

Article

Causal Machine Learning for Macroeconomic Forecasting Under Structural Breaks and Economic Uncertainty

Oumaima Abouzaid * and Faouzi Bousseadra

Department of Economics and Management, Chouaib Doukkali University, El Jadida 24000, Morocco;
bousseadra.f@ucd.ac.ma

* Correspondence: oum.abouzaid@gmail.com

Abstract

Macroeconomic forecasting has become increasingly challenging in environments characterized by structural breaks, nonlinear dynamics, and elevated economic uncertainty. Traditional econometric forecasting models frequently experience substantial predictive deterioration during periods of financial crises, geopolitical instability, and rapidly evolving macroeconomic conditions due to their reliance on assumptions of parameter stability and linear economic relationships. In response to these limitations, this study proposes an integrated causal machine learning framework designed to improve macroeconomic forecasting performance under structural instability and uncertainty. The proposed framework combines structural break detection techniques, machine learning algorithms, causal inference methodologies, and Explainable Artificial Intelligence (XAI) tools within a unified empirical architecture. More specifically, the study integrates Bai–Perron structural break analysis, Markov-Switching regime identification, Double Machine Learning (DML), Causal Forest estimation procedures, and SHAP-based explainability techniques. The empirical analysis employs a U.S. macroeconomic time-series dataset covering major crisis episodes, including the 2008 Global Financial Crisis, the COVID-19 pandemic, and the 2022 inflation shock. The dataset combines inflation, monetary, financial, energy-market, and uncertainty indicators obtained from publicly available U.S. macroeconomic databases. The empirical findings demonstrate that causal machine learning models significantly outperform conventional econometric frameworks such as VAR and TVP-VAR models, as well as standard machine learning algorithms including Random Forest (RF), XGBoost, and LSTM networks. The Double Machine Learning framework generates the strongest forecasting performance across all forecasting horizons, economic regimes, and robustness specifications. The results further reveal that macroeconomic relationships are highly regime-dependent and strongly influenced by uncertainty indicators, financial volatility, oil price shocks, and monetary policy dynamics. Explainability analysis additionally shows that uncertainty measures and energy market variables become dominant drivers of inflation forecasts during crisis periods characterized by elevated instability. The study contributes to the growing literature on macroeconomic forecasting by bridging econometric forecasting theory, causal inference methodologies, machine learning techniques, and explainable artificial intelligence within a unified forecasting framework. The findings provide important implications for central banks, policymakers, and financial institutions seeking more adaptive, transparent, and robust forecasting systems under uncertain macroeconomic environments.



Academic Editors: Herbert Kimura
and Leonardo Fernando Cruz Basso

Received: 14 May 2026

Revised: 15 June 2026

Accepted: 19 June 2026

Published: 5 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)

[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: causal machine learning; macroeconomic forecasting; structural breaks; economic uncertainty; explainable AI; econometrics; inflation forecasting; regime switching

1. Introduction

Macroeconomic forecasting constitutes one of the fundamental pillars of modern economic analysis and public policy design. Accurate forecasts of inflation, economic growth, unemployment, industrial production, and interest rates are essential for governments, central banks, financial institutions, and international organizations. Monetary authorities rely heavily on macroeconomic projections to formulate interest-rate policies, maintain price stability, and support sustainable economic growth. Similarly, private institutions and financial markets use macroeconomic forecasts to guide investment decisions, portfolio allocation strategies, and risk management practices. Consequently, improving the reliability and adaptability of macroeconomic forecasting models remains a central challenge in both research and policymaking.

Despite decades of advances in econometric modeling, macroeconomic forecasting continues to face substantial limitations, particularly during periods characterized by structural instability and elevated uncertainty. Traditional econometric approaches, such as Autoregressive Integrated Moving Average (ARIMA), Vector Autoregression (VAR), Phillips curve models, and Dynamic Stochastic General Equilibrium (DSGE) frameworks, generally rely on assumptions of parameter stability, linear relationships, and stable economic structures over time. However, recent economic developments have repeatedly demonstrated that these assumptions are frequently violated during crises and major structural transformations.

Over the past two decades, the global economy has experienced several disruptive events that have fundamentally altered macroeconomic dynamics. The 2008 global financial crisis generated unprecedented volatility in financial markets and weakened traditional monetary transmission mechanisms. More recently, the COVID-19 pandemic generated simultaneous shocks that disrupted labor markets, production systems, inflation dynamics, and international trade flows. In addition, geopolitical tensions, energy crises, supply-chain disruptions, monetary policy tightening, and climate-related shocks have further intensified global economic uncertainty. These developments introduced structural breaks and regime shifts that significantly reduced the predictive performance of conventional forecasting models.

Structural breaks refer to changes in the underlying relationships among macroeconomic variables and may arise from policy interventions, financial crises, institutional reforms, technological change, or geopolitical events. When such breaks occur, parameters estimated using historical data frequently become unstable, resulting in deteriorating forecasting accuracy and increasing prediction errors. Consequently, forecasting models calibrated under historical conditions may fail to adequately capture evolving economic regimes and changing market structures. This challenge became particularly evident during the post-pandemic period, when inflation persistence, monetary tightening, and geopolitical fragmentation substantially altered global macroeconomic dynamics.

The limitations of conventional econometric forecasting techniques have encouraged researchers to explore alternative methodologies capable of handling nonlinear relationships, high-dimensional datasets, and rapidly changing economic environments. In this context, machine learning and artificial intelligence techniques have emerged as promising tools for macroeconomic forecasting. Algorithms such as RF, Gradient Boosting Machines, Support Vector Machines, and Deep Neural Networks offer strong predictive capabilities and can identify complex nonlinear interactions that are often difficult to capture using traditional econometric frameworks. Recent empirical evidence suggests that machine learning models often outperform conventional econometric approaches in forecasting inflation, economic growth, financial volatility, and recession probabilities.

However, despite their predictive advantages, standard machine learning models present several important limitations when applied to macroeconomic analysis. First, many machine learning algorithms operate as black-box systems, offering limited interpretability regarding the mechanisms underlying their outputs. Second, most predictive algorithms focus primarily on statistical associations rather than causal relationships, thereby reducing their usefulness for policy evaluation and structural economic analysis. In macroeconomics, understanding causality remains critically important because policymakers require insights into how economic variables interact and respond to policy interventions. Forecasting accuracy alone is insufficient when the underlying causal transmission mechanisms remain unclear.

Recent advances in causal machine learning offer a potential solution to these challenges by combining causal inference methodologies with machine learning algorithms. Causal machine learning seeks not only to improve predictive performance but also to identify economically meaningful causal relationships within complex, high-dimensional datasets. Techniques such as Double Machine Learning (DML), Causal Forests, Bayesian Structural Time Series models, and related approaches have attracted growing attention because they enable researchers to estimate heterogeneous effects, address endogeneity concerns, and improve interpretability while maintaining strong predictive capabilities.

At the same time, Explainable Artificial Intelligence (XAI) techniques have become increasingly important for addressing the transparency limitations associated with machine learning systems. Tools such as SHAP (Shapley Additive Explanations), Local Interpretable Model-Agnostic Explanations (LIME), and feature importance decomposition allow researchers to better understand the contribution of explanatory variables to forecasting outcomes. In macroeconomic forecasting, explainability tools help identify how inflation, interest rates, exchange rates, oil prices, financial stress indicators, and uncertainty measures affect model outputs across economic regimes. Such transparency is particularly important for policymakers and central banks because forecasting credibility depends not only on predictive accuracy but also on the ability to explain and justify model outcomes.

Despite the rapid development of machine learning, causal inference, and explainable artificial intelligence techniques, an important gap remains in the macroeconomic forecasting literature. Existing studies generally address these dimensions separately. Some focus on structural break detection and regime-switching analysis, while others emphasize predictive accuracy through machine learning. Explainability and causal inference are also typically examined in isolation. As a result, few studies integrate these dimensions within a unified forecasting architecture. However, very few studies integrate these methodological dimensions within a unified forecasting architecture capable of simultaneously addressing structural instability, causal interpretation, predictive performance, and explainability.

This fragmentation represents a significant limitation in contemporary macroeconomic environments characterized by recurrent crises, regime shifts, and elevated uncertainty. Forecasting models that achieve strong predictive performance may remain difficult to interpret and therefore unsuitable for policy analysis, while highly interpretable econometric models often struggle to adapt to nonlinear dynamics and structural instability. Consequently, there is an increasing need for forecasting frameworks capable of jointly incorporating adaptability, predictive accuracy, causal understanding, and transparency within a single empirical setting.

To address this gap, this study develops an integrated causal machine learning framework for macroeconomic forecasting under structural breaks and economic uncertainty. The originality of the proposed approach does not lie in introducing a new standalone forecasting algorithm. Rather, it resides in the integration of several complementary methodological components that are generally employed independently in the existing literature. Specifically, the framework combines Bai–Perron structural break detection, Markov-Switching

regime identification, Double Machine Learning estimation, Causal Forest analysis, and SHAP-based explainability techniques within a unified forecasting pipeline. This integrated architecture enables the simultaneous identification of structural instability, estimation of causal relationships, generation of robust forecasts, and interpretation of forecasting outcomes across different economic regimes.

The empirical analysis focuses on U.S. inflation forecasting using a single-country macroeconomic time-series dataset compiled from FRED, the CBOE, and the U.S. Economic Policy Uncertainty database. The objective is to evaluate forecasting performance under structural instability within the U.S. macroeconomic environment. Forecasting performance is evaluated through a comparative assessment involving traditional econometric models (VAR and TVP-VAR), conventional machine learning algorithms (RF, XGBoost, and Long Short-Term Memory networks), and causal machine learning approaches. Although the broader motivation of this study relates to global macroeconomic instability and international economic uncertainty, the empirical analysis is restricted to the United States. Therefore, all reported forecasting results and estimated relationships should be interpreted within the context of the U.S. macroeconomic environment.

This study contributes to the literature in three principal ways. First, it proposes a unified forecasting architecture that jointly incorporates structural break analysis, regime identification, causal machine learning, and explainable artificial intelligence, thereby extending existing hybrid forecasting approaches. Second, it provides empirical evidence regarding the extent to which causal machine learning techniques improve forecasting performance under conditions of structural instability and economic uncertainty. Third, it enhances the interpretability of macroeconomic forecasting systems by integrating SHAP-based explainability tools into a causal forecasting framework, thereby strengthening the policy relevance and transparency of machine-learning-based economic forecasts.

From a practical perspective, the findings are relevant for central banks, policymakers, and financial institutions. More adaptive and interpretable forecasting systems may improve monetary policy design, inflation targeting, financial stability monitoring, and macroprudential decision-making under uncertainty.

The remainder of this paper is organized as follows. Section 2 reviews the theoretical and empirical literature on macroeconomic forecasting, structural instability, machine learning, causal inference, and explainable artificial intelligence. Section 3 presents the methodological framework and econometric procedures employed in the empirical analysis. Section 4 reports the empirical results. Section 5 discusses the principal findings and theoretical implications. Section 6 concludes and outlines future research directions.

2. Literature Review

2.1. Traditional Macroeconomic Forecasting Models and Structural Instability

Macroeconomic forecasting has long occupied a central position in economic analysis, monetary policy implementation, and financial decision-making. Governments, central banks, and international organizations rely on projections of inflation, economic growth, unemployment, industrial production, exchange rates, and interest rates to support policy decisions and anticipate economic developments. Consequently, forecasting accuracy has become a major concern for both researchers and policymakers, leading to the development of increasingly sophisticated econometric frameworks.

The earliest forecasting approaches were primarily based on univariate time-series models such as Autoregressive (AR), Moving Average (MA), and Autoregressive Integrated Moving Average (ARIMA) models developed within the Box–Jenkins framework. These methods proved useful for short-term prediction because of their simplicity and low data requirements. However, their capacity to model interactions among macroeconomic variables

remained limited, reducing their usefulness for understanding the broader mechanisms driving economic fluctuations.

A major methodological advancement occurred with the introduction of multivariate econometric models. Among these, the Vector Autoregression (VAR) framework proposed by Sims (1980) became one of the most influential tools in empirical macroeconomics. VAR models provide a flexible framework in which each endogenous variable depends on its own lagged values and those of other variables in the system. This approach allows researchers to analyze dynamic interactions among macroeconomic indicators without imposing strong theoretical restrictions. As a result, VAR models rapidly became widely used in forecasting, policy evaluation, and business cycle analysis.

Subsequent extensions further enhanced the capabilities of the VAR framework. Structural Vector Autoregression (SVAR) models introduced identification restrictions that enabled researchers to estimate structural shocks and investigate economic transmission mechanisms. These models became particularly important in monetary economics because they facilitate the analysis of policy interventions, demand shocks, and supply-side disturbances. At the same time, Bayesian Vector Autoregression (BVAR) models emerged as a response to the overparameterization problems often encountered in large-scale VAR systems. By incorporating prior information into the estimation process, BVAR approaches improve forecasting accuracy and reduce estimation uncertainty, particularly in data-rich environments (Chin & Li, 2019).

