OESTERREICHISCHE NATIONALBANK
EUROSYSTEM

WORKING PAPER 262

Improving disaggregated short-term food inflation forecasts with webscraped data

Christian Beer, Robert Ferstl, Bernhard Graf

The *Working Paper series of the Oesterreichische Nationalbank* is designed to disseminate and to provide a platform for discussion of either work of the staff of the OeNB economists or outside contributors on topics which are of special interest to the OeNB. To ensure the high quality of their content, the contributions are subjected to an international refereeing process. The opinions are strictly those of the authors and do in no way commit the OeNB.

The Working Papers are also available on our website (http://www.oenb.at) and they are indexed in RePEc (http://repec.org/).

Publisher and editor	*Oesterreichische Nationalbank* *Otto-Wagner-Platz 3, 1090 Vienna, Austria* *PO Box 61, 1011 Vienna, Austria* *www.oenb.at* *oenb.info@oenb.at* *Phone (+43-1) 40420-6666* *Fax (+43-1) 40420-046698*
Editor	*Martin Summer*
Cover Design	*Information Management and Services Division*
Data protection information	*www.oenb.at/en/dataprotection*

ISSN 2310-5321 (Print)
ISSN 2310-533X (Online)

Improving disaggregated short-term food inflation forecasts with webscraped data*

Christian Beer† Robert Ferstl‡ Bernhard Graf§

First version: November 11, 2024
This version: January 10, 2025

Most recent version available at:
`https://ssrn.com/abstract=5016579`

Abstract

This study examines the effectiveness of using webscraped data to predict price developments in the Austrian food retail sector. We calculate monthly nowcasts of price changes based on daily price data collected by the OeNB since mid-2020, using Eurostat methodology for price index calculation, along with further details provided by the national statistics office. We assess the quality of our nowcasts by comparing them with various baseline models and more advanced time series methods also covering machine learning approaches. Our findings indicate that webscraped data are a useful way to obtain more accurate nowcasts with a time advantage, amounting to several weeks, over traditional data sources. In addition, we are the first, to our knowledge, to explore the possibility of using the improved accuracy of the nowcasts as a basis for disaggregated short-term forecasts that extend up to one quarter. While direct forecasts at higher levels of aggregation produce slightly more accurate overall metrics, indirect forecasts derived from disaggregated data provide superior insights into the underlying dynamics of specific sub-components. Our results show that more advanced time series models have trade-offs in terms of computational efficiency while performing very similarly to more traditional methods. These findings have implications for policymakers who aim to develop an effective system for real-time monitoring of inflation dynamics at a very granular level.

Keywords: Webscraping, Inflation forecasting, Time series models
JEL Codes: C22, C81, E31, E37

---

*The authors would like to thank seminar and meeting participants of an OeNB internal research seminar in July 2024 and the ECB DPD Power Users Meeting in December 2024 for their valuable comments.
Disclaimer: The views in this paper are those of the authors and do not necessarily reflect the official viewpoint of the Oesterreichische Nationalbank or the Eurosystem.
†Oesterreichische Nationalbank, Business Cycle Analysis Section, Vienna, Austria. Email: christian.beer@oenb.at
‡Oesterreichische Nationalbank, Monetary Policy Section, Vienna, Austria, and University of Regensburg, Department of Finance, Regensburg, Germany. E-mail: robert.ferstl@oenb.at and robert.ferstl@ur.de
§Oesterreichische Nationalbank, Monetary Policy Section, Vienna, Austria. E-mail: bernhard.graf@oenb.at

Non-technical summary

This study presents how to effectively use webscraped data to predict price developments in the Austrian food retail sector. We use daily online price data collected by OeNB since mid-2020 and apply the official methodology of the national statistics office to compile price indices. We compare the quality of month-on-month inflation rates based on these price indices to forecasts of time series models commonly used in the forecasting literature. Moreover, we consider various machine learning approaches for our comparison of inflation forecasts. We also extend the existing literature by being the first to include webscraped data into short-term forecasts of time series models of disaggregated food inflation rates in Austria. Our findings indicate that webscraped data are a useful way to obtain more accurate nowcasts with a time advantage, amounting to several weeks, over traditional data sources. We find that webscraping not only reduces forecast uncertainty on a disaggregated level, but also improves predictions when aggregated to higher COICOP levels. Herein, the careful classification and application of the official methodology used to compile Consumer Price Indices is of utmost importance. Additionally, we observe that employing models beyond traditional ARIMA, such as advanced time series and machine learning models, can offer benefits, although they may involve trade-offs in computational efficiency. For practical inflation forecasting, this means that using state-of-the-art models and software from the machine learning community can help build reliable and automated real-time forecasting tools. These advances have important implications for policymakers and practitioners seeking to develop effective systems for monitoring and forecasting of inflation dynamics at a granular level in real time.

1 Introduction

Prices of food and beverages are a critical yet highly volatile component of the Harmonized Index of Consumer Prices (HICP) inflation, making them notoriously challenging to forecast, especially considering their month-on-month inflation dynamics. Understanding, properly interpreting, and projecting the dynamics of these volatile food components in HICP inflation rates is therefore crucial for central banks and other institutions alike.

We aim to contribute to the existing literature (see section 2) on HICP inflation forecasting using webscraped data by developing a nowcasting and forecasting framework specifically applied to the ECOICOP (European Classification of Individual Consumption according to Purpose) division ”01 Food and non-alcoholic beverages” for Austria. To achieve this, we have created and maintain a database of webscraped price data from Austrian online retailers. With this data we construct price indices that adhere to the official methodologies established by Eurostat and Statistik Austria.

We then nowcast inflation rates and produce forecasts extending up to three months into the future. Our analysis is conducted at the most granular product level used in computing HICP inflation rates and for the aggregate ”Food and non-alcoholic beverages” as well as for the intermediate COICOP aggregates inbetween.

Forecast errors are calculated by comparing the forecasts with ex-post realised inflation rates for each COICOP category and benchmarking them against several univariate time series and two ensemble models. One of our contributions to the literature is the inclusion of more advanced time series and machine learning models as benchmarks for webscraping based nowcasts. In addition, we provide a more detailed assessment of forecasting accuracy at the level of elementary indices below the harmonised ECOICOP aggregation. We also compare direct time series forecasts at higher levels of aggregation to indirect forecasts derived from the lowest level categories and have the benefit that forecasts of indices at higher levels can be decomposed into specific sub-components which drive the results.

Furthermore, we are the first paper, to our knowledge, that proposes to incorporate webscraped data to enhance time series based short-term forecasts of disaggregated inflation rates. For this purpose, we leverage the webscraped nowcast as the basis for the first month of our three-month forecast.

In summary, our analysis shows that incorporating webscraped data significantly enhances nowcasting accuracy for short-term food price inflation and improves the predictive performance of short- and medium-term inflation forecasts in an increasingly volatile economic environment. The use of webscraped data boosts predictability at a highly disaggregated level and yields consistent results across various COICOP categories, particularly those with higher weights in the HICP aggregation. Forecast accuracy is also enhanced by webscraped data when aggregating forecasts to higher COICOP levels. Notably, this improvement can be achieved in a very brief time frame, whereas traditional forecasting methods typically require several years of data input. Furthermore, our comparison reveals that while advanced time series models can match the performance of traditional approaches, they often involve trade-offs in computational efficiency.

The rest of the paper is structured as follows: Section 2 reviews the related literature. Section 3 discusses the official inflation data and the calculation of online price indices. In Section 4, we outline the setup of a nowcasting and short-term forecasting competition. Section 5 presents the results of the analysis, followed by the conclusion in Section 6.

2 Related Literature

The macro-economic analysis of inflation and the price-setting process has recently profited from readily-available high-frequency price data. The structured collection of these data began with the Billion Prices Project (Cavallo and Rigobon, 2016).[1] These data permit the analysis of moments of distributions of price changes and other statistics on the micro level (Gautier et al., 2023). Furthermore, they improve our understanding of inflation by allowing to observe the frequency and size of price changes and help us understand its micro-foundations in macro-economic models such as e.g. in case of the menu costs model of price setting (Carrera de Souza, 2022; Gautier et al., 2023). Given sufficient quantity and quality of micro price data, it has been shown that they can provide timely high-frequency information on aggregate inflation dynamics and may improve nowcasting models of HICP inflation.

In addition, price developments play a crucial role in shaping inflation expectations, which are essential for effective monetary policy. Recent research (Anesti et al., 2024; D'Acunto et al., 2021) highlights the significant impact of food prices on the formation of these expectations, emphasizing the importance of both current and forecasted inflation in this category.

Besides advances in data collection, selecting appropriate models for forecasting inflation is very important in our context. It has been a consensus in this literature that it is quite difficult for inflation forecasts to improve on simple univariate models or expert judgement (Faust and Wright, 2013). While the wider adoption of machine learning models has only marginally challenged this consensus (Medeiros et al., 2019), these models largely depend on the inclusion of a large number of exogenous variables, with a notable exception in Barkan et al. (2023) who focus on explaining the internal time series patterns of inflation rates. Another key point raised by Faust and Wright (2013) is the potential of nowcasts to improve forecasts.

Cavallo and Rigobon (2016) show that indices constructed using online micro prices can effectively track the movements of consumer price indices in the United States and Australia over the period 2008 to 2016 and 2015, respectively. These constructed indices strongly anticipate movements in public CPI data over both long and short time horizons, as shown by an illustrative example of online price movements around the Lehman Brothers bankruptcy on 15 September 2008. While these indices do not necessarily capture the price level well, they effectively track the dynamics of CPI inflation in several countries. Using estimates of impulse response functions, they show that webscraped data are informative for CPI dynamics and yield results that are not statistically significantly different from ex-post realised CPI inflation in a one- to eight-month horizon.

In a similar vein Aparicio and Bertolotto (2020) forecast one-month, two-month, and three-month-ahead headline CPI inflation rates for several countries based on data from Price Stats. They compare various models with an equal-weighted average of forecasts as the baseline specification. The authors'

[1] The Billion Prices Projects (or its commercial spin off PriceStats) do not provide any data for Austria.

findings suggest that online price data closely track CPI data and the inclusion into the proposed models significantly improves forecasts.

Macias et al. (2023) is the research closest to our own. They use data collected by Narodowy Bank Polski (NBP), which has already adopted webscraping price data on a weekly basis since 2009 and expanded to daily scraping in 2018. They nowcast inflation rates for food and non-alcoholic beverages as well as the elementary product groups in this category and the sub-aggregates, comparing the nowcasts from various univariate models as well as the judgemental forecast of the NBP. They show that the incorporation of webscraped data significantly improves nowcasts. In particular, for the aggregate food and non-alcoholic beverages, even the direct use of price changes derived from online data reduces forecast errors by about 29% compared to their benchmark SARMA model. In addition, the inclusion of online data information as an exogenous variable in a time series model can lead to an even greater improvement in forecast accuracy. Macias et al. (2023) emphasize that the benefits of incorporating webscraped data into forecasting models can become evident within just a few months, particularly for aggregate inflation rates. Additionally, Macias et al. (2023) highlight the critical importance of proper classification when utilizing webscraped data for nowcasting purposes. Ensuring that the data used in the models are correctly classified into the product groups used in the calculation of HICP rates is crucial because it ensures that the data used in the nowcasting models accurately reflect the prices of specific food categories. This attention to classification leads to more reliable nowcasts and ultimately strengthens the validity of the forecasts.

Also closely related to our paper is Soybilgen et al. (2023) who collected daily food price data from Turkish retailers and produce 132 indices on the most granular product category for which Turkstat publishes price indices. These subindices are then aggregated into a food price index. They conclude that the online data nowcasts both food and headline inflation rates with a considerable time advantage over official announcements.[2]

An important contribution to the nowcasting of food inflation is also provided by Beck et al. (2023). While the authors do not use webscraped price data, but rather weekly household scanner data, their work is relevant to our paper due to their focus on highly disaggregated levels for nowcasting. For this purpose, they utilize a mixed-frequency data sampling (MIDAS) model[3] combined with machine learning techniques. Their findings demonstrate hat the use of household scanner data can enhance inflation nowcasts within the first week of each month.

In this study, we follow the literature that utilizes webscraped data to forecast official inflation rates (Cavallo and Rigobon, 2016; Aparicio and Bertolotto, 2020) and concentrate on the temporal dynamics of month-on-month changes within individual time series (Barkan et al., 2023). Specifically, we extend the nowcasting of food inflation using webscraped data (Macias et al., 2023; Soybilgen et al., 2023) to short-term forecasts and enhance the benchmark models by adopting more flexible approaches, including machine learning techniques. Furthermore, to our knowledge, this paper is the first to provide a granular comparison of direct versus indirect forecasts at different levels of aggregation of the Harmonized Index

[2]For the data and indices see `https://cefis-food-indices.streamlit.app/`

[3]A MIDAS model also employed by Vicente and Valls Pereira (2022) who use Brazilian data from the Billion Prices Project to nowcast inflation.

of Consumer Prices (HICP).

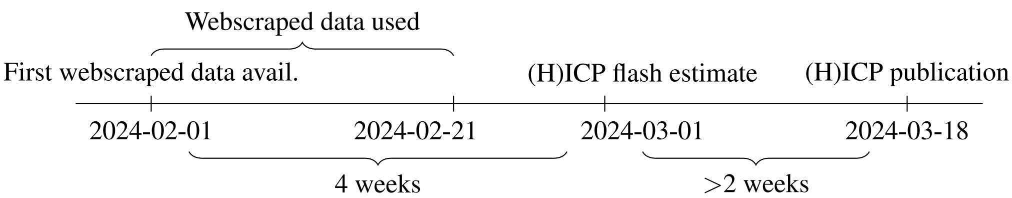
3 Official inflation data and online price indices

We use official data from Statistik Austria alongside internal web-scraped data from OeNB on the ECOICOP category *01 Food and non-alcoholic beverages*.

3.1 Indices calculated by Statistik Austria

For the official inflation data we use index series from Statistik Austria. The agency collects price information through both in-store and scanner data collection methods. Preliminary results for the previous month are published from the 17th to the 22nd of each month.[4] Hence, webscraped data provide a head start of 3 to 6 weeks (Figure 1).[5] The flash estimate at the beginning of the month is only available for headline inflation and the first level of categories. In our case, more detailed indices below the ECOICOP category *01 Food and non-alcoholic beverages* can only be proxied by nowcasts built on webscraping data or time series models.

Figure 1: Exemplary timeline of (H)ICP data availability and publication



Note: This figure shows the release schedule of the HICP in Austria. Webscraped data are collected daily by OeNB and are made available in real-time throughout the month, starting from the beginning of the month. Flash estimates by Statistik Austria at higher levels of aggregation are made available at the end of the month. Detailed indices are published in the third week of the following month.

3.2 Calculation of elementary indices based on webscraping data

The development of webscraping algorithms at Oesterreichische Nationalbank (OeNB) started in the spring of 2020 within the scope of the *ESCB PRISMA* research network.[6] While webscraping can be rapidly implemented, maintaining the infrastructure, managing product classification into the ECOICOP categories, and processing the data require significant resources in a production environment. Given the

[4] `https://www.statistik.at/fileadmin/pages/214/publikationstermine_vpi_und_hvpi_2024.pdf`

[5] It is becoming increasingly common to supplement the in-store obtained data used in HICP compilation by scanner data (as has been implemented by Statistik Austria from January 2021 for the sub-aggregates of "Food and Beverages") as well as webscraped data. Hence, not only does a webscraping project provide timely micro-data, but such data are themselves becoming increasingly relevant in HICP compilation. Going forward, the broader adoption of these methods may indeed necessitate the curation of a webscraping price database on an ongoing basis to accurately form inflation forecasts.

[6] The ECB continues to maintain their Daily Price Dataset (DPD) of webscraped data, but focuses only on the larger euro area countries. The collection of Austrian data is therefore only supported by the OeNB.

high benefits for economic analysis and research at OeNB, we continue to run a webscraping framework for Austrian online retailers. For this purpose, we collect daily data from around 15 Austrian retailers, of which 7 contribute to food and non-alcoholic beverages, with overall 134,000 unique products and 18,000 unique classified products as of March 2024. The collected data include prices and further product information (e.g. product name, shop internal categorisation). Table 1 below presents some descriptive statistics of the webscraped data.

Table 1: Summary statistics of product data by store

Store	Overall Total Products	Classified Products	Classified Total Prices	Mean Price	Mean Days Available
store 1	19,003	4,622	3,892,286	4.4	931
store 2	42,762	762	543,271	4.5	775
store 3	27,717	4,404	2,630,963	6.2	641
store 4	2,567	671	141,729	3.3	254
store 5	36,082	6,338	5,810,560	4.6	1,018
store 6	1,709	362	164,525	7.4	504
store 7	4,707	789	158,039	2.8	219
All Stores	134,547	17,948	13,341,373	4.8	819

Note: This table presents summary statistics for product data across different stores. The first column shows the total number of products per store. The remaining columns provide data on the number of classified products, the total number of observed prices, the mean price, and the average number of days products were available.

We compute monthly time series of elementary indices following Statistik Austria (2022) by the multilateral Gini, Eltetö and Köves, and Szulc (GEKS) index with a 13-month window.[7] We choose a Jevons index as the underlying bilateral index. To avoid revisions of this GEKS-Jevons index we use a rolling window approach (with a movement splice) as recommended by Eurostat (2024). The resulting index is therefore called a rolling window GEKS-Jevons (RWGEKS) with a movement splice.[8]

A detailed description of the index compilation following the official methodology is provided in Appendix A.

3.3 Classification of products and aggregation to higher COICOP levels

To achieve a meaningful forecast, an accurate and granular mapping of our webscraped products to the consumption basket defined by Statistik Austria (see Appendix A.5) is crucial. For the compilation of prices and HICP calculation, national statistical offices in the European Union divide the basket of goods according to the ECOICOP (European Classification of Individual Consumption according to Purpose) system (see Eurostat, 2024). This system is structured hierarchically into divisions (2-digit level, e.g.,

[7]To address data issues and outliers, we perform basic integrity checks, such as ensuring that prices are greater than zero. Furthermore, we exclude products if their price z-scores exceed 5. We refrained from using absolute thresholds for identifying erroneous price changes because legitimate significant price increases can occur, for instance, when a discount period occurs.

[8]The authors of Statistik Austria (2022) decided to use the GEKS method with the Törnqvist index with food scanner data. The different variants are generally very close to each other. A Jevons index is also more appropriate in case of webscraped data, because we do not collect any data on quantities which would be required by the Törnqvist index.

01: Food and non-alcoholic beverages), groups (3-digit level, e.g., 01.1: Food), classes (4-digit level, e.g., 01.1.1: Bread and cereals), and sub-classes (5-digit level, e.g., 01.1.1.3: Bread). This common classification is referred to as Level I classification.

More granular division of the consumption basket are referred to as Level II. In Austria, the most detailed level at which Statistik Austria publishes indices is known as *Indexpositionen*, For example, within the ECOICOP sub-class 01.1.1.3 (Bread), the *Indexpositionen* include specific types of bread such as rye bread, wholemeal bread, and white bread. The elementary indices published at this level are referred to in German as *Bundesmesszahlen*.[9] In response to changing national consumption pattern items at this level may be added or discontinued. However, from 2020 to 2024, they only changes within the "Food and non-alcoholic beverages" division in Austria was the introduction of two additional items in 2024. Since these items were only available for a short time span, we could not take them into account.

We employ a bottom-up approach to categorize products into their respective ECOICOP categories, meaning that products are directly classified at their *Indexposition* level. However, this approach also implies that many products remain uncategorised, as illustrated in Table 1. For simplicity, we refer to this *Indexposition* level as COICOP 6 or level 6 and to the index as *elementary index* or as index (level 6) throughout the text. In a similar vein, we use COICOP 2 (level 2) to COICOP 5 (level 5) for the aggregated categories.

The aggregation of elementary indices to the higher COICOP level is important and requires the appropriate weights and the application of the correct methodology. A detailed description can be found in Appendix A.3.

4 Setting up a nowcasting and short-term forecasting competition

4.1 Scope of the framework

Our primary objective is to evaluate whether the monthly rate of change in webscraped price indices can effectively nowcast the official monthly rates of change for the ECOICOP division *01 Food and Non-Alcoholic Beverages* and the subcategories at all levels below. We therefore select several simple and advanced time series models and compare their forecasting accuracy.

In addition to the one-step ahead nowcast, we provide three-month ahead forecasts that are based on a combination of the webscraped indices and time series models. For the one-step ahead forecast of the next month we will directly use the webscraped nowcast and then predict the official index series for the subsequent two months.

