, while makers of aluminum products (cans, auto parts, etc.) saw higher input costs and mixed results.[4]

Overall, while metal producers experienced short-term gains, many downstream companies faced rising costs and decreased competitiveness, putting pressure on profit margins throughout the manufacturing supply chain.

2.3 Later Evaluations and Overall Assessment

Multiple post 2018 analyses concluded that the steel and aluminum tariffs had a net negative impact on the US economy. A 2020 report by Congress's research arm summarized that most studies found a small but negative effect on the US GDP as a result of the tariffs. These studies highlighted that the American people and companies were the ones bearing nearly the entire cost of the import taxes through an increase in prices of products that had these metals. It could be said that these tariffs were a form of a tax made by the US government for US firms that relied on foreign materials. Other studies led by the Federal Reserve and renowned economists found no meaningful boost to overall man-

[7] The White House. (2025, February) Fact sheet: President Donald J. Trump restores Section 232 tariffs

ufacturing employment due to the fact that the gains in protected metal industries were outweighed by larger losses in manufacturing companies that faced higher input costs or export retaliation. By one estimate, U.S. real income was reduced and about 75,000 manufacturing jobs were lost due to the steel/aluminum tariffs and related retaliation, even before the pandemic hit.[1]

That said, the tariffs were not all negative. In fact, some objectives were accomplished, like the renegotiation of preexisting agreements between the United States and key allies. It is argued that the tariffs were the driving factor for the USMCA deal with Mexico and Canada. In 2023, the American Commerce Department reported the effect the Section 232 tariffs had on Steel production. This policy had managed to strengthen the domestic steel production and capacity utilization, deeming the agency to consider these tariffs as largely successful.[4] On the other hand, the aluminum industry's outcome was less clear-cut, but from a broader economic perspective the costs appeared to outweigh the benefits. The U.S. International Trade Commission and several independent economists lead analyses that led to similar conclusions, which stated that the 2018 tariffs on steel and aluminum had raised input costs, disrupted supply chains, harmed downstream exports and slowed overall economic growth. To better grasp the magnitude of the impact, steel users paid an additional \$5.6 billion for materials in 2018, while producers had gained roughly \$2.4 billion. There's a clear discrepancy which further proves the inefficiency of the policy and the high cost per job saved.

All in all, though President Trump's 2018 tariffs managed to protect a few thousand jobs in metal production, they inadvertently harmed American consumers and businesses via an increase in prices and job losses in other sectors. Subsequent evaluations concluded that the tariffs modestly slowed economic growth and disrupted trade, without resolving the core issue of global overcapacity in steel production.[5]

Chapter 3

Economic Repercussions of the 2025 Tariffs: A Machine Learning Perspective

3.1 Reintroduction of Steel and Aluminum Tariffs in 2025

In February 2025, re-elected President Donald Trump decided to reinstate “True” 25% tariffs on steel and aluminum imports but this time to the whole world with no exemptions. During his last term, he applied 25% steel and 10% aluminum on imports, but a year after the implementation some of America’s key allies were exempt from these. This created a loophole that the Chinese government and others with excess steel and aluminum capacity took advantage of, and caused the tariffs to be easily, which led to a significant decrease in the effectiveness of this policy. [7]

As stated before, the 2018 tariffs outcome was not clear-cut, while some parties claimed it had revitalized the American steel industry, others argued that it stumbled with growth and led to an increase in price for both consumers and businesses. The new section 232 tariffs have upped the rate for aluminum to 25%, claiming that ”the initial tariff of 10 percent ad valorem is not high enough to address the threatened impairment to our national security posed by aluminum imports”. [8] The American Government seems adamant in claiming that the “True” 2025 tariffs will “protect America’s critical steel and aluminum industries, which have been harmed by unfair trade practices and global excess capacity”.[7] Similarly, the official justification for the reapplication of the tariffs is a national security concern. President Trump seeks to abolish unfair trade practices and the global dumping of steel and aluminum.

[8] The White House. (2025, February) Adjusting imports of aluminum into the United States

The 25% tariffs on steel and aluminum imports, announced by President Donald Trump, took effect on March 12, 2025. As of April 2025, investors and businesses are filled with uncertainty, and worried about an increase in costs, supply chain disruption and international retaliation. Precisely due to this uncertainty lies the reason behind the prediction of stock price reaction, considering that by using machine learning techniques, some variance can be captured which provides useful information to all affected parties. Rather than aiming to forecast long-term price movements, the model is designed to assess how stock returns evolved in the days immediately following the event. By capturing these short-term patterns through a panel regression framework, the analysis provides insights into market sensitivity and firm-specific reactions—helping stakeholders better understand the immediate financial impact of such policy shocks.

3.2 Selection of Companies

To build a stock price prediction model, it is necessary to have chosen stocks with meaningful data. The companies chosen for this analysis were:

- **Grupo Simec, S.A.B de C.V ("SIM")**: A Mexican Steel producer that manufactures special bar quality steel and structural steel. Its products are primarily used in construction, automotive and industrial sectors. The reasoning behind this choice was that this company is a non-American steel company with significant exports to the US making it useful to understand the impact of tariffs.

- **Hammond Power Solutions ("HPS-A.TO")**: A Canadian manufacturer of dry-type transformers and related electrical equipment, with a significant portion of its revenue coming from the U.S. market. While this company does not directly sell steel and aluminum, it's still vital to see the impact of tariffs along supply chains considering this company needs steel and aluminum for their products.

- **Tesla ("TSLA")**: An American car manufacturer that relied on global suppliers for raw materials and components like steel and aluminum. During 2018, Tesla was scaling production and couldn't absorb cost increases easily which is why this company was heavily affected by the tariffs considering they elevated both domestic and foreign steel and aluminum costs. Similarly, two other American car manufacturers: General Motors ("GM") and Ford ("F") who also faced similar problems considering steel and aluminum are key components for vehicle manufacturing were chosen.

- **Campbell Soup Company ("CPB")**: An American consumer goods firm that specializes in selling canned goods, like soups, drinks and other shelf stable products. This company was chosen because it was very sensitive to tariffs considering that tariffs increased metal and aluminum costs even for domestically sourced inputs due to the reduced competition.