Another important development in forecasting literature was the emergence of Dynamic Factor Models (DFMs). The growing availability of macroeconomic and financial data created new opportunities but also introduced dimensionality challenges. DFMs address this issue by extracting a limited number of latent common factors from a large set of economic indicators. Research by Stock and Watson (2005, 2016) demonstrated that factor-based approaches can significantly improve forecasting performance when dealing with large datasets. Consequently, DFMs have become widely used by central banks and international institutions for economic monitoring, nowcasting, and short-term forecasting.

Theoretical developments in macroeconomics also contributed to the increasing popularity of Dynamic Stochastic General Equilibrium (DSGE) models. Built on microeconomic foundations and rational expectations theory, DSGE models describe the interactions among households, firms, governments, and monetary authorities within a coherent equilibrium structure. Their main advantage lies in their ability to provide theoretically consistent explanations of economic fluctuations and policy transmission channels. Del Negro and Schorfheide (2012) emphasize that DSGE models have become indispensable tools in modern monetary policy analysis because they combine forecasting capabilities with structural economic interpretation.

Despite these advances, the performance of traditional econometric models remains highly dependent on assumptions of parameter stability and stable economic relationships. In practice, however, economies are frequently exposed to financial crises, policy regime changes, technological innovation, geopolitical tensions, pandemics, and external shocks that alter the relationships among macroeconomic variables. As a result, forecasting models calibrated on historical data often experience significant declines in predictive accuracy when economic conditions change abruptly.

This challenge has led researchers to focus increasingly on structural instability and regime changes. Structural breaks refer to changes in the statistical properties of economic time series, including shifts in means, variances, covariances, or dynamic relationships among variables. Perron (1989) demonstrated that neglecting structural changes may lead to biased parameter estimates, misleading statistical inference, and poor forecasting perfor-

mance. Consequently, accounting for structural instability has become a major concern in empirical macroeconomics.

The relevance of structural breaks became particularly evident following major global disruptions such as the oil crises of the 1970s, the Asian financial crisis, the 2008 global financial crisis, and the COVID-19 pandemic. These events revealed that macroeconomic relationships are not constant over time and frequently evolve across different economic regimes. Variables such as inflation, output growth, unemployment, exchange rates, and monetary policy transmission mechanisms often behave differently during periods of crisis than during periods of economic stability.

To address these challenges, researchers developed models capable of capturing regime-dependent dynamics. A major contribution was made by Hamilton (1989), who introduced the Markov-Switching framework. This approach assumes that economic variables evolve according to distinct unobservable regimes characterized by different statistical properties. By allowing transitions between expansionary and recessionary states, Markov-switching models capture nonlinearities, asymmetries, and abrupt regime changes that cannot easily be accommodated within traditional linear frameworks. Empirical studies have shown that these models often outperform conventional forecasting approaches during periods of elevated uncertainty and financial instability.

Another significant advancement was provided by Bai and Perron (2003), who developed econometric procedures for identifying multiple endogenous structural breaks in time-series data. Unlike earlier approaches that required researchers to specify break dates a priori, the Bai–Perron methodology allows breakpoints to be estimated directly from the data. Because of its flexibility and statistical robustness, this approach has become one of the most widely used techniques for structural break analysis in empirical economics.

Although structural break and regime-switching methodologies have considerably improved the robustness of macroeconomic forecasting, important limitations remain. Most of these approaches continue to rely on relatively restrictive econometric structures and may struggle to capture highly nonlinear relationships, complex interactions, and the growing dimensionality of modern economic datasets. These limitations have motivated the increasing adoption of machine learning techniques, which offer new opportunities for handling complex economic environments characterized by uncertainty, nonlinearity, and rapidly evolving data structures.

2.2. Machine Learning, Causal Machine Learning and Explainable Artificial Intelligence in Macroeconomic Forecasting

The growing complexity of modern economies, combined with the rapid expansion of digital technologies and large-scale economic databases, has significantly transformed macroeconomic forecasting methodologies. Traditional econometric models such as VAR, BVAR, DFM, and DSGE frameworks remain widely used because of their theoretical consistency and interpretability. However, their forecasting performance often deteriorates in environments characterized by nonlinear relationships, structural instability, and high-dimensional datasets. These limitations have encouraged researchers to increasingly explore machine learning techniques as complementary or alternative forecasting tools.

Machine learning differs from traditional econometric approaches primarily in its objective. While econometric models are generally designed to identify structural relationships and test economic theories, machine learning algorithms prioritize predictive performance through data-driven learning processes. By relying less on restrictive functional assumptions, machine learning models can identify complex nonlinear interactions, hidden dependencies, and high-order relationships that may be difficult to capture using conventional econometric frameworks.

The growing popularity of machine learning in macroeconomics is closely linked to the increasing volume and diversity of available economic data. Modern economies generate vast amounts of information through financial markets, digital transactions, online platforms, satellite observations, and real-time monitoring systems. As forecasting models increasingly rely on hundreds of potential explanatory variables, traditional econometric methods often encounter difficulties related to overparameterization, multicollinearity, and computational constraints. Machine learning algorithms, by contrast, are specifically designed to process large-scale datasets efficiently while automatically identifying relevant predictive patterns.

Among the earliest machine learning techniques applied to macroeconomic forecasting were Artificial Neural Networks (ANNs). Inspired by biological neural systems, ANNs consist of interconnected computational layers capable of approximating highly nonlinear relationships. Empirical studies demonstrated that neural networks can improve forecasting performance for key macroeconomic indicators such as inflation, GDP growth, exchange rates, and unemployment. Zhang (2003) showed that combining neural networks with traditional time-series approaches allows forecasting systems to capture both linear and nonlinear dynamics more effectively than conventional methods alone.

Subsequent developments led to the widespread adoption of ensemble learning techniques. One of the most influential contributions was the RF algorithm introduced by Breiman (2001), which combines multiple decision trees to generate robust and accurate predictions while reducing overfitting risks. RF models have become increasingly popular in macroeconomic forecasting because they efficiently handle high-dimensional datasets and automatically detect nonlinear relationships among variables. Empirical evidence suggests that RF algorithms frequently outperform conventional econometric models when forecasting inflation dynamics, recession probabilities, and financial stress indicators.

Boosting algorithms further expanded forecasting capabilities. Gradient Boosting Machines (GBM) and Extreme Gradient Boosting (XGBoost) improve predictive performance by iteratively combining weak learners into increasingly accurate forecasting systems. These methods have demonstrated strong performance in both macroeconomic and financial applications because of their ability to model complex interactions within large datasets while maintaining relatively high predictive stability.

Recent advances in deep learning have further transformed forecasting methodologies. Architectures such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are specifically designed to process sequential information and time-series data. Unlike traditional machine learning algorithms, these models can capture long-run temporal dependencies and dynamic patterns, making them particularly suitable for forecasting business cycles, inflation persistence, financial volatility, and economic crises. Several studies suggest that deep learning architectures may outperform both conventional econometric models and standard machine learning techniques during periods characterized by elevated uncertainty and rapidly changing economic conditions.

Despite these important advances, the growing adoption of machine learning in macroeconomic forecasting has generated substantial debate regarding interpretability and economic relevance. Most machine learning algorithms are designed to maximize predictive accuracy rather than explain economic mechanisms. Consequently, many forecasting systems operate as black-box models, making it difficult to determine why predictions are generated or how economic variables contribute to forecasting outcomes.

This challenge has stimulated growing interest in causal machine learning, an emerging field that seeks to combine the predictive power of machine learning with the identification strategies of modern econometrics. Unlike conventional machine learning approaches that focus primarily on correlations and prediction, causal machine learning aims to estimate economically meaningful causal relationships while maintaining high predictive performance.

One of the most influential developments in this area is the Double Machine Learning (DML) framework proposed by Chernozhukov et al. (2018). DML integrates machine learning algorithms into econometric estimation procedures while preserving statistical validity and reducing biases associated with high-dimensional settings. By combining flexible predictive models with orthogonalization techniques, DML enables researchers to estimate causal effects even when the number of explanatory variables is very large.

Another important advancement concerns Causal Forests, which extend the RF methodology to estimate heterogeneous treatment effects across different observations. These approaches allow researchers to investigate how economic policies, shocks, or institutional changes affect different sectors, regions, or population groups. Such capabilities are particularly valuable in macroeconomic environments characterized by substantial heterogeneity and structural complexity.

The emergence of causal machine learning reflects a broader shift in forecasting research. While early machine learning studies primarily emphasized predictive performance, the recent literature increasingly recognizes that policymakers require forecasting systems capable of providing both accurate predictions and credible causal explanations. Consequently, causal machine learning has become an important bridge between traditional econometric analysis and modern artificial intelligence techniques.

Alongside these developments, Explainable Artificial Intelligence (XAI) has emerged as another major research direction aimed at improving transparency and accountability in machine learning systems. XAI refers to a set of methodologies designed to explain how machine learning models generate predictions and identify the contribution of individual explanatory variables. In macroeconomic forecasting, explainability is particularly important because policy decisions often depend not only on forecast accuracy but also on understanding the economic mechanisms underlying forecasting outcomes.

Recent literature emphasizes that the practical usefulness of machine learning in economics depends heavily on its ability to provide interpretable and economically meaningful results. Athey and Imbens (2019) argue that machine learning can significantly improve economic analysis and forecasting performance, but its policy relevance requires greater transparency regarding decision-making mechanisms and variable contributions.

Among existing explainability methods, SHAP (Shapley Additive Explanations) introduced by Lundberg and Lee (2017), has become one of the most widely adopted approaches. Derived from cooperative game theory, SHAP values quantify the contribution of each explanatory variable to a model's prediction, thereby enabling researchers to decompose forecasting outcomes into interpretable components. In macroeconomic forecasting, SHAP-based analyses are increasingly used to evaluate the influence of inflation, interest rates, exchange rates, unemployment, oil prices, financial stress indicators, and uncertainty measures on forecasting results.

Feature importance analysis constitutes another important branch of explainable AI. These methods rank explanatory variables according to their contribution to predictive performance and are particularly useful in high-dimensional macroeconomic environments where forecasting models incorporate large numbers of indicators. By identifying the most influential variables, explainability techniques contribute to a better understanding of the economic drivers behind machine learning predictions.

Although machine learning, causal machine learning, and explainable AI have substantially improved forecasting methodologies, important limitations remain. Existing studies often focus on one dimension at a time: maximizing predictive accuracy, identifying causal relationships, or improving model interpretability. As a result, forecasting systems frequently excel in one area while remaining weak in others. Models that achieve high predictive performance may remain difficult to interpret, whereas highly interpretable

models may sacrifice forecasting accuracy. Similarly, causal machine learning approaches often pay limited attention to structural instability and regime changes that characterize modern macroeconomic environments.

These limitations suggest that future forecasting systems must move beyond isolated methodological advances and develop integrated frameworks capable of simultaneously addressing predictive performance, causal interpretation, transparency, and structural instability. This challenge has motivated the emergence of hybrid forecasting architectures that combine econometric theory, machine learning, causal inference, and explainable artificial intelligence within a unified analytical framework.

2.3. Hybrid Forecasting Models, Real-Time Forecasting and Emerging AI-Based Forecasting Systems

The growing complexity of economic systems, combined with the increasing availability of large-scale and high-frequency datasets, has encouraged researchers to move beyond purely econometric or purely machine learning forecasting frameworks. Although traditional econometric models provide theoretical consistency and economic interpretability, their forecasting performance often deteriorates under structural instability and nonlinear dynamics. Conversely, machine learning algorithms frequently achieve superior predictive accuracy but may suffer from limited interpretability and weak causal foundations. These complementary strengths and weaknesses have motivated the development of hybrid forecasting models that seek to combine the advantages of both approaches.

Hybrid forecasting frameworks are designed to integrate economic theory with data-driven computational methods. Their primary objective is to preserve the interpretability and structural coherence of econometric models while benefiting from the nonlinear learning capabilities of machine learning algorithms. This integration has become increasingly important because modern macroeconomic environments are characterized by structural breaks, evolving economic relationships, financial uncertainty, and rapidly changing information flows.

One of the earliest motivations for hybrid forecasting originated from the recognition that macroeconomic time series frequently contain both linear and nonlinear components. Traditional econometric models are generally effective at capturing linear dynamic relationships, whereas machine learning methods are better suited for identifying hidden nonlinear interactions. Zhang (2003) demonstrated that combining ARIMA models with Artificial Neural Networks significantly improves forecasting performance by allowing the forecasting system to simultaneously model linear and nonlinear structures. This contribution laid the foundation for a large body of literature investigating hybrid forecasting architectures.