First, we perform the nowcasts and forecasts at the most disaggregated level, the so-called elementary indices. As the disaggregated indices in Austria are one level below the five ECOICOP levels, we also refer to this as level 6. We repeat this exercise for higher levels (up to COICOP 2). To obtain time series

[9] Bundesmesszahlen are subject to revisions, typically available the following month. We do not consider the revised data for several reasons: one-month-ahead nowcasts need to be produced in a timely manner and can only rely on initially reported data. In the subsample of 130 Indexpositionen, there were revisions for only 737 out of 13,970 data points, of which only 238 were larger than 1%. Revisions from 2017 were excluded due to unreliable information.

forecasts for the higher levels, we use two methods of obtaining time series forecasts:

(1) **Indirect**: Time series models are estimated for the elementary index series (level 6) and aggregated using the weights provided by the statistical office, see Appendix A.3.

(2) **Direct**: Time series models are estimated for the official index series at levels 2 to 5 published by Eurostat.

We use data from Statistik Austria for the elementary indices starting in January 2014. We calculate out-of-sample one-step and three-step ahead forecasts from January 2021 until March 2024 and use a a 72 months rolling window.[10] For a limited number of time series, data are only available for less than 72 months. Here we switch to extending window forecasts and start with a minimum window size of 24 months. This means that the first forecasts are available at the same time as the first webscraped data.

4.2 Time series models

We use several time series models in our forecast competition. In addition to traditional benchmarks like random walk or ARIMA we employ more advanced time series models like Prophet, Neuralprophet or ETS and Theta.

We denote by y_t a discrete time series which is observed at $t = 1, 2, \dots T$. In our case, this will be the monthly rate of change of a specific index series for a given level of COICOP-aggregate.

Naive

We begin with a naive forecast, commonly referred to as a random walk model, where the current monthly inflation rate will be equal to the value observed in the previous month:

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

where ε_t is white noise.[11]

ARIMA

We select the best seasonal autoregressive integrated moving average SARIMA$(p,d,q)(P,D,Q,m)$ model based on Akaike's Information Criterion (AIC):

$$y_t = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{P} \Phi_k y_{t-mk} + \sum_{l=1}^{Q} \Theta_l \varepsilon_{t-ml} + \varepsilon_t,$$

[10] We have chosen a rather long time window in order to include enough information on possible seasonalities. The elementary indices at this level of granularity are also not available before 2014. We therefore conclude that with the available sample period and approach, an extending window would have performed similarly. However, if longer time series become available in the future, the use of an extending window may have advantages.

[11] A special version of the random walk model, which averages a fixed number of past observations, is also popular in the literature on inflation forecasting (see Atkeson and Ohanian, 2001). However, we exclude this model from our analysis to retain a single baseline model for benchmarking purposes and to limit the number of competing models.

where α is a constant, p is the number of non-seasonal autoregressive terms, q is the number of non-seasonal lagged forecast errors (moving average terms), P is the number of seasonal autoregressive terms, Q is the number of seasonal lagged forecast errors (seasonal moving average terms), and m is the number of observations per year. Here, y_t denotes a stationary series $\Delta^d \Delta_m^D \tilde{y}_t$, obtained after non-seasonal differencing of order d and seasonal differencing of order D of a potentially integrated series $\tilde{y}_t$. We consider $p,q \leq 5$ and $P,Q \leq 2$. The orders of differencing d and D are determined by testing for stationarity. The parameter $m = 12$ because we have monthly time series.

Lunsford and West (LW)

This model is a special case of an AR(1) model. Initially developed by Lunsford and West (2019) and applied for example by Borio et al. (2022) and Graf (2024), it uses a simple heuristic where the current monthly value is regressed on its first lag and an intercept:

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

where α is the intercept, β is the autoregressive coefficient, and ε_t is the error term. If large outliers are detected (i.e., $|y_t| > 25$), the value of y_t is replaced by its lagged value, y_{t-1}. The autoregressive coefficient estimate β is capped at 0.9 to ensure stationarity and the intercept α is re-estimated if the cap is binding.[12]

Prophet

Prophet (see Taylor and Letham, 2018) is an additive, decomposable time series model:

$$y_t = T(t) + S(t) + H(t) + \varepsilon_t,$$

where $T(t)$ is a trend function, $S(t)$ represents seasonalities, and $H(t)$ is used to include irregular changes such as holidays and ε_t is the error term. It has gained in popularity in recent years because it is very easy to use in an automated forecasting setting. However, there is also some evidence that `AutoARIMA` is faster and more accurate (see Russel, 2022). Therefore, we are interested in evaluating how Prophet performs with our webscraped data. We use Prophet with its default settings: a linear trend with automatic change point detection, seasonalities, no holiday effects.

Neuralprophet

NeuralProphet (see Triebe et al., 2021) is an extension of Prophet that incorporates local context through auto-regression and covariate modules. These modules can be set up as either traditional linear regression models or neural networks, allowing it to handle more complex patterns than Prophet. The model is formulated with the additional components as follows:

[12]To ensure consistency with the AR(1) model, we use the `StatsForecast` implementation of `ARIMA` with a fixed coefficient. It optimizes the conditional sum of squares to find starting values and then uses maximum likelihood estimation.

$$y_t = T(t) + S(t) + H(t) + F(t) + A(t) + L(t) + \varepsilon_t,$$

where $T(t)$ represents trend, $S(t)$ seasonal effects, and $H(t)$ event and holiday effects, ε_t is the error term, similar to Prophet. In addition, $F(t)$ denotes regression effects for future-known exogenous variables, $A(t)$ represents auto-regression effects based on past observations, and $L(t)$ accounts for regression effects for lagged observations of exogenous variables. Similar to Prophet, we do not use any holiday effects. Furthermore, no exogenous variables are present in our model. The learning rate is set to 0.01; weekly and daily seasonalities are turned off, and we include 12 lagged observations.

Theta

The Theta model, as introduced by Assimakopoulos and Nikolopoulos (2000), is defined as the solution to:

$$\Delta^2 Z_t(\theta) = \theta \Delta^2 y_t \quad \text{for } t = 3, \ldots, T,$$

where $y_1, \ldots, y_T$ represent the original time series (non-seasonal or deseasonalized), and Δ^d (with $d = 2$) is the difference operator. The initial values Z_1 and Z_2 are obtained by minimizing the sum of squared deviations $\sum_{t=1}^{T} [y_t - Z_t(\theta)]^2$. An analytical solution to this minimization problem is given by:

$$Z_t(\theta) = \theta y_t + (1 - \theta)(A_T + B_T t) \quad \text{for } t = 1, \ldots, T,$$

where

$$A_T = \frac{1}{T} \sum_{t=1}^{T} y_t - \frac{T+1}{2} B_T,$$

and

$$B_T = \frac{6}{T^2 - 1} \left(\frac{2}{T} \sum_{t=1}^{T} t y_t - \frac{T+1}{T} \sum_{t=1}^{T} y_t \right).$$

Here, A_T and B_T are the coefficients obtained from a linear regression of $y_1, \ldots, y_T$ against $1, \ldots, T$. Fiorucci et al. (2016) propose a dynamic optimized Theta method which is available in the `AutoTheta` method by Garza et al. (2022).[13]

Error-Trend-Seasonality (ETS)

Hyndman et al. (2008) developed a state-space formulation:

[13] From a user perspective, the optimal Theta models are very simple to forecast with `AutoTheta`, because the interface is similar to `AutoARIMA` by Garza et al. (2022).

$$
\begin{aligned}
y_t &= w(\mathbf{x}_{t-1}) + r(\mathbf{x}_{t-1})\varepsilon_t, \\
\mathbf{x}_t &= f(\mathbf{x}_{t-1}) + g(\mathbf{x}_{t-1})\varepsilon_t,
\end{aligned}
$$

encompassing multiple exponential smoothing variations.

The state vector is given by $\mathbf{x}_t = (\ell_t, b_t, s_t, s_{t-1}, \ldots, s_{t-m+1})$ where ℓ_t denotes the level component, b_t represents the trend component, $s_t, s_{t-1}, \ldots, s_{t-m+1}$ are the seasonal components, with m indicating the seasonal period and $\{\varepsilon_t\}$ is a Gaussian white noise process with variance σ^2 and $\mu_t = w(\mathbf{x}_{t-1})$. For models with additive errors $r(\mathbf{x}_{t-1}) = 1$ and for multiplicative errors $r(\mathbf{x}_{t-1}) = \mu_t$. The `AutoETS` function by Garza et al. (2022) selects the best model using the Akaike Information Criterion (AIC).

Complex Exponential Smoothing (CES)

Unlike simple exponential smoothing, which applies a consistent exponential decay of the weights, complex exponential smoothing by Svetunkov et al. (2022) allows for varied weight distributions, including oscillating or harmonic patterns. The CES equation is derived by extending the simple exponential smoothing framework into the complex plane:

$$
\hat{y}_t + i\hat{e}_t = (\alpha_0 + i\alpha_1)(y_{t-1} + ie_{t-1}) + (1 - \alpha_0 + i - i\alpha_1)(\hat{y}_{t-1} + i\hat{e}_{t-1}),
$$

where $\hat{y}_t$ and $\hat{e}_t$ are the estimated value and estimated error term at time t, i represents the imaginary unit (satisfying $i^2 = -1$), α_0 and α_1 are the real and imaginary components of the complex smoothing parameter, y_{t-1} and e_{t-1} are the actual observation and error term at time $t-1$. The `AutoCES` function by Garza et al. (2022) selects the best model using the Akaike Information Criterion (AIC). The types of models included are simple CES, as well as models with simple, partial, or full seasonality.

4.3 Forecasting accuracy

We begin by briefly discussing related work, including studies on online prices, before outlining our approach to measuring prediction accuracy. Faust and Wright (2013) use an AR(1) gap model with a fixed coefficient as a benchmark. They compare the root mean squared prediction errors of all other models with respect to this benchmark. They test whether the prediction errors of the other models are significantly different with a two-sided Diebold and Mariano (1995) test. Aparicio and Bertolotto (2020) compare forecasts with RMSE and perform Diebold-Mariano tests with the null hypothesis that online prices have similar predictive ability compared to each alternative model. Macias et al. (2023) evaluate the accuracy of their nowcasts using mean forecast errors and root mean square forecast errors. They set up three models that include online prices, i.e., a SARIMAX model with online prices as an exogenous variable, the pure nowcasts for the entire month, and the pure nowcasts for the first 15 days of the month. Two-sided Diebold-Mariano tests were run against the SARIMAX specification for all models. Barkan et al. (2023) also use the RMSE for forecast evaluation as well as some correlation metrics. The authors also perform pairwise Diebold-Mariano tests by comparing all models against an AR(1) specification.

The above discussion shows that authors using online prices mainly follow Faust and Wright (2013). However, depending on the data source used and the focus of the study, the choice of benchmark varies. It ranges from simpler models like random walk or AR(1) to more complex models like SARIMAX. For this paper, we are interested in the following questions regarding the accuracy of our nowcasts and short-term forecasts:

- Can disaggregated nowcasts of elementary indices achieve superior or at least similar predictive accuracy compared to various time series models ranging from simple to very flexible specifications?
- Does predictive accuracy remain at least similar to time series models when aggregating webscraping based nowcasts up to higher COICOP levels?
- Which models should be included when forming ensembles?
- Does the predictive accuracy of ensembles improve when including webscraping?
- Given a proper performance of webscraping in one-step ahead nowcasts, can a short-term forecast including the nowcast outperform three-step ahead forecasts of time series models?

We will mainly use the root mean squared error to evaluate the accuracy of our out-of-sample forecasts.

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(y_{t+h} - \hat{y}_{t+h})^2}$$

where y_{t+h} is the actual value h periods ahead and $\hat{y}_{t+h}$ is the predicted value.[14] Furthermore, we consider the mean absolute error $\text{MAE} = \frac{1}{T}\sum_{t=1}^{T}|y_{t+h} - \hat{y}_{t+h}|$ and the mean directional accuracy $\text{MDA} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{I}(\text{sign}(y_{t+h}) = \text{sign}(\hat{y}_{t+h}))$.

5 Results

We provide detailed results for the one-step ahead forecasts in Appendix B.1 and for our three-step ahead short-term forecasts in Appendix B.2. All results are available for each individual index series and at various levels of aggregation. They include the RMSEs of the elementary indices (level 6) and the direct forecasts at the higher COICOP levels 2 to 5, as well as the indirect forecasts of the indices based on the forecasts of the elementary indices. We also created two variants of equally weighted ensembles. The majority of naive forecasts showed a significantly lower predictive accuracy. We therefore decided to exclude them from the ensembles. In the following results, we present an ensemble of all time series models and another ensemble that additionally includes webscraping.[15]

[14]Some authors employ the term root mean squared forecast error (RMSFE) and utilize RMSE to denote the in-sample fit on the training data. As our paper assesses the precision of the out-of-sample forecasts, all the displayed error metrics are derived from forecasts.

[15]We recognize that the composition of the ensembles could be further optimized by tuning the weights, and we believe this is worth exploring in practical applications. In this study, our primary focus is to demonstrate the added value of webscraping and to

5.1 One-step ahead forecasts

We begin the discussion of the results with an evaluation of the performance of webscraping indices as a way to nowcast one-month ahead inflation rates compared to one-step ahead forecasts of various univariate time series models.

5.1.1 Elementary indices

A simple way to aggregate the predictive accuracy of the 130 elementary indices is to rank them by RMSE and count the number of categories in which each model ranks first. Figure 2 shows that for 58 categories the webscraping based nowcasts are the one-step ahead forecasts with the minimum RMSE. The naive forecast never ranks first and is therefore not included in this figure. We also calculated equally-weighted ensemble forecasts with and without the webscraping nowcasts. We did not include the naive forecasts in these ensembles. The ensemble with webscraping achieves 42 first ranks. The univariate time series models show no clear advantage of the more advanced specifications like Prophet. For each index series we provide detailed results in Table A2 in Appendix B.1.1. There we report in the first column the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$). As indicated by the high number of first ranks for the webscraping many of the RMSE ratios are above one indicating higher predictive accuracy. For several indices webscraping also achieves significantly different predictive accuracy according to a Diebold and Mariano (1995) test.

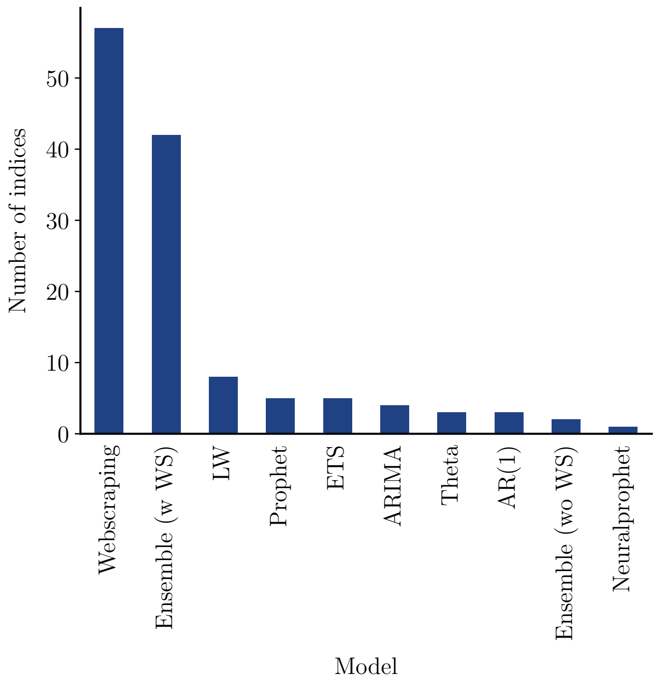
Another way to aggregate the predictive accuracy for the elementary indices (level 6) is to weight the individual RMSE of each index category by the corresponding HICP weight in 2024. The results are shown in Table 2. The rows are sorted by the column RMSE. The ordering of the models is similar to the analysis before. However, on average, the ensemble without webscraping now ranks on second place just before webscraping, which manages to further improve the ensemble. The naive model achieves the worst results measured by any of the provided metrics. The mean directional accuracy of 47% is quite low. So any kind of time series model does better. In particular, simpler models like AR(1) or Lunsford-West perform better on average compared to more flexible specifications like ARIMA or Prophet.

Figure 3a compares the AR(1) and LW results, where LW can be regarded as a more more robust form of an AR(1). In terms of RMSE, the Lunsford and West (2019), Borio et al. (2022) specification by Graf (2024) has a slight advantage. Based on the ranks AR(1) performs a bit better. By definition both models should be very similar. However, the scatterplot (Figure 3a) reveals that there are some time series where the outlier correction and the adjustment of the coefficient in the LW model become binding.

Figure 3b compares the results of Prophet (see Taylor and Letham, 2018) and Neuralprophet (see Triebe et al., 2021). The model specification is very similar, but NeuralProphet has more flexibility due to the inherent neural network. Although there are some noticeable differences in the predictions, these

develop a parsimonious method for combining forecasts from multiple models.

Figure 2: Number of first rank per model for elementary indices (level 6)



Note: This figure shows the number of first ranks per model for one-step ahead forecasts of the elementary indices at level 6.

do not necessarily lead to better forecast accuracy. In terms of first ranks, Neuralprophet achieves the lowest number. Measured by RMSE, Neuralprophet is slightly ahead of Prophet. Given the much higher computation time, we conclude that even more advanced time series models based on neural networks do not lead to a performance gain that justifies the cost.

5.1.2 Higher COICOP levels

Level 2 Starting at the highest level of aggregation, i.e. the level 2 category *01 Food and non-alcoholic beverages*, we see that webscraping performs very similarly to the various time series models and ensembles. For indirect one-step ahead forecasts (see Table A10), only the ARIMA model has a slightly lower RMSE. Direct one-step ahead forecasts (see Table A9) are also compared relative to webscraping, which is always based on the elementary indices. Therefore, the RMSE of 0.764 for webscraping is the same in both tables. This also implies that the relative RMSE for the models can be ranked across the tables for direct and indirect forecasts. Time series models for direct forecasts at level 2 have lower RMSE for Neuralprophet, Theta and the Ensemble (with webscraping). However, none of the pairwise Diebold-Mariano tests showed significant differences in prediction accuracy. We consider these results very promising because using indirect forecasts allows a decomposition into drivers at different levels,

Table 2: Forecast accuracy metrics for elementary indices (level 6).

Type	RMSE	MAE	MDA
Ensemble (w WS)	2.55	1.88	0.65
Ensemble (wo WS)	2.65	1.95	0.63
Webscraping	2.66	2.00	0.64
LW	2.76	2.02	0.60
AR(1)	2.77	2.02	0.61
ETS	2.77	2.04	0.62
Theta	2.84	2.10	0.60
ARIMA	2.89	2.16	0.51
Neuralprophet	2.91	2.18	0.62
CES	2.96	2.21	0.61
Prophet	3.00	2.26	0.59
Naive	4.21	3.17	0.47

Note: This table presents several metrics for assessing the forecast accuracy of the one-step-ahead forecasts for the elementary indices (level 6). The individual metrics are weighted with the official HICP weights for 2024.

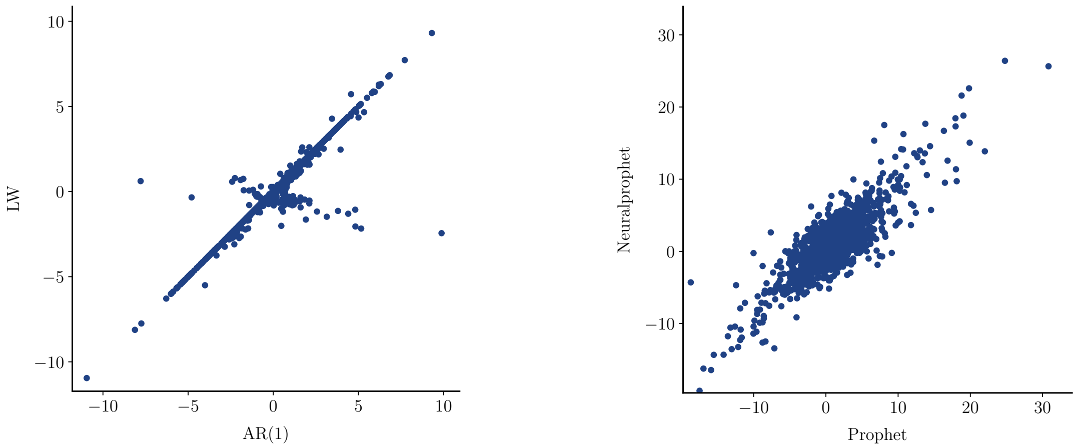
which would not be possible in the case of direct forecasts.

Level 3 For COICOP level 3, in the category *012 Non-alcoholic beverages*, webscraping significantly outperforms all time series models and ensembles. This holds for direct forecasts in Table A7 and for indirect forecasts in Table A8. In contrast, for the category *011 Food*, the results are similar to those at level 2 above. The Diebold-Mariano test does not indicate any significant differences in predictive ability for direct and indirect forecasts. Webscraping has the second lowest RMSE compared to indirect forecasts, where only the ARIMA model minimally outperforms it. For direct forecasts, Theta, ensemble with webscraping, Neuralprophet and CES show lower RMSE metrics.

