From Resilience to Volatility: Economic Narratives of Germany, Italy and Ireland

Vossou Eleni
Karintzi Anastasia
Chatigia Mikelino

June 18, 2024

Executive Summary

This research paper analyzes the economic conditions of three important European countries: Germany, Italy, and Ireland. By examining various economic indicators and applying advanced forecasting models the study aims to provide a clear picture of the current economic performance and anticipate future trends and challenges in these countries.

Purpose and Scope

The main goal is to understand how the economies of Germany, Italy, and Ireland function and evolve over time. The research focuses on constructing a Coincident Economic Indicator to gauge the health of these economies and on forecasting Industrial Production across multiple periods to gain insights into their future states.

Methodology

Data was collected from Eurostat, covering exports, imports, retail sales, and other economic activities. Advanced statistical techniques were used to create economic indicators and forecast future trends, simplifying complex data for accurate economic predictions.

Data Description:

- **Target Variable:** The Industrial Production Index, measuring manufacturing activities and construction output in Italy, Ireland, and Germany.
- **Predictor Variables:** Factors such as exports, imports, unemployment rates, and retail sales were analyzed to understand their impact on industrial production.

Econometric Methods:

- **Principal Component Analysis (PCA):** Reduces data complexity while preserving essential information.
- **Dynamic Factor Models (DFMs):** Incorporate time dynamics to capture changes in economic activity over time.
- **Time Series:** Collects data over time to understand trends, seasonality, and forecast future values.
- **Penalized Regressions:** Build models from time series data, adding a penalty term to balance accuracy and simplicity, helping to avoid overfitting.
- **Model Averaging:** Combines forecasts from multiple models to achieve better accuracy, mitigating individual model weaknesses.

- **Machine Learning Methods:** Advanced methods like Random Forest and Neural Networks are used to improve prediction accuracy.

Key Findings

Germany: Germany is known for its strong industrial sector and robust exports. Despite facing challenges like global trade issues and geopolitical tensions, Germany maintains stability through efficient social policies and investments in infrastructure. Forecasting models that handle complex, non-linear relationships and adapt quickly to recent data trends excel in capturing short-term economic volatility. Additionally, methods that track broad economic trends and fluctuations over the long term align well with Germany's stable yet dynamic economic environment.

Italy: Italy's economy grows more slowly due to bureaucratic inefficiencies and less innovation compared to Germany. Although resilient post-pandemic, Italy faces long-term issues like high debt and an aging population. Short-term forecasting is more effective with models focusing on relevant predictors and adapting to immediate conditions. However, simpler models that fail to capture rapid changes are less effective. Long-term methods attempting to reflect economic recovery and structural shifts struggle with Italy's economic volatility.

Ireland: Ireland's economy is characterized by rapid growth and heavy reliance on multinational corporations. This reliance creates volatility and economic disparity. Forecasting models that adapt to sudden economic shifts and capture the impacts of global crises, geopolitical tensions, and economic fluctuations are particularly effective for Ireland. These methods handle high variability better than smoother, stable predictions, which often fail to reflect actual economic conditions.

Conclusions

- **Germany** maintains stable economic conditions with strong industrial production and effective social policies. Effective short-term forecasting requires models that handle non-linear relationships and adapt to recent trends. Long-term forecasting benefits from methods that track broad economic trends.
- **Italy** shows resilience but needs reforms to address long-term structural challenges. Short-term economic fluctuations are best captured by methods focusing on relevant predictors and adapting quickly to changes. Long-term forecasting remains challenging due to structural issues.
- **Ireland** benefits from foreign investment but faces economic volatility and disparity, requiring balanced growth strategies. High economic volatility necessitates forecasting methods that adapt to sudden shifts and capture the impacts of global economic changes.

Recommendations

To ensure long-term economic stability, policymakers should focus on:

- **Germany:** Continuing investments in infrastructure, education, and advanced technology to maintain industrial competitiveness.
- **Italy:** Implementing reforms to reduce bureaucratic inefficiencies, encouraging innovation, and addressing demographic challenges.
- **Ireland:** Diversifying the economy to reduce dependence on multinational corporations and ensure sustainable growth.

This paper provides valuable insights into the economic conditions of Germany, Italy, and Ireland. By understanding these factors, policymakers can design strategies to enhance economic resilience and achieve sustainable growth.

Abstract

This research paper explores the economic conditions of Germany, Italy and Ireland by analyzing key economic indicators. We employ advanced statistical techniques to provide insights into the economic functions and evolution over time. The findings highlight Germany's economic stability, Italy's need for structural reforms, and Ireland's volatility due to reliance on multinational corporations. Recommendations for policymakers include enhancing productivity, addressing structural challenges, and fostering new industries to ensure long-term economic stability.

Keywords: Descriptive Statistics, Coincident Economic Indicator, Forecasting.

Contents

1 Introduction

In the world of economics, analyzing large datasets provides crucial insights into how different countries' economies function and interact. This project aims to provide a comprehensive analysis of key economic indicators for three significant European economies: Italy, Ireland, and Germany. By employing descriptive statistics and visualizations, we seek to uncover trends, patterns, and relationships within the data, offering a unified view of these economies' current states and historical developments.

Germany, Europe's economic powerhouse, features a diverse industrial sector, advanced technology, a skilled workforce, and strong exports, all contributing to stable industrial production and favorable employment conditions.

Italy, despite being a significant European player, faces challenges such as bureaucratic inefficiencies and slower innovation. This results in stable but slower economic growth and consumer spending compared to Germany.

Ireland's economy, marked by rapid growth in the early 2000s and severe impacts from the 2007-08 financial crisis, remains volatile. Its reliance on multinational corporations and sensitivity to global market changes lead to significant fluctuations in industrial production, trade, and consumer spending.

Our main objective is to perform a comprehensive analysis by describing the data, building coincident indicators using Principal Component Analysis (PCA), and conducting forecasting cross-validation with various models. We focus on several crucial economic indicators: Industrial Production, Unemployment Rate, Money Market Interest Rates, Economic Sentiment Indicator, and Price Trends, which encompass various aspects of economic health, including production efficiency, labor market stability, financial market conditions, consumer sentiment, and inflationary expectations. The essay "Descriptive Analysis of Key Economic Indicators in Italy" provides foundational data and statistical insights for understanding Italy's economic landscape. "Macroeconomic Investment Data in Ireland" captures economic behavior and trends in Ireland, providing critical data for comparison. "Germany's Economic Odyssey" offers a comprehensive dataset for examining Germany's economic conditions and historical context. This unified approach not only facilitates a detailed comparison across different countries but also highlights the unique economic challenges and opportunities each country faces.

2 Navigating Economic Terrain

2.1 Germany's Economic Landscape

Germany has consistently maintained steady growth despite lacking natural resources and not being a major trading hub. It is one of the largest economy globally.

"What is the secret behind Germany's resilience?" The country practices Rhine capitalism, a unique blend of market efficiency and strong social policies. This system not only promotes economic competition but also ensures robust social welfare, contributing to Germany's stability and resilience.

Post-WWII, West Germany developed Rhein capitalism to ensure worker satisfaction and counter the spread of socialism. Today, Germany has a well-educated population and world-leading companies in industries like automobiles and capital investments. A large number of specialized small and medium-sized enterprises (SMEs) play a significant role in strengthening the economy.

Germany's reputation for high-quality engineering and manufacturing enables it to command a premium for its products, strengthening its export-driven economy. This export strength has helped Germany navigate numerous economic challenges, including reunification and the Eurozone crisis.

Current challenges include global trade fragmentation, slowdowns in consumer spending and business investment, the war in Ukraine, and Brexit. Despite these, Germany maintains stability through strong social welfare systems, market controls, and investments in infrastructure, education, and defense.

Germany's GDP of $4.08 trillion makes it the fourth-largest economy, with a GDP per capita of $48,717. It remains a reliable center of international business and is home to numerous world-leading companies, known for high-quality products.

While Germany faces challenges like an aging workforce and global competition, it continues to adapt. Maintaining its reputation for high-quality goods and transitioning towards value-added service roles could help Germany remain an economic powerhouse.

2.2 Italy's Economic Landscape

Italy's economy has a complex and varied history, marked by periods of rapid growth and significant challenges. Historically viewed as an economic "basket case" due to high national debt and frequent political crises, Italy recovered remarkably in the 1980s. Reforms and a boom in small and medium-sized enterprises (SMEs) helped Italy's GDP per capita surpass the UK's in 1987.

However, this growth was unsustainable, with national debt reaching 123% of GDP by 1994.

In recent years, Italy has shown signs of economic resilience post-pandemic, with GDP growth of 4.2% since the end of 2019 and a drop in inflation to 0.75%. Factors include lower energy prices, strong tourism, and a 5% rise in exports in 2023. Policies such as the "super bonus" tax relief on home improvements have boosted property and construction sectors, reducing debt from 155% of GDP in 2020 to 137% in 2023.

However, the road ahead is not without its bumps. Despite improvements, Italy faces ongoing issues like demographic challenges, high debt, and underemployment. An aging population, regional disparities, and bureaucratic inefficiencies complicate stability and long-term planning.

Italy's story highlights the unpredictability of economic growth, offering lessons for other economies. Addressing structural issues, improving productivity, and fostering new industries are key to future prosperity.

2.3 Ireland's Economic Landscape

Ireland's economy has transformed dramatically, with a GDP per capita over $103,983, surpassing nations like Italy and Germany. It is one of the fastest-growing economies, even amid global challenges.

EU membership has benefited Ireland, though it limits control over its currency and interest rates.

"What has driven this economic miracle?". Historically, tax loopholes like the "double Irish with a Dutch sandwich" attracted major corporations, leading to significant investments in local research and development.

The Shannon Free Trade Zone, established in 1959, attracted companies like Intel and Sony, boosting the economy and creating high-paying jobs. Despite closing some tax loopholes, Ireland remains a hub for multinational corporations due to its business-friendly environment.

However, foreign investment has driven up rents and living costs, creating economic disparity. Wealth is concentrated in urban areas like Dublin. Critics argue that Ireland's growth is partly due to tax avoidance, with a few American corporations heavily influencing the economy. These companies have a substantial influence on the Irish economy, raising some concerns about the sustainability and fairness of this growth model.

Despite these challenges, Ireland maintains a strong reputation as a business-friendly environment, essential for attracting multinational companies.

3 Data Description and Variables

The dataset under consideration presents a comprehensive collection of economic indicators obtained from Eurostat. By utilizing all the available variables, we aim to gain insights into the performance, trends, and expectations within different segments of the economies of Italy, Ireland, and Germany. The dataset contains 229 observations for Italy, 220 observations for Ireland, and 230 observations for Germany, offering robust sample sizes for analysis. Our key target variable, Industrial Production, comprises these time series observations across the three countries, while the explanatory variables such as Exports, Imports, Unemployment Rate, Retail Sales, Economic Sentiment Indicator, Financial Situation, and Business Climate Indicator also encompass these time series observations, providing a detailed and multifaceted view of each economy.

After detailed data collection and thorough cleaning, we arrive at our final dataset—a refined compilation that serves as the foundation for our economic analysis. This dataset captures information from key economic indicators, which are vital for understanding the economic landscape of these countries. By transforming complex raw data into a logical and manageable form, our goal is to provide a robust basis for exploratory data analysis (EDA).

The final dataset sets the stage for uncovering meaningful insights. With a focus on key economic drivers, trends, and challenges, we are confident in delving deeper into the economic dynamics of Italy, Ireland, and Germany. In the subsequent sections, we will navigate through this dataset, unraveling patterns and implications that contribute to a comprehensive view of these countries' economic journeys.

For the purpose of the analysis, we use the Eurostat database to collect our variables. The selected variables include:

3.1 Target variable

In this forecasting project, the response variable of interest is the future value of the Industrial Production Index (HD1). This key measure of economic activity encompasses manufacturing activities and construction output for Italy, Ireland, and Germany, providing a comprehensive view of each country's industrial performance. By forecasting the Industrial Production Index, we aim to predict future economic trends and assess the robustness of different forecasting models.

3.2 Predictor Variables

The dataset also includes various key explanatory variables that may affect changes in the industrial production index. Analyzing these variables together allows us to identify relationships and gain insights.

3.2.1 High-Frequency Data (HD)

- **HD2: Exports** - Represents the value or volume of goods and services a country sells to other nations within a given period.
- **HD3: Imports** - Indicates the value or volume of goods and services a country purchases from other nations within a given period.
- **HD4: Production in Construction** - Measures the level of construction activity or output within the economy.
- **HD5: Unemployment Rate** - Represents the percentage of the workforce that is unemployed and actively seeking employment.
- **HD6: Short-term Yields** - Refers to the returns or yields on treasury bills or short-term bonds.
- **HD7: Long-term Yields** - Indicates the returns or yields on government or corporate bonds (long-term financial instruments).
- **HD8: Retail Sales** - Captures the total revenue generated by retail establishments within a specified period.
- **HD9: Euro/ECU (USD) Exchange Rates (Monthly Data)** - Represents the exchange rate between the Euro (or European Currency Unit, ECU) and the U.S. Dollar (USD), typically provided on a monthly basis.
- **HD10: Spread 10Y - 3M (HD7-HD6)** - Calculates the yield spread between long-term (10-year) and short-term (3-month) financial instruments.

3.2.2 Survey Data (SD)

- **SD1: Production Development Observed Over the Past 3 Months**
- **SD2: Production Expectations Over the Next 3 Months**
- **SD3: Employment Expectations Over the Next 3 Months (Retail Survey)**
- **SD4: Assessment of Order-Book Levels (Industry)**
- **SD5: Assessment of Export Order-Book Levels (Industry)**
- **SD6: Assessment of Current Level of Stocks of Finished Products (Industry)**
- **SD7: Building Activity Development Over the Past 3 Months (Construction)**
- **SD8: Evolution of Current Overall Order Books (Construction)**
- **SD9: Employment Expectations Over the Next 3 Months (Construction)**
- **SD10: Price Expectations Over the Next 3 Months (Construction)**
- **SD11: Business Activity (Sales) Development Over the Past 3 Months**
- **SD12: Volume of Stocks Currently Held**
- **SD13: Expectations of the Number of Orders Over the Next 3 Months**
- **SD14: Business Activity Expectations Over the Next 3 Months**
- **SD15: Employment Expectations Over the Next 3 Months**
- **SD16: Business Situation Over the Past 3 Months**
- **SD17: Evolution of Demand Over the Past 3 Months**
- **SD18: Expectation of Demand Over the Next 3 Months**
- **SD19: Evolution of Employment Over the Past 3 Months**
- **SD20: Expectation of Employment Over the Next 3 Months**
- **SD21: Euro-Zone Business Climate Indicator (Monthly Data)** - Measures the overall business climate within the Eurozone region based on monthly data and provides insights into the sentiment, expectations, and perceptions of businesses regarding current and future economic conditions.
- **SD22: Construction Confidence Indicator**
- **SD23: Economic Sentiment Indicator (ESI)** - Gauges the overall economic sentiment or confidence levels within the country, important for assessing economic outlook, sentiment shifts, and potential future economic developments.
- **SD24: Industrial Confidence Indicator**
- **SD25: Retail Confidence Indicator**
- **SD26: Consumer Confidence Indicator**
- **SD27 to SD35: Financial Situation and Consumer Expectations**

3.3 Comparative Insights

In analyzing the economic indicators of Germany, Italy, and Ireland, several underlying assumptions and contextual factors must be considered. Germany and Italy are two major European economies with significant industrial bases and strong positions within the European Union. Germany, often referred to as the economic powerhouse of Europe, boasts a highly diversified industrial sector, advanced technology, a skilled labor force, and robust export markets. Italy, while also a key player in the European economy, faces structural challenges such as bureaucratic inefficiencies and slower innovation, which impact its economic growth and stability.

Ireland, on the other hand, presents a contrasting economic narrative. The country experienced rapid economic growth in the early 2000s, often referred to as the "Celtic Tiger" period. However, the 2007-08 financial crisis severely impacted Ireland's economy, leading to significant job losses, banking sector instability, and long-term economic volatility. Despite subsequent recovery efforts, Ireland remains highly sensitive to global economic fluctuations, largely due to its reliance on multinational corporations and the smaller size of its economy compared to Germany and Italy.

In this section, we analyze the comparative descriptive statistics portrayed in Table 1.

3.3.1 Industrial Stability and Growth

Germany exhibits the most stable and positive industrial production among the three countries, as evidenced by its mean of 0.13 and lower standard deviation of 1.562.

"What makes Germany's industrial production so resilient?" The answer of course lies in its robust manufacturing sector and the country's focus on high-quality production. The stability in Germany's industrial production suggests a well-diversified industrial base and strong economic fundamentals.

Italy shows stability with a mean of 0.03 but less growth compared to Germany. "Why is Italy not growing as fast?" Structural challenges and a less dynamic industrial sector might be the answer to this question. Meanwhile, Ireland's industrial sector appears to be the most volatile and contracting, with a negative mean of -0.196 and a high standard deviation of 8.34. "Could this volatility be due to Ireland's reliance on multinational corporations and global market conditions?" It seems likely.

When we look at the industrial production trends over time in Figure 1, Germany's line is relatively steady, indicating stable industrial output. Although it doesn't show significant growth recently, the stability reflects strong economic fundamentals and a diversified industrial base. In contrast, Ireland's industrial production shows significant fluctuations, reflecting economic sensitivity and volatility. The sharp spikes and dips suggest a high dependency on specific sectors and multinational corporations, making the economy more susceptible to global market changes.

Italy's industrial production line is higher than Germany's for most years, indicating a generally higher level of output. However, it doesn't show significant upward growth, suggesting challenges in achieving further industrial expansion. This stability without growth can be attributed to structural issues within the Italian economy, such as bureaucratic inefficiencies and slower innovation.

Overall, Germany's steady industrial performance underscores its position as an economic powerhouse in Europe. Meanwhile, Ireland's volatility highlights economic vulnerabilities due to reliance on key industries and multinational corporations. Italy's stable yet stagnant growth reflects ongoing structural challenges that need to be addressed.

3.3.2 Trade Dynamics

Germany leads in both exports and imports, with a mean of 0.451 for exports and 0.046 for imports, showing its strong role in global trade. The moderate standard deviations (3.775 for exports and 3.127 for imports) indicate stable trade patterns. This stability is due to Germany's strong industries and trade relationships.

Italy maintains stable trade levels with a mean of 0.25 for exports and 0.12 for imports, and moderate variability (standard deviations of 0.40 and 0.35, respectively). This reflects consistent performance in key sectors like machinery, fashion, and food products. Ireland, however, shows significant fluctuations in trade with a mean of 0.078 for exports and -0.034 for imports, and high standard deviations (8.623 for exports and 6.98 for imports), suggesting a smaller economy highly sensitive to global market changes and reliant on a few large multinational firms.

The positive skewness in Germany's and Italy's exports suggests occasional high export values that raise the mean above the median. Germany's exports are significantly driven by the automotive and machinery sectors, known for their high quality and innovation, contributing to economic stability. In contrast, Italy's exports, dominated by fashion, machinery, and food products, face challenges such as less diversification and slower innovation, limiting trade performance relative to Germany.

3.3.3 Employment Conditions

Germany shows the best employment conditions with a negative mean unemployment rate (-0.033) and low variability (standard deviation of 0.073), reflecting effective labor market policies and a resilient economy. Italy and Ireland have relatively stable but higher unemployment rates, with means of 0.00 and 0.01, respectively, and standard deviations of 0.23 and 0.22. This indicates structural unemployment issues and economic rigidity, which Germany manages more effectively.

The median unemployment rate is 0.00 for both Italy and Germany, indicating stability in their labor markets. The negative skewness in Ireland's unemployment rate (-0.15) suggests occasional low unemployment rates pulling the mean below the median.

3.3.4 Consumer Spending

Retail sales in Germany remain stable and positive, with a mean of 0.115 and a standard deviation of 1.525, indicating high consumer confidence. This stability is due to high employment rates, stable income levels, and a robust social safety net. Italy has steady consumer spending with a mean of 0.15 and a standard deviation of 0.31, due to cultural factors and economic stability. In contrast, Ireland has high variability in retail sales, with a mean of -0.19 and a standard deviation of 7.31, showing uncertain consumer behavior related to volatile industrial output.

Ireland's wide range in retail sales (-25.46 to 23.67) and high standard deviation point to extreme fluctuations, unlike Germany, which has narrower ranges indicating steadier economic performance.

3.3.5 Economic Sentiment and Financial Health

Economic sentiment indicators are positive and relatively stable across all three countries. Germany shows a mean of 0.196 and a standard deviation of 2.191, indicating consistent economic sentiment. Italy has a mean of 0.19 and a standard deviation of 2.61, while Ireland has a mean of 0.203 and a standard deviation of 3.365, reflecting higher variability. These results indicate general optimism about economic prospects, driven by strong economic fundamentals and consumer confidence in Germany, with periodic uncertainties in Italy and Ireland. The financial situation indicators show Germany with the most stability (mean of 0.095, standard deviation of 1.429), indicating better financial health.

3.3.6 Business Climate

The business climate is neutral to slightly positive in all three countries, with minimal variability. Germany has a mean of 0.006 and a standard deviation of 0.194, suggesting a stable business environment. Italy and Ireland show slightly positive means of 0.00 and 0.01, with standard deviations of 0.15 and 0.14. The high kurtosis in Germany's business climate indicator (4.299) and economic sentiment indicator (6.186) suggests a higher likelihood of extreme values, indicating occasional periods of very positive or negative sentiment compared to Italy and Ireland.

Moreover, the COVID-19 pandemic has had varying impacts on these economies. Germany demonstrated resilience due to its diversified industrial base and robust healthcare system, while Italy and Ireland faced more significant disruptions due to their economic structures and reliance on specific sectors.

3.3.7 Comparative Conclusion

The comparative insights indicate that Germany's robust and diversified economy, effective policies, and stable industrial base contribute to its strong performance across various economic indicators. Italy, while stable, faces structural challenges and slower growth. Ireland, recovering from the financial crisis of 2007-2008, shows significant variability due to its smaller size and reliance on multinational corporations. These differences highlight the unique economic contexts and challenges faced by each country.

4 Constructing Coincident Economic Indicator

4.1 Definition and Purpose

"When trying to gauge the health of an economy, one of the most crucial tools in an economist's toolkit is the economic coincident indicator. But what exactly does this mean?" An coincident economic indicators (CEI) is an essential tool in macroeconomic analysis designed to provide a real-time snapshot of the current state of an economy. By integrating multiple macroeconomic variables into a single, comprehensive index, CEIs effectively summarize economic activity and offer a clear picture of overall economic health. This approach captures the common movements in these variables, representing the overall economic conditions accurately.

The concept of Coincident Economic Indicators has evolved significantly since it was first developed. Wesley Mitchell and Arthur Burns at the National Bureau of Economic Research (NBER) were among the first to delve into this area. They focused on understanding the patterns of economic expansions and contractions, which are essential for defining business cycles. Today, more sophisticated methods, like PCA and Dynamic Factor Models (DFMs), are used to construct CEIs. For example, James Stock and Mark Watson, in their work "A Probability Model of The Coincident Economic Indicators," have introduced a formal statistical framework for creating these indicators. They show that CEIs can be built as weighted averages of various key macroeconomic time series, effectively capturing the overall state of economic activity. This modern approach ensures that CEIs provide a reliable and comprehensive picture of economic health.

4.1.1 Data Selection

For each country—Germany, Italy, and Ireland—we selected a comprehensive set of macroeconomic variables indicative of economic performance. This selection includes all 44 explanatory variables and the target variable mentioned earlier in the Data Description and Variables section. These variables were chosen based on their relevance and availability, ensuring they accurately reflect the economic conditions of each country. We utilize all of these variables because we assume, from an economic perspective, that all these variables have meaningful contributions to building an effective Coincident Economic Indicator and eventually capturing the multifaceted nature of economic activity.

4.1.2 Principal Component Analysis (PCA) Approach

Principal Component Analysis (PCA) is a powerful technique widely used in economics to construct Coincident Economic Indicators (CEI).

It simplifies large datasets and panel data by transforming numerous variables into a smaller set of principal components that retain essential information while reducing complexity. By identifying patterns and relationships among variables, PCA helps economists extract key insights from data without assuming specific distributions, thus enhancing the interpretability and reliability of economic indicators. Moreover, PCA's ability to mitigate multicollinearity ensures that the resulting indicators provide robust forecasts by minimizing redundancy among predictor variables.

This methodology will be further analyzed in the subsequent section on econometric methodology, where we will delve into its application and implications in constructing CEIs. Overall, PCA serves as a crucial tool in constructing CEIs, offering a structured approach to distill meaningful economic signals from diverse and intricate datasets.

4.1.3 Dynamic Factor Models (DFMs) Approach

Dynamic Factor Models (DFMs) might sound complex, but in essence they capture the ripple effects in the economy, much like how a drop of water creates waves in a pond. They extend the concept of PCA by incorporating time dynamics into the analysis. While PCA focuses on reducing dimensionality by identifying key components that explain the variance in a dataset, DFMs go a step further by modeling the temporal evolution of these components. Essentially, DFMs assume that a few hidden factors drive the observed economic variables, and these factors change over time, capturing the dynamic nature of economic activity.

In macroeconomics, DFMs are particularly useful for extracting the underlying factors that influence multiple macroeconomic time series. By doing so, DFMs provide a more comprehensive and dynamic Coincident Indicator, allowing economists to better understand the current state of the economy and predict future economic trends. For instance, the work of James Stock and Mark Watson on probability models of Coincident Economic Indicators (CEIs) utilizes DFMs to capture the evolving nature of economic activity. Their approach involves modeling the joint dynamics of multiple time series to extract common factors that represent the state of the economy.

A typical DFM can be described by the following equations:

$$
\begin{aligned}
y_t &= Pf_t + Qx_t + u_t \\
f_t &= Rw_t + A_1 f_{t-1} + \cdots + A_p f_{t-p} + v_t \\
u_t &= C_1 u_{t-1} + \cdots + C_q u_{t-q} + \epsilon_t
\end{aligned}
$$

where:

- y_t represents the observed variables,
- f_t represents the unobserved factors,
- x_t and w_t are exogenous variables,
- $P, Q, R, A_1, \ldots, A_p, C_1, \ldots, C_q$ are matrices of parameters to be estimated,
- P represents the factor loadings,
- ϵ_t and v_t are error terms.

Estimating DFMs can be computationally challenging, especially with large datasets. Standard state-space methods, including the Kalman filter, are often used for estimation. Identification of factors requires assumptions, such as setting certain elements of covariance matrices to zero or fixing some parameters.

4.1.4 Building the Coincident Indicator

In this section, we develop an index to identify trends in the economy by constructing an Economic Coincident Indicator using our extensive macroeconomic dataset.

We employ both Principal Component Analysis (PCA) and Dynamic Factor Models (DFMs) to build the indicators.

On one hand, PCA is straightforward to implement and interpret, as it reduces the dimensionality of the data by capturing the maximum variance in the first principal component. However, PCA is static and does not account for the temporal dynamics in the data. On the other hand, DFMs use multiple time series to extract common factors that explain the co-movements among the variables, capturing both static and dynamic relationships. While DFMs provide a more comprehensive view, they are more complex to implement and require greater computational resources.

The first principal component from PCA is used to generate a time series plot, as it captures the maximum amount of variance compared to subsequent components. This approach ensures that we incorporate the most relevant information for summarizing significant patterns and trends in the data.

To enhance the indicator, DFMs are used to capture the dynamic relationships and time-dependent structures within the data. DFMs assume that observed economic variables are driven by a few unobserved factors following their own stochastic processes, providing a dynamic view of economic activity.

Kernel smoothing is applied to both the first principal component from PCA and the DFM indicator to derive smoothed values. This process averages the values by applying weighted sums, considering both past and future observations. The weights adjust dynamically over time, assigning greater weight to recent observations compared to those from a year ago. This time-sensitive feature enhances the representation of temporal trends in our data, contributing to a clearer depiction of the underlying patterns.

4.2 PCA Coincident Economic Indicator: Application to Each Economy

In this section, we discuss the results of the Coincident Economic Indicators for Germany, Ireland, and Italy, both individually and comparatively. Figures 2 to 5 highlight key economic periods, shaded in pale colors. These indicators are derived from the first principal component (PCA(1)) of 45 economic variables, representing the underlying economic conditions over time. The data spans from around 2005 to 2022, providing a long-term view of each country's economic conditions. The line in each figure represents the coincident indicator, with values moving around zero. Peaks indicate periods of economic expansion, while troughs signify periods of contraction.

4.2.1 Germany's Coincident Economic Indicator

During the period of the global financial crisis (2007 to 2009), the indicator in Figure 2., shows a significant decline, reaching a trough around 2009, reflecting the severe impact of the financial crisis on Germany's economy. The reason for this decline is the strong reliance on exports and the global nature of the crisis, which reduced global demand for German goods. Additionally, the banking sector in Germany faced stress, limiting credit availability and further dampening economic activity.

The gradual recovery from the crisis with smaller fluctuations indicates that the German economy was stabilizing but still faced challenges, preventing a quick return to pre-crisis levels.

During the COVID-19 pandemic (2020 to 2021), the sharp decline in the indicator reflects the immediate and widespread economic disruptions caused by lockdowns, supply chain interruptions, and decreased consumer and business activity. The steep drop in early 2020 is followed by a rapid but fluctuating recovery, indicating the high volatility and uncertainty in economic conditions during the pandemic. The rapid recovery, despite fluctuations, showcases the resilience of the German economy, aided by substantial government support measures, such as financial aid packages and stimulus efforts, which helped to mitigate the pandemic's economic impact.

