



Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Nowcasting food inflation with a massive amount of online prices

Paweł Macias, Damian Stelmasiak, Karol Szafranek*

Narodowy Bank Polski, Poland



ARTICLE INFO

Keywords:

Inflation nowcasting
Online prices
Big data
Nowcasting competition
Web scraping

ABSTRACT

The consensus in the literature on providing accurate inflation forecasts underlines the importance of precise nowcasts. In this paper, we focus on this issue by employing a unique, extensive dataset of online food and non-alcoholic beverages prices gathered automatically from the webpages of major online retailers in Poland since 2009. We perform a real-time nowcasting experiment by using a highly disaggregated framework among popular, simple univariate approaches. We demonstrate that pure estimates of online price changes are already effective in nowcasting food inflation, but accounting for online food prices in a simple, recursively optimized model delivers further gains in the nowcast accuracy. Our framework outperforms various other approaches, including judgmental methods, traditional benchmarks, and model combinations. After the outbreak of the COVID-19 pandemic, its nowcasting quality has improved compared to other approaches and remained comparable with judgmental nowcasts. We also show that nowcast accuracy increases with the volume of online data, but their quality and relevance are essential for providing accurate in-sample fit and out-of-sample nowcasts. We conclude that online prices can markedly aid the decision-making process at central banks.

© 2022 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

1. Introduction

In their influential work, Faust and Wright (2013) underline the importance of providing accurate inflation nowcasts. Following a comprehensive overview of popular methods used for forecasting inflation, the authors state that the quality of the inflation forecast in the long-term hinges considerably upon the accuracy of the short-term prediction of price dynamics. Consequently, inflation nowcasts should have important ramifications for the monetary policy conduct as pinning down the starting point of the headline inflation forecast in a precise manner may determine the success of the entire projection. While a voluminous strand of the literature predicts

inflation developments over the short and medium horizon, the approaches towards improving solely inflation nowcasts are relatively scarce. In this paper, we focus exclusively on this issue. To this end, we employ a unique, extensive dataset of online food and non-alcoholic beverages prices recorded since 2009, and we put forward a hypothesis that the use of online prices can markedly improve food inflation nowcasts.

The interest in analysing online prices has gained a lot of momentum in recent years, with the research spearheaded by studies within the Billion Prices Project, launched at MIT in 2008 and aimed at collecting a massive amount of granular data (Cavallio & Rigobon, 2016, and citations therein). In principle, online prices can be obtained from retailers' webpages via the automatic web scraping procedure (hence the term scraped data). Essentially, this process can be described as extracting information on the prices of products and their other relevant features from websites. While initially this can be technically

* Correspondence to: Narodowy Bank Polski, Economic Analysis and Research Department, Swietokrzyska str. 11/21, 00-919 Warsaw, Poland.

E-mail address: karol.szafranek@nbp.pl (K. Szafranek).

challenging to macroeconomic practitioners, ultimately, the benefits of observing consumer prices in high frequency should outweigh the entry costs related to the launch of data collection projects.

There are several appealing features of scraped data. First, with the automated process of gathering data, the delay in data collection is marginal as prices can be observed and downloaded remotely at any required frequency. Moreover, the researcher can quickly obtain an extensive amount of data. Second, due to their ubiquitous availability and the unlimited access to online prices, the researcher does not incur substantial costs, as is often the case with scanner data offered by private firms (Cavallo & Rigobon, 2016). Third, additional information such as product descriptions, availability, or discounts can complement the analysis on an international micro level regarding price stickiness, market segmentation, the effects of shock transmission, tax changes, or demand and supply disruptions. These research questions are undertaken in only a few recent papers (e.g. Cavallo, 2013, 2017, 2018; Cavallo & Rigobon, 2016; Gorodnichenko et al., 2018a; Gorodnichenko & Talavera, 2017). In turn, to our best knowledge, the only evidence on the usefulness of scraped data in forecasting inflation based on a large-scale project has been very recently provided by Aparicio and Bertolotto (2020).

This paper contributes to the forecasting strand of literature employing scraped data. To evaluate the usefulness of online prices in nowcasting food inflation, we use a unique, massive dataset of food prices in high frequency collected from webpages of major online retailers in Poland since November 2009.¹ To the best of our knowledge, the project launched in Narodowy Bank Polski aimed at collecting online prices in high frequency was among the first endeavors of this kind. To put this into perspective, at the end of December 2020, our food prices database consisted of almost 159 million prices for around 640 thousand products observed over ten years.² These quotations are carefully selected and classified using the methodology of the statistical office before they are aggregated to low-level price indices and used for nowcasting purposes.

We believe that these prices carry important information and can be suitable for nowcasting food inflation even though the Polish e-commerce market is relatively young and still in the development stage.³ Our approach

¹ In describing our dataset as massive, we follow the nomenclature used by Cavallo and Rigobon (2016) to emphasize the amount of data employed.

² These statistics are affected by the varying frequency of web scraping. Until 2017 web scraping was conducted weekly. Starting in 2017, the frequency increased to daily.