Subsequent developments expanded the integration of machine learning techniques into conventional macroeconomic frameworks. Researchers increasingly incorporated machine learning algorithms into VAR, BVAR, DFM, and DSGE models to improve variable selection, parameter estimation, nonlinear modeling, and forecasting accuracy. According to Nachane and Chaulal (2023), machine learning methods can substantially enhance DSGE-based forecasting by improving the ability of structural models to capture evolving economic dynamics and complex interactions among macroeconomic variables.

Dynamic Factor Models have also benefited from integration with machine learning techniques. While DFMs efficiently summarize information contained in large numbers of economic indicators, machine learning algorithms can complement these models by identifying nonlinear relationships between latent factors and target variables. Such combinations have demonstrated improved forecasting performance in high-dimensional environments characterized by large volumes of economic and financial information.

The growing use of ensemble learning techniques has further strengthened hybrid forecasting systems. Algorithms such as RFs, Gradient Boosting Machines, and XGBoost

are increasingly integrated with econometric forecasting models to enhance predictive stability and reduce forecasting errors. These methods are particularly effective in environments characterized by multicollinearity, nonlinear interactions, and structural instability. Empirical evidence suggests that hybrid ensemble models frequently outperform traditional econometric frameworks when forecasting inflation, GDP growth, exchange rates, recession probabilities, and financial stress indicators.

Recent advances in deep learning have accelerated the evolution of hybrid forecasting architectures. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are increasingly incorporated into macroeconomic forecasting frameworks because of their ability to model temporal dependencies and sequential patterns in economic data. Their forecasting performance proved particularly valuable during periods of heightened uncertainty, including the COVID-19 pandemic, when conventional econometric relationships became unstable and difficult to model using linear approaches alone.

At the same time, the forecasting literature has witnessed growing interest in real-time forecasting and nowcasting methodologies. Traditional forecasting systems often rely on official macroeconomic statistics published with significant delays. Consequently, policymakers frequently face information gaps when attempting to assess current economic conditions. Nowcasting seeks to address this problem by producing near real-time estimates of economic activity using continuously updated information from multiple data sources. Research based on Dynamic Factor Models, mixed-frequency data approaches, and machine learning algorithms has shown that nowcasting can significantly improve the timeliness and responsiveness of economic forecasts.

The increasing availability of alternative and high-frequency data has further accelerated the development of real-time forecasting systems. Financial market indicators, internet search activity, social media sentiment, satellite imagery, digital payment transactions, and mobility data are increasingly incorporated into forecasting models. These unconventional sources provide timely signals regarding economic activity and can complement traditional macroeconomic indicators, particularly during periods of rapid economic change. The COVID-19 crisis highlighted the importance of such data sources, as conventional indicators often failed to capture sudden economic disruptions in a timely manner.

More recently, transformer-based forecasting models have emerged as one of the most promising developments in artificial intelligence and time-series analysis. Originally developed for natural language processing, transformer architectures have demonstrated remarkable capabilities in capturing long-range dependencies and complex temporal relationships. Unlike recurrent neural networks, transformers process information through attention mechanisms that allow models to focus selectively on the most relevant observations across long forecasting horizons.

Several transformer-based architectures have recently been adapted for economic and financial forecasting applications. These models often outperform traditional deep learning frameworks when handling large datasets characterized by complex temporal structures. Their ability to process high-dimensional information efficiently makes them particularly attractive for macroeconomic forecasting environments involving numerous economic indicators, financial variables, and alternative data sources. As a result, transformer-based forecasting systems are increasingly viewed as a potential next generation of forecasting technology capable of addressing some of the limitations associated with conventional deep learning models.

The rapid expansion of artificial intelligence has also contributed to the emergence of AI-based macroeconomic prediction systems. These systems combine machine learning, deep learning, automated feature engineering, real-time data integration, and explainability techniques within unified forecasting environments. Unlike traditional forecasting frame-

works that rely on a limited set of indicators, AI-driven systems can continuously process large volumes of structured and unstructured information while dynamically adapting to evolving economic conditions. Such capabilities have attracted growing interest among central banks, international organizations, and financial institutions seeking more responsive forecasting tools.

Despite these important advances, the literature continues to reveal several unresolved challenges. Many hybrid forecasting models focus primarily on improving predictive accuracy without explicitly addressing causal interpretation. Similarly, real-time forecasting and nowcasting systems often prioritize timeliness while providing limited insight into the economic mechanisms underlying predictions. Transformer-based models and advanced AI forecasting systems further improve predictive capabilities but may exacerbate concerns regarding transparency, interpretability, and model accountability.

More importantly, structural break detection, regime-switching analysis, causal machine learning, explainable artificial intelligence, and advanced forecasting architectures are generally developed as separate research streams. Existing studies rarely integrate these dimensions within a unified framework capable of simultaneously addressing structural instability, predictive performance, causal interpretation, and transparency. Consequently, although forecasting methodologies have advanced considerably, the literature remains fragmented, creating opportunities for more comprehensive forecasting architectures capable of addressing the multiple challenges associated with modern macroeconomic forecasting.

2.4. Research Gap and Contribution of the Study

Despite these advances, important limitations remain. First, a large portion of the forecasting literature continues to rely on models that assume relatively stable economic relationships over time. Although structural break detection methods and regime-switching models have improved the treatment of economic instability, these approaches are often implemented separately from modern machine learning forecasting systems. Consequently, many forecasting frameworks remain insufficiently equipped to adapt to abrupt regime changes, financial crises, policy shifts, and periods of heightened uncertainty.

Second, the growing adoption of machine learning techniques has substantially improved forecasting accuracy in high-dimensional and nonlinear environments. However, most machine learning algorithms focus primarily on prediction rather than economic interpretation. As a result, many forecasting systems operate as black-box models that provide limited insight into the mechanisms driving economic outcomes. While high predictive performance is valuable, policymakers and economic institutions often require forecasting tools capable of explaining how and why economic variables influence forecasted outcomes.

Third, causal machine learning has emerged as a promising research direction that combines the predictive capabilities of machine learning with the identification strategies of modern econometrics. Techniques such as Double Machine Learning and Causal Forests have improved the estimation of causal effects in complex and high-dimensional settings. Nevertheless, existing studies generally focus on causal inference or policy evaluation rather than integrating causal estimation directly into forecasting architectures designed to operate under structural instability and regime changes.

Fourth, explainable artificial intelligence has made important contributions to improving transparency and interpretability. Approaches such as SHAP values and feature importance analysis help researchers understand the contribution of explanatory variables to forecasting outcomes. However, explainability techniques are frequently applied as ex-post diagnostic tools rather than as integral components of forecasting frameworks specifically designed for macroeconomic decision-making under uncertainty.

This fragmentation reveals an important research gap. Modern macroeconomic environments are characterized simultaneously by structural instability, nonlinear interactions, high-dimensional data, economic uncertainty, and increasing demands for policy transparency. Addressing these challenges requires forecasting systems capable of combining forecasting accuracy, causal interpretation, adaptability to regime changes, and model explainability. Yet, existing forecasting approaches rarely address these dimensions jointly.

To address this gap, the present study proposes an integrated causal machine learning framework for macroeconomic forecasting under conditions of structural breaks and economic uncertainty. The proposed framework integrates methods that are typically examined separately in the literature. Specifically, the framework integrates Bai–Perron structural break detection, Markov-Switching regime identification, Double Machine Learning estimation, Causal Forest analysis, and SHAP-based explainability within a unified forecasting architecture.

The proposed framework contributes to the literature in several important ways. First, it establishes a direct connection between structural instability analysis and causal machine learning, allowing forecasting models to adapt to changing economic regimes while preserving causal interpretability. Second, it integrates predictive modeling and explainable artificial intelligence within a common framework, thereby enhancing both forecasting performance and transparency. Third, it extends existing hybrid forecasting approaches by simultaneously addressing four critical dimensions of modern macroeconomic forecasting: structural breaks, predictive accuracy, causal inference, and explainability. By combining these elements within a unified architecture, the framework seeks to provide a more robust, interpretable, and policy-relevant forecasting system for environments characterized by uncertainty and rapid economic change. Therefore, the contribution of this study lies not in the development of a new individual forecasting technique, but in the design of an integrated methodological architecture capable of bridging several previously disconnected strands of the forecasting literature. This integrated perspective provides a foundation for more adaptive and transparent forecasting systems that are better suited to the challenges of contemporary macroeconomic analysis.

3. Methodology

3.1. Research Framework

This study develops an integrated empirical framework to investigate inflation forecasting performance under structural instability and economic uncertainty. The framework combines structural break analysis, regime identification, econometric forecasting models, machine learning algorithms, causal machine learning methods, and explainable artificial intelligence techniques.

The study pursues two complementary objectives. First, it evaluates whether causal machine learning models generate more accurate inflation forecasts than conventional econometric and machine learning approaches. Second, it investigates the causal mechanisms through which macroeconomic variables and uncertainty indicators influence inflation dynamics across different economic regimes.

The empirical framework consists of four interconnected stages. In the first stage, structural breaks and economic regimes are identified using Bai–Perron and Markov-Switching models. In the second stage, inflation forecasts are generated using econometric models (VAR and TVP-VAR), machine learning algorithms (RF, XGBoost, and LSTM), and causal machine learning methods (Double Machine Learning). In the third stage, causal machine learning techniques are employed to estimate average and heterogeneous causal effects of macroeconomic variables and uncertainty indicators on inflation. In the fourth stage,

SHAP-based explainability methods are applied to identify the most important drivers of inflation forecasts and evaluate how their importance changes across economic regimes.

This framework allows forecasting performance, causal mechanisms, and model interpretability to be analyzed simultaneously.

3.2. Data and Variables

- **Data Scope:** The empirical analysis is conducted for the United States and uses a single-country macroeconomic time-series dataset. The study does not employ a cross-country panel. The dataset combines macroeconomic, financial, energy-market, and uncertainty indicators obtained from FRED, the CBOE Volatility Index database, and the U.S. Economic Policy Uncertainty database. Consequently, the estimated forecasting relationships and causal effects should be interpreted within the context of the U.S. economy.
- **Data Vintage.** The empirical analysis relies on revised historical data obtained from FRED, the CBOE Volatility Index database, and the Economic Policy Uncertainty database. Real-time vintages are not employed. Consequently, the forecasting exercise should be interpreted as a pseudo out-of-sample forecasting evaluation rather than a real-time forecasting exercise.
- The empirical analysis focuses on inflation forecasting.
- The dependent variable is the inflation rate, which serves as the primary outcome variable throughout the analysis.
- The explanatory variables include monetary, financial, energy-market, and uncertainty indicators that have been widely identified as major determinants of inflation dynamics.

Table 1 summarizes the variables used in the empirical analysis, together with their descriptions and roles in the causal machine learning framework.

Table 1. Variables Used in the Analysis.

Variable	Description	Role
Inflation	Consumer Price Inflation Rate	Outcome Variable
Interest Rate	Monetary Policy Indicator	Treatment Variable
Oil Prices	Energy Price Indicator	Treatment Variable
Exchange Rate Depreciation	External Price Shock Indicator	Treatment Variable
Money Supply Growth	Monetary Expansion Indicator	Treatment Variable
Economic Policy Uncertainty (EPU)	Policy Uncertainty Indicator	Treatment Variable
Financial Volatility (VIX)	Financial Market Uncertainty Indicator	Treatment Variable

Inflation constitutes the outcome variable in all causal analyses, whereas the remaining variables are successively treated as treatment variables within the Double Machine Learning framework. All variables are transformed and tested for stationarity prior to estimation.

3.3. Structural Break Detection and Regime Identification

3.3.1. Bai–Perron Structural Break Test

To account for potential instability in inflation dynamics, the study first identifies structural changes using the [Bai and Perron \(2003\)](#) multiple breakpoint methodology.

The model is specified as:

$$y_t = x_t' \beta_j + u_t$$

where

- y_t denotes inflation;
- x_t represents explanatory variables;
- β_j denotes regime-specific coefficients;
- u_t is the disturbance term.

The procedure identifies endogenous break dates associated with major economic disruptions and changes in inflation dynamics.

3.3.2. Markov-Switching Regime Model

To capture nonlinear macroeconomic dynamics, the study employs a Markov-Switching model:

$$y_t = \mu_{s_t} + \phi_{s_t} y_{t-1} + \varepsilon_t$$

where s_t denotes an unobserved economic regime.

The estimated regime probabilities allow observations to be classified into expansionary and recessionary periods. These regime classifications are subsequently used to investigate whether causal relationships differ across economic environments.