These patterns suggest that while Germany's economy is strong and can recover from crises, it remains vulnerable to significant global disruptions, leading to periods of instability and uncertainty.

In Table 2., we present the top 10 variables with the largest absolute loadings on PCA(1) out of the 45 variables of the dataset for each country. This provide us with useful insights on the correlation be- tween those variables and the first principal component. The absolute loadings provide insight into the magnitude of these relationships, indicating which variables have the strongest influence on the factor.

For Germany, we observe that the top variables influencing PCA(1) are primarily related to survey data (SD), particularly those that gauge business sentiment, financial situation, and order levels. The variable SD23 (Economic Sentiment Indicator - ESI) has a loading of -0.282143, making it a key determinant of the economic conditions. The negative loading indicates that an increase in economic sentiment leads to a decrease in the first principal component. Similarly, SD24 (Industrial Confidence Indicator) has a significant negative loading of -0.267598, suggesting that higher industrial confidence lowers the principal component value. The SD21 (Euro-Zone Business Climate Indicator) also has a negative loading of -0.239699, suggesting a similar inverse relationship. Other significant variables include SD33 (Consumer Confidence Indicator), SD4 (Assessment of Order-Book Levels - Industry), and SD3 (Employment Expectations Over the Next 3 Months - Retail Survey). The predominance of survey data related to confidence and expectations implies that business and consumer sentiments are critical to understanding Germany's economic conditions.

These findings are consistent with the analysis that Germany's economy, which is highly sensitive to global economic conditions, experienced a sharp decline during the financial crisis due to its export-oriented industrial base. The significant impact of sentiment and confidence indicators aligns with the observation of Germany's gradual and relatively stable recovery post-crisis, supported by robust industrial policies and global demand recovery.

4.2.2 Italy's Coincident Economic Indicator

In Italy's Economic Indicator, in Figure 4., we can observe the indicator significantly decline, reaching its lowest point around 2009, which reflects the severe impact of the financial crisis on Italy's economy. This decline is linked to Italy's structural economic weaknesses, problems in the banking sector, austerity measures, and constraints from being part of the Eurozone, leading to a slow recovery that did not quickly return to pre-crisis levels.

During the COVID-19 period, the indicator again shows a significant decline, caused by strict lockdowns, supply chain disruptions, and high levels of uncertainty among consumers and businesses, despite government support measures. These factors highlight Italy's economic vulnerabilities during both financial and health crises.

Italy's top 10 variables show a similar trend but with slightly different emphases. The variable SD23 (Economic Sentiment Indicator - ESI) has a strong negative loading of -0.313050, indicating an inverse relationship with PCA(1). SD24 (Industrial Confidence Indicator) also has a negative loading of -0.310741, suggesting that higher industrial confidence decreases the principal component value. SD4 (Assessment of Order-Book Levels - Industry) has a strong negative loading of -0.280425. Other significant variables include SD2 (Production Expectations Over the Next 3 Months), SD1 (Production Development Observed Over the Past 3 Months), and SD5 (Assessment of Export Order-Book Levels - Industry). For Italy, industrial confidence and order levels seem to be critical in determining economic conditions, similar to Germany but with an even stronger emphasis on immediate production and order book assessments.

This confirms the analysis that Italy's economy, characterized by structural weaknesses and higher public debt, experienced a substantial decline during the financial crisis and a slow, fluctuating recovery. The persistent challenges in Italy's economic structure and the banking sector's issues are reflected in the significant negative impact of these indicators, highlighting the country's prolonged economic difficulties.

4.2.3 Ireland's Coincident Economic Indicator

In Figure 3., we observe that during the Global Financial Crisis, Ireland's indicator shows a significant decline, reflecting the severe impact of the financial crisis on Ireland's economy. The indicator hits its lowest point around 2009, indicating a deep economic downturn.

After the crisis, the indicator demonstrates a gradual recovery. Unlike the steep decline during the crisis, this recovery period features milder fluctuations, suggesting a stabilizing economy.

Unlike the financial crisis, the indicator during the COVID-19 period does not show a sharp decline. Instead, there is an increase in the economic indicator, indicating that Ireland experienced some economic improvement during this period. This is followed by a rapid recovery, showing less volatility compared to the previous crisis.

This behavior can be attributed to several factors. Substantial government stimulus and support measures, including financial aid to individuals and businesses, tax reliefs, and wage subsidies, played a critical role in stabilizing the economy. Additionally, the uneven impact of the pandemic across different sectors benefited Ireland's strong presence in technology and pharmaceuticals, which experienced growth. The swift adaptation to remote work, supported by Ireland's robust IT infrastructure, helped maintain productivity. Coordinated global economic efforts, increased consumer savings leading to pent-up demand, and effective healthcare and vaccination responses further contributed to the swift recovery. Unlike the systemic financial issues underlying the 2007-2009 crisis, the COVID-19 pandemic's disruptions were more short-term, allowing for a quicker rebound once health measures took effect.

We believe that these combined factors explain why Ireland's economic indicator showed improvement during the pandemic, highlighting the resilience of its economy.

For Ireland, in Table 2., the top variables influencing PCA(1) are similarly related to sentiment and expectations. The variable SD23 (Economic Sentiment Indicator - ESI) has a strong positive correlation with the principal component, with a loading of 0.300051, suggesting that higher sentiment increases PCA(1). SD26 (Consumer Confidence Indicator) also shows a strong positive loading of 0.298913, indicating that consumer confidence boosts the economic conditions captured by PCA(1). The SD30 (Business Activity Sales Development Over the Past 3 Months) has a positive loading of 0.277148, indicating a similar positive effect on the principal component. Other significant variables include SD33 (Consumer Confidence Indicator), SD14 (Business Activity Expectations Over the Next 3 Months), and SD28 (Financial Situation Indicator). For Ireland, variables related to both business and consumer confidence and recent business activities play a significant role in defining economic conditions.

This supports the observation that Ireland's economic structure, with a strong presence of multinational corporations in tech and pharmaceutical sectors, shows resilience and quicker recovery patterns. The positive loadings on consumer and business sentiment indicators explain Ireland's relatively faster recovery post-financial crisis and the increase in the indicator during the COVID-19 pandemic, reflecting the benefits from these resilient sectors.

4.3 Comparative Analysis Across Economies

The trend depicted in Figure 5 regarding the economic sentiment indicator for Germany, Italy, and Ireland provides intriguing insights, particularly during significant crises such as the global financial turmoil of 2007-2008 and the COVID-19 pandemic of 2020-2021.

During the global financial crisis (2007-2009), all three countries exhibit significant declines in their economic indicators, reflecting the severe global economic impact. Germany's indicator experiences a sharp drop around 2009, indicating the crisis's profound effect on its export-oriented economy. Italy also shows a substantial decline, reflecting its structural economic weaknesses and banking sector issues. Interestingly, Ireland's recession appears to begin slightly earlier than those of Italy and Germany. However, Ireland's decline is smoother compared to the massive drops seen in the other two countries. This smoother decline may be due to Ireland's different economic structure and more flexible recovery policies, which allowed it to navigate the crisis more steadily.

Following the financial crisis, all three countries exhibit a gradual recovery with fluctuations. Germany's indicator shows a relatively stable recovery, though it does not quickly return to pre-crisis levels. Italy's recovery is slow and marked by ongoing fluctuations, indicating persistent economic challenges. Ireland shows a more volatile recovery, with sharper ups and downs, suggesting both resilience and ongoing economic adjustments.

During the COVID-19 pandemic (2020-2021), the indicators for all three countries show a decline, but the patterns differ. Germany experiences a sharp drop followed by a rapid, albeit fluctuating, recovery, reflecting its strong government support measures and economic resilience. Italy's indicator shows a significant decline, followed by a gradual and fluctuating recovery, highlighting its vulnerabilities and the impact of prolonged restrictions. Ireland's indicator, interestingly, shows an increase during the pandemic period, potentially reflecting the strength of its technology and pharmaceutical sectors, which may have benefited during the crisis.

The reasons for these observed patterns lie in the economic structure and resilience of each country, as well as the effectiveness of government interventions and the specific impacts on different sectors. Germany's strong industrial base and export-oriented economy are highly sensitive to global economic conditions, explaining the sharp declines during crises and the relatively stable recoveries due to robust industrial policies and global demand recovery. Italy's economic structure, characterized by higher public debt and structural weaknesses, results in slower and more fluctuating recoveries. The banking sector's issues also play a role in its prolonged economic challenges. Ireland's economy, with a significant presence of multinational corporations in the tech and pharmaceutical sectors, shows different dynamics. These sectors' resilience during the pandemic explains the increase in the indicator during COVID-19, contrasting with the declines in Germany and Italy.

Government interventions and policies also significantly influence recovery patterns. The scale and effectiveness of interventions, such as stimulus packages and support measures, are critical. Germany's substantial fiscal support and strategic interventions helped stabilize its economy, leading to a rapid recovery during the pandemic. Italy's austerity measures during the financial crisis and challenges in implementing effective support measures during the pandemic resulted in slower and more volatile recoveries. Ireland's proactive policies and the beneficial impact of multinational corporations' activities contributed to its relatively quicker recovery post-financial crisis and resilience during the pandemic.

It seems only fair to assume that, different sectors respond uniquely to crises. Germany's industrial sector was heavily impacted during both crises but benefited from global recovery efforts. Italy's diverse but vulnerable sectors, including tourism and small businesses, faced significant challenges during both crises. Ireland's tech and pharmaceutical sectors likely provided a buffer during the pandemic, explaining the upward trend in its economic indicator.

In Table 2., we observe that Economic Sentiment (SD23) is critical across all three countries, reflecting its importance in gauging overall economic conditions.

Moreover, Consumer Confidence Indicators (SD26 and SD33) and SD24 (Industrial Confidence Indicator) also appear prominently, highlighting the significance of consumer sentiment in shaping economic outlooks.

4.4 PCA Coincident Indicators vs. DFMs Coincident Indicators

As mentioned earlier, selecting the appropriate method for constructing a coincident economic indicator requires careful consideration of the advantages and limitations of each approach.

In this section, we will conduct a comparative analysis of these two methods—PCA and DFMs—to determine which is more suitable for constructing the coincident economic indicator. This section of our analysis aims to provide valuable insights into the strengths and weaknesses of each method, helping to identify the optimal approach for our purposes.

In Figure 6., we present the Economic Coincident Indicator using Dynamic Factor Models (DFMs) for Germany, Italy, and Ireland.

From the comparison of Figures 5 and 6, we observe that Germany's and Italy's PCA indicators show a sharp rise leading up to 2009, reflecting a temporary economic resilience or possibly delayed recognition of the crisis's severity. Finally in 2009, the indicators experience a steep decline, indicating the eventual realization and impact of the financial crisis.

In contrast to Germany and Italy, Ireland's PCA indicator shows a gradual decline leading up to 2009, indicating that Ireland felt the impact of the financial crisis earlier, probably due to the decline in the construction sector. Interestingly, around 2009, Ireland's indicator rises, suggesting an initial phase of recovery or stabilization efforts taking effect sooner than in Germany and Italy.

Interestingly in Figure 6., we unravel a different story. The decline in the DFM indicators of all the three economies is more synchronized across the countries, indicating a common recognition of the crisis period. Ireland's decline is less severe, probably due to technology and pharmaceutical sectors' resilience, compared to Germany and Italy, highlighting a relatively more resilient economy during the crisis or effective early interventions.

All three countries show a rapid recovery post-2009, which is more uniform compared to the PCA indicators.

Moving to the COVID-19 Pandemic (2020-2021), we observe in Figure 5., the both Germany and Italy experience sharp declines, meaning a significant immediate economic impacts due to strict lockdown measures and disruptions in economic activities. Ireland's indicator manages to capture the immediate positive impacts of resilient sectors (e.g., tech and pharmaceuticals) explaining the rise after 2020.

The recovery phases for Germany and Italy are marked by high volatility, with significant fluctuations indicating ongoing economic uncertainty.

The DFMs indicators, narrate an interesting story. Before 2020, key sectors in the two major economies, such as manufacturing in Germany and a mix of manufacturing and services in Italy, might have been performing well, contributing to the rising indicators. Ireland's economy though, might have been undergoing adjustments or facing sector-specific challenges that contributed to a decline in the overall economic indicator (e.g Brexit uncertainties could affect Ireland due to its close economic ties with the UK). Also this indicator could show the impact of the pandemic across all sectors over time.

A reverse pattern is emerging after 2020. The once strong key sectors such as manufacturing, automotive for Germany, and tourism for Italy were heavily affected, and combined with the strict lockdowns and delays in economic recovery measures lead to significant declines in the two major economies. While in Ireland the resilience of the tech and pharmaceutical sectors, effective adaptation to remote work, and strong government support, lead to a rise in the DFMs indicator.

4.4.1 Germany's Case

For Germany, the PCA indicator in Figure 7., shows a sharp rise until 2009, reflecting resilience in industrial and export sectors, followed by a dramatic decline due to global demand plummeting and financial sector stress. The DFM indicator presents a smoother, less volatile trend, capturing broader economic impacts and a more stable recovery post-2009. During the COVID-19 pandemic, the PCA indicator reflects significant disruptions in manufacturing and exports, while the DFM indicator indicates quicker economic stabilization.

4.4.2 Italy's Case

Finally, for the Italian economy, Figure 9., the PCA indicator rises until 2009, reflecting delayed recognition in manufacturing and services, followed by a sharp decline due to structural weaknesses and the banking crisis. Public debt and austerity measures sharpened the downturn. However, the DFM indicator shows a more moderated decline, indicating broader sectoral impacts with less short-term volatility. The slow, fluctuating recovery in both indicators highlights ongoing structural challenges in Italy's economy. During the COVID-19 pandemic, the dramatic decline and the volatile recovery in PCA reflects severe impacts on tourism and services, while the DFM's smoother recovery suggests stabilization by other sectors.

4.4.3 Ireland's Case

In Figure 8., the PCA indicator shows decline during the financial crisis, highlighting gradually impacts on the construction and real estate sectors. The subsequent rise around 2009 and after 2020 and less severe decline in the DFM indicator those years, suggest that Ireland's tech and pharmaceutical sectors, which benefited from global demand, provided a buffer. Despite the banking crisis, proactive policies and restructuring efforts contributed to a quicker recovery, reflected in both indicators. The DFM's smoother trend captures long-term economic stability, while PCA reveals short-term volatility.

4.4.4 Conclusion: The Choice of Indicator

After reviewing our results, we conclude that PCA indicators provide a broad summary but may not capture sector-specific impacts as effectively as DFMs, while DFM indicators filter out short-term noise.

For a detailed and dynamic understanding, especially during periods of economic volatility, DFMs provide deeper insights.

However, while DFMs offer the ability to highlight sector-specific impacts, PCA is well-suited for constructing a broad, inclusive economic indicator that summarizes the overall economic environment.

4.5 Coincident Indicator Process Conclusion

The analysis of Economic Coincident Indicators for Germany, Ireland, and Italy reveals distinct responses to global crises such as the 2007-2008 financial turmoil and the COVID-19 pandemic.

During the financial crisis, all three countries experienced significant declines. Germany's sharp drop highlighted its export dependence, Italy's substantial downturn underscored structural weaknesses, and Ireland, though affected, showed a slightly faster recovery due to different economic structures and policies.

Post-crisis, recovery patterns varied. Germany exhibited a stable but gradual recovery, Italy's was slow and marked by fluctuations, and Ireland's was more volatile, reflecting resilience and ongoing adjustments.

During the COVID-19 pandemic, Germany experienced a sharp decline followed by a rapid, fluctuating recovery due to strong government support. Italy's significant decline and gradual recovery highlighted its vulnerabilities. Ireland's indicator increased, reflecting the strength of its tech and pharmaceutical sectors.

Table 2., confirmed the importance of sentiment indicators (Economic Sentiment SD23, Consumer Confidence SD26 and SD33, Industrial Confidence SD24) across all countries. These findings emphasize the critical role of confidence and expectations in driving economic conditions.

This analysis underscores the importance of testing both PCA and DFM methodologies to gain a comprehensive understanding of economic conditions. PCA is useful for real-time monitoring of economic volatility and in general campuring the general sentiment of the economy, while DFMs offer valuable insights into long-term economic stability and resilience.

In conclusion, recognizing the key factors influencing each country's economic conditions during crises provides valuable insights for policymakers. Understanding the importance of sentiment indicators and targeted interventions can help design strategies to stabilize economies in future global disruptions.

5 Addressing the challenge of Forecasting

"How can we effectively forecast the economic outlooks for Germany, Italy, and Ireland, and ensure our predictions are as accurate as possible?". We're faced with the challenge of uncovering the relationships between industrial production and the explanatory variables. To tackle this, we have two primary approaches: variable selection and variable reduction.

With variable selection methods we try to identify the variables that have the most significant impact on industrial production. By focusing on cherry-picking those predictors that exert the most influence on the outcome variable, we aim to improve model performance. Techniques such as penalized regression (e.g., Rigde or LASSO) help address multicollinearity and select the most relevant variables.

Another questions arise at this point: "Could machine learning techniques more effectively capture the complex, non-linear relationships among variables than traditional econometric models?".

"What if instead of evaluating individual variables, we combine them into factors or composite variables?". This approach condenses the multitude of variables into a more manageable set of factors, enabling a streamlined analysis while preserving the essential information encapsulated in the data. Factor models, like Principal Component Analysis (PCA) and Partial Least Squares (PLS) are commonly used for this purpose.

Both methods offer distinct advantages and trade-offs, necessitating careful consideration to determine the most suitable approach for our analytical objectives.

Building on this thought process, we're considering a different approach—one that leverages the strengths of multiple forecasting models. We wonder: while individual models provide valuable insights, wouldn't it be more robust to combine them? Model combination, or model averaging, effectively creates a new, composite model. This approach has the potential to enhance our predictive accuracy by integrating the strengths of various models, thus providing a more comprehensive and reliable forecast.

Our approach involves setting specific configurations to streamline the forecasting process. We will use both recursive and rolling window methods, adjusting the forecast horizons to 1 month for short-term forecasts, crucial for capturing immediate economic changes (sudden market shifts or policy impacts), 6 months for medium-term to capture more sustained economic trends, and 12 months for long-term forecasts, to predict broader economic trends and structural changes for strategic planning and understanding the potential future state of the economy.

For the rolling window method, we propose using a window size of 24 months for short-term forecasts, in order to capture the most recent trends while maintaining enough historical data to provide a robust forecast, 36 months for medium-term forecasts to incorporate multiple business cycles and reduce the noise from short-term fluctuations and 60 months for long-term forecasts, which captures long-term structural changes, such as demographic shifts, technological advancements, and major policy impacts.

Our primary objective of this analysis is to compare the forecasting performance across the three economies. By applying a consistent methodology, that we will present in the next section, and settings across all models, we can gain a clearer understanding of the relative economic dynamics and stability of Germany, Italy, and Ireland. This comparative analysis will help us draw more reliable and insightful conclusions about their respective economic outlooks.

6 Econometric methodology

In this section, we present some simple models that provide a foundational basis for understanding more complex forecasting methods and serve as useful benchmarks for comparing the performance of advanced techniques. We then proceed to present the advanced techniques, including penalized regression and factor models, which offer greater robustness and accuracy in forecasting.

6.1 Naive Models

Naive models use past data points to predict future values, operating on the assumption that future values will be similar to past values. These models serve as a baseline for comparison with more sophisticated forecasting techniques.

6.2 Averaging Models

Averaging models predict future values based on the average of past observations. This approach helps smooth out noise and highlights underlying trends in the data. A common form is the Moving Average

(MA) model, where the prediction is the average of the past k observations.

$$\hat{y}_{t+1} = \frac{1}{k} \sum_{i=0}^{k-1} y_{t-i}$$

6.3 Linear Time Series Models

6.3.1 Autoregressive Models (AR)

Autoregressive models predict future values based on a linear combination of past values. The order of the model, denoted as AR(p), indicates how many past values are used in the prediction.

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t$$

where ϵ_t is white noise, and ϕ_i are the coefficients of the model.

6.3.2 Autoregressive Moving Average Models (ARMA)

ARMA models combine AR and MA components to capture both trends and noise structure in the data. The AR part regresses the variable on its own previous values, and the moving average (MA) part, which models the error term as a linear combination of past errors.

In an ARMA(p, q) model, p represents the number of lagged observations in the AR part, while q denotes the number of lagged forecast errors in the MA part. The combination of these components allows the ARMA model to capture the dynamic behavior of the time series effectively.

They are used for modeling time series with autocorrelation and they are effective for short-term forecasting.

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=0}^{q} \theta_j \epsilon_{t-j}$$

6.3.3 Limitations of Simple Models

While simple models like AR, MA, and ARMA are useful for basic time series analysis, they have notable limitations:

- **Linearity Assumption**: These models assume linear relationships among variables, which may not always be accurate.
- **Multicollinearity**: High correlation among predictors can lead to unstable estimates and reduce model reliability.
- **Scalability**: Simple models struggle to handle large datasets with numerous predictors, limiting their applicability in more complex forecasting tasks.
- **Variable Selection**: These models do not inherently provide a mechanism for selecting the most relevant predictors, often leading to overfitting or underfitting.

Given the limitations of simple models, we turn to more advanced methodologies like Penalized regression methods and Factor methods that address them.

Penalized regression methods, such as Ridge Regression and LASSO, introduce regularization techniques to handle multicollinearity and improve model interpretability. These methods shrink coefficients and can set some to zero, aiding in variable selection and reducing overfitting. Factor methods like Principal Component Analysis (PCA) and Partial Least Squares (PLS) reduce the dimensionality of the data by creating composite variables (factors) that capture the underlying structure. These techniques are particularly useful in handling high-dimensional data and improving the stability and accuracy of forecasts.

6.4 Penalized Regression Methods

When dealing with numerous explanatory variables, the main objective is to achieve accurate predictions while utilizing as few variables as possible. But why the emphasis on a parsimonious model with only a handful of covariates?

Firstly, a simpler model is easier to interpret, providing better understanding and insight into the underlying relationships between variables. Additionally, a simpler model helps mitigate the risk of overfitting, where the model performs well on the training data but fails to generalize to unseen data. Lastly, reducing the number of covariates helps address multicollinearity issues, where predictor variables are highly correlated, leading to unstable parameter estimates.

Traditionally, selecting covariates involved a trial-and-error process of adding or removing variables based on their significance. However, a more efficient and modern approach, known as penalized regression, has gained popularity. Instead of explicitly including or excluding covariates, penalized regression methods retain all variables but penalize the magnitude of their regression coefficients.

The objective function $\mathcal{L}$ in penalized regression typically combines the ordinary least squares (OLS) loss function with a penalty term:

$$\mathcal{L} = \text{OLS loss function} + \lambda \times \text{Penalty term}$$

where:

- OLS loss function: $\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$, where y_i are the observed values and $\hat{y}_i$ are the predicted values.
- Penalty term: $\lambda \times R(\beta)$, where $R(\beta)$ is a regularization term penalizing the size of the coefficients, and λ is the regularization parameter controlling the strength of the penalty.

In penalized regression, the objective is similar to traditional linear regression, aiming to minimize the sum of squared errors. However, a penalty term is introduced, causing the regression coefficients to shrink towards zero. This regularization or shrinkage mechanism balances the tradeoff between bias and variance in the model.

It's crucial to note that while ordinary least squares (OLS) estimation provides unbiased estimates with minimum variance, penalized regression deliberately introduces bias to reduce variance, ultimately leading to potentially smaller mean squared errors (MSE) in out-of-sample predictions.

While the primary goal for using penalized regression is to improve predictive accuracy, it can also indirectly contribute to enhancing the explanatory performance of the model by promoting simplicity, interpretability, and the selection of relevant covariates.

In our analysis, we will explore how penalized regression works and how it helps tackle the problem of complex models and making better predictions. Techniques like Ridge and Lasso are advanced statistical methods used for regression analysis when dealing with datasets with a large number of predictors, to choose the variables with the largest impact on the response variable. These methods work by adding a penalty term to the regression equation, which helps to shrink the coefficients of less important variables towards zero, effectively reducing model complexity and overfitting.

By integrating penalized regression methods into our forecasting models, we aim to enhance the robustness and reliability of our predictions for Germany, Italy, and Ireland, ultimately providing more insightful economic forecasts.

6.4.1 Ridge Regression

Ridge Regression is a method used in statistics to create a type of linear regression model. We use this method to capture the entire information from the cross-sectional relationships between all the explanatory variables. It adds a penalty to the model that's based on the sum of the squares of the coefficients (the weights applied to each input variable), and this penalty helps shrink the coefficients towards zero, but not exactly to zero. This means that even variables with a minor influence on the outcome will still have some impact on the model, albeit reduced.

In the context of Ridge Regression, we should mention that standardization ensures that all predictors are on the same scale, allowing the ridge penalty to shrink coefficients uniformly and improving the stability and interpretability of the model. Failing to standardize predictors before applying Ridge Regression can lead to biased coefficient estimates and inconsistent interpretations.

The main idea is to find a balance between making accurate predictions and keeping the model simple.

The process of finding the best coefficients involves solving an optimization problem. This problem aims to minimize the difference between the predicted values and the actual values in the dataset, while also adding a penalty term that depends on a parameter called λ. This λ parameter controls how much the coefficients are penalized.

Ridge Regression creates a linear regression model that is penalized with the L2-norm, represented by the term $R(f) = \sum_{i=1}^{K} \beta_i^2 = ||\beta||^2$. This penalty term acts to shrink the coefficient values towards zero, thereby reducing the complexity of the model.

The parameter estimators, $\hat{\beta}_{\text{Ridge}}$, are obtained by solving the following optimization problem:

$$\min_{\beta} \left(\sum_{i=1}^{n} (y_i - x_i^T \beta)^2 + \lambda R(f) \right)$$

Here, λ is the penalty/tuning parameter, which is non negative and x_i^T represents the transpose of the input variables. The solution for $\hat{\beta}_{\text{Ridge}}$ is given by:

$$\hat{\beta}^*(\lambda) = (S + \lambda I)^{-1} X^T Y = (S + \lambda I)^{-1} S \hat{\beta}$$

where $S = X^T X$ and $\hat{\beta} = (X^T X)^{-1} X^T Y = S^{-1} X^T Y$.

It's important to note that $\hat{\beta}^*(\lambda)$ is a shrunken estimator of β. By adding the penalty term, Ridge Regression minimizes the sum of squared residuals while also minimizing the sum of squared beta coefficients. This results in many elements of $\hat{\beta}^*(\lambda)$ being very small (close to zero), but not exactly zero, indicating that Ridge Regression does not perform model selection.

When λ is set to zero, Ridge Regression behaves just like ordinary least squares regression (OLS), which doesn't penalize the coefficients at all. As λ increases, the coefficients get closer to zero, but they never become exactly zero unless λ becomes infinitely large ($\lambda \to \infty$). By increasing λ, the penalty term becomes more dominant in the objective function. As a result, the solution that minimizes the objective function will have smaller coefficient values, leading to a simpler model.

The fact that coefficients are moving towards zero, helps stabilize the model and reduce sensitivity to multicollinearity. This balance between simplicity and accuracy makes Ridge Regression a valuable tool for regression analysis, particularly in datasets with multiple variables.

However, Ridge Regression does not shrink any β_j to zero, preventing exact variable selection. Manual intervention to set coefficients close to zero may be possible, but it's not the ideal solution. To address this limitation, another method called the Lasso is often employed.

6.4.2 LASSO

The Lasso method adds a penalty term equal to the absolute value of the coefficients. This method aims to minimize the difference between predicted and actual values while ensuring that the sum of the absolute values of the coefficients is below a certain limit. This constraint often results in some coefficients being exactly zero, making the model easier to interpret. Studies show that LASSO combines the best of both subset selection and ridge regression: it provides interpretable models like subset selection and the stability of ridge regression.

The Least Absolute Shrinkage and Selection Operator (LASSO) is a regression method developed relatively recently by Robert Tibshirani in 1996 and revolutionized regression analysis by offering a powerful tool for feature selection and model regularization.

The optimization problem remain:

$$\min_{\beta} \left(\sum_{i=1}^{n} (y_i - x_i^T \beta)^2 + \lambda R(f) \right)$$

The difference in LASSO, lies in the penalty term, denoted as $R(f)$, that is based on the sum of the absolute values of the coefficients, represented by the L1-norm: $R(f) = \sum_{i=1}^{p} |\beta_i| = ||\beta||_1$. This penalty encourages sparsity in the model by shrinking some coefficients to exactly zero, effectively performing both regularization and model selection simultaneously.

The optimization problem in LASSO aims to minimize the sum of squared residuals, similar to ordinary least squares regression, while also adding a penalty term that depends on a parameter λ.

However, unlike Ridge Regression, LASSO does not have a closed-form solution. Instead, it relies on convex optimization techniques to find the parameter estimators, $\hat{\beta}^*(\lambda)$.

As a result of the L1-norm penalty, many elements of $\hat{\beta}^*(\lambda)$ will be shrunk towards zero, similar to Ridge Regression. However, LASSO differs in that it explicitly sets some coefficients to zero, effectively performing both regularization and model selection simultaneously.

Simulation studies suggest that LASSO enjoys some of the favorable properties of both subset selection and ridge regression. Compared to Ridge Regression, LASSO offers distinct advantages. While Ridge Regression penalizes the sum of squared coefficients (L2 penalty), LASSO's L1 penalty explicitly sets some coefficients to zero, resulting in sparse models that involve only a subset of the variables. Consequently, models generated from LASSO are generally easier to interpret than those produced by Ridge Regression.

Although explicit formulas for the bias and variance of the LASSO estimator are not available, it is generally observed that the bias increases and the variance decreases as λ increases.