³ Recording impetuous and resilient growth, the revenue in the e-commerce market in the emerging economy of Poland is projected to reach around \$15.2 billion in 2021 and increase on average by 7.2% to approximately \$20.0 billion by 2025 (Statista, 2021). Błażejowska and Szafranek (2020) highlight that the e-commerce market in Poland has also rapidly grown in recent years as the average annual growth rate stood at 22.5% during 2004–2018. Moreover, during the lockdown, online retail sales skyrocketed, and their share in the total retail sales increased to almost 12% in April 2020. According to Eurostat (2020b), the percentage of individuals buying groceries online increased swiftly from 1% in 2005 to 14% in 2019.

is justified by the fact that (i) online prices of consumer goods can be highly similar if not the same as those in the traditional, brick-and-mortar stores (Cavallo, 2017), and (ii) we can successfully approximate food inflation by the development in online prices after a careful classification of products and the subsequent use of expenditure weighing, in line with the methodology of the statistical office. Moreover, the recent COVID-19 pandemic highlighted the importance of the e-commerce channel in a strikingly vivid manner when traditional, brick-and-mortar stores faced stringent confinement measures and collecting prices traditionally became troublesome for statistical offices.⁴

We contribute to the rapidly growing literature on the usefulness of online prices across four margins.

First, we show that incorporating the information on online prices into model-based frameworks substantially increases the nowcast accuracy with respect to both traditional and judgemental approaches that we consider. While we focus solely on food and non-alcoholic beverage prices, we provide evidence across many highly disaggregated components. We report that a marked improvement in the nowcast accuracy can be obtained for many food inflation components, especially those experiencing high volatility throughout the year. On the aggregate level, we show that our model outperforms the benchmark by 34.4% in terms of the RMSFE; a change statistically significant at $\alpha = 0.01$. This increase in the nowcast accuracy is remarkable given that information on online prices is introduced into simple models as an exogenous regressor. Moreover, approximating inflation by the pure change in online prices, at least when the data are properly treated, provides a significant improvement upon traditional benchmarks as the nowcast error falls by 29% compared to the benchmark. This is an encouraging conclusion for forecasting practitioners at central banks. It shows that after only several months of collecting data, the forecaster can already use them to adjust the nowcast before relying on model-based approaches.

Second, we demonstrate that the nowcast accuracy increases with the scale of the web scraping project. Based on nowcasts generated with data for every possible store composition, we conclude that, on average larger samples of online prices offer higher nowcast accuracy. We also show that the marginal gain in the accuracy decreases as the number of stores included in the sample rises. While we conclude that data quantity matters, we also stress the importance of data quality. Specifically, we advocate that employing an exact classification of products that effectively limits the sample is vital. Our experiments with data curation show that the machine-assisted human classification that guarantees accurate assignment of products into respective groups, together with the official weighing of products, provides a generous value added over a framework that uses an approximate approach (both in terms of in-sample fit as well as out-of-sample nowcasts). When the approximate method of

⁴ According to Eurostat (2021), the fraction of imputed prices for the HICP basket peaked at 32% in April 2020 for the euro area during the first wave of the pandemic. The respective share for Poland amounted to 19%.

product classification is considered, the RMSFE for food and non-alcoholic beverages inflation increases by 30%–40% depending on the approach used. Thus, we argue that dispensing with proper data curation severely affects the nowcast precision that online prices grant.

Third, we indicate that combining forecasts from simple models augmented with online prices does not necessarily lead to further improvements in the nowcast accuracy. Our framework, relying on a recursively optimized specification of a simple autoregressive model with distributed lags, produces more precise nowcasts by around 15% in terms of the RMSFE statistic than nowcast combinations, a statistically significant result. In turn, we show that nowcast combinations are more useful when working on a short sample of data.

Fourth, we document that the nowcast precision of the model augmented with online prices improves compared to the traditional approaches following the outbreak of the COVID-19 pandemic. Notably, the quality of the nowcast is still comparable with the judgmental approach, which is extremely useful during unorthodox times. We attribute this to the fact that online prices can be an extremely precise source of information. To substantiate this claim, we show that the inclusion rate of online prices into model-based approaches surged during 2020 and exceeded 90%.

From the policy perspective, we conclude that online prices are an important source of information for CPI development. Thus, they can markedly aid the monetary policy conduct at central banks, given the importance of the inflation nowcasts in macroeconomic projections. They can also monitor price developments in the economy in real-time. This is an important conclusion, both for advanced economies as well as for emerging economies such as Poland, where the share of food expenditures in the consumption basket exceeds 25% and the contribution of food inflation remains substantial (Szafranek & Hałka, 2019). While in advanced economies this share is considerably lower, the importance of the food component has increased due to the COVID-19 outbreak.⁵

The rest of the paper is organized as follows. Section 2 provides a snapshot of the empirical approaches towards forecasting inflation and reviews recent works on the usefulness of online prices in inflation analysis and forecasting. Section 3 introduces the concept of web scraping and describes the E-CPI project that started in Narodowy Bank Polski in 2009 along with the data. Section 4 reports the models entering the real-time nowcasting exercise. Baseline results are reported in Section 5. Section 6 presents the outcomes of the sensitivity analyses, whereas Section 7 provides further results about the importance of sample composition (i.e., data quantity) and product selection (i.e., data quality) for nowcast accuracy. The final section concludes.