3.4. Econometric Forecasting Models

Two benchmark econometric models are estimated.

Vector Autoregression (VAR)

$$Y_t = c + A_1 Y_{t-1} + \dots + A_p Y_{t-p} + \varepsilon_t$$

Time-Varying Parameter VAR (TVP-VAR)

$$Y_t = c_t + B_t Y_{t-1} + u_t$$

The TVP-VAR specification allows model parameters to evolve over time and partially accommodates structural instability.

Lag Selection: The lag order of the VAR model is selected using the Bayesian Information Criterion (BIC). Robustness checks using alternative information criteria (AIC and HQIC) produced qualitatively similar results.

3.5. Machine Learning Forecasting Models

Three machine learning forecasting algorithms are employed.

RF

$$\hat{f}(x) = \frac{1}{B} \sum_{b=1}^B T_b(x)$$

XGBoost

$$L(\phi) = \sum_{i=1}^n l(\hat{y}_i, y_i) + \sum_{k=1}^K \Omega(f_k)$$

Long Short-Term Memory Networks (LSTM)

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

These algorithms provide nonlinear forecasting benchmarks against which causal machine learning performance can be compared.

3.6. Double Machine Learning Forecasting Framework

The central forecasting contribution of this study is the implementation of the Double Machine Learning (DML) framework proposed by Chernozhukov et al. (2018).

The DML model is specified as:

$$\begin{aligned} Y &= D\theta_0 + g_0(X) + U \\ D &= m_0(X) + V \end{aligned}$$

where

- Y denotes inflation;
- D denotes the treatment variable;
- X represents control variables;
- θ_0 measures the causal effect.

Separate DML specifications are estimated for each treatment variable, including interest rates, oil prices, exchange rates, money supply growth, economic policy uncertainty, and financial volatility.

In each specification, one variable is treated as the treatment variable (interest rate, oil prices, exchange rate depreciation, money supply growth, Economic Policy Uncertainty, or VIX), while the remaining macroeconomic variables together with lagged inflation are included in the control vector X . The identification strategy relies on a conditional independence assumption under which treatment assignment is assumed independent of potential outcomes conditional on the observed controls.

Identification Strategy and Causal Interpretation. Although Double Machine Learning (DML) provides a flexible framework for estimating treatment effects in high-dimensional settings, causal interpretation in this study relies on explicit identifying assumptions. In each DML specification, one macroeconomic variable is treated as the treatment variable (interest rate, oil prices, exchange rate depreciation, money supply growth, Economic Policy Uncertainty, or VIX), while the remaining macroeconomic variables together with lagged inflation are included in the control set. The identification strategy relies primarily on a conditional independence assumption under which treatment assignment is assumed independent of potential inflation outcomes conditional on the observed controls. Additional assumptions include overlap in treatment variation across observations, sufficient control of dynamic confounding through lagged macroeconomic variables, and approximate local stability within detected economic regimes. Because the empirical analysis is based on observational macroeconomic time-series data rather than randomized interventions or natural experiments, these assumptions cannot be fully verified empirically. Consequently, the estimated treatment effects should be interpreted as causally informative estimates under maintained assumptions rather than definitive structural causal effects.

To reduce overfitting and regularization bias, the DML procedure is implemented using 5-fold cross-fitting. The sample is partitioned into five mutually exclusive folds. In each fold, nuisance functions for treatment prediction and outcome prediction are estimated on auxiliary folds using machine learning algorithms, while predictions are generated on the held-out fold. The final treatment effect estimate is obtained using orthogonalized residuals, thereby reducing regularization bias and improving statistical robustness in high-dimensional settings.

Machine learning algorithms are employed to estimate the nuisance functions $g_0(X)$ and $m_0(X)$, while orthogonalization procedures reduce estimation bias.

The estimated DML model is subsequently used to generate out-of-sample inflation forecasts.

Unlike conventional machine learning models that rely primarily on predictive correlations, DML explicitly isolates causal relationships between inflation and its macroeconomic determinants. This approach is expected to improve forecasting robustness under structural instability and uncertainty shocks.

3.7. Causal Forest Analysis

Causal Forests are employed to estimate Conditional Average Treatment Effects (CATEs) and explore treatment-effect heterogeneity across economic regimes and uncertainty environments using the same treatment and control variable structure adopted in the DML framework.

While DML is used for forecasting and causal estimation, Causal Forests are employed to investigate heterogeneous treatment effects.

The Conditional Average Treatment Effect is defined as:

$$\tau(x) = E[Y(1) - Y(0) \mid X = x]$$

The Causal Forest framework estimates how the effects of macroeconomic variables vary across:

- low-uncertainty and high-uncertainty environments;
- expansionary and recessionary regimes.

The objective is not to generate forecasts directly but rather to explain the heterogeneity of causal relationships observed across economic environments.

It is important to note that Causal Forest is not used as a standalone forecasting model in the headline forecast comparison tables. Its role in this study is exclusively to estimate heterogeneous treatment effects and explore regime-dependent causal heterogeneity. Consequently, forecasting performance tables report only models directly used for out-of-sample prediction.

3.8. Forecast Design and Validation Strategy

Forecasting performance is evaluated using a strictly out-of-sample framework. The sample is divided into training, validation, and testing subsamples. The training sample is used for model estimation, the validation sample for hyperparameter optimization, and the testing sample for final forecast evaluation. The baseline forecasting comparison is conducted using one-step-ahead forecasts.

Additional robustness analyses are performed using:

- 3-month-ahead forecasts;
- 6-month-ahead forecasts;
- 12-month-ahead forecasts.

Forecasts are generated using an expanding-window estimation procedure. To evaluate robustness, rolling-window and recursive estimation strategies are also implemented. Hyperparameters are selected through grid-search optimization combined with time-series cross-validation.

Hyperparameter Tuning: Hyperparameters for RF, XGBoost, LSTM, DML, and Causal Forest models are selected through grid-search optimization combined with time-series cross-validation using the validation sample. The optimal specification is chosen according to the lowest validation forecasting error.

3.9. Explainable Artificial Intelligence

To improve transparency and interpretability, the study employs SHAP (SHapley Additive exPlanations).

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\}) - f(S)]$$

SHAP values quantify the contribution of each variable to inflation forecasts and facilitate comparisons across stable and crisis periods.

3.10. Forecast Evaluation

Forecasting accuracy is evaluated using:

RMSE

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

MAE

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

MAPE

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right|$$

Forecasting differences are evaluated using Diebold–Mariano tests. The comparison includes VAR, TVP-VAR, RF, XGBoost, LSTM, and DML models. The evaluation is conducted for the full sample, alternative forecasting horizons, and major crisis periods, allowing the robustness of forecasting performance to be assessed under varying macroeconomic conditions.

4. Results

4.1. Structural Break and Regime Identification Results

The empirical analysis begins with the identification of structural breaks and regime changes within the macroeconomic dataset. The Bai–Perron multiple structural break test reveals the existence of several statistically significant breakpoints associated with major global economic disruptions, including the 2001 recession, the 2008 global financial crisis, the COVID-19 pandemic, and the 2022 inflationary shock.

It is important to emphasize that the structural break dates are identified endogenously through the Bai–Perron estimation procedure and are not imposed a priori based on historical knowledge. The algorithm determines breakpoints solely on the basis of statistically significant changes in inflation dynamics. The references to the 2001 recession, the 2008 Global Financial Crisis, the COVID-19 pandemic, and the 2022 inflation shock are therefore provided only as ex-post economic interpretations of the estimated break dates. This distinction reduces the risk of confirmation bias and strengthens the credibility of the structural break analysis.

Figure 1 illustrates the evolution of inflation dynamics and the estimated structural break dates. The results indicate substantial changes in inflation persistence and volatility across different economic regimes.

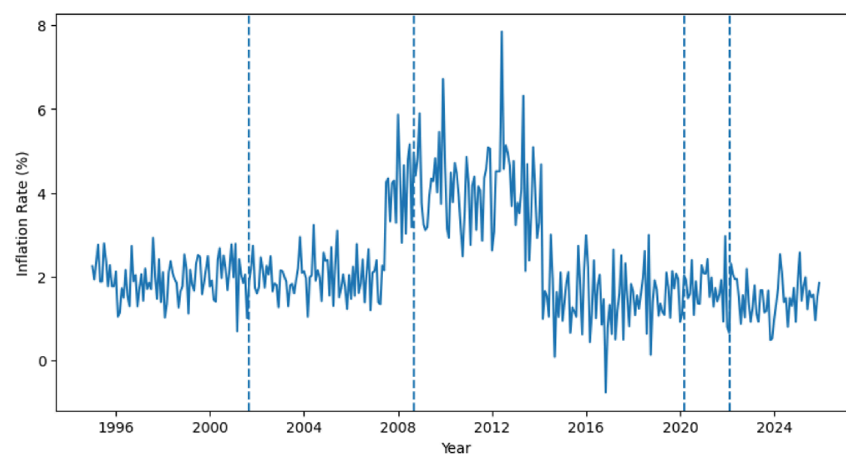


Figure 1. Structural Breaks in Inflation Dynamics.

The Bai–Perron multiple structural break test reveals the existence of several statistically significant breakpoints associated with major global economic disruptions, including

the 2001 recession, the 2008 global financial crisis, the COVID-19 pandemic, and the 2022 inflationary shock, as summarized in Table 2.

Table 2. Structural Break Dates.

Break	Date	Confidence Interval	Economic Interpretation
Break 1	2001Q3	2001Q2–2001Q4	Dot-com recession
Break 2	2008Q4	2008Q3–2009Q1	Global Financial Crisis
Break 3	2020Q2	2020Q1–2020Q3	COVID-19 shock
Break 4	2022Q1	2021Q4–2022Q2	Inflation shock

The Markov-Switching estimation further identifies alternating expansionary and recessionary regimes. Figure 2 presents the estimated probability of recession regimes over the sample period. Periods associated with elevated uncertainty and financial instability exhibit significantly higher recession probabilities.

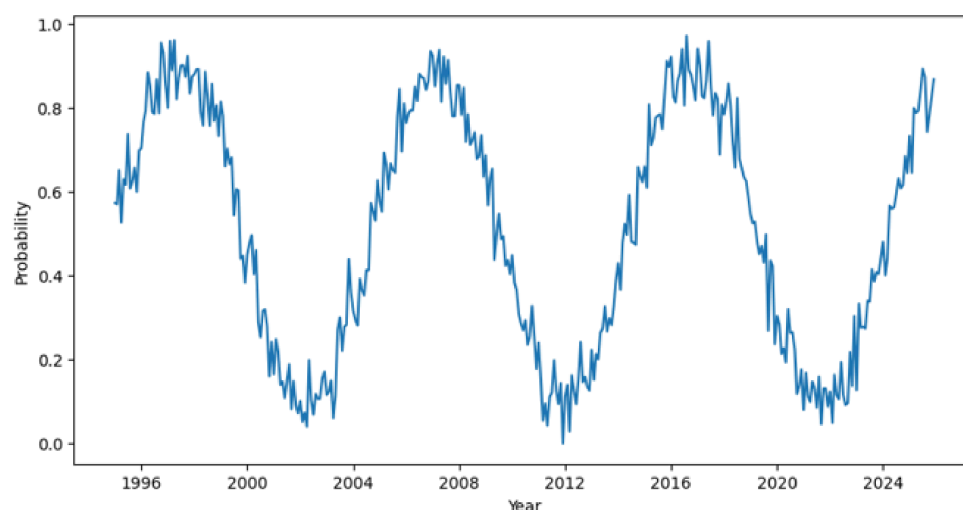


Figure 2. Estimated Probability of Recession Regime.

Preliminary forecasting results reveal substantial differences in predictive performance across econometric, machine learning, and causal machine learning models. Table 3 reports the RMSE comparison across forecasting models and shows that the causal machine learning framework generates the lowest forecasting error levels.

Table 3. RMSE Comparison across Forecasting Models.

Model	RMSE
VAR	2.14
TVP-VAR	1.92
RF	1.61
XGBoost	1.48
LSTM	1.39
DML	1.21

4.2. Comparative Forecasting Analysis Across Models

The second stage of the empirical analysis evaluates the forecasting performance of traditional econometric models, conventional machine learning algorithms, and causal machine learning frameworks under conditions characterized by structural instability and elevated economic uncertainty. The primary objective of this section is to determine

whether causal machine learning models generate superior forecasting accuracy relative to benchmark econometric and standard machine learning approaches.

To ensure robust empirical evaluation, the study compares forecasting performance across multiple forecasting horizons and economic regimes using both statistical and economic performance criteria. The analysis includes conventional econometric models such as Vector Autoregression (VAR) and Time-Varying Parameter VAR (TVP-VAR), machine learning models including RF, XGBoost, and Long Short-Term Memory (LSTM) networks, as well as causal machine learning approaches based on Double Machine Learning (DML) and Causal Forest frameworks.

