Sapienza Università di Roma

Sapienza University of Rome

Department of Economics and Law
PhD in Economics

Thesis For The Degree Of Doctor Of Philosophy

Essays on Machine Learning approaches to Macroeconomic Modeling

Thesis Advisor
Prof. Massimiliano Tancioni

Candidate
Milos Ciganovic

Academic Year MMXIX-MMXXII (XXXV cycle)

To my family

Contents

List of Figures **iv**

List of Tables **v**

1 Introduction **1**

2 The Rise of Machine Learning: A Shift from Econometrics **4**
2.1 Introduction 5
2.2 Machine learning in Economic research 11
2.2.1 Forecasting and prediction with Machine learning 11
2.2.2 Sentiment analysis and Natural Language Processing 13
2.2.3 Deep learning in Economics 16
2.2.4 Causal Machine Learning 19
2.3 Conclusions 22

3 Augmented Coincident Index: Using soft information to track the economy **24**
3.1 Introduction 25
3.2 Text analysis 27
3.2.1 Advancements in Text Analysis: From TF-IDF to Transformers 28
3.3 Data 31
3.4 Economic Index 35
3.4.1 Newspapers sentiment heterogeneity 37
3.5 Real-time indicator evaluation 41
3.6 Conclusions 46

4 Nowcasting inflation with Lasso-regularized vector autoregressions and mixed frequency data **49**
4.1 Introduction 50
4.2 Data and missing data imputation 53
4.2.1 Imputation of high frequency missing data 55
4.3 The econometric model 56
4.3.1 The Lasso-VAR 57
4.3.2 Lasso-VAR time series cross-validation 59
4.4 Results and model comparisons 59
4.4.1 Evaluation of the Lasso-VAR nowcasting performances 60
4.4.2 Model comparisons 63
4.5 Robustness checks 66
4.6 Conclusions 69

Bibliography **71**

A Data sample **82**

B Technical Box: BERT Model - Detailed Architecture **83**

C Technical Box: DeBERTa Model - Detailed Architecture **85**

D Trending words **87**

E Additional results: Augmented Coincident Index **90**

F Dataset **92**

G Additional results: Nowcasting inflation **94**

H Robustness checks **96**

I Trace plots **99**

List of Figures

2.1 Term frequency across top journals . . . 5
2.2 Graphical example of the five-split time series cross-validation . . . 10
2.3 Compute power for training . . . 14
2.4 Transfer learning . . . 17

3.1 Embedding . . . 30
3.2 Wordcloud - Il Sole 24 Ore . . . 32
3.3 Transformers - model architecture by Vaswani et al. (2017) . . . 34
3.4 Economy, Business & Finance articles . . . 35
3.5 Difference between DeBERTa and standard BERT Qian Qian et al. (2022) . . . 37
3.6 Main Italian newspapers sentiment indices . . . 38
3.7 Variance explained . . . 39
3.8 Methodology diagram. EBS is an abbreviation for economics, business & finance topics and ESI for economic sentiment index. . . . 42
3.9 Italian real-time soft information index of economic activity . . . 43
3.10 Italian real-time soft information index of economic activity vs. ITA-Coin . . . 45
3.11 Cross-Correlation with GDP . . . 45
3.12 Cross-Correlation with ITA-Coin . . . 46

4.1 Graphical exemplification of the KF update-prediction procedure. . . . 56
4.2 Graphical example of the five-split time series cross-validation . . . 60
4.3 Nowcasts, actual HICP inflation and SPF estimates . . . 62
4.4 Root Mean Squared Forecast Error . . . 63
4.5 Nowcasting performance per week of the month . . . 64
4.6 Nowcasts: Lasso-VAR vs. FAVAR, BVAR and AR(2) . . . 65
4.7 Models' nowcasting performance per week of the month . . . 67

D.1 Corriere della Sera . . . 88
D.2 Il Fatto quotidiano . . . 88
D.3 La Stampa . . . 88
D.4 I Sole 24 Ore . . . 88
D.5 La Repubblica . . . 89

E.1 Italian stock exchange index - FTSE MIB . . . 91
E.2 Economic policy uncertainty - Italy . . . 91

H.1 Nowcasting Core Inflation: Lasso-VAR vs. FAVAR and BVAR . . . 97

I.1 Trace plot - Pre Covid 19 period . . . 100
I.2 Trace plot - Covid 19 period . . . 101
I.3 Trace plot - Post Covid 19 and high Inflation period . . . 101

List of Tables

3.1 Raw count of total articles and articles included 35
3.2 Estimated effects of Newspaper indices on GDP at different horizons 40

4.1 Model comparisons: RMSFE and Diebold-Mariano test results 65

A.1 Raw count of total articles and articles included 82

F.1 Philips-Perron Tests for nonstationarity 92
F.2 Summary of data, frequency of observations and source 93

G.1 Monthly RMSFE 95

H.1 Model comparisons: RMSFE with and without policy variables 96
H.2 RMSFE across different models and forecast horizons 97
H.3 RMSFE Lasso-ridge-elastic net 97
H.4 RMSFE Core inflation 98

Chapter 1

Introduction

The thesis is an in-depth examination of the potential of machine learning and artificial intelligence techniques to improve the accuracy of macroeconomic forecasting and *real-time* tracking of economic activity. The research seeks to understand how these innovative methods can be used to provide more precise and up-to-date information about the state of the economy, thus allowing for better predictions of macroeconomic trends. The study will focus on different machine learning and econometric approaches, ranging from neural networks to time series models, and incorporate various data types, including economic indicators and financial market data. The aim is to provide economists and policymakers with the tools they need to make informed economic policy decisions and help businesses and investors make wise investments and strategic decisions. This research is a valuable contribution to the field of macroeconomic forecasting, helping to improve the accuracy of predictions and providing stakeholders with the information they need to make well-informed decisions. By using state-of-the-art techniques from machine learning and artificial intelligence, this thesis can significantly enhance the understanding and governance of the economy.

Chapter 2 provides an overview of the existing literature on the use of machine learning and artificial intelligence in the fields of macroeconomic forecasting, policy analysis, and causality. This literature review is not exhaustive, but it covers the most significant and relevant papers in the field, highlighting the key findings and contributions of each study. The chapter begins with the evolution of econometric and machine learning from being separate fields to becoming more and more intertwined and then moving on to more recent developments in the field. This includes a discussion of the different machine learning used for forecasting, text analysis used in economic research, deep learning methods that bring new possibilities, and expanding the world of machine learning to causality.

Chapter 3 examines a new approach to developing an economic index that can provide real-time insights into economic activity. The main idea of using transformer models is introduced. Transformers are a type of neural network architecture that has been successful in natural language processing tasks to analyze unstructured data from various sources such as television, newspapers, and social media. The methodology section of the chapter explains the general approaches used for text analysis in economics and introduces the specific transformer model that will be used in the research. A comprehensive outline of the data collection and inclusion process used to gather information from various sources is presented, followed by a description of the new soft information economic index developed using the transformer model. In the chapter, the analysis results and insights into the accuracy and usefulness of the index are discussed.

Chapter 4 explores the utilization of the Lasso regularization method in mixed-frequency environments for nowcasting inflation. This chapter provides a comprehensive examination of the topic, starting with a discussion of data issues and a description of the data set used in the study. That includes information on the continuous updating process of the data set and how missing values were estimated. The econometric modeling strategy and its validation are also discussed in the chapter, the outline of the method used to model the data and how it was validated for accuracy as well as how the Lasso regularization method was utilized in mixed-frequency environments to nowcast inflation and the metrics used to determine the accuracy of the predictions.

This thesis comprehensively examines the use of machine learning and artificial intelligence in macroeconomic forecasting, policy analysis, and causality. Through its three chapters, the thesis look into existing literature, presents new approaches and ideas, and provides valuable insights into the potential implications of these methods for the field. The thesis can serve as a valuable resource for researchers, policymakers, and anyone interested in the future of macroeconomic forecasting and its potential advancements. Hence, this thesis highlights the importance of this research and its potential to shape the future of macroeconomic forecasting.

Chapter 2

The Rise of Machine Learning: A Shift from Econometrics

Abstract

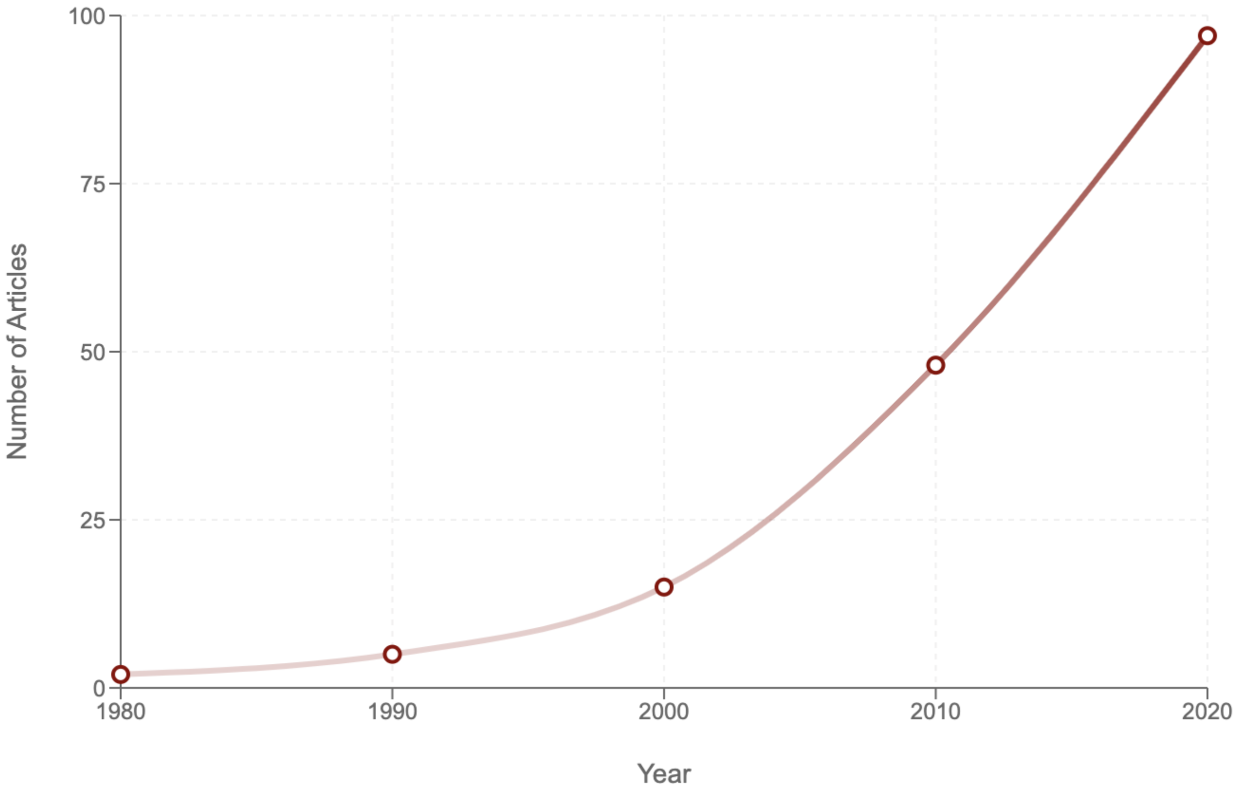
Econometrics and machine learning were once separate disciplines. Econometrics uses statistics to explain economic phenomena, whereas machine learning uses data and algorithms to enhance computer performance. With "big data" and the need to address complex problems, econometrics, and machine learning are being combined. This integration enables machine learning to evaluate and interpret massive amounts of data, solving problems more efficiently. This paper surveys a wide range of studies and highlights the main empirical findings regarding the ever-increasing use of machine learning in economics.

Keywords: Econometrics, Machine learning, Literature review, Natural language processing, Causality.

2.1 Introduction

Econometrics and machine learning have traditionally been two separate disciplines. Machine learning involves using data and algorithms to improve computer performance, whereas econometrics involves using statistical techniques to comprehend economic events. In response to the emergence of "big data" and the need to solve ever-more complex problems, there has been a growing desire to combine econometrics and machine learning. This integration enables machine learning techniques to evaluate and comprehend vast amounts of available data, allowing for more efficient and effective problem-solving. As a result, the lines between econometrics and machine learning are increasingly becoming blurred as the two disciplines collaborate to advance our understanding of the world and solve critical problems. Hence, in recent years this has led to a substantial rise in publications regarding the use of machine learning in economics, as we can see in Figure 2.1.

Figure 2.1: Term frequency across top journals



Notes: Term frequency of bigram *Machine learning* across top journals in economics (The Quarterly Journal of Economics, Econometrica, The American Economic Review, The Review of Economics and Statistics, Journal of Applied Econometrics, Econometrics Journal), source: Constellate.

Econometrics and Machine learning (ML) deal with data analysis and prediction, but their approaches and focus differ. Econometrics is a branch of economics that uses mathematical statistics and probability to understand and quantify the relationship, causes, and dynamics between different variables in data. The models and methodologies of econometrics are built on a robust theoretical foundation and are often expressed using precise mathematical notations and symbols. As a result, econometric models are widely used in policy research, forecasting, and other applications where understanding cause-and-effect relationships are essential. The accuracy of econometric models depends not only on the precision of predictions but also on the validity of underlying assumptions and the statistical properties of results.

Machine learning is a specialized area within the broader field of artificial intelligence, dedicated to the development of computer algorithms and models that learn from data over time without being explicitly programmed. The aim is to create models that generalize well to new inputs and make accurate predictions based on unseen data. This is achieved by training the models on large amounts of data, enabling them to identify patterns and relationships within that data.

The performance of machine learning models is primarily evaluated based on their empirical performance on specific tasks, such as image classification or speech recognition, using well-defined data sets. The power of machine learning lies in its ability to incrementally learn from vast amounts of data, making it an ideal tool for solving complex, data-rich problems across various fields.

By enabling computers to adapt to new information, machine learning serves as a robust tool in numerous applications, facilitating the automation of complex tasks and improving decision-making processes through data-driven insights.

Despite a common quantitative and theoretical basis for machine learning algorithms, their usage may be usually explained by a very simple notation. This is the reason for their big success and popularity if we take one of the most famous algorithms, the XGBoost (See Algorithm 1), its success is primarily due to its exceptional performance in various machine learning competitions, rather than being based on its rigorous mathematical proofs.

Algorithm 1 XGBoost Algorithm

Input: Training set $\{x_i, y_i\}$, $i = 1, 2, \ldots, n$, a differentiable loss function $L(y, \hat{y})$, number of iterations M, and learning rate η.
Initialize: $\hat{y}_i^{(0)} = 0$, for all i.
for $t = 1$ to M **do**
    Compute gradients and second-order derivatives:
    $g_i = \partial_{\hat{y}^{(t-1)}} L(y_i, \hat{y}_i^{(t-1)})$
    $h_i = \partial^2_{\hat{y}^{(t-1)}} L(y_i, \hat{y}_i^{(t-1)})$
    Fit a regression tree to the targets $-g_i/h_i$, resulting in regions R_{jt}, $j = 1, 2, \ldots, J_t$.
    For each region R_{jt}, compute:
    $w^*_{jt} = -\frac{\sum_{i \in R_{jt}} g_i}{\sum_{i \in R_{jt}} h_i + \lambda}$
    $\Omega(f_t) = \gamma J_t + \frac{1}{2}\lambda \sum_{j=1}^{J_t} w^{*2}_{jt}$
    Update model:
    $\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta \sum_{j=1}^{J_t} w^*_{jt} \mathbb{1}(x_i \in R_{jt})$
end for
Output: Final model $\hat{y}_i = \hat{y}_i^{(M)}$.

On the other hand, machine learning is more concerned with the practical side of things, aiming to develop models that can generalize to new data. In contrast, econometrics frequently begins with a model specification able to represent the true data-generating process. After describing the model, its parameters must be estimated using an estimation technique.

Linear regression is frequently employed in econometrics to investigate causal relationships and estimate the coefficients of a model that assumes a linear relationship between a dependent variable and one or more independent variables. This method often uses the Ordinary Least Squares (OLS) technique, which is a straightforward approach to estimating these coefficients (See eq. 2.1). The econometric approach is distinct in its focus on causality and parameter estimation, typically without involving a train-test split, which differentiates it from applications in machine learning. In contrast, when used in machine learning, linear regression is often aimed at predictive accuracy, utilizing techniques such as gradient descent, and commonly includes a train-test split to assess model performance.

$$\begin{aligned}
RSS(\boldsymbol{\beta}) &= \boldsymbol{\varepsilon}^\top \boldsymbol{\varepsilon} \\
&= (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \\
&= \mathbf{Y}^\top \mathbf{Y} - \boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{Y} - \mathbf{Y}^\top \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{X}\boldsymbol{\beta} \to \min_{\beta_0, \beta_1} \\
\Rightarrow \frac{\partial RSS(\widehat{\boldsymbol{\beta}})}{\partial \widehat{\boldsymbol{\beta}}} &= -2\mathbf{X}^\top \mathbf{Y} + 2\mathbf{X}^\top \mathbf{X}\widehat{\boldsymbol{\beta}} = 0 \\
\Rightarrow \widehat{\boldsymbol{\beta}} &= \left(\mathbf{X}^\top \mathbf{X}\right)^{-1} \mathbf{X}^\top \mathbf{Y}
\end{aligned} \tag{2.1}$$

OLS is considered BLUE (Best Linear Unbiased Estimator) under Gauss-Markov theorem if certain assumptions are fulfilled, such as linearity, strict exogeneity, full rank (absence of perfect multi-collinearity) and spherical errors. It produces the smallest variance estimate among other unbiased, linear estimators. As a result, once the model and dataset are defined, there is only one possible solution and combination of estimated coefficients that maximize the likelihood of the given data. On the other hand, building a machine learning model typically begins with selecting an appropriate algorithm for the task. During the training phase, the algorithm is "trained" on a dataset, and the goal is to find the set of parameters that minimize a specific loss function, resulting in the most accurate predictions from the algorithm. The machine learning algorithm's solution is approximate, as similar performance can be achieved with different hyper-parameter configurations. However, it is crucial to note that the system's effectiveness is highly dependent on the specific characteristics of the training dataset, including its size, quality, and representativeness. This dependence underscores the importance of careful dataset selection and preprocessing, as variations in the training data can significantly impact the model's performance and generalizability.

In essence, econometrics seeks a one-of-a-kind solution that maximizes the likelihood of data given a model and a set of assumptions. On the other hand, machine learning focuses on identifying the best approximate solution that matches the data by iteratively modifying the algorithm's hyper-parameters. Econometric models, as a form of statistical models, aim to analyze and forecast economic interactions. They are built on the assumption that the data follows a specific probability distribution, allowing for estimating causal relationships between variables and making predictions for policy decisions. To determine the parameters of these models, econometricians use statistical inference methods such as maximum likelihood estimation. These methods involve deriving the population from a sample of data and using the data's distribution assumptions. Asymptotic theory plays a significant role in examining how the estimator's characteristics change with increasing sample size. That includes the central limit theorem, which states that the distribution of the sample mean becomes closer to a normal distribution as sample size increases, and the law of large numbers, which states that the average of many independent random variables converges to their expected value.

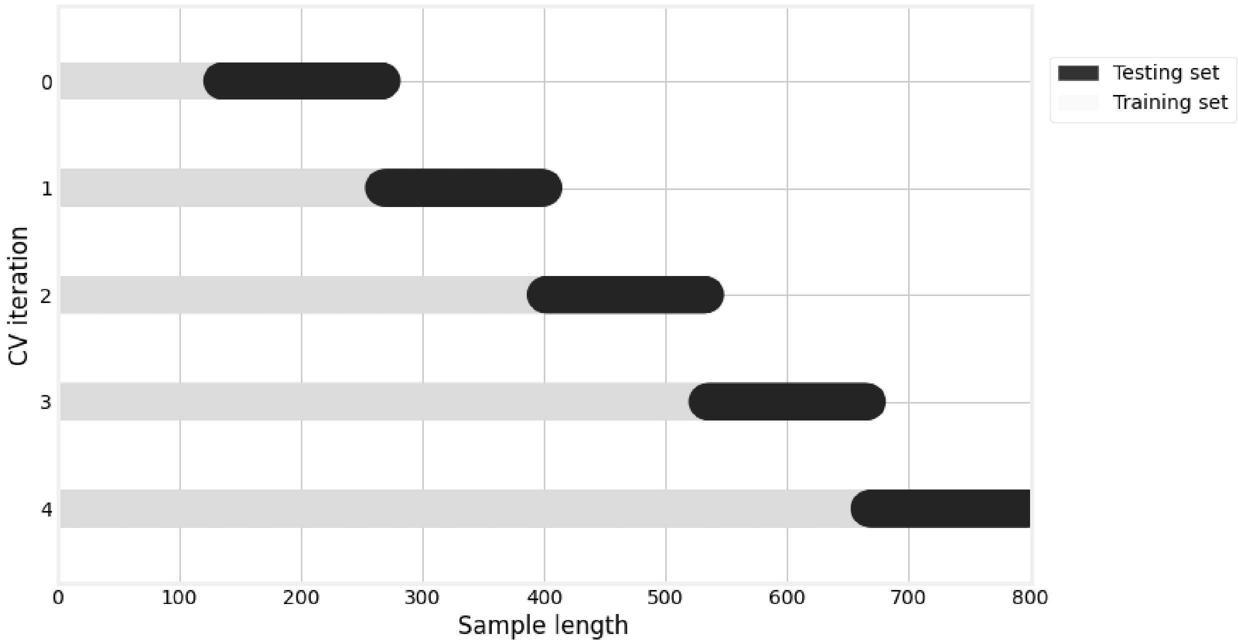
Machine learning algorithms are designed to learn from data without making any prior assumptions about the data's underlying probability distribution as for example Random Forest (see Algorithm 2). Instead of relying on classic statistical inference methods, these algorithms focus on finding the best set of parameters that minimize a specific loss function. This training phase aims to produce accurate predictions for new data points.

Algorithm 2 Random Forest for Classification (RFC)

```
1: for b ← 1, B do
2:     (a) Draw a bootstrap sample ℤ* of size N from the training data.
3:     (b) Grow a random forest tree T_b to the bootstrapped data, by recursively repeating the
       following steps for each terminal node of the tree, until the minimum node size n_min is reached.
4:         (I)   Select m variables at random from the p variables.
5:         (II)  Pick the best variable/split-point among the m.
6:         (III) Split the node into two daughter nodes.
7: end for
8: Output the ensemble of trees {T_b}_1^B
9: Make prediction at new point x:
10: Let Ĉ_b(x) be the class prediction be the class prediction of the bth random forest tree. Then
    Ĉ_r f^B(x) = majority vote {Ĉ_b(x)_1^B}.
```

Machine learning algorithms optimize the algorithm's hyper-parameters to determine the best set of parameters. These hyper-parameters, such as the tree depth in decision tree models or the number of hidden layers in neural networks, control how the algorithm behaves and makes predictions. The optimization process is achieved through cross-validation as can be seen in Figure 2.2 and grid search algorithms, which evaluate different combinations of hyper-parameters to find the set that results in the best model performance. In this way, machine learning algorithms can learn from data, make predictions, and adapt to new data points without making assumptions about the data's underlying probability distribution.

Figure 2.2: Graphical example of the five-split time series cross-validation



In simple terms, econometric and machine learning models serve the same purpose of analyzing relationships and making predictions, but they go about it differently. Econometric models use well-established statistical and mathematical theories to draw conclusions about economics. In contrast, machine learning models concentrate on learning from data to make predictions and use meta-parameter optimization to estimate parameters in their models. Nevertheless, this distinction is becoming weaker as time passes. The remainder of the paper provides a non-exhaustive literature review of the most important publications on machine learning in economics.

2.2 Machine learning in Economic research

2.2.1 Forecasting and prediction with Machine learning

Supervised machine learning has become a popular choice for many applications, particularly in forecasting, where it has shown impressive results. That is due to its ability to adjust its parameters, called hyperparameter tuning, resulting in more accurate predictions. Additionally, techniques like train-validation-test split and advanced optimization algorithms like Adam (See Algorithm 3) have greatly enhanced the performance of these models by making the optimization of parameters more effective.

Algorithm 3 Algorithm: Generalized Adam

```
1: Initialize m_0 = 0 and x_1
2: for t = 1, ..., T do
3:     m_t = β_{1,t} m_{t-1} + (1 − β_{1,t}) g_t
4:     v̂_t = h_t (g_1, g_2, ..., g_t)
5:     x_{t+1} = x_t − α_t m_t / √v̂_t
6: end for
```

Machine learning models have proven to be better and more efficient than traditional statistical methods that often rely on unrealistic assumptions. Unlike these methods, Machine learning models learn from the data at hand and can identify patterns and traits that help increase the accuracy of forecasts. Moreover, the capability of Machine learning models to process large amounts of data means that more historical data can be incorporated into the forecasting process, making it easier for the models to generalize and predict future events with accuracy. Hence, this led to an ever-increasing number of publications.

The papers Kleinberg et al. (2015), and Kleinberg et al. (2018) demonstrate how machine learning can improve predictions and policy decisions. These methods can analyze large datasets, detect patterns, and make accurate predictions. Machine learning can forecast unemployment duration in labor market policy, assisting workers in saving and job search strategies.

Similarly, Chalfin et al. (2016) highlighted the potential social welfare gains that can be realized using machine learning to improve worker productivity estimates. They provided specific examples of this method's effectiveness in two critical applications, namely police employment decisions and teacher tenure decisions. The authors used real-world data from police departments and schools to show how accurate predictions of worker productivity can lead to better hiring decisions, which increases productivity and lowers costs. They show how machine learning can identify candidate characteristics most predictive of future productivity, allowing businesses to make more informed hiring decisions.

Furthermore, their research shows that the approach can be used to determine tenure for instructors, providing evidence of the value a teacher brings to their students. Hence, this can lead to more efficient resource use and equitable resource allocation among teachers.

Chalfin et al. (2016) highlights the potential of machine learning to improve forecasts of worker productivity. Hence, it shows clear benefits it may offer society by improving key decisions such as recruiting and tenure decisions in critical public services such as law enforcement and education. This method is valuable not only for cost savings but also for improving the performance and efficiency of these businesses.

Kapetanios et al. (2017) investigates the various types of big data and provides an overview of several illustrative studies in big data and macroeconomic forecasting. Furthermore, they thoroughly examine several important data approaches currently used in the literature. One of the primary focuses of their research is the use of sparse regression methods, which are used to find and select only the most relevant predictors from a large set of potential predictors. These strategies help to improve the model's interpretability and computational efficiency, which can be very useful when working with large data sets. Another critical aspect of the research is the heuristic optimization of information criteria, which determines the best number of factors or predictors to include in the model. Working with large amounts of data necessitates striking a balance between model simplicity and explanatory power. These strategies aid in achieving this balance.