⁵ In 2020, the expenditure structure changed significantly in most of the EU countries, both advanced and emerging economies. For the whole EU economy, the food weight in the HICP inflation basket increased by 2.0 percentage points in 2021, which is an exceptional change considering past developments of the food expenditure share and its negative long-term tendency.

2. Literature review

Forecasting inflation is perceived as a daunting challenge. Evidence put forward by Atkeson and Ohanian (2001) indicates that a naive benchmark extrapolating past inflation trends is hard to beat. This line of argument is further supported by Stock and Watson (2007), who claim that it is extremely difficult to provide a generous value added over the approximation of the trend inflation. Furthermore, in a comprehensive comparison of models, Faust and Wright (2013) corroborate the fact that simple models accounting for a slowly varying local mean are often hard to beat by more refined approaches. The authors also claim that judgemental forecasts – prepared by professional forecasters and reflecting their subjective perception of economic developments – deliver superior accuracy.

Evident difficulties in providing accurate inflation nowcasts and forecasts have motivated researchers to develop sophisticated models for predicting price developments. Consequently, a plethora of various frameworks have been put forward over the years, and a just review of this strand of literature seems infeasible. In Table 1, we present only the tip of the iceberg of various methodological approaches towards predicting consumer prices.

Our understanding of this literature review is that an overwhelming amount of work is dedicated to forecasting inflation over the short- and medium-term. However, relatively little attention is paid solely to the precision of inflation nowcasts. In turn, the evidence provided by Faust and Wright (2013) indicates that the success of preparing accurate inflation forecasts relies heavily on pinning down inflation nowcasts in an accurate manner, while substantial errors in the nowcast propagate through the whole horizon. Given that Monteforte and Moretti (2013) use a mixed-frequency model for euro area inflation and conclude that the inclusion of daily variables helps to reduce nowcast errors compared to models relying on monthly data. Modugno (2013) employs a trading day frequency factor model and shows that including high-frequency data increases the precision of the nowcast. Departing from the mainstream literature that uses large datasets, Knotek and Zaman (2017) consider a small number of data series at different frequencies to inform the nowcast. They report that the model generates historically lower RMSFEs than MIDAS regression in the spirit of Monteforte and Moretti (2013) and the DFM model based on Modugno (2013) and is generally more accurate than nowcasts of headline CPI prepared by the Blue Chip consensus or the Survey of Professional Forecasters. Most recently, Tallman and Zaman (2020) construct hybrid forecasts by accounting for survey nowcasts, VAR forecasts, and long-horizon survey projections. They show that tilting the forecast towards the survey nowcast significantly improves the inflation forecast accuracy in the nowcast quarter.

We contribute to the literature by discussing how macroeconomic practitioners can improve inflation nowcasts by incorporating online prices from retailers' webpages into simple, univariate models. Online prices gain

Table 1

Approaches towards inflation forecasting.

Methodological approach towards forecasting inflation	Examples of prominent studies
Frameworks relying on the economic premise of the Phillips curve either in its reduced or structural form	Stock and Watson (1999), Atkeson and Ohanian (2001), Rumler and Valderrama (2010), Belmonte et al. (2014), Kichian and Rumler (2014), Kabukcuoğlu and Martínez-García (2018), McKnight et al. (2020).
Time-series, bottom-up approaches exploiting the advantages the disaggregation provides	Duarte and Rua (2007), Tallman and Zaman (2017).
Combination of large sets of models, the use of compressed information from datasets via unobserved factors or both	Forni et al. (2003), Stock and Watson (2006), Eickmeier and Ziegler (2008), Koop and Korobilis (2012), Groen et al. (2013), Korobilis (2017).
Multivariate approaches accounting for time-varying parameters and stochastic volatility, multiple breaks in time series, steady-state assumptions given the inflation targets or seasonality patterns	Villani (2009), Jochmann et al. (2010), Stelmasiak and Szafranski (2016), Mandalinci (2017), Tallman and Zaman (2020).
Frameworks based upon nonlinear models and machine learning methods such as bagging, boosting, artificial neural networks or random forests	McAdam and McNelis (2005), Nakamura (2005), Binner et al. (2006), Inoue and Kilian (2008), Rapach and Strauss (2010), Medeiros et al. (2019), Szafranek (2019).
Structural, large scale DSGE models for policy analysis and forecasting	Edge and Gurkaynak (2010), Kolasa and Rubaszek (2015).
Judgemental approaches based on expert's knowledge	Ang et al. (2007), Faust and Wright (2013).