- **Alcoa Corporation ("AA")**: A major American producer of aluminum, bauxite and alumina. Alcoa operates mines, smelters, and refineries globally. Being an American producer, it was an essential pick, considering the added value of having the stock price reaction of a company which was supposed to benefit from the tariffs.

- **Caterpillar ("CAT")**: A leading U.S. manufacturer of construction and mining equipment such as bulldozers, excavators, etc. All these machines require large amounts of steel and aluminum which led to a direct price increase in caterpillar costs. Their input costs were raised due to the tariffs which also led to a decrease in their profit margins. This pick was also key to further understanding the impact of tariffs on stocks.

3.3 Initial data Analysis

The data importing process began by selecting Yahoo Finance as the data source, due to its provision of real-time and reliable information on global stock markets. The chosen time window spans from 2017-02-20 (one year prior to the tariff announcement) to 2018-07-01 (six months following the announcement). In addition to individual stock prices, market return data and the risk-free rate were also collected, as both are essential inputs for the panel regression model. These variables are widely used in stock return forecasting and help capture broader market behavior. The risk-free rate was divided by 100 to convert it from percentage to decimal form, ensuring consistency with the format of return calculations.

Listing 3.1: Data Collection

```
# Your stock tickers
tickers = ["SIM", 'HPS-A.TO', 'TSLA', 'CPB', 'GM', 'AA', 'F', 'CAT']

# Market and risk-free tickers
market_ticker = "^GSPC"
risk_free_ticker = "^IRX"

# Download data
price_data = yf.download(tickers, start="2017-02-20", end="2018-07-01")["Close"]
market_data = yf.download(market_ticker, start="2017-02-20", end="2018-07-01")["Close"]
risk_free = yf.download(risk_free_ticker, start="2017-02-20", end="2018-07-01")["Close"] / 100
```

Subsequent data preprocessing steps included forward-filling missing values on business days to eliminate gaps and ensure a consistent number of price observations across all stocks. Log returns were then computed, as they are additive over time and thus more suitable for statistical analysis and modeling compared to simple returns.

Date	AA	CAT	CPB	F	GM	HPS-A.TO	SIM	TSLA
2018-06-25	-0.0289	-0.0243	0.0899	-0.0130	-0.0156	0.0044	-0.0035	-0.0019
2018-06-26	0.0342	-0.0076	-0.0220	0.0017	0.0098	-0.0022	0.0217	0.0266
2018-06-27	-0.0013	-0.0069	-0.0350	-0.0087	-0.0157	-0.0698	-0.0091	0.0073
2018-06-28	0.0158	0.0105	0.0213	-0.0123	0.0037	0.0000	0.0488	0.0156
2018-06-29	0.0088	-0.0027	-0.0052	-0.0188	-0.0280	-0.0167	0.0278	-0.0201

Table 3.1: Preview of the log returns data frame: last five business days for selected stocks.

The result is a data frame of daily log returns for each stock, where each row represents a business day and each column corresponds to one of the selected tickers. These log returns are calculated by taking the natural logarithm of the price ratio between consecutive days. This transformation is standard in financial analysis, as log returns are time-additive and more suitable for statistical modeling.

Summary statistics were then applied to the data frame to gain a clearer understanding of its underlying characteristics prior to conducting the analysis. This step provided insight into key metrics such as mean returns, volatility (standard deviation), skewness, and other distributional properties relevant to financial modeling.

Ticker	Mean	Volatility	Min	Max	Skewness	Kurtosis
AA	0.000854	0.024237	-0.145192	0.093758	-0.6503	4.7946
CAT	0.001099	0.015383	-0.064023	0.075671	-0.2597	3.7905
CPB	-0.000905	0.017413	-0.132002	0.089879	-1.6405	14.1335
F	-0.000110	0.012401	-0.072817	0.049534	-0.8547	5.0161
GM	0.000351	0.015385	-0.040451	0.121096	1.2592	11.1868
HPS-A.TO	0.000964	0.015611	-0.069783	0.066249	0.0978	3.3796
SIM	-0.000991	0.019445	-0.072196	0.103752	0.4207	3.3866
TSLA	0.000638	0.025139	-0.090242	0.092987	-0.0980	1.3127

Table 3.2: Summary statistics of log returns for selected stocks.

Mean Return: This represents the average daily return for each stock over the sample period. While not indicative of long-term performance on its own, it gives a general sense of a stock's average direction. Positive means suggest a tendency for upward movement, while negative values (like those for CPB, F and SIM) indicate average declines during the period analyzed.

Volatility (Standard Deviation): Volatility tells us how much a stock's returns fluctuate each day, commonly used as a risk measure. Stocks like TSLA and AA had higher volatility, meaning their prices moved around a lot. In contrast, F and CAT were more stable with lower volatility.

Minimum Return: This shows the worst daily return each stock had. AA had the largest single-day drop at around -14.5%, and CPB also had a particularly bad day with a loss over -13%.

Maximum Return: The best daily return in the dataset. For example, GM had a strong up day with over 12%, and CPB also had a big gain around 9%.

Skewness: Skewness tells us whether big gains or big losses are more common. CPB, F, and AA had negative skewness, which means they were more likely to have big losses than big gains. GM and SIM, on the other hand, had positive skewness, leaning more toward big gains.

Kurtosis: This shows how often extreme values (really big moves) happen. CPB and GM had very high kurtosis, meaning they had more wild swings compared to others. TSLA, interestingly, had the lowest kurtosis, meaning its returns were more evenly distributed and closer to normal.

To understand how the stocks responded to key events during the 2018 steel and aluminum tariffs, a graphical visualization was necessary. Before doing so, the log returns were converted into cumulative returns to allow for a time series representation of performance over the event window.

The key event dates were selected based on the assumption that these announcements and policy actions were likely to have a significant impact on the stock market.