According to Richardson et al. (2021), machine learning methods outperform a simple autoregressive benchmark and a dynamic factor model. They show that applying different Machine learning approaches to economic and financial time series data produces far more accurate projections than benchmark models, demonstrating the machine learning potential to enhance official forecasts.

The results of Medeiros et al. (2021) show that machine learning models with many covariates (predictor variables) can produce systematically more accurate predictions than traditional benchmarks. The authors discover that of the Machine learning models tested, the random forest model consistently outperformed the others. The authors attribute the random forest model's exceptional performance to its unique variable selection method, which enables the identification and use of the most informative predictor variables. According to the authors, another reason for the random forest's superior performance is its ability to discover and capture nonlinear relationships between past key macroeconomic variables and inflation.

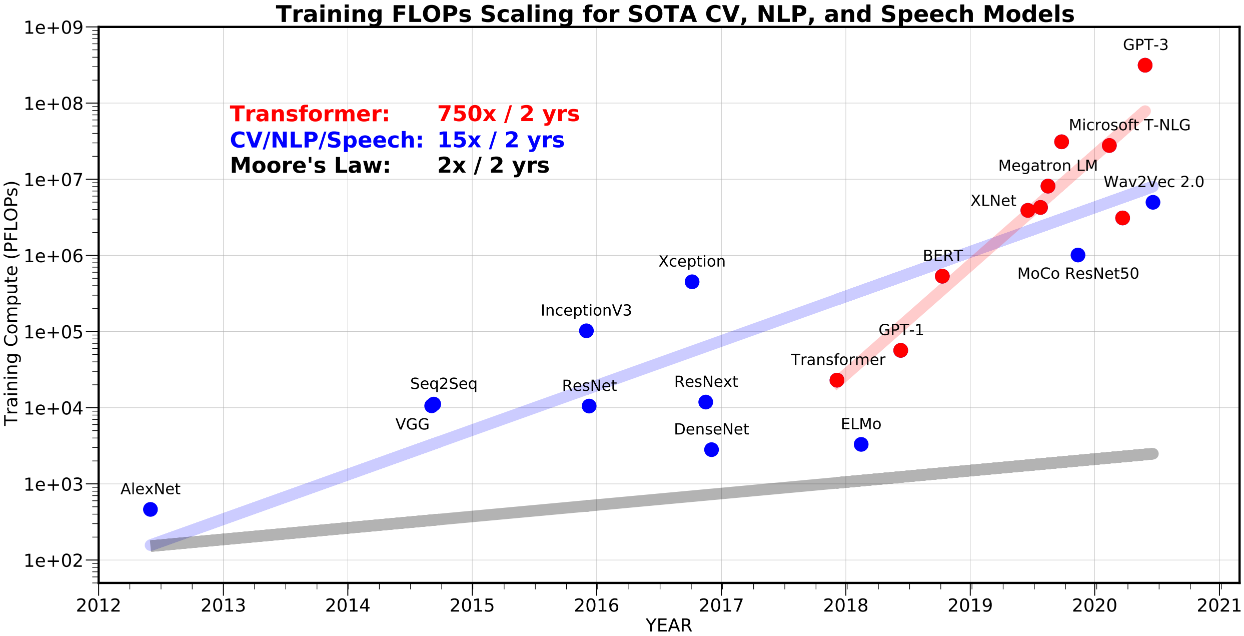
Goulet Coulombe et al. (2022) demonstrates the effectiveness of Machine learning techniques in improving forecasting performance during periods of high macroeconomic uncertainty, financial stress, and housing bubble bursts. According to the study, the ability of Machine learning models to capture nonlinear relationships and patterns in data that traditional linear models may miss makes them a valuable tool for macroeconomic forecasting in uncertain times.

Smyl (2020) presents the winning methodology from the M4 forecasting competition, which integrates exponential smoothing (ES) models with long short-term memory (LSTM) neural networks into a unified hybrid framework for time series forecasting. The hybrid approach leverages the strengths of both statistical and machine learning models by combining ES for capturing local series components like seasonality and level, and LSTM networks for modeling non-linear trends and cross-learning across series. This innovative combination results in improved forecasting accuracy, outperforming traditional statistical and machine learning methods individually.

2.2.2 Sentiment analysis and Natural Language Processing

The growing availability of specialised repositories containing labeled training datasets is likely the cause of the recent interest in applying sentiment analysis to economics. With the assistance of sentiment datasets, researchers have developed classifiers that can be applied to their datasets. Labels like "positive" and "negative" and numerical scales are included in these datasets. Consequently, new economics research has been developed using sentiment analysis, which has, in practice, led to the discovery of new insights and comprehension in various fields. The decade of the 2010s saw significant advances in natural language processing (NLP) and language modeling, which has also been a significant driver of this trend. Because of developments in processing power and advanced NLP techniques (see Figure 2.3, researchers can now examine and analyze massive amounts of textual data in ways that were impossible in the past.

Figure 2.3: Compute power for training



Notes: The amount of compute, measured in Peta FLOPs, needed to train SOTA models, for different CV, NLP, and Speech models, along with the different scaling of Transformer models (750x/2yrs); as well as the scaling of all of the models combined (15x/2yrs) (Gholami et al., 2021)

Hence, this has led to the publication of many new papers in reputable journals, such as the ones written by Algaba et al. (2020) and Gentzkow et al. (2019), respectively. These papers provide an overview of semantic analytic methods used in economic science. These methods use textual, audio, and visual data. Additionally, these papers provide an overview of appropriate statistical methods and a variety of applications to economic problems.

Applying natural language processing techniques is a fruitful way to approach the study of economics. In order to gain insights and improve predictions in their respective fields, researchers have utilized the techniques in various domains, ranging from monetary economics to corporate finance, business economics to labor market policy.

Aruoba and Drechsel (2022) use natural language processing to extract information from the Federal Open Market Committee meeting papers prepared by Fed economists. Academics can learn from these documents how the committee arrived at the policy conclusions based on the information given.

Based on this information, Aruoba and Drechsel (2022) can make an educated guess as to the adjustments that ought to be made to the target interest rate, thanks to machine learning. The residual of the prediction can be used to measure shocks' influences on monetary policy. They make predictions regarding the outcomes of policies and meetings using Machine learning and Natural Language Processing, respectively. The monetary policy shock can be understood as the deviation from what was expected regarding the outcomes instead of what took place.

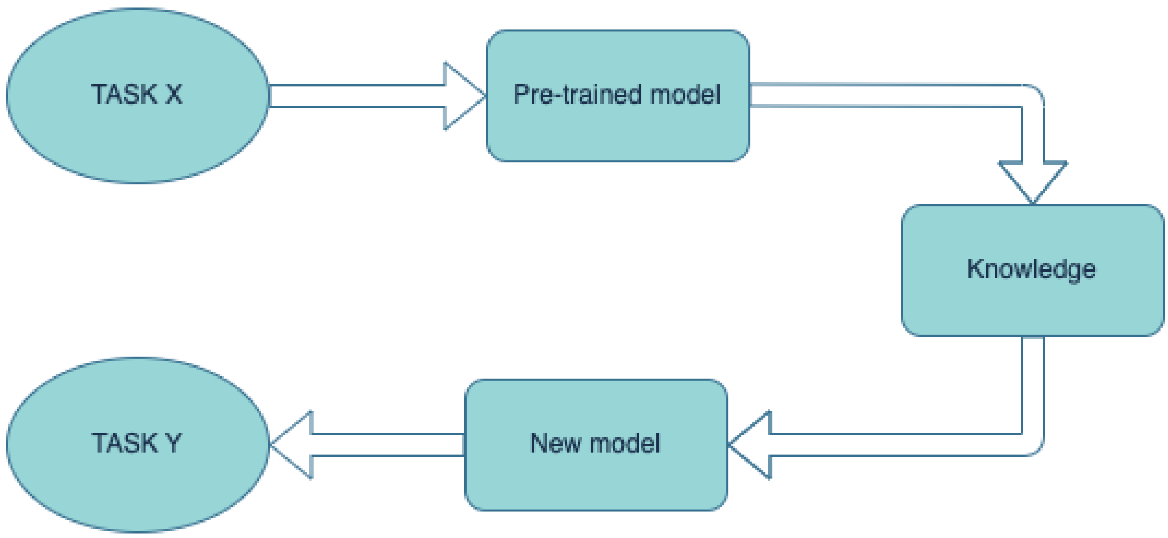
An understanding of monetary policy can be gained by analyzing text-based communication in meetings. In addition, it demonstrates how the Federal Reserve Board uses text information and how this usage influences monetary policy. Another example of notable work in the field of monetary economics is the use of computational linguistics to learn more about the communication patterns of central bankers. Text-based methodologies have been used by Hansen et al. (2018) and Hansen and McMahon (2016), who quantified text data to measure the impact of central bankers' utterances and create rich communication metrics. Hence, this enables a more in-depth comprehension of how central banks communicate with the general public and the markets and how their remarks may influence the outcomes of economic situations. In addition, this makes it possible to understand how their remarks may influence the outcomes of economic situations. Furthermore, strategies derived from the field of NLP have been implemented in other sectors of the economy, such as corporate finance, with notable success. Hoberg and Maksimovic (2015) improved the measurement of financial limitations by employing a text-based measure, which outperforms previous measures used in the literature in predicting investment cutbacks following negative economic shocks. This research highlights the potential for natural language processing techniques to improve the accuracy of established statistical approaches in various fields. The study of business economics is one of the disciplines that stands to gain from implementing NLP strategies such as those developed by Bandiera et al. (2020), which developed a CEO behavior index through the application of Latent Dirichlet Allocation (LDA). Using Natural Language Processing in economic forecasting can provide numerous benefits, such as improved accuracy and efficiency, data collection and analysis automation, and the ability to handle vast amounts of unstructured data. Natural Language Processing can also help overcome limitations in traditional forecasting methods by incorporating various sources of information and providing real-time insights. With its ability to process and analyze large amounts of text data, NLP can significantly enhance the accuracy of economic forecasting and decision-making.

2.2.3 Deep learning in Economics

In recent years, the field of image processing and graphical data processing has seen a substantial increase in the usage of various Convolutional Neural Networks (CNNs) to analyze images and extract useful information. This trend is expected to continue. As a result of their capacity to promptly and effectively analyze large amounts of data, these networks play an essential role in the research of big data. Hence, resulted in a meteoric rise in the number of repositories freely available on the internet for researchers to use. These repositories contain already preprocessed data and models that were trained and ready to be used. *Transfer learning* as seen in Figure 2.4 is a technique that enables researchers to economize both their time and their resources by capitalizing on the labor that other individuals have already performed. This method uses either pre-trained models or preprocessed data in the analysis process.

One field in which CNNs have proven to be quite helpful is examining satellite data. Donaldson and Storeygard (2016) investigate works that use satellite data and provide an extensive literature

Figure 2.4: Transfer learning



Notes: Graphical representation of transfer learning process. Elaborated by Author

review. They provide the following definition of Satellite data: *Information that has been remotely sensed and gathered from orbit by satellites on a smaller scale.* These satellites orbit the Earth at a great distance and take photographs in addition to collecting other data that can be used to investigate various processes taking place on the planet's surface. For instance, Henderson et al. (2012) utilized satellite data to measure economic activity, specifically the increase in GDP, at sub-national and supranational locations. Hence, this is something that cannot be accomplished using any of the other methods available right now.

One more application of CNNs that can be discovered in image processing is the inspection of photographs taken from ground level. Naik et al. (2016) made use of 360-degree panorama pictures of streetscapes to speculate on the factors that led to the physical transformation of five cities in the United States. They used a Support Vector Machine (SVM) to quantify the urban look by using street-level photographs to investigate the relationship between the outward appearance of a city and the behavior and health of its residents to learn more about the relationship between the two. They found a strong correlation between a city's physical appearance and its inhabitants' well-being, and that this correlation could be used to predict the determinants of urban change. In addition, they also found that this correlation can be used to make accurate predictions regarding the health of the people who live in a city. Naik et al. (2017) further expanded on this research and found that researchers can gain a deeper understanding of the relationship between urban environments and the well-being of citizens by quantifying the urban appearance from street-level images.

The application of convolutional neural networks and the use of large amounts of data in processing images and graphical data has, in general, led to a better understanding of the various processes that are taking place on the surface of the Earth. These processes can range from measuring economic activity to comprehending the connection between urban environments and the health and happiness of those living in those environments. As a result of the ongoing advancements in technology and the growing availability of data, these fields will continue to improve and bring new insights into a wide range of phenomena. Hence, this is one reason why collecting as much information as possible is vital.

Deep learning methods are not solely restricted to image analysis. They can also be applied to traditional macroeconomic models such as dynamic stochastic general equilibrium (DSGE).

In their research, Fernandez-Villaverde et al. (2020) offer a deep learning algorithm that, in contrast to earlier methods, does not rely on integral approximations when computing a global solution to a set of problems. Because of this, the system can function at a faster rate. They propose an alternative strategy that, in place of the methods that came before it, efficiently calculates accurate derivatives, which is a significant benefit in comparison to the methods that were traditionally used. To demonstrate the power of the proposed methodology, they use both a multi-location model with 50 continuous state variables and a highly nonlinear migration model with 75 continuous state variables. Both models have a total of 100 continuous state variables. The authors begin by evaluating their methodology through a conventional neoclassical growth model. Then, they demonstrate the power of their methodology through the use of empirical applications. (Fernández-Villaverde et al., 2019; Kahou et al., 2021; Cheela et al., 2022)

One another notable work in Deep learning for economics is by Zheng et al. (2022). In this paper, the authors proposed the concept of an "AI Economist", a policy design framework that utilizes deep reinforcement learning (RL). The framework consists of two levels of agents, where agents and a social planner co-adapt to each other. The AI Economist utilizes structured curriculum learning to address the complexity and difficulty of two-level, co-adaptive learning. The proposed framework is validated in the field of taxation, where it recovered the optimal tax policy as predicted by economic theory in one-step economies. Hence, this was an important step in validating the framework, as it demonstrated the potential of the AI Economist to provide insights into policy design and help improve the efficiency of economic systems. The authors also noted that this framework could be expanded to other fields, such as labor markets, financial markets, and environmental economics. It could be used to analyze and optimize various economic policy issues.

We can now use previously unused alternative data such as images, audio, and video with deep learning. We can enhance the estimation efficiency of standard macroeconomic tools such as Dynamic Stochastic General Equilibrium models. Additionally, we can also simulate and understand economies using Deep Reinforcement learning. As data and computing power expand, the significance of using deep learning in economics will only grow.

2.2.4 Causal Machine Learning

Conventional economic models need help to keep up with the influx of big data. With so many variables and interactions, these traditional models, which typically rely on a limited number of observations, are finding it challenging to manage today's data's sheer volume and complexity. Hence, this presents a significant challenge for the field. To overcome this, a new study area, known as causal machine learning, has emerged. It combines the power of machine learning with the concept of causality to create models that can handle the complexities of modern data while still providing valuable insights. It is a relatively new field of research, with its earliest origins dating back to 2014 and significant advancements made in recent years, especially in 2018. This research aims to create models that can successfully navigate the complexities of big data and deliver precise insights.

One of the first papers is from 2014, Belloni et al. (2014) presented an overview of how the most recent developments in the field of "data mining" can be adapted and modified to improve the inference of model parameters . They concentrated their efforts on locating methods with "true" predictive power, which can protect against false discovery and overfitting. The authors argue that it is essential to distinguish between an in-sample fit and out-of-sample predictive ability and that techniques should accurately account for using the same data to examine a variety of hypotheses or models.

Following with work by Farrell (2015) presenting a method for constructing confidence intervals using a doubly robust estimator. This approach is designed to be robust to model-selection errors, which can significantly challenge causal machine learning. He proves that the intervals have uniform validity over a large class of models, which significantly contributes to the field. Hence, this makes it possible to examine multivalued treatments with heterogeneous effects and selection among more covariates than observations.

In the paper Belloni et al. (2017) proposes efficient estimators and honest confidence bands for various treatment effects in data-rich environments. These treatment effects include local average treatment effects (LATE) and local quantile treatment effects (LQTE). They presented an innovative method for estimating treatment effects in settings that leveraged recent advances in machine learning and causal inference. It allowed them to account for the effects of various treatments more accurately. The authors presented a comprehensive theoretical analysis of their methods. They demonstrated the empirical performance of their methods through a series of simulations and examples taken from the real world.

Hartford et al. (2017) propose a method for using deep learning techniques to analyze the relationship between variables when there are confounding factors, known as instrument variables (IVs). The proposed approach involves two separate prediction tasks, which can be solved using deep neural networks. The first network predicts the treatment, and the second uses this information to estimate the causal effect by incorporating it into the loss function. The authors demonstrate that this "Deep IV" framework improves machine learning methods.

Chernozhukov et al. (2018) published a ground-breaking paper on Double/Debiased Machine Learning for treatment and structural parameters. This paper paved the way for future research in the field. They proposed an algorithm to obtain the non-asymptotic debiased machine learning theorem. As long as a limited number of fundamental interpretable constraints are satisfied, this theorem can be applied to any global or local function of any machine learning algorithm. Consequently, a straightforward set of requirements can be put into practice to convert rates derived from contemporary learning theory into more conventional forms of statistical inference. Hence, this is especially helpful in complex environments because it provides a straightforward and generic method for estimating and inferring the low-dimensional parameters of interest. It is beneficial in situations where there are many variables to consider. It is generally agreed that this paper represents one of the most significant leaps forward in the field of causal inference in recent years.

One of the most recent papers in the field is by Athey et al. (2019), proposing a new method for non-parametric statistical estimation called "Generalized Random Forests." This method uses an adaptive weighting function derived from the forests to overcome the limitations of classical kernel weighting functions. The method is flexible and efficient and produces consistent and asymptotically Gaussian estimates with valid confidence intervals.

The field of causal machine learning is still in its early stages, but it has already shown significant promise in providing precise insights from big data. As more research is conducted and more advancements are made, the models produced by causal machine learning are likely to become increasingly sophisticated and practical. Hence, this presents a promising future for the field and its potential to impact the world of economics significantly and beyond.

2.3 Conclusions

Machine learning is becoming an increasingly important tool for economists as it allows them to analyze large and complex datasets, discover hidden patterns and insights, and make data-driven decisions that are informed and accurate. Applying Natural Language Processing techniques, causal inference in econometrics, extensive data analysis, and Machine Learning in macroeconomics are all active research areas that can revolutionize the field of economics. These proposed methodologies and procedures have a significant potential for enhancing both the researchers' capacity to make data-driven decisions that are informed and accurate, as well as the accuracy and efficiency of the analyses themselves. Machine learning techniques such as deep learning, reinforcement learning, and neural networks can enhance economists' ability to extract valuable insights from large and complex datasets, leading to discoveries and improved decision-making. Furthermore, econometrics has also significantly shifted toward using machine learning for causal inference. Researchers have proposed new methods and techniques, such as sparse regressions, heuristic optimization of information criteria, factor methods, and textual data methods specifically designed to make the most of big data.

In recent years, researchers have also proposed efficient estimators and honest confidence bands in data-rich environments for various treatment effects, such as local average treatment effects (LATE) and quantile treatment effects (LQTE). These methods have been specifically designed to handle the challenges of big data and provide more accurate and reliable results.

Moreover, Chernozhukov et al. (2018) published their groundbreaking Double/Debiased machine learning algorithm, which provides a non-asymptotic debiased machine learning theorem that can be applied to any global or local function of any machine learning algorithm as long as it satisfies a few primary, interpretable constraints. This algorithm is a significant advancement in the field of machine learning in macroeconomics and causal inference, as it enables researchers to make data-driven decisions that are informed, accurate, and efficient.

By leveraging advanced algorithms and computational power, economists can now try to tackle previously intractable problems, make more accurate predictions, and derive deeper understanding from their data. This technological advancement not only enhances the robustness of economic models but also opens new avenues for research and policy-making, ultimately leading to more informed and effective economic decisions.

Chapter 3

Augmented Coincident Index: Using soft information to track the economy

Abstract

The transition from a manufacturing-centric economy to one dominated by services and digital technology is reshaping our economic landscape. This evolution has put forth significant challenges to the established norms of macroeconomic theory and conventional econometric methodologies. In this dynamic environment, the need for real-time data has become paramount for policymakers and forecasters. However, the reporting of GDP growth, arguably the most critical economic statistic, is typically delayed by one to three months. To mitigate these limitations, we propose the incorporation of 'soft information' data, gleaned from social media platforms and leading Italian newspapers, to provide a real-time pulse of the economy. This approach can potentially empower decision-makers to respond swiftly to economic trends and shifts. Our preliminary findings suggest that this approach enhances the accuracy of predictions compared to traditional coincident indices.

Keywords: Real-time data, Economic modeling, Coincident index, Economic activity, GDP, Forecasting, Natural language processing.

3.1 Introduction

The shift from a manufacturing-based economy to a service-oriented one, coupled with the advent of digital technology, has brought about profound changes in our economic structure. These transformations have posed significant challenges to the foundational principles of macroeconomic theory and traditional econometric methods. In this rapidly evolving context, real-time data has become a crucial resource for policymakers and forecasters.

However, the most vital economic indicator, GDP growth, is not immediately available. Typically, GDP is recorded quarterly and reported with a delay of one to three months. This lag can hinder policymakers and analysts from accurately and promptly understanding the economy. To compensate for this, they often rely on immediate indicators related to financial and labor markets and coincident indexes to gauge the current economic state.

The creation of coincident indexes, designed to provide a real-time measure of economic activity, is often limited by the availability of high-frequency data. Developing coincident indexes that not only indicate the direction of the economy but also explain why it is trending upwards or downwards can be challenging.

To address this issue, we propose using text analysis methods to gain a real-time understanding of the economy. Text data, such as news articles, central bank communications, and social media posts, can provide valuable insights into economic activity and sentiment. This data can supplement traditional economic data, often released with a delay, and help capture the complexity of the economy. Techniques such as natural language processing, sentiment analysis, and topic modeling can be used to identify patterns and trends in the data that can inform economic models and forecasts.

Text data offers a more nuanced understanding of the economy compared to traditional economic data. While traditional data like GDP and employment figures can measure broad trends, they may not capture the intricacies of economic activity. Text data, on the other hand, can provide a granular view of the economy and help identify factors driving economic trends. By analyzing the language used in news articles, central bank communications, and social media posts, researchers can identify sentiments and shifts that may indicate changes in economic activity. These shifts in sentiment can serve as leading indicators of future economic outcomes.

Several economic research studies have used text analysis. For instance, Gao and Beling (2003) found that their direct scoring algorithm (DSA) model using the Beige book significantly contributes to the prediction of GDP growth. Dvorák and Novák (2007) discussed using fuzzy logic and evaluative linguistic expressions to automatically model economic texts. Hansen and McMahon (2016) studied the macroeconomic effects of central bank communications, using Latent Dirichlet Allocation to measure the effect of transparency within the Federal Open Market Committee (FOMC). Gentzkow and Kelly (2017) provided an overview of the use of text data in economic research, including appropriate statistical methods and applications. Song and Shin (2019) found that news articles can be a valuable source for generating economic indicators through text mining techniques that extract sentiment information from online economic news articles. Kalamara et al. (2020) found that newspaper text can improve economic forecasts using a new method that combines counts of terms with supervised machine learning techniques, particularly during stressed periods. A summary of relevant methodological approaches can be found in the work of Algaba et al. (2020).

Despite their usefulness, text analysis methods are not without their flaws. One of the significant limitations of text analysis methods is their vulnerability to researcher bias, which can skew the results and undermine the validity of the findings. To address these limitations, we propose using transformers, a state-of-the-art class of models that has recently gained popularity in natural language processing (NLP). These models are designed to analyze language in a more human-like manner. They can reduce bias in the analysis by capturing the nuances and complexities of human language. Transformers are deep learning models trained on massive amounts of data and can capture patterns in language to generate predictions. They have been applied in various NLP tasks, such as sentiment analysis, language translation, and question-answering, and have shown remarkable performance, outperforming other models in several benchmark tests.

Hence, we propose using transformer models to develop an economic index that can provide real-time insights into economic activity. The data for this index is obtained from newspapers and social media. The methodology section outlines the general approaches used for text analysis in economics and introduces the specific transformer model we will use. The data collection and inclusion process is described in the third subsection. The fourth subsection presents the economic index developed using the transformer model. The fifth evaluates the index, and the final one concludes the chapter.

3.2 Text analysis

Text analysis has emerged as an indispensable instrument for deciphering the enormous volume of information produced in our digital era. This procedure entails scrutinizing and categorizing extensive amounts of textual data from a variety of sources, including web pages, emails, scientific journals, e-books, educational materials, news outlets, and social media platforms. The primary objective of text analysis is to extract significant insights and relationships from these sources, thereby yielding valuable information. One of the key advantages of text analysis is its capacity to enable researchers to process large data volumes swiftly and efficiently, thereby facilitating data-driven decision-making.

A crucial component of text analysis is the data source. For instance, web pages offer a rich repository of information on a wide array of subjects, encompassing news, opinions, and trends. Emails can serve as a record of customer interactions, sales, and marketing campaigns. Scientific journals provide detailed insights into specific topics, while e-books can offer educational and entertainment content. Social media platforms, such as Twitter and Facebook, are potent sources of public sentiment and opinion.

Text data can also be scrutinized based on its attributes, such as sentiment, emotion, and mood. This analysis can yield valuable insights into customer attitudes, market trends, and public opinion. Text analysis is an invaluable tool for researchers seeking to glean insights from vast amounts of information. The potential data sources are extensive and can provide a wealth of information on a variety of subjects. Understanding the properties of text data, such as its source and format, is crucial for successful text analysis and for unlocking its full potential. There are numerous approaches to modeling this data, and we will delve into some of these in the subsequent sections.

3.2.1 Advancements in Text Analysis: From TF-IDF to Transformers

Over the years, text analysis methodologies have evolved significantly, transitioning from simple statistical measures to sophisticated deep learning models. This section explores the progression of these techniques, starting with traditional methods such as Term Frequency-Inverse Document Frequency (TF-IDF) and lexicon-based approaches, moving through various machine learning strategies, and culminating in the latest advancements involving word embeddings and transformer models. Each of these methodologies has contributed uniquely to the field, enhancing our ability to process and interpret large volumes of textual data with increasing accuracy and complexity.

Term frequency-inverse document frequency and Lexicon-Based methods

One of the most frequently utilized algorithms in text analysis is Term Frequency-Inverse Document Frequency. This statistical measure assesses the relevance of a word in a document within a larger collection, or corpus. The importance of a term increases with its frequency within a document, but this is offset by the frequency of the term across the entire corpus. In the context of a corpus, focusing on words that are frequently used within a single document is optional. TF-IDF, a measure that combines the works of Luhn (1958) and Jones (1972), is widely used due to its simplicity and effectiveness.