Level 4 Table A6 shows the indirect forecasts of the eleven relevant COICOP 4 categories. Webscraping achieves superior predictive ability for most categories. Time series models perform better for *0116 Fruit* and *0119 Food n.e.c.*. The Diebold-Mariano tests were only significant at the 10% level for better RMSE of ARIMA and the ensemble with webscraping when forecasting indices for fruit. Table A5 presents the direct forecasts. Again, we see a slightly lower RMSE compared to the indirect forecasts. Webscraping also manages here to compete with the time series models and again the only category with significantly (better) different predictability is *0116 Fruit* up to a significance level of 5%. Compared to COICOP level 2 and level 3 forecasts, a gain in predictive ability is already noticeable at level 4, with more categories being outperformed by the nowcasts based on webscraping data.

Level 5 For COICOP level 5, in the case of direct forecasts of the indices in Table A3, webscraping outperforms all time series models, including the ensemble with webscraping, in 22 of the 50 subcategories,

Figure 3: One-step ahead univariate time series forecasts for elementary indices (level 6)



(a) This panel shows the forecasts of the AR(1) model versus the restricted special case by Lunsford and West (LW).

(b) This panel shows the forecasts of the Prophet model versus the more advanced NeuralProphet model.

or 24 when excluding the ensemble. The ensemble with webscraping performs better than webscraping alone in 27 categories. The ensemble without webscraping is superior to the ensemble with webscraping in only one case, i.e. *01161 Fresh or chilled fruit*. Other time series models, such as AR(1) and LW, outperform webscraping in 19 of the 50 categories. Webscraping performs poorly for categories such as *01131 Fresh or chilled fish* and *01132 Frozen fish*, but performs well for *01133 Dried, smoked or salted fish and seafood*. However, all models perform poorly for fish categories. Furthermore, webscraping performs worse for *01161 Fresh or chilled fruit*, *01172 Frozen vegetables other than potatoes and other tubers* (interestingly, it works quite well for *01171 Fresh or chilled vegetables other than potatoes and other tubers*), *01182 Jams, marmalades and honey*, as well as some subcategories of *0119 Food products n.e.c.* Differences between models are often not statistically significant (see table).

Regarding indirect forecasts for COICOP level 5 in Table A4, webscraping achieves superior predictive ability for 46 out of 50 subcategories. *01124 Poultry* shows a lower RMSE for autoregressive models and for both ensembles. In the subcategory *01131 Fresh or chilled fish* the ensemble with webscraping shows superior predictive ability but only at a 10% significance level. For *01161 Fresh or chilled fruit*, the ARIMA model and both ensembles also outperform at the 10% significance level. In the subcategory *01199 Other food products n.e.c.* the lower RMSE of ARIMA and ensemble with webscraping does not indicate a significantly different predictive ability according to the Diebold-Mariano tests.

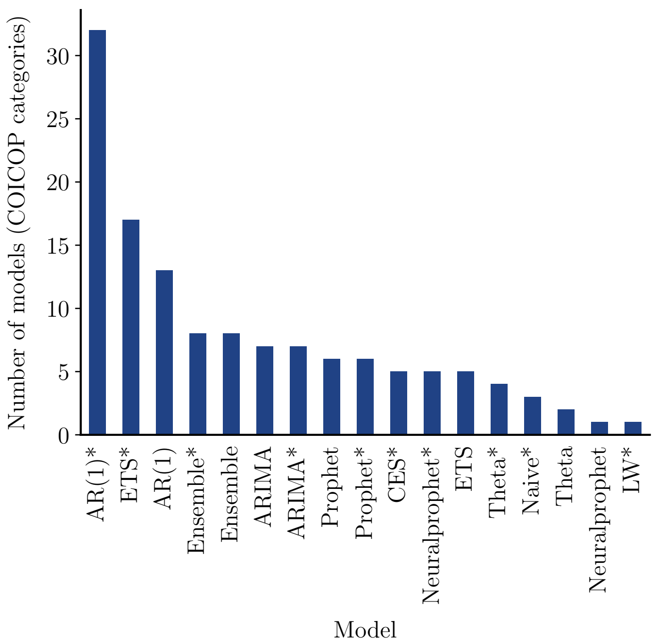
5.2 Three-step ahead forecasts

While the existing literature for disaggregated forecasting with webscraped data (see e.g. Macias et al., 2023; Soybilgen et al., 2023) mainly compares the performance of nowcasts vs. one-step ahead forecasts, we propose to complement multi-step ahead forecasts with webscraping nowcasts. This has be done for headline inflation in Aparicio and Bertolotto (2020) and for disaggregated forecasting, but without webscraping data, in Barkan et al. (2023).

5.2.1 Elementary indices

We will proceed with a similar analysis as for the one-step ahead forecasts. We rank the predictive accuracy of the three-step ahead forecasts for all models by RMSE. Figure 4 displays the number of first ranks per model for each of the 130 elementary indices. Better prediction accuracy is achieved by more parsimonious models such as AR(1)* and ETS*.

Figure 4: Number of first rank per model for elementary indices (level 6)



Note: This figure shows the number of first ranks per model for three-step ahead forecasts of the elementary indices at level 6. The stars * denote models where the webscraping based nowcasts were used for $h = 1$.

In Table 3 models marked with an asterisks contain the webscraping nowcast for $h = 1$. We calculate equally-weighted ensembles (excluding the naive specification) of all models with and without webscraping. The results suggest that the webscraping nowcasts lead to improvements in the error metrics of all individual models and of the ensemble. In line with the one-step ahead forecasts, after the ensemble simpler models like AR(1), LW, ETS have lower RMSE.

5.2.2 Higher COICOP levels

Level 2 Similar to the one-step ahead forecasts, we compare direct and indirect forecasts at the higher COICOP levels. Starting with the level 2 category *01 Food and non-alcoholic beverages* we see in Ta-

Table 3: Forecast accuracy metrics for elementary indices (level 6).

Type	RMSE	MAE	MDA
Ensemble*	4.21	3.19	0.69
LW*	4.26	3.20	0.68
AR(1)*	4.27	3.19	0.69
ETS*	4.29	3.25	0.68
Ensemble	4.30	3.27	0.72
ARIMA*	4.41	3.33	0.66
AR(1)	4.48	3.34	0.69
Theta*	4.49	3.44	0.68
LW	4.50	3.36	0.69
Prophet*	4.53	3.48	0.69
ETS	4.54	3.43	0.72
Theta	4.73	3.64	0.68
ARIMA	4.76	3.61	0.52
Neuralprophet*	4.78	3.71	0.68
CES*	4.81	3.68	0.68
Prophet	4.82	3.71	0.68
Neuralprophet	5.13	4.03	0.69
CES	5.24	4.04	0.67
Naive*	6.29	4.70	0.63
Naive	10.57	7.81	0.52

Note: This table presents several metrics for assessing the forecast accuracy of the three-step-ahead forecasts for the elementary indices (level 6). The individual metrics are weighted with the official HICP weights for 2024.

ble A18 that the inclusion of the webscraping based nowcast for $h = 1$ improves the RMSE of all time series models, indicated by RMSE ratios < 1. For ARIMA, AR(1), Prophet, Neuralprophet, CES and Theta, the results are significantly different up to a 5% significance level according to the Diebold-Mariano tests. Even more importantly, the indirect forecasts in Table A19 also show an significant increase in forecast accuracy for the Naive model as well as improvements of the ETS and Theta models. For all other specifications there are no significant differences in prediction accuracy according to the Diebold-Mariano tests.

Level 3 Moving up to level 3, a similar picture emerges. The indirect forecasts in Table A17 have better accuracy especially for the subcategory *012 Non-alcoholic beverages*. But also for *011 Food* three time series models have RMSE ratios < 1 and the forecasts by the other models are not significantly different. The direct forecasts in Table A16 are all improved by the addition of the webscraping based nowcasts.

Level 4 Also at COICOP level 4, the direct forecasts in Table A14 all have higher accuracy with one exception, the subcategory *0116 Fruit*. Here, only time series based models were able to produce significantly different forecasts with higher accuracy measured by RMSE. The indirect forecasts are presented in Table A15. More than half of the RMSE ratios are below one, indicating a boost in prediction performance when adding webscraping data. The naive forecasts improve almost entirely across all subcategories. Higher RMSE ratios with significant differences are present in the subcategories *0116 Fruit* and *0119 Food products n.e.c.*, *0113 Fish* and *0111 Bread and cereals*.

Level 5 At COICOP level 5 the improved forecast accuracy of the three-step ahead forecasts looks even more visible in the results for the indirect forecasts in Table A13. Of course, mostly the categories without superior outperformance in the nowcasts are not able to achieve a boost in forecast accuracy in the three-step ahead forecasts. These are the subcategories *01131 fresh or chilled fish*, *01161 fresh or chilled fruit*, *01192 alt, spices and culinary herbs*. There are only eight cases where the ensemble including webscraping yields a significantly different performance with RMSE ratios above one compared to an ensemble of time series models without webscraping. Regarding the direct forecast in Table A12 the results are in line with the higher COICOP levels and the increase in forecast accuracy is even more pronounced.

6 Conclusion

Our study demonstrates that incorporating carefully constructed webscraping indices – by ensuring the quality of the classification of products at the elementary index level and applying official index calculation methodologies – significantly boosts the accuracy of one-step ahead food inflation nowcasts and short-term forecasts for a three-month horizon. While direct forecasts at higher levels of aggregation produce slightly more accurate overall metrics, indirect forecasts derived from disaggregated data provide superior insights into the underlying dynamics of specific sub-components. The main advantage of indirect forecasting is that higher level aggregates can be accurately broken down into their sub-components, helping forecasters to communicate exactly what is driving their forecast.

We extend the existing literature by being the first to integrate webscraped data into time series-based short-term forecasts of disaggregated food inflation rates for Austria. Our findings reveal that webscraping not only reduces forecast uncertainty on a disaggregated level, but also enhances predictability when aggregated to higher COICOP levels. Additionally, we observe that employing models beyond traditional ARIMA, such as advanced time series and machine learning models, can offer benefits, although they may involve trade-offs in computational efficiency.

For practical inflation forecasting, this means that using state-of-the-art models and software from the machine learning community can help build reliable and automated real-time forecasting tools. Even without webscraping data available for all ECOICOP categories, our setup can easily incorporate forecasts based on expert judgment or exogenous regressors, thereby providing a flexible and robust framework. These advances have significant implications for policymakers and practitioners seeking to develop effective systems for real-time monitoring and forecasting of inflation dynamics at a granular level.

References

Anesti, N., Esady, V., and Naylor, M. (2024). Food prices matter most: Sensitive household inflation expectations. Technical Report CFM-DP2024-34, LSE Centre for Macro Discussion Paper. 2

Aparicio, D. and Bertolotto, M. I. (2020). Forecasting inflation with online prices. *International Journal of Forecasting*, 36(2):232–247. 2, 3, 10, 16

Assimakopoulos, V. and Nikolopoulos, K. (2000). The theta model: a decomposition approach to forecasting. *International Journal of Forecasting*, 16(4):521–530. The M3- Competition. 9

Atkeson, A. and Ohanian, L. E. (2001). Are Phillips Curves Useful for Forecasting Inflation? *Federal Reserve Bank of Minneapolis Quarterly Review*, 25(1):2–11. Winter 2001. 7

Barkan, O., Benchimol, J., Caspi, I., Cohen, E., Hammer, A., and Koenigstein, N. (2023). Forecasting CPI inflation components with Hierarchical Recurrent Neural Networks. *International Journal of Forecasting*, 39(3):1145–1162. 2, 3, 10, 16

Beck, G. W., Carstensen, K., Menz, J.-O., Schnorrenberger, R., and Wieland, E. (2023). Nowcasting Consumer Price Inflation Using High-Frequency Scanner Data: Evidence from Germany. *Deutsche Bundesbank Discussion Paper No. 34*. 3

Borio, C., Disyatat, P., Juselius, M., and Rungcharoenkitkul, P. (2022). Why So Low for So Long? A Long-Term View of Real Interest Rates. *International Journal of Central Banking*. 8, 12

Carrera de Souza, T. (2022). Reassessing Sticky Price Models through the Lens of Scraped Price Data. *Ensayos Económicos*, 1(79):66–108. 2

Cavallo, A. and Rigobon, R. (2016). The Billion Prices Project: Using Online Prices for Measurement and Research. *Journal of Economic Perspectives*, 30(2):151–178. 2, 3

D'Acunto, F., Malmendier, U., Ospina, J., and Weber, M. (2021). Exposure to grocery prices and inflation expectations. *Journal of Political Economy*, 129(5):1615–1639. 2

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263. 10, 12, 15, 18, 21, 22, 23, 24, 25, 34, 37, 38, 39, 40, 41

Eurostat (2022). *Guide on Multilateral Methods in the Harmonised Index of Consumer Prices: 2022 Edition*. Luxembourg. Publications Office of the European Union. 1

Eurostat (2024). Harmonised index of consumer prices (hicp) methodological manual – 2024 edition. *Eurostat Manual and Guidelines*. 5, 1, 2

Faust, J. and Wright, J. H. (2013). Chapter 1 - Forecasting Inflation. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2 of *Handbook of Economic Forecasting*, pages 2–56. Elsevier. 2, 10, 11

Fiorucci, J. A., Pellegrini, T. R., Louzada, F., Petropoulos, F., and Koehler, A. B. (2016). Models for optimising the theta method and their relationship to state space models. *International Journal of Forecasting*, 32(4):1151–1161. 9

Garza, F., Canseco, M. M., Challú, C., and Olivares, K. G. (2022). StatsForecast: Lightning fast forecasting with statistical and econometric models. PyCon Salt Lake City, Utah, US 2022. 9, 10

Gautier, E., Karadi, P., Conflitti, C., Fabo, B., Fadejeva, L., Fuss, C., Kosma, T., Jouvanceau, V., Martins, F., Menz, J.-O., Messner, T., Petroulas, P., Roldan-Blanco, P., Rumler, F., Santoro, S., Wieland, E., and Zimmer, H. (2023). Price Adjustment in the Euro Area in the Low-Inflation Period: Evidence from Consumer and Producer Micro Price Data. Occasional Paper Series 319, European Central Bank. 2

Graf, B. (2024). Inflation forecasting with the Lunsford-West model. *Internal OeNB presentation*. 8, 12

Hyndman, R., Koehler, A., Ord, J., and Snyder, R. (2008). *Forecasting with Exponential Smoothing: The State Space Approach*. Springer Berlin Heidelberg. 9

Ivancic, L., Diewert, W. E., and Fox, K. J. (2011). Scanner data, time aggregation and the construction of price indexes. *Journal of Econometrics*, 161(1):24–35. Measurement with Theory. 1, 2

Lunsford, K. G. and West, K. D. (2019). Some Evidence on Secular Drivers of US Safe Real Rates. *American Economic Journal: Macroeconomics*, 11(4):113–39. 8, 12

Macias, P., Stelmasiak, D., and Szafranek, K. (2023). Nowcasting food inflation with a massive amount of online prices. *International Journal of Forecasting*, 39(2):809–826. 3, 10, 15

Medeiros, M. C., Vasconcelos, G. F. R., Veiga, Á., and Zilberman, E. (2019). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics*, 39(1):98–119. 2

Russel, R. (2022). Statsforecast: ARIMA and Prophet Adapter. `https://github.com/Nixtla/statsforecast/tree/main/experiments/arima_prophet_adapter`. Accessed: 2024-09-03. 8

Soybilgen, B., Yazgan, M. E., and Kaya, H. (2023). Nowcasting Turkish Food Inflation Using Daily Online Prices. *Journal of Business Cycle Research*, 19(2):171–190. 3, 15

Statistik Austria (2022). *Standard-Dokumentation Metainformationen zum Verbraucherpreisindex und Harmonisierter Verbraucherpreisindex*. 5, 1, 2

Svetunkov, I., Kourentzes, N., and Ord, J. K. (2022). Complex exponential smoothing. *Naval Research Logistics (NRL)*, 69(8):1108–1123. 10

Taylor, S. J. and Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1):37–45. 8, 12

Triebe, O., Hewamalage, H., Pilyugina, P., Laptev, N., Bergmeir, C., and Rajagopal, R. (2021). Neuralprophet: Explainable forecasting at scale. 8, 12

Vicente, H. d. O. R. and Valls Pereira, P. L. (2022). Forecasting Inflation using Online Daily Prices: A MIDAS Approach for Brazil. Technical report. 3

Improving disaggregated short-term food inflation forecasts with webscraped data

Appendix for online publication

Christian Beer, Robert Ferstl, Bernhard Graf

First version: November 11, 2024
This version: January 10, 2025

Most recent version available at:
`https://ssrn.com/abstract=5016579`

A Index compilation

In this appendix, we describe how Eurostat's official methodology (Eurostat, 2024, 2022) is applied to our data. The first step is to calculate the average monthly price for each product. Consistent with Statistik Austria's on-site price collection approach (Statistik Austria, 2022), we include only webscraped product prices from the first three weeks of each month. Let t represent a specific month within a given year, with the price of product i on day d of t denoted as $p_i^{d,t}$. We then compute the geometric average of product prices up to the third week of each month as follows:

$$p_i^t = \left(\prod_{d=1}^{21} p_i^{d,t} \right)^{\frac{1}{21}}.$$

We construct our elementary index as a Fixed Base Jevons Index, which employs the geometric mean of prices as calculated above within each COICOP level 6 category as:

$$P_{\text{Jevons}}^{0,t} = \frac{\left(\prod_{i=1}^{N} p_i^t\right)^{\frac{1}{N}}}{\left(\prod_{i=1}^{N} p_i^0\right)^{\frac{1}{N}}} = \left(\prod_{i=1}^{N} \frac{p_i^t}{p_i^0} \right)^{\frac{1}{N}},$$

where N represents the number of items within a given COICOP level 6 category. The index is calculated for month t and the base month 0, which, following Eurostat's notation, refers to December of the previous year.

A.1 Multilateral GEKS Jevons Index

In the next step, we use these monthly elementary indices $P_{\text{Jevons}}^{0,t}$ to construct GEKS Jevons Indices based on Ivancic et al. (2011). The GEKS index is defined as the geometric mean of the ratios of all bilateral indices, with each period being used as the base period in turn. Let T denote the total number of time periods (i.e., unique combinations of month and year) for which we aim to create a comparison. The GEKS-Jevons index is then defined as

$$P_{\text{GEKS-Jevons}}^{0,t} = \sqrt[T]{\prod_{l=0}^{T-1} \frac{P_{\text{Jevons}}^{l,t}}{P_{\text{Jevons}}^{l,0}}} = \prod_{l=0}^{T-1} \left(P_{\text{Jevons}}^{0,l} P_{\text{Jevons}}^{l,t} \right)^{\frac{1}{T}},$$

using the identity $P_{\text{Jevons}}^{0,t} = \left(\prod_{i=1}^{N} \frac{p_i^t}{p_i^0} \right)^{\frac{1}{N}} = \left(\prod_{i=1}^{N} \frac{p_i^0}{p_i^t} \right)^{-\frac{1}{N}} = \frac{1}{P_{\text{Jevons}}^{t,0}}$ to derive the last term in this equation.

A.2 Rolling window GEKS - Multilateral Jevons Index

A stark disadvantage of this "regular" GEKS index is that as soon as new data become available, the entire index time series would have to be recalculated. To counteract this, the rolling window GEKS index can be employed (Ivancic et al., 2011). The rolling windows GEKS involves the definition of a time window width w, the number of consecutive periods to be used for the index calculation. In alignment with the literature (e.g. Ivancic et al., 2011) and the praxis employed by Statistik Austria (2022) we choose a window lenght of $w = 13$, which should capture seasonalites in monthly data. For the first 12 (i.e. $w-1$) time periods the rolling windows GEKS is calculated in the same way as the GEKS index. Afterwards, for further periods, the window is moved one step further and the index is extended, leaving the index of the previous time periods unchanged. The time series is thus only updated by one value for the "new" period. This method of splicing, the joining of time series, is called movement splice.

The index calculation can then be written as follows denoting the current month with t:

$$P^{0,t}_{\text{RWGEKS-Jevons}} = \begin{cases} P^{0,t}_{\text{GEKS-Jevons}} & \text{if } t < 13 \\ P^{0,t-1}_{\text{RWGEKS-Jevons}} \left(\frac{P^{t-W+1,t}_{\text{GEKS-Jevons}}}{P^{t-W,t-1}_{\text{GEKS-Jevons}}} \right) & \text{otherwise,} \end{cases}$$

where $P^{s,t}_{\text{GEKS-Jevons}} = \left(\prod_{r=s}^{t} \frac{P^{s,r}_{\text{Jevons}}}{P^{t,r}_{\text{Jevons}}} \right)^{\frac{1}{13}}$

A.3 Index aggregation

To perform calculations at more aggregated COICOP levels, we need to aggregate the previously described elementary indices. For clarity, we replace the more generic time index t with (m, y), where m represents the month and y represents the year. Thus, we can express the average price of product i in month m and year y as $p^{m,y}_i$.