1. **Tariff Announcement**: 2018-03-01
2. **Tariff Order Signed**: 2018-03-08
3. **Tariff Effective Date**: 2018-03-23
4. **Tariff Announcement for Canada and Mexico**: 2018-05-31
5. **Canada and Mexico Tariff Effective**: 2018-06-01

```
# Compute cumulative returns
cumulative_returns = (log_returns + 1).cumprod()

# Key dates for tariff-related events
event_info = {
    "2018-03-01": "Tariff Announcement",
    "2018-03-08": "Tariff order signed",
    "2018-03-23": "Tariffs Effective",
```

```
    "2018-05-31": "Announcement Canada/Mexico Tariff",
    "2018-06-01": "Canada/Mexico Tariff Effective"
}
```

The cumulative returns of the selected stocks were then plotted over time to visualize how their performance evolved in relation to the key tariff-related events.

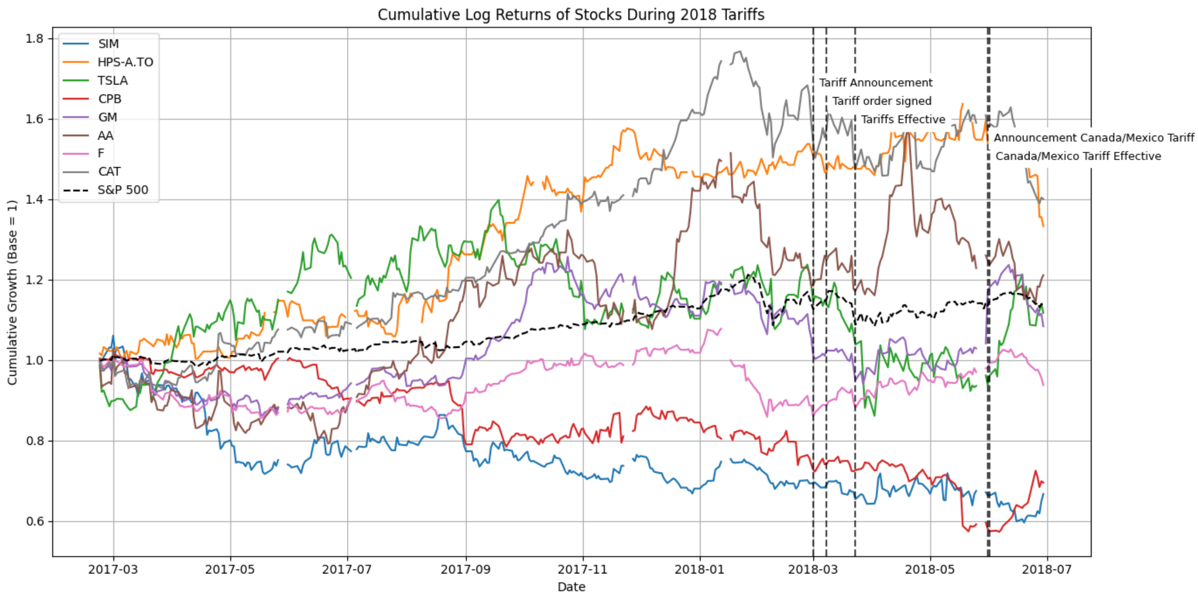


Figure 3.1: Cumulative log returns of selected stocks during the 2018 steel and aluminum tariffs. Key policy dates are marked with vertical lines.

As evidenced by the figure, after the key events, there's a clear downward trend in many of the stocks, which makes sense since most of the companies chosen were negatively affected by the tariffs. However, AA moves in the opposite direction. This is probably because it's a U.S.-based aluminum producer and initially benefited from the price increases caused by the tariffs. The S&P 500 was also included in the graph to serve as a benchmark. It stays relatively stable throughout the period, which shows that the impact of these tariffs was more specific to certain industries rather than the market as a whole.

3.4 Statistical Significance of Event Impacts

To conduct the panel regression model, it is first necessary to establish that the event in question had a statistically significant impact. In this context, statistical significance implies that the event caused an abnormal deviation from the typical distribution of stock returns at the time it occurred. In other words, the event led to a return pattern that differs meaningfully from what would be expected under normal market conditions. To test for this, the CAPM model will first be used to estimate the beta coefficients for each stock in relation to the market, providing a baseline for identifying abnormal returns.

"The Capital Asset Pricing Model (CAPM) is an equilibrium model that underlies all modern financial theory. It describes the relationship between expected return and risk in a market where investors are price takers, have homogeneous expectations, and can borrow and lend at the risk-free rate." [9] This model helps investors understand the trade-off between risk and returns, indicating that the expected return on the asset is proportional to its beta.

The CAPM formula is expressed as:

$$\mathbb{E}(R_i) = R_f + \beta_i \left[\mathbb{E}(R_m) - R_f\right] \tag{3.1}$$

Where:

- $\mathbb{E}(R_i)$ is the expected return on the asset
- R_f is the risk-free rate
- β_i is the beta of the asset
- $\mathbb{E}(R_m)$ is the expected return of the market portfolio

The CAPM beta quantifies an asset's sensitivity to market movements. A beta of 1 implies that the asset moves in line with the overall market; a beta greater than 1 indicates the asset is more volatile than the market, while a beta less than 1 suggests it is less sensitive to market fluctuations, representing lower systematic risk.[10]

In this project, the CAPM model was applied to estimate the beta coefficients of the selected stocks using historical return data. This step serves as the foundation for identifying abnormal returns around the event window. To estimate each stock's beta, log returns were used, market log returns, and the risk-free rate, all sourced from Yahoo Finance. The estimation period covered one year of historical data, ending on February 28, 2018 — the day before the first tariff-related event. Using this data, the CAPM formula was applied, and an ordinary least squares (OLS) regression was performed to calculate the beta coefficients.

