LUISS

Department of Economics and Finance

Chair of Econometric Theory

Forecasting Industrial Production Using Electronic Invoicing Data

Supervisor:
Prof.
Paolo Santucci De Magistris

Co-Supervisor:
Prof.
Megha Patnaik

Candidate:
Francesca Mazzarda
761251

Academic Year 2023/2024

Abstract

This thesis presents a forecasting model for the Italian Industrial Production Index (IPI), one of the key macroeconomic variables tracked by policymakers to steer monetary policy. In recent years, there has been an increasing interest in exploring alternative sources of data to obtain real-time insights into economic activity. The presented forecasting model uses a novel index based on Italian electronic invoicing data as a key predictor to forecast the one-month-ahead IPI. The results indicate that the use of time series techniques (ARMAX models) with invoicing data enables accurate forecasting of industrial production activity and outperforms benchmark models during periods of economic stability.

Keywords: Industrial Production, Business Cycle, Time Series, Forecasting, Nowcasting, Turnover, ARMA, ARMAX

Contents

List of Tables

List of Figures

1 Introduction

In the last 20 years, the global economy has experienced unprecedented disruptions and uncertainty. The Great Recession and the Sovereign Debt Crisis have highlighted the limitations of conventional econometric approaches and the need for more robust forecasting tools (Rodano et al., 2013), which were further consolidated by the COVID-19 outbreak and the energy crisis. The pandemic brought the biggest shock to the global economy ever recorded, putting standard econometric models to the test (Locarno and Zizza, 2020). Traditionally, these models have relied on historical data and statistical relationships that now struggle to capture the complexities of rapidly evolving global economies, especially after extreme disruptions. This is especially relevant to policymakers and economists, who carefully examine the state of the economy by analyzing macroeconomic variables. However, this task has become especially challenging when the horizon is the current and very near future. Official indicators, such as GDP or IPI, are published by national statistics institutes with months of delay, generating interest in explaining economic activity through alternative methods. For example, Altissimo et al. (2010) and Delle Monache et al. (2019) use a combination of unconventional, high-frequency variables to create new indicators for the economy, the EUROCOIN and the Italian Weekly Economic Index (ITWEI).

Many economists are also focusing on the explanatory power of unconventional data to obtain short-term forecasts of traditional macroeconomic variables. The seminal paper by Giannone et al. (2008) already highlighted the strong interest in *nowcasting* macroeconomic variables in advance to assess current and future economic conditions. Since then, the field of nowcasting has seen substantial progress, with the development of techniques such as Dynamic Factor Models (Chernis and Sekkel, 2017), Mixed Data Sampling Models (Galdi et al., 2023), Neural Networks (Fornaro, 2020), and Bayesian Vector Autoregressive Models (Aprigliano, 2020). While some of these models use standard variables, many recent works have highlighted the importance of incorporating alternative data as a way to obtain early forecasts. Internet-based data sources, such as Google data for economic forecasting (D'Amuri and Marcucci, 2017), internet search queries (Götz and Knetsch, 2019), real estate market analyses using housing advertisement datasets (Loberto et al., 2018), and inflation expectations measured through Twitter (Angelico et al., 2022) were able to provide reliable insights. Additionally, traditional media sources like newspapers have been used to construct sentiment and uncertainty indicators using text-mining techniques (Thorsrud, 2020; Aprigliano et al., 2021). Other examples of high-frequency data are payments information (Aprigliano et al., 2017; Ardizzi et al., 2019; Aastveit et al., 2020; Aprigliano et al., 2021), transport data (Fornaro, 2020), GPS data (Delle Monache et al., 2019; Furukawa et al., 2022; Matsumura et al., 2024), and lubricant oils (Fruzzetti and Ropele, 2024).

The contribution of this thesis to the literature is the use of a novel real-time indicator extracted from electronic invoices in a model to forecast the one-month-ahead value of industrial production. The Industrial Production Index (IPI) measures price-adjusted monthly variation in the output of the industrial sector. It is considered one of the key monthly macroeconomic variables, as it constitutes a big share of GDP in many countries: in Italy, it accounts for 18.8% of the value added, as of 2023.[1] It is an essential indicator for assessing business cycle phases, as it is closely linked to other variables such as stock market prices (Choi et al., 1999; Chiang and Chen, 2017), inflation, oil prices (Ewing & Thompson, 2007), and unemployment (Neftçi, 1984). Indeed, it is used to steer and monitor economic and monetary policies in Europe, as it is able to capture the movements of the business cycle and quickly assess production activity in the face of major economic shocks. According to Martínez-García et al. (2015) and Schreiber and Soldatenkova (2016), industrial production is one of the main indicators in predicting turning points of the economy and anticipating recessions. The ability of central banks and governments to respond promptly to large disruptions is crucial to preventing further damages to the economy, as highlighted by the aftermath of the COVID-19 outbreak (Mosser, 2020; Ramos-Francia and García-Verdú, 2022). Indeed, a nowcasting model that delivers data one month in advance of official statistics would allow policymakers' ability to manage and mitigate economic shocks more effectively. The relevance of IPI is further underlined by its ability to nowcast GDP in the short term (Golinelli and Parigi, 2007; Hahn and Skudelny, 2008; Banbura et al., 2011; Baumeister and Guérin, 2021).

Many governments across the world nowadays require electronic invoicing. Italy was the first European country to introduce this system: the regulations were gradually implemented starting in 2014, initially making electronic invoicing mandatory for business-to-government transactions, and later for all public administration. By 2019, it also became mandatory for

[1] Source: ISTAT

all business-to-business and business-to-consumer transactions. The legislation requires all firms to produce electronic invoices up to 12 days after the payment (15 days for deferred invoices).[2] Team System, one of the main electronic invoicing platforms in Italy, and Centro Studi Confindustria elaborate a monthly index, known as Real Time Turnover Index,[3] that includes the sum of all electronic invoices produced in a given area for the month. We will use this dataset to assess whether transaction data can be utilized to forecast IPI, following Galbraith and Tkacz (2018) and Aprigliano et al. (2021).

This thesis proposes an analysis of the predictive power of such industrial turnover index, by assessing the performance of a benchmark ARMA model against two ARMAX models, which include, as exogenous regressors, business survey indexes, production prices, and the industrial turnover index. The empirical results indicate that the ARMAX model, which includes the industrial turnover index as a regressor, performs best in the in-sample analysis. However, in the out-of-sample analysis, this model does not consistently outperform the others overall, but it does show better performance when the COVID-19 period is excluded from the RMSE evaluation. The overall results would suggest that when the economy is not undergoing a serious shock, industrial turnover improves the short-term forecasts of IPI.

The remainder of this thesis is organized as follows. Section 2 introduces the existing literature on forecasting industrial production and describes the methodology; Section 3 presents the data; lastly, Section 4 shows the empirical analysis and the results of this study.

[2]For further details on current legislation, refer to https://ec.europa.eu/digital-building-blocks/sites/display/DIGITAL/eInvoicing+in+Italy

[3]More information available at https://www.confindustria.it/wcm/connect/2b62ba5f-2a64-4813-a42b-5236c32eb4df/Nota_CSC_RTT_Nota_metodologica_290124_Confindustria.pdf?MOD=AJPERES&CACHEID=ROOTWORKSPACE-2b62ba5f-2a64-4813-a42b-5236c32eb4df-oS4k-73

2 Forecasting industrial production

One of the many tasks of forecasters is to find what information can give valuable insight for the variable of interest. For a long time, electricity consumption has been used to forecast IPI. Bodo and Signorini (1987) was the first work to introduce the use of energy consumption data in the models, showing that real-time, high-frequency data provides good forecasts. Similar results have then been confirmed by other studies later on (Bodo et al., 1991; Marchetti and Parigi, 2000; Galdi et al., 2023). In particular, Marchetti and Parigi (2000) accounts for the use of electricity for non-industrial purposes by including meteorological data in the forecasting models. However, Alpino et al. (2023) discusses that due to the 2021-22 energy crisis, the increase in prices of gas and electricity fundamentally disrupted the relationship between energy consumption and industrial production, highlighting the rise of renewable energies. This work suggests, indeed, that energy consumption may not be a suitable explanatory variable for IPI anymore.

Similarly to electricity consumption, another kind of real-time data that can be used as a proxy of production is the transportation of goods. Bruno and Lupi (2001) shows that the number of goods transported by railways can be used to obtain good forecasts of industrial production, while Fornaro (2020) obtains similar results with truck traffic volume. Recent works also include GPS data of factory workers to track production volume (Suimon and Yanai, 2021; Furukawa et al., 2022). According to Matsumura et al. (2024), mobility data recorded on mobile cell phones can be used to quantify the level of production, especially in labor-intensive industries.

Another way to obtain information ahead of time is by using qualitative data. For example, qualitative information related to business conditions can be obtained by surveying households and firms. The recent work by Lehmann (2023) highlights the informational content of metrics extracted by national business surveys. In the context of forecasting IPI, Bodo and Signorini (1987) was one of the earliest studies showing that qualitative indicators from business surveys can be used to explain production; this result has then been confirmed by other works (Bodo et al., 2000; O'Brien and Ladiray, 2003). More recently, Girardi et al. (2016) highlighted the importance of business survey information, particularly in capturing in advance a downturn of the business cycle.

The existing literature proposes several kinds of models to forecast IPI. Many works use single or multivariate linear regressions (Parigi and Schlitzer, 1995; Bodo et al., 2000; Marchetti and Parigi, 2000; O'Brien and Ladiray, 2003; Franses and Van Dijk, 2005), others add square or cubic terms (Bodo and Signorini, 1987; Bodo et al., 1991; Marchetti and Parigi, 2000); the simplicity of these models helps understand the true relationship between explanatory and response variables. Vector Autoregressive (VAR) models are also a popular choice (Bodo et al., 2000; Bruno and Lupi, 2001; Bruno and Lupi, 2003; Schreiber and Soldatenkova, 2016; Chiu et al., 2017).

The models presented so far rely on few exogenous regressors to predict IPI, thus the choice of the covariates is crucial. Another possible approach in econometrics is to supply the models with large datasets, as addressed by Stock and Watson (2002) and Forni et al. (2005). They propose techniques that automatically select and summarize the information of the dataset into a few variables. In the context of industrial production, Bulligan et al. (2009) and Bulligan et al. (2010) question whether it is better to construct models with a few selected indicators (bridge models) or use a factor-based approach (factor model). They conclude that there is a strong trade-off between the two methodologies: the former allows for a better understanding of the relationship between the explanatory variables and industrial production, however, the latter often leads to better forecasts from a statistical point of view. For this reason, Brunhes-Lesage and Darné (2012) and Girardi et al. (2016) combine the two approaches into a single model, the Factor Augmented Bridge Model (FABM). Girardi et al. (2016) concludes that in computing this kind of exercise, partial least squares outperforms principal component analysis. Other examples of factor models include the use of static principal component analysis in Günay (2018) and Costantini (2013), and dynamic principal component analysis to build state-space models in Costantini (2013).

Other studies use more complex environments to forecast IPI, such as Mixed Data Sampling (MIDAS) models (Clements & Galvão, 2008) and Markov-Switching (MS) models (Billio et al., 2012). Galdi et al. (2023) tests MIDAS, MS, and MS-MIDAS models, showing that MS models outperform the other two in forecasting industrial production. Hassani et al. (2009) uses Singular Spectrum Analysis (SSA). Heravi et al. (2004) and Fornaro (2020) test Neural Networks. Aprigliano (2020) employs Bayesian vector autoregression (BVAR) models.

In light of all the options available, our exercise in this thesis only test models with a limited set of variables, in order to evaluate the impact of incorporating electronic invoicing data on short-term forecasts of industrial production.

2.1 Introduction to time series

A time series is a set of observations of a given variable collected over time, formally $\{y_t\}_{t=1}^{T}$, that often exhibits patterns such as trends, seasonal fluctuations, and repeated cycles. The goal of time series analysis is to capture the underlying pattern in the stochastic process in order to predict future values. Time series are modeled through a number of methods; in this section, we introduce the autoregressive moving average (ARMA) model, which was popularized by Box and Jenkins (1970) and has been widely adopted every since. The Box-Jenkins approach refers to the application of three steps to fit an ARMA model. The first step requires checking and adjusting the time series to ensure stationarity (identification). Then we select the optimal lag order and the estimate of the parameters through maximum likelihood (estimation). Lastly, the residual of the models must be analyzed (diagnostic checking). We also introduce an extension to the ARMA model, the autoregressive moving average model with exogenous regressors (ARMAX).

2.1.1 Stationarity

A stochastic process $\{y_t\}$ is said to be strictly stationary if the joint distribution of $(y_{t_1}, y_{t_2}, \dots, y_{t_T})$ is the same as that of $(y_{t_1+k}, y_{t_2+k}, \dots, y_{t_T+k})$ for all k and all sets of time points $t_1, t_2, \dots, t_T$. This definition implies that the entire probability distribution of the process does not change over time. A weaker definition of stationarity requires that a time series has a constant mean, constant variance, and an autocovariance function that depends only on the lag between two time points and not on time itself. Formally, a time series $\{y_t\}$ is weakly stationary if

- $\mathbb{E}[y_t] = \mu$
- $\mathrm{Var}(y_t) = \sigma^2$
- $\mathrm{Cov}(y_t, y_{t-k}) = \mathbb{E}[(y_t - \mu)(y_{t-k} - \mu)] = \gamma(k)$.

There is no unique way to determine whether an observed time series is stationary. One method is to look at the plots of the autocorrelation function (ACF) and partial autocorrelation function (PACF). Introduced by Yule (1927), the ACF, defined as

$$\rho_k = \frac{\mathrm{Cov}(y_t, y_{t-k})}{\mathrm{Var}(y_t)}, \tag{1}$$

evaluates the correlation between an observation in a time series and its previous observations. It shows how much an observation at a given lag is correlated with the current observation. A value close to 1 suggests a strong positive correlation, a value near -1 indicates a strong negative correlation, while a value close to 0 suggests little to no correlation.

The partial autocorrelation function,

$$\phi_{kk} = \mathrm{Corr}(y_t, y_{t-k} \mid y_{t-1}, y_{t-2}, \ldots, y_{t-(k-1)}),$$

measures how much an observation is autocorrelated with its lagged values, ignoring any indirect correlations through intermediate time lags. If the ACF and PACF coefficients decay to zero as lags increase, this would imply that the time series is stationary. When assessing the plots, values that lay outside the confidence interval (typically 95%) are to be considered statistically significant, indicating that there is evidence of autocorrelation or partial autocorrelation at that lag.

Alternatively, unit root tests such as the Dickey-Fuller test (Dickey and Fuller, 1979), the Augmented Dickey-Fuller test (Said and Dickey, 1984), and the Phillips-Perron test (Phillips and Perron, 1988), can be used to assess stationarity. The null hypothesis of the Dickey-Fuller (DF) test posits that an autoregressive time series model contains a unit root, indicating that the series is non-stationary. The alternative hypothesis generally suggests stationarity or trend stationarity, though this can vary depending on the specific version of the test. To understand the DF test, consider a simple first-order autoregressive process, AR(1), given by

$$y_t = \phi y_{t-1} + \varepsilon_t,$$

where ϕ is the parameter that determines the presence of a unit root. If $\phi < 1$, the series is stationary, as shocks to y_t will dissipate over time, and the series will revert to its unconditional mean. If $\phi = 1$, the series is nonstationary, with shocks having a permanent effect. If $\phi > 1$, the series will explode.

To perform the DF test, we subtract y_{t-1} from both sides of the equation,

$$y_t - y_{t-1} = (\phi - 1)y_{t-1} + \varepsilon_t,$$

which can be expressed as

$$\Delta y_t = \delta y_{t-1} + \varepsilon_t,$$

where $\delta = 1 - \phi$. The null hypothesis is then

$$H_0 : \delta = 0$$

indicating the presence of a unit root, while the alternative hypothesis is

$$H_1 : \delta < 0,$$

indicating stationarity. The model can also be extended to account for drift,

$$\Delta y_t = c + \delta y_{t-1} + \varepsilon_t,$$

and trend,

$$\Delta y_t = c + \alpha t + \delta y_{t-1} + \varepsilon_t.$$

The DF test, however, is considered one of the weaker methods for detecting unit roots, as it relies on AR(1) processes. To address this, the Augmented Dickey-Fuller (ADF) test incorporates higher lag orders. We start from an AR(p) with drift and trend

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t.$$

and subtracting y_{t-1} from both sides we get

$$\Delta y_t = c + \alpha t + \delta y_{t-1} + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \ldots + \beta_p \Delta y_{t-p} + \varepsilon_t.$$

Again, we are testing the null hypothesis $H_0 : \delta = 0$.

Including several lags in the ADF aims to eliminate residual autocorrelation. However, the ADF test does not account for structural shifts in the data. To address both autocorrelation and heteroskedasticity in the disturbance process, the Phillips-Perron (PP) test provides a nonparametric correction. The PP test adjusts the t-statistic from the DF equation to correct for these issues in order to obtain robust results.

When a time series is identified as non-stationary, it must often be transformed to achieve stationarity. Taking the difference between consecutive observations can often help stabilize the mean of a time series. Indeed, first-order differencing,

$$\Delta y_t = y_t - y_{t-1},$$

can remove linear trends, while second-order differencing,

$$\Delta^2 y_t = \Delta(\Delta y_t) = \Delta(y_t - y_{t-1}) = \Delta y_t - \Delta y_{t-1} = y_t - 2y_{t-1} - y_{t-2},$$

may be used for quadratic trends. Logarithmic transformations, Box-Cox transformations, and other non-linear transformations can stabilize variance in time series with exponential growth patterns. Seasonal differencing or seasonal decomposition methods such as STL (Seasonal-Trend Decomposition using LOESS) can help in dealing with seasonality, making a series stationary by removing periodic components. All of these adjustments can be used in order to ensure that the time series are stationary.