increasing attention in economic analyses, with this strand of empirical literature quickly expanding. From the policy perspective, such extensive, unique databases containing online prices open vast possibilities for researchers. Prominent papers examine the price formation in the online markets, their characteristics and cross-border comparisons, the degree of price stickiness, the sensitivity to changes in exchange rates, the price synchronization within and between retailers active in online and offline channels, and the frequency of price changes or the bias in official inflation figures (Cavallo, 2013, 2017, 2018; Cavallo & Rigobon, 2016; Gorodnichenko et al., 2018a; Gorodnichenko & Talavera, 2017; Lünemann & Wintr, 2011). With regards to the strand of literature focusing on supermarket prices, both web-scraped data and scanner data are demonstrated to be useful for price dynamics measurement (Goolsbee & Klenow, 2018; Jaworski, 2021), analysis of price sensitivity to shocks (Gorodnichenko et al., 2018b; Jaravel & O'Connell, 2020), or understanding modern pricing policies (Aparicio et al., 2021). Online prices are also increasingly included in the compilation of the consumer price indices by statistical offices (e.g., in the US, the UK, the Netherlands, New Zealand or Norway, Cavallo & Rigobon, 2016) and can complement traditional data sources.⁶

In turn, to the best of our knowledge, the usefulness of web-scraped data in forecasting inflation based on a large-scale project is examined so far only by Aparicio and Bertolotto (2020). Using scraped data from July 2008 to September 2016 for ten advanced economies, the authors show that online prices are useful in forecasting headline inflation in the short-term. Forecasts from the parsimonious model with online prices beat traditional

benchmarks and two leading surveys of professional forecasters. In turn, Powell et al. (2018) demonstrate that even with a limited set of online prices for a narrow group of food and alcoholic beverages products spanning approximately 14 months, one can already produce product-specific daily forecasts helpful in policy analysis.

While our methodological approach is similar to the framework put forward by Aparicio and Bertolotto (2020), our contribution presents complementary results in several aspects. First, though we focus solely on food price inflation, we provide evidence for a large number of highly disaggregated inflation components, while Aparicio and Bertolotto (2020) study exclusively headline inflation. Second, we examine the usefulness of online prices for a small, emerging economy whose online market is much less developed than its counterparts in advanced economies. Third, we report how nowcasting errors are related to the scale of the project (measured with the number of stores being web scraped) and the approach towards data curation (the classification of products into respective groups and the use of expenditure weighing). Finally, taking advantage of the most recent data, we examine the accuracy of nowcasting with online prices during the COVID-19 pandemic.

3. Methodology

3.1. E-CPI project

The E-CPI project aimed at collecting online prices started in Narodowy Bank Polski in December 2009. From the beginning, the project has been dedicated to constructing the food and non-alcoholic beverages index with timely information provided by online prices. Initially, prices were collected from the websites of online stores once a week. As the project advanced, the frequency of data collection was changed to daily from the beginning of 2017. Moreover, while forecasting food prices remains important due to their volatility and their large

⁶ Following the COVID-19 pandemic Eurostat (2020a) encouraged EU Member States to replace field price collection by online sources for the compilation of the HICP. There is also a growing interest in employing scanner data in price index compilation (Lamboray, 2021, and citations therein).