[9] Bodie, Z., Kane, A., & Marcus, A. J. (2014) Investments (10th ed.)
[10] Kenton, W. (2025, April 29) What beta means for investors

Listing 3.2: CAPM beta calculation

```
Estimate Beta for Each Stock
betas = {}

for stock in log_returns_estimation.columns:
    excess_stock_return = log_returns_estimation[stock] - risk_free_estimation.squeeze()
    excess_market_return = market_returns_estimation.squeeze() - risk_free_estimation.squeeze()

    # Combine into one DataFrame and drop NaNs
    combined = pd.concat([excess_stock_return, excess_market_return], axis=1).dropna()
    if len(combined) < 10:
         print(f"Not enough data to estimate beta for {stock}. Skipping.")
         continue

    y = combined.iloc[:, 0]  # Excess stock return
    X = sm.add_constant(combined.iloc[:, 1])  # Excess market return with intercept

    model = sm.OLS(y, X).fit()
    betas[stock] = model.params[1]  # beta is the coefficient on market excess return

# 4. Save Betas into a DataFrame
betas_df = pd.DataFrame.from_dict(betas, orient='index', columns=['Beta'])

# 5. Print Betas
print("\nEstimated Betas:")
print(betas_df)
```

The resulting betas were:

Ticker	Beta
AA	1.324698
CAT	1.276293
CPB	0.583307
F	0.987277
GM	1.121125
HPS-A.TO	0.324094
SIM	0.371896
TSLA	1.199716

The projected returns of the stocks under "normal" market conditions were calculated using the estimated beta values. This allows us to compare the actual observed returns with the expected returns and identify any abnormal returns (AR) that may have occurred around the event. The presence of abnormal returns would suggest that the event had a measurable impact on the market. To quantify this impact, abnormal returns are aggregated over an event window spanning four days before and four days after the event date. These are referred to as cumulative abnormal returns (CARs). The next step was to calculate the CAR for each of the five key tariff-related events that occurred in 2018.

These were the resulting CARs for each of the chosen stocks during each of the events:

Date	AA	CAT	CPB	F	GM	HPS-A.TO	SIM	TSLA
2018-03-01	0.021465	-0.032997	-0.048409	-0.002992	-0.058119	-0.067548	-0.081788	-0.023679
2018-03-08	0.047585	0.038355	-0.023842	0.004311	-0.004934	-0.104889	-0.070158	0.009939
2018-03-23	0.057244	0.037072	-0.040338	0.020670	-0.011564	-0.067517	-0.062436	-0.054339
2018-05-31	0.020736	-0.013416	-0.069735	0.004228	0.143589	-0.077517	-0.061091	0.034507
2018-06-01	0.046209	-0.005716	-0.078975	0.010929	0.136620	-0.096838	-0.109790	0.018378

With the CAR values calculated, the test can now be performed to see whether the events had a statistically significant impact on the selected stocks. This step is essential in determining the relevance of the research. If the events are found to have no significant effect, then using them as the basis for return prediction would not be meaningful. To evaluate this, a one-sample t-test was performed to assess whether the events led to abnormal returns that are statistically different from zero.

Listing 3.3: T-Test on CARs per Event

```
from scipy.stats import ttest_1samp

# Initialize t-test results
t_test_results = {}

for event_date, CARs in CAR_df.iterrows():
    # Perform t-test: Null Hypothesis = CARs have mean 0
    t_stat, p_value = ttest_1samp(CARs.dropna(), 0)
    t_test_results[event_date] = {'t-statistic': t_stat, 'p-value
        ': p_value}

# Convert to DataFrame
t_test_results_df = pd.DataFrame(t_test_results).T
print("\nT-Test Results for CARs per Event:")
print(t_test_results_df)
```

The Output table of our T-test was:

Date	t-statistic	p-value
2018-03-01	-3.022413	0.019320
2018-03-08	-0.702772	0.504891
2018-03-23	-0.885018	0.405519
2018-05-31	-0.090937	0.930090
2018-06-01	-0.337460	0.745663

The results demonstrate that only one of the five important event dates had a statistically significant impact on stock returns. This was March 1, 2018, which was expected given that it was the event that took the entire world by surprise. The t-statistic is -3, indicating a very strong impact on stock returns. This suggests that the market was indeed taken by surprise and reacted negatively. All of the other events had no substantial impact on the companies' returns, which may be attributed to the market having already adjusted to the tariff effects by the time these occurred, highlighting the efficiency of modern financial markets.

Having established that the initial announcement constituted a genuine market shock, the analysis can now proceed to its core objective: predicting the impact of the 2025 tariffs using this historical event as a reference point. The next chapter introduces a panel regression framework that models daily stock returns as a function of market behavior, company characteristics, and interaction with the 2018 event shock. This will allow us to quantify the expected effect of similar policy events on today's market.

Chapter 4

Panel Regression Model for Predicting Stock Returns

First of all, the analysis begins by defining what a panel regression is. Panel regression is "a statistical model that analyzes data containing observations over multiple time periods for the same entities (such as individuals, firms, or countries). This type of model allows researchers to control for variables that change over time and across entities, enabling more accurate modeling of dynamic behaviors and causal relationships". [11]

There are several types of Panel Regression models, common approaches include fixed effects, random effects, and pooled ordinary least squares. Each of these methods makes different assumptions about unobserved heterogeneity—namely, the unmeasured factors that vary between entities and over time.

In this research, the chosen panel regression model was Pooled Ordinary Least Squares (Pooled OLS), which was used to estimate the effect of trade policy announcements, specifically U.S. steel and aluminum tariffs, on stock returns. A machine learning technique called Train-Test split will be put to use.

The train-test split is a fundamental method in machine learning used to evaluate how well predictive models perform. It works by dividing the available dataset into two separate parts:

- **Training Set:** This is the portion of data used to teach the machine learning model. It helps the model identify patterns and relationships within the data. In this paper, the training set consists of stock reactions to the 2018 tariff announcements.

- **Testing Set:** This data is set aside to evaluate how well the model performs on new unseen information. It gives a more realistic sense of the predictive accuracy of the model. Here, the test set includes stock data from 2025 leading up to

[11] Wooldridge, J. M. (2010) Econometric Analysis of Cross-Section and Panel Data (2nd ed.)

the new tariff announcement, allowing the model to generate a forward-looking prediction.