Over the years, various adaptations of the algorithm have been developed, incorporating strategies such as negation (Ramos, 2003; Das and Chakraborty, 2018), stylistic features, or synonyms (Yuntao et al., 2005). TF-IDF has also been used as a methodology for extracting sentiment and polarity indices from Twitter data (Addiga and Bagui, 2022), as well as for real-time monitoring of social phenomena (Alessa and Faezipour, 2018).

While TF-IDF methods provide a quantitative measure of the relevance of each word in a text, dictionary-based techniques require a predefined lexicon of words with positive or negative connotations. A notable work in the sentiment analysis of an Italian corpus using dictionaries is by Aprigliano et al. (2021), which constructs a tailored dictionary that also accounts for valence shifter terms.[1]

Machine Learning

Machine learning methodologies can be broadly classified into supervised and unsupervised categories. In supervised text categorization, once the classification categories have been defined, the system operates on a train-and-evaluate concept. The training process involves providing the algorithm with a set of labeled examples from a knowledgeable external supervisor (Sutton, 2018). During the testing phase, the algorithm is presented with unseen data to process, and it is responsible for determining the category to which the text belongs (Hassoun et al., 1995; Ng and Jordan, 2001; Steinwart and Christmann, 2008).

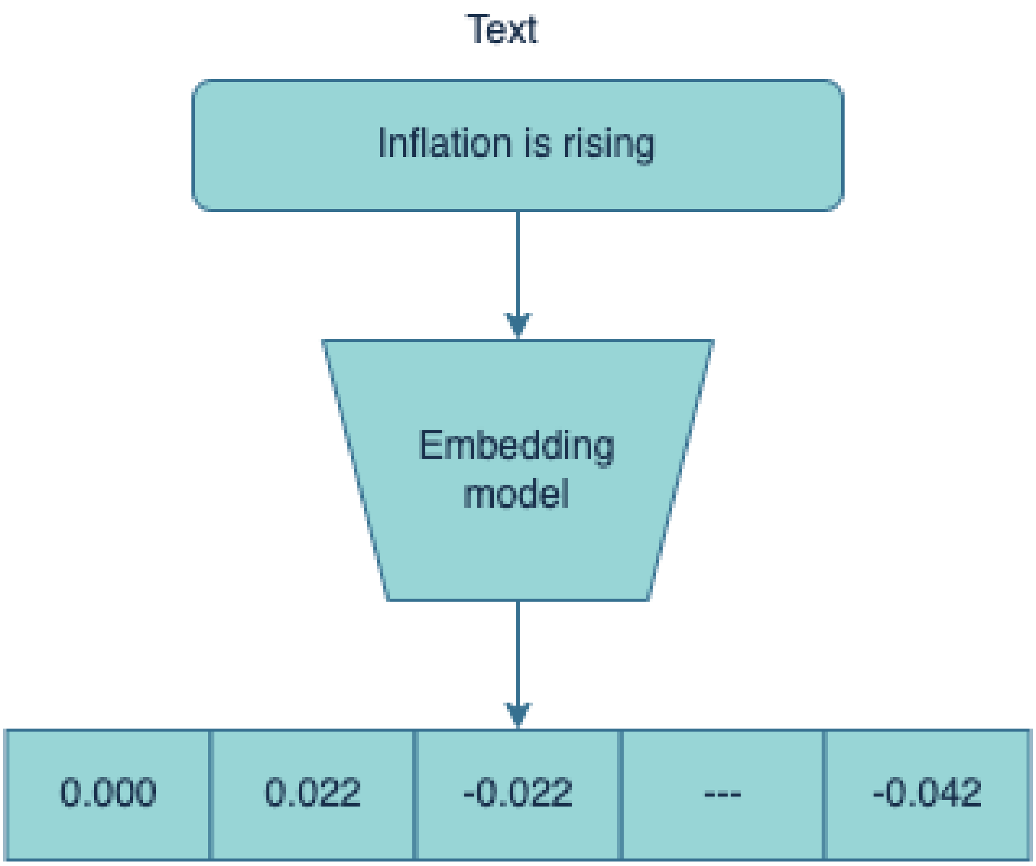
On the other hand, unsupervised classification is a process where the system uncovers hidden structures in a collection of unlabeled data without any external information. The natural division that the system identifies may not always coincide with human logic. Instead, the computer seeks patterns and structures in the data to form clusters. These clusters serve as the basis for categorizing the data. For example, during an internet search, the search term acts as the foundation for the algorithm to create clusters, which are then presented as the search results (Azqueta-Gavaldón, 2017; Azqueta-Gavaldon et al., 2020). In sentiment analysis, unsupervised techniques like Latent Dirichlet Allocation are used to cluster tweets and gain insights about inflation expectations (Angelico et al., 2022).

[1] Valence shifters are words that can modify the polarity of the meaning of other words. These modifiers can either intensify or weaken the semantic value of polarized words and include two main categories: negators and amplifiers. Negators reverse the polarity of the word they modify, while amplifiers reinforce the polarity of the word.

Word embedding

Word embedding is a potent tool in the field of Natural Language Processing (NLP) that converts words into numerical vectors within a high-dimensional space (See Figure 3.1). This transformation encapsulates the meaning and context of words, thereby facilitating their comprehension and utilization by computers for a variety of NLP tasks.

Figure 3.1: Embedding



Notes: Graphical representation of Embedding. Elaborated by Author.

When integrated with deep learning algorithms, word embedding enables the execution of a wide array of NLP tasks, such as language translation, text classification, and sentiment analysis. The combination of word embedding and deep learning algorithms has revolutionized NLP, leading to more precise and effective processing of natural language data. With word embedding, NLP systems can now manage complex language structures and meanings with increased accuracy, offering more human-like insights and solutions to real-world problems (Theil et al., 2018, 2020; Miranda-Belmonte et al., 2023).

Transformers

The self-attention mechanism, as proposed in Vaswani et al. (2017), enables the model to flexibly and efficiently capture relationships between different elements in an input sequence. This is achieved by computing a weighted sum of the input elements, where the weights are determined by the similarity between the elements. This mechanism allows the model to focus on different parts of the input sequence simultaneously, rather than processing the entire sequence in a fixed order.

The feed-forward neural network, a conventional type of neural network, comprises an input layer, one or more hidden layers, and an output layer. It processes input data through a series of linear and nonlinear transformations, enabling it to learn complex patterns in the data.

The combination of the self-attention mechanism and the feed-forward neural network allows the transformer architecture to process input sequences of any length and capture long-range dependencies in the data with only limit related to available computational resources.. This makes it highly suitable for various natural language processing tasks, including language translation, generation, and question-answering (Devlin et al., 2018; Baevski et al., 2020; Brown et al., 2020; Radford et al., 2019).

3.3 Data

In this study, we utilize newspaper articles as data input to gain real-time insights into Italy's economic activity. We have meticulously chosen the top five newspapers in the country based on their distribution.

We utilize data from Twitter, a platform widely used for sharing news and information. Each of the five selected newspapers uses Twitter to disseminate their essential daily news, often providing a summary of the article and a link to the online version. This allows us to access the data through

Twitter's Academic Access, which provides researchers with a large amount of data for analysis. Using this approach, we can preserve any crucial information that may be provided by the news stories while reducing the costs associated with data collection. This approach offers an alternative to traditional methods that rely on Factiva, a commercial news and business information database, which can be costly and time-consuming, especially for large-scale analyses.

Our analysis includes 2.5 million articles from 2012 to 2022. This large sample size allows us to comprehensively understand Italy's economic activity. The selected newspapers were chosen based on their distribution throughout the country, allowing us to capture a representative sample of news and information from different regions. By analyzing the language used in the articles, we can identify trends and patterns that can inform economic models and forecasts. The final data collection contains articles from the Italian newspapers: *Il Sole 24 ore, La Repubblica, Il Fatto Quotidiano, La Stampa, and Corriere della Sera.*. In Figure 3.2 we show simple wordcloud based on frequency of words for *Il Sole 24 Ore*. More wordcolud figures can be found in Appendix D.

Figure 3.2: Wordcloud - Il Sole 24 Ore

Notes: Frequency of words in the whole sample for Il Sole 24 Ore (bigger is more frequent).

After gathering all the information from the selected newspapers, one of our primary objectives is to organize the news articles into categories such as economics, politics, health, sport, and others.

We use a machine learning model known as transformers to do this effectively. These models are beneficial for natural language processing tasks, as they can capture the complex relationships between words and their meanings.

In particular, we align with the findings presented by Vaswani et al. (2017) on the effectiveness of transformers for natural language processing tasks. Figure 3.3 presents the model's specifications, as described in their work. Using transformers, we can accurately classify the news articles into their respective categories, allowing us to understand Italy's overall economic landscape better.

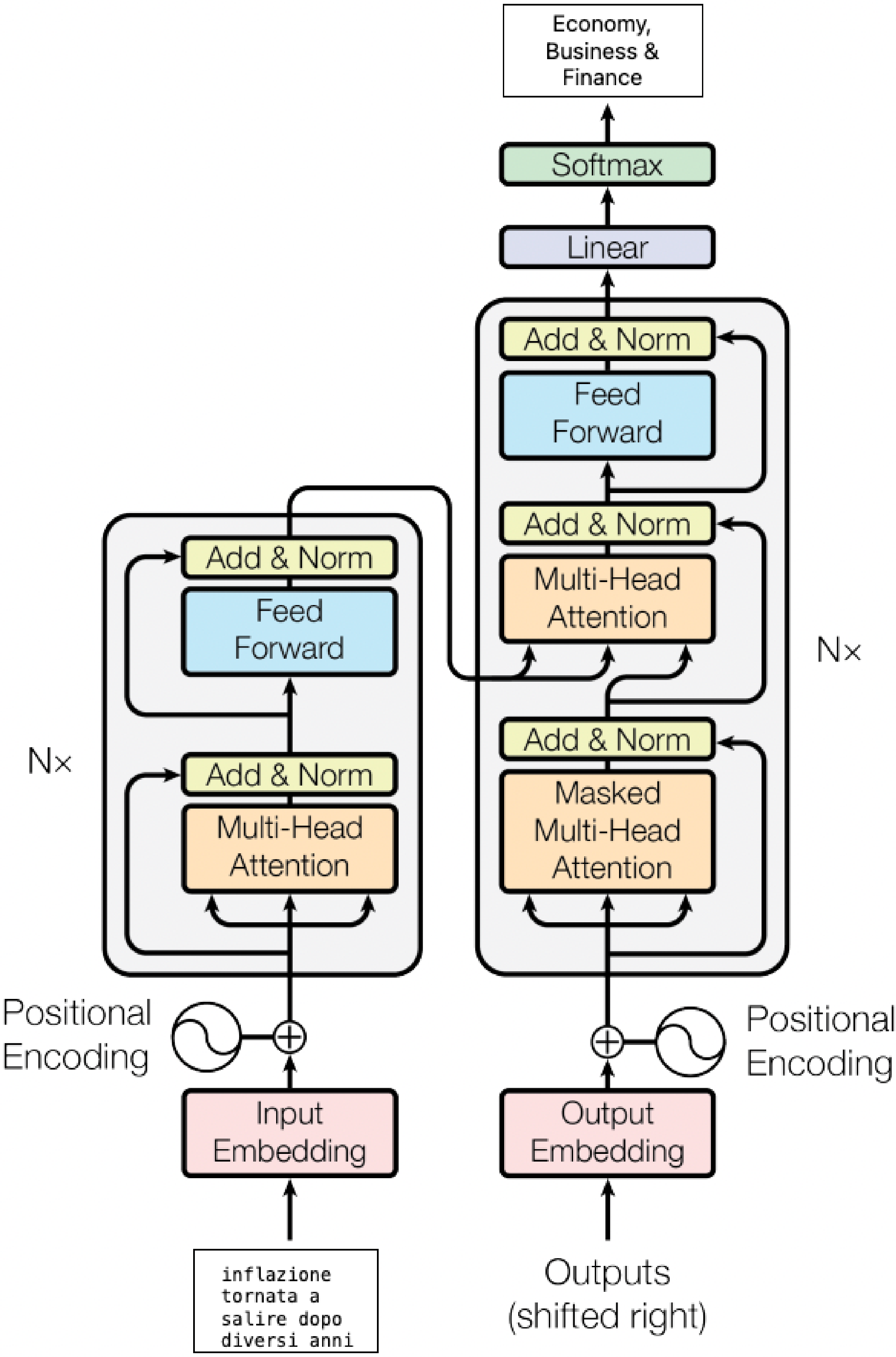
By organizing the news, we can more easily identify trends and patterns relevant to economic modeling and forecasting. Hence, this is particularly important as we seek to gain a real-time understanding of economic activity in the country, supplementing traditional economic data released with a lag and may only partially capture the economy's complexity.

Our analysis focuses exclusively on information regarding Italy's economy, businesses, and financial news. That is because we are interested in gaining a real-time understanding of economic activity and sentiment in the country. News articles related to these topics are particularly relevant for this purpose.

The following Figure 3.4 illustrates how the various newspapers' distribution changed during the period. That gives us a sense of the overall coverage of economic, business, and financial news in Italy over time and may provide insights into trends and patterns relevant to our analysis. By understanding news distribution across different newspapers and periods, we can better contextualize our findings and draw more accurate conclusions about economic activity in the country.

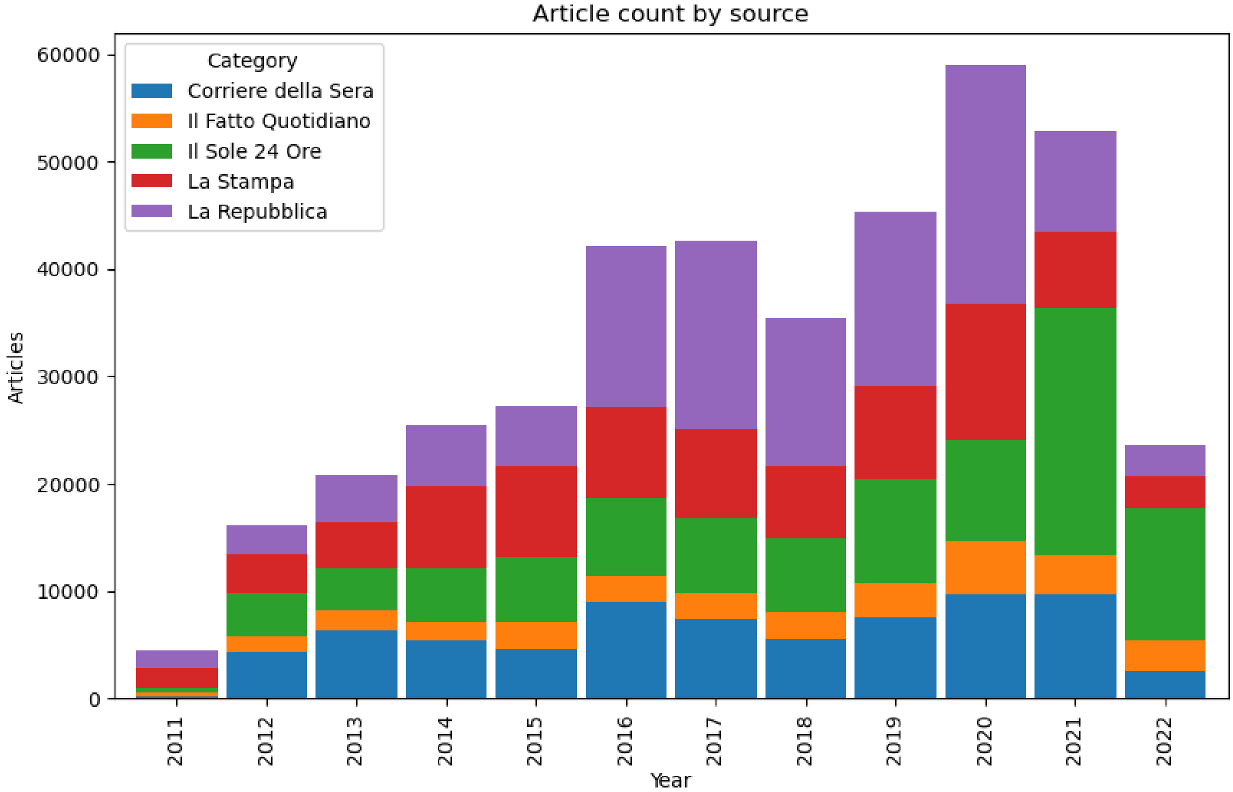
Table 3.1 shows the total number of articles obtained from Twitter for each of the five selected newspapers and the number of relevant articles included in constructing our real-time economic index. This index could be refined by including articles classified as political news, which may provide a more comprehensive view of the economic landscape. By expanding the scope of our analysis to include a broader range of sources, we can gain a more nuanced and accurate understanding of economic developments and trends. However, it is essential to carefully consider the potential biases

Figure 3.3: Transformers - model architecture by Vaswani et al. (2017)



Notes: Detailed explanation of the architecture can be found in Appendix B.

Figure 3.4: Economy, Business & Finance articles



and limitations of different sources and use a diverse range of data to obtain a more balanced and reliable view of the economy.

Table 3.1: Raw count of total articles and articles included

	Il Sole 24 Ore	La Repubblica	Il Fatto Quotidiano	La Stampa	Corriere della Sera
Total Articles	246392	941352	224995	537612	519178
Economics, business and finance	94721	118629	30098	81469	72231

3.4 Economic Index

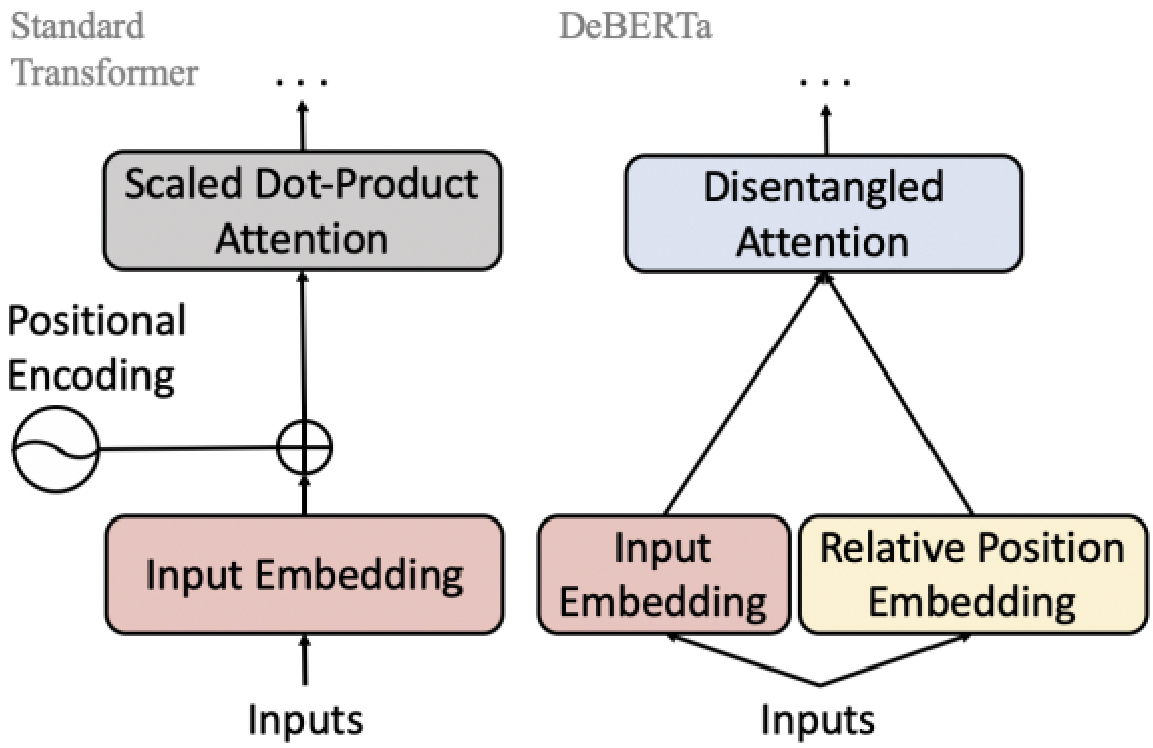
We have devised a method to construct a soft information index using a decoding-enhanced BERT model with disentangled attention (DeBERTa). DeBERTa represents each word in the input layer with two vectors, one encoding its content and the other encoding its position. Attention weights between words are calculated using separate matrices based on content and position rather than a combined vector, as in the BERT model as shown in Figure 3.5. The goal of this approach is to enhance the accuracy and timeliness of economic modeling and forecasting. Specifically, our method employs the Natural Language Inference technique, which is based on the research conducted by He et al. (2020). The proposed method consists of two main stages:

- Class hypotheses: In this stage, the model generates a set of hypotheses about the potential meanings of a given piece of text. These hypotheses are based on the words and phrases present in the text, as well as the context in which they appear. For instance, if the text contains phrases like "strong earnings" or "positive outlook," the model may generate hypotheses suggesting the text expresses a positive sentiment about the economy.
- Model target: In this stage, the model compares the generated hypotheses to a target concept or category and determines which hypothesis is most likely to be true. This is done using a disentangled attention process, which allows the model to focus on specific aspects of the text and analyze them in isolation. For example, the model may concentrate on specific words or phrases indicative of positive or negative sentiment, while ignoring other words or phrases that may be less relevant to the target concept.

Through these two stages, the developed model can generate a soft information index that captures the underlying meaning and sentiment of textual input. This index can play a crucial role in informing economic modeling and forecasting, providing a real-time snapshot of the economic landscape. This allows policymakers and forecasters to make informed decisions and respond proactively to changing circumstances.

The use of a decoding-enhanced BERT model with disentangled attention allows us to overcome many of the limitations inherent in traditional economic indicators. These limitations may include reliance on outdated or incomplete data, and the inability to capture the nuances and complexities of the economy. As a result, our approach can provide a more comprehensive and accurate representation of economic sentiment, which can enhance the effectiveness of economic modeling and forecasting. A small sample of sentiment calculated for individual sentences to showcase accuracy can be viewed in Appendix A.

Figure 3.5: Difference between DeBERTa and standard BERT Qian Qian et al. (2022)



Notes: DeBERTa represents each word in the input layer with two vectors, one encoding its content and the other encoding its position. Attention weights between words are calculated using separate matrices based on content and position rather than a combined vector, as in the BERT model. More details are in Appendix C.

3.4.1 Newspapers sentiment heterogeneity

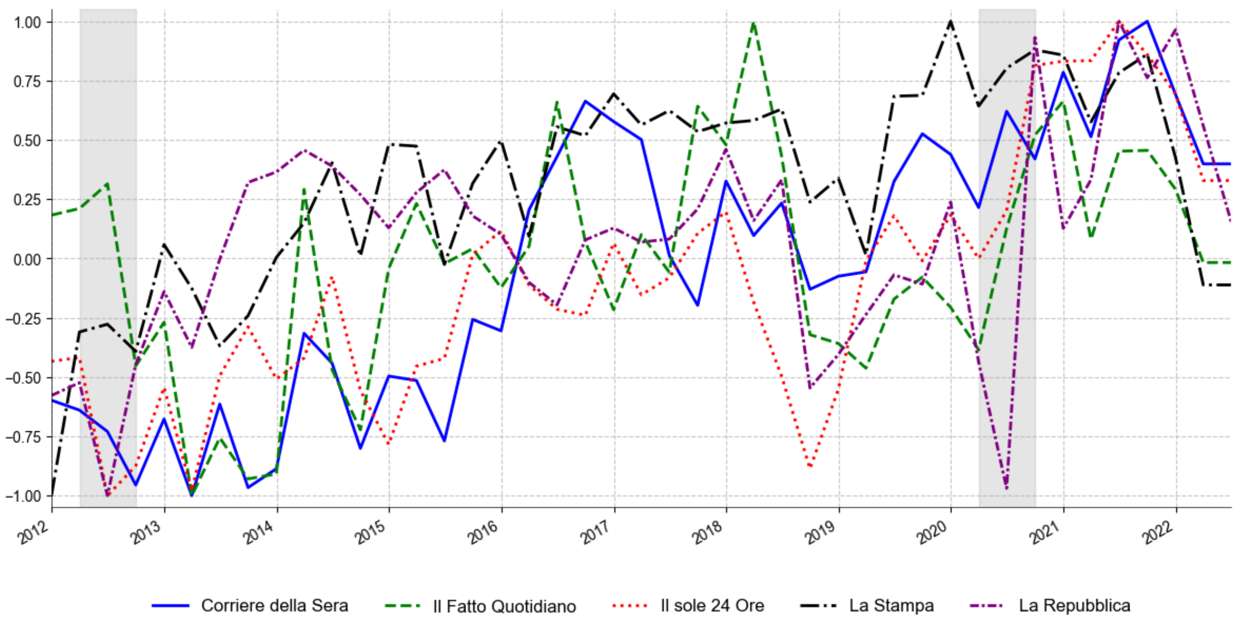
This segment introduces our examination of the sentiment expressed in articles published by Italy's top five newspapers. These articles serve as input for the creation of an economic index, which acts as a real-time indicator of economic sentiment.

In Figure 3.6, we offer a visual representation of individual indices for each of the five newspapers included in our study. These indices mirror the sentiment expressed in the articles from each newspaper, with higher values indicating a more positive sentiment and lower values indicating more negative sentiment. The graph also includes periods of recession, which are marked by shaded areas.

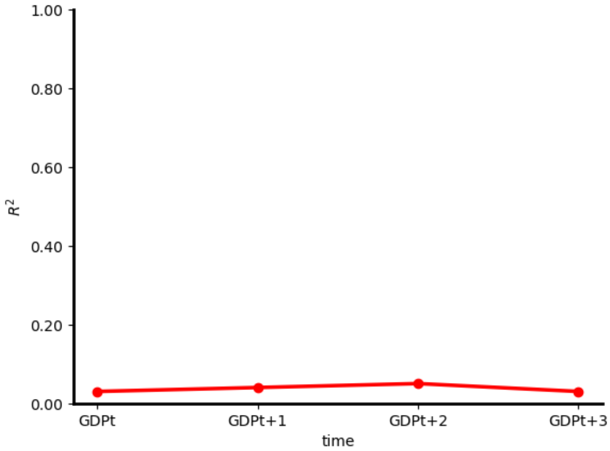
During the initial recessionary period, it's clear that all newspapers, except for La Stampa, predominantly conveyed negative sentiment. This observation suggests that these newspapers mirrored the economic downturn, reporting on negative developments and other unfavorable outcomes. In contrast, La Stampa exhibited a relatively stable trend, with a slight decline in sentiment levels in the middle of the period.