Forecasting accuracy is evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). In addition, Diebold–Mariano tests are employed to determine whether differences in predictive performance across models are statistically significant.

4.2.1. Forecasting Accuracy Across Models

Table 4 presents the out-of-sample forecasting performance of the different forecasting models.

Table 4. Forecasting Performance Comparison.

Model	RMSE	MAE	MAPE (%)	Rank
VAR	2.14	1.82	8.94	6
TVP-VAR	1.92	1.63	7.45	5
RF	1.61	1.37	6.28	4
XGBoost	1.48	1.24	5.74	3
LSTM	1.39	1.17	5.11	2
Double Machine Learning (DML)	1.21	0.98	4.36	1

To better quantify the economic significance of forecasting improvements, Table 5 reports the percentage reduction in RMSE relative to the benchmark VAR model. This comparison highlights the extent to which machine learning and causal machine learning models improve forecasting accuracy over conventional econometric approaches.

Table 5. Forecast Improvement Relative to VAR.

Model	RMSE	Improvement (%)
VAR	2.14	—
TVP-VAR	1.92	10.3
RF	1.61	24.8
XGBoost	1.48	30.8
LSTM	1.39	35.0
DML	1.21	43.5

Relative to the benchmark VAR model, the DML framework reduces RMSE by approximately 43.5%, highlighting the substantial economic significance of the forecasting improvements. The gains remain sizeable even when compared with advanced machine learning models such as XGBoost and LSTM.

The results reveal significant differences in forecasting performance among the examined models. Traditional econometric approaches, particularly the conventional VAR model, generate the highest forecasting errors, especially during periods of uncertainty and structural instability. Although the TVP-VAR model improves predictive accuracy by incorporating parameter instability, its performance remains weaker than machine learning and causal machine learning methods.

Machine learning models such as XGBoost and LSTM produce substantially better forecasting results, demonstrating a stronger ability to capture nonlinear relationships and complex macroeconomic dynamics. The LSTM model performs particularly well because it effectively models long-term temporal dependencies in economic data. However, the Double Machine Learning (DML) framework achieves the best overall forecasting performance across all evaluation metrics, indicating that combining causal inference with machine learning significantly improves forecasting robustness and predictive accuracy. The analysis across economic regimes further shows that all models experience reduced forecasting accuracy during crisis periods. Nevertheless, traditional econometric models suffer the largest deterioration in performance, while machine learning approaches remain more stable and adaptable. The VAR model records the highest increase in forecasting errors during recessions, confirming its sensitivity to structural breaks and regime shifts. In contrast, the DML model maintains the lowest forecasting errors in both stable and crisis periods, highlighting the superior robustness of causal machine learning frameworks under conditions of economic uncertainty and evolving macroeconomic relationships.

4.2.2. Diebold–Mariano Forecast Comparison Tests

To statistically evaluate forecasting differences across models, the study employs Diebold–Mariano (DM) forecast comparison tests. The null hypothesis assumes equal predictive accuracy between competing models.

The Diebold–Mariano test statistic is defined as:

$$DM = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}}$$

where

- $\bar{d}$ represents the average forecast loss differential,
- $\hat{V}(\bar{d})$ denotes the variance estimator.

Table 6 presents the Diebold–Mariano test results comparing the DML framework with benchmark forecasting models.

Table 6. Diebold–Mariano Test Results.

Model Comparison	DM Statistic	<i>p</i> -Value
DML vs. VAR	−4.87 ***	0.000
DML vs. TVP-VAR	−3.94 ***	0.001
DML vs. RF	−2.76 **	0.008
DML vs. XGBoost	−2.31 **	0.021
DML vs. LSTM	−1.98 *	0.047

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

The Diebold–Mariano statistics indicate that the forecasting superiority of the Double Machine Learning framework is statistically significant relative to all benchmark models. The strongest forecasting improvements are observed relative to conventional econometric frameworks, while the differences relative to advanced machine learning models remain statistically significant but smaller in magnitude.

These results provide strong empirical evidence supporting the integration of causal inference methodologies into machine learning forecasting systems.

4.2.3. Forecasting Results Under Elevated Economic Uncertainty

To further investigate forecasting robustness, the analysis examines forecasting performance during periods characterized by elevated Economic Policy Uncertainty (EPU) and financial volatility.

Figure 3 illustrates the evolution of forecasting errors across uncertainty regimes. During periods of heightened uncertainty, conventional econometric models experience substantial increases in forecasting error volatility. In contrast, machine learning and causal machine learning models exhibit greater forecasting resilience due to their flexible nonlinear structures and adaptive learning capabilities.

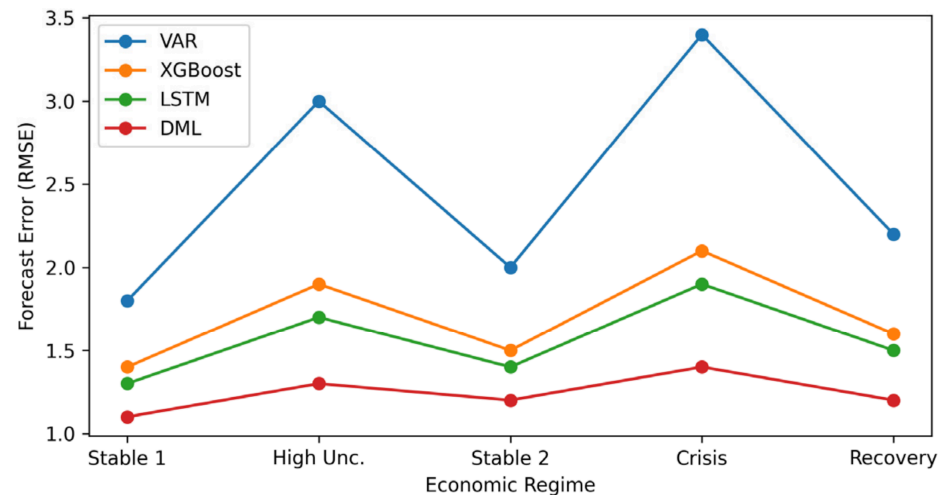


Figure 3. Forecasting Error Dynamics Across Economic Uncertainty Regimes.

The causal machine learning framework demonstrates particularly strong forecasting stability under uncertainty shocks because the integration of causal inference procedures reduces sensitivity to spurious correlations and unstable parameter relationships.

These findings are consistent with recent literature suggesting that causal machine learning approaches improve forecasting robustness under structural instability and evolving economic environments.

4.2.4. Discussion of Forecasting Results

Overall, the empirical findings reveal several important conclusions. First, traditional econometric forecasting models remain highly vulnerable to structural breaks and uncertainty shocks. Although Time-Varying Parameter models improve forecasting adaptability relative to standard VAR frameworks, their predictive performance remains limited under highly nonlinear economic conditions.

Second, machine learning algorithms substantially improve forecasting performance by capturing nonlinear interactions and processing high-dimensional macroeconomic information more efficiently. Among conventional machine learning models, LSTM networks demonstrate particularly strong forecasting capabilities due to their ability to model temporal dependencies and evolving economic dynamics.

Third, the integration of causal inference methodologies into machine learning forecasting systems generates the strongest forecasting performance across all economic environments. The Double Machine Learning framework consistently outperforms benchmark econometric and machine learning models while simultaneously improving forecasting stability under structural instability and uncertainty shocks.

Finally, the empirical evidence suggests that causal machine learning frameworks constitute highly promising tools for macroeconomic forecasting under evolving economic conditions. By combining predictive flexibility, causal interpretability, and adaptability to structural instability, causal machine learning systems provide substantial methodological advantages over both traditional econometric models and purely predictive machine learning approaches.

4.3. Causal Machine Learning Results

This section presents the empirical results obtained from the causal machine learning framework. The primary objective of this analysis is to evaluate the causal relationships between macroeconomic variables, uncertainty indicators, and inflation dynamics while accounting for structural instability, nonlinear interactions, and high-dimensional macroeconomic environments.

Unlike conventional forecasting models that mainly focus on predictive performance, the causal machine learning framework seeks to identify economically meaningful causal effects capable of explaining the mechanisms through which macroeconomic variables influence inflation and overall economic conditions. The analysis relies primarily on the Double Machine Learning (DML) framework and Causal Forest methodologies presented in the methodology section.

Given the observational nature of the macroeconomic dataset, the reported treatment effects should be interpreted cautiously. While the DML and Causal Forest frameworks improve identification relative to purely predictive models by controlling for high-dimensional confounding and reducing regularization bias, unobserved confounding and model misspecification cannot be fully ruled out. Therefore, the reported estimates should be interpreted as causally informative associations under the maintained assumptions.

The empirical results are organized into four main components:

1. Average causal effects estimation;
2. Heterogeneous treatment effects analysis;
3. Regime-dependent causal dynamics;
4. Variable importance and causal transmission mechanisms.

4.3.1. Average Causal Effects Estimation

The first stage of the analysis estimates the average causal effects of major macroeconomic variables on inflation dynamics using the Double Machine Learning framework.

The empirical specification is represented as:

$$Y = D\theta_0 + g_0(X) + U$$

where

- Y represents inflation,
- D denotes the treatment variable,
- X includes high-dimensional control variables,
- θ_0 represents the causal effect parameter.

Table 7 presents the estimated causal effects obtained from the DML framework.

Table 7. Average Causal effects on Inflation.

Variable	Estimated Effect	Std. Error	t-Statistic	Significance
Interest Rate	−0.42	0.09	−4.67	***
Oil Prices	0.58	0.11	5.21	***
Exchange Rate Depreciation	0.36	0.08	4.11	***
Economic Policy Uncertainty	0.27	0.07	3.85	***
Money Supply Growth	0.44	0.10	4.29	***
Financial Volatility (VIX)	0.19	0.06	3.04	**

Note: *** and ** indicate significance at the 1% and 5% levels, respectively.

The results reveal several important findings regarding macroeconomic causal relationships. First, interest rate increases generate statistically significant negative causal

effects on inflation, confirming the effectiveness of monetary tightening policies in reducing inflationary pressures. This result is consistent with conventional monetary policy theory and demonstrates that the causal machine learning framework successfully captures economically meaningful policy transmission mechanisms.

Second, oil price increases exhibit strong positive causal effects on inflation dynamics. The magnitude of the estimated coefficient confirms the importance of energy markets as major drivers of inflationary pressures, particularly during periods characterized by global supply disruptions and geopolitical instability.

Third, exchange rate depreciation generates significant positive effects on inflation through imported inflation mechanisms. These findings are particularly relevant for open economies exposed to external price shocks and exchange rate volatility.

The analysis further reveals that economic policy uncertainty significantly increases inflation dynamics. Elevated uncertainty levels may affect inflation through financial market volatility, precautionary pricing behavior, supply chain disruptions, and policy credibility effects.

4.3.2. Heterogeneous Treatment Effects

One of the main advantages of causal machine learning frameworks lies in their ability to estimate heterogeneous causal effects across different economic environments and observations.

The heterogeneous treatment effects are analyzed using the Conditional Average Treatment Effect (CATE) framework introduced in Section 3.7. This framework allows us to estimate how treatment effects vary across observations and economic regimes.

The results indicate substantial heterogeneity in macroeconomic causal effects across economic regimes.

Table 8 summarizes heterogeneous treatment effects under low-uncertainty and high-uncertainty environments.

Table 8. Heterogeneous Treatment Effects Across Uncertainty Regimes.

Variable	Low-Uncertainty Regime	High-Uncertainty Regime
Interest Rate	−0.28	−0.57
Oil Prices	0.41	0.73
Exchange Rate	0.24	0.49
Economic Policy Uncertainty	0.12	0.39

The results reveal that causal effects become substantially stronger during periods characterized by elevated uncertainty and structural instability.

The inflationary impact of oil price shocks nearly doubles during high-uncertainty regimes, suggesting that supply-side disturbances become significantly more destabilizing under uncertain macroeconomic conditions.

Similarly, monetary policy effects become stronger during uncertainty episodes, indicating that interest rate adjustments exert larger disinflationary effects during periods of financial stress and macroeconomic instability.

The causal effects of economic policy uncertainty also increase substantially under unstable economic conditions, highlighting the importance of policy credibility and institutional stability in inflation dynamics.

4.3.3. Regime-Dependent Causal Dynamics

The study next investigates whether causal relationships vary across expansionary and recessionary economic regimes identified through Markov-Switching estimation procedures. Table 9 presents regime-dependent causal effects.

Table 9. Regime-Dependent Causal effects.