When aggregating to a higher-level COICOP category K, we apply yearly HICP weights provided by Statistik Austria. These weights reflect households' consumption shares for various goods. We use these weights $\bar{w}^y_k$ to sum the elementary indices $k \in K$. This leads to the following expression:

$$P^{m,y,K}_{\text{RWGEKS-Jevons}} = \sum_{k=1}^{K} w^y_k P^{m,y,k}_{\text{RWGEKS-Jevons}} \quad \text{with} \quad w^y_k = \frac{\bar{w}^y_k}{\sum_{k=1}^{K} \bar{w}^y_k},$$

where the last term ensures that $\sum_{k=1}^{K} w^y_k = 1$ for the aggregated COICOP index time series K.

A.4 December as the linking month

We will link the previously introduced RWGEKS-Jevons index over time. For didactic reasons, we will show here how the concept works with a fixed Jevons index. The fixed monthly Jevons Index is calculated using the annual base period of December of the previous year. As mentioned above, all products available in the respective comparison period and in December of the previous year are included in the calculation of the index. In order to obtain a continuous time series, the time series for the year y is rescaled using the value for December $y-1$ in relation to the respective base period (see Eurostat, 2024). For arbitrary ECOICOP category K we arrive at:

$$P_{\text{Jevons}}^{m,y,K} = \underbrace{\sum_{k=1}^{K} w_k^y \left(\prod_{i=1}^{A} \frac{p_i^{m,y}}{p_i^{12,y-1}} \right)^{\frac{1}{A}}}_{\textit{current} \text{ vintage}} \times \underbrace{\sum_{k=1}^{K} w_k^{y-1} \left(\prod_{i=1}^{B} \frac{p_i^{12,y-1}}{p_i^{12,y-2}} \right)^{\frac{1}{B}}}_{\textit{last} \text{ vintage}},$$

with A product prices observed in month m in year y and month 12 of year $y-1$ and likewise for B.

Aggregation weights for year t are employed for product price set of the ”current vintage” while aggregation weights for $t-1$ are used in the case of set of the ”last vintage” in order to arrive at a chain-linked index time series such that we arrive at a chain-linked index.

A.5 Basket 2024 and HICP weights

We provide here the relevant part on food and non-alcoholic beverages of the Harmonized Index of Consumer Prices (HICP) basket from Statistik Austria.[16] The first column shows the E-COICOP classification. Rows that contain an entry in the second column denote an elementary index defined for Austria. The last but one column (items) shows the number of elementary indices relevant for a specific E-COICOP category, while the last column (products) shows the number of products classified in this category in 2024.[17]

Table A1: Basket 2024, Source: Statistik Austria

COICOP	CODE	Text	Weight 2024	Items	Products
01		**FOOD AND NON-ALCOHOLIC BEVERAGES**	**11.90919**	**132**	**10404**
011		**Food**	**10.60782**	**120**	**8448**
0111		*Bread and cereals*	*2.32975*	*21*	*2242*
01111		Rice	0.06261	1	57
011110	001900	Long grain rice	0.06261		57
01112		Flours and other cereals	0.11536	1	76
011120	002000	Wheat flour	0.11536		76
01113		Bread	1.09002	7	317
011130	000600	Rye bread	0.24119		24
011130	000700	Wholemeal bread	0.14218		111
011130	000800	White bread	0.17489		52
011130	000900	Bread roll	0.16415		22
011130	001000	Scone, handmade	0.25052		40
011130	001100	Half baked buns	0.05959		27
011130	001110	Lye bun	0.0575		41
01114		Other bakery products	0.65475	7	291
011140	000200	Yeast dumpling, deep frozen	0.02193		36
011140	001500	Butter biscuits	0.06912		76
011140	001550	Cake	0.10921		65
011140	001600	Wafers with hazelnut cream	0.04273		38
011140	001700	Salty sticks	0.02247		37
011140	002300	Nut cake	0.27159		24
011140	002400	Curd cheese cake	0.1177		15
01115		Pizza and quiche	0.08531	1	173
011150	000100	Pizza, deep frozen	0.08531		173
01116		Pasta products and couscous	0.19477	2	902
011160	000300	Convenience food, deep frozen	0.05759		68

Continued on next page

[16]The full table is available here: `https://www.statistik.at/fileadmin/pages/214/Basket_CPI_HICP_for_2024.pdf`. The weights for previous years is also available on the Statistik Austria homepage.

[17]Mayonnaise and long-life milk were only introduced in 2024 and therefore not taken into account in our analysis.

COICOP	CODE	Text	Weight 2024	Items	Products
011160	002100	Pasta	0.13718		834
01117		Breakfast cereals	0.0864	1	277
011170	001400	Cereals	0.0864		277
01118		Other cereal products	0.04053	1	149
011180	001200	Ready-made dough	0.04053		149
0112		*Meat*	*2.32392*	*23*	*1426*
01121		Beef and veal	0.30429	4	28
011210	002500	Veal cutlet	0.0246		4
011210	003100	Roast beef	0.10646		10
011210	003200	Beef round	0.09529		9
011210	003300	Beef shoulder	0.07794		5
01122		Pork	0.34364	4	79
011220	003500	Pork belly	0.03646		15
011220	003600	Pork chops	0.0783		38
011220	003700	Pork cutlet	0.11972		21
011220	003800	Pork sirloin	0.10916		5
01124		Poultry	0.40981	4	109
011240	004700	Roast chicken	0.08945		41
011240	004800	Turkey breast	0.11026		14
011240	004900	Breaded chicken meat, deep frozen	0.04357		23
011240	005000	Chicken breast	0.16653		31
01127		Dried, salted or smoked meat	1.0252	8	1118
011270	002700	Smoked meat	0.05455		10
011270	002800	Bacon	0.11267		133
011270	004000	Sausages	0.21262		272
011270	004100	Dry sausage	0.09746		143
011270	004200	Pork ham	0.23307		297
011270	004300	Turkey sausage	0.06586		19
011270	004400	Pork sausage	0.1477		38
011270	004600	Salami	0.10127		206
01128		Other meat preperations	0.24098	3	92
011280	002600	Minced meat	0.14748		57
011280	002900	Liver pasty	0.04644		29
011280	003000	Beef goulash canned	0.04706		6
0113		*Fish*	*0.38066*	*5*	*293*
01131		Fresh or chilled fish	0.10865	1	67
011310	005400	Fresh fish	0.10865		67
01132		Frozen fish	0.13989	2	67
011320	005200	Codfish filet, deep frozen	0.07825		34
011320	005300	Fish fingers, deep frozen	0.06164		33

Continued on next page

COICOP	CODE	Text	Weight 2024	Items	Products
01135		Dried, smoked or salted fish and seafood	0.04952	1	37
011350	005500	Smoked salmon	0.04952		37
01136		Other preserved or processed fish and seafood-based preperations	0.0826	1	122
011360	005100	Tunafish	0.0826		122
0114		*Milk, cheese and eggs*	*1.61651*	*17*	*1387*
01141		Whole Milk	0.24142	3	107
011410	005660	Extended shelf life milk	0.14876		43
011410	005661	Fresh milk	0.04561		64
011410	005662	Long-life milk	0.04705		0
01143		Preserved milk	0.04411	1	16
011430	006100	Evaporated milk	0.04411		16
01144		Yoghurt	0.18026	2	341
011440	006200	Yoghurt fruit flavoured	0.09367		224
011440	006400	Yoghurt	0.08659		117
01145		Cheese and curd	0.67221	7	617
011450	006000	Curd cheese	0.02495		35
011450	006600	Emmentaler	0.13874		68
011450	006700	Gouda	0.17494		74
011450	006800	Hard cheese	0.10574		96
011450	006900	Camembert	0.08255		52
011450	007100	Fresh cream cheese	0.06412		202
011450	007200	Mozzarella	0.08117		90
01146		Other milk products	0.2634	3	197
011460	005700	Milk shake	0.11876		114
011460	005800	Sour cream	0.06182		31
011460	005900	Whipped cream	0.08282		52
01147		Eggs	0.21511	1	109
011470	006500	Eggs	0.21511		109
0115		*Oil and fats*	*0.34969*	*4*	*364*
01151		Butter	0.13666	1	100
011510	007300	Butter	0.13666		100
01152		Margarine and other vegetable fats	0.0324	1	25
011520	007400	Margarine	0.0324		25
01153		Olive oil	0.08109	1	78
011530	007800	Olive oil	0.08109		78
01154		Other edible oils	0.09954	1	161
011540	007600	Pure vegetable oil	0.09954		161
0116		*Fruit*	*0.85823*	*16*	*435*
01161		Fresh or chilled fruit	0.6791	13	333

Continued on next page

COICOP	CODE	Text	Weight 2024	Items	Products
011610	008200	Tangerines	0.04605		22
011610	008400	Strawberries	0.06437		26
011610	008440	Berries	0.03554		61
011610	008500	Apples	0.09327		72
011610	008600	Pears	0.01833		29
011610	008800	Peaches/nectarines	0.06553		11
011610	008900	Grapes	0.07821		16
011610	009000	Musk melons/cantaloupes	0.03973		11
011610	009200	Bananas	0.10919		16
011610	009300	Kiwi	0.02126		18
011610	009400	Oranges	0.0471		23
011610	009500	Lemons	0.0372		17
011610	010490	Avocado	0.02332		11
01163		Dried fruit and nuts	0.17913	3	102
011630	007900	Dried fruit mix with nuts	0.12637		51
011630	008000	Peanuts salted	0.02171		34
011630	008100	Raisins	0.03105		17
0117		*Vegetables*	*1.25222*	*15*	*537*
01171		Fresh or chilled vegetables other than potatoes and other tubers	0.70984	9	256
011710	010500	Champignons	0.03904		15
011710	010600	Cucumbers	0.04873		31
011710	010700	Cauliflower	0.07021		6
011710	010800	Carrots	0.05394		23
011710	010900	Iceberg lettuce	0.09231		24
011710	011000	Paprika	0.07872		49
011710	011220	Packed salad	0.05302		10
011710	011400	Tomatoes	0.19532		62
011710	011700	Onions	0.07855		36
01172		Frozen Vegetables other tahn potatoes and other tubers	0.07235	2	119
011720	009900	Mixed vegetables, deep frozen	0.04695		85
011720	010000	Spinach, deep frozen	0.0254		34
01173		Dried vegetables, other preserved or processed vegetables	0.18561	1	20
011730	009700	Pickled cucumbers	0.18561		20
01174		Potatoes	0.1945	2	76
011740	010100	French fries, deep frozen	0.04548		40
011740	011800	Potatoes	0.14902		36
01175		Crisps	0.08992	1	66

Continued on next page

COICOP	CODE	Text	Weight 2024	Items	Products
011750	010200	Potatoe chips	0.08992		66
0118		*Sugar, jam, honey, chocolate and confectionery*	*0.90364*	*10*	*866*
01181		Sugar	0.06944	1	10
011810	013200	Crystal sugar	0.06944		10
01182		Jams, marmalades and honey	0.09697	3	267
011820	012300	Tinned peaches	0.03439		10
011820	013000	Jam	0.03874		168
011820	013100	Natural honey	0.02384		89
01183		Chocolate	0.33159	3	352
011830	012500	Milk chocolate	0.13246		117
011830	012550	Chocolate box	0.16003		116
011830	012600	Chocolate bar	0.0391		119
01184		Confectionery products	0.24835	2	155
011840	012700	Chewing gum	0.10772		93
011840	012800	Fruit drops/ -jelly	0.14063		62
01185		Edible ices and ice cream	0.15729	1	82
011850	012200	Ice cream, family size	0.15729		82
0119		*Food products n.e.c.*	*0.5932*	*9*	*898*
01191		Sauces, condiments	0.18639	4	446
011910	013900	Mustard	0.03709		144
011910	014000	Ketchup	0.05216		125
011910	014100	Vinegar	0.05388		177
011910	014104	Mayonnaise	0.04326		0
01192		Salt, spices and culinary herbs	0.11311	2	44
011920	013700	Salt	0.02316		15
011920	013800	Red pepper	0.08995		29
01193		Baby food	0.08316	1	151
011930	014103	Baby food (milk)	0.08316		151
01194		Ready-made meals	0.1558	1	148
011940	014101	Convenience food, chilled	0.1558		148
01199		Other food products n.e.c.	0.05474	1	109
011990	013400	Soup powder	0.05474		109
012		**Non-alcoholic beverages**	**1.30137**	**12**	**1956**
0121		*Coffee, tea and cocoa*	*0.49973*	*5*	*1070*
01211		Coffee	0.40046	3	669
012110	014300	Coffee	0.19338		329
012110	014301	Coffeepads /-caps	0.14288		282
012110	014302	Instant coffee	0.0642		58
01212		Tea	0.07619	1	374
012120	014200	Tea in bags	0.07619		374

Continued on next page

COICOP	CODE	Text	Weight 2024	Items	Products
01213		Cocoa and powdered chocolate	0.02308	1	27
012130	014500	Cocoa instant drink	0.02308		27
0122		*Mineral waters, soft drinks, fruit and vegetables juices*	*0.80164*	*7*	*886*
01221		Mineral or spring waters	0.15755	1	205
012210	014800	Mineral or table water	0.15755		205
01222		Softdrinks	0.40832	4	554
012220	014600	Mineral water, flavoured	0.04118		91
012220	014700	Energy drink	0.07836		144
012220	014900	Soft drink carbonated	0.10919		159
012220	015000	Cola	0.17959		160
01223		Fruit and vegetable juices	0.23577	2	127
012230	015200	Orange juice	0.14698		55
012230	015300	Apple juice	0.08879		72

B Detailed results

B.1 One-step ahead forecasts

B.1.1 Elementary indices (level 6)

Table A2: One-step ahead forecasts of elementary indices (level 6)

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
001900	Long grain rice	2.064	2.190**	1.296	1.150	1.150	1.245*	1.537**	1.415*	1.543***	1.291*	1.057	1.110
002000	Wheat flour	2.960	1.832**	1.360**	1.348**	1.348**	1.347**	1.333**	1.300*	1.312*	1.301**	1.238*	1.288**
000600	Rye bread	1.680	1.029	0.933	0.799	0.799	0.787	0.826	0.772	0.857	0.756	0.718	0.778
000700	Wholemeal bread	0.949	1.136	1.053	0.957	0.957	0.961	1.052	0.966	0.936	0.884	0.889	0.925
000800	White bread	1.841	1.473*	1.266	0.979	0.979	1.040	1.037	0.967	1.164	0.993	0.957	1.006
000900	Bread roll	0.963	2.094**	1.707**	1.621*	1.621*	1.625**	1.588*	1.559**	1.660**	1.551**	1.435*	1.550**
001000	Scone, handmade	1.462	0.927	0.740	0.692	0.692	0.700	0.781	0.648	0.695	0.641	0.628*	0.664
001100	Half baked buns	2.084	1.181	0.815	0.807	0.807	0.825	0.815	0.778*	0.787	0.792	0.783	0.782
001110	Lye bun	2.326	1.746***	1.231	0.999	0.999	1.058	0.987	1.051	1.221	1.059	0.989	1.001
000200	Yeast dumpling, deep frozen	8.023	1.721***	0.883	0.933	0.933	1.026	0.900	1.005	1.038	1.076**	0.936	0.933
001500	Butter biscuits	2.519	1.245*	0.985	0.875	0.875	0.794	0.848	0.811	0.824	0.818	0.767*	0.771
001550	Cake	1.593	1.291	1.133	1.087	1.087	1.016	1.081	1.021	1.075	0.945	0.905	0.986
001600	Wafers with hazelnut cream	2.154	1.259	1.047	0.930	0.930	0.865	0.957	0.894	0.899	0.877	0.806	0.824
001700	Salty sticks	3.822	1.551***	1.070	0.954	0.954	1.027	1.128	1.103	1.095	1.022	0.895	0.916
002300	Nut cake	1.562	0.892	0.735*	0.658**	0.658**	0.665**	0.723*	0.663**	0.711*	0.618**	0.581***	0.630**
002400	Curd cheese cake	4.325	1.842***	0.990	0.954	0.954	1.056	1.151**	1.149**	1.159**	1.109	1.003	1.011
000100	Pizza, deep frozen	3.047	1.682***	1.028	0.963	0.963	1.010	1.087	1.069	1.053	1.048	0.924	0.936
000300	Convenience food, deep frozen	2.672	1.646**	1.064	0.977	0.977	1.075	1.014	1.045	1.055	1.051	0.942	0.985

Continued on next page

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
002100	Pasta	2.434	1.806***	1.128	1.060	1.060	1.173*	0.933	1.139	1.045	1.189	1.003	1.027
001400	Cereals	1.714	1.817*	0.951	1.040	1.040	1.025	1.190	1.097	1.035	1.079	0.893	0.920
001200	Ready-made dough	1.666	2.007***	1.525*	1.416	1.416	1.366	1.276	1.238	1.340*	1.255	1.201	1.264
002500	Veal cutlet	4.089	1.422**	0.618***	0.609***	0.609***	0.753**	0.750**	0.753**	0.689***	0.743**	0.661***	0.634***
003100	Roast beef	2.131	1.605**	1.211	0.916	0.916	0.924	1.682***	1.213	1.352*	0.964	0.884	0.957
003200	Beef round	3.025	1.350*	0.990	0.962	0.962	0.911	1.178	1.134	0.973	1.022	0.895	0.920
003300	Beef shoulder	3.203	0.963	0.718	0.687*	0.687*	0.624**	0.797	0.772	0.818	0.643**	0.616**	0.647**
003500	Pork belly	2.731	1.836**	1.246	1.170	1.170	1.159	1.208	1.167	1.171	1.169	0.904	0.958
003600	Pork chops	3.465	1.648**	1.011	1.007	1.007	1.034	1.286*	1.285*	1.232	1.078	1.010	1.032
003700	Pork cutlet	4.256	1.516**	1.044	0.899	0.899	0.939	1.437	1.224	1.182	1.079	0.902	0.971
003800	Pork sirloin	6.304	0.633	0.524*	0.442**	0.442**	0.416**	0.482**	0.453**	0.447**	0.430**	0.456**	0.428**
004700	Roast chicken	2.747	1.013	0.855	0.640**	0.640**	0.628**	0.893	0.792	0.792	0.631**	0.649**	0.663**
004800	Turkey breast	1.891	1.275	1.178	0.911	0.911	0.874	0.989	1.008	1.100	0.879	0.814	0.913
004900	Breaded chicken meat, deep frozen	2.734	1.656**	1.497	1.222	1.222	1.132	1.202	1.224	1.549*	1.145	1.037	1.113
005000	Chicken breast	1.986	1.180	0.930	0.792	0.792	0.743*	0.976	0.928	0.827	0.736*	0.728*	0.762
002700	Smoked meat	2.188	0.958	0.844	0.772	0.772	0.669*	0.958	0.765	0.774	0.671*	0.656*	0.714
002800	Bacon	1.167	1.286*	1.041	1.012	1.012	0.953	1.151	1.063	0.973	0.966	0.908	0.973
004000	Sausages	1.084	1.633**	1.303	1.222	1.222	1.159	1.239	1.227	1.162	1.166	1.107	1.168
004100	Dry sausage	0.884	1.555**	1.571	1.485	1.485	1.271	1.836**	1.440	1.464	1.254	1.223	1.365
004200	Pork ham	1.062	1.733***	1.457*	1.333	1.333	1.319*	1.566**	1.465**	1.386**	1.248	1.213	1.295
004300	Turkey sausage	2.142	1.863**	1.363	1.254	1.254	1.319	1.410	1.300	1.516**	1.237	1.130	1.259
004400	Pork sausage	1.943	1.635**	1.562**	1.220*	1.220*	1.164*	1.714***	1.650***	1.541***	1.186**	1.197**	1.253**
004600	Salami	1.326	1.601*	1.411	1.355	1.355	1.093	1.348	1.319**	1.299*	1.100	1.085	1.153
002600	Minced meat	4.109	1.517**	1.325	1.121	1.121	1.023	1.093	1.145	1.239	1.059	1.000	1.040
002900	Liver pasty	2.265	1.804***	1.074	1.064	1.064	1.056	1.158	1.101	1.067	1.096	0.985	1.013
003000	Beef goulash canned	1.849	1.955	1.357	1.166	1.166	1.149	1.314	1.269	1.159	1.154	0.985	1.112
005400	Fresh fish	3.919	0.788	0.535**	0.505**	0.505**	0.510**	0.758	0.628**	0.603**	0.530**	0.520***	0.532**