During the Covid-19 pandemic, all newspapers exhibited negative sentiment, although the magnitude varied. However, this could be due to the different number of articles included in the analysis for each newspaper, as well as the specific focus and coverage of the pandemic by each newspaper. In general, the trends of the newspapers are similar, with some variations in the magnitude of positive or negative sentiment during specific periods. These variations could be related to each newspaper's political perspective and how they present and interpret economic events and developments. For instance, a newspaper with a more conservative perspective might be more inclined to report on negative economic developments and express negative sentiment. Conversely, a newspaper with a more liberal perspective might be more likely to highlight positive developments and express more positive sentiment.

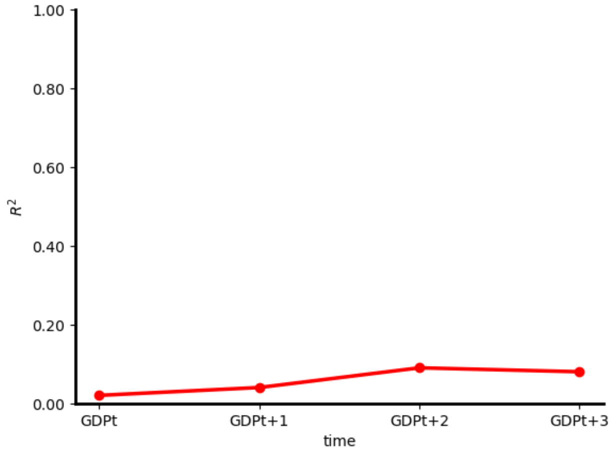
Figure 3.6: Main Italian newspapers sentiment indices



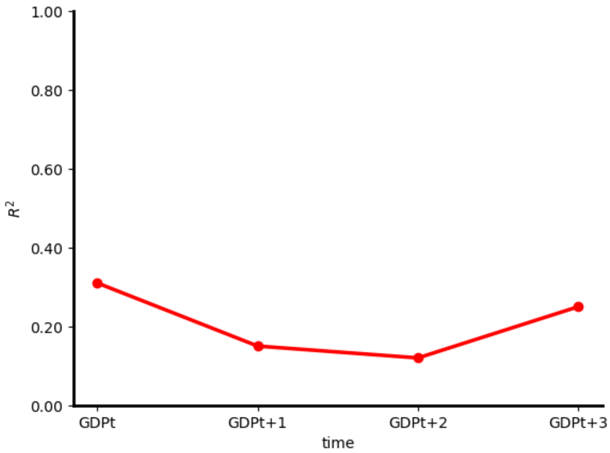
The graph depicted in Figure 3.7 showcases the predictive power of each newspaper-derived index in relation to the current Gross Domestic Product (GDP) and its future projections for t+1, t+2,



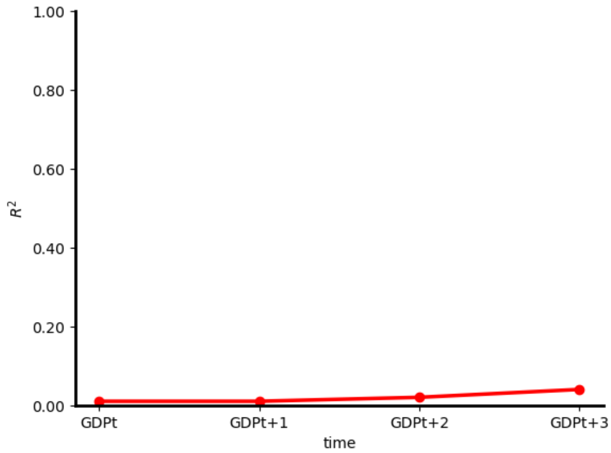
(a) Corriere della Sera



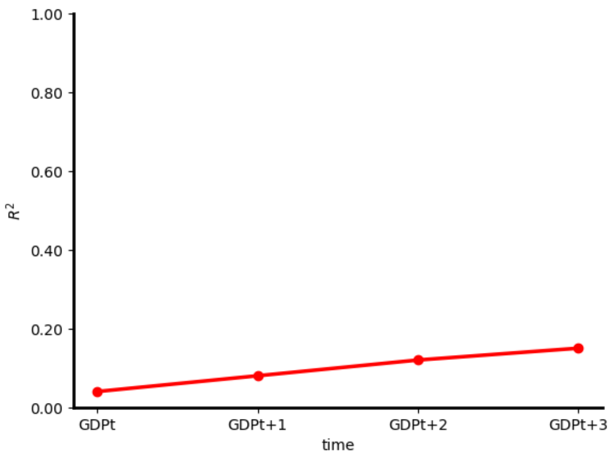
(b) Il Fatto Quotidiano



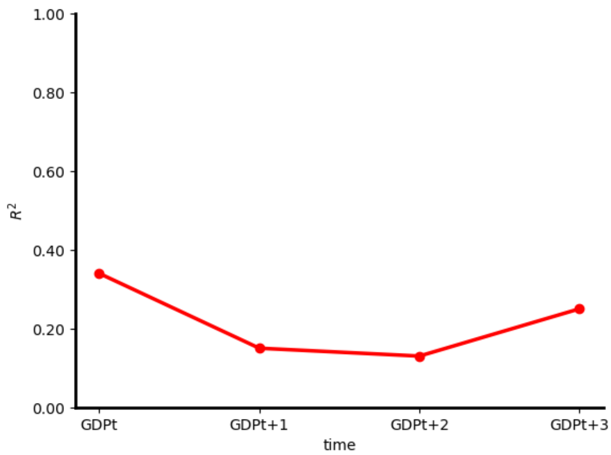
(c) La Repubblica



(d) La Stampa



(e) Il Sole 24 Ore



(f) Combined Sentiment Index

Figure 3.7: Variance explained

and t+3. The graph highlights La Repubblica as the newspaper whose derived index closely aligns with the final index outcome, indicating its superior predictive capability.

Exploring deeper into the high predictive capacity of La Repubblica, it's worth noting that its status as a general newspaper allows it to tap into a wide array of information and insights from various sources. This broad-based approach could potentially enhance its understanding of current affairs, thereby improving its ability to forecast future trends.

Moreover, La Repubblica's wide readership, owing to its general newspaper status, could be a contributing factor to its predictive accuracy. In contrast, Il Sole 24 Ore, with its primary focus on economic news, may not have as wide a readership, but its in-depth coverage of economic and financial news could provide valuable insights into long-term trends and developments.

On the other hand, La Stampa, Il Fatto Quotidiano, and Corriere della Sera, despite showing similar trends in their predictive capacity over time, exhibit relatively limited overall predictive power. This suggests that while these outlets may offer valuable insights into specific events or trends, their coverage may not be as comprehensive or in-depth as that of Il Sole 24 Ore or La Repubblica.

Table 3.2: Estimated effects of Newspaper indices on GDP at different horizons

Variables	GDP_t	GDP_{t+1}	GDP_{t+2}	GDP_{t+3}
Corriere della Sera	0.0112	0.0128	0.0136	0.0106
	(0.010)	(0.010)	(0.009)	(0.009)
Il Fatto Quotidiano	0.0024	0.003	0.0044	0.0041
	(0.002)	(0.002)	0.002	0.002
La Repubblica	0.0210 ***	0.0139 ***	0.0126 **	0.0176***
	(0.005)	(0.005)	(0.005)	(0.005)
Il Sole 24 Ore	0.0104	0.015*	0.0179**	0.0199***
	(0.009)	(0.008)	(0.008)	(0.008)
La Stampa	0.0018	0.0019	0.0022	0.0033
	(0.003)	(0.003)	(0.003)	(0.003)
Economic sentiment	0.0259***	0.0169***	0.0151**	0.0215***
	(0.006)	(0.007)	(0.006)	(0.006)

Notes: The table reports the reaction to the newspaper index using GDP and different time horizons. Robust standard errors in parentheses; $***$, $**$, and $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3.2 displays the estimated effects of the economic sentiment index, derived from various newspapers, on the Gross Domestic Product (GDP) at four different time points: t, t+1, t+2, and t+3. It is important to note that these estimates are derived from a regression analysis, which aims to identify the relationship between the economic sentiment index and GDP.

3.5 Real-time indicator evaluation

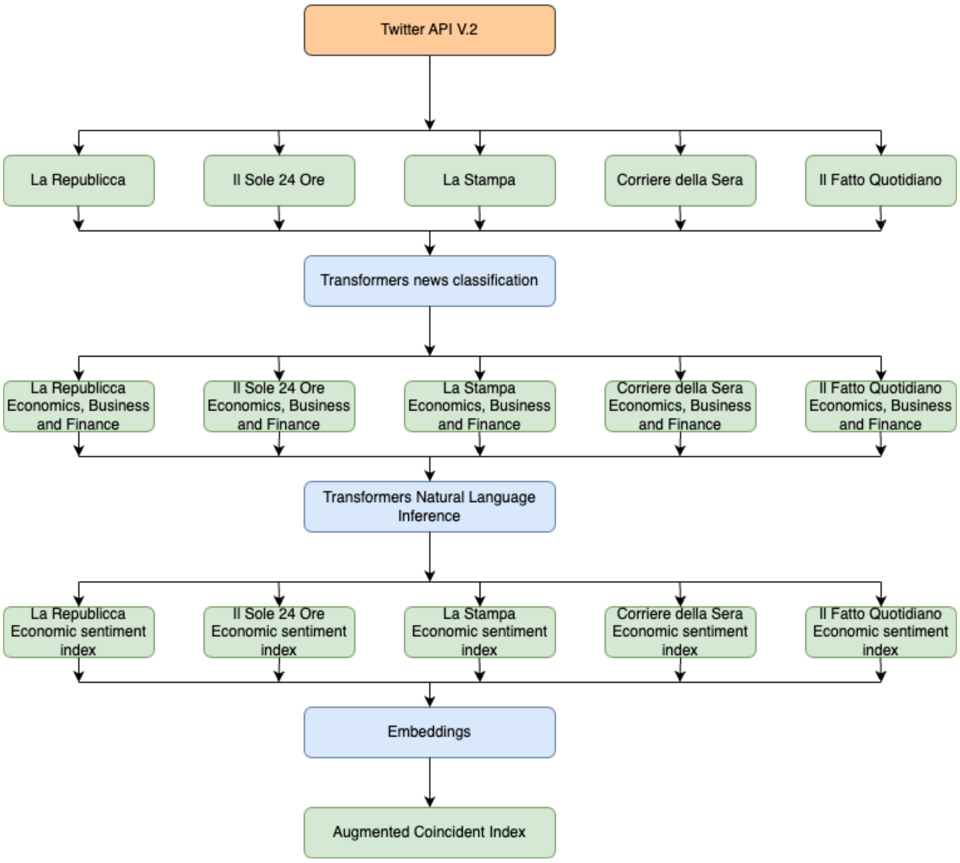
To develop an economic index that provides real-time information using data from five newspapers, we calculate embeddings for each instance of time considering all 5 newspapers as single input. To provide a more in-depth understanding of our methodology, we have included a diagram in Figure 3.8 that outlines the steps we took to create our augmented coincident index. This process begins by collecting articles from the top five newspapers in Italy, which are chosen based on their distribution throughout the country, using Twitter's Academic Access. This allows us to access the data while maintaining the integrity of any important information contained in the news stories.

Once the data has been collected, we use transformers to analyze the articles and classify them according to various topics. We specifically focus on articles related to the economy, businesses, and financial news. After this step, we use natural language inference to calculate the embedding such that for every day we calculate embedding that consider all articles of that day from all five newspapers. This allows us to be agnostic on a way how to aggregate them. This index can be used to track the current state of the economy and anticipate future economic developments.

Once we have embeddings calculated for each day, we use cosine distance to create our index. In essence, our approach involves calculating the embedding vector against which we measure dissimilarity, which we assume represents the hypothesis that the economy is not performing well, essentially reflecting negative economic sentiment. We then measure the cosine distance between each individual forecaster's embedding vector and this embedding vector.

The cosine similarity $\mathrm{CS}(\mathbf{e}_t, \bar{\mathbf{e}}_t)$ between the embedding vector $\mathbf{e}_t$ and the hypothesis embedding

Figure 3.8: Methodology diagram. EBS is an abbreviation for economics, business & finance topics and ESI for economic sentiment index.



vector $\bar{\mathbf{e}}_t$ for period t is computed by:

$$\mathrm{CS}(\mathbf{e}_t, \bar{\mathbf{e}}_t) = \frac{\mathbf{e}_t \cdot \bar{\mathbf{e}}_t}{\|\mathbf{e}_t\| \|\bar{\mathbf{e}}_t\|} \tag{3.1}$$

where $\|\mathbf{e}_t\|$ and $\|\bar{\mathbf{e}}_t\|$ denote the Euclidean norms of $\mathbf{e}_t$ and $\bar{\mathbf{e}}_t$, respectively, and $\mathbf{e}_t \cdot \bar{\mathbf{e}}_t$ is the dot product, indicating the degree of alignment between the individual forecaster's embeddings and the average embedding for that period.

Following this, the cosine distance $\mathrm{CD}(\mathbf{e}_t, \bar{\mathbf{e}}_t)$ is calculated as:

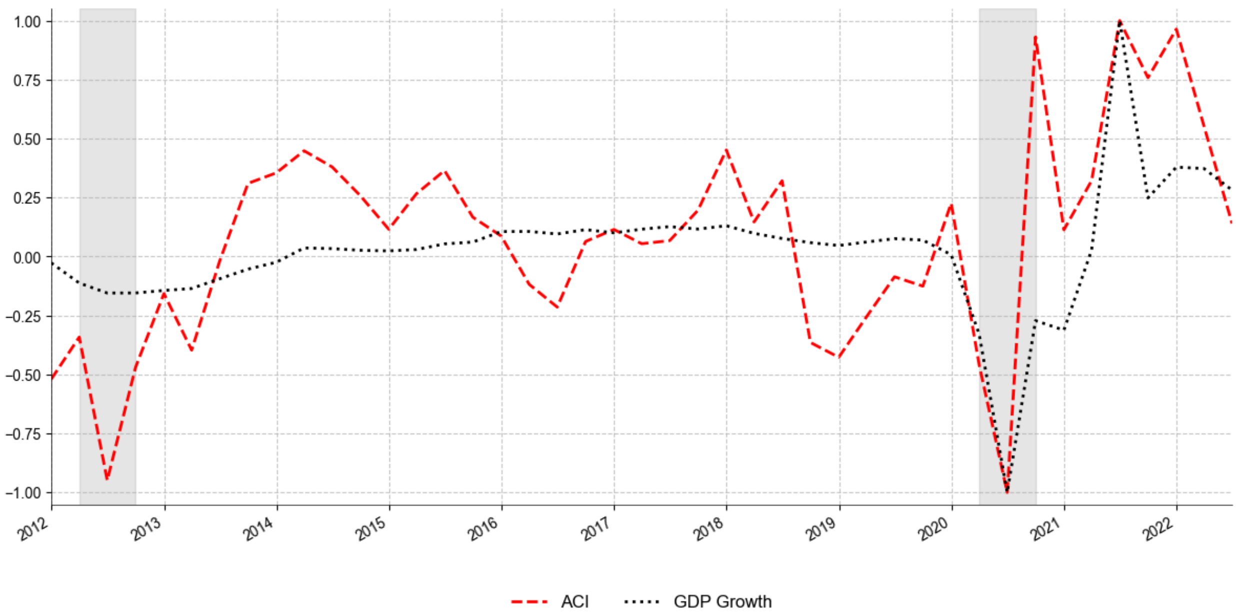
$$\mathrm{CD}(\mathbf{e}_t, \bar{\mathbf{e}}_t) = 1 - \mathrm{CS}(\mathbf{e}_t, \bar{\mathbf{e}}_t) \tag{3.2}$$

Cosine distance measures the dissimilarity between the daily embeddings and the hypothesis embedding.

By considering each newspaper together, we constructed an all-encompassing index that takes into account the different sources of information and their respective weights, thus capturing the overall sentiment conveyed in the articles.

In Figure 3.9, we present the resulting index, which we refer to as the "soft information index", alongside GDP growth. As can be seen, the index accurately reflects the recession phases, indicating that it is a sensitive and reliable measure of economic activity. We aggregate it to quarterly frequency taking average of the period in order to make it comparable with GDP growth. In Appendix E we show evolution of our index in respect to the Stock exchange index and Economic policy uncertainty.

Figure 3.9: Italian real-time soft information index of economic activity



However, there are times when the index shows more pronounced fluctuations than GDP growth, such as during the 2018 Italian elections and the subsequent crisis with the Budget Law. Hence, this indicates that factors other than the state of the economy, such as political events or other developments, may affect the tone of the articles, which only occasionally affects the economy.

We compare our soft information index with the Bank of Italy's Ita-coin coincident cyclical indicator introduced in Aprigliano and Bencivelli (2013) [2] as an additional benchmark. In this analysis,

[2] The Bank of Italy has developed a coincident cyclical indicator for the Italian economy called Ita-coin, which is updated every month and provides a real-time estimate of the trend of economic activity. Ita-coin is based on various variables, including quantitative data (such as industrial production, inflation, retail sales, trade flows, and

we calculate the Mean Directional Accuracy (MDA) as a measure of the performance of our soft information index. MDA is calculated by taking the average of the signs of the differences between consecutive periods for both our index and the benchmark ITA-Coin. Precisely, MDA is calculated as follows:

$$MDA = \frac{1}{N} \sum_{t=\underline{t}}^{\bar{t}} sign(\mathrm{y_{t+1,i,j}} - \mathrm{y_{t,i,j}}) == sign(\mathrm{f_{t+1,i,j}} - \mathrm{f_{t,i,j}})$$

Where N is the number of periods, t is the time, y is the value of our index, and f is the value of ITA-Coin.

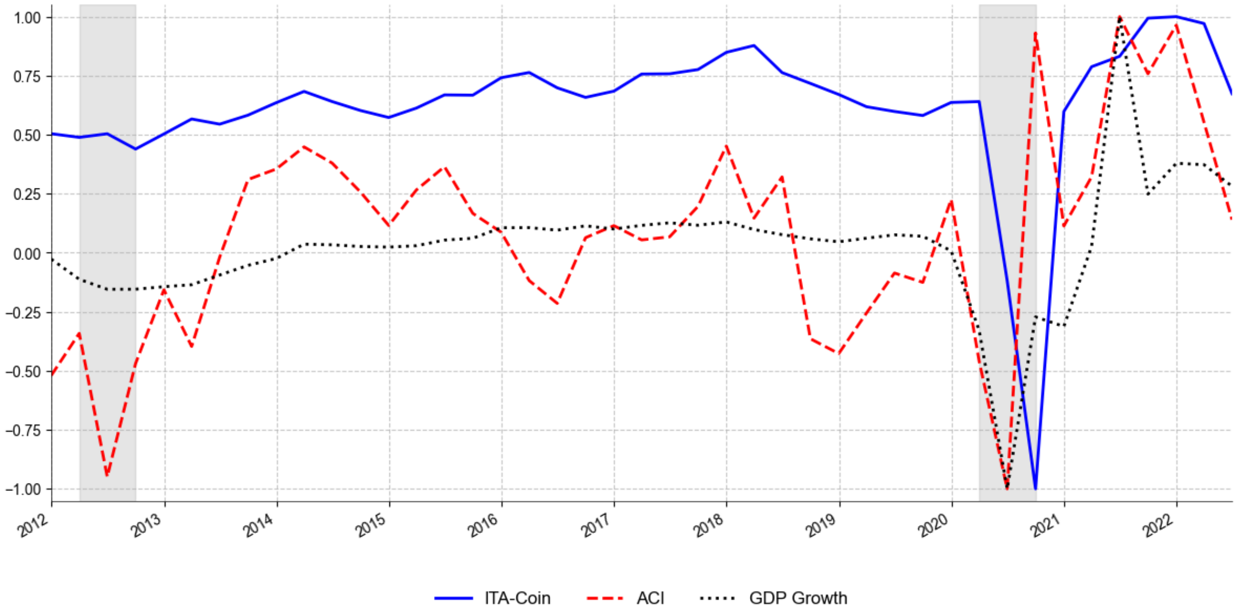
Figure 3.10 compares the two indicators, including the Mean Directional Accuracy score. The Mean Directional Accuracy score measures the indicator's accuracy in predicting the direction of GDP growth. As seen in the figure, our soft information index has a higher Mean Directional Accuracy score than Ita-coin, indicating that it is more precise in predicting the direction of GDP growth. That suggests that our indicator may be able to anticipate changes in the direction of GDP development more accurately than Ita-coin. The results of this calculation show that our index, denoted as ACI, has a Mean Directional Accuracy of 76%, while ITA-Coin has 64%. Hence, this suggests that our index is more precise at predicting the direction of economic activity than ITA-Coin.

As can be seen, our index accurately captures periods of recession. It displays more significant fluctuation than GDP growth during specific events, such as the Italian Elections in 2018 and the subsequent budget crisis. Overall, our index is a valuable tool for predicting the direction of economic activity in real-time, with a performance superior to that of the benchmark ITA-Coin.

To further illustrate the effectiveness of our real-time economic index, we present evidence in the form of cross-correlation plots. These plots show the relationship between our index and other relevant economic indicators, such as gross domestic product (GDP) and the Bank of Italy's coincident

equity indices) and qualitative data (such as household and business confidence, PMI indicators, and other subjective measures). That allows it to capture a broad range of information about the state of the economy and provide a more comprehensive view of economic trends

Figure 3.10: Italian real-time soft information index of economic activity vs. ITA-Coin



cyclical indicator, Ita-coin.

Figure 3.11: Cross-Correlation with GDP

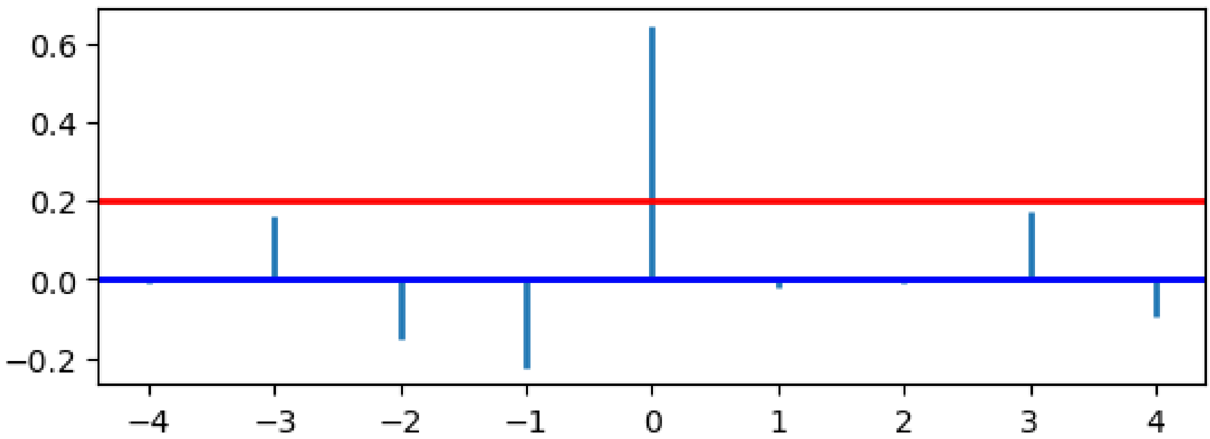
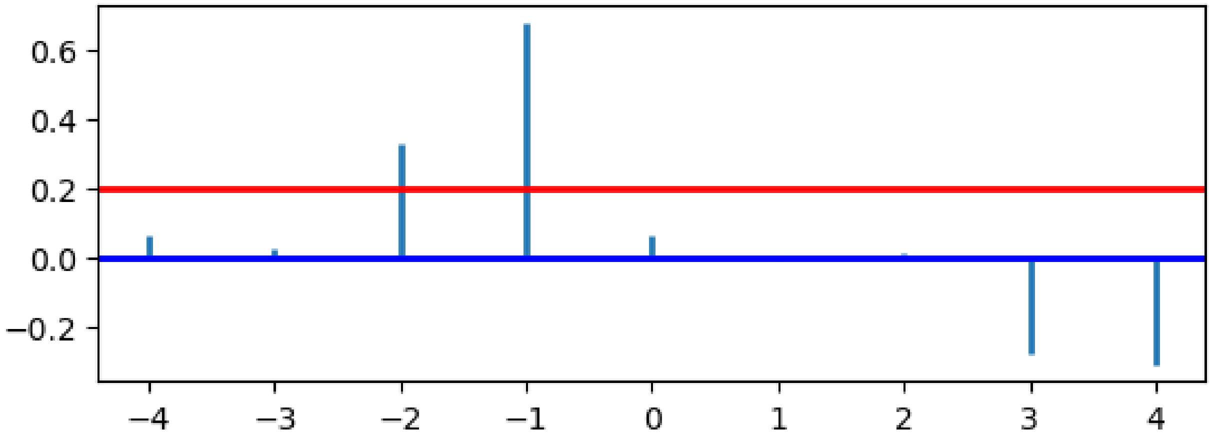


Figure 3.11 shows the cross-correlation between our index and GDP. The value at time 0 is significant, indicating that our index accurately tracks the current state of the economy. Additionally, our index is available at a higher frequency than GDP, which means that we can obtain information about the state of the economy before the GDP aggregate official release, which is usually delayed significantly.

Figure 3.12 shows the cross-correlation between our index and Ita-coin. As can be seen, our index

Figure 3.12: Cross-Correlation with ITA-Coin



leads Ita-coin, indicating that it can capture the current economic situation. Hence, this demonstrates that our index can access the "state of the economy" more quickly and provides evidence of benefits using our index as a real-time economic indicator.

3.6 Conclusions

This paper presents a method for developing a coincident index that uses newspaper articles as a real-time data source for economic modeling and forecasting. We have made some promising results with our approach and are continuing to refine it.

First, we gathered a large dataset of newspaper articles from the top five newspapers in Italy based on their distribution throughout the country. This dataset included a total of 2.5 million articles, spanning from 2012 to 2022. We used Twitter as a source for this data, as each selected newspaper frequently shares news and information on the platform, providing a synopsis of the articles and a link to the online version. By accessing this data through Twitter's Academic Access, we were able to maintain the integrity of any vital information that may be provided by the news stories while also reducing the costs associated with data collection.

We utilized a state-of-the-art machine learning model known as transformers to analyze the data. These models are particularly effective for natural language processing tasks, such as classifying text into different categories or identifying sentiment (Vaswani *et al.*, 2017 Vaswani et al. (2017)). Using transformers, we could accurately model topics and sentiment at the sentence level, which allowed us to gain a more detailed understanding of the economic landscape in Italy. We focused our analysis on economic, business, and financial news articles.

Based on the sentiment expressed in these articles, we constructed a real-time economic index that reflects the overall sentiment of the articles. Higher index values indicate more positive sentiment, while lower values indicate negative sentiment. We compared our index to the Bank of Italy's coincident cyclical indicator, Ita-coin. We found that our index was more precise in predicting the direction of GDP development. To further compare the two indicators, we also used the Mean Directional Accuracy score, which measures the accuracy of predictions made by an index. Our index had a score of 76%, while Ita-coin had a score of 64%.