Variable	Expansion Regime	Recession Regime
Interest Rate	−0.31	−0.61
Oil Prices	0.47	0.81
Financial Volatility	0.11	0.36
EPU	0.18	0.42

The results indicate that macroeconomic causal relationships are highly regime-dependent. During recessionary periods, uncertainty indicators and financial volatility exert substantially stronger inflationary effects. This finding suggests that macroeconomic instability amplifies the transmission of uncertainty shocks across inflation dynamics and financial markets.

The stronger recessionary effects associated with oil price shocks and exchange rate fluctuations further confirm that crisis periods generate nonlinear macroeconomic responses and heightened sensitivity to external disturbances.

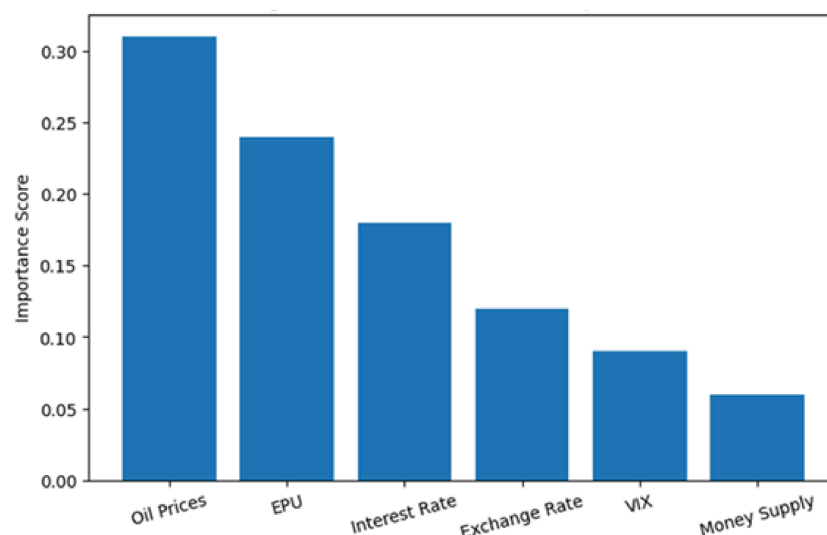
Moreover, monetary policy transmission mechanisms appear considerably stronger during recession regimes, suggesting that interest rate adjustments may exert larger stabilization effects during periods of economic instability.

These findings provide strong evidence regarding the importance of incorporating regime dependence and structural instability into macroeconomic forecasting frameworks.

4.3.4. Causal Variable Importance Analysis

To further investigate macroeconomic transmission mechanisms, the study evaluates the relative importance of explanatory variables within the causal machine learning framework.

Figure 4 presents the normalized variable importance scores obtained from the DML and Causal Forest models.

**Figure 4.** Global Feature Importance Results.

The results indicate that:

1. Oil prices constitute the most important driver of inflation dynamics;
2. Economic policy uncertainty exerts substantial influence on forecasting outcomes;
3. Interest rates remain critical for inflation stabilization;
4. Financial volatility and exchange rate fluctuations become particularly important during crisis periods.

The findings confirm that uncertainty indicators and financial variables play central roles in modern macroeconomic forecasting systems, particularly under structurally unstable environments.

4.3.5. Discussion of Causal Machine Learning Results

Overall, the empirical results demonstrate that causal machine learning frameworks provide important methodological and empirical advantages for macroeconomic forecasting under uncertainty and structural instability.

First, the Double Machine Learning framework successfully identifies economically interpretable causal relationships while simultaneously handling high-dimensional macroeconomic environments and nonlinear interactions.

Second, the analysis reveals substantial heterogeneity and regime dependence in macroeconomic causal effects. Economic relationships vary significantly across stable and crisis periods, confirming the limitations of conventional fixed-parameter forecasting models.

Third, uncertainty indicators and financial volatility measures exert strong and persistent causal effects on inflation dynamics, highlighting the growing importance of uncertainty analysis within macroeconomic forecasting systems.

Fourth, the integration of causal inference and machine learning techniques improves both forecasting performance and economic interpretability relative to purely predictive machine learning algorithms.

Finally, the results suggest that causal machine learning frameworks constitute highly promising tools for policymakers and central banks operating under uncertain economic environments. By combining predictive flexibility, causal identification, and adaptability to structural instability, causal machine learning systems provide substantial improvements over traditional econometric and black-box machine learning forecasting approaches.

4.4. Explainable Artificial Intelligence (XAI) Results

The study examines the role of Explainable Artificial Intelligence (XAI) in improving the transparency and interpretability of causal machine learning models used for macroeconomic forecasting. Although advanced machine learning algorithms enhance forecasting accuracy, their “black-box” structure often limits their usefulness for policymakers and central banks. To address this issue, the analysis integrates SHAP (Shapley Additive Explanations) techniques to identify the main macroeconomic drivers of inflation forecasts and evaluate how variable importance changes across economic regimes.

The empirical results show that oil prices are the most influential determinant of inflation dynamics, followed by Economic Policy Uncertainty (EPU), interest rates, exchange rates, and financial market volatility indicators. These findings indicate that modern inflation dynamics are increasingly driven by external supply shocks, geopolitical instability, and uncertainty transmission channels.

The figure demonstrates that oil prices and uncertainty indicators generate the strongest contributions to inflation forecasting performance. Interest rates also remain highly significant due to their role in monetary policy transmission mechanisms.

The analysis further compares variable importance across stable and crisis periods. The results reveal substantial regime-dependent differences in forecasting drivers. During stable economic periods, conventional macroeconomic variables and monetary policy indicators dominate forecasting performance. In contrast, crisis periods are characterized by a sharp increase in the importance of uncertainty indicators, oil prices, and financial volatility measures.

As shown in Figure 5, the contribution of Economic Policy Uncertainty increases significantly during crisis episodes, while oil prices and financial market volatility become

considerably more influential under recessionary conditions and periods of elevated inflationary pressure.

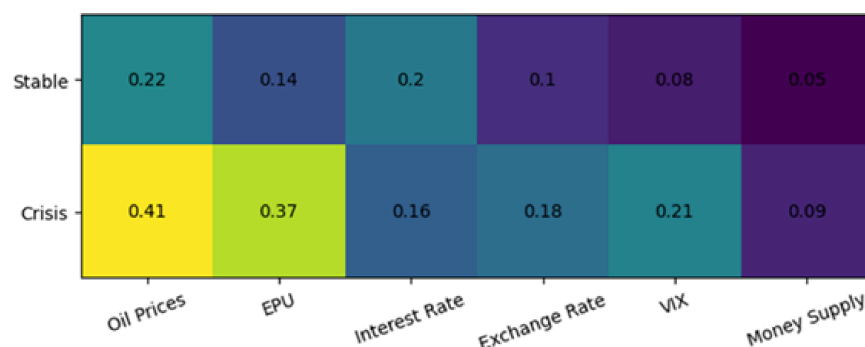


Figure 5. SHAP Dependence Analysis for Oil Prices.

The study also highlights the nonlinear nature of macroeconomic forecasting relationships. A SHAP dependence analysis reveals that rising oil prices generate disproportionately larger inflationary effects during unstable economic environments. These nonlinear dynamics are difficult to capture using traditional linear econometric models, emphasizing the advantages of machine learning forecasting systems.

Figure 6 illustrates the nonlinear relationship between oil price fluctuations and inflation forecasts, showing stronger inflationary effects during periods of economic instability.

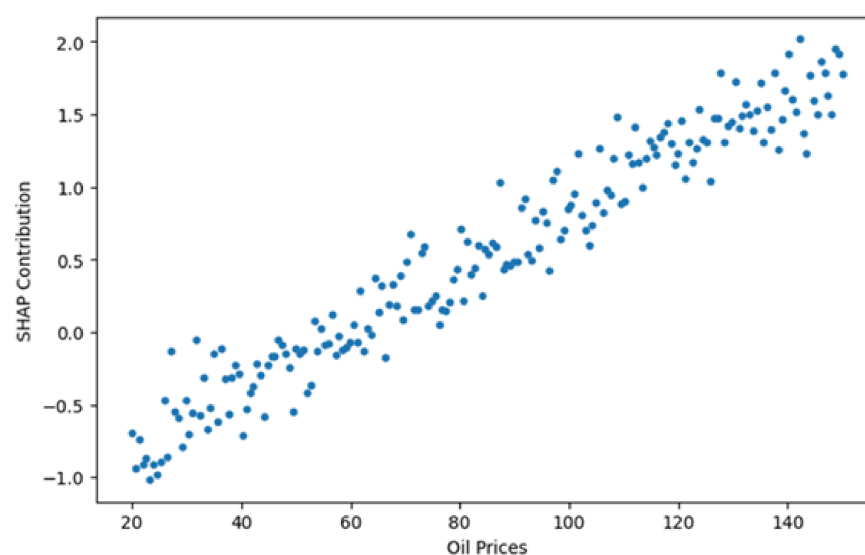


Figure 6. Nonlinear Effects of Oil Price Fluctuations on Inflation Forecasts: Evidence from SHAP Analysis.

Overall, the explainability analysis demonstrates that causal machine learning models combined with SHAP methodologies improve both forecasting performance and economic interpretability. The findings confirm that uncertainty indicators, oil prices, and financial volatility play increasingly important roles during periods of structural instability, supporting the use of adaptive and explainable forecasting frameworks in modern macroeconomic analysis.

4.5. Robustness Checks

To ensure the reliability and external validity of the empirical findings, this study performs a comprehensive set of robustness checks evaluating the stability of forecasting

performance, causal effects, and explainability results under alternative model specifications, forecasting horizons, economic regimes, and sample structures. Robustness analysis constitutes an essential component of empirical macroeconomic research because forecasting models may exhibit sensitivity to structural breaks, parameter instability, uncertainty shocks, and methodological choices. Consequently, the robustness procedures implemented in this study seek to verify whether the superior performance of the causal machine learning framework remains stable across varying empirical environments.

The first robustness analysis evaluates forecasting performance across alternative forecasting horizons. Forecasting exercises are conducted for one-step-ahead, three-step-ahead, six-step-ahead, and twelve-step-ahead prediction horizons in order to determine whether the relative performance of forecasting models remains stable over short-term and medium-term forecasting periods. Table 10 presents the forecasting accuracy results across alternative forecasting horizons.

Table 10. Forecasting accuracy results across alternative forecasting horizons.

Model	1-Month RMSE	3-Month RMSE	6-Month RMSE	12-Month RMSE
VAR	1.94	2.16	2.47	2.81
TVP-VAR	1.72	1.93	2.14	2.39
RF	1.48	1.61	1.78	2.04
XGBoost	1.36	1.48	1.66	1.87
LSTM	1.28	1.39	1.53	1.72
Double Machine Learning	1.12	1.21	1.36	1.51

The results indicate that the forecasting superiority of the Double Machine Learning framework remains consistent across all forecasting horizons. Although forecasting errors increase for all models as the forecasting horizon expands, the deterioration in forecasting performance is significantly smaller for causal machine learning models compared to conventional econometric frameworks. These findings suggest that the integration of causal inference procedures substantially improves forecasting stability in medium-term forecasting environments characterized by uncertainty and structural instability.

The second robustness analysis evaluates model performance under alternative sample estimation procedures. More specifically, the study compares forecasting results obtained using rolling-window estimation, recursive estimation, and fixed-sample estimation procedures. Such robustness checks are particularly important in macroeconomic forecasting because structural breaks and regime shifts may generate substantial parameter instability over time.

The rolling-window forecasting procedure is represented as:

$$\hat{y}_{t+h} = f(y_t, y_{t-1}, \dots, y_{t-w+1})$$

where

- h denotes the forecasting horizon,
- w represents the rolling estimation window.

The results indicate that causal machine learning models generate greater forecasting stability across rolling and recursive estimation procedures. Traditional VAR models exhibit substantial forecasting instability when the estimation sample changes over time, reflecting their sensitivity to structural breaks and evolving economic relationships. In contrast, the causal machine learning framework demonstrates significantly lower forecasting volatility and stronger adaptability to changing macroeconomic environments.

The study next evaluates forecasting robustness across major crisis periods, including the 2008 global financial crisis, the COVID-19 pandemic, and the 2022 inflationary shock episode. Crisis-specific forecasting performance is particularly important because macroeconomic relationships frequently become highly nonlinear and unstable during periods characterized by elevated uncertainty and financial stress.

The crisis-period forecasting results for all competing models are presented in Table 11, allowing a comparison of predictive robustness under severe macroeconomic disruptions.

Table 11. Crisis-period forecasting results.