By separating the data in this manner, researchers can grasp how the model will perform on new real-world data. This method helps prevent overfitting,a major issue when dealing with predictive models, which happens when a model becomes hyper-fixated on training data and fails to perform well on new data. By encouraging the model to generalize, it leads to more accurate and reliable predictions.[12]

4.1 Training Set

As briefly mentioned earlier, the training set for this research consists of the data frame containing the 2018 tariff reaction for the selected stocks. The decision was made to include only three predictors in the model: rolling 20-day volatility, market returns, and an event dummy variable. The dependent variable, the one to be predicted, is stock return. Several operations were performed on the raw data to structure it into a panel format, as shown in the example below:

Index	Date	Ticker	Return	MarketReturn	Volatility	EventDummy
152	2017-03-21	AA	-0.049335	-0.012486	0.033912	0
160	2017-03-22	AA	0.014049	0.001888	0.034178	0
168	2017-03-23	AA	-0.011343	-0.001061	0.031192	0
176	2017-03-24	AA	-0.022773	-0.000844	0.031508	0
184	2017-03-27	AA	-0.002768	-0.001020	0.031392	0

Table 4.1: Stock Return Train Data

These are the predictors of the model were and what they are:

- **Market Returns:** These are the returns of the market (S&P500) aligned to the dates of the 2018 steel and aluminum tariffs. Market returns are often a significant predictor for predictive stock return models.
- **Rolling 20 day volatility:** This variable is a measure of how much stocks returns move over a 20 day window. It helps track changes in risk or uncertainty over time.
- **Event Dummy:** This is a binary variable equal to 1 during the event window surrounding the March 1, 2018 tariff announcement (from 4 business days before to 4 business days after), and 0 otherwise. It is used to isolate the abnormal return effect attributable to the policy shock.

[12] Geron, A. (2019). Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow (2nd ed.)

These predictors were chosen to help the model capture the effects of past tariff announcements and to better understand how markets respond to policy-related events.

After the preparation of the data, the next step was to run the panel regression (Pooled OLS). This was the model:

Listing 4.1: Training Model

```
import statsmodels.formula.api as smf

# Run the OLS regression
model = smf.ols('Return ~ MarketReturn + Volatility + EventDummy'
    , data=returns_long).fit(cov_type='HC1')  # HC1 = robust
    standard errors

# Print summary
print(model.summary())
```

Dependent Variable:	Return
Model:	OLS
Method:	Least Squares
Observations:	2,590
Df Residuals:	2,586
Df Model:	3
R-squared:	0.118
Adj. R-squared:	0.117
F-statistic:	89.03
Prob (F-statistic):	7.97e-55
Log-Likelihood:	6808.9
AIC:	-13,610
BIC:	-13,590
Covariance Type:	HC1

Table 4.2: Model Summary Statistics

Variable	**Coef.**	**Std. Err.**	**z**	**P> $|z|$**	**[0.025, 0.975]**
Intercept	0.0004	0.001	0.360	0.719	[-0.002, 0.002]
MarketReturn	**0.8769**	**0.055**	**15.929**	**0.000**	**[0.769, 0.985]**
Volatility	-0.0193	0.067	-0.289	0.772	[-0.150, 0.111]
EventDummy	**-0.0045**	**0.002**	**-2.254**	**0.024**	**[-0.008, -0.001]**

Notes: Standard errors are heteroscedasticity-robust (HC1). Variables in bold are statistically significant at the 5% level.

Diagnostic Tests:
Omnibus: 383.302 Prob(Omnibus): 0.000
Jarque-Bera (JB): 5459.669 Prob(JB): 0.000
Skew: -0.140 Kurtosis: 10.107
Durbin-Watson: 2.006 Cond. No.: 141

After running our model, the first metric to look at should be the general fit which is indicated by the $R^2 = 0.118$ in the output. R^2 is "the proportion of the total variation in the response variable that is explained by the regression model. It is calculated as the ratio of the regression sum of squares to the total sum of squares."[13]. This R^2 of 11.8% means that our model explains only 11.8% of the variation in stock returns, which might seem low at first. However, daily returns are notoriously noisy and influenced by many idiosyncratic shocks, so when this is taking that into account, the model actually performed quite well overall. As evidenced by the output our model is also highly statistically significant with a f statistic of 89.03 and its p value of ¡ 0.0001.

The next step involves a closer examination of the model variables. The Market Return variable was highly significant with a coefficient of 0.87 and a p value of ¡ 0.0001 meaning that our model correctly captures systematic influence so for each 1% change in the market, the average stock return changes by .88%. the Event dummy variable was significant as well it had a coefficient of -0.0045 and a p value of 0.024. This suggests that the event had a measurable impact on stock returns. Specifically, during the nine-day event window (four days before, the day of the announcement, and four days after), average daily returns were 0.45% lower. On the other hand, our other predictor, rolling 20 day volatility was not significant with a p value of 0.772 which indicates that short term risk does not explain returns in this context and that the event's effect was not driven by elevated firm-specific uncertainty.

This marks the end of the training section, the next part will be solely focused on the prediction (Test).

[13] Draper, N. R., & Smith, H. (1998). Applied Regression Analysis (3rd ed.)