Selecting an appropriate value for the tuning parameter λ is crucial in LASSO regression. Techniques like 10-fold cross-validation are commonly employed to determine the optimal λ value, ensuring the balance between model regularization and sparsity.

In summary, LASSO strikes a balance between regularization and model sparsity, making it particularly useful for feature selection in regression analysis.

6.4.3 Limitations and Decision between Adaptive LASSO and Adaptive Ridge

When selecting the appropriate adaptive method for our forecasting analysis, it's crucial to consider the primary objectives. If our primary goal is to optimize prediction accuracy while maintaining some level of interpretability, Adaptive Ridge might be preferred. Ridge Regression performs well in situations where there is multicollinearity among predictor variables, and it doesn't set coefficients exactly to zero. This leads to more stable predictions, which can be crucial for accurate forecasting.

However, if the goal is to achieve high interpretability and understand the relative importance of predictor variables, Adaptive LASSO is more suitable. Adaptive LASSO sets some coefficients exactly to zero, providing a clear indication of which variables are considered important by the model. This method can help us identify a parsimonious set of predictors, making the model easier to interpret and providing insights into the key drivers of the forecast.

For our forecasting analysis of Germany, Italy, and Ireland, we prioritize both predictive accuracy and interpretability. Therefore, a balanced approach might be to initially use Adaptive Ridge to ensure stable and accurate forecasts, followed by Adaptive LASSO to refine the model and highlight the most significant predictors. This two-step approach allows us to harness the strengths of both methods, ensuring robust and interpretable forecasting models.

By proceeding with Adaptive LASSO, we can identify the most critical variables influencing our forecasts, helping us make informed decisions and providing clearer insights into the economic factors driving the predictions. This method will be instrumental in refining our forecasting models to enhance both their accuracy and interpretability, ultimately leading to more reliable and actionable economic forecasts for Germany, Italy, and Ireland.

6.4.4 Adaptive LASSO

As we observed, the LASSO approach is democratic – it applies the same penalty to all coefficients irrespective of their inherent importance and this can potentially shrink vital coefficients too much.

To address this limitation, we introduce "adaptive weights," tailoring the penalty for each coefficient based on its significance. These weights are calculated from initial estimates, often obtained via Ridge Regression, known for avoiding zero coefficients.

The Adaptive LASSO (A-LASSO) is an extension of the LASSO method introduced by Zou in 2006. It addresses some limitations of traditional LASSO by re-weighting the L1-norm penalty term based on initial estimates of the coefficients. In a linear regression, the Adaptive LASSO seeks to minimize:

$$\min_{\beta} \left(\sum_{i=1}^{n} (y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2 + \lambda \sum_{j=1}^{p} w_j |\beta_j| \right)$$

where: where λ is the tuning parameter (chosen through 10-fold cross validation), w_j is the adaptive weight assigned to the j-th coefficient is are calculated deterministically based on initial estimates of the coefficients and β_j are the estimated p coefficients.

In the A-LASSO optimization problem, the penalty term is adjusted using weights derived from an initial estimator, denoted as β_{init}. These weights, represented as $w_j = 1/|\beta_{\text{init},i}|^{\gamma}$, where $\gamma > 0$, are used to re-scale the penalty for each coefficient.

The initial estimator β_{init} is often obtained using the Ridge Regression method with the constraint parameter tuned through cross-validation. Then, in the second stage, cross-validation is again employed to select the λ parameter in the optimization problem.

Conditions outlined by Haung et al. (2008) ensure the effectiveness of A-LASSO in large samples. These conditions include assumptions about the errors, the consistency of initial estimators, and the adaptive irrepresentable condition, among others. These conditions aim to guarantee the accuracy of variable selection and the asymptotic normality of the A-LASSO estimator in large samples.

The intuition behind the adaptive weights is straightforward. If an initial coefficient is large (indicating a potentially important feature), the corresponding adaptive weight will be small. This effectively reduces the penalty for that coefficient, allowing it to retain a larger value. Conversely, small initial estimates (suggesting less important features) will result in larger weights, amplifying the LASSO penalty and potentially pushing the coefficient closer to zero. This adaptivity enhances the variable selection capabilities of the LASSO method, making it more effective in scenarios where traditional LASSO may struggle.

6.5 Factor Methods

In economics, particularly when dealing with large datasets, the focus often shifts from pinpointing individual important variables to embracing the collective significance of all variables. Instead of selectively choosing variables based on their importance, economists may adopt the perspective that all variables contribute valuable information to the dataset.

To extract insights from the variables, we construct an index that encapsulates the essence of each variable's contribution. This index, derived from a weighted combination of variables, aims to capture the overall information present in the dataset. However, manually assigning weights can be problematic, as it may not accurately reflect their true significance, especially considering variations in scaling among variables.

To address this challenge, we use factor methods such as Principal Component Analysis (PCA) and Partial Least Squares (PLS). These dimensionality reduction techniques simplify high-dimensional datasets for regression analysis by creating new variables, called factors or components, which are combinations of the original variables. These methods condense vast amounts of data into a few key summaries.

PCA identifies loadings (weights) that effectively combine explanatory variables to maximize variance, thereby capturing the inherent variability within the dataset. By standardizing the variables, PCA ensures a fair comparison across all variables, enabling the identification of loadings that best represent the collective information.

PLS aims to find linear combinations of the predictor variables that are most correlated with the response variable, with the objective of maximizing the covariance between the predictors and the response. Unlike PCA, which focuses solely on maximizing variance in the predictor variables, PLS considers both the predictor and response variables simultaneously. By iteratively estimating latent variables (also known as components or factors) that capture the maximum covariance between X and Y, PLS seeks to extract the most relevant information from both sets of variables. Similar to PCA, PLS often involves standardizing the variables to ensure comparability and enhance the stability of the estimation process, ultimately leading to a robust and interpretable model.

Dimension reduction methods, such as PCA and PLS, transform predictors and fit a least squares model using the transformed variables, making the analysis more manageable and interpretable. Let $Z_1, Z_2, \ldots, Z_M$ represent $M < p$ linear combinations of our original p predictors, defined as:

$$Z_M = \sum_{j=1}^{p} \phi_{jm} X_j$$

for constants $\phi_{1m}, \phi_{2m}, \ldots, \phi_{pm}$, $m = 1, \ldots, M$. The linear regression model becomes:

$$y_t = \theta_0 + \sum_{m=1}^{M} \theta_m z_{im} + \epsilon_i$$

where $\theta_0, \theta_1, \ldots, \theta_M$ are the regression coefficients. This approach reduces the problem from estimating $p + 1$ coefficients to $M + 1$ coefficients, thus simplifying the model and potentially improving estimation accuracy.

In summary, PCA focuses on maximizing variance to capture the most critical data variability, while PLS maximizes covariance between predictors and the response variable to improve prediction accuracy. These methods help manage large datasets by reducing dimensionality, ensuring more robust and interpretable models.

6.5.1 Principal Component Analysis - PCA

Principal Component Analysis (PCA) is a powerful technique commonly used for dimensionality reduction, particularly with large datasets and panel data. Unlike methods relying on assumptions of independent and identically distributed (iid) variables, PCA aims to capture the underlying structure and relationships among variables without specific distributional assumptions.

PCA serves several key functions in econometrics. It reduces the number of variables by transforming them into principal components that explain most of the variance, thus simplifying the dataset while retaining its essential information. This dimensionality reduction allows researchers to focus on a few principal components, making it easier to interpret the underlying structure and relationships within the data. Additionally, PCA mitigates multicollinearity in forecasting models, leading to more reliable and robust forecasts by ensuring that the predictor variables are not highly correlated with each other. This makes PCA a valuable tool in both data analysis and predictive modeling.

PCA calculates principal components, which are linear combinations of the original variables capturing the most variation in the data. It identifies patterns in data and simplifies high-dimensional datasets by transforming a large set of variables into a smaller set that still contains most of the original information. For instance, the first principal component (PC1) that we will use on all our variables to build a Coincident Indicator can be likened to a cocktail recipe where we blend parts of explanatory variables, with the mix ratios indicating the importance of each variable in describing data spread.

To ensure effective dimensionality reduction and data analysis, PCA involves several critical steps. The first step is standardization, which ensures that each variable contributes equally by transforming the data to have zero mean and unit variance. This enables a fair comparison across all variables and allows for the identification of loadings that best represent the collective information contained in the dataset. This is followed by the computation of the covariance matrix, a square matrix that shows the covariance between each pair of variables, helping to understand their interactions. Next, eigenvalue and eigenvector calculation is performed to solve for eigenvalues, which indicate the variance captured by each principal component, and eigenvectors, which define the principal components' directions. Finally, principal components selection involves retaining the first few principal components that capture most of the variance in the data, thus simplifying the dataset while preserving its essential information.

Although the first principal component efficiently summarizes a significant portion of the explanatory variable variance, it typically does not encompass all variance. To discern the unaccounted variance, regression techniques are employed. Each explanatory variable is regressed against the first index, treating it as independent and the original variable as dependent. The resulting residuals reveal information that the initial PCA analysis overlooked.

PCA minimizes the sum of squared errors between observed variables and their linear combinations, seeking orthogonal vectors that best represent data variance. The first principal component loading vector maximizes:

$$\max_{\boldsymbol{\varphi}_1} \left\{ \frac{1}{n} \sum_{t=1}^{n} \left(\sum_{j=1}^{p} \varphi_{j1} x_{jt} \right)^2 \right\}$$

subject to:

$$\sum_{j=1}^{p} \varphi_{j1}^2 = 1$$

Here, φ_{j1} is the loading vector for the first principal component, and x_{jt} is the standardized value of the j-th variable at time t.

Subsequently, the second principal component, Z_2, is determined to maximize the variance among linear combinations uncorrelated with Z_1, iterating for subsequent components.

PCA estimates loadings and factors by solving the eigenvalue-eigenvector problem of the data's second moment matrix. This technique endeavors to identify loadings (weights) that effectively combine explanatory variables to maximize variance, thereby capturing the inherent variability within the dataset.

PCA minimizes the sum of squared errors between observed variables and their linear combinations, finding orthogonal vectors representing data variance.

While it simplifies big datasets, identifies patterns, and aids forecasting, it is unsupervised and doesn't consider relationships with response variables. For incorporating response variable information, techniques like Partial Least Squares (PLS) are preferred.

6.5.2 Partial Least Squares - PLS

Partial Least Squares (PLS) is a method used for regression when there are many predictors compared to the number of observations. It's similar to Principal Component Analysis (PCA) in that it creates linear combinations of the original variables to use as regressors.

The basic idea of PLS is to find linear combinations of the predictor variables (x_t) that have the highest covariance with the response variable (y_t), while also being orthogonal to each other. This ensures that the constructed factors capture as much relevant information from the predictors as possible without redundancy.

Here's a simplified version of the algorithm to construct k PLS factors:

1. Initialize $u_t = y_t$ and $v_{i,t} = x_{i,t}$ for $i = 1, \ldots, N$. Set $j = 1$.

2. Determine the vector of weights or loadings (w_j) by computing individual covariances: $w_{ij} = \text{Cov}(u_t, v_{i,t})$ for $i = 1, \ldots, N$. Construct the j-th PLS factor ($f_{j,t}$) by taking the linear combination given by $w_j^T v_t$.

3. Regress u_t and $v_{i,t}$ for $i = 1, \ldots, N$ on $f_{j,t}$. Denote the residuals of these regressions by $\tilde{u}_t$ and $\tilde{v}_{i,t}$ respectively.

4. If $j = k$, stop. Otherwise, set $u_t = \tilde{u}_t$, $v_{i,t} = \tilde{v}_{i,t}$ for $i = 1, \ldots, N$, increment j, and return to step 2.

Once the PLS factors are constructed, the response variable y_t can be modeled or forecasted by regressing it on $f_{j,t}$ for $j = 1, \ldots, k$.

PLS, balancing predictor and response information, offers a robust and interpretable model for regression analysis.

6.5.3 Comparison with PCA

The key difference lies in how these factors are constructed. PCA focuses only on the independent variables to create factors, while PLS considers both the independent and dependent variables, maximizing the covariance between the dependent variable and the independent, making PLS potentially more suitable for regression analysis when there's a strong relationship between the independent and dependent variables. However, we should also mention that PLS hasn't been extensively explored for datasets with a very large number of variables.

6.6 Machine Learning Methods

To enhance the robustness of our analysis, we employ machine learning models due to their ability to capture non-linear relationships that were previously unaccounted for in our models.

Despite the potential challenges in interpreting these models from an economic perspective, their capacity to provide highly accurate forecasts renders them invaluable. Machine learning techniques, such as Random Forest and Gradient Boosting, excel in identifying complex, non-linear relationships among variables, which traditional econometric models may overlook. These techniques aggregate multiple models to improve predictive accuracy, often outperforming single econometric models. In particular, neural networks are adept at modeling non-linear interactions between variables, offering more flexible and precise predictions in intricate economic environments.

Additionally, machine learning methods can automate numerous steps in the model building and validation processes, thereby conserving time and minimizing the potential for human error.

By leveraging automation, these techniques enhance efficiency and contribute to more consistent and reliable forecasting outcomes.

6.6.1 Random Forest

In the preceding sections, we explored approaches centered on defining a parametric model, often involving linear regression, to establish the relationship between the target variable y and a potentially extensive set of explanatory variables x.

Regression trees are constructed by partitioning the space of the dependent variable y into M subsets R_m, with y allocated to each subset according to a specific rule and modeled as a distinct constant c_m within each subset. This approach is particularly powerful as it can accommodate various functional relationships between y and a set of explanatory variables x, such as $y = f(x)$, without imposing the assumptions of linearity or additivity commonly associated with standard linear regression models. Let

$$Y = f(x) = \sum_{m=1}^{M} c_m \mathbf{1}(x \in R_m)$$

where 1 denotes the indicator variable taking value 1 if the condition is satisfied, 0 otherwise.

Then, given a partition, minimising:

$$\|y - f(x)\|_2 = \sum_{i=1}^{N} (y_i - f(y_i))^2$$

A significantly more challenging task is identifying the optimal partition that minimizes the sum of squares. Even in the two-dimensional case, where $N = 2$ and $X = [x_1, x_2]$, determining the best binary partition to minimize the sum of squares is computationally infeasible. Consequently, avaricious algorithms are typically employed. These algorithms perform one split at a time by considering a splitting variable j (where $j = 1, \ldots, k$) and a split point s.

A region $R_1(j, s)$ is then defined as,

$$R_1(j, s) = \{X \mid X_j \leq s\} \; and \; R_2(j, s) = \{X \mid X_j > s\}$$

Then, the sum of squares is minimized with respect to j and s. For each splitting variable, the optimal split point s can be identified. By evaluating all variables X_j, the optimal pair (j, s) can be determined. Upon identifying the best split, the data are partitioned into two resulting regions, and the splitting procedure is repeated for each of the two regions. This process continues recursively for all resulting regions. The depth of the resulting tree is determined by the number of iterations of the algorithm. While shallow trees may fail to capture the underlying structure of the data, deeper trees risk overfitting, potentially leading to poor predictive performance.

Random forests, extend this concept by employing an ensemble approach akin to bagging applied to regression trees. The idea is to grow a large collection of de-correlated trees (hence, the name forest) and then average them. This is achieved by bootstrapping a random sample at each node of every tree. To induce “decorrelation” among trees, when growing them, a subset of the input variables is randomly selected before each split as candidates for splitting. This random selection helps prevent “strong” predictors from imposing too much structure on the trunk of the tree.

Random Forests are made up of decision trees. After creating multiple decision trees, the majority vote is taken to make predictions. Random Forests combine the simplicity and flexibility of decision trees, improving prediction accuracy.

In this exercise, Random Forest is utilized to predict the Industrial Production Index. Its robustness against overfitting and ability to handle high-dimensional data makes it suitable for capturing complex economic interactions. Below, we present the mathematical formula:

$$\hat{f}(x) = \frac{1}{N} \sum_{i=1}^{N} h_i(x) \tag{1}$$

where $h_i(x)$ is the i-th decision tree, and N is the total number of trees. Each tree h_i is trained on a random subset of the data (sampled with replacement). At each split in a tree, a random subset of features is selected to determine the best split:

$$\text{Best Split} = \arg\min_{j,s} \left(\sum_{i \in R_1(j,s)} (y_i - \hat{y}_{R_1})^2 + \sum_{i \in R_2(j,s)} (y_i - \hat{y}_{R_2})^2 \right) \tag{2}$$

where R_1 and R_2 are the two resulting regions from the split based on feature j and threshold s.

6.6.2 Gradient Boosting

In this research, Gradient Boosting is employed to forecast economic indicators by iteratively improving the model's predictions. This technique's ability to handle non-linear relationships and interactions makes it particularly effective for economic data forecasting.

Gradient Boosting for regression starts with an initial prediction based on the average of the target variable. It then iteratively builds decision trees to predict the residuals (the differences between observed and predicted values) from the previous model. Each tree is added to the model in a scaled manner using a learning rate, which ensures incremental improvement in predictions. This process continues until a specified number of trees are built or adding more trees no longer significantly reduces the residuals, thereby refining the model's accuracy with each step.

Gradient Boosting builds models sequentially, where each new model attempts to correct the errors made by the previous models. It combines weak learners to form a strong learner, focusing on the mistakes of previous models. Below we present the additive model:

$$\hat{f}(x) = \sum_{m=1}^{M} \lambda h_m(x) \tag{3}$$

where $h_m(x)$ is the m-th weak learner, M is the number of iterations, and λ is the learning rate.

The method minimizes the loss function $L(y, \hat{f}(x))$:

$$\theta_m = \arg\min_{\theta} \sum_{i=1}^{N} L(y_i, \hat{f}_{m-1}(x_i) + \lambda h_m(x_i, \theta)) \tag{4}$$

where θ_m are the parameters of the m-th weak learner, and N is the number of training samples.

6.6.3 Neural Networks

Neural networks are popular machine learning algorithms often seen as a "black box" because their operations can be hard to understand. They consist of nodes (neurons) organized into layers: the input layer, hidden layers, and output layer. They were named because early researchers thought the nodes and connections resembled neurons and synapses in the brain.

Artificial Neural networks are commonly employed for forecasting. The one-step ahead forecast $\hat{y}_{t+1}$ is computed using inputs that are lagged observations of the time series or other explanatory variables. Let I denote the number of inputs p_i of the ANN. Its functional form can be described as:

$$\hat{y}_{t+1} = \alpha + \sum_{h=1}^{H} \beta_h \cdot g \left(\xi_{0h} + \sum_{i=1}^{I} \xi_{hi} \cdot p_i \right)$$

where $\mathbf{w} = (\beta, \xi)$ are the network weights, $\beta = [\beta_1, \ldots, \beta_H]$ are the output layer weights, $\xi = [\xi_{11}, \ldots, \xi_{HI}]$ are the hidden layer weights, α is the bias term, γ_{0h} are bias terms for each hidden node h, H is the number of hidden nodes, and $g(\cdot)$ is a nonlinear activation function. Commonly used activation functions include:

- Logistic sigmoid: $\sigma(z) = \frac{1}{1+e^{-z}}$
- Rectified Linear Unit (ReLU): $\sigma(z) = \max(0, z)$

In the context of time series forecasting, neural networks can be viewed as nonlinear autoregressive models. They utilize lags of the time series, possibly together with lagged observations of explanatory variables, as inputs to the network. During training, pairs of input vectors and target outputs are presented to the network. The network's output is compared to the target, and the resulting error is used to update the network weights.

Training an ANN is a complex nonlinear optimization problem, often susceptible to getting trapped in local minima of the error surface. To mitigate this, training should be initialized multiple times with different random starting weights and biases to explore the error surface more comprehensively.

For the purpose of this exercise we apply Neural Networks to forecast the Industrial Production Index, leveraging their ability to learn from large datasets and model complex relationships.

6.7 Model Averaging of Different Models

After using a variety of individual models for forecasting, a potentially more robust approach emerges. "Could model combination, specifically model averaging, be a good fit for forecasting?"

6.7.1 Advantages of Model Averaging

Model averaging has several advantages that make it an attractive option for forecasting:

- **Stability**: Model averaging tends to be more stable than individual models. By combining multiple models, the resulting forecasts are less prone to sudden changes and fluctuations that a single model might exhibit. This reduced volatility stems from the averaging process, which smooths out the extremes.
- **Robustness**: Averaging the predictions of several models can provide a more robust forecast. This robustness arises because the strengths of different models can compensate for their individual weaknesses. If one model performs poorly under certain conditions, others might perform better, leading to an overall improved forecast.

6.7.2 Disadvantages of Model Averaging

However, there are also some disadvantages to consider:

- **Slower Adaptation**: While individual models can quickly adapt to new data and changing conditions, model averaging does not adapt as swiftly. The averaged forecast will only change significantly when most of the constituent models have adapted to the new data or circumstances. This lag can be a disadvantage in highly dynamic environments where quick adaptation is crucial.
- **Complexity**: Implementing model averaging adds complexity to the forecasting process. It requires maintaining multiple models and ensuring they are all appropriately tuned and updated, which can be resource-intensive.

6.7.3 Static Averages

We will initially deal with static averages, meaning fixed weights for each model in the averaging process. The predictions will change because, on average, the models have changed. To evaluate the performance of these models, we will use the relative Root Mean Square Forecast Error (relRMSFE). This metric allows us to determine which model has the smallest error.

By comparing the relRMSFE of the model constructed using all individual models with other models, we can assess the effectiveness of model averaging. If the difference in relRMSFE is minimal, it may be preferable to use the model constructed with all the models to achieve more robust predictions. This approach can help reduce volatility in our forecasts, as the static average model leverages the strengths of multiple models, providing a smoother and more reliable prediction.

In our forecasting methodology, we include several models such as AR(1), Random Forest, and ARMA models, etc. Static model averaging involves assigning fixed weights to these models and combining their forecasts. The forecasts are generated using individual models like Penalized Regression models (Ridge, LASSO, Adaptive Ridge and Adaptive LASSO). These forecasts are then combined using static model averaging, where fixed weights are assigned to each model. For instance, the mean of the forecasts from the selected models can be used to produce the final forecast.

Static model averaging reduces volatility by averaging the forecasts from multiple models, thereby minimizing the impact of any single model's volatility. This leads to smoother and more reliable predictions. Additionally, combining models allows us to leverage the strengths of different models, compensating for their individual weaknesses. Overall, static model averaging provides a robust approach to forecasting, as it minimizes the risk associated with relying on a single model.

To evaluate the performance of the static model averaging approach, we use the relative Root Mean Square Forecast Error (relRMSFE). By comparing the relRMSFE of the combined model with other individual models, we can assess its effectiveness. If the relRMSFE difference is minimal, the combined model may be preferable due to its robustness and reduced volatility.

This methodology ensures that our forecasts are more stable and reliable, ultimately leading to better decision-making based on the predicted data. It is also simpler but less responsive to recent changes.

To account for recent changes, we use another method of model averaging that assigns weights to the models differently, making it more complex.

6.7.4 Dynamic Ranking

Dynamic ranking is used to evaluate and rank models based on their performance in recent periods.

The models are evaluated based on the forecast errors in the last period, such as Mean Absolute Error (MAE) and Root Mean Square Forecast Error (RMSFE). Models are ranked in ascending order of their forecast errors. The best-performing model has the smallest error. For instance, if AR(1), the benchmark, has the smallest MAE in the last period, it will be ranked first.

The top-performing models (e.g., Top 5, Top 10) are selected based on their rankings. This selection process helps in focusing on the most accurate models. For example, Top5-MAE-Last3 refers to the top 5 models based on MAE in the last 3 periods.

Weights are assigned to the selected models based on their performance. Models with smaller errors receive higher weights. The weights ensure that models with better recent performance have a greater influence on the combined forecast.

The forecasts from the top models are combined using the calculated dynamic weights to produce a more robust forecast. This approach leverages the strengths of multiple models and reduces the risk of over-reliance on a single model.

The models are also evaluated based on their performance over the last few periods (e.g., last 3 periods). Aggregated performance metrics (e.g., average MAE, RMSFE) over the recent periods are used for ranking. Similar to the single-period ranking, dynamic weights are assigned based on the aggregated metrics, and the forecasts are combined. This helps in capturing the consistent performance of models over a period, rather than relying on a single observation.

6.7.5 Theoretical Robustness

Theoretically, model averaging is considered more robust because it mitigates the risk of large forecast errors that might occur with individual models. An individual model might produce excellent forecasts in some cases but fail in others. By averaging multiple models, we reduce the likelihood of extreme errors, leading to more consistent and reliable forecasts.

In summary, model averaging offers a promising approach to forecasting by leveraging the strengths of multiple models. However, its actual performance must be validated through empirical testing. In our analysis, we include model averaging alongside individual models to determine the most effective forecasting strategy for our data and objectives.

By following this structured approach, in the forecasting method we ensure that the models are continuously evaluated and updated based on their recent performance, making the forecasting framework more adaptive and robust to changes in the data.

This comprehensive procedure helps in achieving a balance between prediction accuracy and stability, leveraging the strengths of individual models while mitigating their weaknesses through averaging and dynamic weighting.

6.8 Recursive and Rolling Window Estimation Methods

To further enhance the robustness of our forecasting models, we explore two estimation options: recursive and rolling methods. By comparing forecasts from both approaches, we aim to conduct a comprehensive analysis for different forecasting horizons (short-term, medium-term, and long-term) under various economic conditions.

6.8.1 Recursive Estimation

Recursive estimation involves updating the forecasting model each time a new data point is received, incorporating all available data up to the current time. This continuous updating process ensures that the model reflects the most recent information.

In this method, the model recalibrates itself whenever new data is introduced, thus continuously improving its predictive accuracy, which is effective in scenarios where past data remains relevant over time, allowing the model to leverage the full historical dataset.

In our case, this method could be beneficial for understanding the underlying stability and long-term dynamics of the economies.

6.8.2 Rolling Window Estimation

Rolling estimation, on the other hand, uses a fixed-size window (Nroll) of the most recent data points to update the model. As new data becomes available, the oldest data within the window is discarded, maintaining a constant window size, which could be ideal for environments where recent changes are more indicative of future trends.

In our case, the rolling window can help us for the short-term forecasting and in understanding how the economies react to new information and changes.

6.9 Forecasting Metrics

To evaluate the effectiveness of our forecasting models, we rely on various metrics. These metrics help us understand not just the accuracy of our predictions but also other important aspects such as consistency and directional correctness.

6.9.1 Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) measures the average magnitude of errors in a set of predictions, without considering their direction. It is calculated as:

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|$$

where:

- y_i represents the observed value for the i-th observation.
- $\hat{y}_i$ represents the predicted value for the i-th observation.
- N is the total number of observations.

6.9.2 Root Mean Squared Forecast Error (RMSFE)

The Root Mean Squared Forecast Error (RMSFE) accounts for both the bias and variance of forecast errors, defined as:

$$\text{RMSFE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$

RMSFE penalizes larger errors more severely than smaller ones, providing a measure of overall forecast accuracy and reliability.

6.9.3 Success Sign Ratio

The success sign ratio is a measure used to evaluate the predictive accuracy of an econometric model, particularly in terms of its ability to correctly predict the direction of change in the variable of interest. It is defined as:

$$\text{Success Sign Ratio} = \frac{\sum_{t=1}^{T} I\left((\hat{y}_t - \hat{y}_{t-1})(y_t - y_{t-1}) > 0\right)}{T}$$

where:

- $\hat{y}_t$ is the predicted value of the dependent variable at time t.
- y_t is the actual value of the dependent variable at time t.
- $I(\cdot)$ is an indicator function that equals 1 if its argument is true and 0 otherwise.
- T is the total number of observations.

By focusing on the direction of changes rather than their magnitude, the success sign ratio provides a complementary measure to traditional accuracy metrics such as Root Mean Squared Error (RMSFE) or Mean Absolute Error (MAE), which assess the size of prediction errors.

6.10 Comparative Analysis

In our comparative analysis, we establish the relative MAE of the AR(1) model as 1. This serves as a reference point: a relative MAE greater than 1 indicates inferior performance compared to the benchmark, while a relative MAE less than 1 suggests superior performance.

6.11 Statistical Assessment

To statistically assess forecasting accuracy, we employ the Diebold-Mariano (DM) Test. This test determines whether differences in predictive accuracy between competing models are statistically significant. Specifically, we conduct DM tests for both MAE and RMSFE metrics to comprehensively compare model performance. A significant result from the DM test indicates that one model exhibits statistically superior predictive accuracy over another.

7 Empirical output discussion

7.1 Germany's Performance Discussion

7.1.1 Recursive Performance Metrics

For Germany, in short-term forecasting with recursive estimation, we evaluated several models. We begin the analysis by comparing the models based on their Mean Absolute Error (MAE) and Root Mean Squared Forecast Error (RMSFE).

In Table 3., in the short-term, the individual models that show the lowest MAE are the Random Forest, closely followed by ARMA(1,1) and the benchmark model AR(1). Notably, Random Forest also shows a high Sign Success Ratio (SSR), indicating superior directional accuracy, which underscores its robustness and effectiveness in short-term forecasting.

When we incorporate model averaging, the Average Machine Learning (Avg-ML) model demonstrates the lowest MAE and RMSFE among the averaged models, while the Random Forest model remains competitive, confirming its consistent performance.

Furthermore, the introduction of dynamic ranking further enhances model performance. The Top10-MAE-Last3 and Top5-MAE-Last3 models, along with their dynamic-weight versions (Top10-MAE-Last3-DYNW and Top5-MAE-Last3-DYNW), dominate in terms of the lowest MAE. Dynamic weighting adjusts model weights based on their recent performance, leading to more adaptive and potentially more accurate forecasting. The better a model performed recently, the higher its weight in the combination.