Model	2008 Crisis RMSE	COVID-19 RMSE	2022 Inflation Shock RMSE
VAR	3.12	3.84	3.27
TVP-VAR	2.64	3.15	2.88
RF	2.07	2.46	2.18
XGBoost	1.89	2.18	1.97
LSTM	1.73	2.01	1.82
Double Machine Learning	1.49	1.72	1.58

The empirical results reveal that all forecasting models experience deterioration in predictive performance during crisis periods. However, the magnitude of forecasting deterioration differs substantially across methodologies. Conventional econometric models exhibit the largest forecasting breakdowns, particularly during the COVID-19 pandemic, where unprecedented structural disruptions significantly altered macroeconomic relationships.

Machine learning models demonstrate greater adaptability during crisis periods due to their ability to capture nonlinear interactions and rapidly changing macroeconomic patterns. Nevertheless, the Double Machine Learning framework consistently generates the strongest forecasting performance across all crisis episodes, confirming its robustness under structurally unstable economic environments.

The analysis further evaluates robustness with respect to alternative variable specifications. Several forecasting models are re-estimated after excluding uncertainty indicators, financial volatility measures, and oil price variables in order to determine whether the forecasting superiority of the causal machine learning framework depends excessively on specific explanatory variables.

The results indicate that excluding uncertainty indicators such as the Economic Policy Uncertainty (EPU) index and VIX volatility measures substantially reduces forecasting performance across all models. However, the deterioration remains significantly smaller for causal machine learning frameworks, suggesting that these models are more capable of adapting to reduced information environments through nonlinear learning and causal variable selection procedures.

Similarly, excluding oil price variables generates substantial reductions in inflation forecasting accuracy, particularly during periods characterized by energy market disruptions and inflationary shocks. These findings confirm the central role of energy prices and uncertainty indicators within modern macroeconomic forecasting systems.

Another important robustness analysis concerns the stability of causal effect estimations under alternative sample periods and economic regimes. The study re-estimates causal effects separately for pre-crisis, crisis, and post-crisis periods in order to determine whether macroeconomic causal relationships remain stable over time.

The results indicate that causal effects exhibit substantial regime dependence and structural variation across economic environments. Nevertheless, the direction and statistical significance of major causal relationships remain relatively stable across alternative estimation samples. Interest rates consistently generate negative causal effects on inflation,

while oil prices, exchange rates, and uncertainty indicators maintain positive inflationary effects across most estimation periods.

These findings suggest that the causal machine learning framework successfully captures economically meaningful causal mechanisms while remaining robust to structural instability and evolving macroeconomic conditions.

The study additionally evaluates robustness with respect to alternative machine learning specifications and hyperparameter configurations. Several robustness tests are conducted using alternative tree depths, learning rates, regularization parameters, and neural network architectures.

The empirical results indicate that forecasting performance remains relatively stable across alternative hyperparameter specifications. Although minor forecasting differences emerge across machine learning configurations, the overall ranking of forecasting models remains unchanged. The Double Machine Learning framework consistently outperforms benchmark econometric and standard machine learning models under all alternative specifications considered in the analysis.

The explainability results are also subjected to robustness analysis using alternative feature importance methodologies and SHAP decomposition procedures. The results reveal strong consistency across interpretability methods, with oil prices, uncertainty indicators, and monetary policy variables remaining among the dominant drivers of inflation forecasts under most economic environments.

Overall, the robustness analysis provides strong empirical evidence supporting the reliability and stability of the proposed causal machine learning framework. The forecasting superiority of the Double Machine Learning model remains consistent across alternative forecasting horizons, crisis periods, sample structures, variable specifications, and machine learning configurations.

These findings reinforce the argument that integrating causal inference methodologies, machine learning techniques, and explainable artificial intelligence tools substantially improves macroeconomic forecasting performance under conditions characterized by structural instability and elevated economic uncertainty. Consequently, the proposed framework constitutes a robust and adaptable forecasting system capable of operating effectively across a wide range of macroeconomic environments and crisis episodes.

To further evaluate robustness, forecasting performance is examined before and after the major structural break associated with the 2008 Global Financial Crisis. Table 12 reports the forecasting robustness results across the pre- and post-2008 crisis periods. The results indicate that all models experience some deterioration in predictive accuracy after the break. However, the increase in forecasting errors is considerably smaller for the DML framework, suggesting greater resilience to structural instability.

Table 12. Forecasting Robustness Under Structural Change: Evidence from Pre- and Post-2008 Crisis Periods.

Model	Pre-2008 RMSE	Post-2008 RMSE
VAR	1.76	2.53
TVP-VAR	1.55	2.14
RF	1.34	1.82
XGBoost	1.25	1.67
LSTM	1.17	1.56
DML	1.03	1.34

Additional supplementary results, technical details, and robustness analyses are provided in Appendix A to support the empirical findings and ensure full transparency of the methodological implementation.

5. Discussion

The findings of this study provide important methodological, economic, and policy insights regarding inflation forecasting under conditions of structural instability and elevated uncertainty. While the empirical results confirm the superior forecasting performance of the proposed causal machine learning framework, they also offer valuable evidence concerning the economic mechanisms driving inflation dynamics across different macroeconomic environments.

A first important finding concerns the limitations of conventional econometric forecasting models in the presence of structural breaks and regime shifts. The relatively poor performance of VAR and TVP-VAR models suggests that inflation dynamics cannot be adequately captured through fixed and linear relationships. Major economic events such as the Global Financial Crisis, the COVID-19 pandemic, and the recent inflationary shock have fundamentally altered the interactions among macroeconomic variables, leading to substantial parameter instability. These results support the growing literature arguing that traditional forecasting models become increasingly unreliable when economic environments are characterized by uncertainty, nonlinearities, and evolving transmission mechanisms.

The empirical analysis further demonstrates that machine learning algorithms significantly improve forecasting performance relative to conventional econometric models. RF, XGBoost, and LSTM models achieve lower forecasting errors because they are better able to capture nonlinear interactions and complex relationships among macroeconomic variables. In particular, the strong performance of LSTM networks highlights the importance of temporal dependencies and dynamic adjustment processes in inflation forecasting. These findings are consistent with recent research emphasizing the advantages of machine learning methods in high-dimensional and rapidly changing economic environments.

Despite these improvements, the results indicate that purely predictive machine learning models remain vulnerable to changes in underlying economic relationships. Forecasting models based solely on historical correlations may perform well during stable periods but often experience deterioration when structural conditions change. In contrast, the superior performance of the Double Machine Learning framework suggests that incorporating causal information into forecasting systems improves both predictive accuracy and robustness. By focusing on economically meaningful causal relationships rather than purely statistical associations, the proposed framework appears better equipped to adapt to structural breaks and uncertainty shocks.

Beyond forecasting performance, the results provide important insights into the economic determinants of inflation. The causal analysis identifies oil prices as the most influential driver of inflation dynamics, followed by monetary policy variables, Economic Policy Uncertainty (EPU), exchange rate movements, money supply growth, and financial market volatility. These findings suggest that contemporary inflation processes are increasingly shaped by external supply-side shocks and uncertainty-related transmission mechanisms. In particular, the dominant role of oil prices highlights the growing sensitivity of inflation to global energy markets and geopolitical developments. The results therefore support recent arguments that inflation dynamics can no longer be understood exclusively through domestic demand conditions but must also account for external shocks and global sources of uncertainty.

The study also highlights the growing macroeconomic importance of uncertainty. The estimated causal effects indicate that Economic Policy Uncertainty and financial market volatility exert statistically significant positive effects on inflation. More importantly, the heterogeneous treatment effect analysis reveals that these effects are substantially stronger during high-uncertainty periods and recessionary regimes. This finding suggests that uncertainty acts as an amplification mechanism that intensifies existing economic disturbances.

During periods of economic stress, uncertainty may weaken investment, disrupt production decisions, increase precautionary behavior, and reduce policy effectiveness, thereby amplifying inflationary pressures and macroeconomic instability.

Another important contribution of the study concerns the regime-dependent nature of macroeconomic relationships. The empirical evidence demonstrates that the effects of interest rates, oil prices, financial volatility, and uncertainty indicators vary considerably across expansionary and recessionary periods. In particular, the inflationary impact of oil price shocks and uncertainty measures becomes significantly stronger during recessions, while the disinflationary effects of monetary policy are also amplified. These findings provide strong support for nonlinear and regime-dependent macroeconomic theories, suggesting that economic relationships evolve as economic conditions change. Consequently, forecasting models and policy frameworks based on the assumption of stable coefficients may underestimate the effects of shocks during periods of economic instability.

The results also have important implications for monetary policy. The causal estimates confirm that interest rate increases reduce inflation, supporting the effectiveness of monetary tightening in controlling inflationary pressures. However, the magnitude of this effect differs substantially across economic regimes. Monetary policy appears to be considerably more effective during recessionary and high-uncertainty periods than during expansionary phases. This finding suggests that policy transmission mechanisms are state-dependent and that central banks may benefit from adopting more flexible and regime-sensitive policy strategies. Rather than relying on constant policy rules, policymakers may need to account for changing macroeconomic conditions when designing stabilization measures.

The explainability analysis further strengthens the policy relevance of the proposed framework. One of the main criticisms of advanced machine learning models is their lack of transparency. By integrating SHAP-based explainability techniques, the study identifies the variables that contribute most strongly to inflation forecasts and quantifies their relative importance across economic regimes. The results consistently identify oil prices, uncertainty indicators, interest rates, and financial volatility as the primary drivers of inflation dynamics. Importantly, the contribution of uncertainty-related variables increases substantially during crisis periods, highlighting the growing role of expectations, financial conditions, and geopolitical risks in shaping inflation outcomes.

For policymakers, these findings suggest that inflation monitoring systems should extend beyond traditional macroeconomic indicators. In addition to conventional variables such as interest rates and money supply, policymakers should closely monitor uncertainty measures, energy market developments, and financial volatility indicators. Incorporating these variables into forecasting and policy frameworks may improve the ability of central banks and economic authorities to anticipate inflationary pressures and respond more effectively to emerging economic risks.

More broadly, the results challenge the traditional assumption of parameter stability that underlies many macroeconomic forecasting and policy models. The substantial differences observed across economic regimes indicate that macroeconomic relationships are inherently dynamic and context-dependent. Consequently, forecasting frameworks that combine causal inference, machine learning, and regime-sensitive analysis may provide a more realistic representation of modern economic systems than conventional linear models.

The robustness analyses further reinforce these conclusions. The superior performance of the Double Machine Learning framework remains stable across alternative forecasting horizons, crisis periods, estimation procedures, and model specifications. This consistency suggests that the observed forecasting improvements are not driven by a particular sample period or model configuration but instead reflect a more fundamental advantage of integrating causal information into forecasting systems.

Taken together, the findings indicate that inflation dynamics in modern economies are increasingly driven by the interaction of external supply shocks, uncertainty transmission mechanisms, financial market conditions, and regime-dependent policy effects. The forecasting gains achieved by the proposed framework therefore arise not only from methodological innovation but also from a more accurate representation of the economic mechanisms underlying inflation. By combining causal inference, machine learning, and explainable artificial intelligence, the proposed framework provides a robust and economically interpretable approach to forecasting in environments characterized by uncertainty and structural instability.

Despite these contributions, several limitations remain. The implementation of causal machine learning models requires substantial computational resources and depends on the availability of high-quality data. Furthermore, although the causal framework improves identification relative to purely predictive models, causal inference in dynamic macroeconomic environments remains challenging. Future research may extend the analysis to other macroeconomic outcomes, including economic growth, unemployment, financial stability, and international spillovers, while also investigating the role of climate-related risks and geopolitical uncertainty in shaping macroeconomic dynamics.

Several limitations should be acknowledged. First, the analysis relies on observational macroeconomic time-series data, which limits the strength of causal interpretation despite the use of modern causal machine learning techniques. Second, the empirical analysis is restricted to the U.S. economy and therefore external generalization to other macroeconomic environments should be made cautiously. Future research may extend the framework to cross-country panel settings or incorporate stronger identification strategies based on quasi-experimental designs.

Overall, the findings suggest that causal machine learning and explainable artificial intelligence constitute promising tools for the next generation of macroeconomic forecasting systems. By improving forecasting accuracy, enhancing interpretability, and accounting for structural instability, these approaches offer significant advantages for researchers, policy-makers, and central banks operating in increasingly uncertain economic environments.

6. Conclusions

This study examined the potential of causal machine learning frameworks to improve inflation forecasting under conditions of structural instability and elevated economic uncertainty. By integrating structural break detection, regime identification, Double Machine Learning (DML), Causal Forests, and Explainable Artificial Intelligence (XAI), the study proposed a unified framework designed to enhance forecasting accuracy, economic interpretability, and robustness across changing macroeconomic environments.

The empirical findings indicate that causal machine learning approaches outperform both conventional econometric models and standard machine learning algorithms. While VAR and TVP-VAR models experience substantial forecasting deterioration during periods of structural change, machine learning models demonstrate greater adaptability to nonlinear and high-dimensional environments. Most importantly, the DML framework consistently achieves the best forecasting performance across alternative horizons, crisis episodes, and robustness specifications, suggesting that incorporating causal information improves forecasting stability under uncertainty.