Our results suggest that our approach to constructing a real-time economic index using newspaper articles as input data is a promising method that policymakers and forecasters could use to respond promptly and effectively to economic developments. In this approach, we relied only on alternative text data without including any standard time series used in constructing coincident indices.

It is essential to ensure that an index is both reliable and robust to provide valuable insights and information for decision-making and analysis. This requires us to consider various factors, including the methodology used for collecting and analyzing the data and the data sources themselves. Additionally, expanding further research can improve the index's comprehensiveness and accuracy. One potential approach for enhancing an index is to evaluate its performance when additional quantitative data is included in the analysis. This common practice can provide valuable insights into the strengths, weaknesses, and potential areas for improvement of the index. Integrating supplementary quantitative data can address potential limitations or biases in the index and produce a more comprehensive representation of the economic activity. In addition to these benefits, incorporating further quantitative data can also limit the index's volatility during certain periods. By including more data points, the index becomes more stable and less prone to fluctuations, leading to a more accurate and reliable representation of the economy.

Moreover, the inclusion of politics-related articles can enhance the index's accuracy by accounting for the impact of political events and decisions on the economy as a whole. Thus, broadening the scope of data sources can ensure that the index is as robust, reliable, and comprehensive as possible.

In future research, we plan to explore deeper into the Twitter data, considering factors like follower count, engagement levels (e.g., clicks), changes over time, and sentiment in headlines versus main articles. We plan to examine the credibility of all articles, considering whether high-profile columnists or changes in editors influence the sentiment expressed in the articles. We will also investigate whether subsequent developments prove sentiment correct (t+4) and how a machine-learning approach treats assertions in the articles.

Carefully considering the methodologies for data collection and analysis, evaluating data sources, and integrating additional quantitative data and political articles can significantly enhance the comprehensiveness, accuracy, and reliability of our soft information index. These improvements can provide valuable insights for decision-making and analysis, making the index a more effective tool for researchers and policymakers.

Chapter 4

Nowcasting inflation with Lasso-regularized vector autoregressions and mixed frequency data

joint with Tesi Aliaj and Massimiliano Tancioni

Abstract

We evaluate the predictive performances of the least absolute shrinkage and selection operator (Lasso) as an alternative shrinkage method for high dimensional vector autoregressions. The analysis extends the Lasso-based multiple equations regularization to a mixed/high-frequency data setting. Very short-term forecasting (nowcasting) is used to target the Euro area's inflation rate. We show that this approach can outperform more standard nowcasting tools in the literature, producing nowcasts that closely follow actual data movements. The proposed tool can overcome information and policy decision problems related to the substantial publishing delays of macroeconomic aggregates.

Keywords: Nowcasting, Inflation, Model shrinkage methods, Lasso-VAR, Mixed frequency Data.

4.1 Introduction

Inflation is back and hitting the economy worldwide. There is a significant concern about how inflation will evolve, given its direct and uncontroversial effect on consumers, businesses, governments, and central banks' expectations and incentives. In principle, everyone makes decisions based on present information and expectations, thus forecasts, of the price levels in the future. Because crucial statistics on key macroeconomic variables are available with a significant delay, real-time (or high-frequency) forecasting of the present and near future (nowcasting) is becoming increasingly relevant in economics (Bańbura et al., 2013).

The anticipation of the present state of the economy adds a critical dimension, timeliness, to the information sets upon which private agents and policymakers inform their decision processes. Unsurprisingly, the practice of nowcasting is becoming increasingly common at central banks and super-national economic institutions.[1]

Possibly because of their high-frequency forecasts, policymakers partially expected the current inflation hype and anticipated its emergence. The availability of increasingly accurate short-term inflation forecasts enriches the information on which policymakers form their decision processes. In this perspective, monitoring the variability of price levels at higher frequencies than the current publication timing is essential.

Our paper contributes to the nowcasting literature on a specific aspect of the high-frequency forecasting practice, i.e. on the model dimensionality reduction issue.

Nowcasting introduces three significant complications to the practical implementation of the analyses. First, including timely information from various sources implies that data are sampled at different frequencies and possibly in an asynchronous manner. Mixed frequency data inherently imply missing values for lower frequency inner observations (gaps) and outer observations (ragged edges). Second, the need to relate standard statistical information to high-frequency information

[1] The Board of Governors of the Federal Reserve implemented the first nowcasting model to provide high-frequency forecasts of GDP (Giannone et al., 2008). Since then, various versions of this seminal model have been built and implemented in other central banks, including the European Central Bank (Bańbura and Saiz, 2008) and the International Monetary Fund (Matheson, 2011).

from different sources implies that the information set is generally extensive. Nowadays, we can get a vast amount of information on the behaviour of individual prices in the distance of a "click", so the issue is how to make the best use of this information. This second point leads to the third analytical complication, i.e. the need to use model shrinkage methods to keep complexity to a computationally tractable level.

Model dimension issues arise particularly in multivariate dynamic settings, as in vector autoregressions (VARs), where the number of coefficients increases with the square of the variables in the VAR. The profligate parameter problem can make estimation infeasible or forecasts inaccurate in cases where the sample size is too small relative to the parameters' space. Here, the "horse race" among different missing data inputting methods and model shrinkage techniques comes to the fore. Several approaches have been proposed for nowcasting in high dimensional information settings, but only some pertain to the model shrinkage issue. On this latter terrain, reference studies in the nowcasting literature adopt either Bayesian shrinkage or dynamic factors convolutions, leading to large Bayesian VAR models (BVAR) and Factor-Augmented VAR models (FAVAR), respectively.

This paper proposes the least absolute shrinkage and selection operator (Lasso) as an alternative (machine learning-based) shrinkage method for high dimensional VARs estimated over mixed frequency data. Compared to other linear regularization methods, such as the Ridge regression and the Elastic Net estimator, the Lasso has the advantage of entirely excluding some information from the model.[2] The missing data inputting issue is instead aligned with the literature reference solutions. To the better of our knowledge, this is the first work in which a Lasso-based VAR regularizer is applied to nowcast Euro area HICP inflation considering a large, mixed frequency information set. Results show that our machine-learning approach is aligned with the BVAR model. At the same time, it can outperform the FAVAR model in the near-term prediction of both inflation and core inflation. Considering forecast sub-samples characterized by different degrees of price variability,

[2]We verify this potential advantage in the robustness checks step. Other machine learning algorithms are not considered since their ability to allow for nonlinear relations would penalize the other benchmark models considered in the analysis, which are inherently linear.

we obtain that the Lasso-VAR tends to outperform all the tested model alternatives in periods of reduced price variability. These results are robust to some evaluation directions, such as the size of the information set (i.e., the inclusion of policy variables) and the consideration of core inflation as a target variable. We also show that other (linear) machine learning-based regularization methods, such as the Ridge and the Elastic Net, do not provide predictive improvements on the Lasso. We interpret this last finding as an indication that the selection step (i.e. the fact that the Lasso entirely excludes some variables/lags from the model by shrinking their parameters to precisely zero) is critical to avoid overfitting and poor forecasting in highly parameterized models. The consideration of extended forecast horizons, i.e. moving from a nowcasting analysis to a forecasting analysis, shows that the BVAR modeling alternative attains the best predictive performances over the other model shrinkage approaches.

Our paper relates to two strands of literature. The first is the literature on short-term forecasting and nowcasting in high-dimensional information settings, which is becoming as vast as its importance for macroeconomic dynamics. To simplify, we can refer to a few significant works in the field. Giannone et al. (2008) evaluate the marginal effect of high-frequency information releases on current period (quarter) forecasts (nowcasts) in a Factor Model. Marcellino (2008) compare the predictive abilities of time-varying models, nonlinear time series models, and artificial neural network models against standard ARMA models in predicting US GDP growth. Modugno (2011) uses mixed frequency data and a Factor modeling framework to separate the effect on forecast revisions due to the inclusions of new data releases from that attached to the high-frequency dimension. Breitung and Roling (2015) propose a non-parametric approach to high-frequency forecasting using mixed-frequency data and equations. Di Filippo (2015) uses dynamic model averaging and dynamic model selection to forecast US and Euro area price inflation, considering a large set of predictors. Hubrich and Skudelny (2017) propose using performance-based forecast combination methods to forecast HICP headline inflation. Cimadomo et al. (2022), following Giannone et al. (2008), apply the mixed frequency data approach to nowcasting within a Bayesian Vector Auto-Regression (BVAR) resembling the modeling approach suggested by Bańbura et al. (2010).[3] Richardson et al. (2021) test the application of machine-learning algorithms in a high dimensional data setting to nowcast New Zealand's GDP. They show that these algorithms can significantly improve over a simple autoregressive benchmark and a dynamic factor model.

The second strand of literature is about using the Lasso as a shrinkage device in VARs. In this perspective, recent literature shows that Lasso-regularized VARs can provide an efficient solution to the "profligate parameterization" issue, as it ensures sparse structures. With the removal of over-parameterization and overfitting, the forecasts generated with these models can outperform those obtainable with benchmark single equation methods and the alternative multivariate methods developed for high dimensional settings (Basu and Michailidis, 2015; Lin and Michailidis, 2017; Nicholson et al., 2017; Messner and Pinson, 2019; Nicholson et al., 2020).

The remainder of the paper is organized as follows. Section two deals with data issues. Here we describe the data set, its continuous updating with mixed frequency information, how missing values are estimated, and the sample is re-balanced. Section three describes the econometric modeling strategy, the estimation, and its validation. Section four describes the nowcasting method and the metrics used to evaluate the out-of-sample model predictions. Here we discuss the main results of the analysis from a comparative perspective. Section five describes some robustness checks. Section six concludes.

4.2 Data and missing data imputation

Our analysis considers mixed frequency data, including weekly and monthly observations from September 2005 to August 2022. The benchmark model considers 31 variables, including inflation indicators for food, commodity, and electricity prices, changes in the broad money aggregate for the euro area (M3), and changes in bilateral and (nominal and real) effective exchange rates. Data are collected from Eurostat, the ECB Statistical Data Warehouse, Bloomberg, and the US Energy

[3] A (Kalman) filtering method has been suggested to estimate general equilibrium models by Canova and Ferroni (2011) in a data-rich environment.

Information Administration (EIA). The complete dataset is described in Appendix F.

In the benchmark model estimates, we purposely left out monetary policy variables to produce nowcasts that are unconditional on information about policy measures. This choice, which is quite unconventional in forecasting, is justified by the fact that we aim to provide real-time forecasts (nowcasts) upon which policy decisions might be anchored. European Central Bank and US Federal Reserve's key policy variables are then included in the nowcasting models to verify the robustness of the main results to extensions of the information set.

Data are published at different time frequencies, on different days, by various institutions that do not coordinate the publishing or the data revision frequency. The first issue is providing a fixed weekly structure for the irregular and unstructured data flow. We attribute all the (working week) daily observations to the specific week they belong. This choice is consistent with the idea that data can be included in the conditioning set if known to the forecaster, no matter the publication day (i.e., their relative weight in the information content of the weekly observations).

We then estimate the unobserved weekly measures from lower frequency (monthly) data. For this purpose, we follow the latent observation VAR method (L-VAR) recently described in Cimadomo et al. (2022). This method treats the missing weekly observations as latent processes that can be inferred using the Kalman Filter (KF).[4] With this strategy, we build a structured weekly dataset such that even an irregular flow of information in the time dimension (i.e. mixed frequencies with ragged edges) can feed any high-dimensional forecasting model.

Provided that the information set is continuously updated in a real-time data environment, we follow the standard practice in the nowcasting literature by introducing data vintages.[5] In this way, we can produce nowcasts at any time by using the data being made available at the same time

[4] CCimadomo et al. (2022) evaluate the nowcasting performances of this approach against two alternative strategies. The first is the blocking VAR (B-VAR), in which the VAR is estimated at the joint (lower) data frequency. The higher-frequency observations are considered separate lower-frequency variables. This approach avoids the latent state definition of the VAR for missing higher-frequency data, limiting the use of Kalman recursions to fill the ragged edges generated by the asynchronous data release. The second alternative is labeled cubic root VAR (C-VAR) in a monthly-quarterly mixed frequency data environment. It starts from a lower-frequency model estimate and then maps it into a higher-frequency model using Kalman filtering techniques, as in the L-VAR.

[5] The vintage is a set of new data available at a particular moment in time, i.e., in which the model is estimated.

nowcasts are calculated.

In the remaining part of this section, we thoroughly explain the missing data imputation process we followed in our analysis.

4.2.1 Imputation of high frequency missing data

In a continuously updating mixed frequency information set by vintages, missing weekly observations emerge within known monthly observations (for data released at the monthly frequency) and at the end of the sample (when the last observation of a monthly frequency variable still needs to be released). In the latter case, the missing data generate a time-evolving sample's ragged edges. In order to get a complete weekly dataset, we need to estimate the motion of the monthly variables at a higher frequency than that available from official sources' releases. The KF can contribute to guessing the missing values between each monthly observation.[6]

In order to fill the missing weekly data, we adopt Cimadomo et al. (2022) latent observation VAR method (L-VAR).[7] The main difference is that our procedure considers a machine learning-regularized VAR instead of a Bayesian VAR.

The proposed strategy relies on KF techniques applied to the state-space representation of the VAR. The missing (weekly) values for the low-frequency observations are conceived as existing - albeit latent - processes for weekly variables that are observed only at the monthly frequency.

Specifically, the missing data imputing procedure goes as follows. First, a preliminary complete weekly dataset is obtained by interpolating monthly observations using splines. The preliminary weekly dataset is then used to initialize the Kalman recursions using a regularized VAR. For the benchmark estimates, the start coefficients for the KF consider a Lasso-VAR.[8] The state-space representation of the VAR is used to iteratively apply the Durbin et al. (2001) simulation smoother,

[6] The KF is designed to filter out the best guess for the latent state of a system in an environment characterized by the presence of a given level of noise.

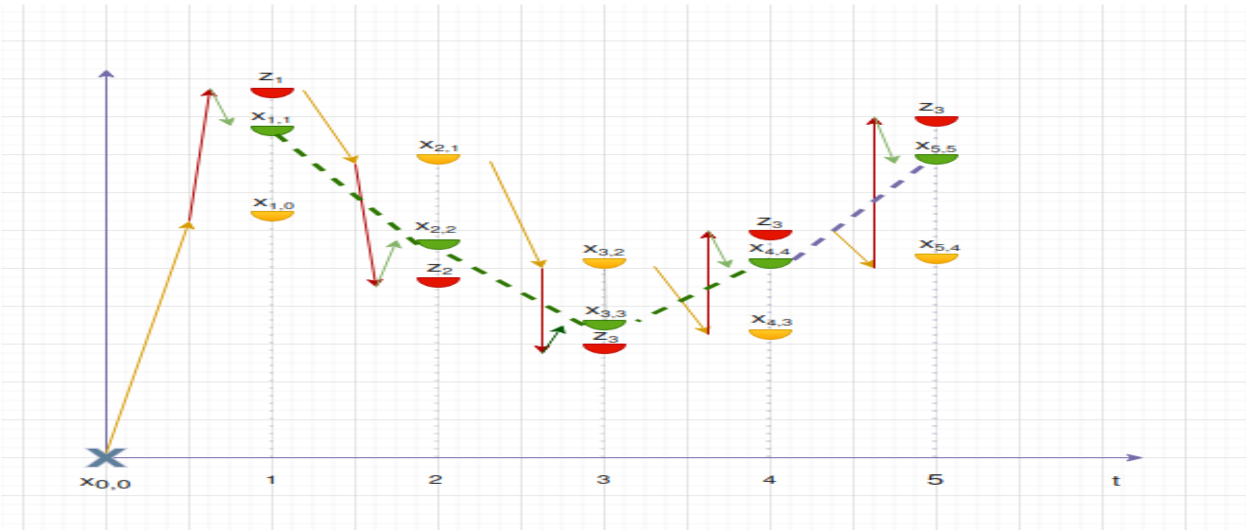
[7] In a frequentist setting, this strategy has been applied Giannone et al. (2008), Mariano and Murasawa (2010), Kuzin et al. (2011) and) Foroni et al. (2015). Eraker et al. (2014), Schorfheide and Song (2015), Cimadomo and D'Agostino (2016), Brave et al. (2019) apply the latent variable approach within Bayesian settings.

[8] Other regularization strategies for linear models (Ridge and Elastic Net) are also considered to evaluate the robustness of results to the use of alternative machine learning-based model shrinkage methods.

obtaining improved VAR and missing data estimates in each repetition of the procedure. The process is rerun until convergence - and thus the final complete weekly dataset - is achieved.

Since this process also yields nowcasts/forecasts conditional on the final dataset, a first one-month (four weeks) ahead forecast can be made. The procedure is then repeated for each new weekly data observation, obtaining weekly nowcasts. In fact, following each emission of new information, the KF is re-booted in order to provide corrected estimates. Figure 4.1 depicts the update-prediction procedure graphically. The red semi-circles are the observations released by official sources every four-time steps (weeks). The orange ones denote the output of the correction step, whereas the green semi-circles and lines result from the prediction step (this is where the data imputation occurs). The arrows show the direction of the KF updating process, starting from defining initial conditions to the final prediction step.

Figure 4.1: Graphical exemplification of the KF update-prediction procedure.



4.3 The econometric model

Since the seminal work of Sims (1980), the VAR has become the most used methodological approach to empirical macroeconomic modeling. Despite their popularity, VAR models are often at risk of over-parameterization, leading to overfitting and poor forecast performances. Such a drawback of the methodology becomes particularly stringent in high-dimensional data settings.

The over-parameterization problem can be solved by artificially penalizing the model coefficients. Prior structures (or hyperparameters) in Bayesian settings or common factors-based structures are often employed to reach this goal. In the nowcasting literature, two are, in fact, the main modeling approaches to the high dimensionality issue: *i)* large Bayesian VARs (BVAR); *ii)* Dynamic Factor and Factor Augmented VAR Models (DFM-FAVAR).

A non-exhaustive summary of literature in which these approaches are proposed is, for DFMs-FAVARs, Stock and Watson (2002a), Stock and Watson (2002b), Giannone et al. (2008), Aruoba et al. (2009) and Cascaldi-Garcia et al. (2021). On the side of BVARs (2021), some significant contributions are Bańbura et al. (2010), Koop and Onorante (2019) and, more recently, Cimadomo et al. (2022). Higgins (2014) proposes the joint consideration of the two approaches in a high-dimensional forecasting setting.

We propose a Lasso-based regularization as an alternative approach to overfitting issues in high-dimensional VARs, seeking to improve their real-time predictive performances (Basu and Michailidis, 2015; Lin and Michailidis, 2017; Nicholson et al., 2017; Messner and Pinson, 2019; Nicholson et al., 2020).

4.3.1 The Lasso-VAR

The Lasso is a method for automatic variable selection and parameter shrinkage used to select the most informative predictors of a target variable from a large set of variables and parameters. A peculiarity of the approach is that the information set (i.e. number of variables) might be even larger than the sample size. This peculiarity makes high-dimensional modeling and forecasting feasible for any degree of model dimension and complexity.

The Lasso has been initially developed for single equation settings by Tibshirani (1996). The Lasso approaches curve fitting as a quadratic programming problem, where the objective function penalizes the total size of the regression coefficients based on the value of a tuning parameter, λ.

In doing so, the Lasso can drive the coefficients of irrelevant variables to zero, thus performing the automatic variable selection. The strength of the penalty must be tuned. The stronger the penalty, the higher the number of coefficients shrunk to zero. The model is thus forced to select only the most important predictors, i.e. those with the highest contribution to the prediction of the target variable.

Let $\{x_t\}_{t=1}^{T}$ be a k dimensional vector including time series that follow a VAR process of order p. All the variables are entered in first differences in the VAR, such that it considers aggregate, commodity-specific and currency-specific relative prices inflation rates. We have verified with Phillips-Perron tests that all the time series included in the VAR are stationary (non-stationarity test results are provided in the Appendix F).

We fix the maximum order of lag p to 12 periods (thus one quarter). The chosen lag order is higher than the one indicated by the Akaike Information Criterion, suggesting a four-week lag order. This choice allows the model to capture economically plausible lags in the transmission dynamics from specific commodity price variations to other prices and aggregate inflation. [9]

$$\begin{aligned} X_t = A_1 X_{t-1} + \ldots + A_p X_{t-p} + u_t \\ u_t \sim \mathsf{N}(0, \Sigma_u) \end{aligned} \tag{4.1}$$

where each A_i is a $k \times k$ matrix of coefficients for the endogenous variables, and $\mathbf{u}_t \sim (\mathbf{0}, \mathbf{\Sigma}_u)$ is the vector of reduced-form errors. Since we standardize data before modeling, the VAR does not consider the k-dimensional intercept vector.

The LASSO objective function is minimized as follows:

$$\hat{\mathbf{A}}(\lambda) = \arg\min_{\mathbf{A}} \frac{1}{T} \|\mathbf{AZ} - \mathbf{Y}\|_2^2 + \lambda \|\mathbf{A}\|_1, \tag{4.2}$$

where λ is the shrinkage parameter, whose calibration targets the out-of-sample model's predictive

[9] Nicholson et al. (2020) fix the maximum order of lag of the VAR to the number of periods included in the lower frequency dating, such as four weeks or 22 trading days in a month.

ability, i.e. the minimization of the forecast error. The optimization problem is solved by applying a coordinate descent numerical procedure, as explained in Kim et al. (2007) and Friedman et al. (2010).

4.3.2 Lasso-VAR time series cross-validation

As immediately evident from Equation 4.2, λ is the most critical parameter in the Lasso framework. Its calibration is based on selecting the best predicting model, which should not be sample-specific. To minimize the risk of a sample-specific calibration, a cross-validation stage, based on sample splitting, is thus employed for getting the "optimal" value for λ.

In this respect, the data set is divided into a training and a test sample. The test set is for final evaluation, whereas the training set is split into five subsets. We follow an expanding window (more precisely, an "anchored walk forward") approach to cross-validation (Carta et al., 2021).[10]

The anchored walk forward (five-fold) cross-validation method implies a gradually expanding training set, pushing forward a fixed dimension test set. The 5-split time series cross-validation method is exemplified in Figure 4.2, where the length of training and test sets is shown on the horizontal axis. In contrast, the five cross-validation iterations are shown on the vertical axis.

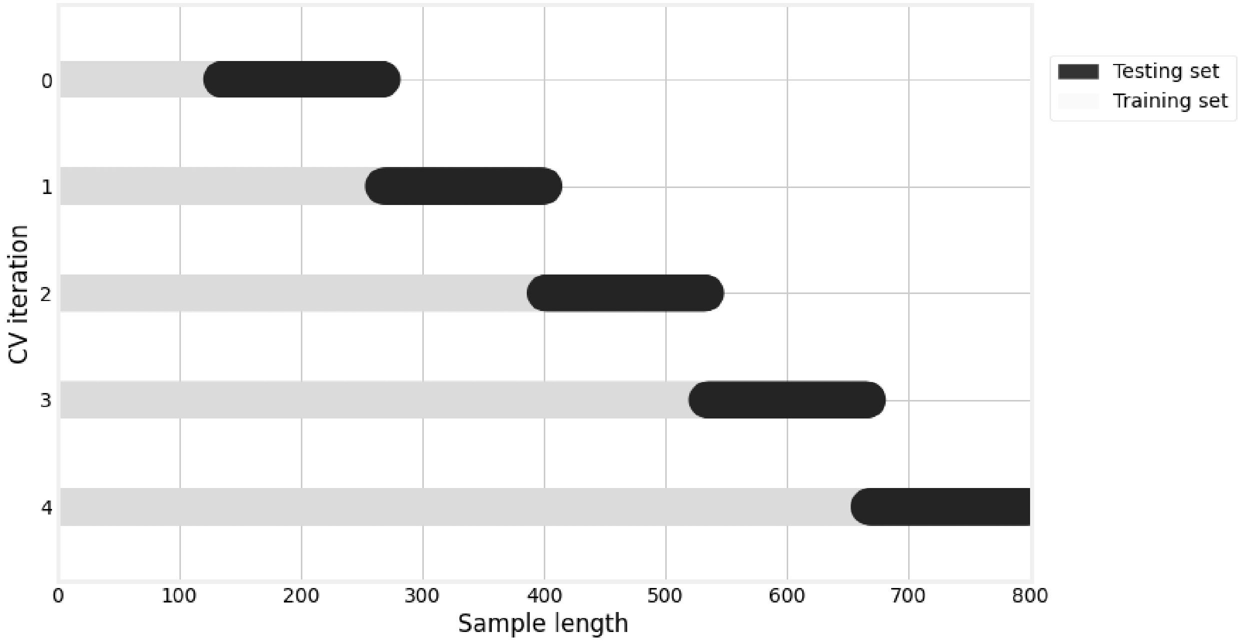
4.4 Results and model comparisons

After estimating the sparse model resulting from the cross-validated Lasso-VAR, we proceed to short-term forecasting, considering a sample spanning from July 2019 to August 2022. Consistent with the idea of nowcasting (i.e., forecasting the present or very near future), we are interested in a four periods-ahead forecast exercise, thus covering one month given the weekly update of the high-frequency information. In the robustness checks, we also consider extensions of the forecasting window to eight and 12 periods ahead, moving from a typical nowcasting analysis to forecasting.

We thus restrict the information set to the one that would be available during the week in which

[10]Time series cross-validation essentially considers time dependence, such that the training set includes observations temporally preceding those in the test set.

Figure 4.2: Graphical example of the five-split time series cross-validation



the nowcast is performed. Following Diebold (2020) approach to using information vintages, we build 165 weekly data releases (the weeks between July 2019 and August 2022). These releases are imputed using the methodology described in section 4.2 to deal with the missing data issues related to mixed frequencies and ragged edges. During the weekly nowcasting exercises, we consider one other vintage per nowcast up to the date in which one additional monthly information is released. We then move to the next vintage of data.

4.4.1 Evaluation of the Lasso-VAR nowcasting performances

A standard dynamic (recursive) forecasting method is applied to calculate the near-term out-of-sample forecast:

$$X_{T+h|T} = c + A_1 X_{T+h-1|T} + ... + A_p X_{T+h-p|T} \tag{4.3}$$

In order to provide a first evaluation of the performances of a real-time approach to inflation forecasting, the Survey of Professional Forecasters (SPF) estimate of the Euro area HICP inflation is first taken into account.

