Machine Learning Policy Simulation of Section 301 Tariff Exposure

Alfredo Sosa

2026-03-22

1 Introduction

This memo presents a machine learning framework for simulating the effects of tariff exposure on U.S. vehicle production and sales. Building on the empirical analysis in Sosa (2026), the objective is to quantify the contribution of trade policy exposure to observed economic outcomes using a predictive model that incorporates policy intensity directly as an input.

The analysis focuses on four outcomes: commercial vehicle production, light vehicle production, commercial vehicle sales, and light vehicle sales. For each outcome, the model generates predictions under both observed tariff exposure and a counterfactual scenario in which exposure is set to zero. The difference between these predictions provides a model-based estimate of the contemporaneous effect of tariff exposure.

2 Data

The dataset is derived from the monthly HS6-level panel constructed in Sosa (2026), which is collapsed to a national time series for each outcome. The data include both macroeconomic controls and a policy exposure index that captures the intensity of tariff exposure over time.

Each outcome is denoted by (Y_t), and the feature vector includes the exposure variable along with macroeconomic controls, time trends, seasonal indicators, and lagged values of the outcome. Formally, the feature vector is given by:

$$X_t = \left(\text{Exposure}_t,\ Z_t,\ Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-6}, Y_{t-12}, \overline{Y}_t^{(3)}, \overline{Y}_t^{(6)}, \overline{Y}_t^{(12)}, S_t\right)$$

where (Z_t) represents macroeconomic variables such as unemployment, inflation, and interest rates, and (S_t) captures seasonal patterns. The exposure variable reflects the degree of tariff intensity affecting the sector at each point in time.

This structure allows the model to jointly account for macroeconomic conditions and policy exposure when predicting production and sales dynamics.

3 Methodology

The analysis implements a walk-forward machine learning framework in which the model is recursively estimated using only past information. At each period, predictions are generated using the available information set, ensuring that the simulation reflects real-time forecasting conditions:

$$\hat{Y}_t = f_t(X_t)$$

where (f_t( )) is approximated using a Random Forest model.

The key feature of this framework is the policy simulation exercise. For each time period, two predictions are generated:

$$\hat{Y}_t^{(1)} = f_t(\text{Exposure}_t, X_t)$$

$$\hat{Y}_t^{(0)} = f_t(0, X_t)$$

The first prediction uses the observed exposure level, while the second sets exposure to zero, holding all other variables constant. The machine learning policy effect is defined as:

$$\text{ML Effect}_t = \hat{Y}_t^{(1)} - \hat{Y}_t^{(0)}$$

This quantity represents the model-implied effect of tariff exposure on the outcome, conditional on observed macroeconomic conditions and historical dynamics.

It is important to note that this is a one-step simulation. Lagged outcome variables remain fixed at their observed values, so the estimated effect should be interpreted as a contemporaneous partial effect rather than a fully dynamic structural response.

To assess statistical significance, the average policy effect is evaluated using heteroskedasticity and autocorrelation consistent (HAC) standard errors, allowing for serial correlation in the estimated effect series.

4 Results

The results indicate that tariff exposure has a systematic and economically meaningful effect on both production and sales outcomes. The comparison between predictions with observed exposure and those with zero exposure reveals persistent differences following the introduction of Section 301 tariffs.

In the post-policy period, the model frequently predicts lower levels of production and sales when exposure is present, suggesting that tariff intensity contributes negatively to economic performance. These effects are particularly informative when examined outside the COVID-19 window, where the influence of broader macroeconomic shocks is less dominant.

The walk-forward structure ensures that these findings are based on real-time predictions rather than ex post fitting. As a result, the estimated policy effects reflect the information available to economic agents at each point in time.

The use of HAC standard errors further supports the statistical reliability of the results by accounting for serial dependence in the policy effect estimates.

5 Contributions

This analysis extends the empirical framework of Sosa (2026) by incorporating policy exposure directly into a machine learning model. While traditional econometric approaches identify average treatment effects, the present framework provides a complementary simulation-based perspective that isolates the contribution of policy intensity within a predictive setting.

The methodology allows for a transparent decomposition of observed outcomes into components attributable to macroeconomic fundamentals and those associated with tariff exposure. This provides a flexible tool for evaluating the economic impact of trade policy in a high-frequency time-series context.

In addition, the project demonstrates the integration of machine learning, policy variables, and statistical inference within a unified analytical pipeline. This approach highlights the potential for combining modern data science techniques with applied economic analysis to address policy-relevant questions.

6 Interactive Dashboard

An interactive Tableau dashboard accompanies this analysis and presents the simulated policy effects across all four outcomes. The dashboard allows users to compare predicted outcomes under observed and zero-exposure scenarios, visualize the resulting policy effects, and explore how these effects evolve across different time periods.

The dashboard is available at:

https://public.tableau.com/app/profile/alfredo.sosa/viz/dashboard_ml_policy_simulation/PolicySimulationMain

This interactive component enhances the interpretability of the results by providing a visual representation of the policy simulation framework. It also illustrates the integration of data engineering, machine learning, and visualization tools within a cohesive workflow.

7 Conclusion

The machine learning policy simulation framework provides a flexible and informative approach to evaluating the effects of tariff exposure on production and sales. By generating predictions under alternative policy scenarios, the model isolates the contribution of exposure to observed economic outcomes.

The results suggest that tariff exposure plays a meaningful role in shaping production and sales dynamics in the U.S. automotive sector. When combined with the causal estimates in Sosa (2026) and the counterfactual analysis presented in the previous memo, this framework contributes to a comprehensive understanding of the economic impact of Section 301 tariffs.