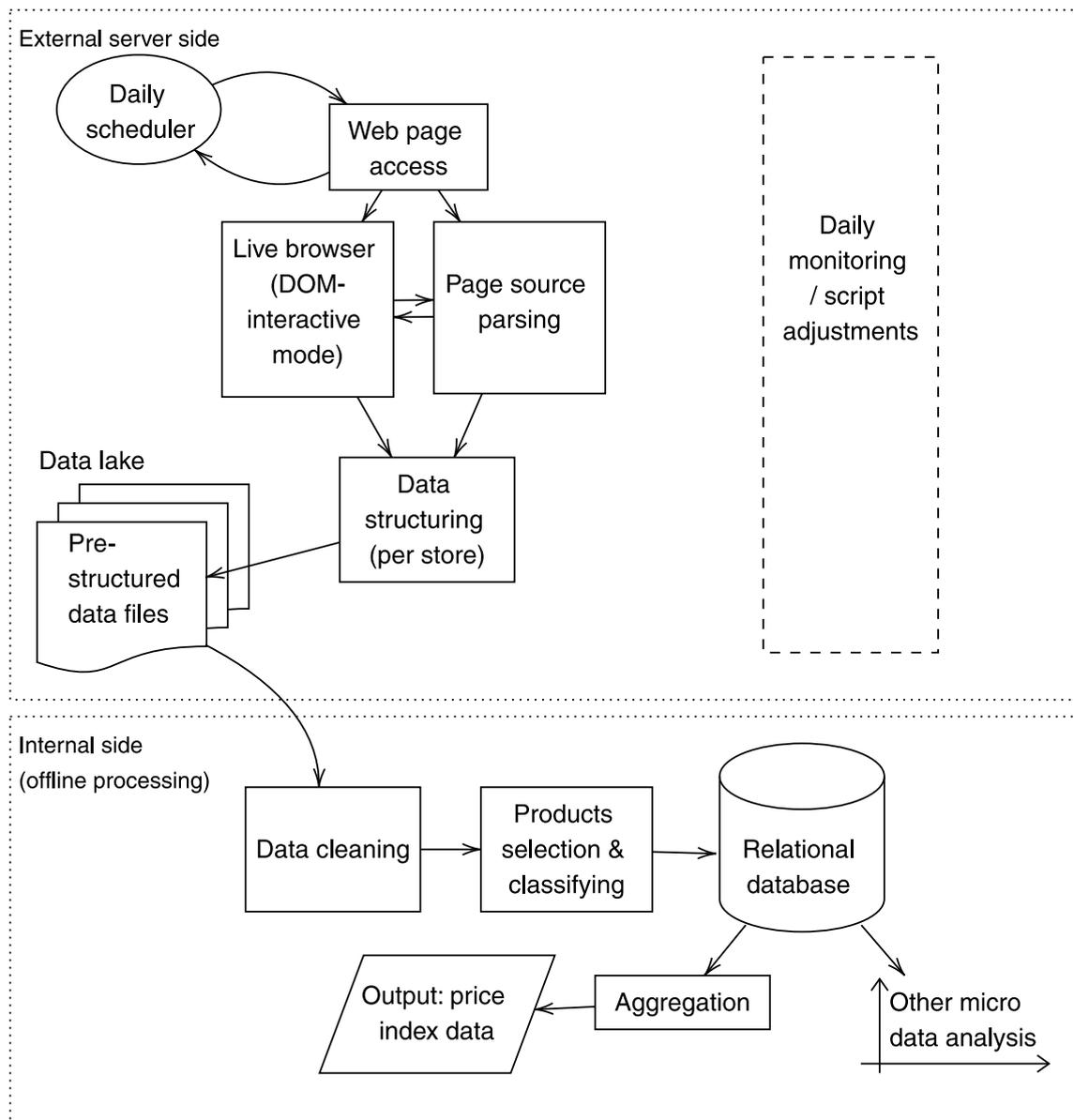


Fig. 1. Schematic representation of the web scraping process. Note: Fig. 1 depicts the process of obtaining and processing data from online retailers.

contributions to headline inflation in Poland (Szafranek & Hałka, 2019), the database is systematically expanded to account for clothing, footwear, products from home-improvement do-it-yourself stores, electronics, drugs, and airplane tickets.

3.2. The approach towards web scraping

Fig. 1 depicts our approach towards obtaining and processing data from online retailers. We use an external server (cloud solution) like other large-scale web scraping or crawling projects. Our pipeline on the external server side is as follows. A scheduler distributes daily tasks that execute Python scripts for specific web stores. Every script is tailored to a given web store as it considers

the webpage structure, its mechanisms and available data. Our objective is to maximize the data retrieval efficacy while ensuring that all the data are the same as displayed for webpage viewers (consumers). In production, we use either a live browser approach (primarily Google Chrome with browser automation tools), a page source parsing, or both. The webpage is rendered in the web browser approach, and we interact with DOM objects as real users would. Instead, page source parsing may be done completely offline; thus, it is considerably faster and utilized wherever possible to optimize the whole process (see Table 3 for a comparison of those techniques). Regardless of the choice of the web scraping technique, we minimize the burden on web servers and respect the Code of Practice for Statistics (ONS, 2020) by delaying accessing pages.

When the data are already scraped, daily web-store-specific data files are transferred to our internal infrastructure and processed entirely offline. The internal side pipeline consists of data cleaning and products identification per store, thus producing panel price data (see Macias & Stelmasiak, 2019, for more details on data cleaning). In the next step, we classify new items to respective COICOP groups. Available data from different stores are mapped, unified, and digested into a relational database. We use that fully-structured data to produce monthly price indices for nowcasting needs and other micro price data analyses.

3.3. Product classification and expenditure weighing

In contrast to traditional, survey-based collecting methods, web scraping allows using information about all items available in online stores quickly and inexpensively. However, there is a potential trade-off between data quantity and data quality. While including all available products may help to better track overall price dynamics in an economy (Cavallo, 2013), we find that for CPI nowcasting the best approach is to choose only the products similar enough to those chosen by the statistical office. We will substantiate this claim later in the text.

We distinguish two main factors that may lead to discrepancies between our online price index and the CPI on the micro price level. The first one is *product classification error*. By this, we mean a situation when a product is clearly and objectively assigned to the wrong class (group) of products (e.g., apple classified to pears). In this scenario, a price index may suffer from the price mismatch and the improper assignment of the expenditure weight. The second factor, i.e., *product selection bias*, is more subtle as it results from a mismatch of product varieties, even if the product itself is properly classified to the particular group of goods. While our dataset contains various types of products, including the one monitored by the statistical office, selection bias occurs when we choose a different one. There are sporadic cases of varieties tracked by the statistical office that are not included in our dataset.