2.1.2 Maximum likelihood estimation

Maximum likelihood estimation (MLE) is a method used to estimate the parameters of a model. The main idea of MLE is to find the parameter values that maximize the likelihood of observing the given data. Given a model defined by a probability density function $f(y \mid \boldsymbol{\theta})$, where y represents the observed data and $\boldsymbol{\theta}$ denotes the parameters, the likelihood function is defined as

$$L(\boldsymbol{\theta} \mid y) = \prod_{t=1}^{T} f(y_t \mid \boldsymbol{\theta}),$$

where T is the number of observations. To simplify calculations, we typically work with the log-likelihood function, which is obtained by taking the natural logarithm of the likelihood function, $\ell(\boldsymbol{\theta} \mid y) := \log L(\boldsymbol{\theta} \mid y)$. Then, the log-likelihood function is

$$\ell(\boldsymbol{\theta} \mid y) = \sum_{t=1}^{T} \log f(y_t \mid \boldsymbol{\theta}).$$

Since the logarithm is a monotonic transformation, the MLE of the parameters $\boldsymbol{\theta}$ is then defined as

$$\hat{\boldsymbol{\theta}}_{MLE} = \arg\max_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta} \mid y).$$

2.1.3 Residual diagnostics

To ensure that the model adequately describes the data, it is important to conduct goodness-of-fit checks through residual diagnostics. This process involves visualizing the residuals to verify whether they are normally distributed and serially uncorrelated. Useful plots for this purpose include quantile-quantile (Q-Q) plots, histograms of the residuals, and autocorrelation function (ACF) plots. Making residual diagnostic plots is an informal, but useful, way to assess the violation of model assumptions: it is necessary to test the residuals for autocorrelation using statistical tests such as the Ljung-Box Q-test. These checks help confirm the validity of the model assumptions and their suitability for the data.

One assumption of many time series models, including ARMA and ARMAX, is that the residuals, i.e. the difference between the observed values and the predicted values, are normally distributed,

$$\hat{\varepsilon}_t = y_t - \hat{y}_t \sim N(0, \sigma^2)$$

Normality of residuals can be assessed by looking at the histogram of the residuals and the quantile-quantile (Q-Q) plot. The Q-Q plot compares the quantiles of the sample data to the quantiles of a theoretical normal distribution. If the residuals are normally distributed, it indicates that the ARMA model has captured the majority of the systematic patterns in the data, and the remaining differences are random noise that can be well approximated by a normal distribution. In other words, if the residuals of an ARMA model are well approximated by a normal distribution, it suggests that the model is a good fit for the data.

The Q-tests proposed by Box and Pierce (1970) and Ljung and Box (1978) verifiy the null hypothesis that there is no autocorrelation in the residuals up to a certain number of lags L,

$$H_0 : \rho_j = 0 \quad \text{for} \quad j = 1, 2, \ldots, L$$

where ρ_j is the autocorrelation at lag j, as defined in Equaiton 1. The results of the test are sensitive to the choice of L. If the test rejects the null hypothesis, it suggests that the residuals exhibit significant autocorrelation. The Ljung-Box-Q-test statistic is computed as follows,

$$Q = T(T+2) \sum_{j=1}^{L} \frac{\hat{\rho}_j^2}{T-j}$$

where $\hat{\rho}_j = \frac{\sum_{t=j+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}}{\sum_{t=1}^{T} \hat{\varepsilon}_t^2}$ is the sample autocorrelation. Under the null hypothesis of no autocorrelation, Q follows a chi-squared distribution with L degrees of freedom,

$$Q \sim \chi^2(L).$$

2.1.4 Forecasting

The reliability of a time series model's forecast should be evaluated based on its out-of-sample performance. Having observed a time series $\{y_1, \ldots, y_t\}$, we are interested in forecasting a future value y_{t+h}, where $h > 0$ is the forecasting horizon. The forecasted value will be denoted as $\hat{y}_{t+h}$. The purpose of the forecasting exercise presented in this thesis is to compare different models based on how well the forecasted values, $\hat{y}_t$, align with the observed values, y_t. The h-step-ahead forecast of y_{t+h} is given by

$$\hat{y}_{t+h} = \mathbb{E}_t[y_{t+h}].$$

Suppose we are only interested in predicting the next value of the time series. Forecasting the one-step-ahead value of y_t can be achieved through an expanding window or a rolling window of data points. A rolling window of size w is simply a subset of w consecutive observations from a time series y_t, $\{y_{t-w+1}, y_{t-w+2}, \ldots, y_t\}$. At each time step t, we first fit the model through the values $\{y_{t-w+1}, y_{t-w+2}, \ldots, y_t\}$, thus obtaining the estimated parameters, and then we predict the value of y_{t+1}. Suppose in our data sample there are T observations, $\{1, 2, \ldots, w, \ldots, T\}$. We can repeat these steps $T - w + 1$ times to obtain a set of forecasted values, $\{\hat{y}_{w+1}, \hat{y}_{w+2}, \ldots, \hat{y}_T\}$, which we can compare to the observed values, $\{y_{w+1}, y_{w+2}, \ldots, y_T\}$. We would like predictions to be as close as possible to actual values.

The choice of the size of the rolling window, w, is crucial and depends on several aspects. First of all, the size should be chosen according to the sample size available and the number of regressors. In general, a smaller window size allows the model to quickly adapt to changes, making the forecast exercise more suitable when we want to capture short-term fluctuations. On the other hand, a larger window size smooths out short-term volatility. When we deal with extraordinary events and shocks to the target variable, it might be best to choose a larger size to smooth out the effects of the shock.

To evaluate and compare our forecasting models, we use the Root Mean Squared Error (RMSE) to measure how close the predictions are to the true data. The RMSE is defined as

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(e_t)^2},$$

where $e_t = y_t - \hat{y}_t$ is the forecast error. The advantage of using RMSE is that squaring the individual errors before averaging them gives more weight to larger errors. Indeed, RMSE is more sensitive to outliers or cases where the model significantly deviates from the observed values. The lower the RMSE, the better the model's ability to predict the data. Conversely, a higher RMSE indicates a greater discrepancy between the predicted and actual values of y_t.

2.2 ARMA model

ARMA models are a class of models that describe the evolution of a variable as a linear combination of its own past values (autoregressive component) and a linear combination of past errors (moving average component). A moving average (MA) process of order q in which the variable of interest, y_t, can be modeled as a function of past and current error terms, ε_{t-i} for $i = 0, \ldots, q$, and a constant term, c. It can be expressed as

$$\begin{aligned} y_t &= c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \\ &= c + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t, \end{aligned}$$

where the errors are a Gaussian white noise, $\varepsilon_t \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$. Using the lag operator B ($Bx_t = x_{t-1}$), we can rewrite MA(q) as

$$y_t = c + \Theta(B)\varepsilon_t, \tag{2}$$

where $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ is the characteristic polynomial.

The autoregressive (AR) process of order p explains y_t as a function of its own past values, y_{t-i} for $i = 1, \ldots, p$, a constant term c, and a random shock ε_t. It can be written as

$$\begin{aligned} y_t &= c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \\ &= c + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t \end{aligned}$$

Using the lag operator B, we can rewrite AR(p) as

$$\Phi(B) y_t = c + \varepsilon_t, \tag{3}$$

where $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ is the characteristic polynomial.

Combining Equations 2 and 3 together, we obtain an ARMA(p,q) process for which y_t is a function of both its past values (up to lag p) and its present and past shock values (up to lag q). We can write it as

$$\Phi(B) y_t = c + \Theta(B)\varepsilon_t, \tag{4}$$

or

$$\begin{aligned} y_t &= c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \\ &= c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t. \end{aligned} \tag{5}$$

2.2.1 Stationarity of ARMA processes

One of the key assumptions of ARMA models is that the time series is stationary, i.e. its statistical properties do not change over time. Let us first analyze the stationarity of MA(q) and AR(p) processes.

Assume an MA(q) process with $c = 0$; it is stationary by definition since,

$$\begin{aligned}
\mathbb{E}[y_t] &= \mathbb{E}[\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \cdots + \theta_q\varepsilon_{t-q}] \\
&= \mathbb{E}[\varepsilon_t] + \theta_1\mathbb{E}[\varepsilon_{t-1}] + \cdots + \theta_q\mathbb{E}[\varepsilon_{t-q}] \\
&= 0, \qquad (6) \\
\mathrm{Var}(y_t) &= \mathbb{E}[(y_t)^2] + \mathbb{E}[y_t]^2 \\
&= \mathbb{E}[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q})^2] \\
&= (1 + \theta_1^2 + \cdots + \theta_q^2)\sigma^2, \qquad (7)
\end{aligned}$$

and

$$(8)$$

$$\begin{aligned}
\mathrm{Cov}(y_t, y_{t-k}) &= \mathbb{E}[(y_t)\cdot(y_{t-k})] \\
&= \mathbb{E}[(\varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q})\cdot(\varepsilon_{t-k} + \theta_1\varepsilon_{t-k-1} + \cdots + \theta_q\varepsilon_{t-k-q})] \\
&= \mathbb{E}[\theta_k\varepsilon_{t-k}^2 + \theta_1\theta_{k+1}\varepsilon_{t-k-1}^2 + \cdots + \theta_q\theta_{q-k}\varepsilon_{t-k-q}^2] \\
&= (\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_q\theta_{q-k})\sigma^2. \qquad (9)
\end{aligned}$$

An MA(∞) process,

$$y_t = \sum_{j=0}^{\infty}\theta_j\varepsilon_{t-j},$$

is stationary if $\sum_{j=0}^{\infty}\theta_j^2 < \infty$.

On the other hand, an AR(p) process is stationary only under some condition. Consider the AR(p) process as defined in Equation 3,

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)y_t = \varepsilon_t.$$

The stationarity condition requires the roots of the characteristic polynomial $(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p) = \Phi(B) = 0$ to lie outside the unit circle. That is, $|\phi_i| < 1$.

Remember that an ARMA process is the combination of a MA process and an AR process. Hence, for the ARMA(p,q) process given by $\Phi(B)y_t = \Theta(B)\varepsilon_t$, y_t is stationary if only if the roots of $\Phi(B) = 0$ have all modulus greater than 1.

2.2.2 Estimation

One way to estimate the parameters of the ARMA model, $\boldsymbol{\theta} = (\phi_1, \cdots, \phi_p, \theta_1, \cdots, \theta_q, c, \sigma^2)$, is through maximum likelihood. The maximum likelihood estimator of the parameters of the ARMA model, $\boldsymbol{\theta} = (\phi_1, \cdots, \phi_p, \theta_1, \cdots, \theta_q, c, \sigma^2)$, is

$$\hat{\boldsymbol{\theta}}_{MLE,T} = \arg\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \ell_T(\boldsymbol{\theta})$$

where $\boldsymbol{\Theta}$ is the parameter space, and $\ell_T(\boldsymbol{\theta})$ is the log-likelihood function. Under the assumption of Gaussian errors the likelihood function is equal to

$$L_T(\boldsymbol{\theta}) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\varepsilon_t(\boldsymbol{\theta})^2}{2\sigma^2}\right),$$

and $\ell_T(\boldsymbol{\theta})$ is

$$\ell_T(\boldsymbol{\theta}) = -\frac{T}{2}\log(2\pi) - \frac{T}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(\varepsilon_t(\boldsymbol{\theta}))^2.$$

2.2.3 Information criteria

Before an ARMA(p,q) model can be estimated, we need to select the optimal choice of p and q. The two most well-known methods to select the optimal p and q are based on information criteria. These criteria balance how well the model predicts the data with the model complexity, i.e. they maximize the likelihood function while penalizing the number of parameters included.

The Akaike Information Criterion (Akaike, 1974) and the Bayesian Information Criterion (Schwarz, 1978) are equal to

$$\text{AIC} = -2\log(L) + 2(p+q)$$

and

$$\text{BIC} = -2\log(L) + (p+q)\log(T),$$

where L is the likelihood function. The idea is to calculate AIC and BIC for a set of pairs of p and q and then choose the model with the lowest value.

The AIC and BIC test often provide different results. The AIC favors models that fit the data well, even if they are more complex, because the penalty for adding parameters is relatively lower. On the other hand, the BIC imposes a stricter penalty on additional parameters, especially when the sample size is large. This difference can lead to selecting simpler models using BIC as opposed to AIC, particularly when the number of observations (T) is large.

2.2.4 Forecasting

Let us start by clarifying the following concepts. First, the observation y_{t+l} is known at time t if l is negative (indicating the past), and it is unknown and must be forecasted at time t if l is positive (indicating the future). Formally,

$$\mathbb{E}_t[y_{t+l}] = \begin{cases} y_{t+l} & \text{if } l \leq 0 \\ \hat{y}_{t+l} & \text{if } l > 0 \end{cases}.$$

Second, the error term is known in the past and assumed to be 0 in the future, under the assumption of normality of the errors, that is

$$\mathbb{E}_t[\varepsilon_{t+l}] = \begin{cases} \varepsilon_{t+l} & \text{if } l \leq 0 \\ 0 & \text{if } l > 0 \end{cases}.$$

Consider the ARMA(p,q) process given by Equation 5. The one-step-ahead forecast is the value y_{t+1} computed with information available at time t,

$$\begin{aligned} \hat{y}_{t+1} = \mathbb{E}_t[y_{t+1}] &= \mathbb{E}_t\left[c + \sum_{i=1}^{p} \phi_i y_{t+1-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t+1-j} + \varepsilon_{t+1}\right] \\ &= \mathbb{E}_t[c] + \mathbb{E}_t\left[\sum_{i=1}^{p} \phi_i y_{t+1-i}\right] + \mathbb{E}_t\left[\sum_{j=1}^{q} \theta_j \varepsilon_{t+1-j}\right] + \mathbb{E}_t[\varepsilon_{t+1}] \\ &= \hat{c}_t + \sum_{i=1}^{p} \hat{\phi}_{i,t} y_{t+1-i} + \sum_{j=1}^{q} \hat{\theta}_{j,t} \varepsilon_{t+1-j}, \end{aligned} \tag{10}$$

where

$$\hat{\boldsymbol{\theta}}_{MLE,t} = (\hat{\phi}_{1,t}, \cdots, \hat{\phi}_{p,t}, \hat{\theta}_{1,t}, \cdots, \hat{\theta}_{q,t}, \hat{c}_t, \hat{\sigma}_t^2)$$

is the maximum likelihood estimator computed at time t. The two-step-ahead forecast is

$$\begin{aligned} \hat{y}_{t+2} = \mathbb{E}_t[y_{t+2}] &= \mathbb{E}_t\left[c + \sum_{i=1}^{p} \phi_i y_{t+2-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t+2-j} + \varepsilon_{t+2}\right] \\ &= \mathbb{E}_t[c] + \mathbb{E}_t\left[\sum_{i=1}^{p} \phi_i y_{t+2-i}\right] + \mathbb{E}_t\left[\sum_{j=0}^{q} \theta_j \varepsilon_{t+2-j}\right] \\ &= \hat{c}_t + \left(\hat{\phi}_{1,t}\hat{y}_{t+1} + \sum_{i=2}^{p} \hat{\phi}_{i,t} y_{t+2-i}\right) + \sum_{j=2}^{q} \hat{\theta}_{j,t} \varepsilon_{t+2-j}, \end{aligned} \tag{11}$$

where $\hat{y}_{t+1}$ is the value computed at Equation 10. Repeating this step forward, we can achieve the h-step-ahead forecast,

$$
\begin{aligned}
\hat{y}_{t+h} = \mathbb{E}_t[y_{t+h}] &= \mathbb{E}_t\left[c + \sum_{i=1}^{p}\phi_i y_{t+h-i} + \sum_{j=1}^{q}\theta_j \epsilon_{t+h-j} + \varepsilon_{t+h}\right] \\
&= \mathbb{E}_t[c] + \mathbb{E}_t\left[\sum_{i=1}^{p}\phi_i y_{t+h-i}\right] + \mathbb{E}_t\left[\sum_{j=0}^{q}\theta_j \varepsilon_{t+h-j}\right] \\
&= \hat{c}_t + \left(\sum_{i=1}^{h-1}\hat{\phi}_{i,t}\hat{y}_{t+h-i} + \sum_{i=h}^{p}\hat{\phi}_{i,t} y_{t+h-i}\right) + \sum_{j=h}^{q}\hat{\theta}_{j,t}\varepsilon_{t+h-j}, \qquad (12)
\end{aligned}
$$

where $\sum_{i=1}^{h-1}\hat{y}_{t+h-i}$ are the forecasts obtained at the previous steps.