We acknowledge that the comparison with the SPF is only partially legitimate since the latter is not based on nowcasts. The reference to the SPF is only to highlight the potential improvements in forecasting macro-aggregates coming from the use of higher frequency information and efficient model shrinkage methods, as compared with an established methodology within the operation of central banks.[11]

Figure 4.3 compares Lasso-VAR-based nowcasts, realized HICP inflation and SPF estimates. The figure also reports the nowcasts' 95% confidence intervals, obtained from bootstrapped forecast errors.[12] We also include the forecasts obtainable by a simple linear extrapolation of inflation from its past, obtained by a naive AR(2) process (Cimadomo et al., 2022). Unsurprisingly, the nowcasting approach outperforms the AR(2) model, since the latter neglects the information embedded in the higher frequency variables. The predictive improvement obtained with the Lasso-VAR nowcasts on the naive benchmark is particularly evident when relevant changes in dynamics are building up, possibly due to unexpected shocks.

In order to evaluate the forecasting model performances over objective metrics, the Root Mean Squared Forecast Error (RMSFE) is adopted and calculated on out-of-sample residuals.

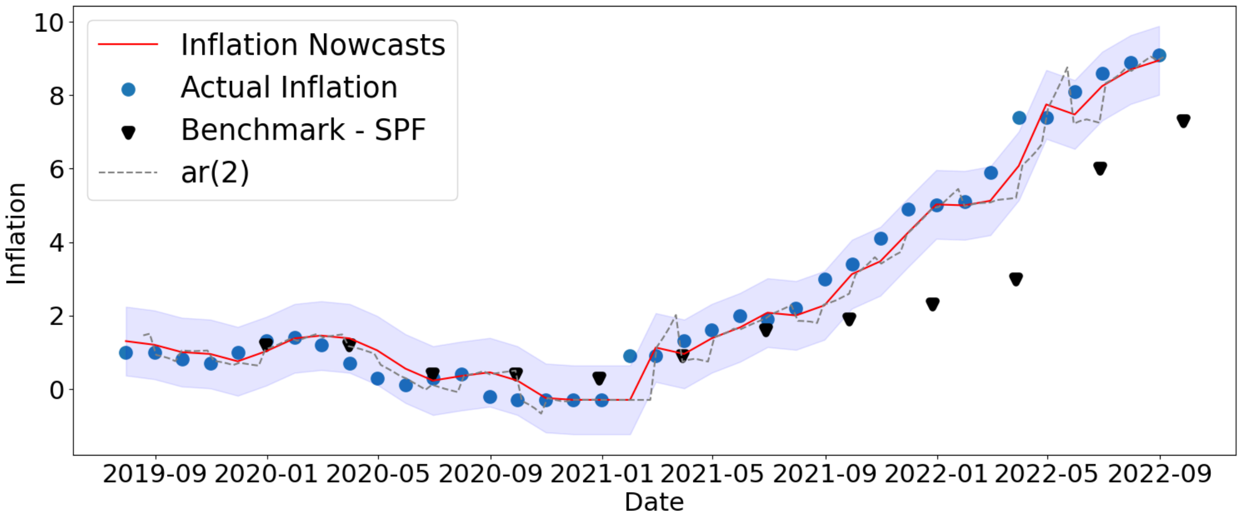
$$RMSFE = \sqrt{(\frac{1}{H})\sum_{i=1}^{n}(x_i - \hat{x}_i)^2} \tag{4.4}$$

A critical step in evaluating the nowcasting results is deciding to which values these nowcasts are to be compared. We do not have the actual values of the weekly inflation rate within two consequent

[11] The SPF is a quarterly survey conducted by the European Central Bank, reflecting the average perception of approximately 90 experts regarding their expectations about the dynamics of a set of leading macroeconomic indicators. In the first month of each quarter, the participants (experts affiliated with financial or non-financial institutions within the Euro area) declare their expectations regarding the HICP inflation rate, GDP growth and unemployment (and their degree of uncertainty). Expectations are formed after having been provided with all the information available up to that date.

[12] Since forecast errors are normally distributed according to the Shapiro-Wilk, D'Agostino and Pearson, and Anderson-Darling tests for normality, we have verified that analytical and bootstrapped confidence intervals are only marginally different.

Figure 4.3: Nowcasts, actual HICP inflation and SPF estimates



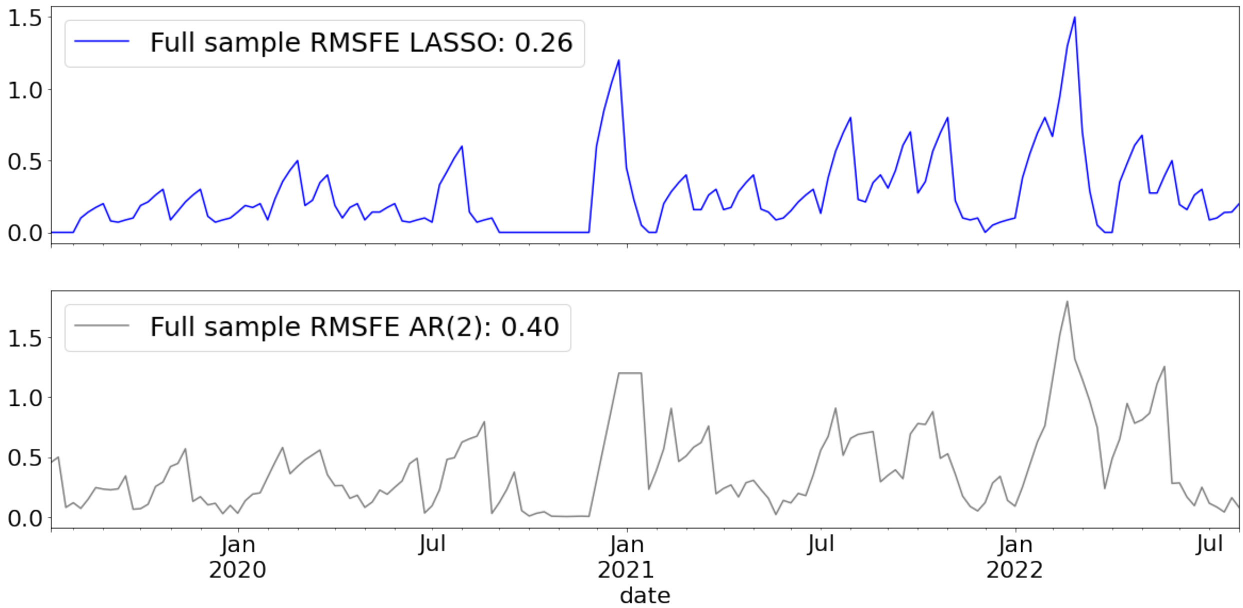
Notes: Nowcasts' 95% confidence intervals are obtained from bootstrapped forecast errors.

months, and the model is trained and tested on subsets including both actual inflation data (at the monthly frequency) and imputed data (at the weekly frequency). Consistent with the standard approach in the literature, we calculate the ex-post RMSFE considering weekly point HICP inflation nowcasts and the realized HICP inflation, which is observed at the monthly frequency. This approach highlights the role of the arrival of new high-frequency information for forecasting performances.

Figure 4.4 depicts the dynamics of the RMSFE for the Lasso-VAR-based nowcasts and for the AR(2) benchmark over the entire (weekly) sample spanning from July 2019 to August 2022. The graphs highlight that the forecast errors increase with time, possibly reflecting the build-up of inflationary pressures at the end of 2021.

The behavior of the forecast errors over the weekly sample indicates that the forecasting accuracy improves as more weekly information becomes available over the month. This result is better depicted in Figure 4.5, in which we report, for every week in a month, the average RMSFEs for the point HICP inflation nowcasts in the month, calculated over the entire sample (July 2019 - August 2022). The behavior of the average forecast error over subsequent weeks indicates that the proposed Lasso-VAR is a valid nowcasting tool. As more information becomes available each week of the month, the forecast error decreases. Such a decrease does not materialize for the naive AR(2)

Figure 4.4: Root Mean Squared Forecast Error



benchmark, denoting an RMSFE slightly lower than the Lasso-VAR at the beginning of the month, then increasing with the (weekly) forecast horizon. This difference in the average RMSFE behavior over the weeks in the month highlights the informational advantage arising from the efficient use of higher-frequency information.

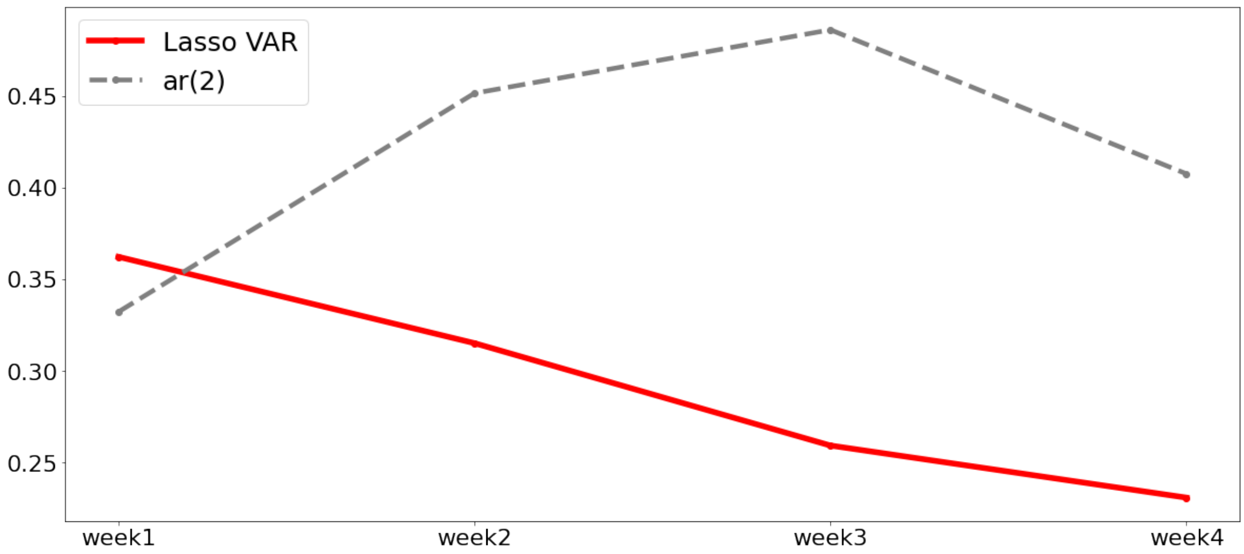
4.4.2 Model comparisons

We compare the performance of our Lasso-VAR-based nowcasting approach to two modeling benchmarks in the literature, the FAVAR and the BVAR [13]. Since our main focus is to evaluate the relative forecasting abilities of the methodologies, the model comparisons are performed on standardized experimental features. The information set and the missing values imputation strategy are the same across model specifications, and the maximum lag order is fixed to 12 in all model shrinkage alternatives. For the FAVAR model, the number of top factors identified by the largest eigenvectors is defined such that they jointly explain about 85 percent of the variation in the data. In our analysis, this threshold is reached with three factors.

Even in this case, the forecast performances of our nowcasting strategy are tested considering an

[13] We implement BVAR from Cimadomo et al. (2022)

Figure 4.5: Nowcasting performance per week of the month



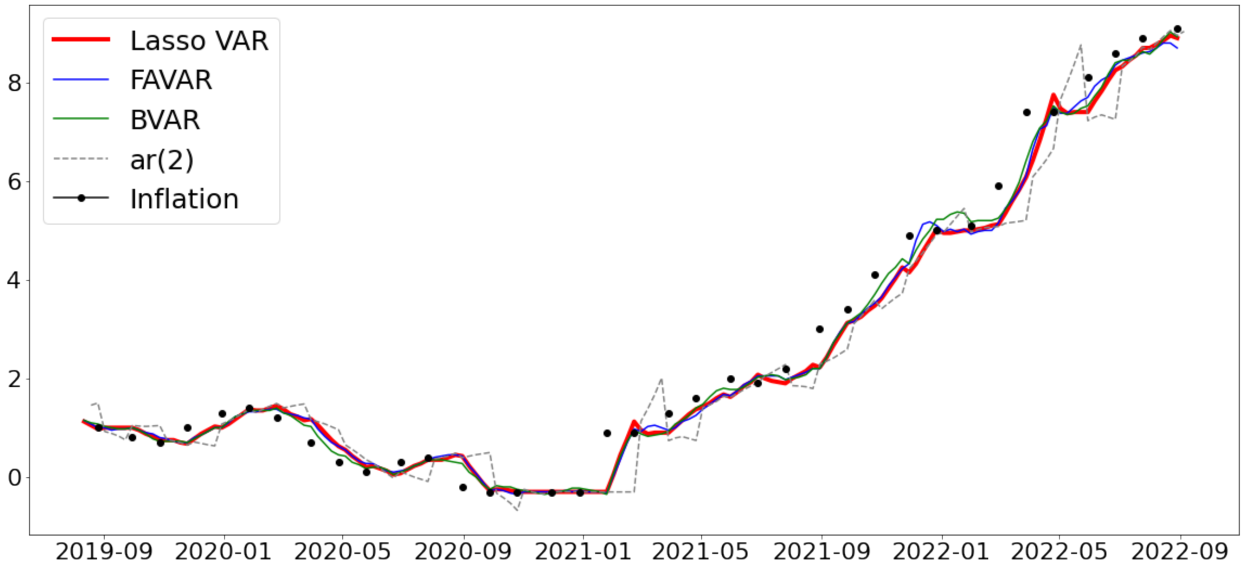
Notes: The line denotes average monthly RMSFEs for weekly point HICP inflation nowcasts for the sample July 2019 - August 2022.

extended period spanning from July 2019 to August 2022. Reference to an extended sample allows us to verify whether model comparison results are sample-dependent, an issue that emerged in recent literature (de Bondt et al., 2021; Dauphin et al., 2022). To better highlight this point, the forecasting period over which model comparisons are evaluated is divided into three reference periods: a pre-COVID-19 period, spanning from July 2019 to February 2020; a COVID-19 crisis period, spanning from March 2020 to September 2021 (i.e. when the European mass vaccination campaign was almost completed, allowing for a generalized relaxation of non-pharmaceutical containment measures); a post-COVID-19/Energy crisis period, starting from October 2022 and still ongoing. In Appendix I, we also show trace plots for these three periods to illustrate how the lagged variables considered by our Lasso-VAR-based nowcasting approach change across different periods.

Figure 4.6 depicts the Lasso-VAR, the FAVAR and BVAR-based average weekly nowcasts for the sample from July 2019 to August 2022, along with the realized inflation values. The nowcasts are compared to those obtained with the naive AR(2) benchmark.

The figure shows that all the nowcasting model alternatives can closely follow the actual inflation rate, providing policymakers with valuable real-time information about inflation.

Figure 4.6: Nowcasts: Lasso-VAR vs. FAVAR, BVAR and AR(2)



To give an idea of the relative nowcasting abilities of the three model competitors, Table 4.1 summarizes their RMSFEs (and that of the naive AR(2) model benchmark) in the three sub-samples. The table also reports, with bold values, whether the results from the model's predictive accuracy comparison tests are significant. These tests are based on the corrected Diebold and Mariano (1995) statistics (Harvey et al., 1997), taking the Lasso-VAR as a reference.

Table 4.1: Model comparisons: RMSFE and Diebold-Mariano test results

		pre-COVID-19 sample	COVID-19 sample	Energy crisis sample	Full sample
h=4	Lasso-VAR	0.144	0.239	0.379	0.261
	FAVAR	**0.153**	**0.262**	0.401	**0.281**
	BVAR	**0.152**	0.235	0.364	0.256
	AR(2)	**0.200**	**0.386**	**0.546**	**0.407**

Notes: Bold values indicate that the predictive performance of the Lasso VAR is statistically different from those obtained by the other models. The reference test is the Diebold and Mariano (1995) Diebold and Mariano (1995) test of differences in the model's predictive accuracy. The reference statistics consider the (Harvey et al., 1997) correction for small samples.

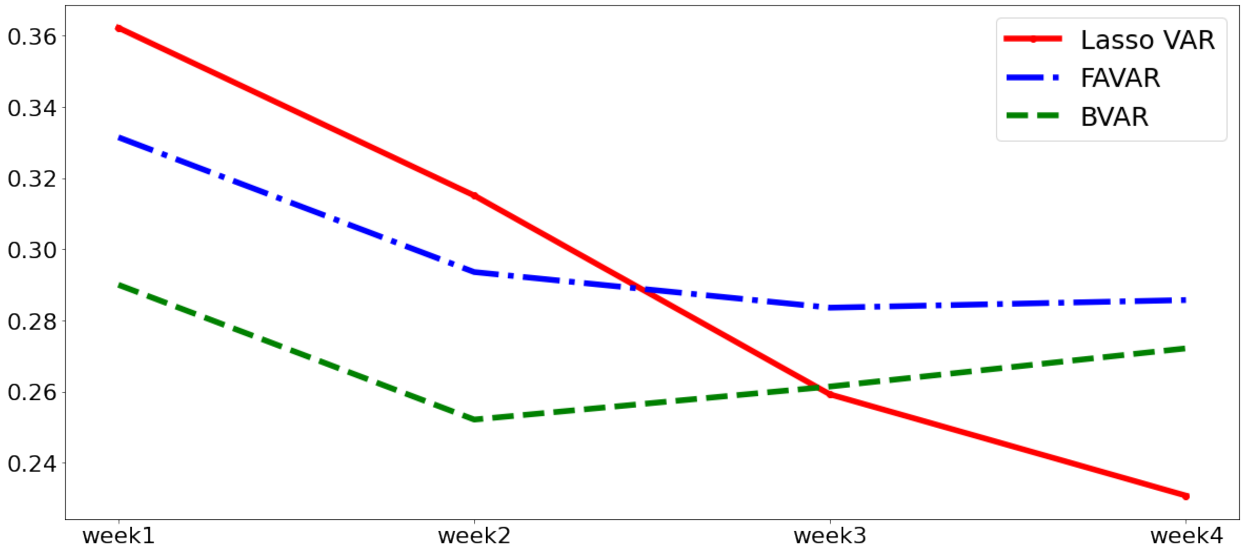
The table shows that the three nowcasting alternatives have similar predictive properties, with the Lasso-VAR significantly outperforming the FAVAR and the AR(2) benchmark in the total nowcast sample. In "normal times", the Lasso-VAR outperforms all the alternative methods, reaching the minimum RMSFE across model competitors. With the increase of the HICP volatility registered during the COVID-19 and the energy crisis periods, the BVAR marginally outperformed the Lasso-VAR in terms of RMSFE (the difference in the predictive performances is not statistically significant). The latter, however, continues to perform significantly better than the FAVAR and the AR(2). The relative decrease in the nowcasting abilities of the Lasso-VAR in the heightened price volatility periods is entirely due to a loss of accuracy in two specific episodes. One between February 2021 and March 2021 and one just before May 2022 (see Figures 4.3, 4.4, and 4.6). We speculate that the Lasso shrinkage (which sets to exactly zero some of the VAR coefficients) might become inaccurate when some sudden shocks hit the economy (possibly the COVID-19 shock in February-March 2021 and the energy prices turmoils from the Russian-Ukrainian war in May 2022). The three nowcasting tools outperform the AR(2) model, whose RMSFE is significantly higher by 45% to 59%.

A further comparison of the models' nowcasting properties can be obtained from the average RMSFEs for the weekly point HICP inflation nowcasts in the month. Figure 4.7 replicates the information included in Figure 4.5 for the Lasso-VAR (and for the AR(2) benchmark), adding the same information for the FAVAR and the BVAR. The performance of the three models is, on average, similar. For the Lasso-VAR and the FAVAR, it shows monotonic improvements in accuracy as new information becomes available during the month. The Lasso-VAR attains the minimum (and least) average RMSFE at the end of the month. The BVAR attains its minimum average RMSFE in the second week of the month.

4.5 Robustness checks

The robustness of the results described in the previous section can be evaluated in several ways. Here, we focus on four significant aspects of the analysis: *i*) the extension of the information set to the inclusion of high-frequency (weekly) and standard-frequency (monthly) policy variables; *ii*) the extension of the forecasting window up to 12 weeks (approximately one quarter); *iii*) the use of alternative machine-learning-based regularization tools in the class of linear models; *iv)* the

Figure 4.7: Models' nowcasting performance per week of the month



Notes: The line denotes average monthly RMSFEs for weekly point HICP inflation nowcasts for the sample July 2019 - August 2022.

application of the Lasso-VAR procedure to perform core inflation nowcasts.

The first robustness check enriches the information set by including the monetary policy rates. Three of six policy rates are observed at the highest frequency (daily, thus moved to weekly): the ECB's Marginal lending facility rate, the Deposit facility rate and the Main refinancing operations rate. The other rates are available at a monthly frequency: the ECB's shadow interest rate (Wu and Xia, 2020), the US Federal Funds rate and its shadow rate (Wu and Xia, 2016).[14] With this data extension, the number of variables (thus equations) considered in the VARs increases from 31 to 37. The missing data imputing procedures and the dynamic model specifications are fixed to those used for the no-policy variables model estimates. Results, summarized in Table H.1 in Appendix H, show that including policy variables does not alter the near-term (four weeks ahead) models' predictive properties. This result is constant across model alternatives.

With the second check, we verify whether our methodology, specifically designed for very short-term

[14]The shadow rates are included in the model as a measure of the monetary policy stance during the zero-lower-bound periods. They are defined as the monthly interest rates implied by a multi-factor shadow rate term structure (yield curve) model. Wu and Xia's (Wu and Xia, 2020, 2016) rates are obtained as a linear combination of three latent factors following a VAR(1) process. An extended Kalman filter estimates the latent factors and the shadow rate. The main characteristic of the shadow rate is that it is not bounded to zero, while it equals the policy rate when this is above its lower bound (0.25 % for the US and 0 for the Eurozone).

forecasts, maintains its predictive properties even at larger forecasting horizons, i.e. moving from a nowcasting analysis to a forecasting analysis. We also consider eight and 12-week-ahead forecasts along with the four-week ahead forecasts. Results are summarized in Table H.2 in Appendix H. Unsurprisingly, the predictive abilities of the Lasso-VAR, the FAVAR and the BVAR, summarized in the values of the RMSFE, worsen with the size of the forecasting periods. There are, however, signals that the predictive abilities of the Lasso-VAR tend to be outperformed by the alternative methods when the forecast period is extended, with the BVAR prevailing on the other model shrinkage methods.

The third robustness check verifies whether we can improve the nowcasting properties by considering alternative regularization methods for the VAR. In the class of linear models, the Ridge and the Elastic Net estimators are "natural" machine learning competitors for the Lasso. The former replaces a quadratic penalty to the Lasso coefficients' mass, i.e. it relies on the following optimization problem:

$$\hat{\mathbf{A}}(\lambda) = \arg\min_{\mathbf{A}} \frac{1}{T}\|\mathbf{AZ} - \mathbf{Y}\|_2^2 + \lambda\|\mathbf{A}\|_2, \tag{4.5}$$

By penalizing the sum of squared coefficients (the so-called L2 penalty) instead of the sum of their absolute values (L1 penalty) as in the Lasso, the VAR coefficients are shrunk but not set to exactly zero.

The Elastic Net estimator combines the L1 and L2 penalties to minimize the following loss function:

$$\hat{\mathbf{A}}(\lambda) = \arg\min_{\mathbf{A}} \frac{1}{2T}\|\mathbf{AZ} - \mathbf{Y}\|_2^2 + \lambda(\frac{1-\alpha}{2}\ \|\mathbf{A}\|_2 + \alpha\|\mathbf{A}\|_1), \tag{4.6}$$

where α is the mixing parameter between Ridge (α=0) and Lasso (α=1). The Ridge/Lasso penalty parameter λ is cross-validated over five-time series folds, with α calibrated over a grid of values between 0 and 1.

Results, summarized in Table H.3 in Appendix H, show that, for our nowcasting sample and the 31 variables dataset, the Lasso-VAR ensures a lower RMSFE than the Ridge and Elastic Net-regularized

VARs. This result holds irrespective of the particular sub-sample being considered.

With the fourth robustness check, we verify whether the predictive abilities of the different high-dimensional VAR methods are also confirmed for Euro area core inflation. This variable is within the information set used for the estimates. The nowcast graph and the related RMSFEs, depicted in Figure H.1 and Table H.3 of Appendix H, respectively, show that the tested VAR-based methodologies perform very well in nowcasting core inflation in "normal" times (the pre-COVID-19 period) while worsening in the COVID-19 period. This result is likely related to the specificity of the information set employed for the estimates, in which 10 out of 31 variables are energy prices, the most important predictors of the increase in HICP price variability. The models' predictive performances denote an improvement in the following energy crisis period, signalling that the rise of energy prices is increasingly embedded in the other components of the price level, thus augmenting their predictive content for core inflation.

4.6 Conclusions

Nowcasting tools are becoming increasingly popular in real-time predicting macroeconomic aggregates such as industrial production, gross domestic product and inflation, particularly within the central banks' research offices.

This work contributes to the nowcasting approach by evaluating the performances of the Least absolute shrinkage and selection operator Vector Auto Regression (Lasso-VAR) in the near-term prediction of aggregate Euro area inflation. The Lasso-VAR performances are compared to well-established model shrinkage strategies adopted in high dimensional, mixed frequency data settings; the Factor Augmented Vector Auto Regression (FAVAR) and the Bayesian Vector Auto Regression (BVAR) models. Emerging literature shows that the Lasso strategy efficiently handles dimensionality reduction, generating the sparsity through which the resulting adequately-fitted VAR can outperform both FAVARs models and BVARs in high-dimensional settings.

We merge real-time high-frequency data and standard-frequency data released by official sources.

We describe the different stages of the analysis, from the imputation of the missing high-frequency data to the estimation of the sparse structure and the comparative evaluation of the model's forecasting performances.

The modelling approach being proposed performed relatively well as a nowcasting tool. We show that the Lasso-VAR can closely follow the actual inflation rate and effectively handle real-time information. The forecasting accuracy improves as more high-frequency data become available over time. In "normal times" environments (pre-COVID-19 sample), the Lasso-VAR outperforms the alternative methods, reaching the lowest forecast error across model competitors. With the recent increase in the price volatility registered during the COVID-19 and the energy crisis periods, the Lasso-VAR cannot significantly outperform the BVAR, even if it continues to perform better than the FAVAR and the naive AR(2) model benchmark. For the complete nowcast sample considered in our study (July 2019 - August 2022), the Lasso-VAR continues to outperform the FAVAR and the AR(2) benchmark, and its nowcasting ability remains statistically aligned with that of the BVAR. These results suggest that machine learning-based model shrinkage methods provide a valid and efficient alternative to well-established methods used in nowcasting. A possible advantage is that they can handle high-dimensional information sets, a feature that becomes increasingly appealing with the availability of real-time information. That could significantly improve the forecasting abilities of these methods. The inclusion of real-time "soft" information to detect the drivers of nowcast revisions, the application of the proposed approach to nowcast country-specific information and inflation components are possible avenues for future research.