This shows that in the short-term, these combined models outperform others, while the Average Machine Learning models and the Random Forest model are strong contenders.

In the medium-term, individual models become more prominent. The ARMA(Auto) model outperforms others with the lowest MAE and RMSFE, followed closely by the Average ARMA model and penalized regression models like Ridge and LASSO. Interestingly, the first appearance of a Top 10 model by lower MAE over the last three periods comes much later in the list (15th out of 20 models), suggesting a preference for individual models over averaged models for medium-term forecasting. This indicates a shift in the forecasting strategy where more weight is given to individual models that can adapt better over a medium-term horizon.

In the long-term, the Random Forest model stands out, leading with the lowest MAE, followed by ARMA(Auto) and averaging models. The average of all models, despite having a slightly larger MAE and RMSFE, ranks third. This suggests that choosing a more robust averaged model of all models can be beneficial compared to relying solely on individual models like Random Forest or ARMA(Auto). The average of all models offers robustness by leveraging the strengths of the multiple models, ensuring more stable and reliable performance over a longer horizon.

In addition to these observations, several other aspects are worth considering to provide a comprehensive analysis. The stability and robustness of each model over different time horizons are crucial. Models like Random Forest and ARMA(Auto) demonstrate consistent performance, indicating their reliability.

Moreover, the complexity and computational demands of each model are important for practical implementation. For instance, while Random Forest is highly accurate, it may require more computational resources compared to simpler models. The impact of dynamic weights in model averaging is another critical aspect. Dynamic weights significantly enhance forecasting accuracy by adapting to recent performance, making models like Top10-MAE-Last3-DYNW more reliable.

Additionally, analyzing the error distribution helps in understanding how different models handle outliers and variance. Models with lower RMSFE are better at managing variance, which is an important consideration for forecasting accuracy. Lastly, evaluating model performance under various economic scenarios or shocks can provide insights into their reliability and robustness, aiding in better decision-making and strategy formulation.

Interesting insights from the Table 5., reveal that in the short-term period, the Top5-MAE-Last3 and Avg-ML models maintain strong performance relative to the benchmark AR(1) based on the relative MAE (relMAE) and relative RMSFE (relRMSFE) metrics, that are below 1.

In the medium-term, the ARMA(Auto) model stands out with a relMAE of 0.9656 and a relRMSFE of 0.9732, showing superior performance compared to the benchmark AR(1), while Ridge and Lasso models also show competitive performance with relMAE values of 0.9980 and 0.9992.

Finally in the long-term, AR(2), AR(3), ARMA(Auto) and Avg-ARMA perform slightly better than AR(1) but without significant improvements.

In conclusion, short-term forecasting favors Averaged Models, which show excellent performance metrics. This could be happening because short-term data can be more volatile and prone to random fluctuations, so averaged models can filter out some of this noise, leading to better performance metrics.

In the medium-term, there is a noticeable shift towards individual models such as ARMA(Auto) and penalized regression models like Ridge and LASSO. Averaged models might perform well in the short-term due to their ability to handle noise, but as the forecast horizon extends, they might suffer from overfitting to short-term patterns. This can reduce their effectiveness compared to well-calibrated individual models.

For long-term forecasting, the sophisticated Random Forest model again demonstrates superior performance, probably due to capturing complex interactions and dependencies, but the average of all models, despite a slightly higher MAE and RMSFE, provides a robust and reliable alternative, when dealing with long-term uncertainties, emphasizing the importance of considering model averaging to achieve stable and accurate forecasts.

7.1.2 Rolling Window Performance Metrics

In this section, we present the forecasting metrics results with rolling estimation.

In Table 4., in the short-term, models like PLS(2), AR(2), and Random Forest exhibit strong performance. The PLS(2) model, in particular, achieves the lowest MAE, suggesting it handles short-term volatility well. AR(2) and Random Forest also show competitive results, though Random Forest has a higher Sign Success Ratio (SSR), indicating some trade-offs in directional accuracy.

Models such as PLS(1), Adaptive Ridge, ARMA(Auto) and Lasso struggle with higher RMSFE, suggesting they might not handle short-term noise as effectively.

Moving forward to the medium term, individual models like PLS(1), AdRidge, and Ridge emerge as top performers. PLS(1) leads with the lowest MAE, closely followed by Adaptive Ridge and Ridge. This shift highlights the effectiveness of penalized regression models in capturing medium-term patterns and dependencies.

While the AR(1) model remains competitive, it's outperformed by these specialized regression models. Random Forest and ARMA(Auto) also perform well, though their higher RMSFE values suggest some difficulty in managing medium-term variations.

When it comes to the long-term forecasting, models such as ARMA(Auto), linear regression models like LinReg-HD10 and Avg-ARMA stand out. ARMA(Auto) has the lowest MAE and RMSFE, indicating its robustness in long-term forecasting scenarios. Although Random Forest maintains good performance, its higher RMSFE and SSR point to challenges in sustaining accuracy over extended periods.

The average of all models comes to the 14th place out of the 20 models, not reinforcing the point of view that model averaging of all models achieves stable and reliable forecasts. This is particularly crucial for long-term predictions where individual model performance can be more erratic.

In Table 6., we observe that in the short-term, AR(2) and Random Forest demonstrate relative MAE and RMSFE values below 1, outperforming the benchmark AR(1).

In the medium-term, AR(2) maintains strong performance with a relative MAE just below 1 and statistically significant enhancements over AR(1), underscoring the preference for individual models that adeptly capture medium-term trends.

For long-term forecasts, ARMA(Auto) excels with statistically significant improvements in both MAE and RMSFE compared to AR(1). Models like AR(2) and AR(3) also show marginal improvements.

In summary, short-term forecasting with rolling estimation favors individual models like PLS(2), AR(2), and Random Forest, which excel in managing short-term volatility. This success can be attributed to their robust handling of noise and fluctuations. In the medium-term, the focus shifts towards penalized regression models like Ridge and AdRidge, which effectively capture medium-term dependencies. For long-term forecasts, sophisticated models like ARMA(Auto) demonstrate superior performance by capturing complex interactions and dependencies over extended periods.

7.1.3 Performance Conclusion for Germany

Recursive estimation, as we mentioned earlier, updates the model parameters by incorporating new data as it becomes available, typically without discarding old data. This method allows the model to adapt gradually to new information while retaining a long-term view of the data.

Therefore in our case, in Figure 10. we observe that in the short-term, combined models (such as Top 10 by MAE/ Top 5 by MAE) benefit from the aggregation of multiple forecasts, which helps mitigate short-term volatility and noise, leading to better performance. In the medium-term, we notice the individual models starting to capture more stable patterns, and by the long-term, both individual and averaged models strike a balance.

In the rolling estimation a moving window of fixed size is used to update the model parameters. This means that only the most recent data within the window are used for parameter estimation, and older data are discarded. This allows the model to be more responsive to recent changes, as it focuses exclusively on the most recent observations.

This is the reason that, in the Figure 11., we observe that individual models benefit more from rolling estimation, as they can quickly adapt to recent trends and patterns without being influenced by potentially outdated information. This enhances their ability to capture current dynamics effectively.

Therefore, we conclude that in the rolling estimation, individual models outperform averaged models like Top 10 by MAE over the last 3 or 5 periods due to their ability to rapidly adapt to the most recent data. The focus on a fixed, recent data window allows individual models to quickly align with current trends and patterns, making them more effective in this context.

On the other hand, the averaged models perform really well in recursive estimation because they leverage the stability and robustness provided by a broader historical data context, which helps smooth out short-term volatility and enhances predictive accuracy over longer horizons.

Understanding these important dynamics is essential for selecting the most appropriate models for different forecasting horizons, ensuring more precise and dependable predictions.

7.1.4 Model Choice for Germany

In our analysis for Germany, we have chosen to visually examine the short-term performance of four models. Figure 16., illustrates the performance of these models against the actual observed data (represented by the black line).

Notably, the Top10 by MAE over the last three periods model using recursive estimation and the PLS(2) model using rolling estimation perform relatively well, closely aligning with the actual events. These models, shown in red and yellow respectively, capture the fluctuations and trends effectively.

Additionally, we included the Average Machine Learning model using recursive estimation (in purple) and the Lasso model using rolling estimation (in blue). These models were selected due to their high SSR. As depicted, both models track the direction of the actual data relatively well, indicating their robustness in capturing the general trends and movements. In Figure 19., the vertical axis lists the indices of the variables selected by the Lasso model and the horizontal axis represents the timeline (early 2022 to mid-2023). In the early 2022, three variables were selected initially, so the model was focusing on a limited set of predictors, probably the HD2 variable and two other variables form Survey Data. Towards the July of 2022, the model considered a broader set of predictors, eventually heading towards the beginning of 2023 where variables like Exports, Unemployment Rates, Retail Sales, Economic Sentiment Indicator (ESI), Industrial Confidence Indicator, etc., are consistently being chosen, reflecting their importance in predicting the Industrial Production.

This visual comparison underscores the efficacy of these models in short-term forecasting, with the recursive Top10-MAE-Last3 and rolling PLS(2) models showing particularly strong performance. The loadings in Figure 20., reveal that variables like Exports, Unemployment Rate and Retail Sales have significant and dynamic contributions to the model. We observe that for the Exports (HD2) from 2021-09-30 to 2022-04-30 higher export values correspond to a decrease in the predicted industrial production. Around 2022-05-30, a transition phase is noticeable where the loadings are moving around zero, indicating a period of uncertainty. From 2022-11-30 to 2023-08-30 there is a change. Now higher export values correspond to an increase in industrial production. The loadings for Unemployment Rate (HD5) show fluctuations with a notable increase in the negative impact around November 2022. This suggests that as unemployment fluctuated, it had varying impacts on industrial production, with higher unemployment generally having a negative influence. For Retail Sales (HD8) the loadings are mostly negative until mid-2022, indicating that retail sales were a drag on industrial production. However, towards the end of 2022 and at the beginning of 2023, the impact becomes positive, showing an improvement in retail sales supporting industrial production.

In Figure 17., we analyze the medium-term performance of selected models based on their low MAE, comparing them to the actual observed data represented by the black line.

Upon examining the graph, it is evident that neither the ARMA(Auto) model using recursive estimation (depicted in red) nor the PLS(1) model using rolling estimation (shown in yellow) effectively captures the trend of the actual data. Both models exhibit a smoother trajectory compared to the actual data, which displays more pronounced fluctuations and variability.

The ARMA(Auto) model demonstrates some alignment with the overall direction of the data but in general, fails to accurately capture the significant peaks and troughs observed in the actual series.

Similarly, the PLS(1) model provides a stable forecast but lacks responsiveness to the abrupt changes seen in the actual data.

Therefore, they may not be suitable for scenarios requiring high sensitivity to short-term fluctuations and abrupt changes in the data.

In Figure 18., we examine the long-term forecasting performance of selected models for Germany by comparing them against the actual data, represented by the black line. The models portrayed include the Random Forest model using recursive estimation (in red), the averaged model of all models (Avg-ALL) using recursive estimation (in yellow), and the ARMA(Auto) model using rolling estimation (in blue).

It is evident from the graph that none of the models effectively captures the pronounced variability and fluctuations present in the actual data. While the Random Forest model shows some responsiveness to the data's peaks and troughs, it still fails to accurately mirror the actual trend. Both the averaged model and the ARMA(Auto) model exhibit smoother lines that do not align well with the significant changes in the actual data. The models tend to smooth out the pronounced peaks and troughs, resulting in forecasts that miss the width of the variations. This indicates a potential limitation in the long-term forecasting capabilities of these models when dealing with data characterized by high variability and abrupt changes.

Overall, while the selected models provide a general trend, their inability to closely follow the actual data's movements suggests that they may not be fully reliable for capturing the true dynamics over the long-term, highlighting the necessity for continuous model refinement and possibly the integration of more adaptive models to improve forecast accuracy.

7.2 Italy's Performance Discussion

7.2.1 Recursive Performance Metrics

For Italy, applying the same methodologies as mentioned, we tried to evaluate several models by comparing them based on their Mean Absolute Error (MAE), Root Mean Squared Forecast Error (RMSFE) and Success Sign Ratio (SSR) .

In Table 7., regarding short term period, the lowest MAE obtained by individual models are those of ARMA(Auto) and AR(1), closely followed by ARMA(1,1). Surprisingly, linear regression with only HD10 used as a predictor scores exceptionally low on MAE and RMSFE in the short term. This suggests that the spread between long and short-term yields (HD10) is a particularly potent indicator of economic activity, effectively capturing essential dynamics that influence industrial production. The strong performance of this simple model underscores the predictive power of yield spreads, a well-known leading economic indicator, and highlights the advantage of avoiding overfitting associated with more complex models. Thus, focusing on key economic predictors like HD10 may yield robust and reliable short-term forecasts.

Regarding model averages, we observe moderate performance, with most of them providing average MAE and RMSFE contrary to individual models. A slight edge is noted for the average of penalized regression models and the average of ARMA models, which score almost similar results. In contrast, the average of the PCA regressions and the factor models, in general, PCA and PLS, underperformed. This indicates that, in short term horizon, factor models may introduce noise by emphasizing variance rather than predictive relevance.

Interestingly, no machine learning models were selected based on recursive performance metrics. This may happen because these models may struggle with the evolving nature of the data, leading to overfitting or poor generalization. Simpler models like ARMA, AR, and linear regression with HD10 tend to be more robust, effectively capturing key economic relationships without overfitting, thus providing better short-term predictive performance in a recursive framework.

In the medium-term, factor models became more prominent. Specifically, the best two performances were both from PLS(1) and PLS (2) regression and their averages, followed closely by PCA(1) to (4) models and their averages, indicating that factor models were superior to any other category. This suggests that factor models, such as PLS and PCA, are effective at capturing the underlying structure and common influences in the data, thus providing robust predictive power over medium-term horizons. Surprisingly, their dominance was interrupted, again, by a linear regression with just two variables this time: HD10 (the spread between long and short-term yields), as previously and SD23 (an economic sentiment indicator). This indicates that these two variables together capture critical economic signals that significantly influence medium-term industrial production and confirms the importance of yield spread, once again. Additionally, penalized regression models continued their moderate performances, showing their robustness in various time horizon forecasts.

Notably, no time series models made the top 20 performances. This may happen because time series models, while effective for short-term forecasting, may not adequately capture the broader economic trends and structural shifts relevant in the medium-term, which factor models and carefully selected economic indicators can better address.

In the long term, factor models excelled once again. This time, PCA(5) and PCA(1) performed best, followed by the average of all PCA models, while PLS(1) also had a strong performance, highlighting the dominance of factor models in this category. This may happen because PCA and PLS are adept at distilling complex datasets into principal components that capture the most significant underlying variations. In the long-term horizon, these components likely encapsulate the broad economic trends and structural shifts that influence industrial production. The preference for PCA models in the long term, compared to PLS models in the medium-term, might be due to PCA's ability to emphasize variance across a more extended period, which is crucial for capturing long-term trends, whereas PLS balances variance and covariation with the target variable, which is particularly beneficial for medium-term forecasting.

AR(1) and AR(2) models re-emerged with commendable performances based on MAE. When evaluated by SSR, these time series models did not perform as good as factor models. This suggests that while they achieved lower MAE, they performed poorer in predicting the up or down movement (increase or decrease) of industrial production. This can occur because AR models, while excellent at minimizing overall error, may struggle with accurately capturing directional changes, especially in a volatile or rapidly changing environment. Factor models, on the other hand, might better capture underlying trends and shifts, making them more adept at predicting the direction of changes even if their overall error is higher.

Penalized regression models continued their moderate but steady performance, reinforcing their robustness and adaptability across different time horizons. Notably, in the long term, Lasso and Ridge regression placed higher on the list, contrary to the short and medium-term periods where adaptive Ridge and adaptive Lasso were more prominent. This shift may occur because the simpler regularization approaches of Lasso and Ridge are better suited for capturing the long-term underlying relationships without overfitting, whereas adaptive methods, which adjust regularization based on data characteristics, may perform better with shorter horizons due to their ability to adapt more dynamically to immediate changes.

Average models, once again, underperformed relative to individual models. This underperformance may be due to the averaging process diluting the strengths of the best-performing individual models, leading to less precise forecasts.

Finally, linear regression models utilizing HD10 (the spread between long and short-term yields) and SD23 (an economic sentiment indicator), along with their average, continued to show strong contender status. Although they may not perform as well in the long term as the top factor models, their consistent inclusion in top-performing lists across different time horizons underscores their significant predictive power and relevance in capturing key economic indicators.

In conclusion, in the short-term, the simplest models performed best, with ARMA(1,1), AR(1), and linear regression using only HD10 (the spread between long and short-term yields) achieving the lowest MAE. This suggests that HD10 is a particularly powerful predictor for short-term industrial production, likely due to its ability to capture immediate economic conditions effectively. Complex models and machine learning approaches did not perform well, potentially due to overfitting and the limited size and dynamic nature of the short-term data.

In the medium-term, there is a noticeable shift towards factor models, with PLS regression and PCA models, along with their averages, leading in performance. PLS models were particularly effective, possibly due to their ability to balance variance and covariation with the target variable, making them well-suited for capturing medium-term economic trends. Similarly, linear regression model using two variables, HD10 and SD23 (an economic sentiment indicator), performed well, indicating these variables' strong predictive power. Penalized regression models maintained moderate but consistent performance, demonstrating their robustness, while the remarkable fact is that no time series models made the top 20, suggesting they may not capture medium-term economic dynamics as effectively as factor models and key economic indicators.

For long-term forecasting, factor models once again excelled. The long-term superiority of PCA models may be attributed to their ability to emphasize long-term variance, capturing broader economic trends. AR(1) and AR(2) models returned with strong performance based on MAE, highlighting their stability and consistent error patterns over longer horizons. However, these models had a lower Success Sign Ratio (SSR) compared to the factor models. This suggests that factor models, while having higher overall error, are better at capturing directional trends. Additionally, the linear regression model using HD10 and SD23 variables showed a bit poorer, but still good performance, indicating its relevance in long-term forecasting.

7.2.2 Rolling Window Performance Metrics

Moving on to the forecasting metrics results with rolling estimation, we observe in Table 8. a diversification among top-performing models based on Mean Absolute Error (MAE).The top performers come from a variety of model categories, demonstrating that no single approach is universally superior. Specifically, models such as Lasso, ARMA(Auto), AR(3), PLS(1), and PCA(1) were all placed among the best performances. This diversity indicates that different models capture various aspects of the underlying data effectively, reflecting the complex and multifaceted nature of industrial production dynamics.

Surprisingly, a Random Forest model also enters the fray with impressive performance, leveraging its capability to model non-linear relationships effectively. The introduction of dynamically weighted model averages, such as Top5-RMSFE-Last3-DYNW and Top10-MAE-Last3-DYNW, reflects a strategic adaptation to recent performance trends. These models dynamically adjust weights based on the recent forecasting accuracy of individual models, thereby improving the overall reliability and precision of forecasts over time. This adaptive approach ensures that models responding well to current economic conditions carry more weight, enhancing the overall predictive performance in dynamic environments.

Remarkably, simple average models like Avg6 and Avg12, which aggregate the last few observations, exhibit robust performance in short-term forecasting according to MAE. However, their low Success Sign Ratio (SSR) suggests that despite their simplicity and ability to smooth out short-term fluctuations, they perform poorer in predicting the direction of changes in industrial production. This indicates that while these models provide consistent forecasts and mitigate large deviations in forecast errors, they are less effective at capturing directional trends. Nonetheless, their reliability in uncertain economic landscapes makes them valuable forecasting tools.

In the context of medium-term forecasting, the landscape remains characterized by diversity among top-performing models, reflecting their respective strengths in capturing medium-term economic trends. Notably, average models such as Avg6 and Avg12 continue to demonstrate robust performance similar to the short term, underscoring their stability and reliability.

In contrast to its performance in the short term, Random Forest did not make the list of top-performing models in the medium-term. This may be attributed to Random Forest's inherent strength in capturing complex patterns and interactions in data, which is more beneficial in capturing short-term variations rather than medium or long term trends. The absence of Random Forest in medium-term forecasting underscores the importance of selecting models that align closely with the temporal dynamics and underlying structure of the data for each forecasting horizon.

However, a notable shift in medium-term forecasting is the dominance of ARMA models among the top performers. Specifically, ARMA(Auto), ARMA(1,1), and Avg-ARMA emerged as the top three models. AR models also featured prominently, indicating their effectiveness in capturing autocorrelation and temporal dependencies within the medium-term horizon. The strong performance of ARMA models can be attributed to their robustness in modeling both the autoregressive and moving average components of time series data, which are essential for forecasting medium-term trends in industrial production.

Although the ARMA(Auto) model shows precedence based on MAE, this lead seems to be deceptive. The SSR metric of 0.1 indicates that ARMA(Auto) produced flat forecasts with no significant movement, failing to capture the actual fluctuations in industrial production. As confirmed in Figure 22, the model's predictions remained largely static, suggesting that while it minimized overall errors, it did so at the expense of accurately predicting directional changes. This highlights a critical limitation in its effectiveness for capturing medium-term economic dynamics.

In long-term forecasting, ARMA(auto) achieves, once again, low MAE, but the previously mentioned problem still occurs, as it fails to score notable SSR. This indicates that while it minimizes overall errors, it does not effectively capture the directional movements in the data, resulting in less reliable predictions for actual fluctuations in industrial production..

Time series models, including ARMA variants such as Avg-ARMA, ARMA(1,1), and AR models like AR(1) and AR(2), continue to dominate the top performing models in long-term forecasting, based both on MAE and SSR. These models excel in capturing the temporal dependencies and autocorrelation present in economic time series data, making them particularly suited for forecasting over longer horizons.

The Avg12 model also performs well, indicating their stability and effectiveness in smoothing out short-term fluctuations and providing reliable forecasts in rolling estimation scenarios. The robust performance of Avg12 specifically underscores its resilience across different forecasting horizons, maintaining consistent predictive accuracy.

Dynamic weighted average models, also demonstrate strong performance in long-term forecasting. While Ridge and Lasso models showed steady performance across recursive and rolling estimation for industrial production forecasting, adaptive Ridge and adaptive Lasso were not selected in rolling estimation. This is likely due to their computational complexity in dynamically adjusting regularization parameters, which contrasts with the straightforward static regularization of Ridge and Lasso. The stable nature of industrial production data may also render adaptive methods unnecessary, as simpler techniques already prevent overfitting effectively. In practice, models like Ridge and Lasso offer a balance of computational efficiency and predictive accuracy suitable for rolling estimation, where robust performance over time is crucial.

In Table 9., regarding recursive estimation, ARMA(Auto) model's precedence based on MAE is depicted, such short-term as medium-term although, although this image is diluted due to its extremely low SSR value. Despite this, ARMA(1,1) seems to perform pretty good in terms of both SSR and MAE. For long-term forecasts, PCA(1) and PLS(1) models stand out with low MAE and decent SSR values. This reflects their ability to extract essential data components, reducing noise and enhancing predictive accuracy over extended periods by focusing on significant predictors and underlying trends in industrial production data.

In Table 10., which focuses on rolling estimation, the Lasso model emerges as highly effective across all forecasting horizons, showcasing remarkably low SSR values. This reaffirms its proficiency in accurately capturing trends and fluctuations in industrial production data.

Alongside Lasso, notable performances in the short term are observed from models like Avg6, Avg12, and Lasso, which exhibit low MAE. These models' success in rolling estimation can be attributed to their respective approaches: average models (Avg6, Avg12) aggregate predictions over recent historical data points within the rolling window, smoothing out short-term variability and enhancing predictive stability. Meanwhile, Lasso regularization penalizes less influential predictors, promoting model simplicity and improving forecast accuracy by focusing on significant variables.

Although many models achieved low MAE, they performed poorly based on SSR. This discrepancy indicates that these models, despite their accuracy in minimizing overall errors, struggled to correctly predict the direction of changes in industrial production. Notably, the highest SSR was obtained by the Random Forest model, which scored relatively poorly on MAE but achieved an SSR value of 0.5833. This suggests that while Random Forest might not have been the most precise in terms of absolute error, it was more effective in capturing the directional movements of the data. This can occur because Random Forest, as an ensemble learning method, can be better at identifying complex, non-linear relationships in the data, which helps it anticipate the direction of changes even if it doesn't always minimize the absolute error as effectively as other models.

7.2.3 Performance Conclusion for Italy

Recursive and rolling estimation methods offer distinct advantages in forecasting industrial production data, each tailored to handle varying data dynamics and forecast horizons effectively. Recursive estimation updates model parameters continuously as new data becomes available, allowing models to adapt gradually while maintaining a long-term perspective. In Figure 12., we observe that simpler models, such as ARMA(auto) and LinReg-HD10, excel in the short term due to their ability to quickly adapt to and capture immediate trends and patterns in the data without the computational overhead of more complex methodologies. As the forecast horizon extends, simpler models in the form of factors like PCA(1) and PLS(2) excel because they efficiently extract and leverage essential data components, reducing noise and focusing on significant predictors to maintain predictive accuracy over longer periods.

Conversely, rolling estimation employs a fixed-size moving window to update parameters using only the most recent data, discarding older observations. This approach enhances responsiveness to recent changes, as seen in Figure 13, where time series models benefit significantly by quickly adapting to current trends without outdated influences. The ARMA(Auto) models consistently demonstrate their superiority across all forecasting horizons, based on MAE. In the short term, models that utilize the average of last observations also exhibit strong performance, leveraging recent data points to capture immediate trends and patterns effectively. However, as the forecasting horizon extends to medium and long terms, the prominence of time series models becomes increasingly evident. These models excel by leveraging historical data to capture and forecast complex relationships and trends over extended periods, highlighting their robustness and reliability in capturing the evolving dynamics of industrial production data.

7.2.4 Model Choice for Italy

Trying to take a better grasp of our model's performance, for Italy, we have chosen to visually examine the short-term performance of four models.

Figure 21., illustrates the performance of these models against the actual observed data (represented by the black line).

Notably, no model seem to perform well, since they fail to capture the volatility range. Though, the trend seems to be captured, a bit better by the ARMA(Auto) model. This model, shown in red, captures the fluctuations and trends effectively, but not their intensity and volatility, adequately. Additionally, we included the Linear regression with HD10 variable (yield spreads) using recursive estimation (in green) and the Lasso model using rolling estimation (in purple). These models were selected due to their low MAE. As depicted, while Lasso makes larger forecast errors, most of the times it captures the trend correctly, which explains the high values both on MAE and SSR. For example, for the period from June of 2021 to approximately June of 2022, the model successfully predicts major peaks and troughs observed in the actual data.

From Figure 24., it is evident that the Lasso model utilized most of the variables during the period from June of 2021 to approximately June of 2022. This extensive use of variables indicates the lingering effects of COVID-19, which blurred the variance explained by each predictor and introduced significant uncertainty. Following this period, the importance of each variable became clearer, as fewer variables were selected by the Lasso model. Notably, the economic sentiment indicator (SD23) was chosen almost every period, highlighting its consistent relevance in the model. This transition underscores the evolving clarity in variable significance as the impact of the pandemic diminished.

From Figure 25., the PLS(2) loadings used for forecasting the period from September of 2020 to March of 2023 demonstrate the consistency of these loadings across different periods. This consistency indicates robustness, as the loadings for almost every variable remain approximately the same over time. Notably, the most significant variable, SD23 (economic sentiment indicator), produced negative loadings across all time periods. This contrasts with our expectations regarding the correlation between industrial production and the economic sentiment indicator. This unexpected result may be due to the complex economic environment, where negative economic sentiment could be influenced by factors such as the long-standing effects of COVID-19, supply chain disruptions, shifts in consumer behavior, and other residual impacts. These factors could have led to a complex and inverse correlation between industrial production and economic sentiment for Italy over this period.

We also observe similar negative loadings of the Consumer Confidence Indicator from September 2020 to March 2023, reaching as low as -0.3. Italy, like many countries, experienced significant disruptions, like job losses and reduced income levels, due to the COVID-19 pandemic, leading probably to a more cautious approach on spending habits. The economic instability of the period led consumers to postpone major purchases and adopt a more conservative financial stance, while the delayed results of government stimulus packages or even the psychological impact of the prolonged crisis could have heightened consumer anxiety and reduced their optimism about the future.

In Figure 22., we analyze the medium term performance of selected models based on their low MAE, comparing them to the actual observed data represented by the black line. Upon examining the graph, it is evident that the PCA(1) model using rolling estimation (depicted in yellow) does not effectively capture the trend of the actual data. The model exhibits a smoother trajectory compared to the actual data, which displays more pronounced fluctuations and variability. The other models ARMA(Auto), PLS(2) and LinReg-HD10-SD23 demonstrate some alignment with the overall direction of the data, especially PLS(2) using recursive estimation, but in general, they fail to accurately capture the significant peaks and troughs observed in the actual series.

The low SSR value of ARMA(Auto) can be confirmed by Figure 22. and 23., where we observe almost plain forecasts, with movements occurring only during high fluctuations in actual data. This indicates that ARMA(Auto) captures significant trends, but fails to predict smaller, more frequent changes. Consequently, while it effectively identifies major trends, its predictions remain relatively static, reflecting its limitation in capturing finer, short-term dynamics.

Therefore, they may not be suitable for scenarios requiring high sensitivity to short-term fluctuations and abrupt changes in the data.