Under the assumption that forecast errors ε_t are normally distributed, the $100(1-\alpha)\%$ confidence interval is given by

$$
\hat{y}_{t+h} \pm z_{\alpha/2}\cdot \mathrm{SE}(\hat{y}_{t+h}),
$$

where $z_{\alpha/2}$ is the critical value from the standard normal distribution corresponding to the desired confidence level (e.g., $z_{0.025} \approx 1.96$ for a 95% confidence interval), and $\mathrm{SE}(\hat{y}_{t+h})$ is the standard error of the forecast. The standard error is computed as

$$
\mathrm{SE}(\hat{y}_{t+h}) = \sqrt{\mathrm{Var}(\hat{y}_{t+h})} = \hat{\sigma}\sqrt{\sum_{j=0}^{h-1}\hat{\psi}_j^2}, \qquad (13)
$$

where ψ_j are the coefficients from the MA(∞) representation of the ARMA process (see Appendix A.1). The proof of Equation 13 can be found in Brockwell and Davis (2016). Note that, as the forecast horizon h increases, the confidence interval widens, i.e. uncertainty increases. Consider a larger forecast horizon $H > h$, then

$$
\hat{\sigma}\sqrt{\sum_{j=0}^{h-1}\hat{\psi}_j^2} < \hat{\sigma}\sqrt{\sum_{j=0}^{H-1}\hat{\psi}_j^2}.
$$

2.3 ARMAX model

ARMAX (Autoregressive Moving Average with Exogenous Inputs) models extend the ARMA framework to include external variables. These models are particularly useful when the time series of interest is influenced by external factors that are not captured by the autoregressive and moving average components alone.

An ARMAX(p,q,r) model can be specified as

$$
\begin{aligned}
y_t &= c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \boldsymbol{\gamma}_1 \mathbf{X}_{t-1} + \cdots + \boldsymbol{\gamma}_r \mathbf{X}_{t-r} \\
&= c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{r} \boldsymbol{\gamma}_k \mathbf{X}_{t-k} + \varepsilon_t, \qquad (14)
\end{aligned}
$$

where $\mathbf{X}_t$ represents the vector of exogenous variables, and $\boldsymbol{\gamma}_i$ are the coefficients associated with those exogenous variables.

2.3.1 Stationarity of ARMAX processes

Consider Equation 14. For the process to be stationary, we need all components to be stationary. First, the MA part is stationary by definition (see Equation 6). Second, the AR part is stationary if the roots of the characteristic polynomial $(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) = \Phi(B) = 0$ to lie outside the unit circle. Third, we require the stationarity of each of the exogenous variables in $\mathbf{X}_t$, which can be assessed with the methods proposed in Section 2.1.1.

2.3.2 Estimation

For an ARMAX model we follow the same steps explained in Section 2.2.2, considering that we have to estimate $\boldsymbol{\theta} = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_r, c, \sigma^2)$.

2.3.3 Forecasting

Consider the ARMAX(p,q,r) process expressed Equation 14. The one-step-ahead forecast, $\hat{y}_{t+1}$, is computed with information available at time t,

$$
\begin{aligned}
\hat{y}_{t+1} = \mathbb{E}_t[y_{t+1}] &= \mathbb{E}_t\left[c + \sum_{i=1}^{p} \phi_i y_{t+1-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t+1-j} + \sum_{k=1}^{r} \boldsymbol{\gamma}_k \mathbf{X}_{t+1-k} + \varepsilon_{t+1}\right] \\
&= \mathbb{E}_t[c] + \mathbb{E}_t\left[\sum_{i=1}^{p} \phi_i y_{t+1-i}\right] + \mathbb{E}_t\left[\sum_{j=1}^{q} \theta_j \varepsilon_{t+1-j}\right] + \mathbb{E}_t\left[\sum_{k=1}^{r} \boldsymbol{\gamma}_k \mathbf{X}_{t+1-k}\right] + \mathbb{E}_t[\varepsilon_{t+1}] \\
&= \hat{c}_t + \sum_{i=1}^{p} \hat{\phi}_{i,t} y_{t+1-i} + \sum_{j=1}^{q} \hat{\theta}_{j,t} \varepsilon_{t+1-j} + \sum_{k=1}^{r} \hat{\boldsymbol{\gamma}}_{k,t} \mathbf{X}_{t+1-k}, \qquad (15)
\end{aligned}
$$

where

$$
\hat{\boldsymbol{\theta}}_{MLE,t} = (\hat{\phi}_{1,t}, \cdots, \hat{\phi}_{p,t}, \hat{\theta}_{1,t}, \cdots, \hat{\theta}_{q,t}, \hat{\boldsymbol{\gamma}}_{1,t}, \cdots, \hat{\boldsymbol{\gamma}}_{r,t}, \hat{c}_t, \hat{\sigma}_t^2)
$$

is the maximum likelihood estimator at time t. The two-step-ahead forecast is

$$\begin{aligned}
\hat{y}_{t+2} = \mathbb{E}_t[y_{t+2}] &= \mathbb{E}_t\left[c + \sum_{i=1}^{p}\phi_i y_{t+2-i} + \sum_{j=1}^{q}\theta_j \epsilon_{t+2-j} + \sum_{k=1}^{r}\boldsymbol{\gamma}_k \mathbf{X}_{t+2-k} + \varepsilon_{t+2}\right] \\
&= \mathbb{E}_t[c] + \mathbb{E}_t\left[\sum_{i=1}^{p}\phi_i y_{t+2-i}\right] + \mathbb{E}_t\left[\sum_{j=0}^{q}\theta_j \varepsilon_{t+2-j}\right] + \mathbb{E}_t\left[\sum_{k=1}^{r}\boldsymbol{\gamma}_k \mathbf{X}_{t+2-k}\right] \\
&= \hat{c}_t + \left(\hat{\phi}_{1,t}\hat{y}_{t+1} + \sum_{i=2}^{p}\hat{\phi}_{i,t} y_{t+2-i}\right) + \sum_{j=2}^{q}\hat{\theta}_{j,t}\varepsilon_{t+2-j} + \left(\hat{\boldsymbol{\gamma}}_{1,t}\hat{\mathbf{X}}_{t+2-k} + \sum_{k=2}^{r}\hat{\boldsymbol{\gamma}}_{k,t}\mathbf{X}_{t+2-k}\right). \quad (16)
\end{aligned}$$

The h-step-ahead forecast is

$$\begin{aligned}
\hat{y}_{t+h} = \mathbb{E}_t[y_{t+h}] &= \mathbb{E}_t\left[c + \sum_{i=1}^{p}\phi_i y_{t+h-i} + \sum_{j=1}^{q}\theta_j \epsilon_{t+h-j} + \sum_{k=1}^{r}\boldsymbol{\gamma}_k \mathbf{X}_{t+h-k} + \varepsilon_{t+h}\right] \\
&= \mathbb{E}_t[c] + \mathbb{E}_t\left[\sum_{i=1}^{p}\phi_i y_{t+h-i}\right] + \mathbb{E}_t\left[\sum_{j=0}^{q}\theta_j \varepsilon_{t+h-j}\right] + \mathbb{E}_t\left[\sum_{k=1}^{r}\boldsymbol{\gamma}_k \mathbf{X}_{t+h-k}\right] \\
&= \hat{c}_t + \left(\sum_{i=1}^{h-1}\hat{\phi}_{i,t}\hat{y}_{t+i} + \sum_{i=h}^{p}\hat{\phi}_{i,t} y_{t+h-i}\right) + \sum_{j=h}^{q}\hat{\theta}_{j,t}\varepsilon_{t+h-j} + \left(\sum_{k=1}^{h-1}\hat{\boldsymbol{\gamma}}_{k,t}\hat{\mathbf{X}}_{t+h-k} + \sum_{k=h}^{p}\hat{\boldsymbol{\gamma}}_{k,t}\mathbf{X}_{t+h-k}\right).
\end{aligned} \quad (17)$$

Note that the values of the exogenous variables in the future are unknown, hence we have called them $\hat{\mathbf{X}}_{t+h-k}$. There are several options to deal with this issue. For instance, one could assume that it behaves like a random walk, i.e. $\mathbf{X}_t = \mathbf{X}_{t-1} + \varepsilon_t$. Alternatively, the value could be set as constant, $\mathbb{E}_t[\mathbf{X}_{t+h}] = \cdots = \mathbb{E}_t[\mathbf{X}_t] = \mathbf{X}_{t-1}$. Another option would be to forecast $\mathbf{X}_t$ using an AR, MA, or ARMA process.

Under the assumption that forecast errors ε_t are normally distributed, the $100(1-\alpha)\%$ confidence interval is given by

$$\hat{y}_{t+h} \pm z_{\alpha/2} \cdot \mathrm{SE}(\hat{y}_{t+h}),$$

where $z_{\alpha/2}$ is the critical value from the standard normal distribution and $\mathrm{SE}(\hat{y}_{t+h})$ is the standard error of the forecast.

3 Data

The target variable of our analysis is the Industrial Production Index. The exogenous regressors that are included in the ARMAX models consist of indexes based on the sentiment reported by manufacturing firms on ISTAT surveys, the Producer Price Index (PPI), and a novel index of industrial turnover computed from electronic invoices.

3.1 Industrial Production Index

The index of industrial production helps identify the different phases of the business cycle by tracking the increases and decreases of the value added of sectors B through E of NACE, the four-digit classification of economic activity: (B) mining and quarrying, (C) manufacturing, (D) electricity, gas, steam, and air conditioning supply, and (E) water supply, sewerage, waste management, and remediation activities.[4] In the simplest form, an index is set to be 100 in the base year, 0, and all the other observations are computed as $\text{index}_t = \frac{\text{value}_t}{\text{value}_0} \cdot 100$.

Total industrial production cannot be directly observed, hence it is computed using different proxies: the actual quantity of goods produced (72.7% of the total index, as of 2023), number of hours worked (12.2%), and the value of production (15.1%). Each month, a panel of 5700 Italian firms provides ISTAT with the level of production (quantity, hours worked, or value, depending on the type of good) related to specific goods, which belong to a basket that is built to be representative of the Italian industry.

The index is compiled at different stages. The first stage requires aggregating products into a product group index,

$$P_{j,t} = \frac{\sum_i p_{i,j,t} q_{i,j,t}}{\sum_i p_{i,j,t} q_{i,j,0}} \cdot 100,$$

where $p_{i,j,t}$ and $q_{i,j,t}$ are the prices and quantity produced of good i (belonging to the product group j) at time t; $q_{i,j,0}$ is the quantity of good i in the base period 0. For those goods that are computed according to the number of hours worked the formula is

$$P_{j,t} = \frac{\sum_i h_{i,j,t}}{\sum_i h_{i,j,0}} \cdot 100,$$

where $h_{i,j,t}$ and $h_{i,j,0}$ are the number of hours worked to produce good i, belonging to sector j.

[4] For further details, read European Commission, Eurostat, NACE Rev. 2 - Statistical classification of economic activities in the European Community, Publications Office, 2008

For the value of production, we compute

$$P_{j,t} = \frac{\sum_i v_{i,j,t}}{\sum_i v_{i,j,0}} \cdot 100,$$

where $v_{i,j,t}$ and $v_{i,j,0}$ are the values of production at time t and 0.

After the indexes for all product groups have been calculated, the IPI can be computed at the 4-digit level of NACE. At this stage, the weights for the product groups in the base year are derived from the share of the gross production value,

$$w_j = \frac{PV_{j,0}}{\sum_{j \in k} PV_{j,0}},$$

where PV stands for "gross production value" and $\sum_{j \in k} w_j = 1$. The index is then computed as

$$I_{k,t} = \sum_{j \in k} w_j \cdot P_{j,t}.$$

To obtain a representative IPI of the entire industry, it is important to measure the distribution of value added between the sectors. Using Laspeyres formula, the weights of the individual sectors (k) for the base year (0) are calculated as

$$w_k = \frac{VA_{k,0}}{\sum_{k=1}^{K} VA_{k,0}},$$

where VA represents the value added and $\sum_{k=1}^{K} w_k = 1$. The IPI can now be computed as

$$\text{IPI}_t = \sum_{k=1}^{K} w_k \cdot I_{k,t},$$

where w_k is the weight associated to each economic sector.

The index is then adjusted for calendar days and seasonality. Calendar effects are corrected using regression methods (TRAMO procedure), accounting for factors such as working days, leap years, and national holidays by introducing specific variables into the statistical model. The TRAMO-SEATS+ procedure is used to obtain seasonally adjusted indices. This model-based method assumes that each time series consists of unobservable components: cycle-trend (medium and long-term trend), seasonal (annual periodic movements), and irregular (erratic factors). The methodology involves additive and multiplicative decomposition of raw data. Indices of industrial production are seasonally adjusted separately for each economic sector and overall index.

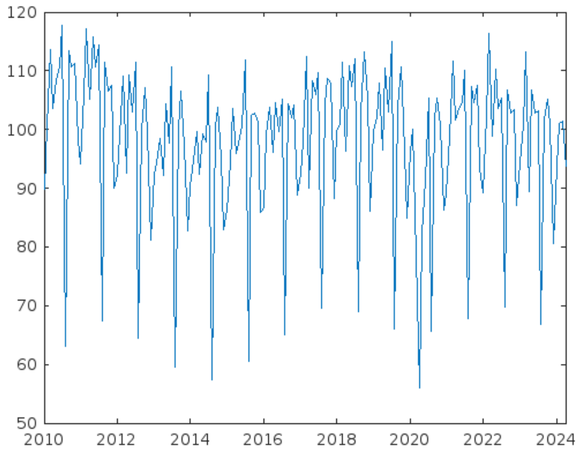
Adjusting for calendar days and seasonality is essential to accurately interpret IPI because these factors can introduce regular and predictable fluctuations that do not reflect the

Index of Industrial Production for Italy				
	Min	Max	Mean	SD
No adjustment	56.0	117.8	98.0	13.2
Calendar adjusted	54.9	119.0	98.5	13.1
Seasonally adjusted*	55.6	107.4	98.0	5.3

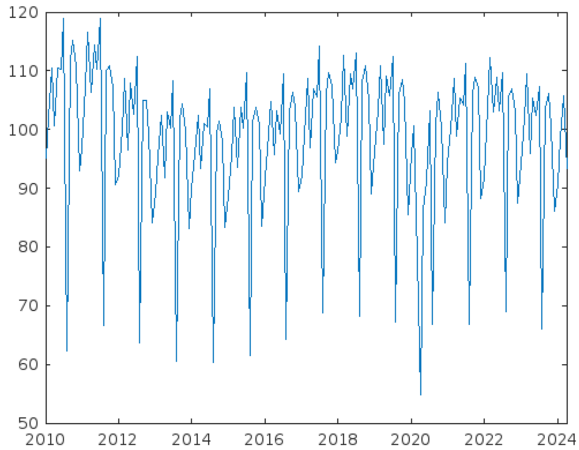
* Seasonally adjusted data also accounts for calendar adjustments

Table 1: Descriptive statistics for IPI, showing the effects of calendar and seasonal adjustments on the observed minimum, maximum, mean, and standard deviation. Sample: January 2010 - April 2024 (172 observations)

underlying economic trend. Many economic variables exhibit regular patterns that occur at the same time each year due to factors like weather, holidays, and production cycles. By removing these seasonal effects, it is easier to identify true changes in the industry's output rather than variations caused by predictable factors. The same idea applies to calendar days: production will be lower in shorter months, but this is an expected pattern that we wish to ignore as it is repeated every year and does not provide any interesting insight. Figure 1 shows the three specifications for Italy from January 2010 to April 2024: the cyclical patterns are clearly visible when the data is not adjusted and when it is adjusted only for the number of calendar days. As shown in Table 1, the seasonally adjusted data–corrected for both seasonal patterns and calendar days–exhibits the lowest level of volatility. Specifically, its standard deviation is 5.3, compared to 13.2 and 13.1 for the unadjusted and calendar-adjusted data, respectively. Whether it is beneficial to use the adjusted value to forecast industrial production is unclear. According to Mir and Osborn (2004), economists should be aware that the seasonal adjustment of IPI could potentially distort the analysis of the business cycle, however, Menezes et al. (2006) argues that even though the adjustment leads to a slower identification of the business cycle phase, it minimizes the false detections of turning points.



(a) No adjustment



(b) Calendar adjusted

(c) Seasonally adjusted

Fig. 1: Plots of IPI for Italy from January 2010 to April 2024.

3.2 Explanatory variables

3.2.1 Business survey

European national statistics institutes conduct monthly surveys to investigate the sentiments of consumers and business owners on current and future economic conditions. The information collected, mainly qualitative, is considered significant because it can detect changes in individual behaviors and current or future shifts in the surveyed economic sectors (Lehmann, 2023). The business surveys usually cover four economic sectors: manufacturing, construction, services, and retail. The survey aims to gather opinions (judgments and expectations) from all economic agents regarding specific variables related to their future behavior and the economic environment in which they operate. The opinions expressed by the interviewed units on key variables, once quantified and processed, also provide an indication of the level of confidence in the surveyed sectors.

In Europe, business surveys are harmonized and follow the same principles: they contain mostly qualitative questions, to which the firms can respond positively, neutrally, or negatively. From these questions, information about firms' perceptions of current and future business conditions are turned into indexes ranging from -100 to 100. A question may have three options, e.g. "up", "unchanged", and "down". The percentages of respondents are hence divided into P (positive), E (neutral), and N (negative), such that $\mathrm{P} + \mathrm{E} + \mathrm{N} = 100$. The net balance is then computed as

$$\text{Net balance} = \mathrm{P} - \mathrm{N}.$$

Alternatively, a question may have three possible answers, e.g. "increased sharply", "increased", "unchanged", "decreased", and "decreased sharply". In this case, the answers are divided as PP (very positive), P (positive), E (neutral), N (negative), NN (very negative), such that $\mathrm{PP} + \mathrm{P} + \mathrm{M} + \mathrm{N} + \mathrm{NN} = 100$. The net balance is then computed as

$$\text{Net balance} = (\mathrm{PP} + \frac{1}{2}\mathrm{P}) - (\frac{1}{2}\mathrm{N} + \mathrm{NN}).$$

Business surveys contain several questions related to the assessment or expectations of orders, production, prices, liquidity, investments, employment, and the general state of the economy. In our dataset, we will only use the indicators derived from the survey conducted by ISTAT on Italian manufacturing firms. The sample comprises 4000 firms of different sizes (5-9 employees; 10-49 employees; 50-249 employees; 250-999 employees; at least 1,000 employees), geographical distribution (Northwest; Northeast; Center; South), and main activity. The first variable, PROEX, is related to production expectations, and it is extracted from the following question,

In the next three months, production will:

(a) Increase (b) Stay unchanged (c) Decrease

It is calculated as

$$\mathrm{PROEX}_t = (\% \text{ of firms that responded (c)})_t - (\% \text{ of firms that responded (a)})_t.$$

Using the same formula, the indicator related to the assessment of orders, ORDAS, is obtained with the question

How is the current level of orders?