Bibliography

A. Addiga and S. Bagui. *Journal of Computer and Communications*, 2022.

A. Alessa and M. Faezipour. *Tweet Classification Using Sentiment Analysis Features and TF-IDF Weighting for Improved Flu Trend Detection*, pages 174–186. 07 2018. ISBN 978-3-319-96135-4. doi: 10.1007/978-3-319-96136-1_15.

A. Algaba, D. Ardia, K. Bluteau, S. Borms, and K. Boudt. Econometrics meets sentiment: An overview of methodology and applications. *Journal of Economic Surveys*, 34(3):512–547, 2020.

C. Angelico, J. Marcucci, M. Miccoli, and F. Quarta. Can we measure inflation expectations using twitter? *Journal of Econometrics*, 228(2):259–277, 2022. ISSN 0304-4076. doi: https://doi.org/10.1016/j.jeconom.2021.12.008. URL https://www.sciencedirect.com/science/article/pii/S0304407622000227.

V. Aprigliano and L. Bencivelli. Ita-coin: a new coincident indicator for the italian economy. *Bank of Italy Temi di Discussione (Working Paper) No*, 935, 2013.

V. Aprigliano, S. Emiliozzi, G. Guaitoli, A. Luciani, J. Marcucci, and L. Monteforte. The power of text-based indicators in forecasting the Italian economic activity. Temi di discussione (Economic working papers) 1321, Bank of Italy, Economic Research and International Relations Area, Mar. 2021. URL https://ideas.repec.org/p/bdi/wptemi/td_1321_21.html.

S. B. Aruoba and T. Drechsel. Identifying monetary policy shocks: A machine learning approach. 2022.

S. B. Aruoba, F. X. Diebold, and C. Scotti. Real-time measurement of business conditions. *Journal of Business & Economic Statistics*, 27(4):417–427, 2009.

S. Athey, J. Tibshirani, and S. Wager. Generalized random forests. *The Annals of Statistics*, 47(2): 1148–1178, 2019.

A. Azqueta-Gavaldón. Developing news-based economic policy uncertainty index with unsupervised machine learning. *Economics Letters*, 158:47–50, 2017.

A. Azqueta-Gavaldon, D. Hirschbühl, L. Onorante, and L. Saiz. Economic policy uncertainty in the euro area: an unsupervised machine learning approach. *Available at SSRN 3516756*, 2020.

A. Baevski, Y. Zhou, A. Mohamed, and M. Auli. wav2vec 2.0: A framework for self-supervised learning of speech representations. *Advances in Neural Information Processing Systems*, 33:12449–12460, 2020.

S. R. Baker, N. Bloom, and S. J. Davis. Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636, 2016.

M. Bańbura and L. Saiz. Short-term forecasts of economic activity in the euro area. *Monthly Bulletin*, (April):69–74, 2008.

M. Bańbura, D. Giannone, and L. Reichlin. Large bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92, 2010.

M. Bańbura, D. Giannone, M. Modugno, and L. Reichlin. Now-casting and the real-time data flow. In *Handbook of Economic Forecasting*, volume 2, pages 195–237. Elsevier, 2013.

O. Bandiera, A. Prat, S. Hansen, and R. Sadun. Ceo behavior and firm performance. *Journal of Political Economy*, 128(4):1325–1369, 2020.

S. Basu and G. Michailidis. Regularized estimation in sparse high-dimensional time series models. *The Annals of Statistics*, 43(4):1535–1567, 2015.

A. Belloni, V. Chernozhukov, and C. Hansen. High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2):29–50, 2014.

A. Belloni, V. Chernozhukov, I. Fernández-Val, and C. Hansen. Program evaluation and causal inference with high-dimensional data. *Econometrica*, 85(1):233–298, 2017.

S. A. Brave, R. A. Butters, D. Kelley, et al. A new" big data" index of us economic activity,". *Economic Perspectives, Federal Reserve Bank of Chicago*, 1, 2019.

J. Breitung and C. Roling. Forecasting inflation rates using daily data: A nonparametric midas approach. *Journal of Forecasting*, 34(7):588–603, 2015.

T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

F. Canova and F. Ferroni. Multiple filtering devices for the estimation of cyclical dsge models. *Quantitative Economics*, 2(1):73–98, 2011.

S. M. Carta, S. Consoli, L. Piras, A. S. Podda, and D. R. Recupero. Explainable machine learning exploiting news and domain-specific lexicon for stock market forecasting. *IEEE Access*, 9:30193–30205, 2021.

D. Cascaldi-Garcia, T. RT Ferreira, D. Giannone, and M. Modugno. Back to the present: Learning about the euro area through a now-casting model. *International Finance Discussion Paper*, (1313), 2021.

A. Chalfin, O. Danieli, A. Hillis, Z. Jelveh, M. Luca, J. Ludwig, and S. Mullainathan. Productivity and selection of human capital with machine learning. *American Economic Review*, 106(5):124–27, 2016.

B. Cheela, A. DeHon, J. Fernández-Villaverde, and A. Peri. Programming fpgas for economics: An introduction to electrical engineering economics. Technical report, National Bureau of Economic Research, 2022.

V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins. Double/debiased machine learning for treatment and structural parameters, 2018.

J. Cimadomo and A. D'Agostino. Combining time variation and mixed frequencies: An analysis of government spending multipliers in italy. *Journal of Applied Econometrics*, 31(7):1276–1290, 2016.

J. Cimadomo, D. Giannone, M. Lenza, F. Monti, and A. Sokol. Nowcasting with large bayesian vector autoregressions. *Journal of Econometrics*, 231(2):500–519, 2022.

B. Das and S. Chakraborty. An improved text sentiment classification model using TF-IDF and next word negation. *CoRR*, page 6, 2018.

M. J.-F. Dauphin, M. K. Dybczak, M. Maneely, M. T. Sanjani, M. N. Suphaphiphat, Y. Wang, and H. Zhang. *Nowcasting GDP-A Scalable Approach Using DFM, Machine Learning and Novel Data, Applied to European Economies*. International Monetary Fund, 2022.

G. J. de Bondt, E. Hahn, and Z. Zekaite. Alice: Composite leading indicators for euro area inflation cycles. *International Journal of Forecasting*, 37(2):687–707, 2021.

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

G. Di Filippo. Dynamic model averaging and cpi inflation forecasts: A comparison between the euro area and the united states. *Journal of Forecasting*, 34(8):619–648, 2015.

F. X. Diebold. Real-time real economic activity: Exiting the great recession and entering the pandemic recession. Technical report, National Bureau of Economic Research, 2020.

X. Diebold and S. Mariano. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 1995.

D. Donaldson and A. Storeygard. The view from above: Applications of satellite data in economics. *Journal of Economic Perspectives*, 30(4):171–98, 2016.

J. Durbin, S. Koopman, et al. An efficient and simple simulation smoother for state space time series analysis. Technical report, Society for Computational Economics, 2001.

A. Dvorák and V. Novák. Towards automatic modeling of economic texts. *Mathware and Soft Computing*, 14(3):217–231, 2007.

B. Eraker, C. W. Chiu, A. T. Foerster, T. B. Kim, and H. D. Seoane. Bayesian mixed frequency vars. *Journal of Financial Econometrics*, 13(3):698–721, 2014.

M. H. Farrell. Robust inference on average treatment effects with possibly more covariates than observations. *Journal of Econometrics*, 189(1):1–23, 2015.

J. Fernández-Villaverde, S. Hurtado, and G. Nuno. Financial frictions and the wealth distribution. Technical report, National Bureau of Economic Research, 2019.

J. Fernandez-Villaverde, G. Nuno, G. Sorg-Langhans, and M. Vogler. Solving high-dimensional dynamic programming problems using deep learning. *Unpublished working paper*, 2020.

C. Foroni, M. Marcellino, and C. Schumacher. Unrestricted mixed data sampling (midas): Midas regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pages 57–82, 2015.

J. Friedman, T. Hastie, and R. Tibshirani. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1, 2010.

L. Gao and P. A. Beling. Machine quantification of text-based economic reports for use in predictive modeling. In *SMC'03 Conference Proceedings. 2003 IEEE International Conference on Systems,* *Man and Cybernetics. Conference Theme-System Security and Assurance (Cat. No. 03CH37483)*, volume 4, pages 3536–3541. IEEE, 2003.

M. Gentzkow and B. T. Kelly. Text As Data. *SSRN Electronic Journal*, 2017. doi: 10.2139/ssrn.2934001. URL `http://dx.doi.org/10.2139/ssrn.2934001`.

M. Gentzkow, B. Kelly, and M. Taddy. Text as data. *Journal of Economic Literature*, 57(3):535–74, 2019.

A. Gholami, Z. Yao, S. Kim, M. W. Mahoney, and K. Keutzer. Ai and memory wall. *RiseLab Medium Post*, 2021.

D. Giannone, L. Reichlin, and D. Small. Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676, 2008.

P. Goulet Coulombe, M. Leroux, D. Stevanovic, and S. Surprenant. How is machine learning useful for macroeconomic forecasting? *Journal of Applied Econometrics*, 37(5):920–964, 2022.

S. Hansen and M. McMahon. Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99:S114–S133, 2016.

S. Hansen, M. McMahon, and A. Prat. Transparency and deliberation within the fomc: a computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870, 2018.

J. Hartford, G. Lewis, K. Leyton-Brown, and M. Taddy. Deep iv: A flexible approach for counterfactual prediction. In *International Conference on Machine Learning*, pages 1414–1423. PMLR, 2017.

D. Harvey, S. Leybourne, and P. Newbold. Testing the equality of prediction mean squared errors. *International Journal of forecasting*, 13(2):281–291, 1997.

M. H. Hassoun et al. *Fundamentals of artificial neural networks*. MIT press, 1995.

P. He, X. Liu, J. Gao, and W. Chen. Deberta: Decoding-enhanced bert with disentangled attention. *arXiv preprint arXiv:2006.03654*, 2020.

J. V. Henderson, A. Storeygard, and D. N. Weil. Measuring economic growth from outer space. *American economic review*, 102(2):994–1028, 2012.

P. C. Higgins. Gdpnow: A model for gdp'nowcasting'. 2014.

G. Hoberg and V. Maksimovic. Redefining financial constraints: A text-based analysis. *The Review of Financial Studies*, 28(5):1312–1352, 2015.

K. Hubrich and F. Skudelny. Forecast combination for euro area inflation: a cure in times of crisis? *Journal of Forecasting*, 36(5):515–540, 2017.

K. S. Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 1972.

M. E. Kahou, J. Fernández-Villaverde, J. Perla, and A. Sood. Exploiting symmetry in high-dimensional dynamic programming. Technical report, National Bureau of Economic Research, 2021.

E. Kalamara, A. Turrell, C. Redl, G. Kapetanios, and S. Kapadia. Making Text Count: Economic Forecasting Using Newspaper Text. *SSRN Electronic Journal*, 2020. doi: 10.2139/ssrn.3610770. URL `http://dx.doi.org/10.2139/ssrn.3610770`.

D. Kapetanios, Buono, G. Mazzi, Gian Luigi, M. Marcellino, and F. Papailias. Big data types for macroeconomic nowcasting. *Eurostat Review on National Accounts and Macroeconomic Indicators*, 1(2017):93–145, 2017.

S.-j. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky. An interior-point method for large-scale 1-regularized logistic regression. In *Journal of Machine Learning Research*. Citeseer, 2007.

J. Kleinberg, J. Ludwig, S. Mullainathan, and Z. Obermeyer. Prediction policy problems. *American Economic Review*, 105(5):491–95, 2015.

J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan. Human decisions and machine predictions. *The quarterly journal of economics*, 133(1):237–293, 2018.

G. Koop and L. Onorante. Macroeconomic nowcasting using google probabilities. In *Topics in Identification, Limited Dependent Variables, Partial Observability, Experimentation, and Flexible Modeling: Part A*. Emerald Publishing Limited, 2019.

V. Kuzin, M. Marcellino, and C. Schumacher. Midas vs. mixed-frequency var: Nowcasting gdp in the euro area. *International Journal of Forecasting*, 27(2):529–542, 2011.

J. Lin and G. Michailidis. Regularized estimation and testing for high-dimensional multi-block vector-autoregressive models. *Journal of Machine Learning Research*, 18, 2017.

H. P. Luhn. The automatic creation of literature abstracts. *IBM Journal of research and development*, 2(2):159–165, 1958.

M. Marcellino. A linear benchmark for forecasting gdp growth and inflation? *Journal of Forecasting*, 27(4):305–340, 2008.

R. S. Mariano and Y. Murasawa. A coincident index, common factors, and monthly real gdp. *Oxford Bulletin of economics and statistics*, 72(1):27–46, 2010.

T. Matheson. New indicators for tracking growth in real time. *IMF Working Papers*, 11/43, 2011.

M. C. Medeiros, G. F. Vasconcelos, Á. Veiga, and E. Zilberman. Forecasting inflation in a data-rich environment: the benefits of machine learning methods. *Journal of Business & Economic Statistics*, 39(1):98–119, 2021.

J. W. Messner and P. Pinson. Online adaptive lasso estimation in vector autoregressive models for high dimensional wind power forecasting. *International Journal of Forecasting*, 35(4):1485–1498, 2019.

H. U. Miranda-Belmonte, V. Muñiz-Sánchez, and F. Corona. Word embeddings for topic modeling: an application to the estimation of the economic policy uncertainty index. *Expert Systems with Applications*, 211:118499, 2023.

M. Modugno. Nowcasting inflation using high frequency data. Working Paper Series 1324, European Central Bank, 2011. URL `https://ideas.repec.org/p/ecb/ecbwps/20111324.html`.

N. Naik, R. Raskar, and C. A. Hidalgo. Cities are physical too: Using computer vision to measure the quality and impact of urban appearance. *American Economic Review*, 106(5):128–32, 2016.

N. Naik, S. D. Kominers, R. Raskar, E. L. Glaeser, and C. A. Hidalgo. Computer vision uncovers predictors of physical urban change. *Proceedings of the National Academy of Sciences*, 114(29): 7571–7576, 2017.

A. Ng and M. Jordan. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Advances in neural information processing systems*, 14, 2001.

W. B. Nicholson, D. S. Matteson, and J. Bien. Varx-l: Structured regularization for large vector autoregressions with exogenous variables. *International Journal of Forecasting*, 33(3):627–651, 2017.

W. B. Nicholson, I. Wilms, J. Bien, and D. S. Matteson. High dimensional forecasting via interpretable vector autoregression. *Journal of Machine Learning Research*, 21(166):1–52, 2020.

J. Qian, H. Wang, Z. Li, S. Li, and X. Yan. Limitations of language models in arithmetic and symbolic induction. *arXiv preprint arXiv:2208.05051*, 2022.

A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, et al. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9, 2019.

J. Ramos. Using tf-idf to determine word relevance in document queries. page 4, 2003.

A. Richardson, T. van Florenstein Mulder, and T. Vehbi. Nowcasting gdp using machine-learning algorithms: A real-time assessment. *International Journal of Forecasting*, 37(2):941–948, 2021.

F. Schorfheide and D. Song. Real-time forecasting with a mixed-frequency var. *Journal of Business & Economic Statistics*, 33(3):366–380, 2015.

C. A. Sims. Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, pages 1–48, 1980.

S. Smyl. A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, 36(1):75–85, 2020.

M. Song and K. Shin. Forecasting economic indicators using a consumer sentiment index: Survey-based versus textbased data. *Journal of Forecasting*, apr 7 2019. doi: 10.1002/for.2584. URL `http://dx.doi.org/10.1002/FOR.2584`.

I. Steinwart and A. Christmann. *Support vector machines.* Springer Science & Business Media, 2008.

J. H. Stock and M. W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179, 2002a.

J. H. Stock and M. W. Watson. Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162, 2002b.

Sutton, Richard S and Barto. *Reinforcement learning: An introduction.* MIT press, 2018.

C. K. Theil, S. Štajner, and H. Stuckenschmidt. Word embeddings-based uncertainty detection in financial disclosures. In *Proceedings of the First Workshop on Economics and Natural Language Processing*, pages 32–37, 2018.

C. K. Theil, S. Štajner, and H. Stuckenschmidt. Explaining financial uncertainty through specialized word embeddings. *ACM Transactions on Data Science*, 1(1):1–19, 2020.

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

J. C. Wu and F. D. Xia. Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291, 2016.

J. C. Wu and F. D. Xia. Negative interest rate policy and the yield curve. *Journal of Applied Econometrics*, 35(6):653–672, 2020.

Z. Yun-tao, G. Ling, and W. Yong-cheng. An improved tf-idf approach for text classification. *Journal of Zhejiang University-Science A*, 6(1):49–55, 2005.

S. Zheng, A. Trott, S. Srinivasa, D. C. Parkes, and R. Socher. The ai economist: Taxation policy design via two-level deep multiagent reinforcement learning. *Science advances*, 8(18):eabk2607, 2022.

Appendix A

Data sample

Table A.1 presents a selection of sentences from our dataset and their corresponding scores, which have been divided into three categories: positive, negative, and neutral. The scores are based on the sentiment expressed in each sentence and were determined using the transformers model and natural language inference techniques. By examining this table, it is possible to see the diversity of sentiments expressed in the articles and how the proposed method is able to accurately capture and quantify them. Please note that this is just simple excercise showing performance on single sentences. Our proposed framework consider all articles published in one day as parts of unique space on which embedding is calculated.

Table A.1: Raw count of total articles and articles included

text	Positive	Negative	Neutral
Tra vent'anni 6,8 milioni di lavoratori in meno, 3,8 milioni di pensionati in pi l'italia in recessione demografica.	0.130711898	-0.77282989	0.096458204
Prezzi, dal burro alle uova ecco la stangata sul carrello della spesa.	0.365407795	-0.549198389	0.085393801
Il pil italiano accelera a +1% nel secondo trimestre.	0.955184698	-0.015128067	0.029687295
A settembre il bonus trasporti 60 euro per gli abbonamenti ai mezzi pubblici.	0.754406989	-0.092518248	0.153074756
Crisi di governo, le dimissioni di draghi e quando a dimettersi fu charles de gaulle.	0.023530997	-0.930478156	0.045990806
Le navi del grano ferme a odessa.	0.013422744	-0.952030838	0.034546442
SPcabbassa le prospettive sull'italia riforme a rischio, allarme dell'fmi sul 2023.	0.017197166	-0.89275378	0.090049036
Guerra, inflazione, gas, fmi il mondo e' sull'orlo della recessione, rischio disordini sociali.	0.014254248	-0.966132343	0.019613372
I timori del financial times dopo l'addio di draghi in gioco le risorse del recovery fund.	0.018940613	-0.933990002	0.047069423
Un decreto aiuti bis da 13 miliardi per prorogare gli sconti sulle bollette .	0.817359448	-0.075284727	0.107355826
Gas, nuovo taglio al nord stream 1 il flusso scende al 20%, il prezzo vola.	0.071652457	-0.912863612	0.015483935

Appendix B

Technical Box: BERT Model - Detailed Architecture

The BERT (Bidirectional Encoder Representations from Transformers) model is a sophisticated architecture designed for natural language processing tasks. Here are the key components of the BERT architecture:

- Input Embeddings: The input to BERT is a sequence of tokens, which are first converted into vectors and then processed in the neural network. BERT uses WordPiece embeddings with a 30,000 token vocabulary.
- Positional Encoding: To account for the order of words in a text, BERT includes positional encodings added to the input embeddings. These positional encodings are learnable parameters that allow the model to understand the relative positions of words in a sentence.
- Transformer Encoders: The core of BERT's architecture is a set of Transformer encoders. A Transformer encoder reads the entire sequence of words at once, which is a significant departure from previous models that read input sequences sequentially. This global perspective allows BERT to use the context from both the left and the right of a word in understanding its meaning.

- Self-Attention Mechanism: Within each Transformer encoder, the self-attention mechanism allows the model to weigh the importance of words in the sequence. In essence, it allows the model to focus on different words when processing each word in the sentence. This mechanism is particularly useful in understanding the context and eliminating ambiguity in natural language.

- Feed-Forward Neural Networks: Each Transformer encoder also contains a feed-forward neural network, applied identically to each position. This network consists of two linear transformations with a ReLU activation in between.

- Output: The final hidden state corresponding to the special [CLS] token is used as the aggregate sequence representation for classification tasks. For each token in the sequence, the output is the corresponding hidden state from the final layer of the Transformer encoders.

BERT's architecture, with its bidirectional approach and fine-tuning capability, makes it highly versatile and powerful for a wide range of NLP tasks, outperforming many existing models.

Appendix C

Technical Box: DeBERTa Model - Detailed Architecture

DeBERTa (Decoding-enhanced BERT with Disentangled Attention) is an advanced variant of the BERT model, designed to improve the performance of natural language processing tasks. Here are the key components of the DeBERTa architecture:

- Input Embeddings: Similar to BERT, DeBERTa takes a sequence of tokens as input. These tokens are converted into vectors using WordPiece embeddings, which are then processed in the neural network.

- Positional Encoding: DeBERTa also includes positional encodings to account for the order of words in a text. These encodings are added to the input embeddings and are learnable parameters, allowing the model to understand the relative positions of words in a sentence.

- Disentangled Attention Mechanism: The core innovation in DeBERTa is the disentangled attention mechanism. Unlike BERT, which uses a single attention mechanism, DeBERTa disentangles the content and position to compute attention scores. This disentanglement allows the model to capture more nuanced dependencies between words and their contexts.

- Enhanced Mask Decoder: DeBERTa introduces an enhanced mask decoder that includes a casual mask in the self-attention mechanism. This mask allows the model to use the context from the left of a word but not the right, which is particularly useful for language modeling tasks.

- Transformer Encoders: DeBERTa uses a set of Transformer encoders, similar to BERT. However, due to the disentangled attention mechanism and the enhanced mask decoder, these encoders can capture more complex and nuanced relationships between words and their contexts.

- Output: The final hidden state corresponding to the special [CLS] token is used as the aggregate sequence representation for classification tasks. For each token in the sequence, the output is the corresponding hidden state from the final layer of the Transformer encoders.

DeBERTa's architecture, with its disentangled attention mechanism and enhanced mask decoder, makes it a powerful tool for a wide range of NLP tasks, often outperforming BERT and other existing models.

Appendix D

Trending words

This appendix provides analysis of word usage across five prominent newspapers. Our analysis aims to gain a deeper understanding of how these newspapers covered two major global events: the Euro crisis and the Coronavirus pandemic. To visually showcase our findings, we have created word clouds for each of the five newspapers, with a separate plot for each event (Figures D.1 to D.5). Our word clouds clearly show noticeable differences in the language used to describe these two crises, not just within the same newspaper but also between different newspapers. These differences offer us valuable insights into the distinct perspectives and framing of these events by the media. Interestingly, the way the Euro crisis was represented in the media was much more uniform than the varying approaches used to cover the Coronavirus pandemic. That highlights the media's critical role in shaping public perception and understanding of events. Please note that wordclouds are simple frequencies.

(a) Euro crisis

(b) Coronavirus Pandemic

Figure D.1: Corriere della Sera

(a) Euro crisis

(b) Coronavirus Pandemic

Figure D.2: Il Fatto quotidiano

(a) Euro crisis

(b) Coronavirus Pandemic

Figure D.3: La Stampa

(a) Euro crisis

(b) Coronavirus Pandemic

Figure D.4: I Sole 24 Ore

(a) Euro crisis

(b) Coronavirus Pandemic

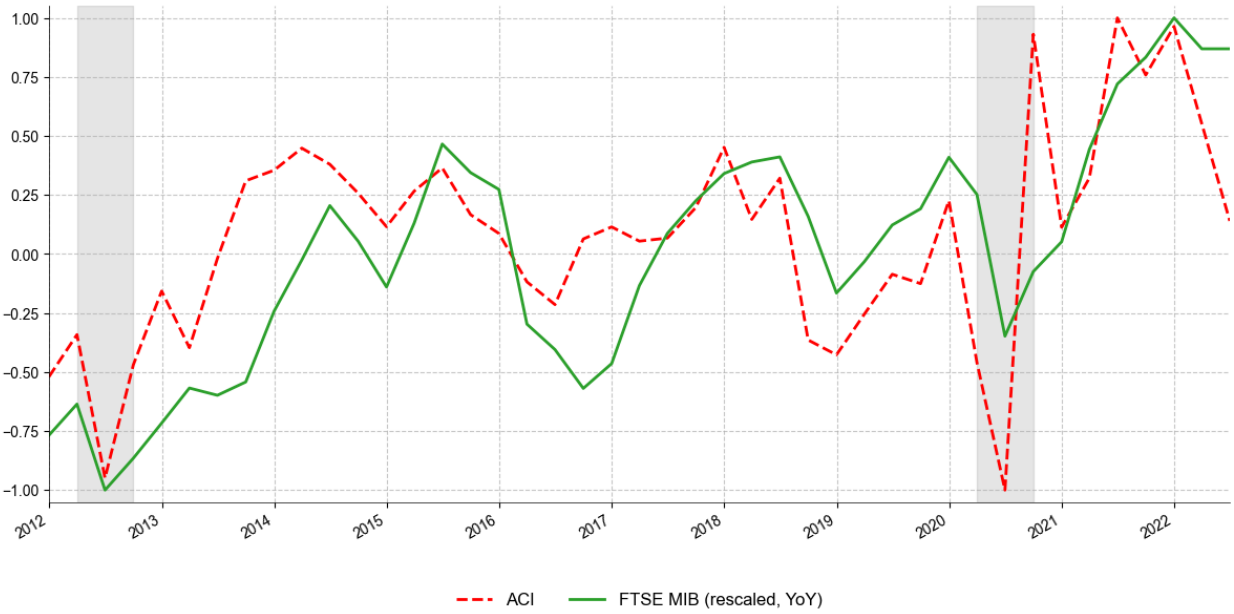
Figure D.5: La Repubblica

Appendix E

Additional results: Augmented Coincident Index

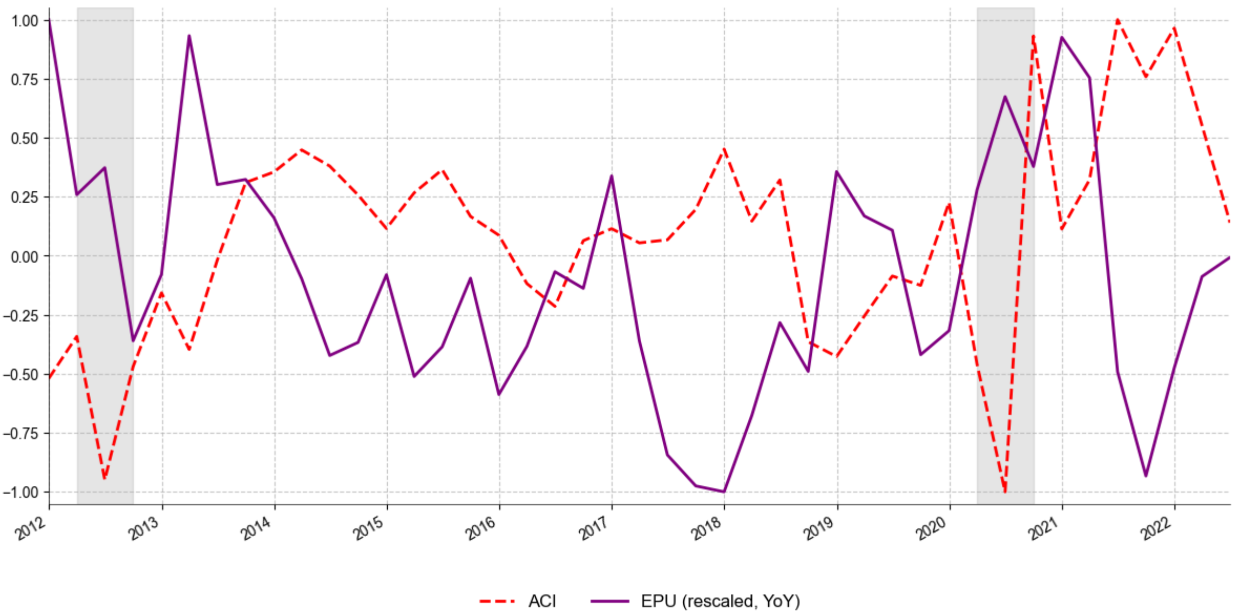
Here we present additional results in the figures below. Figure E.1 displays the plot of the stock exchange index FTSE MIB, which exhibits a positive correlation of 0.658 with our index of real activity (gap growth). Conversely, Figure E.2 illustrates the Economic Policy Uncertainty (EPU) index, demonstrating a negative correlation of -0.426 with our index. The negative correlation with the EPU index can be attributed to the fact that not all economic policy uncertainty translates directly into changes in real activity.