4.2 Test Set

In order to start the prediction, The first step is to download the data. This is how the operation was performed using yahoo finance:

Listing 4.2: Test Data Download

```
# Your stock tickers
tickers = ["SIM", 'HPS-A.TO', 'TSLA', 'CPB', 'GM', 'AA', 'F', '
    CAT']
market_ticker = "^GSPC"

# Define download period: January 1 to March 1, 2025
start_date = "2024-10-01"
end_date = "2025-03-01"

# Download stock prices
price_data_2025 = yf.download(tickers, start=start_date, end=
    end_date)["Close"]

# Download market index prices
market_data_2025 = yf.download(market_ticker, start=start_date,
    end=end_date)["Close"]

# Save to CSV
price_data_2025.to_csv("price_data_2025.csv")
market_data_2025.to_csv("market_data_2025.csv")

print("Data successfully downloaded and saved to CSV.")
```

Following this step, several operations identical to those applied to the training data were performed to format the dataset into a panel structure suitable for model implementation. After these transformations, the resulting data frame appeared as follows:

Index	Date	Ticker	Return	MarketReturn	Volatility	EventDummy
150	2024-10-29	AA	-0.008509	0.001613	0.023498	0
158	2024-10-30	AA	-0.009814	-0.003306	0.023650	0
166	2024-10-31	AA	-0.011655	-0.018790	0.022978	0
174	2024-11-01	AA	0.011902	0.004084	0.022509	0
182	2024-11-04	AA	0.003936	-0.002816	0.022455	0

Table 4.3: Stock Return Test Data

The next step was modeling:

Listing 4.3: Training Model

```
# Predict 2025 returns
returns_long_2025['PredictedReturn'] = model.predict(
    returns_long_2025[['MarketReturn', 'Volatility', 'EventDummy'
    ]])

# Show results
returns_long_2025[['Date', 'Ticker', 'Return', 'PredictedReturn',
     'MarketReturn', 'Volatility', 'EventDummy']].head()
returns_long_2025[returns_long_2025['EventDummy'] == 1]
```

Index	Date	Ticker	Return	MarketReturn	Volatility	EventDummy	PredictedReturn
682	2025-02-10	AA	0.021907	0.006690	0.028231	1	0.001200
690	2025-02-11	AA	0.006749	0.000340	0.028169	1	-0.004368
698	2025-02-12	AA	-0.022858	-0.002728	0.028569	1	-0.007066
706	2025-02-13	AA	0.001375	0.010372	0.028469	1	0.004424
683	2025-02-10	CAT	-0.001815	0.006690	0.019702	1	0.001364

Table 4.4: Predicted and Actual Returns for Selected Observations

By observing the data, the predicted returns differ from the actual returns by a substantial amount. However, this does not imply that the model was ineffective. It is important to first examine error metrics and plot the predicted values against the actual returns to assess whether the model was able to capture any meaningful patterns in the data.

These were the error metrics:

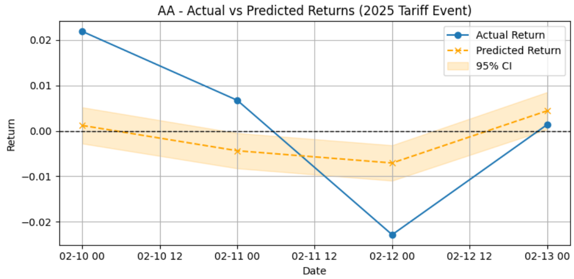
Metric	Value
Mean Prediction Error	-0.000862
Root Mean Squared Error (RMSE)	0.023969
Mean Absolute Error (MAE)	0.016559
Adjusted R-squared	0.052190

Table 4.5: Prediction Error Metrics

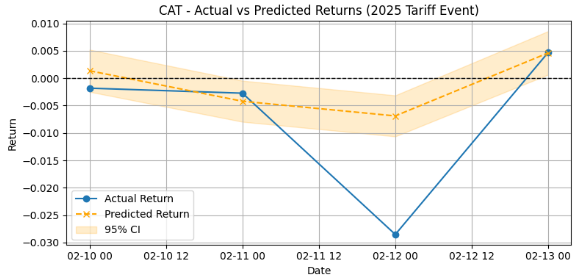
Index	Date	Ticker	Return	PredictedReturn	Error
682	2025-02-10	AA	0.021907	0.001200	0.020707
690	2025-02-11	AA	0.006749	-0.004368	0.011117
698	2025-02-12	AA	-0.022858	-0.007066	-0.015792
706	2025-02-13	AA	0.001375	0.004424	-0.003049
683	2025-02-10	CAT	-0.001815	0.001364	-0.003180

Table 4.6: Predicted vs. Actual Returns with Error

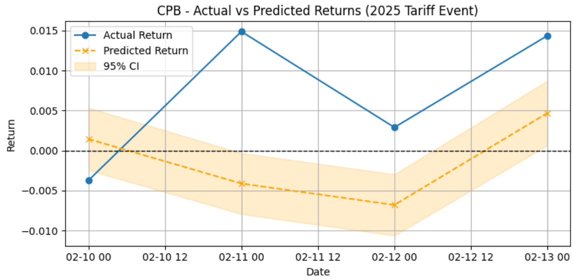
Let's analyze the model's error metrics. The mean prediction error indicates that on average this model slightly under predicts daily returns by 0.0862%. By interpreting this, the models prediction bias is minimal, it does not over estimate or under estimate returns. In the context of daily return prediction, this result is expected and acceptable. The mean absolute error which indicates the prediction's error, regardless of the direction, was 1.6% daily. This result is also fairly accurate and acceptable in daily returns considering all the volatility and noise in them. The root mean squared error was 2.3% per day is relatively modest for financial forecasting models indicating good accuracy. The fact that the RMSE ¿ MSE shows that there are large outliers present which is common in financial data. Finally, the $R^2 = 5.2\%$ suggests that the model captures a meaningful but partial share of the daily returns in the event window. All in all, the error metric were fairly common for daily stock predicting returns. It is time, to see graphically how the predictions matched against the actual returns.