(a) High (b) Normal (c) Low

Hence,

$$\text{ORDAS}_t = (\% \text{ of firms that responded (c)})_t - (\% \text{ of firms that responded (a)})_t.$$

The indicator related to the assessment of production, PROAS, is derived from

How is the current level of production?

(a) High (b) Normal (c) Low.

and thus it is calculated as

$$\text{PROAS}_t = (\% \text{ of firms that responded (c)})_t - (\% \text{ of firms that responded (a)})_t.$$

3.2.2 Producer Price Index

The Producer Price Index (PPI) tracks the monthly changes in output prices for industrial goods produced in Italy and sold both domestically and internationally. ISTAT publishes a system of monthly indicators broken down into five categories: (1) the domestic market, (2) the Euro area foreign market, (3) the non-Euro area foreign market, and two summary indices for (4) the foreign market (Euro area plus non-Euro area) and for (5) the total (domestic market plus foreign market). Here we will use the total producer price index, as the prices of both domestic and foreign goods impact the cost and level of production of Italian firms. The products included in the calculation are from the extractive, manufacturing, and electricity, gas, and water sectors for the domestic market, and from the extractive and manufacturing sectors for the foreign market (excluding in both markets the sectors related to shipbuilding, aerospace, railways, armaments, and industrial services). The prices are net of VAT.

PPI is computed at different stages of aggregation. First, data is aggregated for each observation unit h,[5]

$$I_{h,t} = \frac{\sum_i (v_{i,0} \cdot \frac{p_{i,t}}{p_{i,0}})}{\sum_i v_{i,0}},$$

[5] The definition of "observation unit" can be found at https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Observation_unit

where $v_{i,0}$ is value of sales of product/service i in the base period, while $p_{i,t}$ and $p_{i,0}$ are the prices of good i at time t and 0. The second step is to compute an index for each NACE sector by summing the observation units,

$$I_{k,t} = \frac{\sum_i (v_{h,0} \cdot I_{h,t})}{\sum_i v_{h,0}}.$$

Then, the index for the domestic market ($m = D$) and for the foreign market ($m = F$) can be computed as

$$\text{PPI}_{m,t} = \frac{\sum_{k=1}^{K} (v_{k,0} \cdot I_{k,t})}{\sum_i v_{k,0}}.$$

The overall Producer Price Index then is the weighted average of the index for the domestic market and the foreign market,

$$\text{PPI}_t = b \cdot \text{PPI}_{D,t} + (1 - b) \cdot \text{PPI}_{F,t},$$

where b is the weight associated with the domestic market.

3.2.3 Industrial Turnover

Turnover can be defined as the sum of invoices of a company. A turnover index is simply an indicator developed to track the sales of a given economic sector as a whole (industry, services, etc.). In this context, it is important to note the difference between industrial turnover and industrial production: turnover can be used to track the volume of actual sales, while production measures the amount of goods manufactured, whether they are sold in that same month or not. In this sense, we do not expect them to match perfectly; as Figure 2 shows, there is a close relationship between the two variables, as they often move up and down together. However, following the beginning of 2021, industrial production has been stagnant, while turnover increased above its pre-pandemic levels, indicating that the correlation between them has changed.

As already mentioned, the main novelty of this dataset is the Real Time Turnover Index (RTT). This monthly index of industrial turnover is a real-time economic activity indicator for Italy, developed by TeamSystem and Centro Studi Confindustria, based on electronic invoicing data from businesses. The RTT Index uses data from electronic invoices issued by around 200,000 corporations ("societá di capitali") that are clients of TeamSystem. These companies, varying in size and sector, represent about 20% of all corporations in Italy.

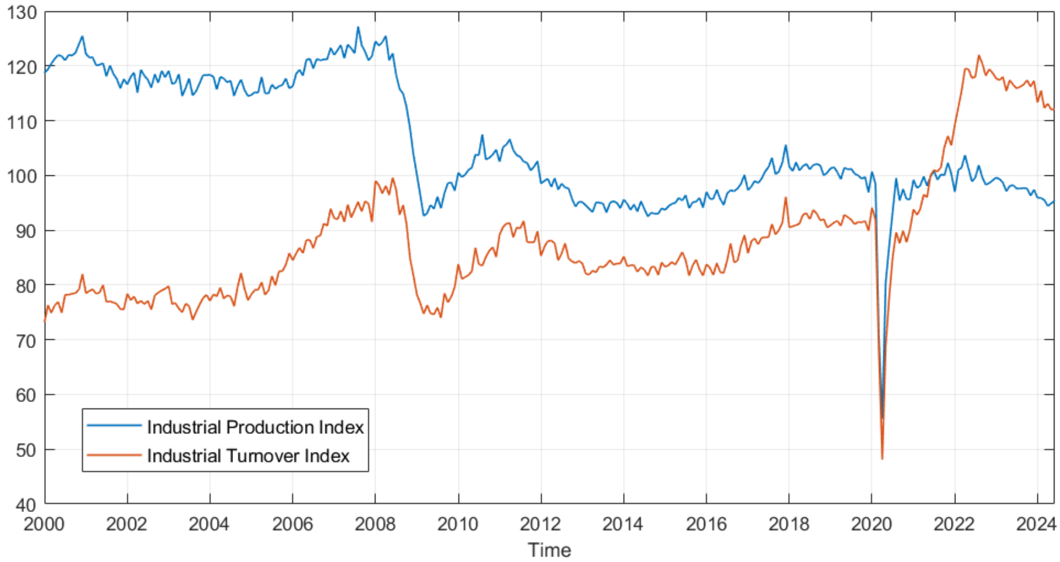


Fig. 2: Monthly Italian Industrial Production vs. Italian Industrial Turnover (2000–2024; 2021=100).

Source: ISTAT

Monthly turnover data are collected, starting from January 2020, and updated regularly. The index is calculated as the average amount of electronic invoices produced to other businesses, public administrations, and consumers (when invoices are issued) over a month. Corrections are made for credit notes and debit notes to ensure accuracy. The RTT Index is then processed by removing anomalies and adjusting for the number of working days, seasonal variations, and inflation.

Apart from the aggregate turnover index for the Italian economy, there are also three types of disaggregation available for this data, resulting in a total of 11 detailed indexes: (1) geographical location (North-West, North-East, Center, South-Islands), (2) sector of activity (agriculture, industry, construction, services), and (3) company size (small, medium, large). In this thesis, we will use the disaggregated index for the Italian industry, computed as

$$\text{TUR}_t = \frac{1}{N} \sum_{i=1}^{N} T_{i,t}$$

where $T_{i,t}$ is the total amount of electronic invoices (in terms of Euros) emitted by industrial firm i in the month t, and N is the number of firms in the sample.

The RTT Index aligns closely with official ISTAT industrial turnover index and services turnover index, showing a strong correlation (98% for industry, 49% for services from 2021-

2023). It also aligns well with data on construction production. Unlike ISTAT, which provides turnover data with significant delays and only for specific sectors, the RTT Index offers timely and comprehensive monthly data for the entire economy, across regions, sectors, and company sizes. It fills a gap in official data availability.

All the variables are summarized in Table 2.

Variable	Name	Type	Source	Adjustment	Frequency	Sample period
IPI	Industrial Production	Index	ISTAT	Seasonal adjustment	Monthly	01-2010: 04-2024
PROEX	Production Expectations	Net balance	Business Survey (ISTAT)	Seasonal adjustment	Monthly	01-2010: 04-2024*
ORDAS	Assessment of Orders	Net balance	Business Survey (ISTAT)	Seasonal adjustment	Monthly	01-2010: 04-2024*
PROAS	Assessment of Production	Net balance	Business Survey (ISTAT)	Seasonal adjustment	Monthly	01-2010: 04-2024*
PPI	Producer Price Index	Index	ISTAT	No adjustment	Monthly	01-2010: 04-2024
TUR	Industrial Turnover	Average value	Centro Studi Confindustria	Seasonal adjustment	Monthly	03-2020: 04-2024

* Data for April 2020 not available

Table 2: Overview and description of variables used in the empirical analysis.

3.3 Stationarity and unit root tests

Figure 3 depicts the plots of the raw time series, while Figure 4 shows their first differences, covering the period from February 2010 to April 2024 (except for TUR, which is available from March 2020). The level plots of the variables show clear trends and fluctuations, suggesting potential non-stationarity in the raw data. For instance, industrial turnover (TUR) appears to follow a growing trend over time, while ORDAS, PROEX, and

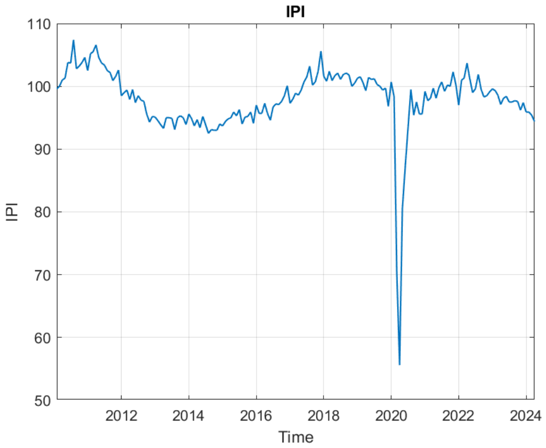
PROAS exhibit more cyclical movements. PPI seems to be sable for most of the sample but with a growing trend starting from 2021. This is coherent with the energy crisis: the sudden spike in overall producer prices can be explained by the increase in gas prices, arguably one of the most relevant factors determining the costs of production. On the other hand, when examining the first-differenced series (Δ variables), the plots exhibit more stable behavior around zero, confirming that differencing these variables induces stationarity. The plots of the ACF and PACF seem to confirm the absence of trends and seasonal patterns (see Appendix A.2), hence there is no need to make other transformations such as taking natural logarithms of the variables. This result aligns with the results of the ADF and PP tests (Table 3), which indicate that most of these variables are non-stationary in levels but become stationary when differences.

	ADF test*		**PP test**	
Variable	p-value	stat	p-value	stat
IPI	0.446	-0.559	0.530	-0.328
PROEX	0.024	-2.250	0.014	-2.470
ORDAS	0.139	-1.444	0.045	-1.990
PROAS	0.136	-1.455	0.024	-2.243
PPI	0.875	0.751	0.943	1.221
TUR	0.924	1.084	0.976	1.679
ΔIPI	0.001	-4.430	0.001	-13.732
ΔPROEX	0.001	-4.550	0.001	-16.921
ΔORDAS	0.001	-3.475	0.001	-12.305
ΔPROAS	0.006	-2.773	0.001	-10.356
ΔPPI	0.001	-3.932	0.001	-4.170
ΔTUR	0.030	-1.272	0.001	-9.192

0.05 significance level

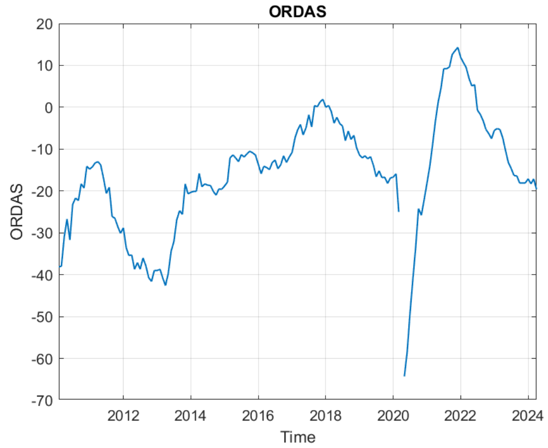
* 12 lags

Table 3: Augmented Dickey-Fuller and Phillips-Perron Test results: a variable is stationary if its p-value is below the significance level (0.05).

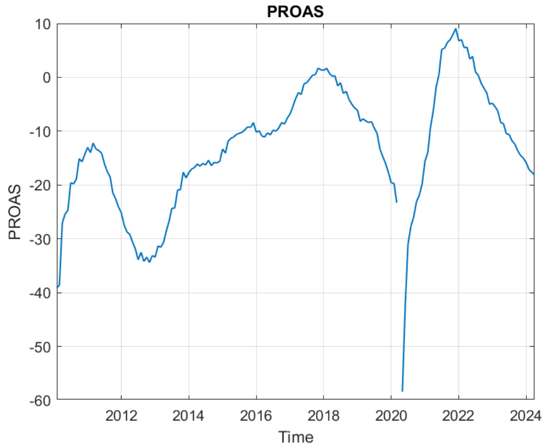


(a) IPI

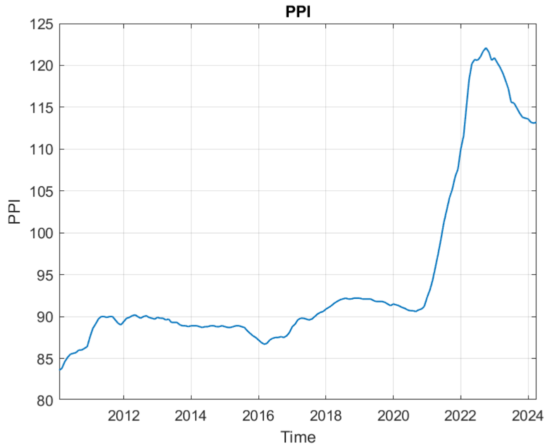
(b) PROEX



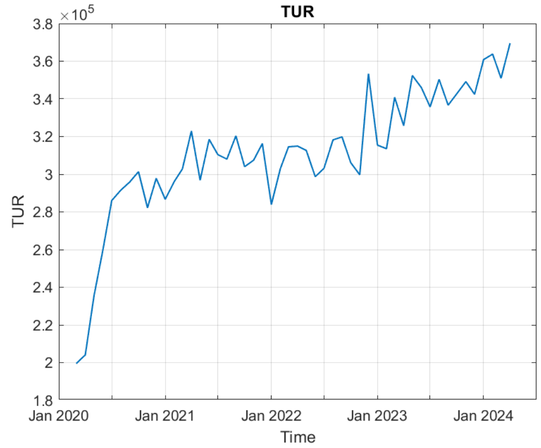
(c) ORDAS



(d) PROAS

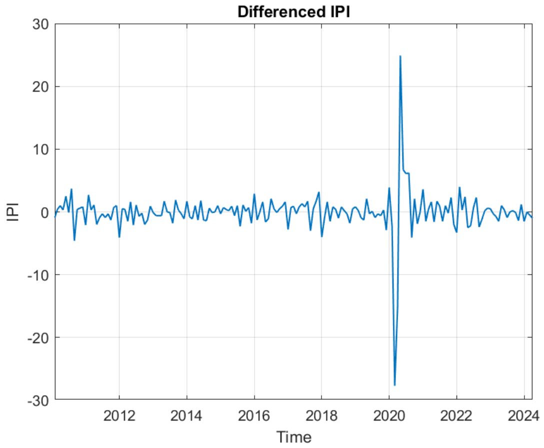


(e) PPI

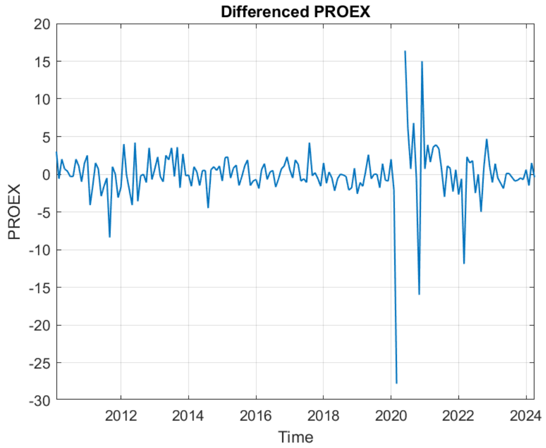


(f) TUR

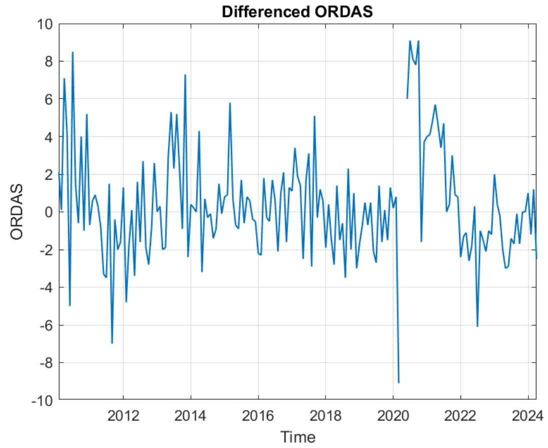
Fig. 3: Plots of raw series for Italy. Sample period: February 2010–April 2024.