Continued on next page

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
005200	Codfish filet, deep frozen	4.687	1.494**	1.062	0.961	0.961	0.957	0.947	0.985	0.988	0.994	0.936	0.950
005300	Fish fingers, deep frozen	5.539	1.798***	0.858	0.812	0.812	1.005	0.895	0.970	0.927	1.036	0.864*	0.861
005500	Smoked salmon	3.314	1.861**	1.073	1.132	1.132	1.145	1.040	1.039	1.023	1.069	1.004	1.027
005100	Tunafish	2.778	1.509**	0.836*	0.784**	0.784**	0.880	1.056	1.045	0.966	0.930	0.861*	0.864
005660	Extended shelf life milk	1.347	2.413**	1.884*	1.769	1.769	1.857*	1.890*	1.675	2.156**	1.790*	1.645	1.785
005661	Fresh milk	1.465	2.152**	1.794*	1.605	1.605	1.644	1.733*	1.536	1.701*	1.615	1.498	1.618
006100	Evaporated milk	2.333	1.879	1.332*	1.210	1.209	0.965	1.339	1.176	1.352	0.982	0.896	0.974
006200	Yoghurt fruit flavoured	1.941	1.686***	1.166	1.098	1.098	1.134	1.148	1.163*	1.246**	1.119	1.080	1.112
006400	Yoghurt	1.445	1.863**	1.508*	1.461*	1.461*	1.415*	1.584**	1.461**	1.388	1.459*	1.322	1.405
006000	Curd cheese	1.604	1.946**	1.590*	1.498	1.498	1.602*	1.674	1.572	1.699*	1.580*	1.433	1.540
006600	Emmentaler	1.993	1.370*	1.154	1.108	1.108	1.072	1.334*	1.251	1.280	1.095	1.054	1.113
006700	Gouda	1.971	1.814***	1.416**	1.338**	1.338**	1.368**	1.521**	1.409*	1.452**	1.315**	1.261*	1.343**
006800	Hard cheese	2.105	1.533**	1.124	0.996	0.996	0.954	1.112	0.991	1.098	0.979	0.933	0.970
006900	Camembert	3.046	1.947***	1.169	1.083	1.083	1.157**	1.299**	1.333**	1.203*	1.204**	1.095	1.123
007100	Fresh cream cheese	2.067	1.737*	1.073	0.996	0.996	0.997	1.059	0.980	0.940	0.994	0.932	0.949
007200	Mozzarella	1.912	1.633**	1.420***	1.298*	1.298*	1.265*	1.328*	1.237	1.185	1.279**	1.181	1.239*
005700	Milk shake	1.820	1.679**	1.110	1.069	1.069	1.088	1.089	1.082	1.043	1.116	0.984	1.029
005800	Sour cream	2.100	1.972***	1.455*	1.466*	1.466*	1.489**	1.594**	1.496**	1.613**	1.521**	1.394*	1.468*
005900	Whipped cream	1.749	2.163**	1.557	1.531	1.531	1.528	1.726*	1.590*	1.888**	1.578	1.473	1.554
006500	Eggs	2.079	1.145	1.058	0.954	0.954	1.008	1.046	0.968	1.029	0.880	0.912	0.949
007300	Butter	2.684	1.349**	1.257	1.126	1.126	1.182	1.471**	1.447**	1.290*	1.201	1.143	1.178
007400	Margarine	2.547	1.658*	1.117	1.108	1.108	1.132	1.057	1.100	1.132	1.149	1.057	1.095
007800	Olive oil	2.563	1.765***	1.124	1.324**	1.324**	1.142	1.248**	1.147*	1.150	1.090	1.040	1.069
007600	Pure vegetable oil	2.396	1.573**	1.266**	1.257**	1.257**	1.219***	1.210***	1.265**	1.317**	1.168***	1.150***	1.186***
008200	Tangerines	7.741	1.164	0.845	0.854	0.855	0.826	0.841	0.861	0.828	0.865	0.676*	0.686
008400	Strawberries	8.532	1.337*	1.155	1.051	1.082	1.056	1.064	1.132	1.169	1.138	0.911	0.948

Continued on next page

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
008440	Berries	4.836	1.381***	1.035	0.963	0.963	0.956	1.002	1.091	1.054	1.083	0.919	0.918
008500	Apples	2.288	1.217	0.774*	0.804*	0.804*	0.793*	0.979	0.971	0.887	0.809	0.753**	0.762*
008600	Pears	5.395	1.695**	1.132	1.171	1.171	1.118	1.193	1.207	1.178	1.204	1.076	1.124
008800	Peaches/nectarines	10.765	1.213	0.793*	0.857	0.816	0.851	0.958	0.885	0.927	0.922	0.767*	0.753*
008900	Grapes	5.963	2.086***	0.931	1.460***	1.467***	1.203	1.343*	1.178	1.207	1.225	1.083	1.122
009000	Musk melons/cantaloupes	5.700	2.268***	1.352*	1.680***	1.688***	1.273*	1.303**	1.403**	1.352*	1.388**	1.236*	1.332**
009200	Bananas	3.622	2.524	1.128	1.259	1.257**	1.263**	1.385*	1.273**	1.377*	1.458***	0.971	1.013
009300	Kiwi	4.382	1.760**	1.248*	1.234	1.234	1.199	1.211	1.198	1.186	1.268	1.140	1.191
009400	Oranges	4.601	1.441**	1.121	1.164	1.164	1.102	1.103	1.211	1.154	1.171	1.005	1.046
009500	Lemons	4.971	1.701***	1.375***	1.394**	1.394**	1.174	1.220	1.340*	1.226	1.360*	1.115	1.191
010490	Avocado	4.578	2.043***	1.264	1.305**	1.305**	1.340***	1.312	1.366*	1.222	1.349**	1.102	1.147
007900	Dried fruit mix with nuts	1.753	1.447**	0.848	0.852	0.852	0.881	1.070	0.970	0.939	0.897	0.844	0.854
008000	Peanuts salted	2.598	1.197	0.839	0.844	0.844	0.817	1.081	0.940	0.920	0.846	0.808	0.826
008100	Raisins	2.559	1.923**	1.427	1.305	1.305	1.469	1.405*	1.336*	1.322	1.422	1.245	1.325
010500	Champignons	2.909	1.392	0.983	0.910	0.910	0.897	1.063	1.069	1.006	0.933	0.902	0.916
010600	Cucumbers	5.980	1.789***	1.171	1.226***	1.226***	0.959	1.051	1.047	1.011	1.041	0.974	0.993
010700	Cauliflower	4.337	2.333**	1.441*	1.413**	1.413**	1.393*	1.602**	1.608**	1.564*	1.512**	1.287	1.390*
010800	Carrots	3.848	1.910**	1.264*	1.409**	1.409**	1.158	1.200*	1.218**	1.237**	1.220**	1.193*	1.236**
010900	Iceberg lettuce	4.509	2.023***	1.531**	1.432**	1.426**	1.332**	1.341**	1.481**	1.483**	1.473**	1.205	1.263
011000	Paprika	5.789	1.496**	1.143	1.182	1.182	1.023	1.138	1.024	1.049	1.207	1.034	1.070
011220	Packed salad	1.758	1.668***	0.916	0.903	0.903	0.988	1.005	1.025	0.989	1.025	0.916	0.917
011400	Tomatoes	2.786	1.791***	1.205	1.227	1.227	1.160	1.031	1.093	1.142	1.141	1.040	1.093
011700	Onions	3.380	1.722***	1.304**	1.281**	1.281**	1.338**	1.324**	1.250*	1.274**	1.326**	1.215**	1.260**
009900	Mixed vegetables, deep frozen	3.459	1.574***	0.816	0.767	0.767	0.880	0.802	0.847	0.810	0.862	0.770*	0.777
010000	Spinach, deep frozen	4.846	1.615***	0.924	1.020	1.020	1.015	0.905	0.982	0.944	0.988	0.912	0.921

Continued on next page

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
009700	Pickled cucumbers	2.186	2.071**	1.058	0.941	0.941	1.090	1.295	1.078	1.029	1.130	0.888	0.938
010100	French fries, deep frozen	2.117	1.394**	1.115	1.080	1.080	1.029	1.151	1.176	1.169	1.029	1.009	1.052
011800	Potatoes	2.137	1.391*	1.094	1.017	1.017	1.375	1.457**	1.333*	1.357	1.379*	1.016	1.093
010200	Potatoe chips	4.018	1.581**	0.967	0.889	0.889	0.939	1.017	0.983	0.950	0.976	0.867	0.885
013200	Crystal sugar	5.307	1.769*	1.589	1.256	1.243	1.288	1.337	1.251	1.642*	1.281	1.208	1.250
012300	Tinned peaches	3.572	1.113	0.793	0.770	0.770	0.758	0.751	0.767	0.778	0.784	0.755	0.756
013000	Jam	1.768	1.631**	1.273	1.189	1.189	1.129	1.215	1.191	1.182	1.170	1.049	1.124
013100	Natural honey	1.792	1.557**	0.963	0.877	0.877	0.927	1.144	1.115	1.079	0.967	0.834	0.854
012500	Milk chocolate	4.409	1.696	0.906	0.934	0.934	1.008	1.083*	1.038	1.092	1.068	0.924	0.917
012550	Chocolate box	3.427	1.872**	1.185**	1.183*	1.183*	1.215**	1.227*	1.163	1.183	1.290*	1.108	1.147
012600	Chocolate bar	2.402	1.467**	0.863	0.754	0.754	0.624*	0.735**	0.693*	0.680	0.851	0.672**	0.662*
012700	Chewing gum	1.157	1.350*	1.176	1.127	1.127	1.052	1.135	1.118	1.139	1.007	1.011	1.047
012800	Fruit drops/ -jelly	1.740	1.316**	1.016	0.987	0.987	0.942	1.024	0.991	1.036	0.918	0.931	0.948
012200	Ice cream, family size	2.736	1.458**	1.008	0.966	0.966	0.959	1.007	1.035	1.036	0.996	0.899	0.929
013900	Mustard	2.424	1.891***	1.349	1.253	1.253	1.195*	1.393**	1.379**	1.413**	1.234**	1.141	1.196
014000	Ketchup	2.758	1.959***	1.241	1.129	1.129	1.193	1.456***	1.448**	1.332**	1.234	1.119	1.179
014100	Vinegar	1.860	1.991***	0.893	0.823	0.823	1.068	1.133	1.011	0.918	1.123	0.861*	0.870
013700	Salt	1.666	1.509*	1.201	1.145	1.145	1.173	1.126	1.115	1.234	1.132	1.033	1.128
013800	Red pepper	1.991	1.160	0.885	0.922	0.922	0.890	0.859	0.844	0.859	0.849	0.837	0.846
014103	Baby food (milk)	1.052	1.470*	1.191	1.074	1.074	1.093	1.145	1.054	1.107	1.049	1.015	1.069
014101	Convenience food, chilled	1.702	1.102	0.915	0.887	0.887	0.884	0.988	0.957	0.923	0.916	0.873	0.897
013400	Soup powder	2.396	1.093	0.871	0.775	0.775	0.749*	0.868	0.789	0.834	0.777	0.714*	0.760*
014300	Coffee	3.460	1.610**	0.922	0.842	0.842	0.934	1.190	1.054	1.088	0.968	0.898	0.908
014301	Coffeepads /-caps	2.527	2.404***	1.085	1.189	1.189	1.319**	1.156	1.289**	1.320*	1.379**	1.098	1.135
014302	Instant coffee	5.558	1.491***	0.738**	0.682***	0.682***	0.845**	0.901	0.972	1.028	0.883	0.788**	0.780**
014200	Tea in bags	1.713	1.771***	1.173	1.225	1.225	1.212	1.188	1.229	1.158	1.155	1.089	1.151

Continued on next page

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
014500	Cocoa instant drink	2.080	2.280***	1.315	1.189	1.189	1.333	1.291	1.211	1.075	1.352	1.116	1.182
014800	Mineral or table water	2.667	1.637**	1.086	0.992	0.992	0.993	0.971	1.000	1.005	1.019	0.944	0.958
014600	Mineral water, flavoured	3.229	1.406**	0.950	0.859	0.859	0.887	0.958	0.936	0.932	0.912	0.857	0.878
014700	Energy drink	4.124	2.150***	0.902	0.913	0.913	1.182	1.144	1.123	1.062	1.238*	0.960	0.977
014900	Soft drink carbonated	1.954	1.729***	1.339	1.138	1.138	1.122	1.115	1.202	1.307	1.140	1.041	1.079
015000	Cola	3.889	1.922***	1.074	1.152	1.152	1.137	0.994	1.040	1.072	1.155	1.026	1.045
015200	Orange juice	2.120	1.451***	1.241*	1.169	1.169	1.118	1.229	1.199	1.244*	1.077	1.051	1.112
015300	Apple juice	1.442	1.521**	1.093	1.055	1.055	1.050	1.072	1.148	1.096	1.071	0.999	1.029

Note: This table presents the root mean squared errors of one-step ahead forecasts of the elementary indices. The first column contains the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.1.2 Direct forecasts of indices at level 5

Table A3: Direct one-step ahead forecasts of indices at level 5

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
01111	Rice	2.066	2.197**	1.286	1.149	1.149	1.246*	1.541**	1.421*	1.550***	1.293*	1.057	1.110
01112	Flours and other cereals	2.958	1.831**	1.361**	1.348**	1.348**	1.301**	1.334**	1.300*	1.307*	1.301**	1.232*	1.282*
01113	Bread	0.813	1.005	1.048	0.955	0.955	1.006	1.151	0.983	0.951	0.929	0.881	0.936
01114	Other bakery products	1.248	1.455**	1.188	1.041	1.041	1.079	1.093	0.998	1.054	1.004	0.956	0.991
01115	Pizza and quiche	3.032	1.683***	1.033	0.966	0.966	1.012	1.084	1.068	1.060	1.049	0.926	0.938
01116	Pasta products and couscous	2.133	1.819***	1.161	1.030	1.030	1.137	0.894	1.136	1.069	1.211**	1.001	1.022
01117	Breakfast cereals	1.698	1.824*	0.952	1.061	1.061	1.035	1.170	1.092	1.053	1.082	0.896	0.926
01118	Other cereal products	1.671	2.011***	1.555**	1.414	1.414	1.372	1.295	1.243	1.329**	1.260	1.210	1.274
01121	Beef and veal	1.898	1.672**	1.185	1.071	1.071	1.001	1.357	1.223	1.222	1.033	0.965	1.027
01122	Pork	2.850	1.239	1.012	0.928	0.928	0.859	1.194	1.057	1.059	0.890	0.868	0.913
01124	Poultry	1.277	1.107	1.322**	1.106	1.106	0.953	1.069	0.932	1.203	0.805	0.834	0.924
01127	Dried, salted or smoked meat	0.860	1.540**	1.565*	1.456	1.456	1.356	1.565**	1.429*	1.354	1.278	1.262	1.374
01128	Other meat preparations	2.635	1.476**	1.263	1.176	1.176	1.077	1.093	1.106	1.027	1.103	1.024	1.063
01131	Fresh or chilled fish	3.925	0.783	0.532**	0.502**	0.502**	0.507**	0.756	0.626**	0.598**	0.527**	0.518***	0.530**
01132	Frozen fish	4.120	1.630***	0.931	0.844	0.844	0.943	0.875	0.929	0.907	0.977	0.853	0.856
01135	Dried, smoked or salted fish and seafood	3.328	1.859**	1.076	1.130	1.130	1.143	1.038	1.038	1.029	1.068	1.004	1.027
01136	Other preserved or processed fish and seafood-based preparations	2.791	1.508**	0.836*	0.784**	0.784**	0.880	1.053	1.044	0.942	0.930	0.857*	0.860

Continued on next page

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
01141	Whole Milk	1.317	2.433**	1.883	1.783	1.783	1.849*	1.944*	1.716	1.912*	1.813*	1.655	1.795
01143	Preserved milk	2.346	1.868	1.359**	1.205	1.205	0.963	1.338	1.174	1.348	0.981	0.904	0.981
01144	Yoghurt	1.488	1.784***	1.449**	1.266	1.266	1.279	1.377**	1.318**	1.316*	1.273	1.210	1.271
01145	Cheese and curd	1.207	1.871**	1.534**	1.460**	1.460**	1.378*	1.742**	1.542**	1.508**	1.378*	1.299	1.400*
01146	Other milk products	1.425	1.968**	1.410*	1.448	1.448	1.437	1.570*	1.443	1.846*	1.467*	1.343	1.424
01147	Eggs	2.085	1.141	1.064	0.953	0.953	1.009	1.041	0.967	1.028	0.878	0.913	0.950
01151	Butter	2.686	1.349**	1.248	1.127	1.127	1.186	1.467**	1.451**	1.333**	1.202	1.145	1.180
01152	Margarine and other vegetable fats	2.554	1.655*	1.139	1.107	1.107	1.130	1.057	1.095	1.125	1.147	1.052	1.090
01153	Olive oil	2.571	1.764***	1.124	1.320**	1.320**	1.142	1.248**	1.148*	1.151	1.090	1.039	1.068
01154	Other edible oils	2.395	1.576**	1.320**	1.261**	1.261**	1.219***	1.214***	1.267**	1.294**	1.169***	1.146***	1.180***
01161	Fresh or chilled fruit	7.751	0.389**	0.203**	0.229**	0.229**	0.246**	0.260**	0.257**	0.244**	0.263**	0.269**	0.221**
01163	Dried fruit and nuts	1.501	1.585**	0.884	0.919	0.919	1.037	1.065	1.047	1.089	1.031	0.912	0.930
01171	Fresh or chilled vegetables other than potatoes and other tubers	1.770	1.977***	1.492***	1.451**	1.451**	1.348**	1.365**	1.286*	1.391**	1.425***	1.257**	1.331***
01172	Frozen Vegetables other tahn potatoes and other tubers	2.750	1.230	0.849	0.850	0.850	0.839*	0.742**	0.849	0.832*	0.873	0.792**	0.807**
01173	Dried vegetables, other preserved or processed vegetables	2.189	2.063**	1.069	0.937	0.937	1.085	1.299*	1.077	1.007	1.124	0.884	0.933
01174	Potatoes	1.741	1.479*	1.154	1.077	1.077	1.449*	1.526**	1.412**	1.468*	1.470**	1.100	1.175
01175	Crisps	4.019	1.580**	0.964	0.887	0.887	0.938	1.019	0.983	0.959	0.975	0.868	0.885
01181	Sugar	5.324	1.771*	1.664*	1.255	1.240	1.285	1.339	1.253	1.654*	1.255	1.206	1.249
01182	Jams, marmalades and honey	1.750	1.186	0.858	0.760	0.760	0.764	0.923	0.896	0.879	0.785	0.743	0.779
01183	Chocolate	3.174	1.637	1.172	1.016	1.016	1.043	1.104	1.028	1.133	1.104	0.971	0.981

Continued on next page

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
01184	Confectionery products	1.324	1.368**	1.053	1.018	1.018	0.984	1.070	1.030	1.037	0.943	0.949	0.971
01185	Edible ices and ice cream	2.738	1.456**	1.008	0.966	0.966	0.959	1.018	1.044	1.027	0.996	0.895	0.925
01191	Sauces, condiments	1.546	1.504*	1.214	1.068	1.068	1.025	1.296*	1.284*	1.203	1.034	1.011	1.058
01192	Salt, spices and culinary herbs	1.697	1.073	0.945	0.922	0.922	0.948	0.830	0.816	0.824	0.809	0.822	0.841
01193	Baby food	1.057	1.464*	1.105	1.054	1.054	1.073	1.164	1.060	1.114	1.040	0.999	1.050
01194	Ready-made meals	1.703	1.115	0.910	0.893	0.893	0.888	0.992	0.959	0.954	0.914	0.876	0.900
01199	Other food products n.e.c.	2.393	1.100	0.872	0.776	0.776	0.750*	0.869	0.794	0.803	0.778	0.710*	0.756*
01211	Coffee	2.799	1.909***	0.972	0.941	0.941	1.085	1.157	1.089	1.131	1.097	0.965	0.990
01212	Tea	1.711	1.762***	1.173	1.222	1.222	1.209	1.190	1.225	1.167	1.149	1.087	1.148
01213	Cocoa and powdered chocolate	2.073	2.287***	1.314	1.190	1.190	1.326	1.293	1.214	1.094	1.354	1.117	1.184
01221	Mineral or spring waters	2.654	1.635**	1.081	0.994	0.994	0.994	0.976	1.004	0.999	1.019	0.942	0.956
01222	Softdrinks	2.305	1.972***	1.207	1.205	1.205	1.202*	1.094	1.138	1.127	1.192*	1.101	1.131
01223	Fruit and vegetable juices	1.595	1.573***	1.260*	1.156	1.156	1.123	1.236*	1.258**	1.258**	1.115	1.084	1.141