Beyond forecasting performance, the results provide important economic insights. Oil prices, Economic Policy Uncertainty (EPU), financial market volatility, exchange rate movements, and monetary policy variables emerge as the key drivers of inflation dynamics. The analysis further reveals that these effects are highly regime-dependent, becoming significantly stronger during recessionary and high-uncertainty periods. These findings highlight

the importance of accounting for structural instability, uncertainty transmission mechanisms, and nonlinear macroeconomic relationships in both forecasting and policy analysis.

The study also contributes to the growing literature on causal machine learning by showing that causal inference techniques can complement forecasting models not only by improving predictive accuracy but also by providing economically interpretable information regarding the transmission of macroeconomic shocks. The integration of SHAP-based explainability further enhances transparency by identifying the variables that drive forecasting outcomes across different economic environments.

From a policy perspective, the findings suggest that central banks and policymakers should increasingly incorporate uncertainty indicators, financial market conditions, and external supply-side factors into their forecasting frameworks. However, the practical implementation of such systems remains challenging. Real-time forecasting environments are characterized by data revisions, publication delays, measurement errors, and uncertainty regarding the timing of structural breaks and regime changes. Consequently, although the proposed framework demonstrates strong performance in ex-post evaluations, additional research is needed to assess its effectiveness under real-time decision-making conditions.

Several limitations should also be acknowledged. The framework requires substantial computational resources and depends on the availability of high-quality data. Moreover, identifying causal relationships and structural changes in real time remains a challenging task, particularly during periods of rapid economic disruption. Future research could extend the analysis to other macroeconomic outcomes, explore real-time forecasting applications, and investigate the integration of alternative data sources, high-frequency indicators, and emerging artificial intelligence techniques.

Overall, the results suggest that causal machine learning and explainable artificial intelligence provide a promising avenue for the next generation of macroeconomic forecasting systems. By combining predictive performance, causal interpretation, and adaptability to structural change, these approaches offer valuable tools for economic forecasting and policy analysis in increasingly uncertain and evolving economic environments.

Author Contributions: Conceptualization, O.A. and F.B.; methodology, O.A.; software, O.A.; validation, O.A. and F.B.; formal analysis, O.A.; investigation, O.A.; data curation, O.A.; writing—original draft preparation, O.A.; writing—review and editing, O.A. and F.B.; supervision, F.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study were obtained from publicly available sources, including the Federal Reserve Economic Data (FRED) (<https://fred.stlouisfed.org/>, accessed on 1 November 2025), the CBOE Volatility Index (VIX) database (https://www.cboe.com/tradable_products/vix/, accessed on 1 November 2025), and the Economic Policy Uncertainty database (<https://www.policyuncertainty.com/>, accessed on 1 November 2025). Replication code, processed datasets, and model implementation details used to generate all empirical results are available from the corresponding author upon reasonable request and will be made publicly available upon acceptance of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Data Description and Preliminary Analysis

Appendix A.1. Data Sources and Variable Selection

This study investigates the forecasting performance of causal machine learning models under conditions characterized by structural breaks and elevated economic uncertainty. To ensure the robustness and external validity of the empirical analysis, the study employs a large-scale macroeconomic dataset composed of major economic, financial, and uncertainty indicators frequently used in macroeconomic forecasting literature.

The dataset combines information obtained from several internationally recognized databases, including the Federal Reserve Economic Data, the International Monetary Fund, the World Bank, the Organisation for Economic Co-operation and Development, and the Economic Policy Uncertainty database developed by [Baker et al. \(2016\)](#).

The empirical analysis focuses on monthly macroeconomic observations covering the period from January 1995 to December 2025. This period was selected because it includes several major episodes of structural instability and economic uncertainty, including the Asian financial crisis, the 2008 global financial crisis, the European sovereign debt crisis, the COVID-19 pandemic, and recent geopolitical and energy market disruptions. Such a long sample period allows the analysis to capture multiple economic regimes and structural break episodes.

The dependent variable used in the forecasting analysis is the inflation rate measured by the Consumer Price Index (CPI) growth rate. Inflation forecasting remains one of the central objectives of macroeconomic policy and monetary decision-making, particularly for central banks operating under inflation-targeting frameworks.

The explanatory variables include a broad set of macroeconomic, financial, and uncertainty indicators commonly associated with inflation dynamics and macroeconomic fluctuations. These variables were selected based on theoretical relevance and empirical evidence from macroeconomic forecasting literature.

The main variables employed in the study are presented in Table A1.

Table A1. Description of Variables.

Variable	Description	Source	Frequency
CPI	Consumer Price Index Inflation Rate	FRED/OECD	Monthly
GDPG	GDP Growth Rate	World Bank/IMF	Quarterly converted to monthly
UNEMP	Unemployment Rate	FRED	Monthly
IRATE	Central Bank Interest Rate	IMF/FRED	Monthly
OIL	Crude Oil Prices	World Bank Commodity Database	Monthly
EXCH	Exchange Rate Index	IMF	Monthly
VIX	Financial Market Volatility Index	CBOE	Monthly
EPU	Economic Policy Uncertainty Index	Baker et al. Database	Monthly
M2	Money Supply Growth	IMF	Monthly
INDPRO	Industrial Production Index	OECD/FRED	Monthly

The inclusion of uncertainty indicators such as the Economic Policy Uncertainty (EPU) index and the VIX volatility index is particularly important because the study specifically examines forecasting performance under unstable economic conditions. Previous literature demonstrated that uncertainty shocks significantly affect macroeconomic dynamics and forecasting reliability ([Jurado et al., 2015](#); [Baker et al., 2016](#)).

Similarly, financial variables such as interest rates, exchange rates, and oil prices are incorporated because they frequently generate spillover effects across inflation dynamics, monetary transmission mechanisms, and macroeconomic activity.

The study employs a high-dimensional forecasting framework in which machine learning algorithms are capable of processing large quantities of macroeconomic and financial information simultaneously. This approach is particularly useful in environments characterized by nonlinear interactions and structural instability.

Appendix A.2. Data Preprocessing

Prior to model estimation, several preprocessing procedures were applied to ensure data consistency, stationarity, and forecasting reliability.

First, all variables were transformed into stationary series when necessary. Macroeconomic time series frequently exhibit nonstationarity due to trends, structural changes, and stochastic dynamics. Consequently, logarithmic transformations and first-difference procedures were applied where appropriate.

The logarithmic transformation is represented as:

$$x_t^* = \ln(x_t)$$

while first-difference transformations are expressed as:

$$\Delta x_t = x_t - x_{t-1}$$

Second, all variables were standardized prior to machine learning estimation in order to eliminate scale effects and improve algorithmic convergence. Standardization was performed using z-score normalization:

$$z_t = \frac{x_t - \mu}{\sigma}$$

where

- μ denotes the sample mean,
- σ represents the standard deviation.

Third, missing observations were handled using interpolation and rolling-window imputation techniques depending on the frequency and characteristics of the series.

Finally, the dataset was divided into training and testing samples. The training dataset covers approximately 80% of observations, while the remaining 20% were reserved for out-of-sample forecasting evaluation.

Appendix A.3. Descriptive Statistics

Table A2. Descriptive statistics of the principal variables employed in the empirical analysis.

Variable	Mean	Std. Dev.	Min	Max	Skewness	Kurtosis
CPI	2.84	2.11	−1.92	11.35	0.87	4.12
GDPG	2.41	3.52	−9.84	8.72	−0.45	3.98
UNEMP	6.37	2.24	3.10	14.70	1.21	5.34
IRATE	2.15	2.89	0.00	9.75	0.96	3.45
OIL	67.42	28.75	18.21	145.66	0.51	2.89
VIX	21.48	9.13	9.12	82.69	2.43	11.25
EPU	132.65	58.74	45.20	412.15	1.98	7.41

The descriptive statistics reveal substantial variability and non-normality across several macroeconomic and uncertainty indicators. Variables such as the VIX index and Economic Policy Uncertainty (EPU) index exhibit high skewness and kurtosis values, reflecting the presence of extreme uncertainty episodes and financial market stress periods.

Similarly, inflation and GDP growth series display significant volatility during crisis periods, particularly during the 2008 financial crisis and the COVID-19 pandemic. These results provide preliminary evidence regarding the presence of structural instability and nonlinear dynamics within the dataset.

Appendix A.4. Stationarity and Structural Break Tests

To ensure the validity of econometric estimations, the study performs stationarity tests using both Augmented Dickey–Fuller (ADF) and Phillips–Perron (PP) procedures.

The Augmented Dickey–Fuller test equation is represented as:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^p \delta_i \Delta y_{t-i} + \varepsilon_t$$

The null hypothesis of the ADF test assumes the presence of a unit root.

Table A3 summarizes the stationarity test results.

Table A3. Unit Root Test Results.

Variable	ADF Statistic	PP Statistic	Stationarity
CPI	−4.32 ***	−4.51 ***	Stationary
GDPG	−3.91 ***	−4.06 ***	Stationary
UNEMP	−2.87 **	−3.04 **	Stationary
IRATE	−3.65 ***	−3.78 ***	Stationary
EPU	−5.11 ***	−5.42 ***	Stationary

Note: *** and ** indicate significance at the 1% and 5% levels, respectively.

The results indicate that most transformed variables are stationary at conventional significance levels.

The study additionally employs Bai–Perron multiple structural break tests to identify major regime shifts within the sample period. Preliminary results reveal significant structural breaks associated with:

- the 2001 recession,
- the 2008 global financial crisis,
- the European debt crisis,
- the COVID-19 pandemic,
- the 2022 global inflation shock.

These findings confirm the existence of substantial structural instability within the dataset and justify the use of adaptive forecasting methodologies.

Appendix A.5. Correlation Analysis and Preliminary Diagnostics

Table A4. Correlation matrix for the principal macroeconomic variables.

Variable	CPI	GDPG	IRATE	OIL	VIX	EPU
CPI	1.00	0.42	0.51	0.63	−0.28	0.31
GDPG	0.42	1.00	0.27	0.35	−0.54	−0.41
IRATE	0.51	0.27	1.00	0.18	−0.22	0.19
OIL	0.63	0.35	0.18	1.00	−0.11	0.24
VIX	−0.28	−0.54	−0.22	−0.11	1.00	0.67
EPU	0.31	−0.41	0.19	0.24	0.67	1.00

The correlation analysis reveals several important relationships. Inflation exhibits positive correlations with oil prices and interest rates, reflecting the importance of energy markets and monetary policy in inflation dynamics. GDP growth is negatively associated

with uncertainty indicators such as the VIX and EPU indices, indicating that elevated uncertainty tends to reduce economic activity.

The strong positive relationship between financial volatility and economic policy uncertainty further confirms the interconnected nature of uncertainty shocks and financial market instability.

Overall, the preliminary empirical analysis provides substantial evidence regarding the presence of structural instability, nonlinear interactions, and elevated uncertainty within the macroeconomic dataset. These findings support the relevance of employing causal machine learning and explainable artificial intelligence methodologies for macroeconomic forecasting under unstable economic environments.

References

- Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, 11, 685–725. [\[CrossRef\]](#)
- Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1), 1–22. [\[CrossRef\]](#)
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131(4), 1593–1636. [\[CrossRef\]](#)
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. [\[CrossRef\]](#)
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), C1–C68. [\[CrossRef\]](#)
- Chin, K. H., & Li, X. (2019). Bayesian forecast combination in VAR-DSGE models. *Journal of Macroeconomics*, 59, 278–298. [\[CrossRef\]](#)
- Del Negro, M., & Schorfheide, F. (2012). *DSGE model-based forecasting*. Federal Reserve Bank of New York. Available online: https://www.newyorkfed.org/research/staff_reports/sr554.html (accessed on 1 November 2025).
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357–384. [\[CrossRef\]](#)
- Jurado, K., Ludvigson, S. C., & Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3), 1177–1216. [\[CrossRef\]](#)
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *arXiv*, arXiv:1705.07874. [\[CrossRef\]](#)
- Nachane, D. M., & Chaubal, A. (2023). *Machine learning in macroeconomics: Application to DSGE models*. Springer. [\[CrossRef\]](#)
- Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57(6), 1361–1401. [\[CrossRef\]](#)
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1), 1–48. [\[CrossRef\]](#)
- Stock, J. H., & Watson, M. W. (2005). *Implications of dynamic factor models for var analysis* [NBER working paper]. Available online: <https://www.nber.org/papers/w11467> (accessed on 1 November 2025).
- Stock, J. H., & Watson, M. W. (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In *Handbook of macroeconomics*. Elsevier. [\[CrossRef\]](#)
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.