Fig. 2 illustrates the in-sample comparison between the monthly and the annual rate of change in the official CPI for food and non-alcoholic beverages (grey line) and in the index based on online prices (red line). Panels A and B clearly show that using the information on prices of all available products without any weighing scheme nor variety selection results in a highly inaccurate approximation of the official CPI. For instance, this measure does not pick up important seasonality patterns. Moreover, the annual rate of change develops fairly independently of the official measure.⁷ Simple correlation between the official

⁷ The unsatisfactory performance of the unweighted online index originates from the under-representation of some basic goods like fruits, vegetables, and other unprocessed or low-level processed products that are available in one variant only. At the same time, many processed products are offered in different flavors, sizes, and brands. Moreover, maintaining a wide offer in online stores is especially easy due to the lower cost of keeping products in stock and bigger area of activity than in traditional shops. Therefore, some varieties of products gain importance in the unweighted price index, while their price dynamics are not representative of the average household consumption.

Table 2
Summary of the E-CPI database.

Store ID	q	q_s	n_p	\bar{n}_p	\bar{n}_d	$med\{n_d\}$
1	82,995	19,250	13,394,437	7,327	161	42
2	56,739	16,368	14,721,627	8,544	260	76
3	76,652	13,007	25,600,013	17,534	334	222
4	50,954	15,935	17,196,321	11,495	341	164
5	132,543	26,175	43,387,750	23,709	327	187
6	21,658	7,381	6,346,966	9,544	302	266
7	103,234	34,054	16,708,475	9,170	175	87
8	119,168	27,544	21,630,105	13,213	182	34
Total	643,943	159,714	158,985,694	86,687	250	102

Note: Table 2 reports basic characteristics of the E-CPI database across 8 online stores providing food and non-alcoholic beverages. Abbreviations used: q – number of products, q_s – number of selected products, n_p – number of observed prices, \bar{n}_p – average number of prices per day, \bar{n}_d – mean number of days a price is observed, $med\{n_d\}$ – median number of days a price is observed.

monthly inflation and the E-CPI inflation, in this case, amounts to 0.42, and the RMSE equals 0.72.⁸ Panels C and D indicate that using an approximate product classification⁹ together with the official weighing of products improves our in-sample fit. The correlation between the official monthly inflation and the E-CPI counterpart increases to 0.64, the RMSE falls to 0.62, and the annual rate of change in online prices reflects a greater extent the official change in prices. Finally, panels E and F show the exact product classification¹⁰ – supervised by the team members – together with the official weighing of products results in the further improvement in the in-sample fit. The correlation for the monthly indices increases to 0.77, while the RMSE falls to 0.50. Moreover, both the monthly and the annual change in the online price index resemble much better the official CPI, though some discrepancies are still present. Therefore, our results indicate that online and official price changes are similar. We also study the impact of product classification on the out-of-sample nowcast performance in Section 7.

3.4. Data

In this paper, we focus solely on food and non-alcoholic beverages prices. Our total sample of online prices collected within the E-CPI project spans the period December 2009 – December 2020. As of December 2020, the database contains almost 159 million quotations for around 650 thousand food products (identifiable items

⁸ The RMSE (in pp) is calculated as follows:

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

where y_t is the monthly rate of change of the official food and non-alcoholic beverages inflation and \hat{y}_t is its E-CPI counterpart. Calculations are carried out for January 2010 – December 2020 (132 observations).

⁹ This classification is based on the dictionary approach. It constitutes the initial step of the classification process we perform monthly. More details on the classification procedure are provided in the Supplementary Material, available online (Section S1).

¹⁰ Here, we refer to the final and verified classification results. They are based on machine-generated suggestions manually revised (if necessary) by the team members. More details on the classification procedure are provided in the Supplementary Material, available online (Section S1).

Table 3
Comparison of web scraping techniques.

Method	Processing speed	Processing difficulty	Stability	Availability	Data amount
Page source	Fast	High	Good	High	Sometimes more than on the screen
Live browser	Slow	High	Medium	Very high	On the screen
API	Very fast	Low	Very good	Very low	Often more than on the screen

Note: based on authors' experience.