In Figure 23., we examine the long-term forecasting performance of selected models for Italy by comparing them against the actual data, represented by the black line. The models portrayed include the PCA(1) using recursive estimation (in red), the averaged model of all 12 last observation (Avg12) using rolling estimation (in yellow), and the ARMA(Auto) model using rolling estimation (in purple) and the PCA(5) model using recursive estimation (in green). It is evident from the graph that none of the models effectively captures the pronounced variability and fluctuations present in the actual data. While the PCA models show some responsiveness to the data's peaks and troughs, it still fails to accurately mirror the actual trend. Almost all models provide forecasts, with values proximately zero, indicating strong uncertainty in long time period, possibly due to the lingering effects of COVID-19. Overall, while the selected models provide a general trend, their inability to closely follow the actual data's movements suggests that they may not be fully reliable for capturing the true dynamics over the long term, highlighting the necessity for continuous model refinement and possibly the integration of more adaptive models to improve forecast accuracy.

7.3 Ireland's Performance Discussion

7.3.1 Recursive Performance Metrics

As we notice in Table 11., in the short term period, the RandomForest model demonstrated the best performance with the lowest MAE (9.51) and RMSFE (12.87), and it showed high SSR (0.625), showing a good directional accuracy, compared to the other models. The Top10-Last-DYNW and Top10-Last models also performed relatively well, with slightly higher MAEs and RMSFEs and lower SSRs. The Avg-ML model showed good performance as well, with a balanced trade-off between MAE (9.61) and RMSFE (12.77) while maintaining very high SSR (0.67), indicating a superior directional accuracy. The Top10 by MAE and RMSFE over the last 3 periods models, show the highest directional accuracy (0.7) among the 20 models we are examining.

The presence of models like RandomForest, Top10-Last-DYNW, Avg-ML and Top10-MAE-Last3 in the top 5 of the table, with the least MAE and RMSFE, can be attributed to their ability to capture complex relationships and patterns in the data. RandomForest, for instance, leverages ensemble learning and bagging techniques, which enhance its robustness and accuracy. The Top10-Last-DYNW model benefits from dynamic weighting, adapting to recent performance trends and improving forecasting accuracy. Avg-ML, by averaging multiple machine learning models, balances the strengths and weaknesses of individual models, resulting in better overall performance.

The large values of MAE and RMSFE in Ireland's models, compared to the values observed for Germany and Italy, may be due to several factors. One possible reason is the inherent volatility and complexity of the data being analyzed for Ireland, which might exhibit more pronounced fluctuations and non-linear patterns. Additionally, differences in economic structure, external shocks we mentioned earlier could contribute to the higher error metrics. This highlights the importance of selecting and tuning models according to the specific characteristics of the country.

Based on the relative measures in Table 13., we observe that many models outperform the AR(1) benchmark model. For instance, the RandomForest model outperforms AR(1) with a relative MAE of 0.919 and a relative RMSFE of 0.947, indicating superior predictive accuracy and robustness. The Top10-Last-DYNW and Avg-ML models also show strong performance, with relative MAEs of 0.920 and 0.928, and relative RMSFEs of 0.973 and 0.940, respectively.

The fact that many models (simple regression models, penalised regression models, factor models, average models) outperform AR(1) suggests that these advanced models are more effective at capturing the underlying patterns and complexities in the data, leading to better forecasting accuracy. This underscores the potential advantages of using machine learning techniques and model averaging strategies over traditional econometric models for short-term forecasting.

Moving forward to the medium-term analysis, the Avg12 model excelled, achieving the lowest MAE (9.23) and a relatively low RMSFE (12.12), paired with a good SSR (0.64), indicating a balance of accuracy and consistency. The Avg-ML and RandomForest models followed, with comparable MAEs and RMSFEs, but RandomForest showed the same SSR (0.64). The LinReg-HD10-SD23 model also performed well with a low RMSFE (11.97) and a good SSR (0.68), suggesting robust performance over the medium-term period, along with the Top5-MAE-Last3 model. Interestingly, NeuralNetwork model is among the top 6 models based on these metrics, with the highest directional accuracy of 0.67.

From the comperative metrics in Table 13., we notice that the models that outperform the AR(1) are the Avg6, Avg12, RandomForest, LinReg-HD10, and Top5-MAE-Last3, among others.

The superior performance of these models compared to AR(1) can be attributed to their ability to capture more complex relationships within the data as AR(1) is a relatively simple model and it may not adequately capture such complexities, leading to higher error metrics.

Finally, for the long-term, the Avg6 model provided the best performance, exhibiting a low MAE (9.31) and RMSFE (12.01), along with a high SSR (0.68). The Top10-Last-DYNW and Top5-MAE-Last3-DYNW models also showed strong performance, with low MAEs and RMSFEs, and SSRs indicating consistent forecasting. We observe that there are mostly averaging models that performed well in the long-term period, which makes sense since they combine the strengths of multiple models, which helps mitigate the biases inherent in individual models. Also the dynamic adaptation of models like Top10-Last-DYNW and Top5-MAE-Last3-DYNW allows them to adapt to recent performance trends, which is crucial in long-term forecasting where data patterns can evolve over time.

In the comparative performance, according to Table 13., we observe that models like Avg6 and Avg12 outperform the AR(1) benchmark and Top5-Last-DYNW and Top10-Last-DYNW utilize dynamic weighting to adapt to recent trends, which significantly enhances their forecasting accuracy, in comparison with the benchmark. The low relative MAE and RMSFE, coupled with significant DMpvals, indicate their robustness.

7.3.2 Rolling Window Performance Metrics

From Table 12., using rolling estimation, we unravel a different story. The values of MAE and RMSFE are still large, which indicates volatility in the data, but now, in the sort-term the principal component analysis models (PCA(3), PCA(1), PCA(2)) perform well, indicating their effectiveness in capturing the underlying structure of the data by reducing dimensionality and noise. Also, Ridge and Adaptive Ridge models show relatively good performance, suggesting that regularization techniques help in improving forecasting accuracy by preventing overfitting.

Compared to the benchmark, in Table 14., we observe that these models have smaller relative MAE and RMSFE values than AR(1). These models outperform AR(1) indicating better accuracy in short-term forecasts.

In the medium-term we observe that models like LinReg-HD10 and Avg-LinReg (Average of Linear Regression models) perform exceptionally well, highlighting the effectiveness of simple linear models in medium-term forecasting. The models AR(2) and AR(3) also perform well, suggesting that autoregressive models can capture temporal dependencies effectively. We also notice that these models outperform AR(1), in the comparative performance metrics.

For the long-term, dynamic models like Top5-Last and Top10-Last, Top5-RMSFE-Last3, Top10-MAE-Last3, Top10-MAE-Last3-DYNW consistently perform well, indicating that adaptive weighting based on recent performance enhances long-term forecasting accuracy. Moreover, Avg12 and Avg6 also perform relatively well, supporting the effectiveness of model averaging in capturing diverse data patterns over longer horizons.

In Table 14., the comparative performance metrics show that there are many models, especially dynamic and averaging models, that significantly outperform AR(1), demonstrating the importance of using simpler (Linear Regression) and advanced techniques (Averaging, Penalised Regressions, Factor Methods, Machine Learning, Dynamic Ranking) for accurate predictions in the long-term forecasting.

7.3.3 Performance Conclusion for Ireland

In the short-term with recursive estimation, the RandomForest model emerged as the top performer with the lowest MAE (9.51) and RMSFE (12.87), coupled with a high SSR (0.625), indicating superior directional accuracy, as we notice in Figure 14. Models such as Top10-Last-DYNW, Top10-Last, and Avg-ML also performed well, showcasing their ability to capture complex relationships and patterns in the data. The high SSR values of these models underscore their robustness in forecasting the correct direction of changes.

The high error metrics for Ireland, compared to Germany and Italy, highlight the unique volatility of the Irish economic data, necessitating tailored modeling approaches.

For the medium-term forecasting with recursive estimation, in Figure 14. we observe that the Avg12 model excelled with the lowest MAE (9.23) and a relatively low RMSFE (12.12), paired with a good SSR (0.64). The Avg-ML, RandomForest, and LinReg-HD10 models also demonstrated robust performance, effectively balancing error minimization and directional accuracy. The superiority of these models over the AR(1) benchmark, as shown by their relative metrics, underscores their ability to capture more intricate data relationships, making them well-suited for medium-term forecasting in Ireland.

In the long-term with recursive estimation, Figure 14. shows that the Avg6 model provided the best performance with a low MAE (9.31), RMSFE (12.01), and a high SSR (0.68). Dynamic models such as Top10-Last-DYNW and Top5-MAE-Last3-DYNW also demonstrated strong performance, leveraging adaptive weighting to enhance long-term forecasting accuracy. The prevalence of averaging models among the top performers underscores their effectiveness in mitigating biases and capturing diverse data patterns, essential for accurate long-term forecasting in Ireland.

In Figure 15., in the short-term with rolling estimation, we highlight that the PCA models (PCA(3), PCA(1), PCA(2)) demonstrated strong performance, effectively capturing the underlying structure of the data through dimensionality reduction. Ridge and Adaptive Ridge models also showed good performance, indicating the benefits of regularization techniques in improving forecasting accuracy by mitigating overfitting. Compared to the AR(1) benchmark, these models exhibited smaller relative MAE and RMSFE values, highlighting their superior accuracy in short-term forecasts for Ireland.

For the medium-term, models such as LinReg-HD10, Avg-LinReg and Adaptive Ridge performed exceptionally well in the metrics, highlighting the efficacy of simple linear models in medium-term forecasting. These models' superior performance compared to AR(1), as evidenced by their relative metrics, underscores the importance of leveraging penalised and linear models for accurate medium-term forecasts.

Finally, for long-term forecasting with rolling estimation, in Figure 15. we observe that dynamic models like Top5-Last and Top10-Last performed well, highlighting the benefits of adaptive weighting based on recent performance trends. Models such as Avg12 and Avg6 also performed robustly, supporting the effectiveness of model averaging in capturing complex, evolving data patterns over longer horizons. Compared to the AR(1) benchmark, these models showed significantly better performance, demonstrating the importance of employing advanced techniques like averaging, penalized regressions, and dynamic ranking for long-term forecasting accuracy in Ireland.

For short-term with recursive estimation, we conclude that dynamic weighting and model combination strategies help the models to adapt to recent trends by emphasizing the most accurate models, which enhances their ability to capture complex, non-linear relationships in the data. On the other hand, the rolling estimation approach continuously updates model parameters with new data, making these models highly responsive to recent developments and trends, which is crucial for accurate short-term forecasting.

Under recursive estimation for the medium-term, models like RandomForest, Avg-ML, LinReg-HD10, Top5-MAE-Last3, excel in capturing medium-term data patterns due to their robust methodologies, such as ensemble learning and dynamic weighting. Models such as AdRidge, Avg12, LinReg-HD10-SD23, Avg-LinReg, benefit from the rolling window approach.

In both long-term recursive and rolling forecasting for Ireland, dynamic models like Top10-Last-DYNW, Top10-Last, Top10-Last, etc., consistently perform well. These models excel due to their adaptive weighting mechanisms, which adjust to recent performance trends.

7.3.4 Model Choice for Ireland

In our analysis for Ireland, we have chosen to compare the actual values with forecasts from four models.

In Figure 26., with the recursive estimation we chose the RandomForest model (orange line) that captures complex relationships in the data through ensemble learning techniques, and the Top10-MAE-Last3 (green line) that uses the top 10 models based on mean absolute error over the last three periods, with dynamic weighting to adapt to recent trends and shocks. Under rolling window estimation we chose PCA(3) and PCA(1).

We observe that the recursive models generally show better alignment with the actual data trends compared to rolling models, capturing sudden changes and volatility in the data more effectively.

As mentioned earlier, the fluctuations in the actual data in early 2022 could be due to the residual effects of the COVID-19 pandemic, including recovery efforts and initial vaccine rollouts. From mid 2022 to early 2023, the volatility we observe might be driven by global supply chain issues, ongoing geopolitical tensions (Russia-Ukraine conflict), inflationary pressures, energy price shocks and adjustments post-Brexit. The sharp rise and subsequent drop in mid-2023 can be tied to factors such as economic recovery momentum, fiscal policy interventions, and possibly sector-specific developments in technology and pharmaceuticals, which are significant for Ireland.

RandomForest, due to its ability to capture complex non-linear relationships shows good performance in capturing the direction of changes but tends to have larger deviations at certain points, such as in late 2022. The Top10-MAE-Last3 balances between capturing trend directions and maintaining a smoother forecast, because of its dynamic weighting approach's effectiveness.

Both PCA(3) and PCA(1) rolling models exhibit smoother trends and lower volatility in their forecasts. They are better at capturing general patterns but we notice that they miss some sharp fluctuations present in the actual data.

The PCA(3) loadings in Figure 29., illustrate how various economic indicators influenced Ireland's economic conditions during different periods. Two very important variables are the Economic Sentiment Indicator and the Industrial Confidence Indicator.

The Economic Sentiment Indicator (ESI) - SD23 is a significant variable that captures overall economic confidence. In the loadings, the fluctuation around zero indicates varying degrees of sentiment influence over time, highlighting its role during periods of economic stability and volatility. The initial positive sentiment in late 2020 to early 2021 reflects optimism as the economy began to recover from the initial pandemic shock earlier in 2020. However, subsequent waves and uncertainties led to fluctuating and often negative sentiment. In late 2021 to early 2022 there is a noticeable decline in economic sentiment. The new COVID-19 variants, such as Delta and Omicron, and the resulting disruptions may have contributed to increased pessimism. Moreover, concerns over potential lockdowns, supply chain issues, and inflationary pressures also negatively impacted sentiment. The gradual improvement around late 2022 in sentiment towards the end of the period highlights the effectiveness of government interventions and the adaptability of the economy.

The resilience and recovery of sectors such as technology and pharmaceuticals played a crucial role in stabilizing and eventually improving economic sentiment.

Upon examining the loadings of the Industrial Confidence Indicator - SD24, that reflects the overall sentiment within manufacturing and production industries, we observe that the initial fluctuations in late 2020 - early 2021 probably reflect the industrial sector's struggle to stabilize post-pandemic, with ongoing disruptions causing swings in confidence.

The negative sentiment mid-2021 indicates the impact of supply chain issues and increased costs on industrial production, a common challenge globally during this period. While the sharp recovery spike in early 2022 could be due to easing restrictions, improved supply chain conditions, or positive economic news.

The mid-2022 stabilization of the indicator and the slight decline towards the end demonstrate a more resilient sector that has adapted to the new normal, finding ways to operate effectively despite ongoing challenges, while the slight decline in the late 2022 could be due to concerns over inflation, geopolitical tensions, or other economic uncertainties.

Moving forward to Figure 27., that displays the performance of selected medium-term forecasting models for Ireland, we compare the actual values with predictions from the recursive Avg12 model and the rolling Avg-LinReg model.

We observe that the actual data line shows considerable fluctuations, with sharp increases and decreases, indicating periods of economic shocks and recoveries, that the chosen models fail to capture effectively. These two models had the smallest MAE among the models, but they are still not suitable for capturing the medium-term dynamics of Ireland's economy during periods of substantial volatility. Their almost horizontal lines indicate a failure to respond to the significant economic fluctuations. This also serves as a reminder that a lower MAE value does not always equate to a model's practical usefulness in capturing economic dynamics.

Finally, in Figure 28., we display the long-term performance of four chosen models for recursive and rolling estimation for the country.

The actual data (black line) shows significant fluctuations, indicating volatility in Ireland's long-term economic trends.

The Rec-Avg6 model (orange line) and Rec-Top10-Last-DYNW (purple line) show similar trends to the rolling models Rol-Top5-Last (green line) and Rol-Top10-Last (red line).

When there are big declines or spikes in the economy (e.g. in the early 2021, mid-2022, early 2023 or late 2023, Avg6 shows more gradual changes due to the averaging effect. For instance, during the sharp drop around late 2021, the line shows a decline but at a much more tempered rate compared to the actual data, and does not react as drastically to the steep drop in actual data just before 2024, indicating its smoothing characteristic which averages out the impact over six periods.

The Top10-Last-DYNW adapts to the data more closely than the simple average model so it can better capture sudden shifts in the economy but still smooths out some of the noise.

The Top5-Last tends to be less volatile than the actual data but more responsive than the simple average model. It adapts quickly to changes in the economy, as we observe, showing more immediate responses to big declines or spikes compared to the recursive models.

Similar to the Top5-Last model but using the top 10 indicators, the Top10-Last It tends to capture the trends of the actual data closely while still providing a smoother prediction line. It responds quickly to economic changes but incorporates more information from the additional indicators.

We notice that each model exhibits a trade-off between smoothness and responsiveness. The rolling models generally adapt more quickly to changes in the Irish economy in long-term, while the recursive models provide smoother, less volatile predictions. The choice between these models depends on the specific requirements for prediction accuracy versus the need to smooth out fluctuations.

7.4 Summary of Findings

This section provides a comprehensive discussion on the overall performance of models across short-term, medium-term, and long-term forecast horizons for Germany, Italy, and Ireland. The aim is to identify patterns, strengths, and weaknesses of different models in capturing economic trends over the three periods we have examined and to understand why certain models perform better or worse in different countries.

7.4.1 Short-Term Forecasting Summary Findings

For Germany, the Top10-MAE-Last3 and Avg-ML models excel in capturing short-term economic volatility of supply chain disruptions and industrial production changes, due to their ability to handle non-linear relationships and adapt to recent data trends, leading to more stable predictions. PLS(2) is effective in capturing short-term economic trends by effectively summarizing the information from multiple predictors, while Lasso focuses on the most relevant predictors, enhancing its ability to capture short-term economic fluctuations.

ARMA(Auto) and Linear Regression with HD10 are not well-suited for Italy's short-term forecasting as they fail to effectively capture immediate economic conditions, such as short-term fluctuations in industrial production, consumer sentiment, and business activity. At the beginning of 2022 up until the middle of 2022, Lasso, focusing on the most relevant predictors, manages to follow the fluctuations of the economy more closely. However, in the following years, Lasso and Avg6 make a decent effort to follow the trend of the actual fluctuations in the short-term economy, whereas ARMA(Auto) and Linear Regression with HD10 follow a horizontal line trend, missing the shocks.

Ireland's short-term economic conditions are effectively captured by models like Random Forest and dynamic models such as Top10-Last-DYNW due to their capacity to adapt to sudden economic shifts, such as the impacts of the global financial crises, Brexit-related uncertainties, geopolitical tensions, inflationary pressures, energy price shocks and the economic fluctuations during the COVID-19 pandemic. However, PCA models fail to capture these trends effectively, as they produce more horizontal and smooth predictions that do not align with the actual fluctuations.

Overall, we observe that short-term forecasting often benefits from models capable of handling volatility and non-linear relationships, such as Random Forest and dynamic models like Top10-Last-DYNW. In contrast, simpler models like Lasso are effective for capturing immediate economic conditions in Germany and Italy, including short-term fluctuations in industrial production, consumer sentiment, and business activity. However, models like PCA often fall short in capturing the actual economic trends due to their tendency to produce smoother, more horizontal predictions.

7.4.2 Medium-Term Forecasting Summary Findings

For medium-term forecasting, Germany benefits from models like ARMA(Auto) and PLS(1) that balance variance and covariation with the target variable. Despite their ability to provide stable predictions, these models show limitations in capturing the actual economic trends accurately, as evidenced by the deviation from the actual data in the graph. Both of the models struggle to align closely with the actual fluctuations.

Italy's medium-term forecasts are captured by models such as PLS(2), Linear Regression with HD10 and SD23, ARMA(Auto), and PCA(1). These models also face challenges in closely following the actual economic data. While they reduce dimensionality and emphasize significant variations, the deviations in the graph indicate that they do not fully capture the medium-term economic dynamics.

In Ireland, models like Avg12 and Rolling Avg-LinReg are considered the best for medium-term forecasting due to their lower MAE. However, these models appear almost as horizontal lines in the graph, indicating their failure to capture the actual economic trends accurately. Despite their lower error metrics, their predictions remain relatively static and do not reflect the observed volatility in the medium-term economic data.

Overall, while these models demonstrate theoretical strengths and lower error metrics, their practical application in accurately forecasting medium-term economic trends remains limited. The consistent deviation from actual economic data across all three countries underscores the need for more dynamic and adaptable forecasting models that can better capture the complexities of medium-term economic fluctuations.

7.4.3 Long-Term Forecasting Summary Findings

For long-term forecasting, Germany benefits from robust models like Random Forest and ARMA(Auto), which maintain consistent performance over extended periods. The Avg-ALL model, which combines the strengths of multiple models, helps mitigate individual biases and captures broader economic trends. Among these, the Random Forest model aligns more closely with the actual data trends compared to the Avg-ALL and ARMA(Auto) models, demonstrating its superior ability to track the long-term fluctuations in the economy, as it likely captures the cyclical nature of industrial production, including periods of growth and contraction, fluctuations in global trade conditions, tariffs, and international relations and spending patterns, critical for predicting long-term economic trends.

We try to address Italy's long-term forecasts by PCA models, which distill complex datasets into principal components, encapsulating possibly significant long-term variations. These models aim to capture broader economic trends and structural shifts for extended forecasting horizons. However, the PCA models, hover around zero with minimal fluctuations, failing to effectively capture the true economic trend of Italy over the long term. ARMA(Auto) shows some minor fluctuations but still struggles to accurately follow the trend. The Avg12 model, while also moving mostly around zero like ARMA(Auto), displays a slight upward movement from mid-2021 to mid-2022 (economic recovery post-COVID-19 period) with smaller fluctuations. This indicates that it might capture the trend better, possibly due to the combination of the twelve models it comprises. However, overall, these models, even though they we chosen for their lowest MAE and RMSFE values, do not fully capture the long-term fluctuations of Italy's economy, that were caused by government stimulus packages after mid-2021, public debt, slow productivity growth, demographic issues like an aging population, supply chain disruptions and the slower recovery of tourism sector compared to manufacturing.

The chosen models for Ireland, identified for their lowest MAE and RMSFE, exhibit a similar trend where they fail to effectively capture the actual data's trend. Although averaging and dynamic models attempt to track some movements, they fall short of capturing the significant fluctuations observed in the actual data. The dramatic drops and spikes, indicative of sudden economic shifts in Ireland, caused by Brexit impact, Tech and Pharma sector booms and Energy price fluctuations, remain inadequately captured by these models. The substantial economic changes and volatility in Ireland's economy are not reflected accurately, highlighting the limitations of the selected models in long-term forecasting.

Therefore, despite utilizing models designed to capture broader trends and structural shifts for long-term forecasting, they often fail to effectively reflect the actual long-term trends. This failure is primarily due to the unpredictability and magnitude of the economic shocks from 2021 to 2023, which significantly impact each economy and render forecasting models inadequate in accounting for these sudden, substantial changes.

8 Conclusion

Our study provides a thorough analysis of the economic conditions in Germany, Italy and Ireland by leveraging advanced econometric and machine learning techniques. The objective was to understand the economic dynamics of these countries and forecast future trends with greater accuracy. The findings shed light on the unique economic landscapes, challenges, and strengths of each country.

For Germany, the Random Forest model was identified as the best forecasting tool in the short-term and in the long-term. This model effectively handles non-linear relationships and captures the cyclical nature of industrial production, global trade conditions, and international relations. Its robustness in dealing with complex interactions and variability makes it the most suitable for short-term and long-term forecasting Germany's economic indicators. Germany's economy remains robust, characterized by a strong industrial sector and high export levels. Despite facing global trade disruptions and geopolitical tensions, Germany maintains economic stability through effective social policies and continuous investments in infrastructure and technology. For medium-term forecasting, despite the limitations in closely following actual economic data, the ARMA(Auto) model was found to be the most effective, because it balances variance and covariation.

In the case of Italy, the Lasso focuses on the most relevant predictors, enhancing its ability to capture the short-term economic fluctuations while Partial Least Squares (PLS(2)) model proved to be the most effective in the medium-term. Partial Least Squares model reduces the dimensionality of the data and emphasizes significant variations in economic data. However, it still faces challenges in accuracy due to Italy's structural economic issues and high volatility in this period. Italy's economy shows resilience in its post-pandemic recovery but is hindered by long-standing structural issues such as high debt and bureaucratic inefficiencies. The need for reform is evident to boost productivity and innovation, essential for sustained economic growth. In the long term, the Avg6 model proves to be the most effective for Italy, because it might capture the trend better, possibly due to the combination of the twelve models it comprises, making the long-term forecasting more stable.

Ireland's economic forecasting benefitted from using both the Random Forest and Top10-Last-DYNW (Dynamic Weights) models in short-term and Top10-Last-DYNW in long-term. These models are highly adaptable to sudden economic shifts such as Brexit, changes in the tech and pharma sectors, and energy price fluctuations. The combination of these models provides a robust framework for capturing Ireland's volatile economic conditions. Ireland's rapid economic growth is driven by its business-friendly environment and the presence of multinational corporations, particularly in the technology and pharmaceutical sectors. However, this reliance on foreign investment leads to economic volatility and disparity. Efforts to diversify the economy are essential to ensure long-term stability and equitable growth. The medium-term economic forecasting for Ireland is challenging due to the unique and rapidly changing economic conditions. The Avg12 and Avg-LinReg may not be dynamic or sophisticated enough to capture these complexities, leading to predictions that do not align well with actual economic trends.

The study also employed Principal Component Analysis (PCA) and Dynamic Factor Models (DFMs) to construct Coincident Economic Indicators (CEIs) for each country. These indicators provided valuable insights into the real-time economic conditions and highlighted the economic trends over time. In Germany, the CEIs reflected the stable yet fluctuating industrial production, showcasing the country's resilience in times of global economic disruptions. Italy's CEIs indicated significant economic volatility, underscoring the structural weaknesses and the need for robust reforms. For Ireland, the CEIs captured the rapid economic shifts, emphasizing the country's dependency on multinational corporations and the volatile nature of its economic growth.

To achieve sustainable economic development, policymakers in each country should focus on several key areas. For Germany, continuing to invest in infrastructure, education, and advanced technology will help maintain industrial competitiveness and support economic resilience. Italy needs to implement comprehensive reforms to reduce bureaucratic inefficiencies, encourage innovation, and address demographic challenges. Enhancing productivity is key to overcoming Italy's structural issues. In Ireland, diversifying the economic base to reduce reliance on multinational corporations and mitigate economic volatility is crucial. Policies should aim to promote balanced growth across different sectors.

The practical application of these models highlights their strengths in theoretical settings but also exposes their limitations in accurately forecasting economic trends across different time horizons due to the unpredictable nature of economic shocks. Future research should focus on sector-specific analysis to identify targeted strategies for improvement, examining the long-term impact of global events such as pandemics and geopolitical changes on economic stability and growth, and exploring new econometric and machine learning models to improve the accuracy of economic forecasts and better understand economic trends.

This study underscores the importance of tailored policy measures and strategic investments in enhancing the economic stability and growth of Germany, Italy and Ireland. By leveraging their unique strengths and addressing their specific challenges, these countries can navigate future economic uncertainties and achieve sustainable development.