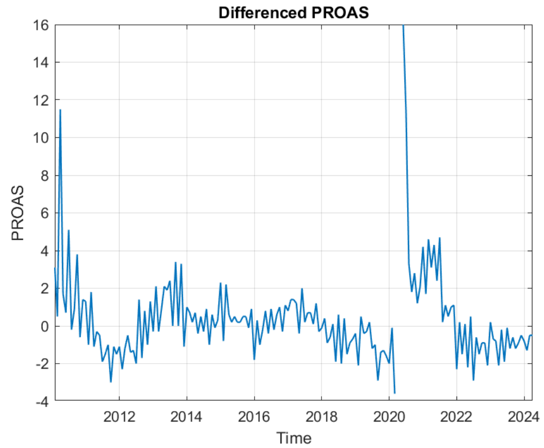
(a) ΔIPI



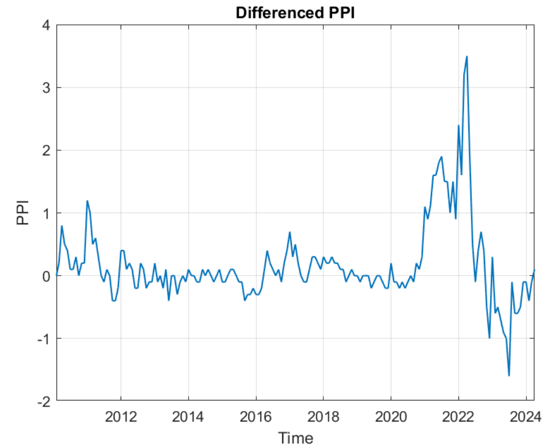
(b) ΔPROEX



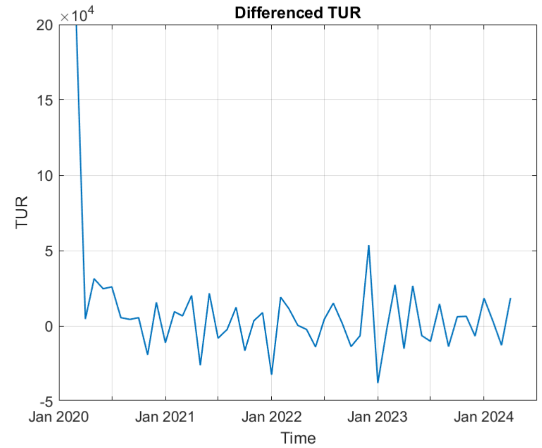
(c) ΔORDAS



(d) ΔPROAS



(e) ΔPPI



(f) ΔTUR

Fig. 4: Plots of differenced series for Italy. Sample period: February 2010–April 2024.

4 Empirical analysis

This Section focuses on whether including electronic invoicing data, i.e. the industrial turnover index, in the model leads to a higher forecast accuracy. We first introduce the three models over the entire sample to compare the in-sample goodness of fit. Then we forecast the one-step-ahead value of industrial production using a rolling window and compare the out-of-sample performance of these models.

4.1 In-sample performance

We propose an analysis of IPI over the entire sample, from February 2010 to April 2024, using the first difference variables to ensure stationarity. Let us define $\Delta y_t = y_t - y_{t-1} = \mathrm{IPI}_t - \mathrm{IPI}_{t-1}$ as the independent variable. The first model proposed is an ARMA(p,q) model used as a benchmark, where the choice of p and q is determined by the information criteria.

The second model is an ARMAX(p,q,r),

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{r} \boldsymbol{\gamma}_k \mathbf{X}_{t-k} + \varepsilon_t \quad \text{for } t = 1, \dots, T,$$

where the exogenous variables are

$$\mathbf{X_t} = \begin{bmatrix} \Delta\mathrm{PROEX}_t \\ \Delta\mathrm{ORDAS}_t \\ \Delta\mathrm{PROAS}_t \\ \Delta\mathrm{PPI}_t \end{bmatrix}. \tag{18}$$

The third model is an ARMAX(p,r,q) where we introduce the index of industrial turnover, starting from February 2020 (t^*),

$$y_t = \begin{cases} c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{r} \gamma_k \mathbf{X_{t-k}} + \varepsilon_t & \text{for } t \in [1, t^*], \\ c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{r} \gamma_k \mathbf{X_{t-k}} + \sum_{k=1}^{r} \delta_k \Delta TUR_{t-k} + \varepsilon_t & \text{for } t \in (t^*, T]. \end{cases} \tag{19}$$

where $\mathbf{X_t}$ is defined in Equation 18 and d_t is a dummy variable such that

$$\begin{cases} d_t = 0 & \text{for } t \in [1, t^*] \\ d_t = 1 & \text{for } t \in (t^*, T] \end{cases}.$$

Combining the two equations in 19 we obtain the following formulation,

$$
\begin{aligned}
y_t = & c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{r} \gamma_k \mathbf{X_{t-k}} + \varepsilon_t \\
& + d_t \cdot (c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{r} \gamma_k \mathbf{X_{t-k}} + \sum_{k=1}^{r} \delta_k \Delta TUR_{t-k} + \varepsilon_t) \quad \text{for } t = 1, \ldots, T.
\end{aligned}
$$

To account for the sharp decline in industrial production between February and March 2020 (refer to Figure 3) due to the outbreak of COVID-19, we include as an additional regressor in the models listed above a dummy variable that takes value 1 in March 2020 and 0 in all the other observations,

$$
\begin{cases}
d_t(\text{COVID}) = 1 \text{ if } t = \text{March 2020} \\
d_t(\text{COVID}) = 0 \text{ if } t \neq \text{March 2020}
\end{cases}.
$$

4.1.1 ARMA

We estimate several ARMA models using both the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to identify the optimal combination of autoregressive and moving average terms, (p, q). The results suggest that the two criteria point to the same choice. As Table 4 suggests, the lowest value of AIC is obtained when $p = 1$ and $q = 3$ (886.0026). BIC criterion confirms the same result: according to Table 5, $(1, 3)$ minimize BIC, with a value equal to 899.0284.

		q				
		0	1	2	3	4
	0	919.0559	917.5483	892.7251	890.7298	891.8482
	1	919.8435	910.1257	889.5794	886.0026	887.7954
p	2	900.7721	897.2238	887.4172	887.9971	888.4043
	3	902.6846	903.4105	888.1997	888.2113	889.1603
	4	894.8661	894.6636	890.1908	890.2013	889.4243

Table 4: Akaike information criterion (AIC) for ARMA models of order p, q, where $p = 0, \ldots, 4$, and $q = 0, \ldots, 4$.

Table 6 presents the results of the estimated maximum likelihood coefficients for the ARMA(1,3) model. The AR(1) coefficient is positive (0.792) and its associated p-value is

		q				
		0	1	2	3	4
	0	921.6611	922.7586	900.5406	901.1505	904.8741
	1	925.0538	917.9412	900.0001	899.0284	903.4264
p	2	908.5876	907.6445	900.4431	903.6281	906.6405
	3	913.1053	916.4363	903.8307	906.4475	910.0016
	4	907.8919	910.2946	908.4270	911.0427	912.8708

Table 5: Bayesian information criterion (BIC) for ARMA models of order p, q, where $p = 0, \ldots, 4$, and $q = 0, \ldots, 4$.

very close to zero. This indicates a strong autoregressive relationship: industrial production at time t depends strongly on its own past value at $t - 1$. The MA coefficients are all strongly significant: the first two lags are negative (-0.775 and -0.489) and significant, while the third is positive (0.264).

	Estimate	SE	t-stat	p-value
AR(1)	0.792	0.044	18.149	0.000
MA(1)	-0.775	0.066	-11.702	0.000
MA(2)	-0.489	0.071	-6.934	0.000
MA(3)	0.264	0.055	4.814	0.000
AIC	886.003			
BIC	899.028			

Table 6: Summary of the estimated parameters of ARMA(1,3) through maximum likelihood, standard errors, t-statistic, and p-value.

Table 7 presents the results of the Ljung-Box Q-test for the ARMA model with lags 6, 12, and 24. The results suggest that the p-values are all very close to 1, suggesting that there is no significant autocorrelation remaining in the residuals. Hence the model is well specified.

4.1.2 ARMAX(1,1,1)-A

Let us define the following model as ARMAX(1,1,1)-A. The model includes one lag of the moving average component ($q = 1$), one lag of the autoregressive component ($p = 1$), and one

	Ljung-Box Q-Test		
	$L = 6$	$L = 12$	$L = 24$
stat	1.502	4.730	11.796
p-value	0.959	0.966	0.982
c-value	12.592	21.026	36.415

$\alpha = 0.05$

Table 7: Results of the Ljung-Box Q-test for ARMA(1,3).

lag for the exogenous variables ($r = 1$).The ARMAX(1,1,1)-A model is estimated with and without the inclusion of the COVID-19 dummy variable. When COVID-19 is not accounted for (left side of Table 8), the results show that besides the Producer Price Index and the constant term, all variables are statistically significant at the 95% confidence level. Once we add the dummy for March 2020 (right side of Table 8), the significance of the expectations on production (PROEX) drops. The COVID-19 dummy itself is highly significant, which suggests that it is indeed necessary to include it in the model. In terms of goodness of fit, including the COVID-19 dummy reduces both the AIC and BIC values: from 611.406 to 600.189 for AIC, and from 636.398 to 628.305 for BIC.

	Not accounting for COVID				Accounting for COVID			
	Estimate	SE	t-stat	p-value	Estimate	SE	t-stat	p-value
Constant	-0.001	0.053	-0.017	0.986	-0.020	0.039	-0.507	0.612
AR(1)	0.231	0.105	2.191	0.028	0.316	0.084	3.762	0.000
MA(1)	-0.632	0.102	-6.224	0.000	-0.727	0.075	-9.653	0.000
ORDAS	-0.118	0.045	-2.603	0.009	-0.089	0.052	-1.729	0.084
PROEX	-0.098	0.031	-3.136	0.002	-0.006	0.044	-0.142	0.887
PROAS	0.454	0.066	6.839	0.000	0.294	0.082	3.604	0.000
PPI	-0.096	0.068	-1.415	0.157	-0.016	0.055	-0.298	0.766
d(COVID)					5.683	1.333	4.264	0.000
AIC	611.406				600.189			
BIC	636.398				628.305			

Table 8: Summary of the estimated parameters of ARMAX(1,1,1) through maximum likelihood, standard errors, t-statistic, and p-value.

According to Ljung-Box Q-test (Table 9), there is evidence of autocorrelation in the residuals when we do not account for COVID. For all lag lengths (L=6, 12, and 24) the Q statistics exceed the critical values; for L=24, the p-value is somewhat higher, but still below 0.05. Hence we reject the null hypothesis of no autocorrelation at the 5% significance level. On the other hand, the results indicate that once the dummy for March 2020 is introduced, the model exhibits lower test statistics and higher p-values. For L=6 and 12, the test still suggests the rejection of the null hypothesis, while for L=24, the p-value increases to 0.103, suggesting that the null hypothesis of no autocorrelation cannot be rejected at this longer lag. This implies that accounting for COVID reduces autocorrelation in the residuals, especially at longer lags.

	Ljung-Box Q-Test					
	Not accounting for COVID			Accounting for COVID		
	$L=6$	$L=12$	$L=24$	$L=6$	$L=12$	$L=24$
stat	32.470	33.501	37.015	28.107	29.254	33.044
p-value	0.000	0.001	0.044	0.000	0.004	0.103
c-value	12.592	21.026	36.415	12.592	21.026	36.415

$\alpha = 0.05$

Table 9: Results of the Ljung-Box Q-test for ARMAX(1,1,1)-A.

The overall results suggest that this model is better at estimating the in-sample values of IPI than the benchmark ARMA(1,3) model, especially if we account for the COVID-19 outbreak.

4.1.3 ARMAX(1,1,1)-B

Model ARMAX(1,1,1)-B extends ARMAX(1,1,1)-A by introducing the index of industrial turnover, starting from time t*. We present the results in Table 10. The AIC and BIC values are lower than the other models presented so far, indicating a superior goodness of fit for this model. Again, accounting for the COVID shock of March 2020 seems to increase the fit. However, Table 10 suggests that the dummy variable has a large p-value, thus it is not statistically significant. The introduction of such a dummy also alters the estimates and significance of the other variables. If we look at the left side of Table 10, we see that industrial turnover is highly significant, while on the right side it is not.

	Not accounting for COVID				Accounting for COVID			
	Estimate	SE	t-stat	p-value	Estimate	SE	t-stat	p-value
Constant	-0.064	0.040	-1.604	0.109	-0.044	0.042	-1.051	0.293
AR(1)	0.086	0.096	0.894	0.371	0.029	0.119	0.245	0.807
MA(1)	-0.648	0.083	-7.801	0.000	-0.576	0.101	-5.695	0.000
ORDAS	0.068	0.064	1.056	0.291	0.052	0.054	0.962	0.336
PROEX	-0.027	0.068	-0.393	0.694	-0.007	0.057	-0.119	0.905
PROAS	0.124	0.088	1.401	0.161	0.136	0.079	1.731	0.084
PPI	0.471	0.209	2.251	0.024	0.462	0.205	2.253	0.024
d(COVID)					29.477	75.410	0.391	0.696
d_t·IPI	0.481	0.095	5.083	0.000	0.985	1.600	0.616	0.538
d_t·ORDAS	-0.182	0.110	-1.658	0.097	-0.053	0.884	-0.060	0.952
d_t·PPI	-0.511	0.237	-2.157	0.031	-0.442	0.875	-0.505	0.614
d_t·PROAS	0.184	0.145	1.271	0.204	-0.144	1.504	-0.096	0.924
d_t·PROEX	-0.078	0.086	-0.904	0.366	0.010	0.544	0.018	0.985
d_t·TUR	0.000	0.000	6.130	0.000	0.000	0.000	0.012	0.991
AIC	550.499				521.700			
BIC	594.235				568.560			

Table 10: Summary of the estimated parameters of ARMAX(1,1,1) through maximum likelihood, standard errors, t-statistic, and p-value.

The Ljung-Box Q-test presented in Table 11 suggests there is evidence of autocorrelation in the residuals for both models at all lags. The Q statistics are well above the critical values for all tests. Unlike the ARMAX(1,1,1)-A model, which showed some improvement when accounting for COVID-19, particularly at longer lags, the ARMAX(1,1,1)-B model exhibits residual autocorrelation both with and without the COVID adjustment.

4.2 Out-of-sample performance

As explained in Section 2, ARMA and ARMAX models are widely recognized as key tools to obtain macroeconomic forecasts. We conduct a dynamic out-of-sample forecast analysis to predict the future behavior of IPI. The analysis is based on monthly data from February 2010 to April 2024, with the forecast period set from June 2018 to April 2024. We present three

	Ljung-Box Q-Test					
	Not accounting for COVID			Accounting for COVID		
	$L = 6$	$L = 12$	$L = 24$	$L = 6$	$L = 12$	$L = 24$
stat	45.956	46.912	50.150	47.910	48.725	51.943
p-value	0.000	0.000	0.001	0.000	0.000	0.001
c-value	12.592	21.026	36.415	12.592	21.026	36.415

$\alpha = 0.05$

Table 11: Results of the Ljung-Box Q-test for ARMAX(1,1,1)-B.

models, an ARMA(1,3) model, an ARMAX(1,0,1)-A model with

$$\mathbf{X_t} = \begin{bmatrix} \Delta\text{PROEX}_t \\ \Delta\text{ORDAS}_t \\ \Delta\text{PROAS}_t \\ \Delta\text{PPI}_t \end{bmatrix}, \tag{20}$$

and an ARMAX(1,0,1)-B model with the same exogenous regressors proposed in 20 up to February 2020, and afterwards

$$\mathbf{X_t} = \begin{bmatrix} \Delta\text{PROEX}_t \\ \Delta\text{ORDAS}_t \\ \Delta\text{PROAS}_t \\ \Delta\text{PPI}_t \\ \Delta\text{TUR}_t \end{bmatrix}.$$

Unlike the previous analysis, there is no need to use a temporal dummy to account for the structural change at t^*. Indeed, the rolling window analysis should already account for the time varying parameters by itself. The rolling window size chosen is 100, in order to smooth out the effect of the disruption of COVID-19.

The results of the out-of-sample dynamic forecast analysis presented in Table 12 provide a comparison of the forecasting performance for three different models. The evaluation metrics compare the results for the entire sample and a subset of the data excluding April-May 2020, which represents a period of extreme volatility and abnormal economic activity due to the COVID-19 outbreak. Over the entire sample, the ARMAX(1,0,1)-A model shows the lowest RMSE of 4.639, outperforming both ARMA(1,3) and ARMAX(1,0,1)-B, with RMSE values of

6.140 and 6.070, respectively. Hence we deduce that ARMAX(1,0,1)-A explains the patterns of industrial production better than the other models, producing fewer large forecasting errors.

	RMSE		
	ARMA(1,3)	ARMAX(1,0,1)-A	ARMAX(1,0,1)-B
Entire sample	6.140	4.639	6.070
Excluding Apr-May 2020	5.345	4.556	4.478

Table 12: Measuring forecast error for the models, $w = 100$

The second part of the analysis excludes April and May 2020. The first thing to notice is that the reduction in RMSE values, particularly for ARMA(1,3) and ARMAX(1,0,1)-B, confirms that the large errors observed earlier were influenced by those volatile months. The second result is that ARMAX(1,0,1)-B improves significantly and even outperforms the other two models. This result suggests that, aside from periods of extreme disruption, incorporating electronic invoices as an exogenous regressor could improve forecast accuracy.