Note: This table presents the root mean squared errors of one-step ahead forecasts of the indices. The first column contains the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.1.3 Indirect forecasts of indices at level 5

Table A4: Indirect one-step ahead forecasts of indices at level 5

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
01111	Rice	2.066	4.207***	1.305**	1.481***	1.481***	2.204**	2.401***	2.288**	1.984**	2.323**	1.537**	1.649**
01112	Flours and other cereals	2.958	3.169***	1.802**	1.821***	1.821***	1.902**	1.841***	1.891***	1.864***	1.872***	1.669**	1.801**
01113	Bread	0.813	1.801*	1.120	1.017	1.017	1.107	1.004	1.073	1.205	1.074	0.961	1.040
01114	Other bakery products	1.248	2.620***	1.285*	1.337*	1.337*	1.419**	1.418**	1.471***	1.496**	1.527***	1.256**	1.329**
01115	Pizza and quiche	3.032	3.076***	1.188*	1.345**	1.345**	1.700***	1.863***	1.874***	1.496**	1.787***	1.373***	1.455***
01116	Pasta products and couscous	2.133	3.434***	1.500***	1.461***	1.461***	1.955***	1.449**	1.830***	1.634***	1.943***	1.498***	1.582***
01117	Breakfast cereals	1.698	3.913**	1.133	1.404	1.404	1.860**	1.907***	1.917***	1.863***	1.993**	1.338	1.430*
01118	Other cereal products	1.671	3.976***	1.951**	1.778**	1.778**	2.220***	1.896**	2.062***	1.998***	2.151***	1.703**	1.869**
01121	Beef and veal	1.898	3.439**	1.151	1.245	1.245	1.686**	2.421***	2.176**	1.711**	1.852**	1.382	1.499*
01122	Pork	2.850	2.272***	1.068	1.145	1.145	1.236	1.802***	1.593**	1.428*	1.406	1.127	1.230
01124	Poultry	1.277	2.548	0.840	0.801	0.801	1.116	1.365	1.344	1.319	1.188	0.900	0.980
01127	Dried, salted or smoked meat	0.860	2.935***	1.643***	1.415**	1.415**	1.608**	1.787***	1.713***	1.584**	1.639**	1.390**	1.508**
01128	Other meat preparations	2.635	2.653***	1.264*	1.312**	1.312**	1.494**	1.519**	1.667**	1.365*	1.572**	1.177	1.241*
01131	Fresh or chilled fish	3.925	1.427*	0.784	0.764	0.764	0.792	1.027	0.923	0.848	0.830	0.724*	0.798
01132	Frozen fish	4.120	3.062***	1.332**	1.260**	1.260**	1.625***	1.460***	1.628***	1.505***	1.703***	1.339***	1.403***
01135	Dried, smoked or salted fish and seafood	3.328	3.663**	1.617*	1.676*	1.676*	1.883**	1.663*	1.773*	1.653*	1.810*	1.518*	1.639*
01136	Other preserved or processed fish and seafood-based preperations	2.791	2.799**	1.247	1.249	1.249	1.517**	1.770**	1.814**	1.625**	1.590**	1.338*	1.430*

Continued on next page

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
01141	Whole Milk	1.317	4.249**	2.645***	2.468**	2.468**	2.720**	2.439**	2.350**	2.874***	2.587**	2.295**	2.509**
01143	Preserved milk	2.346	4.489	1.211	1.477	1.473	1.890	2.304*	2.244*	1.804**	2.022	1.302	1.404
01144	Yoghurt	1.488	3.071***	1.867***	1.817**	1.817**	1.895***	1.835***	1.855***	1.864**	1.929**	1.688**	1.828**
01145	Cheese and curd	1.207	3.591**	1.694**	1.652**	1.652**	1.911**	2.262**	2.147**	2.117**	2.017**	1.696**	1.841**
01146	Other milk products	1.425	3.558**	1.915**	1.947**	1.947**	2.029**	1.938**	2.013**	2.162**	2.112**	1.777**	1.935**
01147	Eggs	2.085	2.081**	1.269	1.134	1.134	1.380	1.350	1.299	1.462	1.216	1.152	1.239
01151	Butter	2.686	2.274***	1.849***	1.690**	1.690**	1.666**	1.634***	1.559***	1.914***	1.662**	1.547**	1.696***
01152	Margarine and other vegetable fats	2.554	2.840**	1.650*	1.583*	1.583*	1.743*	1.588	1.731*	1.724*	1.751*	1.514*	1.641*
01153	Olive oil	2.571	3.364***	1.321***	1.254**	1.254**	1.833***	1.946***	1.979***	1.744***	1.867***	1.437***	1.523***
01154	Other edible oils	2.395	2.702***	1.553**	1.568**	1.568**	1.704**	1.517***	1.791**	1.775**	1.679**	1.480**	1.593**
01161	Fresh or chilled fruit	7.751	1.001	0.393*	0.540	0.599	0.523	0.544	0.532	0.524	0.560	0.485*	0.495*
01163	Dried fruit and nuts	1.501	2.855***	1.359**	1.402**	1.402**	1.625**	1.851***	1.824***	1.806***	1.626***	1.435**	1.532***
01171	Fresh or chilled vegetables other than potatoes and other tubers	1.770	3.596***	1.734***	1.944***	1.941***	1.735***	1.835***	1.779***	1.897***	1.903***	1.636***	1.776***
01172	Frozen Vegetables other tahn potatoes and other tubers	2.750	2.075***	1.206	1.117	1.117	1.258	1.123	1.306*	1.216	1.257	1.077	1.137
01173	Dried vegetables, other preserved or processed vegetables	2.189	4.376**	1.328*	1.393**	1.393**	2.085**	2.415***	2.045**	1.750***	2.192**	1.511**	1.641**
01174	Potatoes	1.741	2.788**	1.680*	1.677**	1.677**	1.702**	1.919**	1.894**	1.950**	1.844**	1.471	1.626*
01175	Crisps	4.019	2.951**	1.219	1.294*	1.294*	1.591**	1.673**	1.650***	1.486**	1.681**	1.312**	1.399**
01181	Sugar	5.324	3.418*	2.321*	1.512*	1.619	2.168	1.945*	1.832*	2.313	2.071	1.747	1.877
01182	Jams, marmalades and honey	1.750	2.259***	1.052	1.055	1.055	1.205	1.438**	1.462**	1.355*	1.265	1.114	1.174
01183	Chocolate	3.174	2.865	1.544*	1.675	1.675	1.641	1.760**	1.716*	2.004	1.738	1.538	1.650

Continued on next page

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
01184	Confectionery products	1.324	2.480***	1.308**	1.211	1.211	1.404**	1.512**	1.562**	1.551**	1.451**	1.259*	1.334**
01185	Edible ices and ice cream	2.738	2.627***	1.320***	1.224**	1.224**	1.468**	1.617***	1.693***	1.471***	1.540**	1.268**	1.350**
01191	Sauces, condiments	1.546	2.702***	1.388*	1.394	1.394	1.516*	1.934***	1.896***	1.833**	1.586**	1.370*	1.480*
01192	Salt, spices and culinary herbs	1.697	1.730***	1.100	1.110	1.110	1.154	1.157	1.137	1.107	1.122	1.038	1.068
01193	Baby food	1.057	2.654**	1.724*	1.445	1.445	1.567*	1.566	1.497	1.611	1.551*	1.415	1.521
01194	Ready-made meals	1.703	1.737*	1.195	1.162	1.162	1.127	1.161	1.249	1.219	1.248	1.068	1.165
01199	Other food products n.e.c.	2.393	1.956**	0.978	1.035	1.035	1.104	1.375	1.205	1.295	1.170	0.995	1.082
01211	Coffee	2.799	3.548***	1.263**	1.429**	1.429**	1.869***	1.961***	1.908***	1.845***	1.958***	1.527***	1.628***
01212	Tea	1.711	3.245***	1.558**	1.588***	1.588***	1.784***	1.884***	1.968***	1.852***	1.823***	1.544**	1.682***
01213	Cocoa and powdered chocolate	2.073	4.395***	1.738**	1.643**	1.643**	2.312***	2.241***	2.122***	1.759***	2.426***	1.756**	1.909***
01221	Mineral or spring waters	2.654	3.054***	1.463**	1.389**	1.389**	1.671**	1.608**	1.699**	1.737**	1.749**	1.437**	1.523**
01222	Softdrinks	2.305	3.608***	1.516***	1.636***	1.636***	1.969***	1.728***	1.876***	1.800***	2.089***	1.577***	1.680***
01223	Fruit and vegetable juices	1.595	2.712***	1.607***	1.474***	1.474***	1.599***	1.890***	1.931***	1.900***	1.644***	1.471***	1.594***

Note: This table presents the root mean squared errors of one-step ahead forecasts of the indices. The first column contains the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.1.4 Direct forecasts of indices at level 4

Table A5: Direct one-step ahead forecasts of indices at level 4

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
0111	Bread and cereals	0.709	1.482**	1.322*	1.416**	1.416**	1.290*	1.369*	1.150	1.136	1.122	1.017	1.082
0112	Meat	0.948	1.578***	1.680	1.564	1.564	1.308	1.617*	1.383	1.379	1.263	1.292	1.400
0113	Fish	1.807	1.525**	1.076	0.946	0.946	1.004	1.093	1.039	1.035	0.997	0.940	0.980
0114	Milk, cheese and eggs	0.900	2.062**	1.802**	1.700*	1.700*	1.690*	1.942**	1.728**	1.688***	1.628*	1.491	1.616*
0115	Oil and fats	1.651	1.413**	1.255**	1.151	1.151	1.215**	1.399***	1.448**	1.305**	1.232**	1.163**	1.197**
0116	Fruit	6.194	0.422*	0.200**	0.245**	0.245**	0.256**	0.265**	0.265**	0.255**	0.262**	0.274**	0.229**
0117	Vegetables	1.237	2.065***	1.403**	1.410*	1.410*	1.364**	1.402*	1.241	1.424*	1.341**	1.211*	1.286**
0118	Sugar, jam, honey, chocolate and confectionery	1.437	1.436	1.110	1.019	1.019	0.999	1.131	1.052	1.084	1.002	0.950	0.980
0119	Food products n.e.c.	1.058	0.946	0.879	0.824	0.824	0.826	0.986	0.870	0.859	0.754	0.781	0.810
0121	Coffee, tea and cocoa	2.286	2.046***	1.089	1.013	1.013	1.161**	1.231**	1.163	1.223**	1.186**	1.028	1.064
0122	Mineral waters, soft drinks, fruit and vegetables juices	1.397	2.005***	1.341*	1.396*	1.396*	1.326**	1.144	1.223*	1.222*	1.290**	1.184	1.238

Note: This table presents the root mean squared errors of one-step ahead forecasts of the indices. The first column contains the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.1.5 Indirect forecasts of indices at level 4

Table A6: Indirect one-step ahead forecasts of indices at level 4

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
0111	Bread and cereals	0.709	3.018**	0.992	1.116	1.116	1.509**	1.419**	1.543**	1.442**	1.531**	1.164	1.232
0112	Meat	0.948	3.274**	1.438**	1.424**	1.424**	1.555***	1.938***	1.762**	1.481**	1.710**	1.331**	1.456**
0113	Fish	1.807	2.893***	1.351	1.320	1.320	1.577**	1.566**	1.764***	1.611**	1.690**	1.343*	1.445*
0114	Milk, cheese and eggs	0.900	4.014**	1.896**	1.897**	1.897**	2.133**	2.105***	2.133**	2.239**	2.207**	1.812**	1.971**
0115	Oil and fats	1.651	2.466***	1.677***	1.586**	1.586**	1.637***	1.615***	1.702***	1.975***	1.663***	1.503**	1.642***
0116	Fruit	6.194	1.058	0.415*	0.571	0.630	0.546	0.568	0.559	0.547	0.580	0.506*	0.520
0117	Vegetables	1.237	4.140**	1.576**	1.880***	1.879***	1.860**	2.045***	1.889***	1.955***	2.042**	1.652**	1.798**
0118	Sugar, jam, honey, chocolate and confectionery	1.437	2.643**	1.528*	1.425	1.423	1.463	1.743***	1.646**	1.936*	1.548	1.391	1.491
0119	Food products n.e.c.	1.058	1.648**	0.929	0.901	0.901	0.949	1.220	1.203	1.157	1.000	0.921	0.968
0121	Coffee, tea and cocoa	2.286	3.828***	1.356**	1.542**	1.542**	2.012***	2.094***	2.050***	1.969***	2.088***	1.633***	1.751***
0122	Mineral waters, soft drinks, fruit and vegetables juices	1.397	3.611***	1.623**	1.679***	1.679***	2.019***	1.693***	1.869***	1.829***	2.136***	1.616***	1.729***

Note: This table presents the root mean squared errors of one-step ahead forecasts of the indices. The first column contains the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.1.6 Direct forecasts of indices at level 3

Table A7: Direct one-step ahead forecasts of indices at level 3

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
011	Food	0.809	1.219	1.290	1.236	1.236	0.989	1.246	0.984	0.985	0.924	0.934	1.009
012	Non-alcoholic beverages	1.123	2.143***	1.497*	1.465*	1.465*	1.450**	1.452**	1.393**	1.367**	1.351**	1.284*	1.361*

Note: This table presents the root mean squared errors of one-step ahead forecasts of the indices. The first column contains the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.1.7 Indirect forecasts of indices at level 3

Table A8: Indirect one-step ahead forecasts of indices at level 3

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
011	Food	0.809	3.073	0.930	1.089	1.077	1.264	1.268	1.358	1.242	1.317	1.006	1.100
012	Non-alcoholic beverages	1.123	3.874***	1.688***	1.788***	1.788***	2.125***	2.274***	2.332***	2.268***	2.333***	1.809***	1.956***

Note: This table presents the root mean squared errors of one-step ahead forecasts of the indices. The first column contains the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.1.8 Direct forecasts of indices at level 2

Table A9: Direct one-step ahead forecasts of indices at level 2

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
01	FOOD AND NON-ALCOHOLIC BEVERAGES	0.764	1.267	1.311	1.277	1.277	1.005	1.257	0.989	1.022	0.939	0.934	1.014

Note: This table presents the root mean squared errors of one-step ahead forecasts of the indices. The first column contains the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.1.9 Indirect forecasts of indices at level 2

Table A10: Indirect one-step ahead forecasts of indices at level 2

COICOP	Text	Web-scraping	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble (w WS)	Ensemble (wo WS)
01	FOOD AND NON-ALCOHOLIC BEVERAGES	0.764	3.176	0.948	1.132	1.131	1.332	1.329	1.461	1.292	1.404	1.049	1.149

Note: This table presents the root mean squared errors of one-step ahead forecasts of the indices. The first column contains the RMSE in levels for the webscraping based nowcast, the other columns contain ratios of the time series models' RMSEs relative to the first column. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.2 Three-step ahead forecasts

B.2.1 Elementary indices (level 6)

Table A11: Three-step ahead forecasts of elementary indices (level 6)

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
001900	Long grain rice	0.406**	0.785**	0.831*	0.831*	0.835*	0.975	0.922	0.702**	0.807*	0.901
002000	Wheat flour	0.471***	0.848**	0.896*	0.896*	0.870**	0.842***	0.815**	0.805***	0.879**	0.877**
000600	Rye bread	1.192	1.009	1.028	1.028	1.080	1.097	1.233	1.094	1.182	1.145
000700	Wholemeal bread	0.888	0.858*	0.903	0.903	0.939	0.918	1.032	1.089	1.030	0.989
000800	White bread	0.787	0.911	0.957	0.957	0.975	0.896	1.023	0.967	1.027	1.001
000900	Bread roll	0.522**	0.761***	0.823**	0.823**	0.727***	0.795***	0.732***	0.693***	0.778***	0.780***
001000	Scone, handmade	1.037	0.834	0.906	0.906	0.842	0.873	0.961	0.904	0.981	0.931
001100	Half baked buns	0.894	1.013	1.049	1.049	1.122	0.984	1.011	1.080	1.079	1.090
001110	Lye bun	0.409***	1.006	0.949	0.949	0.839	0.999	0.918	0.945	0.856	0.951
000200	Yeast dumpling, deep frozen	0.373***	1.074	1.046	1.046	0.978	1.067	1.002	1.000	0.933**	1.041
001500	Butter biscuits	0.593**	0.917	0.966	0.966	1.029	1.051	1.048	1.043	1.034	1.099
001550	Cake	0.850	0.821**	0.809*	0.809*	0.947	0.859	0.940	0.944	0.940	0.922
001600	Wafers with hazelnut cream	0.598*	0.830**	0.898	0.898	0.914	0.933	0.901	0.920	0.920	0.945
001700	Salty sticks	0.542***	0.947	1.012	1.012	1.035	1.075	0.995	1.070	0.923	1.032
002300	Nut cake	1.467**	1.100	1.112	1.112	1.181	1.250*	1.519***	1.216*	1.460**	1.341**
002400	Curd cheese cake	0.333***	0.932	1.000	1.000	0.965	0.919*	0.915**	0.926**	0.924*	0.976
000100	Pizza, deep frozen	0.442***	0.905*	0.959	0.959	0.903	0.899	0.819**	0.863	0.883	0.953
000300	Convenience food, deep frozen	0.614*	0.816**	0.958	0.958	0.907	0.897	0.912	0.837	0.955	0.920
002100	Pasta	0.381**	0.921*	0.959	0.959	0.913*	1.009	0.922	1.024	0.908	0.980
001400	Cereals	0.327**	0.893	0.963	0.963	0.932	0.993	0.982	0.977	0.985	1.025
001200	Ready-made dough	0.518**	0.817**	0.849**	0.849**	0.832	0.854**	0.841	0.848**	0.866	0.885

Continued on next page

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
002500	Veal cutlet	0.794	1.102	1.281**	1.281**	1.285*	1.250*	1.240	1.301**	1.279*	1.291**
003100	Roast beef	0.745	0.938	0.970	0.970	0.978	0.976	0.954	0.815*	1.009	0.965
003200	Beef round	0.639**	1.082	1.060	1.060	1.017	0.988	0.989	0.936	0.979	1.039
003300	Beef shoulder	0.922	0.939	0.924	0.924	0.995	0.922	0.977	0.801	1.026	0.966
003500	Pork belly	0.529*	0.985	1.018	1.018	1.060	1.029	1.047	1.076	0.899	1.095
003600	Pork chops	0.561**	1.012	1.080	1.080	1.075	1.045	1.038	1.080	1.034	1.090
003700	Pork cutlet	0.891	1.019	1.089	1.089	1.145	1.087	1.149	1.148	1.121	1.120
003800	Pork sirloin	1.581	1.374	1.260	1.260	1.472*	1.373*	1.402	1.496*	1.528*	1.457*
004700	Roast chicken	1.026	0.899	0.977	0.977	1.051	0.972	0.988	1.017	1.127	1.040
004800	Turkey breast	1.016	0.848	0.905	0.905	1.021	0.888	0.985	0.952	1.005	0.966
004900	Breaded chicken meat, deep frozen	0.665	0.872*	0.962	0.962	0.998	0.946	0.916	0.793	1.038	1.007
005000	Chicken breast	0.808	0.890	0.887	0.887	0.933	0.930	0.870	0.967	0.955	0.941
002700	Smoked meat	1.078	0.981	0.993	0.993	1.088	1.037	1.131	1.272	1.149	1.127
002800	Bacon	0.761	0.899	0.924	0.924	1.167	0.896	0.929	1.086	0.953	0.989
004000	Sausages	0.557**	0.822**	0.867*	0.867*	0.869	0.824*	0.829	0.899	0.859	0.875
004100	Dry sausage	0.594	0.757*	0.768	0.768	0.747*	0.782**	0.760*	0.766*	0.747*	0.773
004200	Pork ham	0.697**	0.855*	0.868	0.868	0.885	0.841	0.863	0.890	0.922	0.906
004300	Turkey sausage	0.999	0.909	0.858	0.858	0.930	0.927	0.955	0.840*	0.985	0.938
004400	Pork sausage	0.527***	0.918	0.889	0.889	0.895	0.935	0.865	0.835	0.889	0.897
004600	Salami	0.540*	0.917	0.862	0.862	0.883	0.893	0.922	0.937	0.879	0.938
002600	Minced meat	0.558**	0.921	0.896	0.896	0.916	0.884	0.843*	0.801*	0.883	0.893
002900	Liver pasty	0.313**	0.857*	0.865*	0.865*	0.828**	0.817**	0.832*	0.878	0.805***	0.854*
003000	Beef goulash canned	0.675	0.746**	0.778**	0.778**	0.772**	0.833*	0.751**	0.732**	0.793**	0.797**
005400	Fresh fish	1.720**	1.283*	1.507**	1.507**	1.567**	1.019	1.394**	1.542***	1.603**	1.585**
005200	Codfish filet, deep frozen	0.512**	0.936	0.960	0.960	0.958	0.993	0.939	0.866*	0.929	0.952
005300	Fish fingers, deep frozen	0.356***	0.961	0.980	0.980	0.965	1.050	1.001	0.899	0.916	1.027