9 Figures

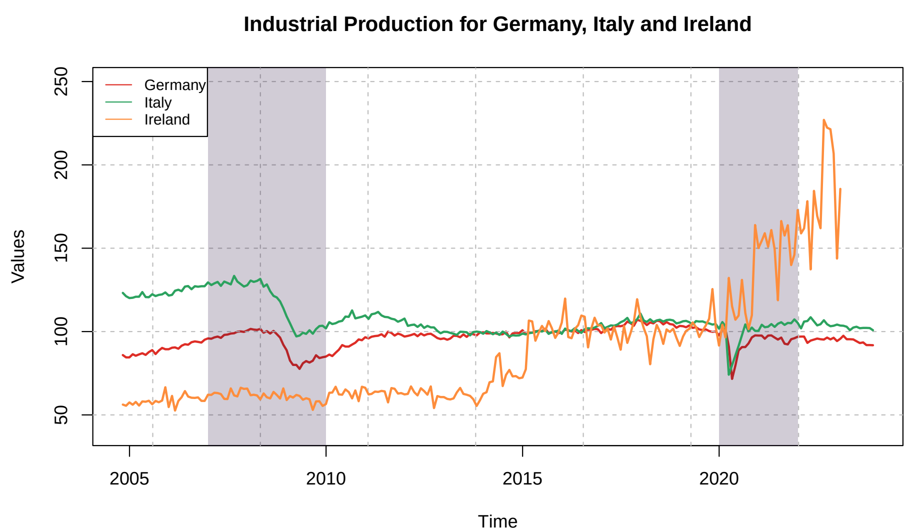

Figure 1: Industrial Production for Germany, Italy and Ireland

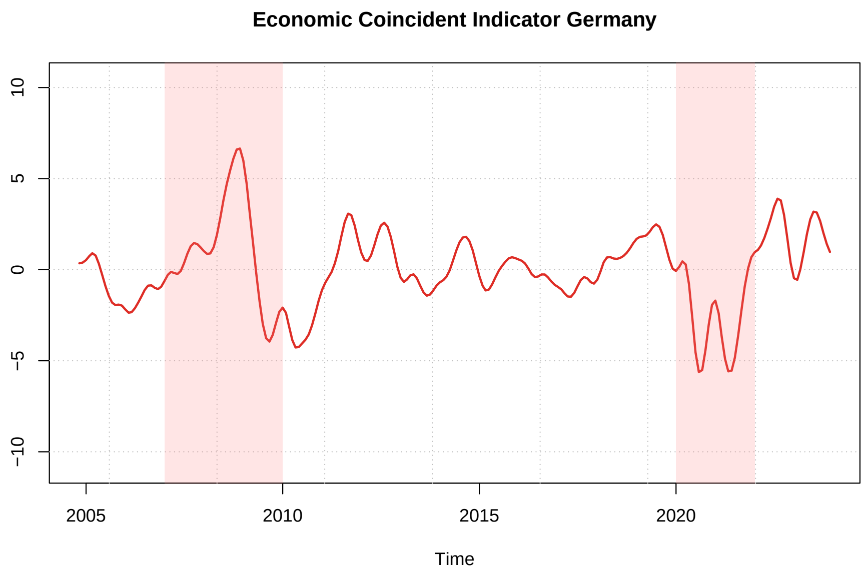

Figure 2: PCA - Smoothed First Principal Components of 45 Economic Variables of Germany

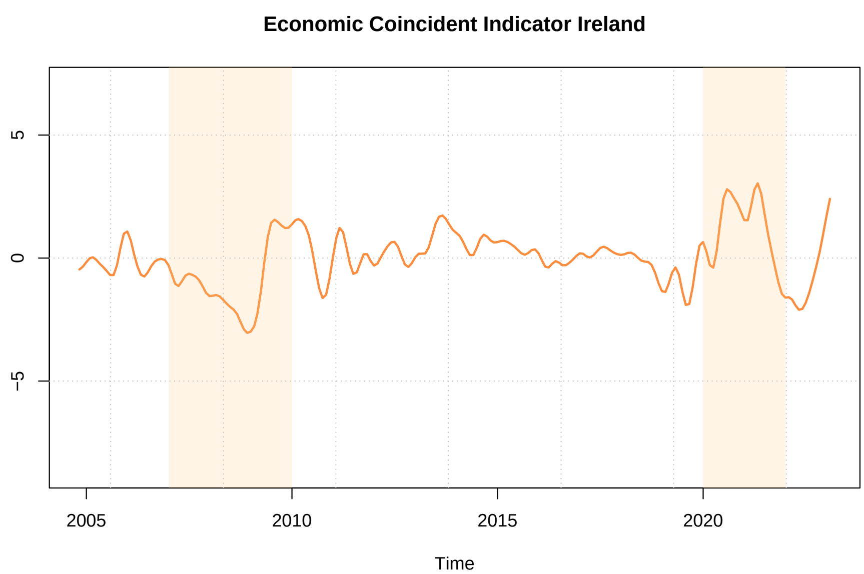

Figure 3: PCA - Smoothed First Principal Components of 45 Economic Variables of Ireland

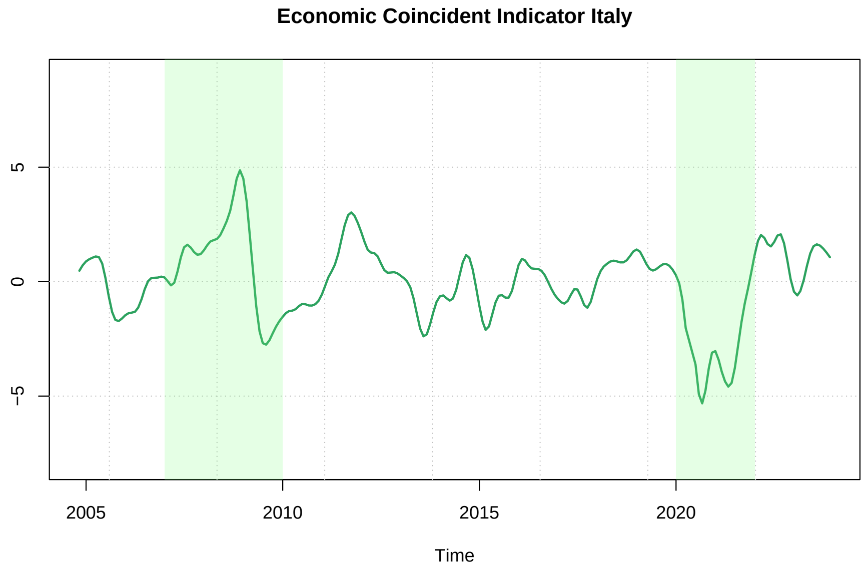

Figure 4: PCA - Smoothed First Principal Components of 45 Economic Variables of Italy

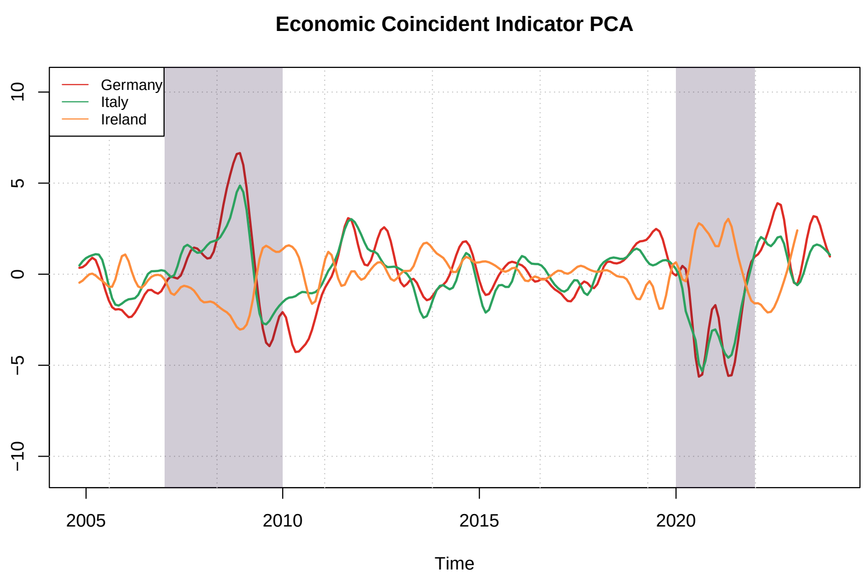

Figure 5: PCA - Smoothed First Principal Components of 45 Economic Variables of Italy, Ireland and Germany

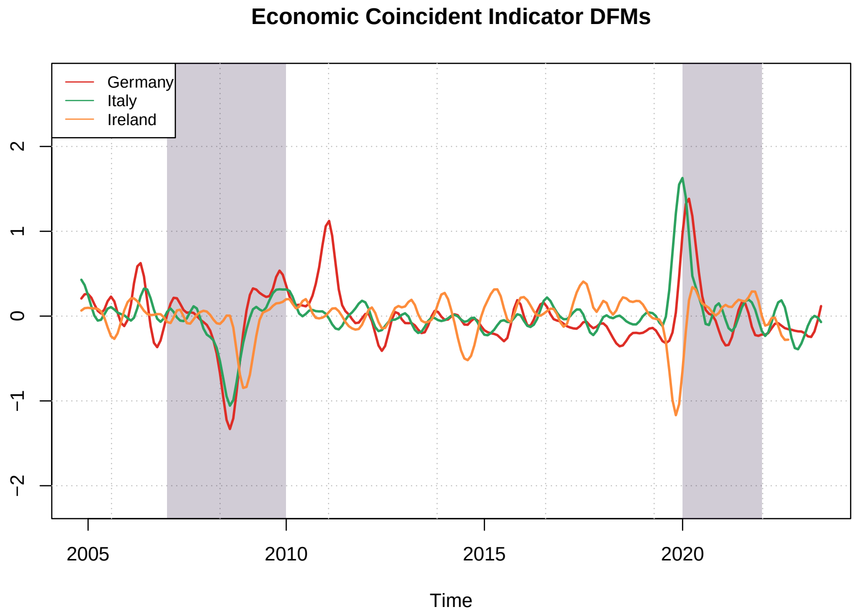

Figure 6: DFMs - Smoothed DFM Indicators of 45 Economic Variables of Italy, Ireland and Germany

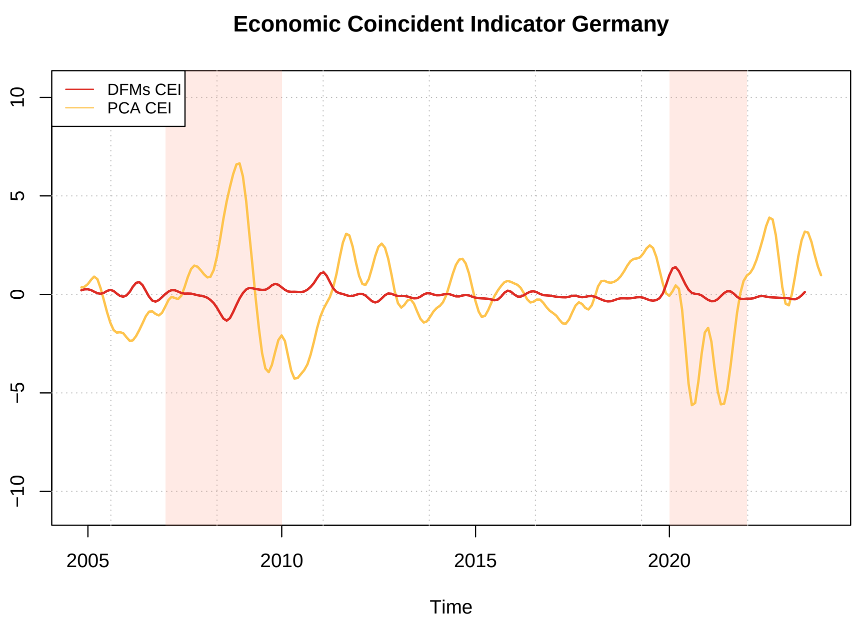

Figure 7: Smoothed DFM Indicator vs. Smoothed PCA Indicator for Germany

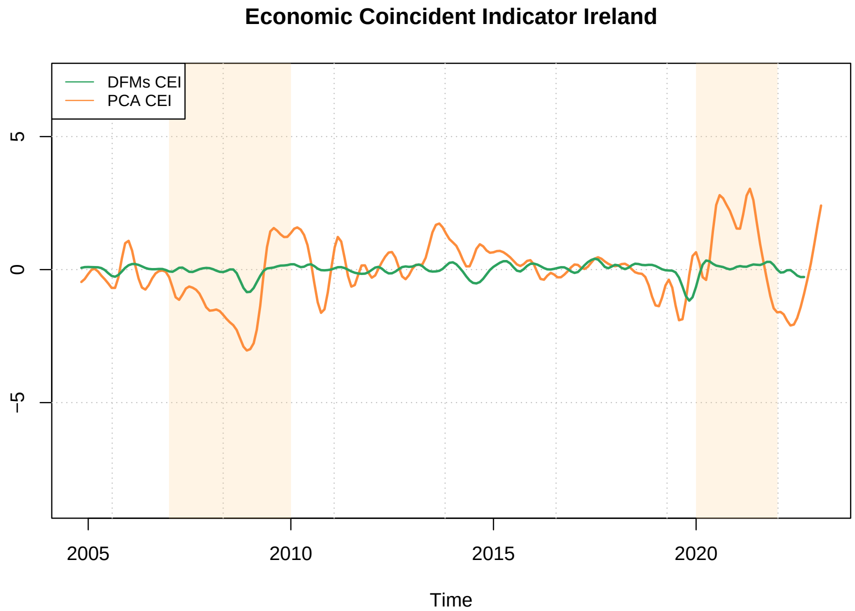

Figure 8: Smoothed DFM Indicator vs. Smoothed PCA Indicator for Ireland

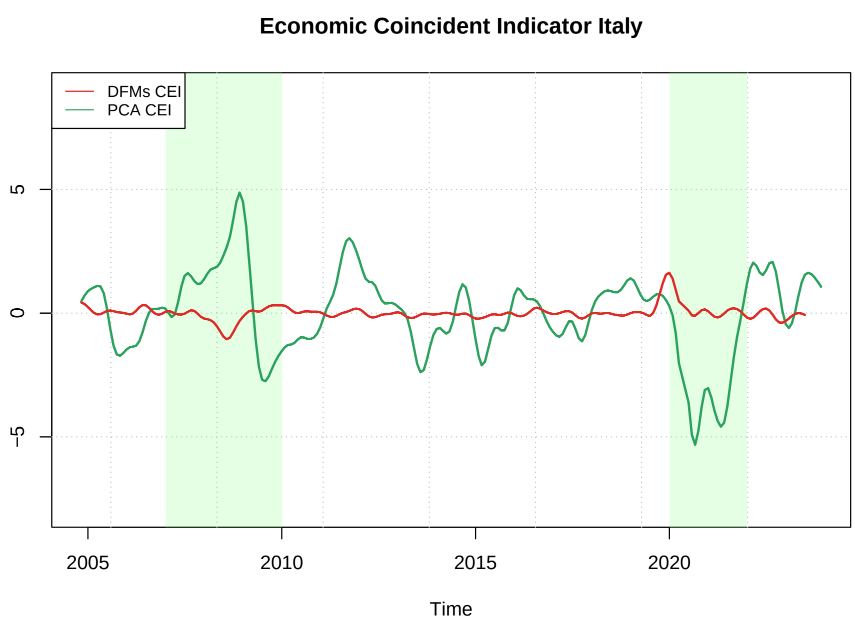

Figure 9: Smoothed DFM Indicator vs. Smoothed PCA Indicator for Italy

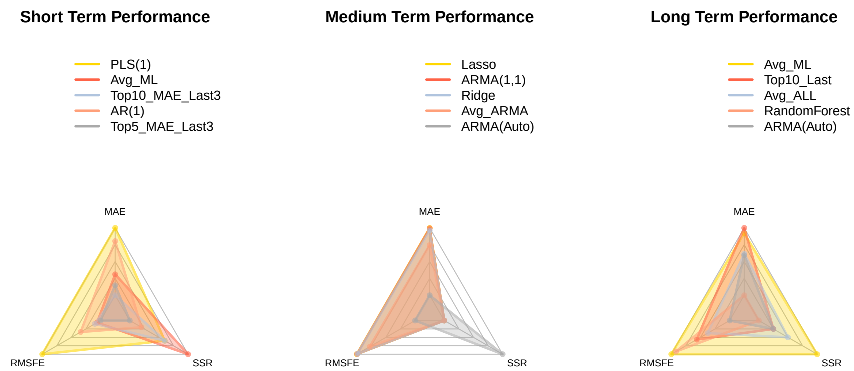

Figure 10: Recursive Performance Metrics for 5 models on each period for Germany

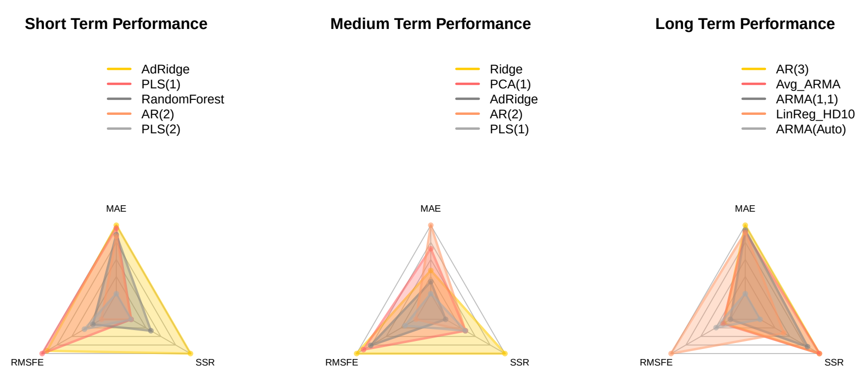

Figure 11: Rolling Performance Metrics for the first 5 models on each period for Germany

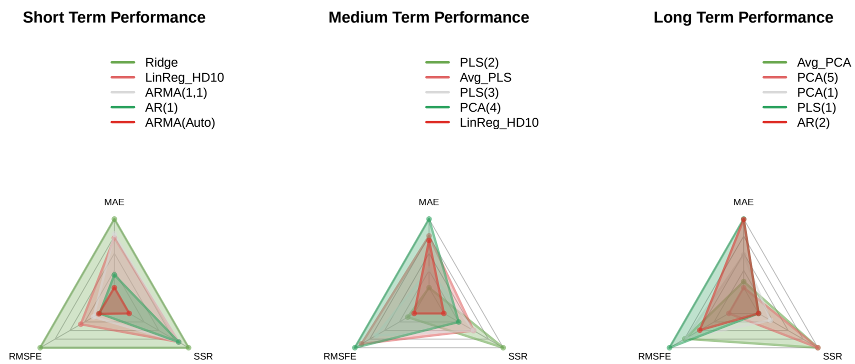

Figure 12: Recursive Performance Metrics for the first 5 models on each period for Italy

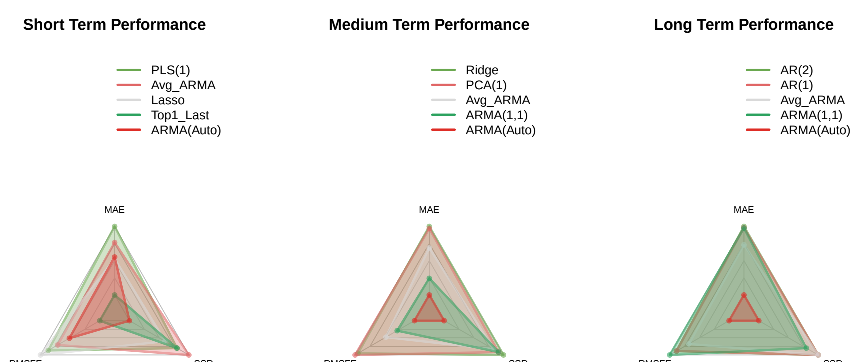

Figure 13: Rolling Performance Metrics for the first 5 models on each period for Italy

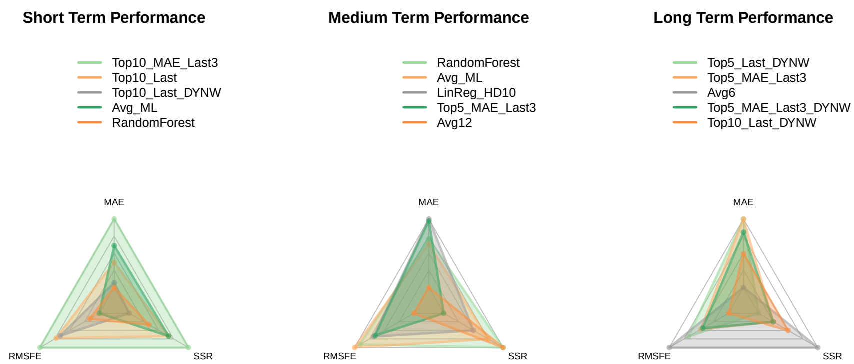

Figure 14: Recursive Performance Metrics for the first 5 models on each period for Ireland

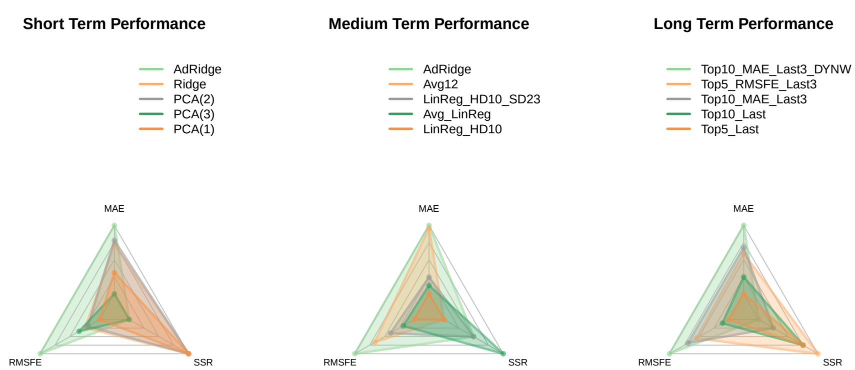

Figure 15: Rolling Performance Metrics for the first 5 models on each period for Ireland

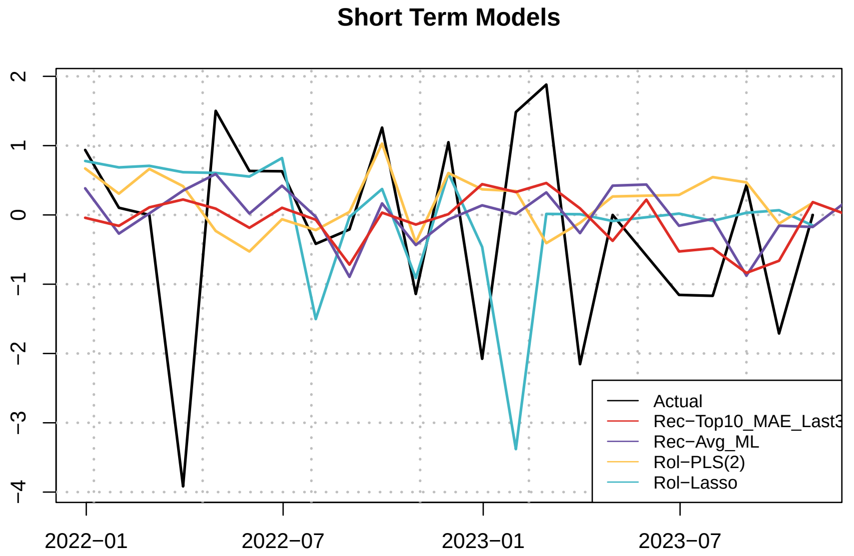

Figure 16: Chosen Short-Term Models for Recursive and Rolling Estimation for Germany

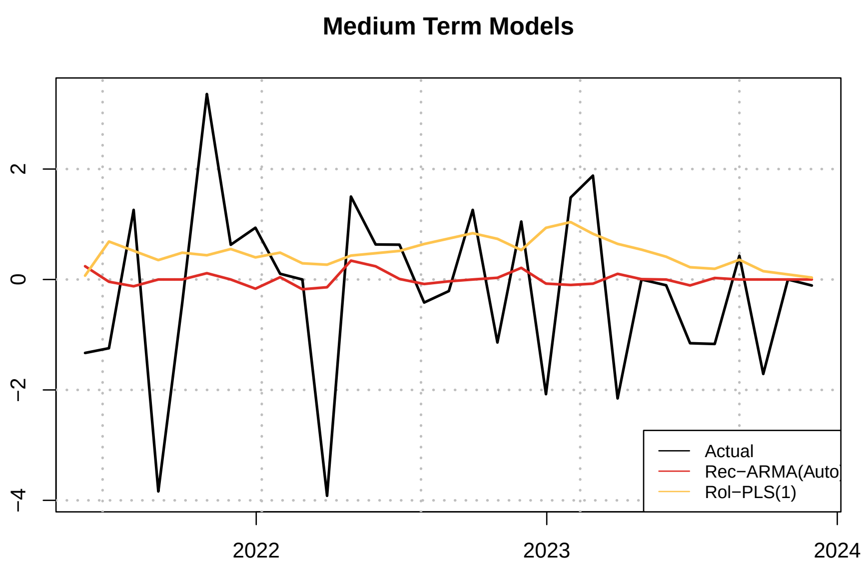

Figure 17: Chosen Medium-Term Models for Recursive and Rolling Estimation for Germany

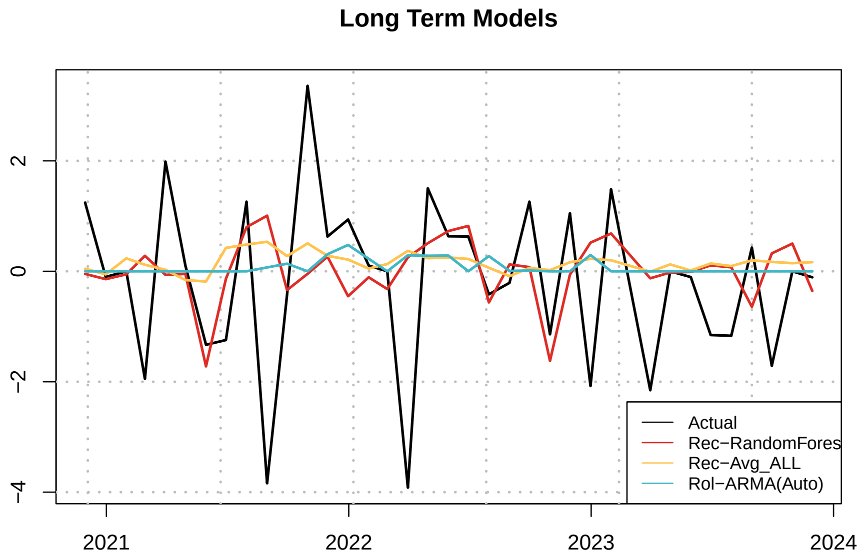

Figure 18: Chosen Long-Term Models for Recursive and Rolling Estimation for Germany

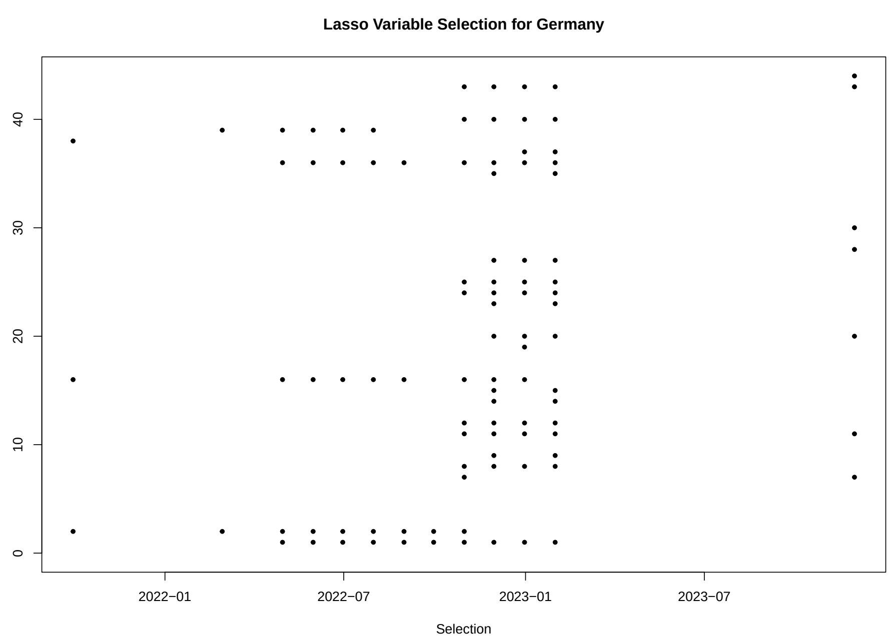

Figure 19: Lasso Variable Selection graph for Germany

Figure 20: PLS(2)'s loadings for Germany

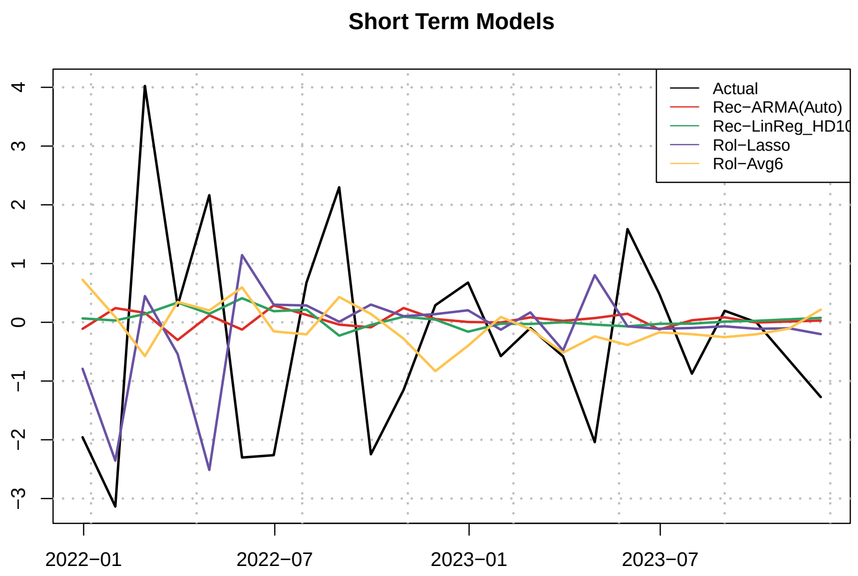

Figure 21: Chosen Short-Term Models for Recursive and Rolling Estimation for Italy

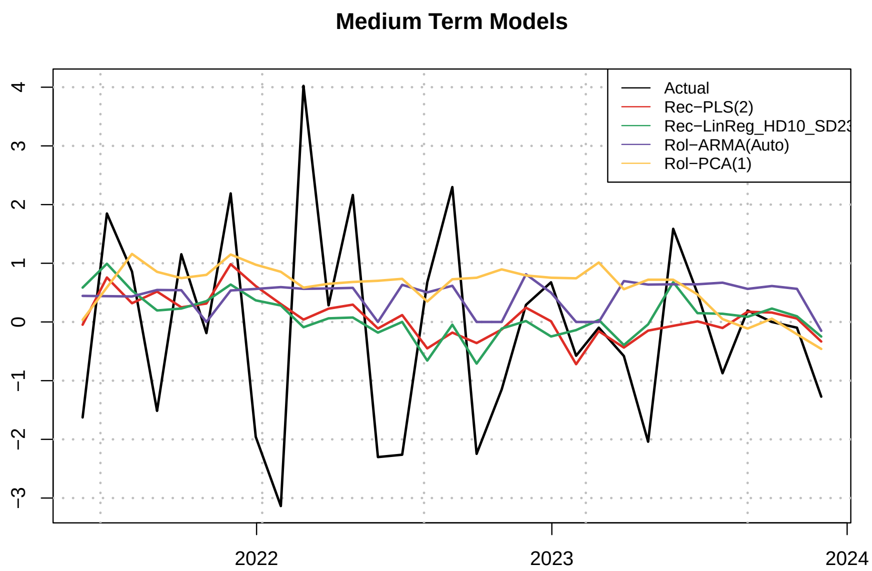

Figure 22: Chosen Medium-Term Models for Recursive and Rolling Estimation for Italy