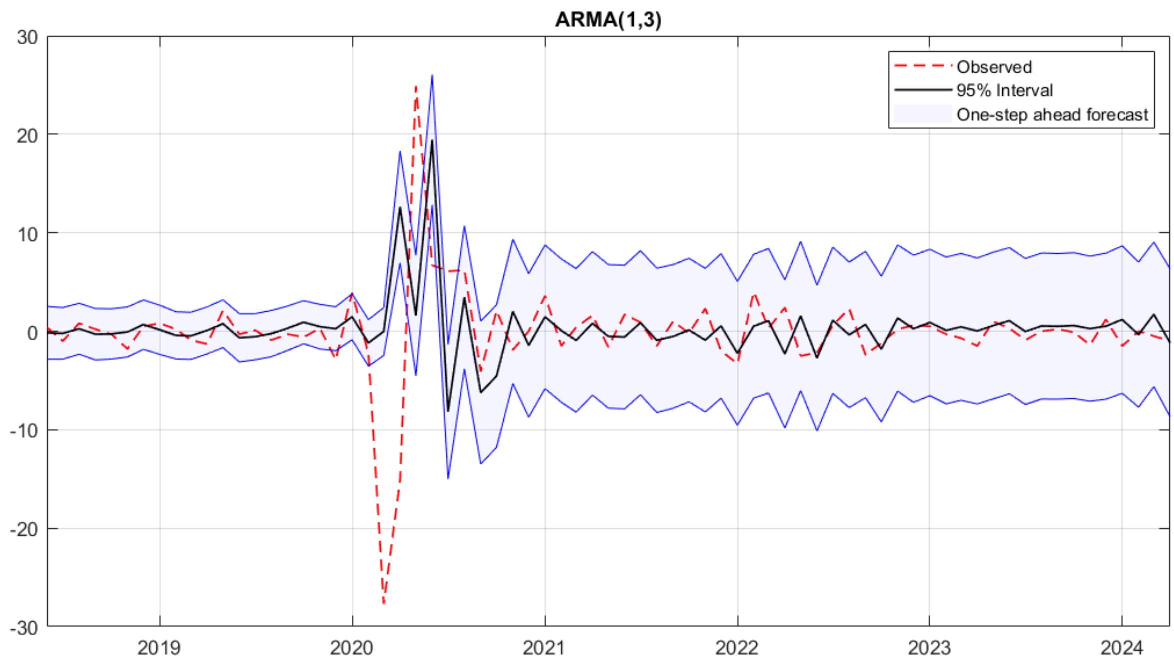


Fig. 5: One-step-ahead forecasts of ARMA(1,3) model vs. observed values of IPI. 95% confidence interval.

These trends are also reflected visually in Figures 5, 6, and 7. In Figure 5, the ARMA(1,3) forecast exhibits larger deviations from actual data, especially during the first half of 2020, as evidenced by the larger errors observed around April-May 2020. Indeed, the ARMA(1,3) was not able to capture the large disruption of March 2020 at all. In contrast, Figure 6 shows that ARMAX(1,0,1)-A was able to produce the most accurate forecasts during the beginning of the COVID-19 pandemic. This would suggest that the sentiment of the manufacturing sector was able to predict with accuracy the shock to production. On the contrary, Figure 7 shows that model ARMAX(1,0,1)-B did not predict well the values of industrial production during the first part of 2020; indeed, it looks like there was some delay in the signal. This can be explained by the fact that, as we discussed in the previous Section, industrial turnover and industrial production do not move exactly together. Indeed, a plausible explanation is that turnover decreased significantly with some lag, due to the fact that invoices may be related to goods that were produced previously.

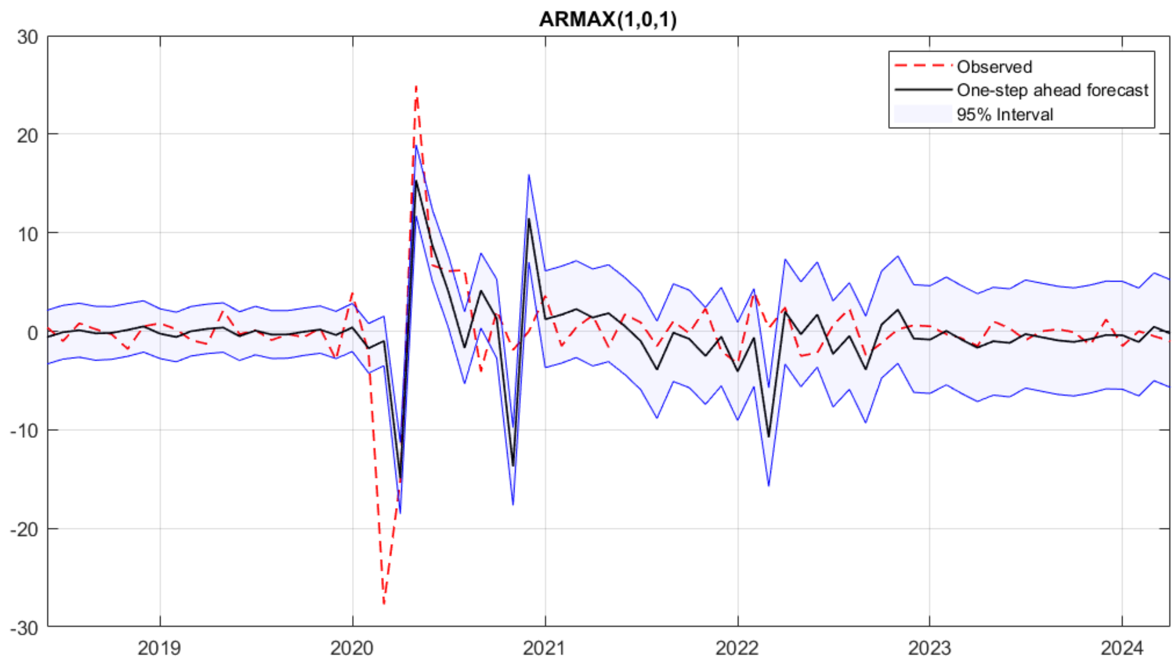


Fig. 6: One-step-ahead forecasts of ARMAX(1,0,1)-A model (industrial turnover is not included) vs. observed values of IPI. 95% confidence interval.

In general, it is clear that none of the models could predict the large shock associated to the beginning of the COVID-19 outbreak in Italy. The widening of the 95% confidence intervals observed in all three Figures following the beginning of the pandemic reflect the increased uncertainty in forecasts during this time of economic volatility, and hence the models' difficulty to adapt new dynamics emerging post-pandemic. Despite the ARMAX models generally performing better than ARMA(1,3), all models exhibit an expansion in confidence intervals. Overall, ARMAX(1,0,1)-A generally outperforms the other models in terms of both magnitude and directional accuracy, despite the challenges posed by the COVID-19 pandemic. Compared to the ARMA benchmark, the ARMAX models, which incorporate external data, appear more resilient in the face of economic shocks, and their improved performance suggests the value of incorporating additional information in forecasting. Lastly, the exclusion of April-May 2020 highlights the potential explanatory power of electronic invoices.

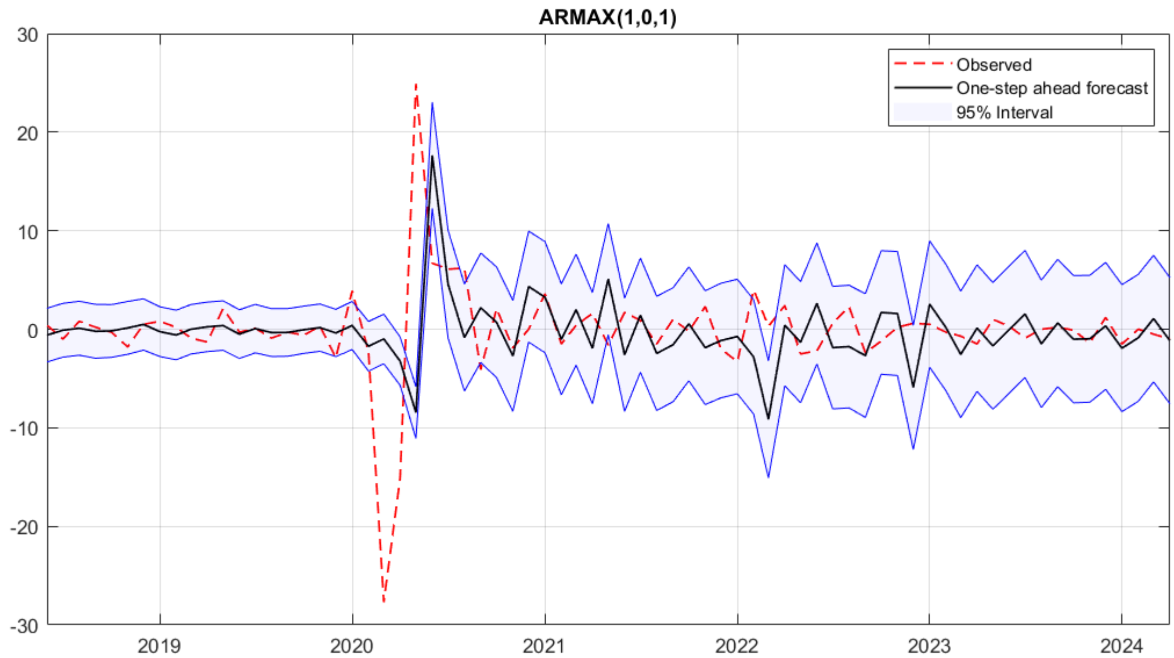


Fig. 7: One-step-ahead forecasts of ARMAX(1,0,1)-B model (industrial turnover is included) vs. observed values of IPI. 95% confidence interval.

5 Conclusion

Producing timely and reliable estimates of the current and future level of industrial production is essential to have a clear picture of the state of the economy for policymakers. Industrial production data reflects the output of manufacturing, mining, and electricity firms, which is a significant component of the gross domestic product, especially in Italy. This thesis investigates a method to forecast the one-step-ahead values of industrial production using data from electronic invoices of Italian firms.

The analysis consists of comparing a benchmark autoregressive moving average (ARMA) model with two autoregressive moving average models with exogenous inputs (ARMAX): one incorporating business survey indexes and production prices, and another that includes the industrial turnover index as an additional explanatory variable. The results indicate two main findings. First, the industrial turnover index increases the in-sample fit. Second, the forecasting exercise conducted using a rolling window to predict the one-step-ahead value shows that the model incorporating industrial turnover is able to predict out-of-sample values of IPI with some accuracy. In particular, this thesis finds that the ARMAX model with industrial turnover outperforms the other models only if we exclude March to April 2020 to our analysis. Indeed, this suggests that industrial turnover is a good predictor of IPI during stable conditions, while the information provided by business surveys and producer prices alone is better for turbulent conditions.

These findings align with the growing interest in forecasting economic activity using high-frequency, real-time data, such as electronic invoice data. Given the timeliness and granularity of this data, its integration into economic models could provide more accurate results. However, this exercise cannot be considered exhaustive: our analysis was possible due to the availability of electronic invoicing data in Italy. In the future, further research could explore the applicability of this methodology in different countries.

Appendix

A.1 MA(∞) representation of ARMA(p, q) process

Consider an ARMA(p, q) process:

$$y_t - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q},$$

where y_t represents the observed time series, $\phi_1, \ldots, \phi_p$ are the autoregressive (AR) coefficients, ε_t is a white noise error term, and $\theta_1, \ldots, \theta_q$ are the moving average (MA) coefficients.

To express this process we can use use the backshift operator B, defined as $B^k y_t = y_{t-k}$, to rewrite the ARMA(p, q) model as:

$$\Phi(B) y_t = \Theta(B) \varepsilon_t,$$

where $\Phi(B)$ and $\Theta(B)$ are polynomials of degrees p and q, respectively. These polynomials are given by:

$$\Phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p,$$

and

$$\Theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q.$$

To derive the moving average representation of y_t, we divide both sides of the equation by $\Phi(B)$, obtaining,

$$y_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t = \Psi(B) \varepsilon_t.$$

Here, $\Psi(B)$ is an infinite-degree polynomial representing the moving average coefficients of the error terms, where

$$\Psi(B) = \sum_{j=0}^{\infty} \psi_j B^j.$$

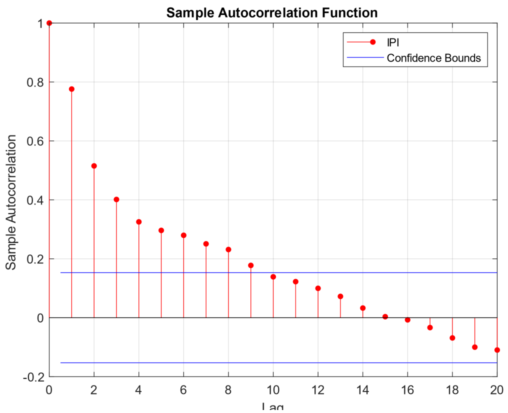
Thus, the ARMA(p, q) process can be expressed as an infinite sum of past errors:

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}.$$

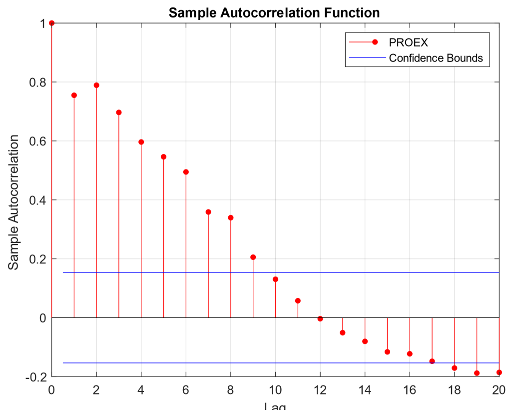
The coefficients ψ_j are determined by the recursive relationship between the AR and MA coefficients. Specifically, the recursion is given by

$$
\begin{aligned}
1 &= \psi_0 \\
\theta_1 &= \psi_1 - \psi_0\phi_1 \\
\theta_2 &= \psi_2 - \psi_1\phi_1 - \psi_0 \\
&\vdots \\
\theta_j &= \psi_j - \sum_{k=1}^{p} \phi_k \psi_{j-k}.
\end{aligned}
$$