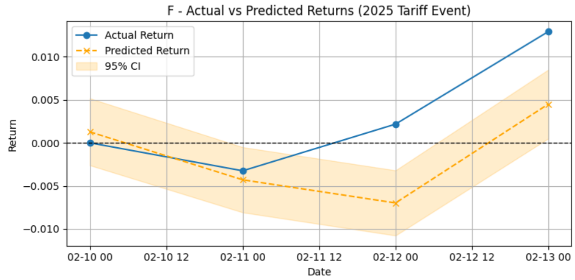
AA - Actual vs Predicted



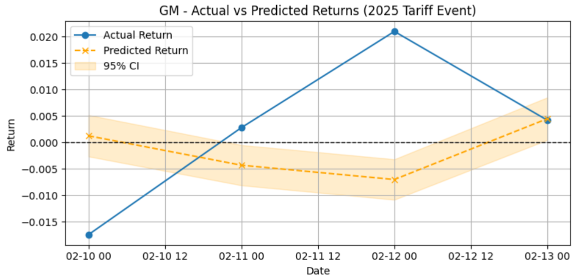
CAT - Actual vs Predicted



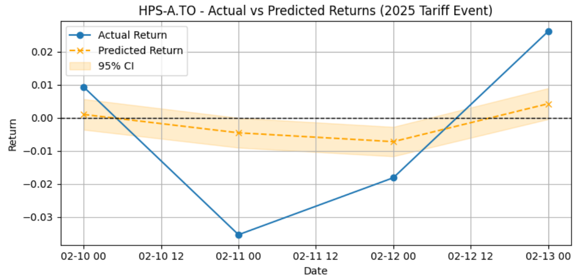
CPB - Actual vs Predicted



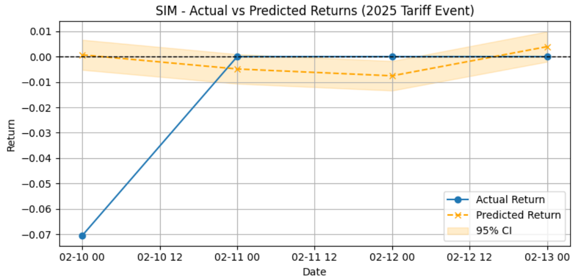
F - Actual vs Predicted



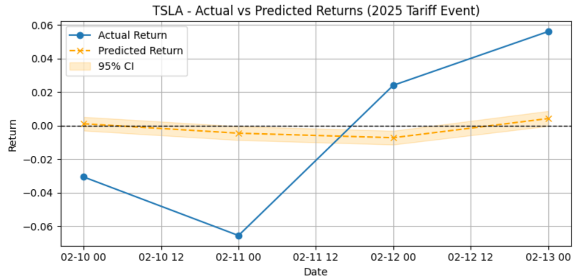
GM - Actual vs Predicted



HPS-A.TO - Actual vs Predicted



SIM - Actual vs Predicted



TSLA - Actual vs Predicted

Figure 4.1: Actual vs Predicted Returns for Selected Stocks (2025 Tariff Event)

At first glance, the model does not seem to predict returns very accurately, which is consistent with the relatively low R^2 value of 5.2%. However, when looking more closely, the model does manage to capture the general trend in returns for most of the stocks. For example, in the case of AA, CAT, F, and HPS.TO, the predicted and actual returns moved in the same direction. That said, the model struggled with the volatility of these movements—its predictions tended to move less sharply than the actual returns. For other companies like GM, TSLA, and SIM, the model failed to capture both the direction and the volatility of the returns, which may be explained by firm-specific developments in 2025 that the model could not account for. Another reason to perhaps explain this deviation is the fact that in some stocks perhaps the return materializes one or two days after the event. In any case, the panel regression ,with a fairly low variance capture of 5.2%, still delivered insightful information regarding patterns when dealing with policy shocks on stock returns that should be taking into account when assessing stock risk for example.

Chapter 5

Conclusion

5.1 Summary

The primary goal of this study was to examine the impact of the American steel and aluminum tariffs and determine whether similar events can be predicted based on past occurrences. This objective was achieved by using a model trained on the 2018 stock reactions to predict the returns of the newly implemented 2025 tariffs. The approach used in this study was a panel regression combined with the machine learning method, train-test split. This method proved effective, as it helped capture the impact of the 2018 tariffs and integrate this information into the model, allowing for the prediction of returns related to the 2025 tariffs.

The use of the panel regression model proved to be extremely valuable for this study, considering this type of model captured variations in returns across different firms and over time. This is due to its unique ability to include both cross sectional and time series data. As a result, the model choice had a crucial role in the outcome of this study and it helped estimate the impact of 2025 tariff affected returns with higher accuracy.

Similarly, the train test split was also essential, due to the fact that it allowed the panel model to be trained on the historical data (2018 tariff stock reactions) while reserving the 2025 data as a validation set. This approach not only helped to mitigate overfitting but it also increased the reliability of the results.

5.2 Key findings

As demonstrated by the rigorous testing conducted in this study, the 2018 tariff announcement was a statistically significant event that disrupted the normal pattern of returns associated with the 2018 tariffs. The event had a p-value of 0.002, indicating that it effectively captured variance in the stock returns during the event impact day range. Furthermore, it proved to be a reliable predictor, accurately capturing the trend in stock returns throughout the event window.

However, despite its ability to capture the direction of the returns, the panel regression model did not achieve high predictive accuracy with an R^2 of 11.8% in the training set and a 5.2% in the test. This outcome is understandable given that the objective was to predict daily returns, which are inherently noisy and characterized by high volatility. When considering this factor, the model performed well.

An important insight uncovered during the methodology was the model's ability to capture the trend in stock returns, which became evident through visualization plots. This finding highlights the importance of visualizing data, as error metrics or R^2 values alone may not fully convey the model's capacity to identify patterns or trends.

5.3 Result interpretation

The performance of the model indicates that while it is possible to predict the directional impact of trade policy shocks on stock returns, accurately forecasting exact daily return values remains a significant challenge. This difficulty arises from the inherent characteristics of daily returns, which are highly volatile and influenced by numerous firm-specific and external factors beyond the scope of the model. As Burton G. Malkiel (2011) articulates in A Random Walk Down Wall Street, "The random walk theory holds that the stock market is efficient and that prices reflect all known information. Therefore, future price movements are determined solely by information not yet known and are thus unpredictable." [14]

This perspective underscores the notion that daily stock returns are largely independent and unpredictable, rendering precise short-term forecasting exceedingly difficult.

The results of the study align with the notion that "history repeats itself," but not in exact proportion. While the model correctly identified whether the stock returns were moving up or down within a four-day range of the event, it did not accurately predict the precise return values and tended to underestimate the volatility. This outcome is understandable, given the inherent unpredictability of daily returns as discussed earlier.

[14] Malkiel, B. G. (2011). A random walk down Wall Street: The time-tested strategy for successful investing (11th ed.).