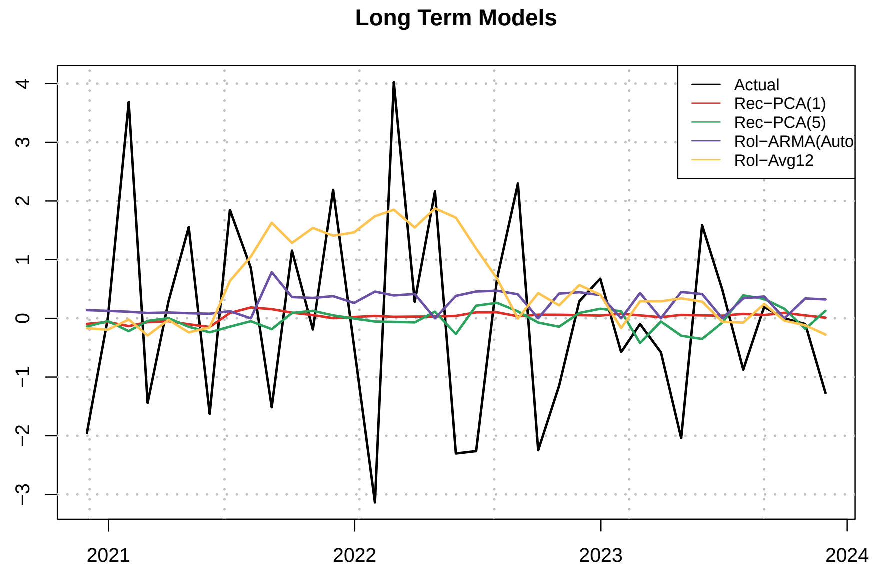

Figure 23: Chosen Long-Term Models for Recursive and Rolling Estimation for Italy

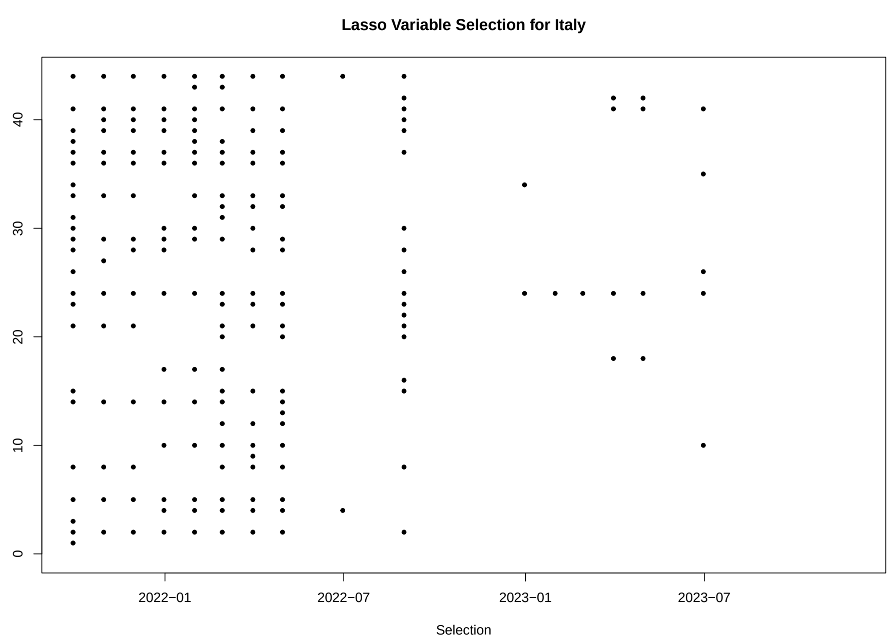

Figure 24: Lasso Variable Selection graph for Italy

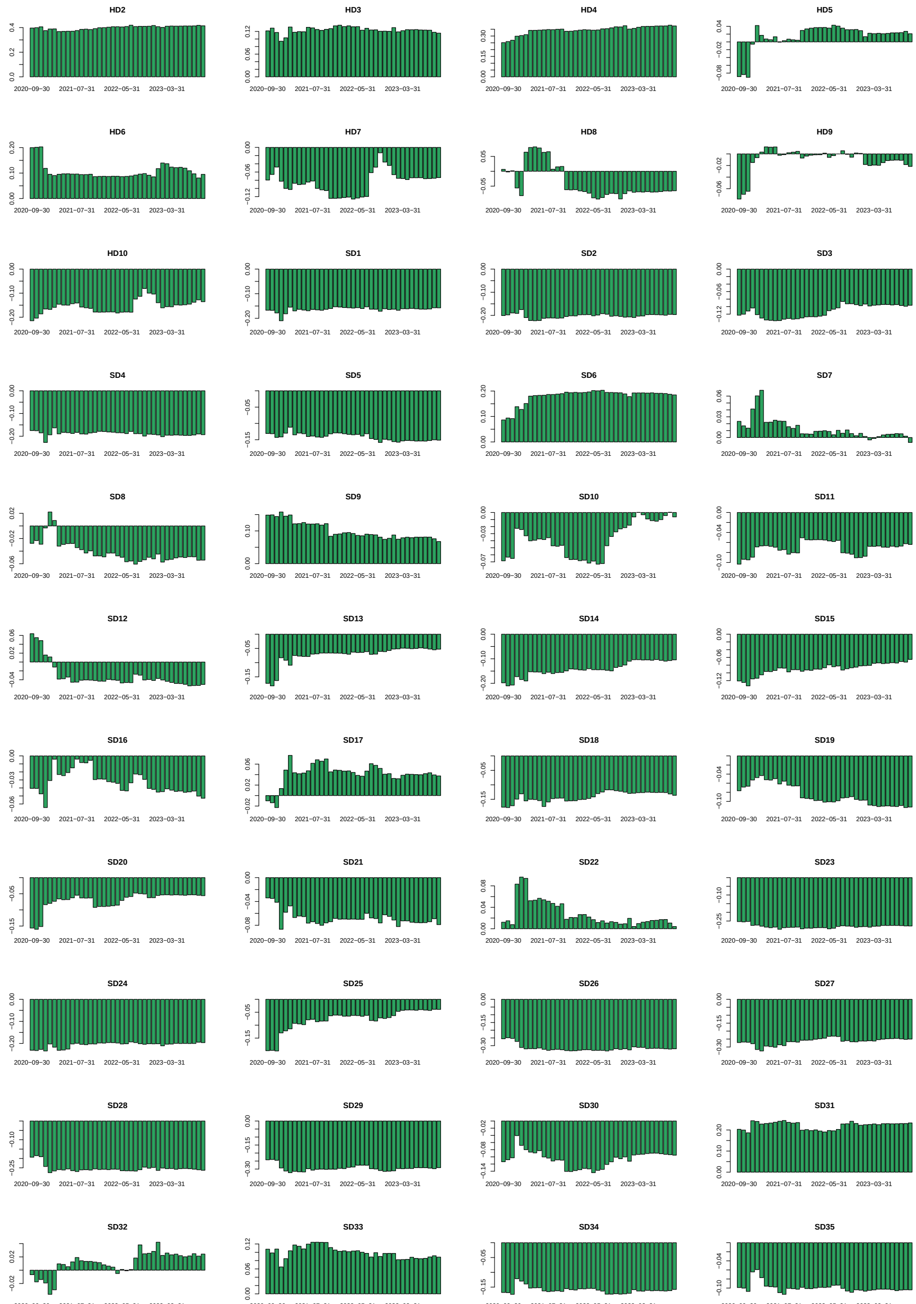

Figure 25: PLS(2)'s loadings for Italy

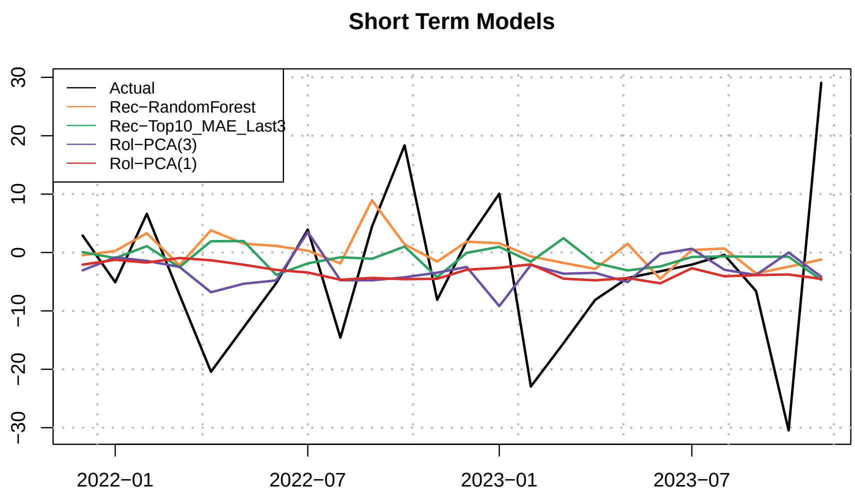

Figure 26: Chosen Short-Term Models for Recursive and Rolling Estimation for Ireland

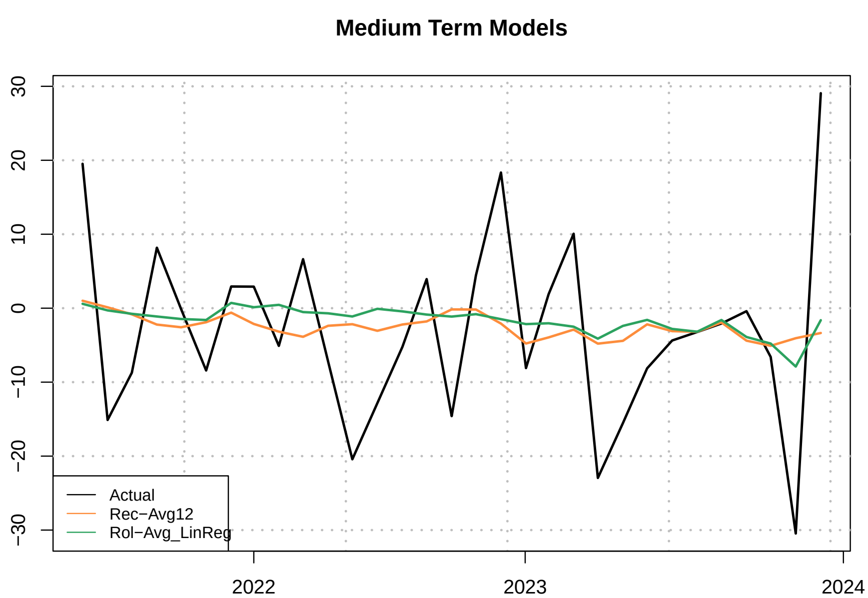

Figure 27: Chosen Medium-Term Models for Recursive and Rolling Estimation for Ireland

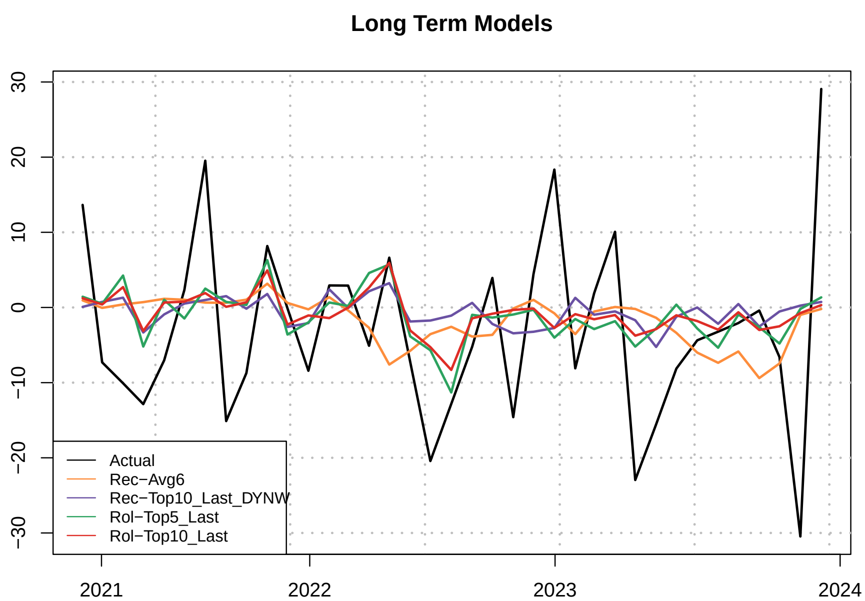

Figure 28: Chosen Long-Term Models for Recursive and Rolling Estimation for Ireland

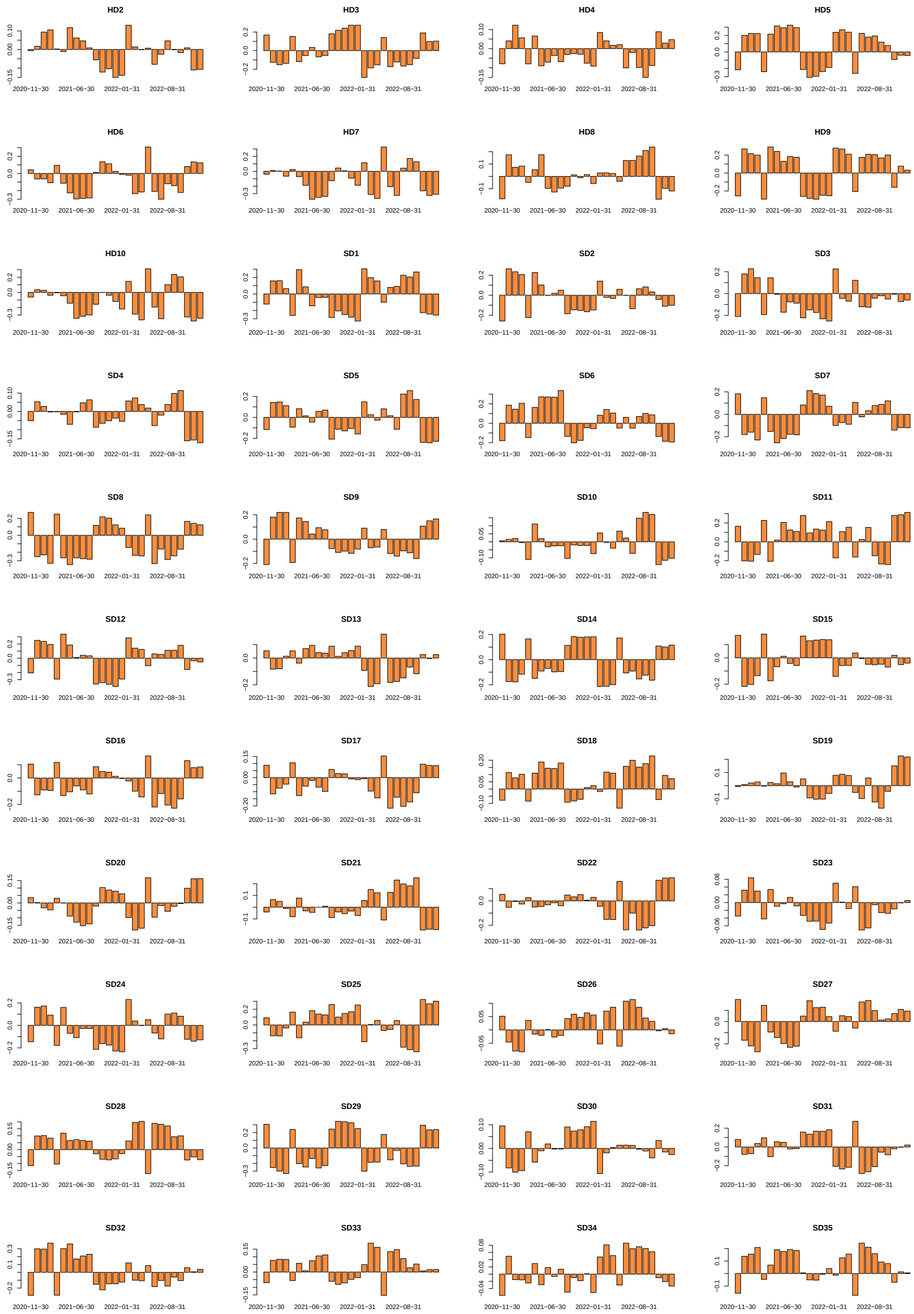

Figure 29: PCA(3)'s loadings for Ireland

10 Tables

Variable	Country	Mean	Median	Std. Dev.	Min	Max	Skewness	Kurtosis
Industrial Production (HD1)	Italy	0.03	0.19	1.86	-4.30	8.17	0.55	2.16
	Ireland	-0.196	-0.314	8.34	-30.46	29.06	-0.157	4.29
	Germany	0.13	0.201	1.562	-6.982	4.381	-0.509	4.67
Exports (HD2)	Italy	0.25	0.22	0.40	-1.22	1.52	0.33	2.45
	Ireland	0.078	-0.051	8.623	-21.87	20.908	0.0906	0.691
	Germany	0.451	0.256	3.775	-8.444	11.015	0.307	2.872
Imports (HD3)	Italy	0.12	0.11	0.35	-1.02	1.25	0.29	2.37
	Ireland	-0.034	-0.042	6.98	-20.45	19.84	0.10	0.70
	Germany	0.046	-0.095	3.127	-9.693	9.200	-0.079	3.707
Unemployment Rate (HD5)	Italy	0.00	0.00	0.23	-0.80	0.80	0.21	0.71
	Ireland	0.01	0.02	0.22	-0.60	0.70	-0.15	0.68
	Germany	-0.033	0.00	0.073	-0.20	0.20	-0.065	3.061
Retail Sales (HD8)	Italy	0.15	0.14	0.31	-0.90	1.23	0.23	2.18
	Ireland	-0.19	-0.22	7.31	-25.46	23.67	-0.19	4.12
	Germany	0.115	0.103	1.525	-5.461	5.146	-0.308	4.870
Economic Sentiment Indicator (SD23)	Italy	0.19	0.00	2.61	-8.67	9.86	-0.01	2.10
	Ireland	0.203	0.373	3.365	-11.81	10.563	-0.238	4.085
	Germany	0.196	0.096	2.191	-7.224	7.709	0.417	6.186
Financial Situation (SD28)	Italy	0.05	0.00	1.52	-4.00	4.50	0.20	2.75
	Ireland	0.07	0.00	1.43	-3.50	4.30	0.15	2.61
	Germany	0.095	0.00	1.429	-4.900	4.200	0.203	3.683
Business Climate Indicator (SD21)	Italy	0.00	0.00	0.15	-0.50	0.40	0.14	2.21
	Ireland	0.01	0.00	0.14	-0.40	0.35	0.12	2.15
	Germany	0.006	0.00	0.194	-0.70	0.60	-0.234	4.299

Table 1: Comparative Descriptive Statistics for Key Economic Indicators

Variable	Country	PCA(1) Loadings	Absolute PCA(1) Loadings
SD23	Germany	-0.282143	**0.282143**
SD24	Germany	-0.267598	**0.267598**
SD21	Germany	-0.239699	0.239699
SD33	Germany	0.232712	**0.232712**
SD4	Germany	-0.224581	0.224581
SD3	Germany	-0.215539	0.215539
SD5	Germany	-0.207447	0.207447
SD25	Germany	-0.200939	0.200939
SD26	Germany	-0.195194	**0.195194**
SD6	Germany	0.195134	0.195134
SD23	Ireland	0.300051	**0.300051**
SD26	Ireland	0.298913	**0.298913**
SD30	Ireland	0.277148	0.277148
SD33	Ireland	-0.244635	**0.244635**
SD14	Ireland	0.241423	0.241423
SD28	Ireland	0.228761	0.228761
SD25	Ireland	0.215516	0.215516
SD29	Ireland	0.213551	0.213551
SD13	Ireland	0.195486	0.195486
SD24	Ireland	0.191362	**0.191362**
SD23	Italy	-0.313050	**0.313050**
SD24	Italy	-0.310741	**0.310741**
SD4	Italy	-0.280425	0.280425
SD2	Italy	-0.268961	0.268961
SD1	Italy	-0.242891	0.242891
SD5	Italy	-0.237153	0.237153
SD21	Italy	-0.233829	0.233829
SD3	Italy	-0.208025	0.208025
SD33	Italy	0.207569	**0.207569**
SD17	Italy	-0.190652	0.190652

Table 2: Top 10 Variables with the Largest Absolute Loadings for Each Country

Short Term (Nout = 24 + 3, h = 1)				Medium Term (Nout = 36 + 3, h = 6)				Long Term (Nout = 48 + 3, h = 12)			
Model	MAE	RMSFE	SSR	Model	MAE	RMSFE	SSR	Model	MAE	RMSFE	SSR
Top10_MAE_Last3	1.0146	1.3495	0.5	ARMA(Auto)	1.1542	1.5449	0.4516	RandomForest	1.0834	1.5933	0.4444
Top10_MAE_Last3_DYNW	1.0166	1.3508	0.4583	Avg_ARMA	1.1841	1.5772	0.4194	ARMA(Auto)	1.1050	1.5048	0.4722
Top5_MAE_Last3	1.0172	1.3409	0.375	Ridge	1.1929	1.5859	0.4194	Avg_ALL	1.1058	1.5408	0.5
Top5_MAE_Last3_DYNW	1.0181	1.3456	0.375	ARMA(1,1)	1.1942	1.5861	0.4194	Avg_ML	1.1168	1.6011	0.5556
Avg_ML	1.0199	1.3462	0.5833	Lasso	1.1944	1.5864	0.4194	Top10_Last	1.1197	1.5592	0.4722
RandomForest	1.0219	1.4259	0.5417	AR(2)	1.1949	1.5864	0.4194	Top10_Last_DYNW	1.1200	1.5401	0.5
Top1_MAE_Last3	1.0229	1.4041	0.5417	AR(1)	1.1953	1.5873	0.4194	Avg_ARMA	1.1202	1.5209	0.4167
ARMA(1,1)	1.0245	1.3676	0.4167	PCA(1)	1.2018	1.5939	0.4194	PLS(2)	1.1216	1.5618	0.4167
Top10_RMSFE_Last3	1.0258	1.3637	0.5	PLS(1)	1.2063	1.6031	0.4194	AR(3)	1.1234	1.5248	0.4167
AR(1)	1.0283	1.3706	0.4167	AR(3)	1.2088	1.5924	0.3871	Ridge	1.1250	1.5255	0.4167
AR(2)	1.0300	1.3807	0.5	PLS(2)	1.2118	1.6070	0.4194	ARMA(1,1)	1.1270	1.5264	0.4167
Top5_RMSFE_Last3	1.0309	1.3639	0.4583	PLS(3)	1.2128	1.6141	0.4194	AR(2)	1.1282	1.5273	0.4167
Top10_RMSFE_Last3_DYNW	1.0311	1.3698	0.5	Avg_PenReg	1.2155	1.6107	0.3871	AR(1)	1.1285	1.5275	0.4167
PLS(1)	1.0317	1.4280	0.5	AdRidge	1.2209	1.6333	0.3871	LinReg_HD10	1.1289	1.5321	0.4167
Top5_RMSFE_Last3_DYNW	1.0339	1.3742	0.4583	Top10_MAE_Last3	1.2248	1.6073	0.3871	Top10_MAE_Last3	1.1295	1.5501	0.4167
PLS(4)	1.0440	1.4183	0.4583	Avg_PLS	1.2272	1.6360	0.4194	Top10_MAE_Last3_DYNW	1.1311	1.5524	0.4167
Avg_PLS	1.0452	1.4248	0.4167	Top10_MAE_Last3_DYNW	1.2352	1.6151	0.4194	Top10_RMSFE_Last3	1.1323	1.5519	0.4167
PCA(2)	1.0487	1.4544	0.4167	Top10_RMSFE_Last3	1.2357	1.6165	0.4516	Lasso	1.1333	1.5529	0.4444
PLS(3)	1.0489	1.4222	0.4167	PLS(4)	1.2371	1.6365	0.3871	Top10_RMSFE_Last3_DYNW	1.1340	1.5560	0.4167
PLS(5)	1.0495	1.4188	0.4583	LinReg_HD10	1.2396	1.6568	0.3871	Top5_Last	1.1353	1.5626	0.5556

Table 3: Recursive Performance Metrics for Germany

Short Term (Nout = 24 + 3, h = 1, Nroll = 12*2)				Medium Term (Nout = 36 + 3, h = 6, Nroll = 12*3)				Long Term (Nout = 48 + 3, h = 12, Nroll = 12*5)			
Model	**MAE**	**RMSFE**	**SSR**	**Model**	**MAE**	**RMSFE**	**SSR**	**Model**	**MAE**	**RMSFE**	**SSR**
PLS(2)	1.0938	1.4981	0.4545	**PLS(1)**	1.2584	1.7212	0.4286	**ARMA(Auto)**	1.0401	1.4208	0.2941
AR(2)	1.1268	1.4766	0.4545	**AdRidge**	1.2669	1.7941	0.3929	**LinReg_HD10**	1.0616	1.4319	0.3529
RandomForest	1.1277	1.4874	0.5	**Ridge**	1.2750	1.8241	0.5	**Avg_ARMA**	1.0618	1.4190	0.4412
PLS(1)	1.1315	1.5514	0.4545	**PCA(1)**	1.2905	1.8095	0.4286	**ARMA(1,1)**	1.0624	1.4172	0.4118
AdRidge	1.1331	1.5453	0.5909	**AR(2)**	1.3073	1.6977	0.4286	**AR(3)**	1.0641	1.4188	0.4412
ARMA(Auto)	1.1404	1.4994	0.1818	**AR(1)**	1.3097	1.7085	0.4286	**AR(2)**	1.0706	1.4223	0.4412
Lasso	1.1438	1.7170	0.6364	**ARMA(1,1)**	1.3267	1.8022	0.4643	**AR(1)**	1.0716	1.4235	0.4412
Top5_Last	1.1438	1.6325	0.5	**Avg_ARMA**	1.3310	1.7302	0.4286	**Avg12**	1.0846	1.4642	0.4706
AR(1)	1.1445	1.4890	0.5455	**RandomForest**	1.3330	1.8092	0.5357	**RandomForest**	1.1540	1.5365	0.3824
Avg_ARMA	1.1468	1.4942	0.4091	**Avg_PenReg**	1.3477	1.9020	0.4643	**Avg_LinReg**	1.1551	1.5431	0.4118
Avg_PenReg	1.1513	1.6088	0.5909	**AR(3)**	1.3502	1.7286	0.4286	**Ridge**	1.1666	1.5151	0.4118
ARMA(1,1)	1.1521	1.4884	0.2273	**Avg_ALL**	1.3524	1.8335	0.4286	**PLS(1)**	1.1890	1.5320	0.4118
Top10_Last	1.1670	1.6203	0.4091	**Top10_MAE_Last3**	1.3788	1.8475	0.3929	**Lasso**	1.2106	1.5324	0.4412
PLS(3)	1.1685	1.5917	0.5455	**Top10_RMSFE_Last3**	1.3836	1.8382	0.3929	**Avg_ALL**	1.2310	1.5854	0.4412
Avg_PLS	1.1695	1.5678	0.4545	**ARMA(Auto)**	1.3872	1.8013	0.2143	**Avg6**	1.2340	1.5845	0.4412
PCA(1)	1.1741	1.5376	0.3636	**Top10_MAE_Last3_DYNW**	1.3889	1.8458	0.3929	**Avg_PenReg**	1.2363	1.5744	0.4412
Top5_MAE_Last3	1.1801	1.5810	0.5	**Avg12**	1.3946	1.8333	0.3929	**Avg_Simple**	1.2536	1.5890	0.4118
Avg12	1.1808	1.4873	0.2273	**Top10_RMSFE_Last3_DYNW**	1.3964	1.8279	0.3929	**Top10_MAE_Last3**	1.2744	1.6954	0.3824
Avg_ALL	1.1826	1.5670	0.4091	**PCA(2)**	1.3985	1.9174	0.4286	**Top10_RMSFE_Last3**	1.2834	1.6990	0.3824
Ridge	1.1846	1.5551	0.4091	**Top5_MAE_Last3**	1.4062	1.8357	0.4286	**Top10_Last**	1.2899	1.6919	0.4412

Table 4: Rolling Performance Metrics for Germany

Short Term (Nout = 24 + 3, h = 1)				Medium Term (Nout = 36 + 3, h = 6)				Long Term (Nout = 48 + 3, h = 12)			
Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR
Top5_MAE_Last3	0.9891	0.9783	0.375	**AR(1)**	**1**	**1**	**0.4194**	**AR(1)**	**1**	**1**	**0.4167**
Top5_MAE_Last3_DYNW	0.9900	0.9817	0.375	**AR(2)**	0.9996	0.9994	0.4194	**AR(2)**	0.9998	0.9998	0.4167
Avg_ML	0.9918	0.9822	0.5833	**ARMA(Auto)**	0.9656	0.9732	0.4516	**AR(3)**	0.9955	0.9982	0.4167
Top10_MAE_Last3	0.9867	0.9846	0.5	**ARMA(1,1)**	0.9990	0.9993	0.4194	**ARMA(Auto)**	0.9792	0.9851	0.4722
Top10_MAE_Last3_DYNW	0.9886	0.9855	0.4583	**Ridge**	0.9980	0.9991	0.4194	**ARMA(1,1)**	0.9987	0.9993	0.4167
Top10_RMSFE_Last3	0.9976	0.9949	0.5	**Lasso**	0.9992	0.9994	0.4194	**Ridge**	0.9969	0.9987	0.4167
Top5_RMSFE_Last3	1.0025	0.9951	0.4583	**Avg_ARMA**	0.9907	0.9936	0.4194	**Avg_ARMA**	0.9927	0.9957	0.4167
ARMA(1,1)	0.9962	0.9978	0.4167								
NeuralNetwork	1.0942	0.9993	0.6667								
Top10_RMSFE_Last3_DYNW	1.0026	0.9994	0.5								
AR(1)	**1**	**1**	**0.4167**								