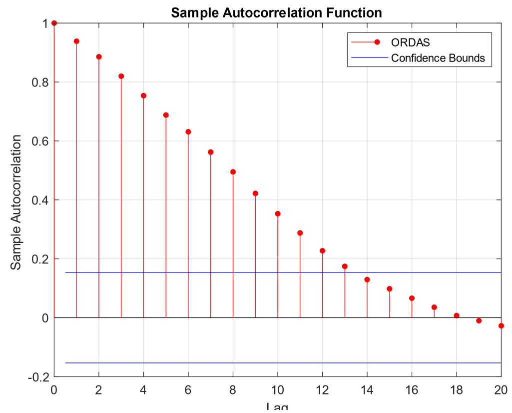
A.2 Stationarity



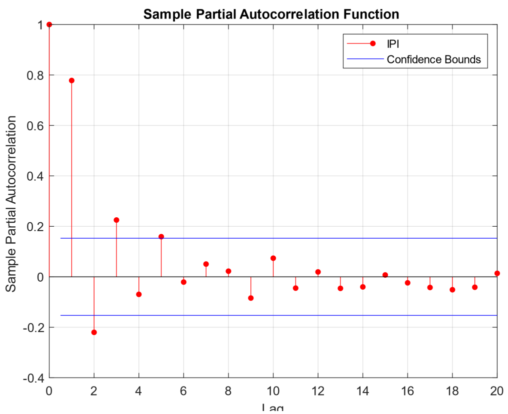
(a) IPI ACF



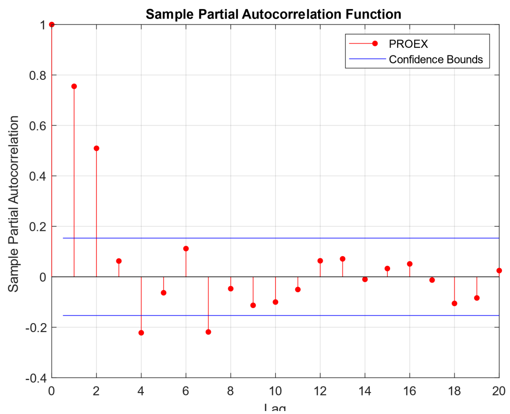
(c) PROEX ACF



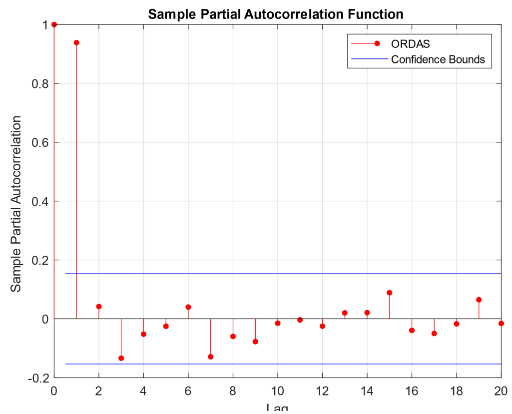
(e) ORDAS ACF



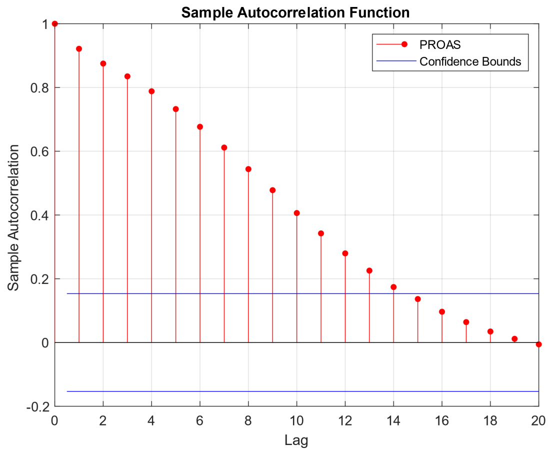
(b) IPI PACF



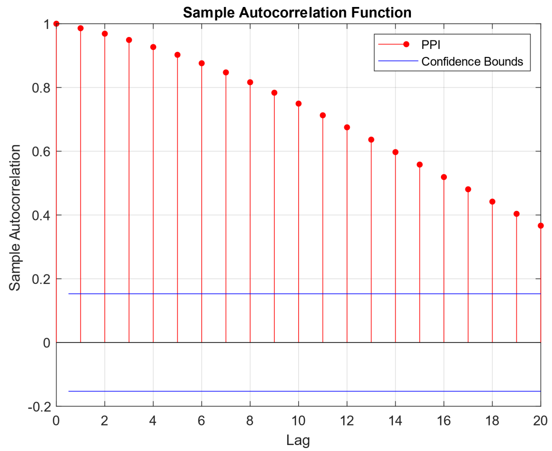
(d) PROEX PACF



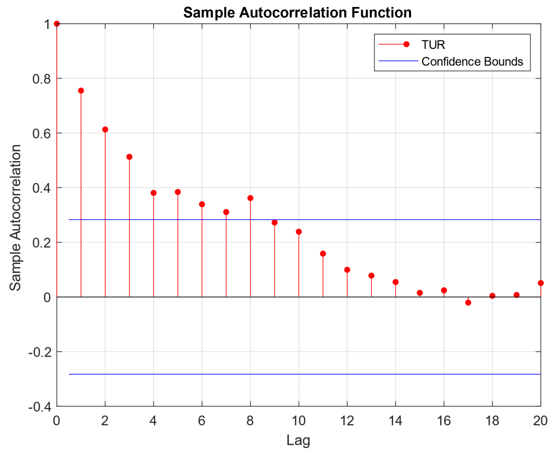
(f) ORDAS PACF



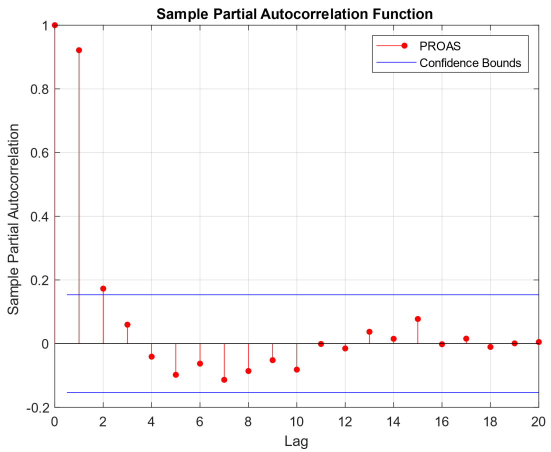
(a) PROAS ACF



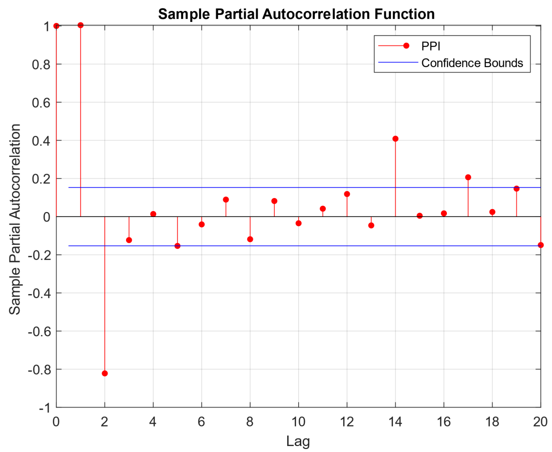
(c) PPI ACF



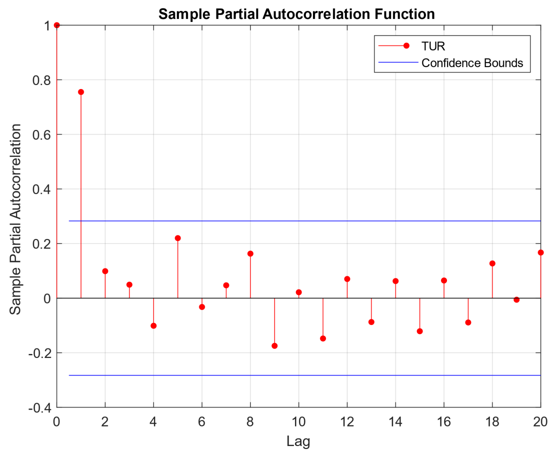
(e) TUR ACF



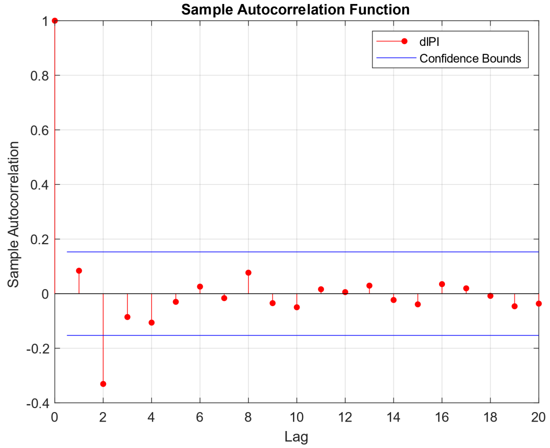
(b) PROAS PACF



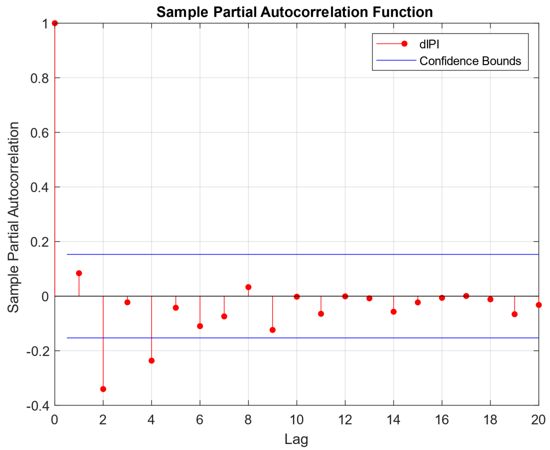
(d) PPI PACF



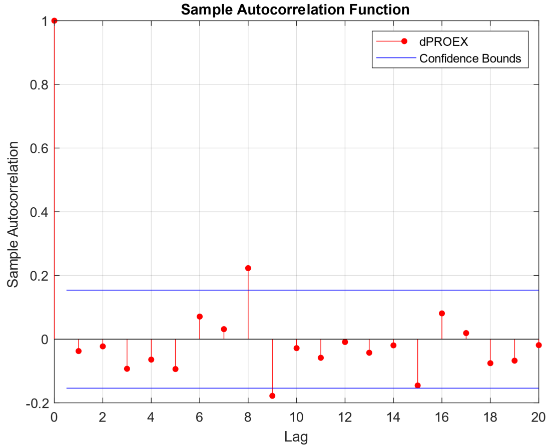
(f) TUR PACF



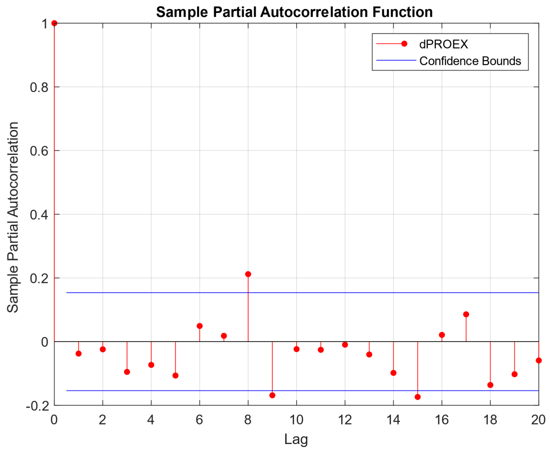
(a) ΔIPI ACF



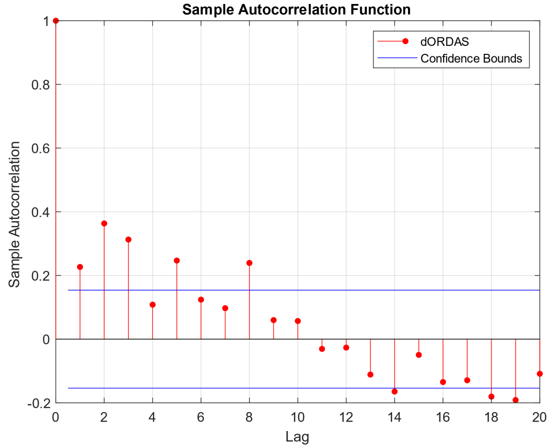
(b) ΔIPI PACF



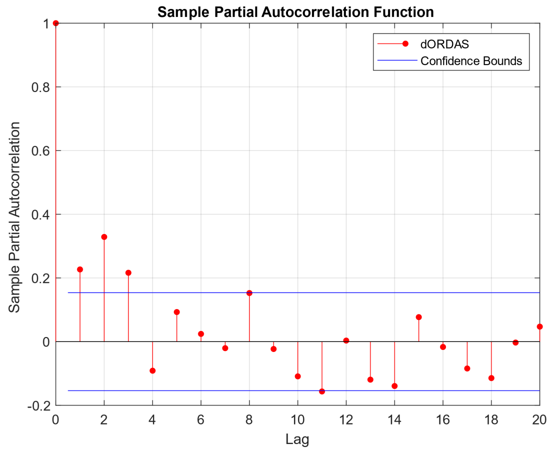
(c) $\Delta PROEX$ ACF



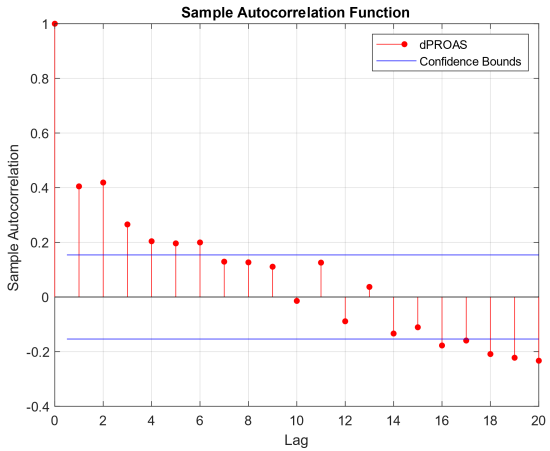
(d) $\Delta PROEX$ PACF



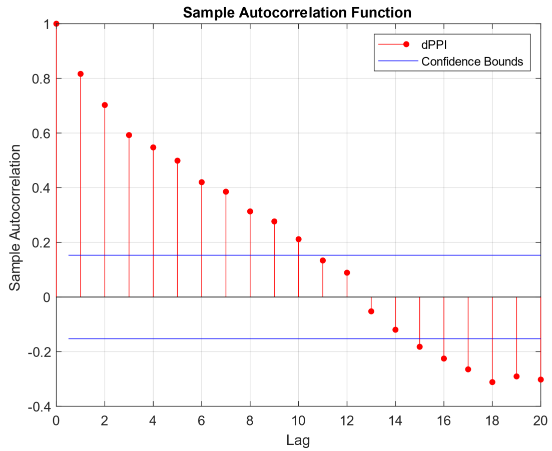
(e) $\Delta ORDAS$ ACF



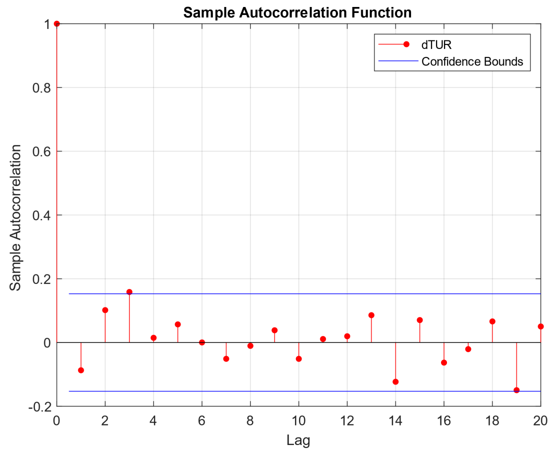
(f) $\Delta ORDAS$ PACF



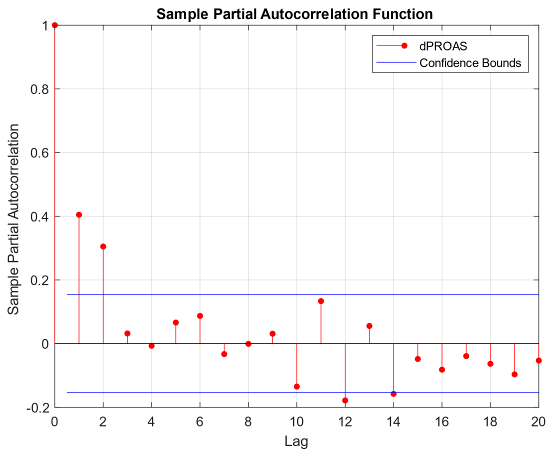
(a) $\Delta PROAS$ ACF



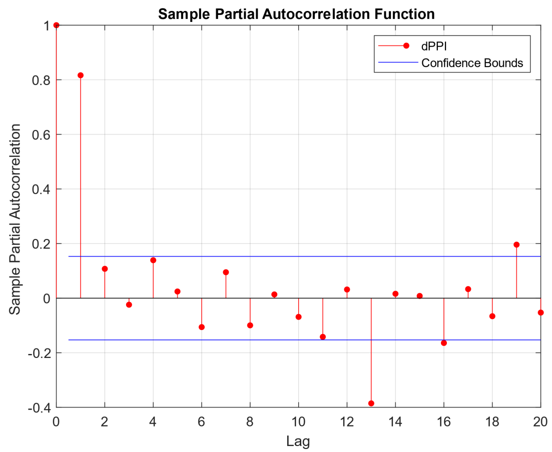
(c) ΔPPI ACF



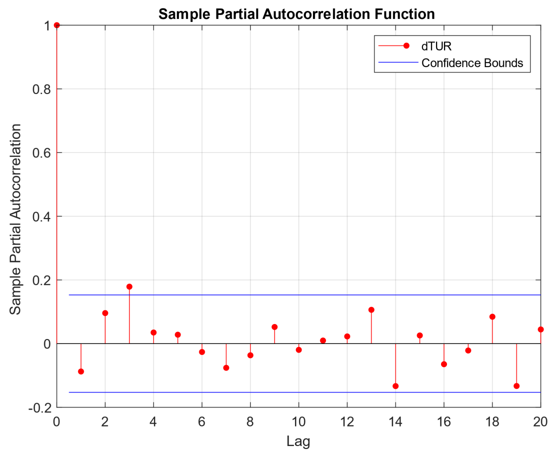
(e) ΔTUR ACF



(b) $\Delta PROAS$ PACF

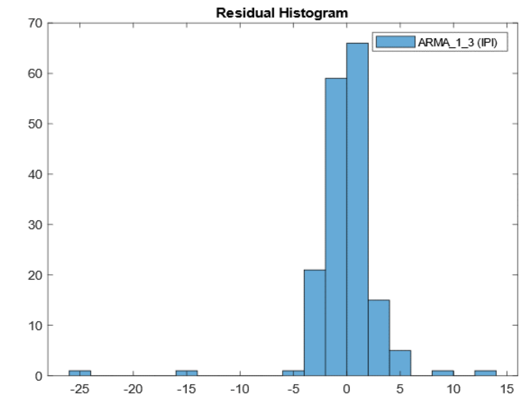


(d) ΔPPI PACF

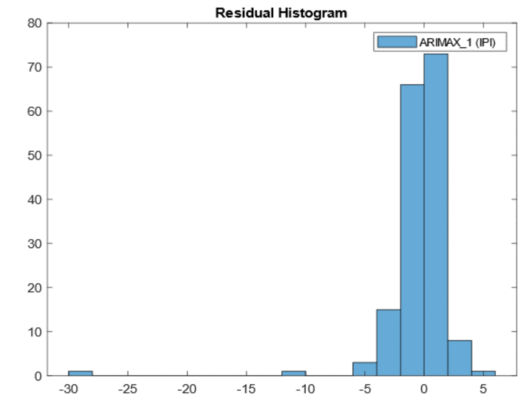


(f) ΔTUR PACF

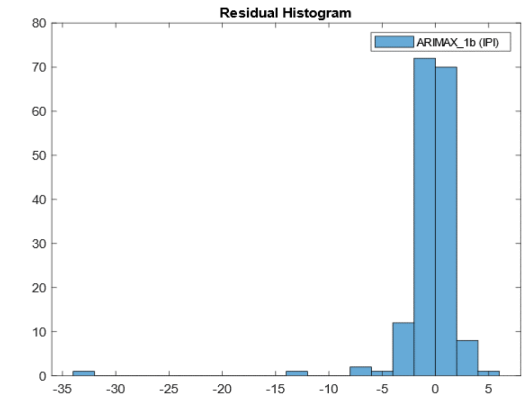
A.3 Diagnostics checking



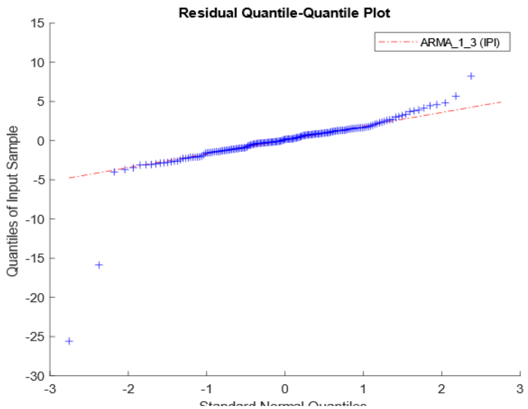
(a) ARMA(1,3) Histogram



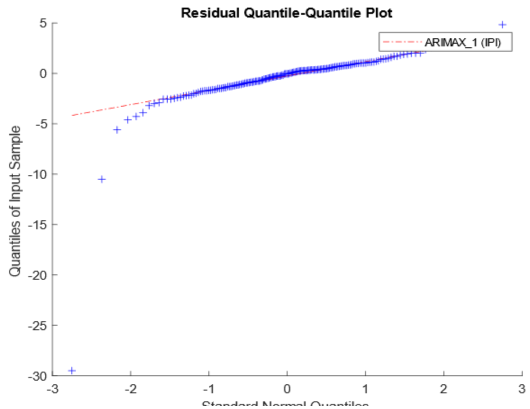
(c) ARMAX(1,0,1)-A Histogram



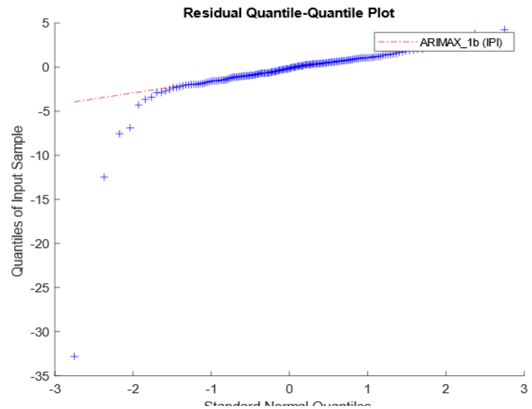
(e) ARMAX(1,0,1)-A* Histogram



(b) ARMA(1,3) Q-Q Plot

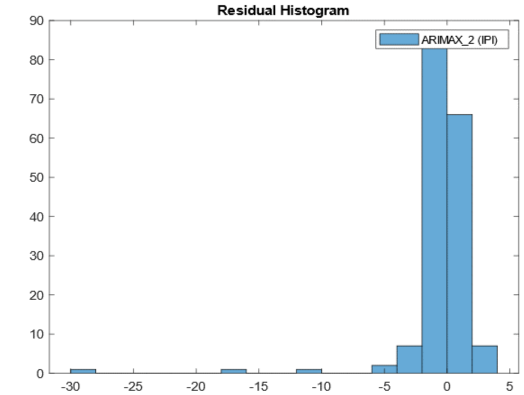


(d) ARMAX(1,0,1)-A Histogram

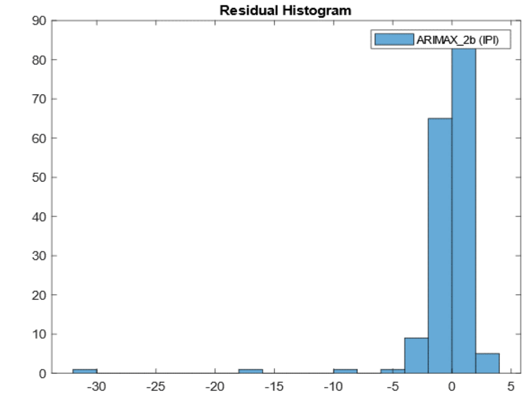


(f) ARMAX(1,0,1)-A* Histogram

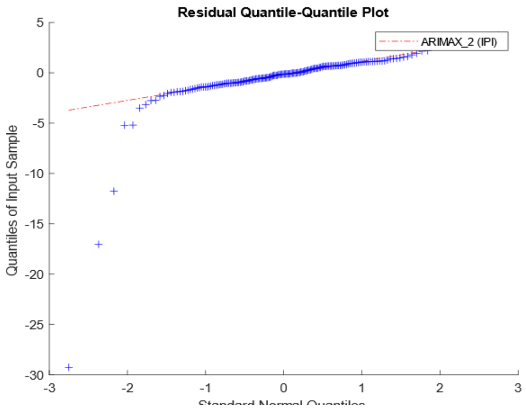
*Accounting for COVID-19



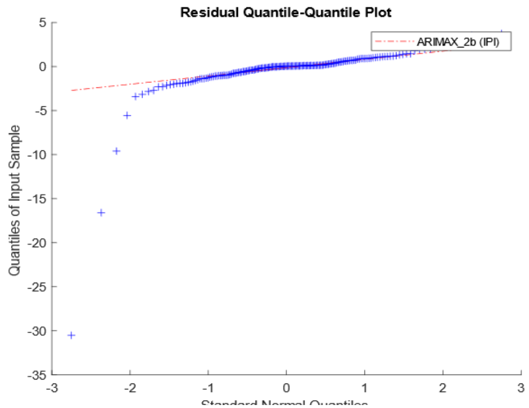
(g) ARMAX(1,0,1)-B Histogram



(i) ARMAX(1,0,1)-B* Histogram



(h) ARMAX(1,0,1)-B Histogram



(j) ARMAX(1,0,1)-B* Histogram

*Accounting for COVID-19

References

Aastveit, K. A., Fastbø, T. M., Granziera, E., Paulsen, K. S., & Torstensen, K. N. (2020). Nowcasting Norwegian household consumption with debit card transaction data. In *Working paper*. Norges Bank.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Alpino, M., Citino, L., & Frigo, A. (2023). The effects of the 2021 energy crisis on medium-sized and large industrial firms: evidence from Italy. In *Occasional papers*. Bank of Italy.

Altissimo, F., Cristadoro, R., Forni, M., Lippi, M., & Veronese, G. (2010). New eurocoin: Tracking economic growth in real time. *The Review of Economics and Statistics*, *92*(4), 1024–1034.

Angelico, C., Marcucci, J., Miccoli, M., & Quarta, F. (2022). Can we measure inflation expectations using Twitter? *Journal of Econometrics*, *228*(2), 259–277.

Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L., & Rünstler, G. (2011). Short-term forecasts of euro area GDP growth. *Econometrics Journal*, *14*, 25–44.

Aprigliano, V., Emiliozzi, S., Guaitoli, G., Luciani, A., Marcucci, J., & Monteforte, L. (2021). The power of text-based indicators in forecasting the Italian economic activity. In *Working papers* (Vol. 1321). Bank of Italy.

Aprigliano, V. (2020). A large Bayesian VAR with a block-specific shrinkage: a forecasting application for Italian industrial production. *Journal of Forecasting*, *39*(8), 1291–1304.

Aprigliano, V., Ardizzi, G., Cassetta, A., Cavallero, A., Emiliozzi, S., Gambini, A., Zizza, R., & Renzi, N. (2021). Exploiting payments to track Italian economic activity: The experience at Banca d'Italia. In *Occasional papers* (Vol. 609). Bank of Italy.

Aprigliano, V., Ardizzi, G., & Monteforte, L. (2017). Using the payment system data to forecast the Italian GDP. In *Working papers* (Vol. 1098). Bank of Italy.

Aprigliano, V., & Liberati, D. (2019). Using credit variables to date business cycle and to estimate the probabilities of recession in real time. In *Working papers* (Vol. 1229). Bank of Italy.

Ardizzi, G., Emiliozzi, S., Marcucci, J., & Monteforte, L. (2019). News and consumer card payments. In *Working papers*. Bank of Italy.

Aruoba, S. B., Diebold, F. X., & Scotti, C. (2009). Real-Time Measurement of Business Conditions. *Journal of Business & Economic Statistics*, *27*(4), 417–427.

Banbura, M., Giannone, D., & Reichlin, L. (2011). Nowcasting. In *The Oxford Handbook of Economic Forecasting*. Oxford University Press.

Baumeister, C., & Guérin, P. (2021). A comparison of monthly global indicators for forecasting growth. *International Journal of Forecasting*, *37*(3), 1276–1295.

Billio, M., Casarin, R., Ravazzolo, F., & Van Dijk, H. (2012). Combination schemes for turning point predictions. *The Quarterly Review of Economics and Finance*, *52*(4), 402–412.

Bodo, G., Cividini, A., & Signorini, L. F. (1991). Forecasting the Italian industrial production index in real time. *Journal of Forecasting*, *10*(3), 285–299.

Bodo, G., Golinelli, R., & Parigi, G. (2000). Forecasting industrial production in the euro area. *Empirical Economics*, *25*, 541–561.

Bodo, G., & Signorini, L. (1987). Short-term forecasting of the industrial production index. *International Journal of Forecasting*, *3*, 245–259.

Box, G., & Pierce, D. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, *65*(332), 1509–1526.

Box, G., & Jenkins, G. (1970). *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting*. Springer International Publishing.

Brunhes-Lesage, V., & Darné, O. (2012). Nowcasting the french index of industrial production: A comparison from bridge and factor models. *Economic Modelling*, *29*(6), 2174–2182.

Bruno, G., & Lupi, C. (2001). Forecasting industrial production and the early detection of turning points. *ISAE Working Paper*.

Bruno, G., & Lupi, C. (2003). Forecasting euro-area industrial production using (mostly) business surveys data. *ISAE Working Paper*.

Bulligan, G., Golinelli, R., & Parigi, G. (2009). Forecasting monthly industrial production in real-time: From single equations to factor-based models. *Empirical Economics*, *39*(2), 303–336.

Bulligan, G., Golinelli, R., & Parigi, G. (2010). Forecasting industrial production: The role of information and methods. In *The IFC's contribution to the 57th ISI Session, Durban, August 2009* (pp. 227–235, Vol. 33). Bank for International Settlements.

Chernis, T., & Sekkel, R. (2017). A dynamic factor model for nowcasting Canadian GDP growth. *Empirical Economics*, *53*, 217–234.

Chiang, T., & Chen, X. (2017). Stock market activities and industrial production growth: Evidence from 20 international markets. In *Advances in pacific basin business economics and finance* (pp. 39–75, Vol. 5). Emerald Publishing Limited, Leeds.

Chiu, C., Mumtaz, H., & Pintér, G. (2017). Forecasting with var models: Fat tails and stochastic volatility. *International Journal of Forecasting*, *33*(4), 1124–1143.

Choi, J., Hauser, S., & Kopecky, K. (1999). Does the stock market predict real activity? time series evidence from the g-7 countries. *Journal of Banking & Finance*, *23*(12), 1771–1792.

Chow, G. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, *28*(3), 591–605.

Clements, M., & Galvão, A. (2008). Macroeconomic forecasting with mixed-frequency data. *Journal of Business & Economic Statistics*, *26*(4), 546–554.

Commission, E., & Eurostat. (2003). *NACE Rev. 2 - Statistical classification of economic activities in the European Community*. European Commission.