Figure E.1: Italian stock exchange index - FTSE MIB



Note: Italian stock exchange index resampled at quarterly frequency.

Figure E.2: Economic policy uncertainty - Italy



Note: Economic policy uncertainty index resampled at quarterly frequency. For more details see Baker et al. (2016)

Appendix F

Dataset

We provide a summary of the dataset used in the analyses. Table F.1 describes the time series data. Information about their frequency and the source and codes for retrieving the data are provided. Table F.2 summarizes the results of the Phillips-Perron nonstationarity tests. The tests are performed on the series entered in the VARs, thus considering differenced price levels (inflation), while interest rates are tested in levels.

Table F.1: Philips-Perron Tests for nonstationarity

Variable name	Test Statistic	p-value	Variable name	Test Statistic	p-value
HICP	-7.9074	0.05	Brent	-18.8200	0.00
HICP - Processed food incl. alcohol and tobacco	-4.4884	0.00	Power - ITA	-30.6786	0.00
HICP - Unprocessed food	-7.6619	0.00	Power - FRA	-37.6345	0.00
HICP - Industrial goods excluding energy	-7.7836	0.00	Power - DE	-39.2505	0.00
HICP - Energy	-6.4242	0.00	Power - ESP	-28.4744	0.00
HICP - Services	-9.2710	0.00	USD/EUR	-17.0097	0.00
Food, import weighted	-7.3260	0.733	YEN/EUR	-18.2911	0.00
Non food, import weighted	-6.9617	0.696	GBP/EUR	-18.1924	0.00
Total non-energy commodity, import weighted	-7.1914	0.719	RMB/EUR	-17.1134	0.00
Food, use weighted	-7.2363	0.724	Nominal effective exchange rate, eurozone	-7.0695	0.00
Non-food, use weighted	-6.9343	0.693	Real effective exchange rate, eurozone	-7.2795	0.00
Total non-energy commodity, use weighted	-7.1080	0.711	Marginal lending facility	-20.7241	0.02
Global Natural Gas	-6.3165	0.00	Deposit facility	-22.2143	0.03
EUA CO2	-24.3585	0.00	Main refinancing operations	-19.9126	0.02
Euro Super 95	-15.3554	0.00	Shadow Rate ECB	-8.2463	0.00
Diesel	-14.6733	0.00	Shadow rate US	-8.1027	0.02
Gas oil	-18.4698	0.00	Federal Funds Rate	-5.3728	0.00
Heating Oil	-19.8784	0.00	Monetary aggregate M3	-7.6345	0.00
LPG	-13.1760	0.00			

Table F.2: Summary of data, frequency of observations and source

Variable name	frequency	source	code
HICP	Monthly	ECB	ICP.M.IT.N.000000.4.ANR
HICP - Processed food incl. alcohol and tobacco	Monthly	ECB	ICP.M.IT.N.FOODPR.4.ANR
HICP - Unprocessed food	Monthly	ECB	ICP.M.IT.N.FOODUN.4.ANR
HICP - Industrial goods excluding energy	Monthly	ECB	ICP.M.IT.N.IGXE00.4.ANR
HICP - Energy	Monthly	ECB	ICP.M.IT.N.NRGY00.4.ANR
HICP - Services	Monthly	ECB	ICP.M.IT.N.SERV00.4.ANR
Food, import weighted	Monthly	ECB	STS.M.I8.N.ECPE.CFOOD0.3.000
Non food, import weighted	Monthly	ECB	STS.M.I8.N.ECPE.CNFOOD.3.000
Total non-energy commodity, import weighted	Monthly	ECB	STS.M.I8.N.ECPE.CTOTNE.3.000
Food, use weighted	Monthly	ECB	STS.M.I8.N.UWIE.CFOOD0.3.000
Non-food, use weighted	Monthly	ECB	STS.M.I8.N.UWIE.CNFOOD.3.000
Total non-energy commodity, use weighted	Monthly	ECB	STS.M.I8.N.UWIE.CTOTNE.3.000
Global Natural Gas	Monthly	EC	EC Weekly Oil Bulletin
EUA CO2	Weekly	Bloomberg	-
Euro Super 95	Weekly	EC	EC Weekly Oil Bulletin
Diesel	Weekly	EC	EC Weekly Oil Bulletin
Gas oil	Weekly	EC	EC Weekly Oil Bulletin
Heating Oil	Weekly	EC	EC Weekly Oil Bulletin
LPG	Weekly	EC	EC Weekly Oil Bulletin
Brent	Weekly	EIA	EC Weekly Oil Bulletin
Power - ITA	Weekly	Bloomberg	-
Power - FRA	Weekly	Bloomberg	-
Power - DE	Weekly	Bloomberg	-
Power - ESP	Weekly	Bloomberg	-
USD/EUR	Weekly	ECB	EXR.D.USD.EUR.SP00.A
YEN/EUR	Weekly	ECB	EXR.D.JPY.EUR.SP00.A
GBP/EUR	Weekly	ECB	EXR.D.GBP.EUR.SP00.A
RMB/EUR	Weekly	ECB	EXR.D.CNY.EUR.SP00.A
Nominal effective exchange rate, eurozone	Monthly	EC	ERT_EFF_IC_M
Real effective exchange rate, eurozone	Monthly	EC	ERT_EFF_IC_
Marginal lending facility	Weekly	ECB	FM.B.U2.EUR.4F.KR.MLFR.LEV
Deposit facility	Weekly	ECB	M.B.U2.EUR.4F.KR.DFR.LEV
Main refinancing operations	Weekly	ECB	FM.B.U2.EUR.4F.KR.MRR_FR.LEV
Shadow Rate ECB	Monthly	Wu-Xia shadow rates	ECB_WU_XIA_M
Shadow rate US	Monthly	Wu-Xia shadow rates	US_WU_XIA_M
Federal Funds Rate	Monthly	FRED	FEDFUNDS
Monetary aggregate M3	Monthly	ECB	BSI.M.U2.Y.V.M30.X.I.U2.2300.Z01.A

Appendix G

Additional results: Nowcasting inflation

Table G.1 summarizes the RMSFE of the different models at the end of each month for different forecasting horizons.

Table G.1: Monthly RMSFE

	h=4				h=8			h=12		
	lasso	favar	lbvar	ar(2)	lasso	favar	lbvar	lasso	favar	lbvar
date										
2019-09-30	0.172	0.142	0.108	0.208	0.167	0.125	0.121	0.231	0.222	0.246
2019-10-31	0.084	0.131	0.056	0.216	0.153	0.166	0.094	0.202	0.194	0.181
2019-11-30	0.254	0.156	0.188	0.179	0.138	0.119	0.184	0.108	0.088	0.125
2019-12-31	0.184	0.260	0.270	0.346	0.403	0.355	0.388	0.422	0.332	0.432
2020-01-31	0.107	0.169	0.137	0.083	0.281	0.272	0.244	0.421	0.396	0.381
2020-02-29	0.164	0.119	0.087	0.138	0.113	0.122	0.078	0.123	0.125	0.104
2020-03-31	0.313	0.350	0.245	0.429	0.423	0.445	0.300	0.435	0.463	0.317
2020-04-30	0.294	0.344	0.181	0.475	0.522	0.519	0.313	0.667	0.696	0.444
2020-05-31	0.165	0.200	0.094	0.213	0.347	0.350	0.078	0.521	0.565	0.238
2020-06-30	0.159	0.125	0.135	0.172	0.145	0.135	0.135	0.175	0.250	0.200
2020-07-31	0.084	0.106	0.119	0.316	0.209	0.147	0.250	0.198	0.060	0.331
2020-08-31	0.369	0.390	0.290	0.383	0.313	0.418	0.285	0.260	0.395	0.205
2020-09-30	0.092	0.269	0.200	0.537	0.416	0.553	0.234	0.481	0.700	0.319
2020-10-31	0.000	0.069	0.100	0.193	0.125	0.238	0.144	0.273	0.529	0.104
2020-11-30	0.000	0.005	0.025	0.017	0.012	0.050	0.105	0.068	0.202	0.152
2020-12-31	0.000	0.031	0.056	0.003	0.000	0.066	0.019	0.008	0.110	0.104
2021-01-31	0.922	0.738	0.737	0.752	0.756	0.703	0.700	0.763	0.663	0.760
2021-02-28	0.160	0.388	0.375	0.958	0.775	0.778	0.759	0.925	0.871	0.879
2021-03-31	0.256	0.180	0.290	0.566	0.360	0.240	0.448	0.687	0.578	0.730
2021-04-30	0.225	0.281	0.300	0.538	0.375	0.306	0.488	0.456	0.308	0.646
2021-05-31	0.295	0.290	0.175	0.251	0.368	0.405	0.355	0.505	0.465	0.595
2021-06-30	0.112	0.150	0.144	0.134	0.122	0.131	0.091	0.223	0.248	0.227
2021-07-31	0.221	0.088	0.100	0.209	0.138	0.022	0.047	0.175	0.060	0.062
2021-08-31	0.516	0.535	0.560	0.662	0.615	0.497	0.588	0.643	0.440	0.543
2021-09-30	0.296	0.431	0.438	0.599	0.744	0.700	0.741	0.921	0.775	0.875
2021-10-31	0.497	0.513	0.394	0.437	0.772	0.700	0.622	1.017	0.965	0.917
2021-11-30	0.551	0.545	0.405	0.689	0.838	0.735	0.453	1.097	0.975	0.732
2021-12-31	0.128	0.169	0.169	0.166	0.469	0.316	0.153	0.733	0.281	0.175
2022-01-31	0.066	0.125	0.195	0.193	0.135	0.175	0.428	0.307	0.238	0.345
2022-02-28	0.572	0.600	0.412	0.519	0.453	0.778	0.344	0.485	0.608	0.315
2022-03-31	1.080	1.113	1.000	1.448	1.331	1.484	1.244	1.433	1.800	1.025
2022-04-30	0.260	0.381	0.294	0.775	0.866	0.872	0.716	1.248	1.477	1.131
2022-05-31	0.412	0.325	0.415	0.736	0.410	0.340	0.390	0.773	0.703	0.662
2022-06-30	0.329	0.362	0.444	0.879	0.694	0.331	0.616	0.712	0.340	0.615
2022-07-31	0.241	0.231	0.337	0.197	0.347	0.259	0.456	0.612	0.302	0.633
2022-08-31	0.124	0.235	0.225	0.094	0.210	0.273	0.460	0.285	0.242	0.542

Appendix H

Robustness checks

In this section, the results from robustness checks are reported. The table summarizes the RMSFE results of the different models, considering the inclusion of the six monetary policy measures detailed in Section 4.5. For the Eurozone, these are the Marginal lending facility, the Deposit facility, the Main refinancing operations rates, and the Wu and Xia (2020)'s shadow interest rate. The Federal Funds rate and the Wu and Xia (2016)'s shadow rate are considered for the US. Table H.2 reports the RMSFE results for different forecast horizons. Table H.3 summarizes the RMSFEs obtained considering alternative machine learning regularization methods for the VAR (Ridge and Elastic Net). Figure H.1 displays the nowcasts of the model alternatives for Euro area core inflation. Table H.4 summarizes the RMSFEs obtained by the different models for Euro area core inflation.

Table H.1: Model comparisons: RMSFE with and without policy variables

		pre-COVID-19 sample		COVID-19 sample		Energy crisis sample		Full sample	
		no policy	policy	no policy	policy	no policy	policy	no policy	policy
	Lasso-VAR	0.144	0.146	0.239	0.240	0.379	0.387	0.261	0.264
h=4	FAVAR	**0.153**	**0.153**	**0.262**	**0.262**	0.401	0.401	**0.281**	**0.281**
	BVAR	**0.152**	**0.152**	0.235	0.235	0.364	0.364	0.256	0.256

Notes: Bold values indicate that the predictive performance of the Lasso VAR is statistically different from those obtained by the other models. The reference test is the Diebold and Mariano (1995) test of differences in the model's predictive accuracy. The reference statistics consider the Harvey et al. (1997) correction for small samples.

Table H.2: RMSFE across different models and forecast horizons

		pre-COVID-19 sample	COVID-19 sample	Energy crisis sample	Full sample
	Lasso-VAR	0.144	0.239	0.379	0.261
h=4	FAVAR	**0.153**	**0.262**	0.401	**0.281**
	BVAR	**0.152**	0.235	0.364	0.256
	Lasso-VAR	0.204	0.353	0.577	0.388
h=8	FAVAR	0.198	0.351	0.541	0.376
	BVAR	0.208	0.310	0.498	0.345
	Lasso-VAR	0.234	0.437	0.777	0.496
h=12	FAVAR	0.220	0.440	0.683	0.467
	BVAR	0.252	0.396	0.620	0.433

Notes: Bold values indicate that the predictive performance of the Lasso VAR is statistically different from those obtained by the other models. The reference test is the Diebold and Mariano (1995) test of differences in the model's predictive accuracy. The reference statistics consider the Harvey et al. (1997) correction for small samples.

Table H.3: RMSFE Lasso-ridge-elastic net

		pre-COVID-19 sample	COVID-19 sample	Energy crisis sample	Full sample
	Lasso-VAR	0.144	0.239	0.379	0.261
h=4	Ridge-VAR	**0.147**	**0.251**	0.38	0.268
	Elastic Net-VAR	**0.154**	**0.258**	0.425	**0.286**

Notes: Bold values indicate that the predictive performance of the Lasso VAR is statistically different from those obtained by the other models. The reference test is the Diebold and Mariano (1995) test of differences in the model's predictive accuracy. The reference statistics consider the Harvey et al. (1997) correction for small samples.

Figure H.1: Nowcasting Core Inflation: Lasso-VAR vs. FAVAR and BVAR

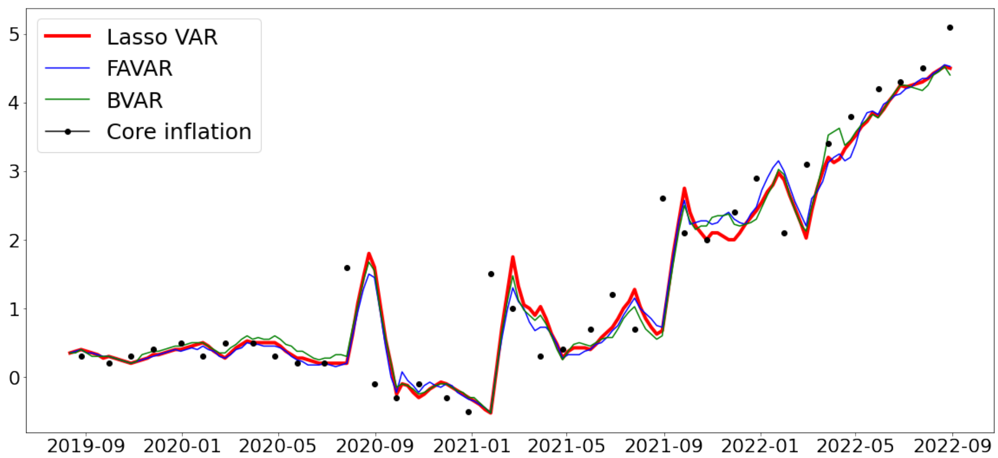


Table H.4: RMSFE Core inflation

		pre-COVID-19 sample	COVID-19 sample	Energy crisis sample	Full sample
	Lasso-VAR	0.089	0.420	0.306	0.320
h=4	FAVAR	**0.087**	0.408	**0.332**	0.321
	BVAR	**0.073**	0.433	0.317	0.327

Notes: Bold values indicate that the predictive performance of the Lasso VAR is statistically different from those obtained by the other models. The reference test is the Diebold and Mariano (1995) test of differences in the model's predictive accuracy. The reference statistics consider the Harvey et al. (1997) correction for small samples.

Appendix I

Trace plots

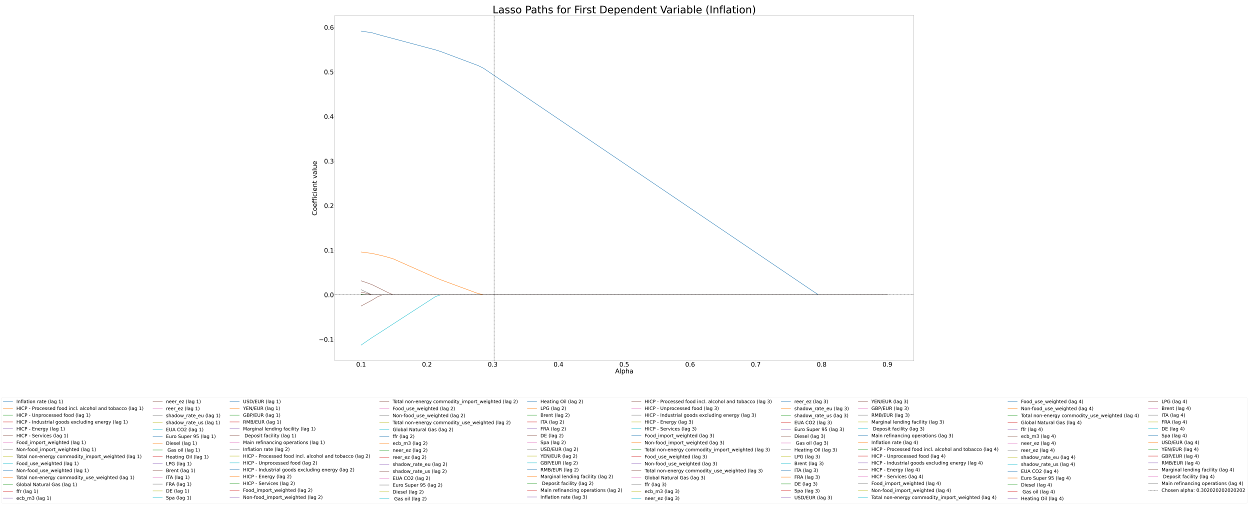
The detailed analysis of Lasso paths for inflation across pre-COVID, COVID, and post-COVID periods highlights significant shifts in the underlying determinants of inflation, aligning with the distinct economic conditions prevalent during each period. In the pre-COVID period (see Figure I.1), the inflation rate was predominantly influenced by its own lag, suggesting a relatively stable and self-persistent inflationary environment with minimal external disruptions. This stability is indicative of predictable inflation dynamics, driven largely by historical inflation values without substantial external shocks.

However, during the COVID period (see Figure I.2), the complexity of inflation dynamics increased markedly. At the optimal alpha, several predictors emerged as significant, including various components of the Harmonized Index of Consumer Prices (HICP), such as processed food, unprocessed food, industrial goods excluding energy, energy, and services. This diversification of significant predictors reflects the broad and multifaceted economic disruptions caused by the pandemic, including supply chain interruptions, shifts in consumer behavior, and sector-specific economic impacts. The inclusion of these diverse HICP components underscores the widespread and heterogeneous nature of the pandemic's economic effects.

In the post-COVID period (see Figure I.3), the Lasso paths reveal a continuation of the complex inflationary landscape observed during the pandemic, with several HICP components remaining significant. However, this period also marks the re-emergence of more traditional economic influences alongside the persistent pandemic-induced factors. The optimal alpha in the post-COVID period captures a blend of both old and new significant predictors, indicating an adjustment towards a new economic equilibrium. This transitional phase suggests that while the economy is stabilizing from the acute impacts of the pandemic, it is simultaneously incorporating the ongoing effects of COVID-19 alongside traditional inflation determinants.

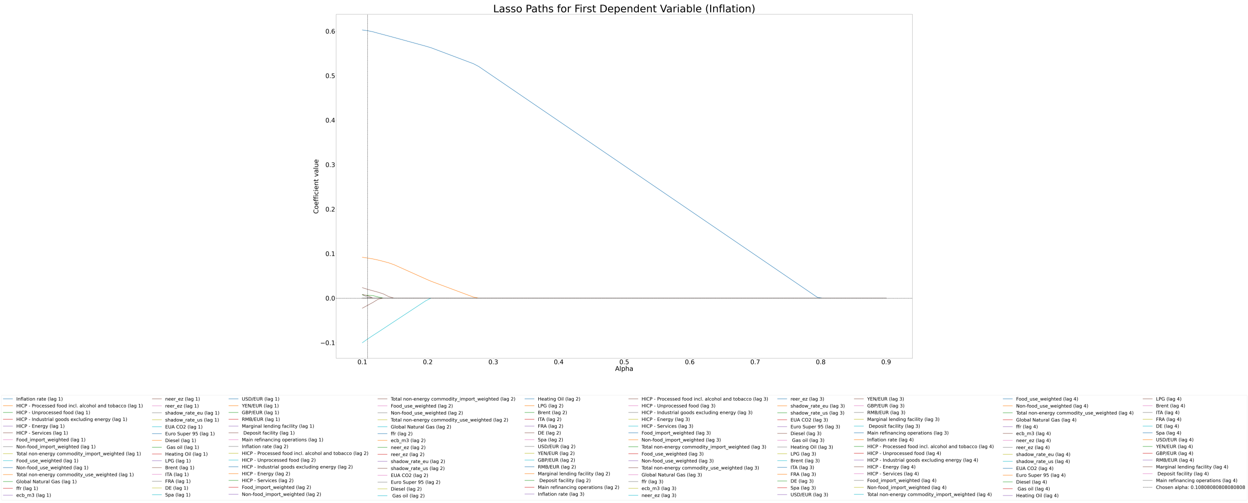
Overall, the evolution of significant predictors across these three periods underscores the dynamic and adaptive nature of inflation determinants in response to external shocks and recovery phases. The pre-COVID period's stability, the COVID period's heightened complexity, and the post-COVID period's transitional blend of influences collectively illustrate the intricate and evolving inflationary processes in varying economic contexts.

Figure I.1: Trace plot - Pre Covid 19 period



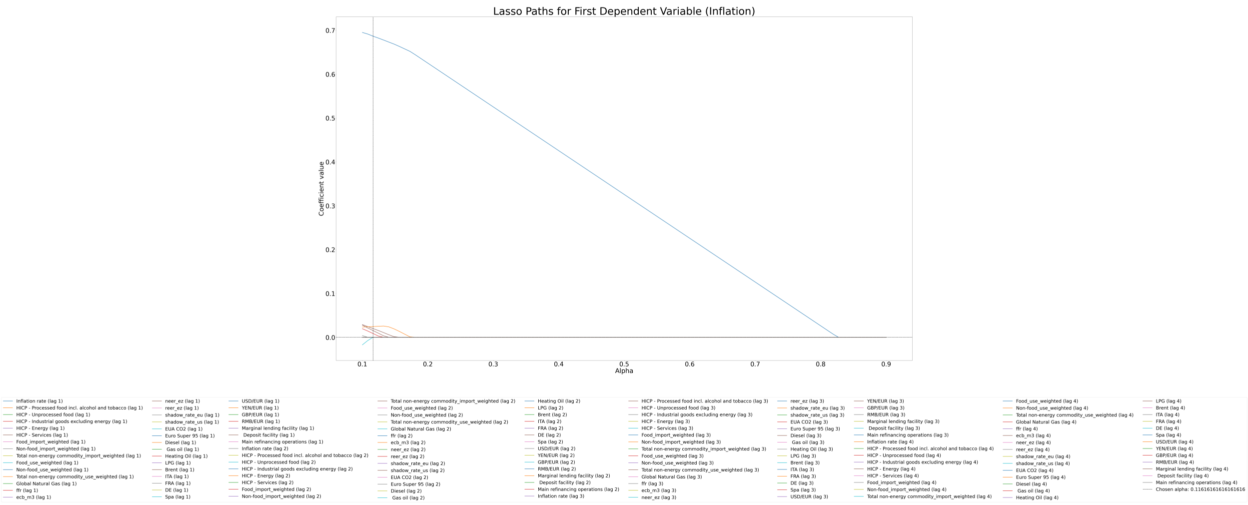
Note: The figure shows the Lasso regression trace plot for the inflation dependent variable, highlighting significant predictors at the optimal alpha for the specified period

Figure I.2: Trace plot - Covid 19 period



Note: The figure shows the Lasso regression trace plot for the inflation dependent variable, highlighting significant predictors at the optimal alpha for the specified period

Figure I.3: Trace plot - Post Covid 19 and high Inflation period



Note: The figure shows the Lasso regression trace plot for the inflation dependent variable, highlighting significant predictors at the optimal alpha for the specified period

Acknowledgements

I would like to extend my sincerest gratitude to the individuals who contributed to the realization of this research project. First and foremost, my gratitude goes to my supervisor, Professor Massimiliano Tancioni, for his invaluable guidance and unwavering support throughout the project. Additionally, I express my appreciation to Professor Giuseppe Ragusa for providing insightful comments that greatly aided the direction of the research.

Furthermore, I wish to acknowledge the valuable feedback and assistance received from my colleagues and individuals I met at various conferences. Their contributions have been instrumental in shaping and refining this research project.

Moreover, I cannot overstate my appreciation for the love, encouragement, and patience of my girlfriend, who has been an immense source of support and inspiration.

Lastly, I must mention my parents and friends for their constant encouragement and cooperation throughout my academic journey. Their unwavering support has been an invaluable source of strength and encouragement.

Milos Ciganovic, Rome, June 2024