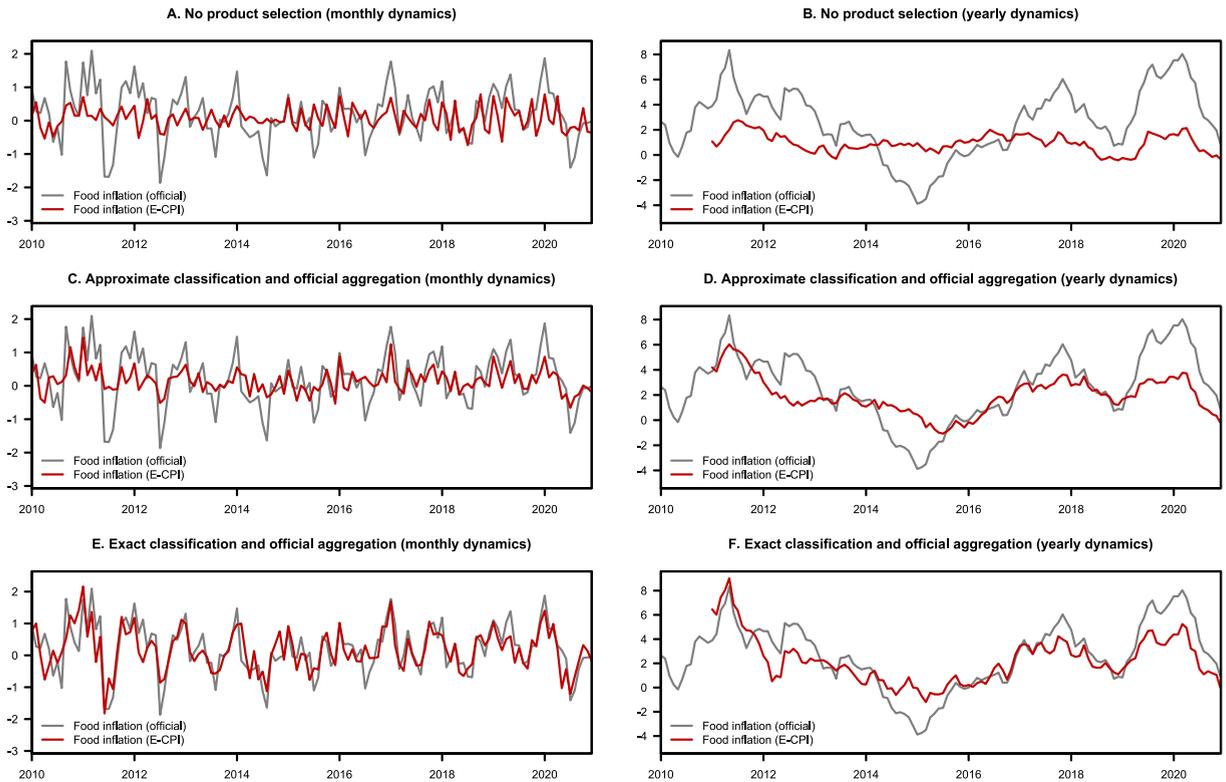


Fig. 2. Comparison of official CPI and online prices. Note: Fig. 2 provides various comparisons of the official food and non-alcoholic beverages inflation (grey line) and the respective measure calculated using online prices (red line). Panels A and B compare monthly and yearly dynamics when all products are selected to calculate the change in prices. In turn, panels C and D provide monthly and yearly dynamics when products are selected using the approximate classification and subsequently weighted using the official inflation basket weights. Panels E and F show the monthly and yearly dynamics when products are classified into respective groups with the means of the exact method and subsequently weighted using the official inflation basket weights. (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this article).

either by their unique ID or product description) from 8 grocery stores in Poland (Table 2). The effective sample size used to compile our baseline online price index, obtained with the exact classification of products, is considerably smaller as we discard 75% of all products. Still, it is substantial as it amounts to 160 thousand unique items. In turn, the average number of price quotations per day equals around 87 thousand in total, as not all prices are observed each day. Moreover, there is a substantial degree of heterogeneity across individual stores and across time (Fig. 3). To give some intuition into the sparse nature of web-scraped prices, we also provide very basic turnover figures. The individual price is observed for 250 days on average, while due to the heavy asymmetries, the corresponding median amounts to only 102 days. This figure varies a lot across the stores (roughly 30–300 days), resulting from different product ranges and pricing policies.

4. Forecasting competition

Throughout the paper, we define the inflation rate as the monthly, non-seasonally adjusted change in prices.¹¹ We denote the change in the online price index as $o_{c,t}$ for the c th inflation component, i.e., either 1 of the 84 elementary COICOP groups, 1 of the ten food inflation subaggregates, or the food and non-alcoholic beverages aggregate. Online price changes $o_{c,t}$ based on the E-CPI dataset are expected to approximate changes in official consumer price indices denoted as $p_{c,t}$.

In calculating changes in online prices for elementary groups we follow closely the official methodology. Let

¹¹ For forecasting experiments, we have also tested the logarithmic transformation of price indices. As we fail to find significant differences, we proceed with simple price growth rates (as communicated by the statistical office). These results are available upon request.