Costantini, M. (2013). Forecasting the industrial production using alternative factor models and business survey data. *Journal of Applied Statistics*, *40*(10), 2275–2289.

D'Amuri, F., & Marcucci, J. (2017). The predictive power of Google searches in forecasting US unemployment. *International Journal of Forecasting*, *33*(4), 801–816.

Delle Monache, D., Emiliozzi, S., & Nobili, A. (2019). Tracking Economic Growth during the Covid-19: A Weekly Indicator for Italy. In *Note Covid-19*. Bank of Italy.

Dickey, D., & Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366), 427–431.

Engle, R. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, *50*(4), 987–1007.

Eurostat. (2023). *Industrial production (volume) index overview*. Retrieved July 17, 2024, from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Industrial_production_(volume)_index_overview

Ewing, B., & Thompson, M. (2007). Dynamic cyclical comovements of oil prices with industrial production, consumer prices, unemployment, and stock prices. *Energy Policy*, *35*(11), 5535–5540.

Fornaro, P. (2020). Nowcasting industrial production using unconventional data sources. *ETLA Working Papers*, (80).

Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2005). The generalized dynamic factor model. *Journal of the American Statistical Association*, *100*(471), 830–840.

Franses, P., & Van Dijk, D. (2005). The forecasting performance of various models for seasonality and nonlinearity for quarterly industrial production. *International Journal of Forecasting*, *21*(1), 87–102.

Fruzzetti, M., & Ropele, T. (2024). Nowcasting Italian industrial production: the predictive role of lubricant oils. In *Occasional papers*. Bank of Italy.

Furukawa, K., Hisano, R., Minoura, Y., & Yagi, T. (2022). A Nowcasting Model of Industrial Production using Alternative Data and Machine Learning Approaches. In *Working paper series*. Bank of Japan.

Galbraith, J. W., & Tkacz, G. (2018). Nowcasting with payments system data. *International Journal of Forecasting*, *34*(2), 366–376.

Galdi, G., Casarin, R., Ferrari, D., Fezzi, C., & Ravazzolo, F. (2023). Nowcasting industrial production using linear and non-linear models of electricity demand. *Energy Economics*, *126*, 107006.

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

Girardi, A., Guardabascio, B., & Ventura, M. (2016). Factor-augmented bridge models (fabm) and soft indicators to forecast Italian industrial production. *Journal of Forecasting*, *35*(6), 542–552.

Golinelli, R., & Parigi, G. (2007). The use of monthly indicators to forecast quarterly GDP in the short run: an application to the G7 countries. *Journal of Forecasting*, *26*(2), 77–94.

Götz, T. B., & Knetsch, T. A. (2019). Google data in bridge equation models for German GDP. *International Journal of Forecasting*, *35*(1), 45–66.

Greene, W. (2012). *Econometric analysis* (7th ed.). Pearson Education.

Günay, M. (2018). Forecasting industrial production and inflation in turkey with factor models. *Central Bank Review*, *18*(4), 149–161.

Hahn, E., & Skudelny, F. (2008). Early estimates of euro area real GDP growth: A bottom up approach from the production side. In *ECB working paper series no. 975*. European Central Bank.

Hamilton, J. (1994). *Time series analysis*. Princeton University Press.

Hassani, H., Heravi, S., & Zhigljavsky, A. (2009). Forecasting european industrial production with singular spectrum analysis. *International Journal of Forecasting*, *25*(1), 103–118.

Heravi, S., Osborn, D., & Birchenhall, C. (2004). Linear versus neural network forecasts for european industrial production series. *International Journal of Forecasting*, *20*(3), 435–446.

Lehmann, R. (2023). The Forecasting Power of the ifo Business Survey. *Journal of Business Cycle Research*, (19), 43–94.

Ljung, G., & Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297–303.

Loberto, M., Luciani, A., & Pangallo, M. (2018, April). *The potential of big housing data: an application to the Italian real-estate market* (tech. rep. No. 1171). Bank of Italy, Economic Research and International Relations Area.

Locarno, A., & Zizza, R. (2020). Previsioni ai tempi del coronavirus. In *Note Covid-19*. Bank of Italy.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer-Verlag Berlin.

Marchetti, D., & Parigi, G. (2000). Energy consumption, survey data and the prediction of industrial production in Italy: A comparison and combination of different models. *Journal of Forecasting*, *19*, 419–440.

Martínez-García, E., Grossman, V., & Mack, A. (2015). A contribution to the chronology of turning points in global economic activity (1980-2012). *Journal of Macroeconomics*, *46*, 170–185.

Matsumura, K., Oh, Y., Sugo, T., & Takahashi, K. (2024). Nowcasting Economic Activity with Mobility Data. *Journal of the Japanese and International Economies*, (73).

Menezes, Z., McLaren, C., Von Sanden, N., Zhang, X., & Black, M. (2006). Timely detection of turning points: Should i use the seasonally adjusted or trend estimates? In *Luxembourg: Office for official publications of the european communities*. European Commission.

Mir, A., & Osborn, D. (2004). Seasonal adjustment and the detection of business cycle phases. In *ECB working paper series no. 357*. European Central Bank.

Mosser, P. (2020). Central bank responses to COVID-19. *Business Economics*, *55*, 191–201.

Neftçi, S. (1984). Are economic time series asymmetric over the business cycle? *Journal of Political Economy*, *92*(2), 307–328.

Newey, W., & West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703–708.

O'Brien, D., & Ladiray, D. (2003). Nowcasting eurozone industrial production. In *Luxembourg: Office for official publications of the european communities*. European Commission.

Parigi, G., & Schlitzer, G. (1995). Quarterly forecasts of the Italian business cycle by means of monthly economic indicators. *Journal of Forecasting*, *14*, 117–141.

Phillips, P., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335–346.

Ramos-Francia, M., & García-Verdú, S. (2022). Central Bank Response to COVID-19. *Latin American Journal of Central Banking*, *3*(3).

Rodano, L., Siviero, S., & Visco, I. (2013). Inversioni cicliche e previsioni macroeconomiche: Racconto di due recessioni. In *Economic multisectoral modelling between past and future: a tribute to Maurizio Grassini and a selection of his writings* (pp. 229–247). Firenze university press.

Said, S., & Dickey, D. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, *71*(3), 599–607.

Schlitzer, G. (1996). Business cycles in Italy: A statistical investigation. *European Journal of Political Economy*, *11*(4), 683–698.

Schreiber, S., & Soldatenkova, N. (2016). Anticipating business-cycle turning points in real time using density forecasts from a var. *Journal of Macroeconomics*, *47*, 166–187.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

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

Suimon, Y., & Yanai, M. (2021). Using mobile phone gps data to nowcast auto production activity and equity investment strategy based on the activity nowcasting. *2021 10th International Congress on Advanced Applied Informatics (IIAI-AAI)*, 680–683.

Thorsrud, L. A. (2020). Words are the New Numbers: A Newsy Coincident Index of the Business Cycle. *Journal of Business & Economic Statistics*, *38*(2), 393–409.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, *48*(4), 817–838.

Yule, G. (1927). On a method of investigating periodicities in disturbed series, with special reference to Wolfer,Äôs sunspot numbers. *Philosophical Transactions of the Royal Society*, (226), 636–646.