Table 5: Comparative Recursive Performance Metrics for Germany

Short Term (Nout = 24 + 3, h = 1, Nroll = 12*2)				Medium Term (Nout = 36 + 3, h = 6, Nroll = 12*3)				Long Term (Nout = 48 + 3, h = 12, Nroll = 12*5)			
Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR
Avg12	1.0316	0.9989	0.2273	**AR(1)**	**1**	**1**	**0.4286**	**AR(1)**	**1**	**1**	**0.4412**
AR(1)	**1**	**1**	**0.5455**	**AR(2)**	0.9982	0.9937	0.4286	**AR(2)**	0.9991	0.9991	0.4412
AR(2)	0.9845	0.9917	0.4545					**AR(3)**	0.9930	0.9967	0.4412
ARMA(1,1)	1.0066	0.9996	0.2273					**ARMA(Auto)**	0.9706	0.9981	0.2941
RandomForest	0.9853	0.9990	0.5					**ARMA(1,1)**	0.9914	0.9955	0.4118
								Avg_ARMA	0.9908	0.9969	0.4412

Table 6: Comparative Rolling Performance Metrics for Germany

Short Term (Nout = 24 + 3, h = 1)				Medium Term (Nout = 36 + 3, h = 6)				Long Term (Nout = 48 + 3, h = 12)			
Model	**MAE**	**RMSFE**	**SSR**	Model	**MAE**	**RMSFE**	**SSR**	Model	**MAE**	**RMSFE**	**SSR**
ARMA(Auto)	1.3556	1.7158	0.3354	**PLS(2)**	1.2257	1.6014	0.6774	**PCA(5)**	1.3247	1.6758	0.5833
AR(1)	1.3603	1.7144	0.4783	**PLS(3)**	1.2346	1.6131	0.6129	**Avg_PCA**	1.3251	1.6845	0.5833
LinReg_HD10	1.3742	1.7525	0.4783	**LinReg_HD10_SD23**	1.2483	1.5990	0.5484	**PCA(1)**	1.3270	1.6843	0.4722
ARMA(1,1)	1.3756	1.7266	0.4783	**Avg_PLS**	1.2505	1.6185	0.6129	**AR(2)**	1.3295	1.6815	0.4444
Ridge	1.3812	1.8350	0.5652	**PCA(4)**	1.2586	1.6210	0.5806	**PLS(1)**	1.3295	1.6875	0.4444
AdRidge	1.3829	1.8524	0.6087	**PCA(1)**	1.2600	1.6150	0.5806	**AR(1)**	1.3295	1.6814	0.4444
Avg_PenReg	1.3864	1.8471	0.5652	**PLS(4)**	1.2662	1.6310	0.5806	**Ridge**	1.3296	1.6816	0.4444
Avg_ARMA	1.4000	1.7528	0.3913	**PCA(2)**	1.2691	1.6256	0.6129	**Lasso**	1.3318	1.6830	0.4444
AdLasso	1.4045	1.8817	0.5652	**PLS(5)**	1.2730	1.6380	0.6129	**PCA(4)**	1.3324	1.6904	0.5278
Avg_LinReg	1.4150	1.8312	0.3913	**Avg_PCA**	1.2737	1.6267	0.5806	**AR(3)**	1.3326	1.6854	0.4444
Lasso	1.4324	1.8498	0.4783	**AdRidge**	1.2760	1.6395	0.5806	**Avg_ALL**	1.3326	1.7211	0.4722
AR(2)	1.4407	1.7913	0.3043	**PCA(3)**	1.2776	1.6311	0.5484	**Avg_ARMA**	1.3329	1.6850	0.4444
Avg_ALL	1.4476	1.8695	0.3913	**AdLasso**	1.2791	1.6254	0.5484	**ARMA(Auto)**	1.3332	1.6860	0.2833
PCA(4)	1.4508	1.9109	0.4348	**Avg_PenReg**	1.2797	1.6301	0.6129	**PCA(3)**	1.3358	1.6890	0.4722
PCA(2)	1.4514	1.8938	0.3913	**Avg_LinReg**	1.2891	1.6144	0.5161	**LinReg_HD10**	1.3369	1.6906	0.4722
PCA(3)	1.4539	1.8992	0.3043	**PLS(1)**	1.2934	1.6309	0.5806	**PCA(2)**	1.3377	1.6920	0.4444
Avg_PCA	1.4577	1.9111	0.3913	**Avg_ALL**	1.2935	1.6439	0.4839	**ARMA(1,1)**	1.3397	1.6919	0.4167
LinReg_HD10_SD23	1.4579	1.9266	0.3913	**Lasso**	1.2940	1.6333	0.5484	**Avg_LinReg**	1.3397	1.6879	0.3611
PLS(1)	1.4587	1.8932	0.2174	**PCA(5)**	1.3044	1.6497	0.4839	**Avg12**	1.3443	1.8402	0.5278
PLS(5)	1.4589	1.8815	0.4783	**Ridge**	1.3056	1.6430	0.4839	**LinReg_HD10_SD23**	1.3451	1.6863	0.4167

Table 7: Recursive Performance Metrics for Italy

Short Term (Nout = 24 + 3, h = 1, Nroll = 12*2)				Medium Term (Nout = 36 + 3, h = 6, Nroll = 12*3)				Long Term (Nout = 48 + 3, h = 12, Nroll = 12*5)			
Model	MAE	RMSFE	SSR	Model	MAE	RMSFE	SSR	Model	MAE	RMSFE	SSR
Top1_Last	1.3118	1.6401	0.5417	**ARMA(Auto)**	1.3017	1.6273	0.1	**ARMA(Auto)**	1.3373	1.6886	0.0892
Lasso	1.3450	1.8571	0.5	**ARMA(1,1)**	1.3258	1.6788	0.4667	**Avg_ARMA**	1.3777	1.7254	0.4595
ARMA(Auto)	1.3467	1.7513	0.332	**Avg_ARMA**	1.3698	1.7125	0.4333	**ARMA(1,1)**	1.3912	1.7430	0.4054
Avg_ARMA	1.3601	1.7938	0.5833	**PCA(1)**	1.3979	1.8045	0.4667	**AR(1)**	1.3914	1.7367	0.4595
PLS(1)	1.3749	1.8286	0.5417	**Ridge**	1.4009	1.7960	0.5	**AR(2)**	1.3925	1.7370	0.4595
RandomForest	1.3781	1.7339	0.5	**Avg_ALL**	1.4387	1.8314	0.5333	**Avg12**	1.4004	1.9004	0.5135
PCA(1)	1.4043	1.8505	0.4583	**AR(3)**	1.4429	1.7807	0.4333	**AR(3)**	1.4020	1.7480	0.4595
Avg6	1.4065	1.8205	0.4167	**AR(1)**	1.4465	1.7823	0.4333	**Top10_Last_DYNW**	1.4029	1.7959	0.5135
AR(3)	1.4157	1.8647	0.5833	**AR(2)**	1.4487	1.7837	0.4333	**Ridge**	1.4033	1.7523	0.4595
ARMA(1,1)	1.4164	1.8449	0.5	**Lasso**	1.4509	1.8037	0.4667	**Top10_Last**	1.4196	1.8021	0.5405
PLS(2)	1.4190	1.8745	0.5	**Avg6**	1.4511	1.8596	0.4667	**Avg_ALL**	1.4247	1.7878	0.4865
Top5_RMSFE_Last3_DYNW	1.4192	1.8671	0.5417	**PLS(1)**	1.4614	1.8510	0.4667	**Lasso**	1.4270	1.7656	0.4324
Top5_RMSFE_Last3	1.4219	1.8580	0.5417	**PCA(3)**	1.4620	1.9785	0.5333	**PCA(2)**	1.4399	1.8280	0.6216
Avg12	1.4223	1.8208	0.375	**Avg_PenReg**	1.4889	1.8894	0.4667	**AdRidge**	1.4408	1.7822	0.5135
Avg_ALL	1.4231	1.8724	0.625	**Top10_RMSFE_Last3**	1.4928	1.8636	0.5333	**Avg_PenReg**	1.4463	1.7796	0.4595
AR(1)	1.4270	1.8716	0.5417	**AdRidge**	1.5006	1.9007	0.4667	**Top5_Last_DYNW**	1.4473	1.8481	0.4865
Top10_Last	1.4286	1.8559	0.4167	**Top10_MAE_Last3**	1.5067	1.8587	0.5667	**PCA(1)**	1.4757	1.7914	0.3784
AdRidge	1.4346	1.8455	0.4167	**Avg12**	1.5110	1.9178	0.4	**Top5_Last**	1.4765	1.8374	0.4865
Top10_MAE_Last3_DYNW	1.4355	1.8446	0.5	**PLS(2)**	1.5138	2.0233	0.5333	**Top10_RMSFE_Last3**	1.4978	1.8580	0.5135
Top5_MAE_Last3	1.4358	1.8784	0.625	**Top10_MAE_Last3_DYNW**	1.5152	1.8824	0.5667	**PLS(1)**	1.5074	1.8744	0.4324

Table 8: Rolling Performance Metrics for Italy

Short Term (Nout = 24 + 3, h = 1)				Medium Term (Nout = 36 + 3, h = 6)				Long Term (Nout = 48 + 3, h = 12)			
Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR
AR(1)	**1**	**1**	**0.4783**	**AR(1)**	**1**	**1**	**0.4516**	**AR(1)**	**1**	**1**	**0.4444**
ARMA(Auto)	0.9965	1.0008	0.3354	**ARMA(Auto)**	0.9964	0.9971	0.185	**AR(2)**	1.0000	1.0001	0.5333
				ARMA(1,1)	0.9992	0.9994	0.4516	**PCA(1)**	0.9981	1.0017	0.4722
				LinReg_HD10	0.9483	0.9745	0.5484	**PCA(5)**	0.9964	0.9967	0.5833
				Ridge	0.9918	1.0013	0.4839	**PLS(1)**	1.0000	1.0037	0.4444
				Lasso	0.9830	0.9954	0.5484	**Avg_PCA**	0.9967	1.0018	0.5833
				AdRidge	0.9693	0.9992	0.5806				
				AdLasso	0.9717	0.9906	0.5484				
				PCA(1)	0.9571	0.9843	0.5806				
				PCA(2)	0.9640	0.9907	0.6129				
				PCA(3)	0.9705	0.9941	0.5484				
				PCA(4)	0.9561	0.9879	0.5806				
				PCA(5)	0.9909	1.0054	0.4839				
				PLS(1)	0.9826	0.9939	0.5806				
				PLS(2)	0.9311	0.9760	0.6774				
				PLS(3)	0.9379	0.9831	0.6129				
				PLS(4)	0.9619	0.9940	0.5806				
				PLS(5)	0.9671	0.9983	0.6129				
				Avg_ARMA	0.9995	0.9996	0.4516				
				Avg_LinReg	0.9793	0.9839	0.5161				
				Avg_PenReg	0.9722	0.9935	0.6129				
				Avg_PCA	0.9675	0.9914	0.5806				
				Avg_PLS	0.9499	0.9864	0.6129				
				Avg_ALL	0.9826	1.0019	0.4839				

Table 9: Comparative Recursive Performance Metrics for Italy

Short Term (Nout = 24 + 3, h = 1, Nroll = 12*2)				Medium Term (Nout = 36 + 3, h = 6, Nroll = 12*3)				Long Term (Nout = 48 + 3, h = 12, Nroll = 12*5)			
Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR
Avg6	0.9856	0.9727	0.4167	**AR(1)**	**1**	**1**	**0.4333**	**AR(1)**	**1**	**1**	**0.4595**
Avg12	0.9967	0.9728	0.375	**AR(3)**	0.9976	0.9991	0.4333	**ARMA(Auto)**	0.9611	0.9723	0.0892
AR(1)	**1**	**1**	**0.5417**	**ARMA(Auto)**	0.9000	0.9130	0.1	**Avg_ARMA**	0.9902	0.9935	0.4595
AR(3)	0.9921	0.9963	0.5833	**ARMA(1,1)**	0.9166	0.9419	0.4667				
ARMA(Auto)	0.9437	0.9357	0.332	**Avg_ARMA**	0.9470	0.9609	0.4333				
ARMA(1,1)	0.9926	0.9857	0.5								
Lasso	0.9426	0.9922	0.5								
AdRidge	1.0053	0.9861	0.4167								
PCA(1)	0.9841	0.9887	0.4583								
PLS(1)	0.9635	0.9770	0.5417								
RandomForest	0.9658	0.9264	0.5								
Avg_ARMA	0.9531	0.9584	0.5833								
Avg_PLS	1.0223	0.9885	0.4583								
Top1_Last	0.9193	0.8763	0.5417								
Top10_Last	1.0012	0.9916	0.4167								
Top10_MAE_Last3	1.0072	0.9833	0.5								
Top10_MAE_Last3_DYNW	1.0060	0.9856	0.5								
Top5_RMSFE_Last3	0.9964	0.9927	0.5417								
Top5_RMSFE_Last3_DYNW	0.9946	0.9976	0.5417								

Table 10: Comparative Rolling Performance Metrics for Italy

Short Term (Nout = 24 + 3, h = 1)				Medium Term (Nout = 36 + 3, h = 6)				Long Term (Nout = 48 + 3, h = 12)			
Model	**MAE**	**RMSFE**	**SSR**	**Model**	**MAE**	**RMSFE**	**SSR**	**Model**	**MAE**	**RMSFE**	**SSR**
RandomForest	9.5164	12.8736	0.625	**Avg12**	9.2326	12.1239	0.6452	**Avg6**	9.3137	12.0100	0.6757
Top10_Last_DYNW	9.5264	13.2112	0.5833	**Avg_ML**	9.6143	12.7709	0.6129	**Top10_Last_DYNW**	9.3945	11.8606	0.6216
Top10_Last	9.5699	13.2591	0.6667	**RandomForest**	9.6588	12.7127	0.6452	**Top5_MAE_Last3_DYNW**	9.4464	11.9257	0.5946
Avg_ML	9.6054	12.7661	0.6667	**Top5_MAE_Last3**	9.8156	12.5467	0.5161	**Top5_Last_DYNW**	9.4776	11.9620	0.5676
Top10_MAE_Last3	9.6620	13.4447	0.7083	**LinReg_HD10**	9.8335	12.5634	0.5806	**Top5_MAE_Last3**	9.4780	11.9305	0.5946
Top10_RMSFE_Last3	9.6936	13.3425	0.7083	**NeuralNetwork**	9.8430	12.9472	0.6774	**Avg12**	9.5291	12.2396	0.5676
Top10_MAE_Last3_DYNW	9.7037	13.4713	0.6667	**Avg6**	9.8597	12.4727	0.5806	**Top10_MAE_Last3_DYNW**	9.5319	12.0191	0.5676
Top10_RMSFE_Last3_DYNW	9.7850	13.4566	0.6667	**Top5_MAE_Last3_DYNW**	9.8646	12.6307	0.5484	**Top10_RMSFE_Last3_DYNW**	9.5419	12.0275	0.5676
Top5_RMSFE_Last3	9.8172	13.8338	0.6667	**Avg_ALL**	9.8712	12.7292	0.5484	**Top10_MAE_Last3**	9.5599	12.0288	0.5405
PLS(4)	9.8903	13.2161	0.5833	**Avg_LinReg**	9.8903	12.6699	0.6129	**Top10_RMSFE_Last3**	9.5649	12.0288	0.5676
PLS(3)	9.9624	13.3107	0.5	**Avg_PenReg**	9.8992	12.9186	0.5484	**Top10_Last**	9.6137	12.0467	0.5676
Avg12	9.9692	13.5404	0.6667	**AdRidge**	9.9073	12.9617	0.6129	**Top5_Last**	9.6570	12.0995	0.5676
Top5_RMSFE_Last3_DYNW	9.9719	13.8420	0.6667	**AR(2)**	9.9140	12.7211	0.4516	**Top5_RMSFE_Last3**	9.6747	12.1897	0.4865
PLS(5)	9.9853	13.3900	0.4583	**Lasso**	9.9209	12.9408	0.5484	**Top5_RMSFE_Last3_DYNW**	9.6955	12.1776	0.4865
GradientBoosting	10.0141	13.4477	0.5833	**AdLasso**	9.9410	13.0762	0.5806	**LinReg_HD10**	9.7971	12.3490	0.4054
Top5_Last_DYNW	10.0253	14.1412	0.5417	**AR(1)**	9.9431	12.7395	0.4516	**ARMA(Auto)**	9.8008	12.3127	0.1892
Avg_PLS	10.0718	13.2388	0.5833	**Ridge**	9.9441	12.7622	0.5806	**AR(1)**	9.8080	12.3418	0.4054
Top5_Last	10.0871	14.2972	0.5	**LinReg_HD10_SD23**	9.9472	12.8073	0.5484	**Avg_ARMA**	9.8087	12.3367	0.4054
Avg_ALL	10.0924	13.2031	0.5833	**AR(3)**	9.9486	12.7542	0.4839	**Avg_LinReg**	9.8091	12.2874	0.4595
Lasso	10.1318	13.1996	0.5833	**ARMA(1,1)**	9.9515	12.7458	0.3871	**AR(2)**	9.8093	12.3413	0.4054

Table 11: Recursive Performance Metrics for Ireland

Short Term (Nout = 24 + 3, h = 1, Nroll = 12*2)				Medium Term (Nout = 36 + 3, h = 6, Nroll = 12*3)				Long Term (Nout = 48 + 3, h = 12, Nroll = 12*5)			
Model	**MAE**	**RMSFE**	**SSR**	**Model**	**MAE**	**RMSFE**	**SSR**	**Model**	**MAE**	**RMSFE**	**SSR**
PCA(3)	9.4843	13.1421	0.625	**LinReg_HD10**	9.1947	11.7310	0.6452	**Top5_Last**	8.7055	11.3708	0.6486
PCA(1)	9.5974	12.9473	0.6667	**Avg_LinReg**	9.1996	11.8391	0.7097	**Top10_Last**	8.7823	11.3906	0.6486
Ridge	9.7453	13.0827	0.6667	**LinReg_HD10_SD23**	9.2044	11.9674	0.6774	**Top5_RMSFE_Last3**	8.8895	11.4700	0.6757
PCA(2)	9.7675	13.0575	0.6667	**Avg12**	9.2326	12.1239	0.6452	**Top10_MAE_Last3**	8.9162	11.4954	0.5946
AdRidge	9.8473	13.5121	0.625	**AdRidge**	9.2344	12.3232	0.6774	**Top10_MAE_Last3_DYNW**	9.0176	11.5523	0.5676
PLS(1)	9.8557	13.1989	0.625	**AR(2)**	9.5340	12.4541	0.7097	**Top5_MAE_Last3**	9.0609	11.5871	0.6757
PLS(2)	9.9050	13.4332	0.6667	**AR(3)**	9.5717	12.4839	0.6129	**Top10_Last_DYNW**	9.0628	11.4246	0.6216
PLS(3)	9.9200	13.6018	0.6667	**AR(1)**	9.6043	12.4638	0.6129	**Top10_RMSFE_Last3**	9.0815	11.7008	0.6216
Avg_PLS	9.9271	13.6834	0.6667	**Avg_PenReg**	9.6154	12.6202	0.6129	**Top5_Last_DYNW**	9.0977	11.5477	0.5946
Avg_ALL	9.9402	13.7510	0.6667	**PCA(1)**	9.6360	12.6406	0.6452	**Top5_RMSFE_Last3_DYNW**	9.1474	11.6524	0.6486
Top5_Last	9.9478	14.3888	0.6667	**Avg_ARMA**	9.6963	12.5967	0.6452	**Top5_MAE_Last3_DYNW**	9.2057	11.6684	0.6757
Avg12	9.9692	13.5404	0.6667	**Ridge**	9.6993	12.4415	0.5806	**Top10_RMSFE_Last3_DYNW**	9.2436	11.7776	0.5946
Avg_PCA	10.0683	13.8511	0.625	**PLS(2)**	9.7805	12.9696	0.6129	**Avg6**	9.3137	12.0100	0.6757
RandomForest	10.0748	14.0657	0.625	**Top10_RMSFE_Last3_DYNW**	9.7931	12.5427	0.5806	**PLS(1)**	9.3394	11.8814	0.5676
Top10_Last	10.1296	13.8775	0.6667	**AdLasso**	9.8079	12.8512	0.6129	**PLS(2)**	9.3894	11.8743	0.6216
Top10_RMSFE_Last3	10.1776	13.9217	0.6667	**ARMA(1,1)**	9.8296	12.7229	0.3871	**Avg_ALL**	9.4973	12.0023	0.5946
Top10_Last_DYNW	10.1911	13.9799	0.5833	**Avg6**	9.8597	12.4727	0.5806	**Avg12**	9.5291	12.2396	0.5676
PLS(4)	10.2039	13.2022	0.5417	**Top10_Last**	9.8638	12.5287	0.6452	**LinReg_HD10**	9.5425	12.0780	0.5135
AR(1)	10.2290	13.6523	0.625	**Top10_RMSFE_Last3**	9.8645	12.5764	0.6129	**Avg_LinReg**	9.5446	11.9944	0.5676
Avg_PenReg	10.2583	13.4286	0.6667	**Avg_ALL**	9.8699	12.6886	0.5806	**Avg_PenReg**	9.6052	12.0876	0.5676

Table 12: Rolling Performance Metrics for Ireland

Short Term (Nout = 24 + 3, h = 1)				Medium Term (Nout = 36 + 3, h = 6)				Long Term (Nout = 48 + 3, h = 12)			
Model	**relMAE**	**relRMSFE**	**SSR**	**Model**	**relMAE**	**relRMSFE**	**SSR**	**Model**	**relMAE**	**relRMSFE**	**SSR**
Avg12	0.9632	0.9970	0.6667	**Avg6**	0.9916	0.9791	0.5806	**Avg6**	0.9496	0.9731	0.6757
AR(1)	**1**	**1**	**0.4583**	**Avg12**	0.9285	0.9517	0.6452	**Avg12**	0.9716	0.9917	0.5676
LinReg_HD10	1.0000	0.9849	0.4583	**AR(1)**	**1**	**1**	**0.4516**	**AR(1)**	**1**	**1**	**0.4054**
LinReg_HD10_SD23	0.9991	0.9893	0.4583	**AR(2)**	0.9971	0.9986	0.4516	**AR(2)**	1.0001	1.0000	0.4054
Ridge	0.9836	0.9742	0.5417	**LinReg_HD10**	0.9890	0.9862	0.5806	**ARMA(Auto)**	0.9993	0.9976	0.1892
Lasso	0.9789	0.9719	0.5833	**RandomForest**	0.9714	0.9979	0.6452	**LinReg_HD10_SD23**	1.0037	0.9934	0.4595
AdRidge	0.9928	0.9908	0.5417	**Avg_LinReg**	0.9947	0.9945	0.6129	**PLS(1)**	1.0097	0.9907	0.3243
AdLasso	0.9843	0.9844	0.5833	**Avg_ALL**	0.9928	0.9992	0.5484	**PLS(4)**	1.0235	0.9985	0.4865
PCA(1)	0.9886	0.9788	0.5	**Top5_MAE_Last3**	0.9872	0.9849	0.5161	**NeuralNetwork**	1.0323	0.9972	0.5135
PCA(2)	0.9974	0.9860	0.375	**Top10_MAE_Last3**	1.0033	0.9957	0.5161	**Avg_ARMA**	1.0001	0.9996	0.4054
PCA(3)	0.9999	0.9960	0.375	**Top5_MAE_Last3_DYNW**	0.9921	0.9915	0.5484	**Avg_LinReg**	1.0001	0.9956	0.4595
PCA(4)	1.0042	0.9971	0.375	**Top10_MAE_Last3_DYNW**	1.0044	0.9971	0.4839	**Top5_Last**	0.9846	0.9804	0.5676
PLS(1)	0.9863	0.9746	0.5	**Top5_RMSFE_Last3**	1.0105	0.9893	0.4839	**Top10_Last**	0.9802	0.9761	0.5676
PLS(2)	0.9963	0.9743	0.5417	**Top10_RMSFE_Last3**	1.0069	0.9933	0.5161	**Top5_Last_DYNW**	0.9663	0.9692	0.5676
PLS(3)	0.9625	0.9801	0.5					**Top10_Last_DYNW**	0.9578	0.9610	0.6216
PLS(4)	0.9556	0.9731	0.5833					**Top5_MAE_Last3**	0.9664	0.9667	0.5946
PLS(5)	0.9648	0.9859	0.4583					**Top10_MAE_Last3**	0.9747	0.9746	0.5405
RandomForest	0.9194	0.9479	0.625					**Top5_MAE_Last3_DYNW**	0.9631	0.9663	0.5946
GradientBoosting	0.9675	0.9902	0.5833					**Top10_MAE_Last3_DYNW**	0.9719	0.9739	0.5676
NeuralNetwork	0.9875	0.9353	0.4583					**Top5_RMSFE_Last3**	0.9864	0.9877	0.4865
Avg_LinReg	0.9995	0.9868	0.5417					**Top10_RMSFE_Last3**	0.9752	0.9746	0.5676
Avg_PenReg	0.9842	0.9798	0.5					**Top5_RMSFE_Last3_DYNW**	0.9885	0.9867	0.4865
Avg_PCA	0.9991	0.9914	0.375					**Top10_RMSFE_Last3_DYNW**	0.9729	0.9745	0.5676
Avg_PLS	0.9731	0.9748	0.5833								
Avg_ML	0.9280	0.9400	0.6667								
Avg_ALL	0.9751	0.9721	0.5833								
Top1_Last	1.0359	0.9884	0.4583								
Top10_Last	0.9246	0.9763	0.6667								
Top10_Last_DYNW	0.9204	0.9727	0.5833								
Top1_MAE_Last3	0.9902	0.9896	0.5417								
Top10_MAE_Last3	0.9335	0.9899	0.7083								
Top10_MAE_Last3_DYNW	0.9375	0.9919	0.6667								
Top1_RMSFE_Last3	1.0228	0.9940	0.5417								
Top10_RMSFE_Last3	0.9366	0.9824	0.7083								
Top10_RMSFE_Last3_DYNW	0.9454	0.9908	0.6667								

Table 13: Comparative Recursive Performance Metrics for Ireland

Short Term (Nout = 24 + 3, h = 1, Nroll = 12*2)				Medium Term (Nout = 36 + 3, h = 6, Nroll = 12*3)				Long Term (Nout = 48 + 3, h = 12, Nroll = 12*5)			
Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR	Model	relMAE	relRMSFE	SSR
Avg12	0.9746	0.9918	0.6667	**Avg12**	0.9613	0.9727	0.6452	**Avg6**	0.9661	0.9831	0.6757
AR(1)	**1**	**1**	**0.625**	**AR(1)**	**1**	**1**	**0.6129**	**AR(1)**	**1**	**1**	**0.5676**
ARMA(Auto)	1.0144	0.9935	0.2083	**AR(2)**	0.9927	0.9992	0.7097	**AR(2)**	1.0011	0.9998	0.5676
Ridge	0.9527	0.9583	0.6667	**LinReg_HD10**	0.9573	0.9412	0.6452	**AR(3)**	1.0006	0.9997	0.5676
Lasso	1.0171	0.9787	0.5833	**LinReg_HD10_SD23**	0.9584	0.9602	0.6774	**LinReg_HD10**	0.9899	0.9887	0.5135
AdRidge	0.9627	0.9897	0.625	**Ridge**	1.0099	0.9982	0.5806	**LinReg_HD10_SD23**	1.0040	0.9983	0.5676
PCA(1)	0.9383	0.9484	0.6667	**AdRidge**	0.9615	0.9887	0.6774	**Ridge**	0.9984	0.9974	0.5676
PCA(2)	0.9549	0.9564	0.6667	**Avg_LinReg**	0.9579	0.9499	0.7097	**AdRidge**	1.0031	0.9922	0.5405
PCA(3)	0.9272	0.9626	0.625					**PLS(1)**	0.9688	0.9726	0.5676
PLS(1)	0.9635	0.9668	0.625					**PLS(2)**	0.9740	0.9720	0.6216
PLS(2)	0.9683	0.9839	0.6667					**RandomForest**	0.9983	0.9908	0.6216
PLS(3)	0.9698	0.9963	0.6667					**Avg_LinReg**	0.9901	0.9818	0.5676
PLS(4)	0.9975	0.9670	0.5417					**Avg_PenReg**	0.9964	0.9894	0.5676
Avg_PenReg	1.0029	0.9836	0.6667					**Avg_ML**	1.0463	0.9880	0.5946
								Avg_ALL	0.9852	0.9825	0.5946
								Top5_Last	0.9030	0.9308	0.6486
								Top10_Last	0.9110	0.9324	0.6486
								Top5_Last_DYNW	0.9437	0.9452	0.5946
								Top10_Last_DYNW	0.9401	0.9352	0.6216
								Top5_MAE_Last3	0.9399	0.9485	0.6757
								Top10_MAE_Last3	0.9249	0.9410	0.5946
								Top5_MAE_Last3_DYNW	0.9549	0.9551	0.6757
								Top10_MAE_Last3_DYNW	0.9354	0.9456	0.5676
								Top5_RMSFE_Last3	0.9221	0.9389	0.6757
								Top10_RMSFE_Last3	0.9420	0.9578	0.6216
								Top5_RMSFE_Last3_DYNW	0.9489	0.9538	0.6486
								Top10_RMSFE_Last3_DYNW	0.9589	0.9641	0.5946

Table 14: Comparative Rolling Performance Metrics for Ireland