5.4 Limitations of analysis

For the purpose of inference, this model was intentionally kept relatively simple to facilitate an understanding of the interactions between the chosen variables and their impact on the predictions. The goal was to explore how past tariff reactions could help investors, companies and policy makers anticipate certain movements in the stock market. However, this simplification also introduced some limitations.

The assumption that macroeconomic conditions in 2018 and 2025 were comparable was one obvious drawback of the model. In reality, these situations were very different, especially considering the major economic upheavals caused by the COVID-19 pandemic. This difference could have made it more difficult for the model to capture more variance in stock return estimates. In retrospect, including additional variables such as a volatility index (like the VIX), the Fear and Greed Index, or industry-specific returns could have improved the model's accuracy and brought anticipated returns closer to actual values. Furthermore, considering a larger number of companies could have resulted in a more comprehensive analysis.

Exploring alternative modeling approaches, such as using random forests, could have also yielded higher accuracy. Furthermore, the study assumed an average reaction to the tariffs across all stocks, whereas a stock-specific interaction term might have provided more nuanced insights and improved the model's predictive performance.

5.5 Practical Implications

This study shows that using historical data to analyze policy shocks can be a practical, data-driven way to understand how markets react. By giving a clearer picture of how stocks respond to policy changes, this approach can help businesses, investors, and governments make more informed decisions. Plus, using past data for predictions can reduce some of the risks associated with stock market investments.

The study also demonstrated that combining machine learning techniques with panel regression can give valuable insights that aren't immediately obvious when just looking at historical price reactions. For example, understanding how past tariff reactions influenced stock returns can help investors better assess the risks linked to similar future events, making investment decisions more grounded and reliable.

This information is especially important for policymakers when they consider the potential effects of introducing new tariffs or regulations. Likewise, knowing how previous policy shocks impacted stock performance helps investors make more accurate risk assessments when deciding whether to buy or hold. These insights can also support businesses in planning for changes that might come from new government policies, like tariffs.

One practical example of this can be seen in portfolio management. When building a portfolio, it might be useful to include the policy effect on stocks as a variable to better estimate the real beta of a stock, making risk assessments more precise.

For future research, it would be helpful to expand the model by adding more predictors and extending the time window, maybe to three months after a policy event, to capture more variability. This is especially important since daily returns tend to be noisy and unpredictable.

Bibliographic References

[1] Li, J., Liang, Y., Wang, J., & Zhou, X. (2022). A survey of multi-view clustering: Current techniques, emerging challenges and future directions. *Expert Systems with Applications*, 189, 116126. https://doi.org/10.1016/j.eswa.2021.116126

[2] Tang, Z. (2024). Risk-return analysis of equity portfolios: Comparison between CAPM and Fama-French three-factor model. In S. Tehseen et al. (Eds.), *Proceedings of ICEDBC 2023 (AEBMR 258, pp. 227–237)*. Atlantis Press. https://doi.org/10.2991/978-94-6463-246-0_28

[3] Michailidis, G., & de Andrade e Silva, M. (2004). The effect of beta estimation on the performance of the CAPM: An empirical study. *Applied Financial Economics*, 14(15), 1127–1134. https://doi.org/10.1080/0960310042000187351

[4] PolitiFact. (2023, October 5). Fact-check: Trump's repeated falsehood that he saved the U.S. steel industry. https://www.politifact.com/factchecks/2023/oct/05/donald-trump/fact-check-trumps-repeated-falsehood-that-he-saved/#:~:text=In%20March%202018%2C%20Trump%20imposed,was%20intended%20to%20punish%20China

[5] U.S. Department of Commerce. (2018). Section 232 investigations: The effect of steel imports on the national security. U.S. Department of Commerce. https://www.commerce.gov/issues/trade-enforcement/section-232-steel#:~:text=,steel%20imports%20on%20government%20revenue

[6] Naylor, B. (2018, March 8). Trump expected to formally order tariffs on steel, aluminum imports. *NPR*. https://web.archive.org/web/20191231034502/https://www.npr.org/2018/03/08/591744195/trump-expected-to-formally-order-tariffs-on-steel-aluminum-imports

[7] U.S. Government Accountability Office. (2020). Steel and aluminum tariffs: Commerce should improve its exclusion request process and economic impact reviews (GAO-20-517). https://www.gao.gov/products/gao-20-517

[8] Durante, A. (2022). How Section 232 tariffs harmed the economy. *Tax Foundation*. https://taxfoundation.org/research/all/federal/how-section-232-tariffs-harmed-the-economy/

[9] Bohan, C., & Lawder, D. (2018, March 5). Trump's trade tariffs: Long on rhetoric, short on impact? *Reuters*. https://www.reuters.com/article/us-usa-trade-trump-rhetoric-analysis/trumps-trade-tariffs-long-on-rhetoric-short-on-impact-idUSKBN1GH37N

[10] Kelly, K. (2020, October 14). Trump steel tariffs bring job losses to swing state Michigan. *Reuters*. https://www.reuters.com/article/economy/trump-steel-tariffs-bring-job-losses-to-swing-state-michigan-idUSKBN26V06M

[11] The White House. (2025, February). Fact sheet: President Donald J. Trump restores Section 232 tariffs. *The White House*. https://www.whitehouse.gov/fact-sheets/2025/02/fact-sheet-president-donald-j-trump-restores-section-232-tariffs/

[12] The White House. (2025, February). Adjusting imports of aluminum into the United States. *The White House*. https://www.whitehouse.gov/presidential-actions/2025/02/adjusting-imports-of-aluminum-into-the-united-states/

[13] Bodie, Z., Kane, A., & Marcus, A. J. (2014). *Investments* (10th ed.). McGraw-Hill Education.

[14] Kenton, W. (2025, April 29). What beta means for investors. *Investopedia*. https://www.investopedia.com/terms/b/beta.asp

[15] Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press.

[16] Geron, A. (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media.

[17] Draper, N. R., & Smith, H. (1998). *Applied Regression Analysis* (3rd ed.). Wiley-Interscience.