Continued on next page

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
005500	Smoked salmon	0.377**	0.863*	0.872	0.872	0.888	0.832*	0.862	0.804**	0.839	0.851
005100	Tunafish	0.525**	0.946	1.079	1.079	1.021	0.887	0.882	0.840	0.980	0.987
005660	Extended shelf life milk	0.597*	0.882	0.869	0.869	0.793*	0.833**	0.902	0.896	0.883	0.877
005661	Fresh milk	0.663	0.891	0.932	0.932	0.899	0.885	0.958	0.839*	0.930	0.921
006100	Evaporated milk	0.476	0.906	0.836*	0.837*	0.893**	0.880**	0.820*	0.918	0.915	0.959
006200	Yoghurt fruit flavoured	0.581**	0.926	0.949	0.949	0.946	0.956	0.865	0.824**	0.984	0.967
006400	Yoghurt	0.454**	0.833**	0.876*	0.876*	0.878	0.835**	0.835**	0.827*	0.838**	0.858**
006000	Curd cheese	0.653*	0.860	0.872	0.872	0.758***	0.814***	0.790**	0.701*	0.820**	0.833**
006600	Emmentaler	0.676**	0.825***	0.878**	0.878**	0.939	0.863**	0.824**	0.896	0.917	0.905
006700	Gouda	0.557***	0.798**	0.857**	0.857**	0.770*	0.848**	0.795**	0.828*	0.846*	0.841*
006800	Hard cheese	0.551**	0.816**	0.821**	0.821**	0.839**	0.847**	0.822*	0.849**	0.851**	0.858**
006900	Camembert	0.391***	0.856*	0.910	0.910	0.902*	0.868*	0.860	0.958	0.875**	0.908
007100	Fresh cream cheese	0.413*	0.894**	0.922	0.922	0.903**	0.888***	0.869**	0.895	0.934	0.933*
007200	Mozzarella	0.593**	0.784***	0.852**	0.852**	0.811***	0.822***	0.796**	0.855	0.841**	0.851**
005700	Milk shake	0.542**	0.844*	0.874	0.874	0.859	0.835	0.791**	0.820*	0.861	0.864
005800	Sour cream	0.431**	0.848*	0.835**	0.835**	0.856**	0.832**	0.818**	0.875*	0.792***	0.846**
005900	Whipped cream	0.403*	0.833*	0.847*	0.847*	0.856*	0.809**	0.780***	0.785*	0.817**	0.825**
006500	Eggs	0.687	0.813	0.846	0.846	0.728	0.823	0.831	0.792	0.848	0.827
007300	Butter	0.657**	0.890	0.946	0.946	0.808**	0.872**	0.817***	0.901	0.812**	0.914
007400	Margarine	0.546	0.971	0.992	0.992	0.935	0.943	0.982	0.913	0.964	0.984
007800	Olive oil	0.498***	0.914	0.850***	0.850***	0.857*	0.883***	0.977	0.807	0.980	0.973
007600	Pure vegetable oil	0.492***	0.808***	0.870***	0.870***	0.809***	0.835***	0.788***	0.693**	0.848***	0.843***
008200	Tangerines	0.821	1.239*	1.127	1.123	1.147	1.060	1.116	1.084	1.128	1.358**
008400	Strawberries	0.482**	1.012	0.911	0.874*	0.970	0.988	0.929	0.897	0.938	0.989
008440	Berries	0.491***	0.911*	1.031	1.031	0.927	0.871	0.827**	0.839*	0.833***	0.959
008500	Apples	0.868	1.190*	1.214*	1.214*	1.206	1.221*	1.064	1.369**	1.176	1.368**
008600	Pears	0.557*	0.963	0.886*	0.886*	1.023	0.959	1.004	0.953	0.922	0.946
008800	Peaches/nectarines	1.075	1.121*	1.225	1.106	1.059	1.013	1.025	0.970	1.032	1.117
008900	Grapes	0.470***	0.972	0.909*	0.910*	0.814**	0.800**	0.783**	0.783**	0.793**	0.906

Continued on next page

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
009000	Musk melons/cantaloupes	0.592**	1.043	0.923	0.919	1.011	1.024	0.936	0.980	0.994	1.026
009200	Bananas	0.300	0.943	1.007	0.876	0.919	0.979	0.875	0.793**	0.917	0.980
009300	Kiwi	0.574**	1.024	0.960	0.960	0.926	0.905	0.809**	0.895	0.869**	0.939
009400	Oranges	0.825	0.944	0.987	0.987	1.068	1.020	0.980	1.102	1.035	1.118
009500	Lemons	0.752*	0.993	0.948	0.948	0.933	0.880	0.831	1.003	0.881	1.016
010490	Avocado	0.398***	0.909	0.878**	0.878**	0.891*	0.929	0.772*	0.938	0.827*	0.889
007900	Dried fruit mix with nuts	0.571**	0.891*	0.955	0.955	0.971	0.919	0.968	1.083	0.903	0.998
008000	Peanuts salted	0.561	1.039	1.143	1.143	1.150	1.200	1.179	1.139	1.166	1.211*
008100	Raisins	0.638*	0.912*	0.926	0.926	0.951	0.894	0.962	0.952	1.043	0.963
010500	Champignons	0.674*	1.031	1.021	1.021	1.053	1.044	1.002	1.188	1.057	1.089
010600	Cucumbers	0.447**	1.041	0.971	0.971	1.214*	1.087	1.159	1.104	1.145	1.120*
010700	Cauliflower	0.607*	0.921	0.899	0.899	0.897	0.893	0.877	0.854	0.875	0.892
010800	Carrots	0.383*	0.965	0.899*	0.899*	0.969	0.983	0.943	0.905	0.925	0.942
010900	Iceberg lettuce	0.600**	0.930	1.026	1.020	1.033	0.931	0.887	0.935	0.947	1.045
011000	Paprika	0.527***	0.972	0.884**	0.884**	1.022	1.128	1.014	0.975	0.871	0.981
011220	Packed salad	0.411***	0.932	0.953	0.953	0.936	0.928	0.838**	0.887*	0.909*	0.932
011400	Tomatoes	0.602**	0.989	0.989	0.989	1.027	1.065	1.078	1.088	1.059	1.064
011700	Onions	0.451***	0.811*	0.852	0.852	0.835*	0.897*	0.804**	0.842	0.803**	0.854*
009900	Mixed vegetables, deep frozen	0.578**	1.005	1.057	1.057	1.009	1.108	1.009	1.076	1.039	1.065
010000	Spinach, deep frozen	0.510***	1.125	1.088	1.088	1.145	1.150	1.089	1.079	1.107	1.128
009700	Pickled cucumbers	0.570*	0.916	1.129	1.129	1.045	0.939	0.972	0.834	1.098	1.103
010100	French fries, deep frozen	0.643**	0.845**	0.888**	0.888**	0.895*	0.859*	0.819**	0.842*	0.882	0.887
011800	Potatoes	0.755	0.931	1.007	1.007	0.848*	0.807**	0.780**	0.852	0.801**	0.888
010200	Potatoe chips	0.619*	0.917	0.999	0.999	1.034	1.058	1.045	0.995	1.046	1.051
013200	Crystal sugar	0.456	0.805	0.915	0.916	0.911	0.887	0.897	0.608*	0.898	0.895

Continued on next page

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
012300	Tinned peaches	0.843	1.198	1.166	1.166	1.393	1.055	1.091	1.337	1.311	1.226
013000	Jam	0.835	0.960	0.968	0.968	0.993	0.897	0.953	1.045	1.000	1.001
013100	Natural honey	0.589**	0.931*	1.037	1.037	0.998	0.861	0.946	0.939	0.968	1.013
012500	Milk chocolate	0.255	0.965	1.089	1.089	0.983	0.937	0.941	1.167	0.916	1.057
012550	Chocolate box	0.422	0.958	0.988	0.988	0.933	0.988	0.945	0.960	0.906	0.969
012600	Chocolate bar	0.418**	0.992	1.082	1.082	1.031	1.145	1.046	0.977	0.990	1.140
012700	Chewing gum	0.649*	0.985	0.896	0.896	0.850*	0.892	0.852	0.891	0.929	0.937
012800	Fruit drops/ -jelly	0.575***	0.865**	0.889**	0.889**	0.894**	0.920	0.871**	0.874	0.907*	0.922
012200	Ice cream, family size	0.818	0.982	1.041	1.041	1.104	1.076	1.076	0.970	1.110	1.157
013900	Mustard	0.542**	0.872*	0.925	0.925	0.936	0.876	0.852	0.860*	0.947	0.947
014000	Ketchup	0.576**	0.888***	0.918	0.918	0.908	0.930	0.831	0.904	0.895	0.901
014100	Vinegar	0.396***	0.924	1.147*	1.147*	0.894	0.923	1.102	1.038	0.840*	1.082
013700	Salt	0.978	1.054	1.080	1.080	1.034	0.939	0.958	0.989	1.047	1.041
013800	Red pepper	0.734*	1.077	1.012	1.012	1.061	1.062	1.214*	1.114	1.205*	1.125
014103	Baby food (milk)	0.810	0.898	0.910	0.910	0.938	0.909	0.964	0.904	0.973	0.947
014101	Convenience food, chilled	0.837	0.877***	0.891**	0.891**	0.912*	0.848**	0.873	0.944	0.915	0.908*
013400	Soup powder	0.911	0.976	1.041	1.041	1.063	1.083	1.087	0.981	1.051	1.075
014300	Coffee	0.501***	0.915	0.986	0.986	0.971	0.892	0.920	0.964	0.995	0.998
014301	Coffeepads /-caps	0.396***	0.941	0.930	0.930	0.929	0.919	0.907	0.897	0.894	0.980
014302	Instant coffee	0.624***	1.059	1.220**	1.220**	1.136*	1.113	1.053	1.052	1.158*	1.170**
014200	Tea in bags	0.673**	0.887	0.893	0.893	0.913	1.018	0.991	0.993	1.022	0.961
014500	Cocoa instant drink	0.505**	0.788**	0.819*	0.819*	0.799	0.831	0.859	0.827	0.792	0.830
014800	Mineral or table water	0.444**	0.860*	0.888*	0.888*	0.892*	0.924	0.863**	0.856**	0.901	0.911*
014600	Mineral water, flavoured	0.674*	0.954	1.047	1.047	1.035	1.142	1.090	1.053	1.063	1.104
014700	Energy drink	0.309***	1.050	1.046	1.046	0.809*	0.874	0.844*	0.893*	0.759**	0.941
014900	Soft drink carbonated	0.518***	0.847***	0.942	0.942	0.939	0.931	0.900	0.976	0.955	0.986
015000	Cola	0.309***	0.977	0.982	0.982	0.920	0.997	0.976	0.935	0.929	0.993

Continued on next page

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
015200	Orange juice	0.746**	0.890	0.844**	0.844**	0.867*	0.895	0.885	0.853	0.927	0.912
015300	Apple juice	0.586**	0.882	0.925	0.925	0.923	0.952	0.920	0.887	0.920	0.941

B.2.2 Direct forecasts of indices at level 5

Table A12: Direct three-step ahead forecasts of indices at level 5

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
01111	Rice	0.404**	0.812**	0.834*	0.834*	0.836*	0.979	0.925	0.709**	0.807*	0.905
01112	Flours and other cereals	0.471***	0.849**	0.895*	0.895*	0.892*	0.843***	0.817**	0.816**	0.877**	0.881**
01113	Bread	0.907	0.850	0.881	0.881	0.789*	0.796**	0.861*	0.917	0.906	0.898
01114	Other bakery products	0.563**	0.834**	0.867*	0.867*	0.840**	0.880	0.942	0.910	0.904	0.923
01115	Pizza and quiche	0.443***	0.905*	0.959	0.959	0.902	0.902	0.820**	0.864	0.883	0.951
01116	Pasta products and couscous	0.369***	0.869**	0.958	0.958	0.896*	0.976	0.915	0.946	0.852**	0.957
01117	Breakfast cereals	0.326**	0.900	0.955	0.955	0.931	1.001	0.988	0.967	0.988	1.031
01118	Other cereal products	0.515**	0.804**	0.849**	0.849**	0.834	0.846**	0.836	0.840**	0.866	0.881
01121	Beef and veal	0.559*	0.884	0.910	0.910	0.870	0.809	0.836	0.836*	0.870	0.891
01122	Pork	0.777	0.941	0.990	0.990	1.020	1.039	1.117	1.028	1.003	1.041
01124	Poultry	0.831	0.727**	0.788*	0.788*	0.741**	0.812**	0.813**	0.676**	0.847	0.801**
01127	Dried, salted or smoked meat	0.686	0.915	0.828	0.828	0.848	0.816*	0.834	0.921	0.848	0.870
01128	Other meat preperations	0.540**	0.873*	0.861	0.861	0.863	0.883	0.873	0.925	0.843	0.886
01131	Fresh or chilled fish	1.725**	1.278*	1.507**	1.507**	1.565**	1.020	1.397**	1.540***	1.605**	1.568**
01132	Frozen fish	0.483***	0.905	0.967	0.967	0.966	1.047	0.957	0.923	0.946	0.996
01135	Dried, smoked or salted fish and seafood	0.376**	0.859*	0.868	0.868	0.875	0.833*	0.863	0.803*	0.838	0.848
01136	Other preserved or processed fish and seafood-based preperations	0.524**	0.947	1.079	1.079	1.021	0.893	0.887	0.871	0.980	0.994
01141	Whole Milk	0.604*	0.838**	0.878	0.878	0.877	0.836**	0.905	0.871	0.891	0.882

Continued on next page

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
01143	Preserved milk	0.477	0.890	0.836*	0.837*	0.887**	0.879**	0.819*	0.912	0.915	0.951
01144	Yoghurt	0.541**	0.843*	0.894	0.894	0.844*	0.885*	0.854**	0.853*	0.880	0.889
01145	Cheese and curd	0.568*	0.854*	0.802***	0.802***	0.862*	0.799***	0.734***	0.841	0.813**	0.843**
01146	Other milk products	0.473**	0.805*	0.833**	0.833**	0.819**	0.793***	0.768**	0.754**	0.803**	0.810**
01147	Eggs	0.688	0.797*	0.847	0.847	0.727	0.824	0.832	0.777	0.849	0.823
01151	Butter	0.657**	0.894	0.947	0.947	0.803**	0.872**	0.817***	0.896	0.807**	0.916
01152	Margarine and other vegetable fats	0.545	0.968	0.993	0.993	0.931	0.943	0.988	0.916	0.963	0.983
01153	Olive oil	0.497***	0.923	0.850***	0.850***	0.894**	0.882***	0.977	0.804	0.980	0.981
01154	Other edible oils	0.492***	0.810***	0.869***	0.869***	0.797***	0.835***	0.791***	0.666**	0.849***	0.847***
01161	Fresh or chilled fruit	2.862*	3.299**	2.998**	2.998**	3.374**	2.891**	3.063**	3.703**	3.819**	3.515**
01163	Dried fruit and nuts	0.658**	1.071	1.001	1.001	0.956	0.998	1.104	1.017	1.066	1.091
01171	Fresh or chilled vegetables other than potatoes and other tubers	0.517***	0.917	0.906*	0.906*	0.976	1.042	1.027	0.905	0.939	0.977
01172	Frozen Vegetables other tahn potatoes and other tubers	0.721**	1.039	1.054	1.054	1.105	1.126	1.094	1.072	1.082	1.102
01173	Dried vegetables, other preserved or processed vegetables	0.571*	0.884*	1.129	1.129	1.047	0.934	0.970	0.873	1.098	1.108
01174	Potatoes	0.656**	0.955	0.932	0.932	0.857*	0.809**	0.773**	0.799**	0.800**	0.876
01175	Crisps	0.620*	0.929	1.000	1.000	1.033	1.056	1.041	0.988	1.046	1.050
01181	Sugar	0.454	0.927	0.915	0.916	0.917	0.887	0.897	0.609*	0.929	0.911
01182	Jams, marmalades and honey	0.997	0.967	1.018	1.018	1.133	0.902	1.091	1.004	1.124	1.081
01183	Chocolate	0.340	1.010	1.073	1.073	0.956	0.968	0.962	0.951	0.905	1.013
01184	Confectionery products	0.587**	0.851**	0.870**	0.870**	0.913*	0.888*	0.848*	0.871*	0.898*	0.908*

Continued on next page

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
01185	Edible ices and ice cream	0.820	0.991	1.041	1.041	1.105	1.073	1.077	0.966	1.111	1.161
01191	Sauces, condiments	0.641**	0.820**	0.895	0.895	0.918	0.847	0.797**	0.851	0.936	0.888
01192	Salt, spices and culinary herbs	0.841*	0.963	1.013	1.013	1.032	1.012	1.134	1.200	1.204	1.105
01193	Baby food	0.812	0.947	0.914	0.914	0.950	0.897*	0.958	0.957	0.974	0.959
01194	Ready-made meals	0.827	0.884**	0.891**	0.891**	0.911*	0.849**	0.876	0.944	0.926	0.912*
01199	Other food products n.e.c.	0.911	0.975	1.040	1.040	1.062	1.084	1.086	1.061	1.050	1.086
01211	Coffee	0.482***	0.983	0.944	0.944	0.914	0.842*	0.890	0.931	0.915	0.971
01212	Tea	0.677**	0.878	0.894	0.894	0.914	1.018	0.996	0.984	0.976	0.954
01213	Cocoa and powdered chocolate	0.505**	0.799**	0.819*	0.819*	0.799	0.833	0.860	0.844	0.792	0.834
01221	Mineral or spring waters	0.446**	0.871*	0.888*	0.888*	0.897*	0.923	0.862**	0.857**	0.902	0.917
01222	Softdrinks	0.366***	0.980	0.942	0.942	0.931	0.955	0.941	0.977	0.966	0.975
01223	Fruit and vegetable juices	0.662**	0.859**	0.847**	0.847**	0.840**	0.872	0.839	0.802*	0.908	0.883

Note: This table presents the root mean squared errors of three-step ahead forecasts of the indices. All columns contain ratios of RMSEs of the time series models incl. webscraping relative to the time series models without webscraping. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.2.3 Indirect forecasts of indices at level 5

Table A13: Indirect three-step ahead forecasts of indices at level 5

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
01111	Rice	0.282***	1.183**	1.179**	1.179**	0.850	1.146	1.164	1.129	0.787*	1.121*
01112	Flours and other cereals	0.430***	0.960	0.960	0.960	0.926	0.982	0.950	0.928	0.926	0.951
01113	Bread	0.943	1.038	1.034	1.034	1.010	1.069	1.096	1.030	1.051	1.050
01114	Other bakery products	0.561***	1.112	1.118	1.118	1.109	1.109	1.138	1.165	1.047	1.131
01115	Pizza and quiche	0.438***	1.115*	1.144**	1.144**	0.985	0.977	0.956	1.178	0.961	1.080
01116	Pasta products and couscous	0.319***	1.091*	1.128**	1.128**	0.963	1.173**	1.043	1.155*	0.969	1.084*
01117	Breakfast cereals	0.277*	1.235**	1.252*	1.252*	1.030	1.169**	1.107*	1.018	1.009	1.166**
01118	Other cereal products	0.438**	0.834	1.008	1.008	0.814	0.961	0.918	0.912	0.904	0.932
01121	Beef and veal	0.529**	1.160**	1.175	1.175	1.030	0.882	0.870	1.006	1.001	1.023
01122	Pork	0.772	1.122	1.138	1.138	1.186	1.198	1.191	1.220	1.167	1.180
01124	Poultry	0.773	1.164	1.230*	1.230*	1.201	1.115	1.236	1.184	1.231*	1.213*
01127	Dried, salted or smoked meat	0.590**	0.993	0.997	0.997	0.980	0.962	0.936	0.951	0.966	0.977
01128	Other meat preparations	0.532**	1.026	1.061	1.061	1.025	1.092	1.069	1.075	0.999	1.073
01131	Fresh or chilled fish	1.585***	1.733***	1.703***	1.703***	1.740***	1.444**	1.699***	1.771***	1.830***	1.754***
01132	Frozen fish	0.398***	1.129*	1.178*	1.178*	1.068	1.201*	1.111	1.106	1.017	1.136
01135	Dried, smoked or salted fish and seafood	0.267**	0.968	0.955	0.955	0.912*	0.953	0.897	0.907	0.863	0.923
01136	Other preserved or processed fish and seafood-based preperations	0.525**	1.166*	1.192*	1.192*	1.064	0.920	0.941	0.909	1.044	1.043
01141	Whole Milk	0.448**	0.909	0.904	0.904	0.818	0.904	0.922	0.890	0.865	0.884

Continued on next page

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
01143	Preserved milk	0.338	1.310*	1.263	1.265	1.056	0.989	0.998	1.347*	1.022	1.220*
01144	Yoghurt	0.462***	0.913	0.934	0.934	0.906	0.962	0.929	0.918	0.876*	0.921
01145	Cheese and curd	0.444**	0.978	1.028	1.028	0.997	0.963	0.938	0.974	0.952	0.984
01146	Other milk products	0.386**	0.893	0.888	0.888	0.890	0.915	0.886	0.843	0.850*	0.879
01147	Eggs	0.811	1.038	1.124	1.124	0.975	1.094	1.134	0.904	1.167	1.079
01151	Butter	0.539**	0.859	0.873	0.873	0.820*	0.921	0.939	0.820*	0.829*	0.876
01152	Margarine and other vegetable fats	0.532*	1.030	1.031	1.031	0.910	1.056	0.993	0.976	0.948	0.999
01153	Olive oil	0.394***	1.118	1.125**	1.125**	0.947	0.947	0.999	1.148	0.976	1.063
01154	Other edible oils	0.393***	1.011	0.983	0.983	0.883	0.972	0.928	0.948	0.920	0.964
01161	Fresh or chilled fruit	0.431	1.212*	1.025	0.911	1.099**	1.117*	1.172**	1.174**	1.122**	1.113**
01163	Dried fruit and nuts	0.482***	1.022	1.141**	1.141**	1.122*	1.011	1.036	1.201*	1.090	1.108
01171	Fresh or chilled vegetables other than potatoes and other tubers	0.403***	0.991	0.881*	0.883*	1.007	0.994	0.988	0.967	0.937	0.956
01172	Frozen Vegetables other tahn potatoes and other tubers	0.646**	1.179*	1.147	1.147	1.144	1.243**	1.168	1.261*	1.136	1.183
01173	Dried vegetables, other preserved or processed vegetables	0.372*	1.129	1.294*	1.294*	0.950	0.845	0.892	0.918	0.917	1.019
01174	Potatoes	0.609*	1.061	1.015	1.015	1.081	1.003	1.041	1.047	1.028	1.061
01175	Crisps	0.482*	1.058	1.151*	1.151*	1.057	1.113	1.123	1.181	1.027	1.121
01181	Sugar	0.343	1.119	1.039	0.998	0.816	0.928	0.954	0.581	0.860	0.900
01182	Jams, marmalades and honey	0.728**	1.517*	1.429	1.429	1.572	1.089	1.129	1.408	1.478	1.387
01183	Chocolate	0.372*	1.182	1.069	1.069	1.008	1.008	1.004	1.051	0.958	1.041
01184	Confectionery products	0.565**	1.080	1.125**	1.125**	1.090	1.091	1.079	1.077	1.066	1.102

Continued on next page

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
01185	Edible ices and ice cream	0.540**	1.032	1.079	1.079	1.008	0.954	0.996	1.124	0.988	1.054
01191	Sauces, condiments	0.583**	1.061	1.095	1.095	1.062	1.022	1.022	1.051	1.061	1.066
01192	Salt, spices and culinary herbs	0.758**	1.291*	1.235*	1.235*	1.230**	1.212*	1.273**	1.315**	1.298**	1.264**
01193	Baby food	0.601**	0.853	1.026	1.026	0.990	1.046	1.137	0.952	1.009	1.009
01194	Ready-made meals	0.849	1.005	1.046	1.046	1.062	0.968	1.037	1.092	1.071	1.038
01199	Other food products n.e.c.	0.839	1.175	1.151	1.151	1.156	1.138	1.187	1.099	1.152	1.159
01211	Coffee	0.416***	1.136	1.162	1.162	0.968	0.945	0.987	1.027	0.940	1.043
01212	Tea	0.560**	1.015	0.998	0.998	0.977	1.046	1.008	1.008	1.055	1.014
01213	Cocoa and powdered chocolate	0.380***	1.021	1.010	1.010	0.817**	0.851*	0.845**	1.062	0.769**	0.913
01221	Mineral or spring waters	0.393***	1.061	1.108*	1.108*	0.989	1.060	1.035	0.988	0.982	1.049
01222	Softdrinks	0.266***	1.062	1.054	1.054	0.899***	0.987	0.949	0.970	0.858***	0.986
01223	Fruit and vegetable juices	0.659**	1.007	1.024	1.024	1.079	0.948	0.935	1.004	1.009	1.017

Note: This table presents the root mean squared errors of three-step ahead forecasts of the indices. All columns contain ratios of RMSEs of the time series models incl. webscraping relative to the time series models without webscraping. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.2.4 Direct forecasts of indices at level 4

Table A14: Direct three-step ahead forecasts of indices at level 4

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
0111	Bread and cereals	0.626**	0.835*	0.774***	0.774***	0.792**	0.768***	0.820**	0.804**	0.835**	0.875
0112	Meat	0.548*	0.816*	0.811	0.811	0.818	0.816*	0.831	0.839	0.806	0.834
0113	Fish	0.762	0.864	0.955	0.955	0.942	0.849*	0.966	0.983	0.962	0.979
0114	Milk, cheese and eggs	0.553**	0.809**	0.796**	0.796**	0.815**	0.783***	0.756***	0.778**	0.799**	0.810**
0115	Oil and fats	0.573***	0.844**	0.911*	0.911*	0.851**	0.851***	0.798***	0.891**	0.840***	0.889**
0116	Fruit	2.619**	3.038**	2.780**	2.780**	3.161**	2.891**	2.923**	3.299**	3.752**	3.334**
0117	Vegetables	0.451***	0.875**	0.862**	0.862**	0.872**	0.993	0.955	0.861	0.917	0.928
0118	Sugar, jam, honey, chocolate and confectionery	0.490	0.847**	0.902	0.902	0.919	0.874**	0.861	0.935	0.916	0.929
0119	Food products n.e.c.	0.914	0.867*	0.851**	0.851**	0.995	0.773**	0.864	0.916	0.946	0.915
0121	Coffee, tea and cocoa	0.476***	0.900	0.909	0.909	0.885	0.836*	0.885	0.801	0.898	0.919
0122	Mineral waters, soft drinks, fruit and vegetables juices	0.468***	0.942	0.862**	0.862**	0.858*	0.920	0.910	0.876	0.899	0.914

Note: This table presents the root mean squared errors of three-step ahead forecasts of the indices. All columns contain ratios of RMSEs of the time series models incl. webscraping relative to the time series models without webscraping. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.2.5 Indirect forecasts of indices at level 4

Table A15: Indirect three-step ahead forecasts of indices at level 4

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
0111	Bread and cereals	0.563**	1.102	1.102*	1.102*	1.027	1.106*	1.115*	1.136**	1.048	1.101*
0112	Meat	0.526*	1.003	1.043	1.043	1.011	0.995	0.985	1.022	0.983	1.018
0113	Fish	0.666**	1.159	1.225*	1.225*	1.151	1.133	1.124	1.058	1.139	1.158
0114	Milk, cheese and eggs	0.452**	0.958	0.978	0.978	0.937	0.947	0.947	0.934	0.923	0.953
0115	Oil and fats	0.471***	0.905	0.915	0.915	0.863	0.935	0.949	0.844*	0.857	0.904
0116	Fruit	0.427*	1.193**	1.025	0.918	1.099**	1.120*	1.166**	1.187**	1.118**	1.113**
0117	Vegetables	0.331**	1.025	0.925	0.926	0.993	0.979	0.975	0.978	0.923	0.966
0118	Sugar, jam, honey, chocolate and confectionery	0.483**	1.106	1.138	1.132	1.061	1.028	1.052	0.984	1.046	1.083
0119	Food products n.e.c.	0.807	1.156*	1.201**	1.201**	1.219**	1.116	1.240**	1.242**	1.223**	1.214**
0121	Coffee, tea and cocoa	0.395***	1.117	1.115	1.115	0.928	0.917	0.953	1.007	0.909	1.008
0122	Mineral waters, soft drinks, fruit and vegetables juices	0.356***	1.011	1.033	1.033	0.917	0.980	0.949	0.945	0.878*	0.976

Note: This table presents the root mean squared errors of three-step ahead forecasts of the indices. All columns contain ratios of RMSEs of the time series models incl. webscraping relative to the time series models without webscraping. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.2.6 Direct forecasts of indices at level 3

Table A16: Direct three-step ahead forecasts of indices at level 3

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
011	Food	0.634	0.839*	0.784**	0.784**	0.938	0.819**	0.868**	0.902	0.830*	0.878
012	Non-alcoholic beverages	0.556**	0.804**	0.838**	0.838**	0.831**	0.833**	0.849	0.782**	0.868	0.862

Note: This table presents the root mean squared errors of three-step ahead forecasts of the indices. All columns contain ratios of RMSEs of the time series models incl. webscraping relative to the time series models without webscraping. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.2.7 Indirect forecasts of indices at level 3

Table A17: Indirect three-step ahead forecasts of indices at level 3

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
011	Food	0.394*	1.038	1.054	1.053	0.997	1.013	1.015	1.038	0.981	1.028
012	Non-alcoholic beverages	0.428***	0.965	0.993	0.993	0.898	0.857*	0.839*	0.838**	0.848*	0.911

Note: This table presents the root mean squared errors of three-step ahead forecasts of the indices. All columns contain ratios of RMSEs of the time series models incl. webscraping relative to the time series models without webscraping. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.2.8 Direct forecasts of indices at level 2

Table A18: Direct three-step ahead forecasts of indices at level 2

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
01	FOOD AND NON-ALCOHOLIC BEVERAGES	0.655	0.817*	0.781**	0.781**	0.877	0.810**	0.853**	0.837*	0.834*	0.863

Note: This table presents the root mean squared errors of three-step ahead forecasts of the indices. All columns contain ratios of RMSEs of the time series models incl. webscraping relative to the time series models without webscraping. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

B.2.9 Indirect forecasts of indices at level 2

Table A19: Indirect three-step ahead forecasts of indices at level 2

COICOP	Text	Naive	ARIMA	AR(1)	LW	ETS	Prophet	Neural-prophet	CES	Theta	Ensemble
01	FOOD AND NON-ALCOHOLIC BEVERAGES	0.409*	1.035	1.051	1.049	0.997	1.006	1.001	1.023	0.980	1.023

Note: This table presents the root mean squared errors of three-step ahead forecasts of the indices. All columns contain ratios of RMSEs of the time series models incl. webscraping relative to the time series models without webscraping. The stars indicate different levels of significance of a Diebold and Mariano (1995) test for equal predictive accuracy: * significant at the 10% level ($p < 0.10$), ** significant at the 5% level ($p < 0.05$) and *** significant at the 1% level ($p < 0.01$).

C Further robustness analysis

C.1 Possible structural break in food prices

We compare RMSEs of out-of sample one-step ahead forecasts of the elementary indices (level 6) before and after the outbreak of the Russia-Ukraine conflict in February 2022. Despite the potential structural break in the underlying food price series, the distributions of RMSEs remain largely unchanged.

Table A20: Summary statistics of out-of sample forecast until February 2022

Type	Mean	Median	SD	Min	Q1	Q3	Max
AR(1)	2.87	2.26	2.08	0.36	1.57	3.33	13.73
ARIMA	2.70	2.37	1.53	0.30	1.63	3.25	8.35
CES	2.88	2.27	1.78	0.62	1.54	3.80	9.65
ETS	2.81	2.25	1.73	0.35	1.59	3.75	9.28
Ensemble (w WS)	2.57	2.07	1.61	0.32	1.37	3.27	9.95
Ensemble (wo WS)	2.64	2.11	1.64	0.36	1.45	3.34	9.86
LW	2.89	2.26	2.13	0.36	1.57	3.33	14.19
Naive	4.58	3.59	3.31	0.54	2.53	5.88	21.13
Neuralprophet	2.97	2.26	1.81	0.57	1.61	4.05	9.74
Prophet	2.85	2.25	1.72	0.64	1.62	3.82	9.90
Theta	2.93	2.34	1.85	0.34	1.64	3.93	10.03
Webscraping	2.78	2.25	1.86	0.50	1.54	3.50	12.68

Note: This table presents summary statistics for the RMSEs of the one-step-ahead forecasts of the elementary indices (level 6). The out-of-sample forecasts cover the period from January 2021 to February 2022, ending just before the outbreak of the Russia-Ukraine conflict.

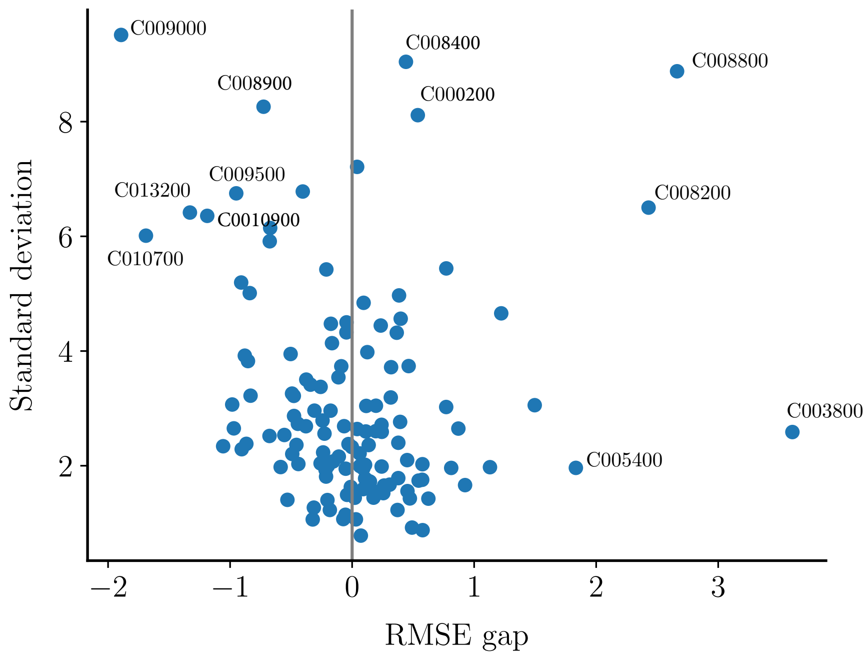
Table A21: Summary statistics of out-of sample forecast from March 2022 onwards

Type	Mean	Median	SD	Min	Q1	Q3	Max
AR(1)	3.15	2.53	1.83	0.87	1.99	3.53	9.89
ARIMA	3.38	2.82	1.88	0.95	2.19	3.84	10.50
CES	3.38	2.79	1.96	0.88	2.10	3.81	11.10
ETS	3.15	2.56	1.82	0.91	1.96	3.44	10.13
Ensemble (w WS)	2.91	2.41	1.64	0.80	1.87	3.23	8.88
Ensemble (wo WS)	3.03	2.51	1.68	0.84	1.91	3.36	8.69
LW	3.13	2.53	1.81	0.87	1.99	3.49	9.95
Naive	4.67	3.86	2.81	1.09	2.78	5.45	15.50
Neuralprophet	3.31	2.82	1.89	0.93	2.02	3.75	10.43
Prophet	3.46	3.01	1.85	1.01	2.22	4.06	11.51
Theta	3.24	2.56	1.98	0.81	1.94	3.49	11.06
Webscraping	2.97	2.53	1.68	0.86	1.90	3.38	10.88

Note: This table presents summary statistics for RMSEs of the one-step-ahead forecasts of the elementary indices (level 6). The out-of-sample forecasts cover the period from March 2022 to March 2024, starting shortly after the outbreak of the Russia-Ukraine conflict.

C.2 RMSE gaps

Figure A1: RMSE gaps for elementary indices (level 6)



Note: This figure displays the RMSE gaps between *Webscraping* and *Ensemble (wo WS)* for the elementary indices (level 6), plotted against the standard deviation of each elementary index series in the out-of-sample forecast window from March 2021 to March 2024.

Index of Working Papers:

January 13, 2021	Maximilian Böck, Martin Feldkircher, Burkhard Raunig	233	A View from Outside: Sovereign CDS Volatility as an Indicator of Economic Uncertainty
May 20, 2021	Burkhard Raunig	234	Economic Policy Uncertainty and Stock Market Volatility: A Causality Check
July 8, 2021	Thomas Breuer, Martin Summer, Branko Uroševic	235	Bank Solvency Stress Tests with Fire Sales
December 14, 2021	Michael Sigmund, Kevin Zimmermann	236	Determinants of Contingent Convertible Bond Coupon Rates of Banks: An Empirical Analysis
February 14, 2022	Elisabeth Beckmann, Christa Hainz, Sarah Reiter	237	Third-Party Loan Guarantees: Measuring Literacy and its Effect on Financial Decisions
February 16, 2022	Markus Knell, Reinhard Koman	238	Pension Entitlements and Net Wealth in Austria
May 9, 2022	Nicolás Albacete, Pirmin Fessler, Peter Lindner	239	The Wealth Distribution and Redistributive Preferences: Evidence from a Randomized Survey Experiment
June 20, 2022	Erwan Gautier, Cristina Conflitti, Riemer P. Faber, Brian Fabo, Ludmila Fadejeva, Valentin Jouvanceau, Jan-Oliver Menz, Teresa Messner, Pavlos Petroulas, Pau Roldan-Blanco, Fabio Rumler, Sergio Santoro, Elisabeth Wieland, Hélène Zimmer	240	New Facts on Consumer Price Rigidity in the Euro Area
June 29, 2022	Svetlana Abramova, Rainer Böhme, Helmut Elsinger, Helmut Stix	241	What can CBDC designers learn from asking potential users? Results from a survey of Austrian residents

July 1, 2022	Marcel Barmeier	242	The new normal: bank lending and negative interest rates in Austria
July 14, 2022	Pavel Ciaian, Andrej Cupak, Pirmin Fessler, d’Artis Kancs	243	Environmental-Social-Governance Preferences and Investments in Crypto-Assets
October 18, 2022	Burkhard Raunig, Michael Sigmund	244	The ECB Single Supervisory Mechanism: Effects on Bank Performance and Capital Requirements
April 5, 2023	Norbert Ernst, Michael Sigmund	245	Are zombie firms really contagious?
May 8, 2023	Richard Sellner, Nico Pintar, Norbert Ernst	246	Resource Misallocation and TFP Gap Development in Austria
September 5, 2023	Katharina Allinger, Fabio Rumler	247	Inflation Expectations in CESEE: The Role of Sentiment and Experiences
October 16, 2023	Pietro Saggese, Esther Segalla, Michael Sigmund, Burkhard Raunig, Felix Zangerl, Bernhard Haslhofer	248	Assessing the Solvency of Virtual Asset Service Providers: Are Current Standards Sufficient?
October 20, 2023	Pirmin Fessler, Severin Rapp	249	The subjective wealth distribution: How it arises and why it matters to inform policy?
October 27, 2023	Burkhard Raunig, Michael Sigmund	250	Watching over 21,000 Billion Euros: Does the ECB Single Supervisory Mechanism Affect Bank Competition in the Euro Area?
December 5, 2023	Markus Knell	251	Housing and the secular decline in real interest rates
December 14, 2023	Niko Hauzenberger, Florian Huber, Thomas O. Zörner	252	Hawks vs. Doves: ECB’s Monetary Policy in Light of the Fed’s Policy Stance
February 28, 2024	Lukas Olbrich, Elisabeth Beckmann, Joseph W. Sakshaug	253	Multivariate assessment of interviewer-related errors in a cross-national economic survey

March 5, 2024	Nicolas Albacete, Pirmin Fessler, Atanas Pekanov	254	The Role of MPC Heterogeneity for Fiscal and Monetary Policy in the Euro Area
March 11, 2024	Martin Schneider	255	What contributes to consumer price inflation? A novel decomposition framework with an application to Austria
April 9, 2024	Konstantin M. Wacker	256	Investment incomes vs. the trade balance: Is the current account still a meaningful concept?
August 12, 2024	Michael Sigmund, Johannes Wächtler, Philip Schuster, Robert Ferstl, Maria Teresa Valderrama	257	Take it and leave it: Banks' balance sheet optimization and targeted longer-term refinancing operations
August 20, 2024	Maximilian Boeck, Alina Steshkova, Thomas O. Zörner	258	The Impact of Currency Carry Trade Activity on the Transmission of Monetary Policy
October 3, 2024	Tabea Bucher-Koenen, Pirmin Fessler, Maria Silgoner	259	Households' risk perceptions, overplacement, and financial literacy
December 4, 2024	Valentin Voith, Sandra Mauser	260	Making Sense of Financial Vulnerability: Between Sensitivity, Resilience, and Exposure
January 10, 2025	Alessandra Agati, Michael Sigmund	261	Banking in the Negative: A Vector Error Correction Analysis of Bank-Specific Lending and Deposit Rates
January 15, 2025	Christian Beer, Robert Ferstl, Bernhard Graf	262	Improving disaggregated short-term food inflation forecasts with webscraped data