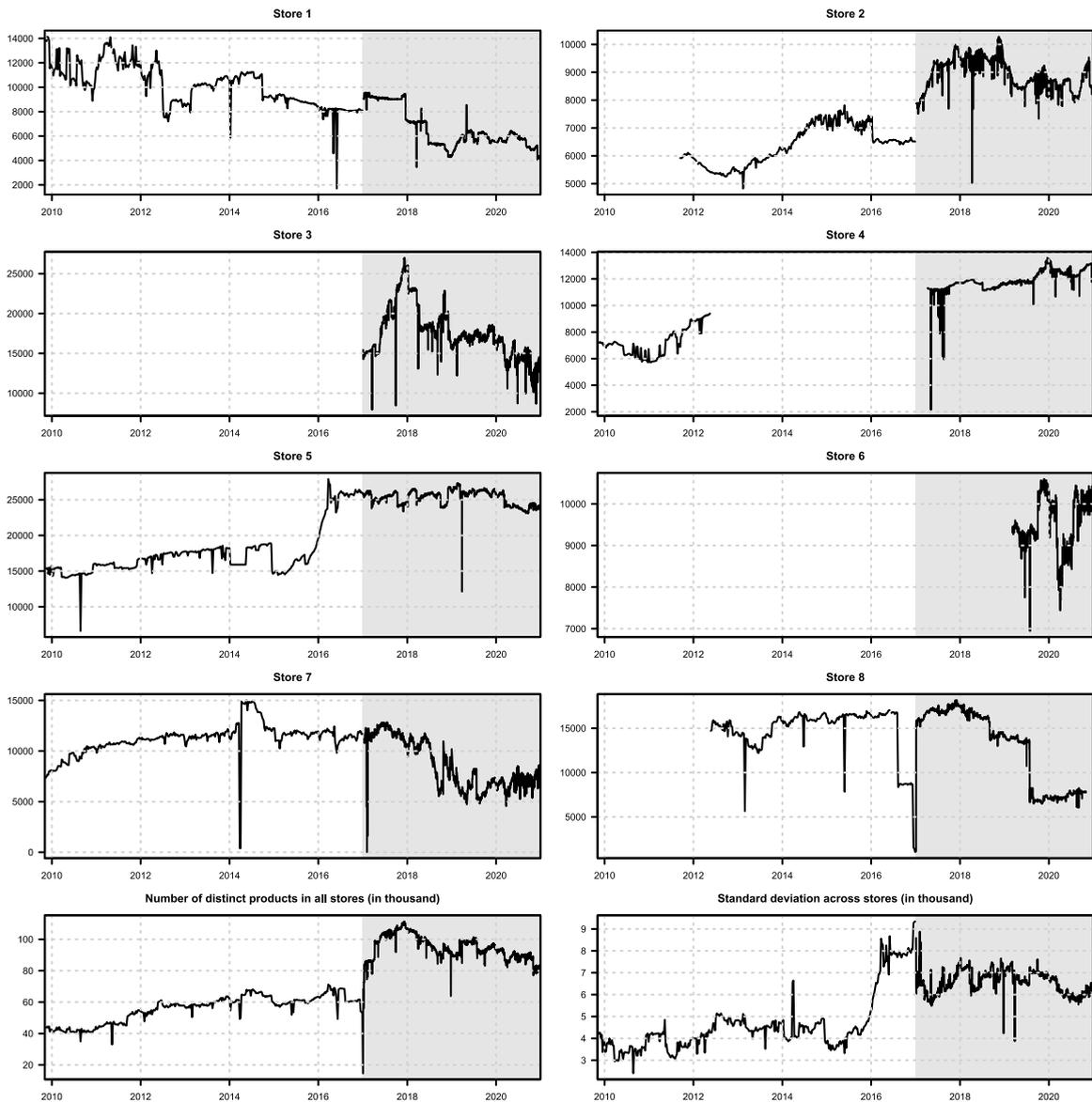


Fig. 3. Number of products across stores and in the sample. Note: Fig. 3 shows how the number of distinct products evolves across stores and throughout time. Missing data occurred due to technical problems related to web scraping (e.g., Store 4 data collection was limited by the technology of our choice) or unavailable stores' servers. The shaded area indicates a time period when data collection is performed daily instead of weekly.

$\bar{S}_{i,k,t}$ denote the price of an individual product i posted on a website on day k of month t . In each month, we calculate a simple average of prices $\bar{S}_{i,t} = K^{-1} \sum_{k=1}^K S_{i,k,t}$. For each product we obtain the monthly change in average prices $s_{i,t} = \bar{S}_{i,t} / \bar{S}_{i,t-1}$. Next, using the geometric mean, we aggregate monthly dynamics at the product level into the monthly change in prices at the elementary level $o_{c,t} = (\prod_{i=1}^I s_{i,t})^{\frac{1}{I}}$, where I reflects the number of products within the group c . For higher levels of aggregation we use the official CPI weights based on the expenditure of households. For instance, any aggregate price index is calculated as $\sum_{c \in M} o_{c,t} w_c^*$, where M represents some set of elementary groups and weight w_c^* is expressed as a share of subcomponent c in the aggregate M (for

simplicity we omitted time index in the weights, which change only once a year).

To assess the quality of the inflation nowcasts based on online prices, we establish a nowcasting experiment among the following simple, univariate linear models¹² across all considered inflation components:

¹² Despite the popularity of the mixed-frequency models, we do not consider any of these in this paper. Our forecasting framework's primary assumption is to mimic the official CPI methodology closely. A price index is calculated using a ratio of price means on a fixed interval (month). This is in line with the statistical office's formulae. Nonlinear weighting schemes employed, i.e., by MIDAS models, are incompatible with that methodology. We strongly believe that the MIDAS approach may be beneficial when accounting for high frequency variables from other sectors of the economy, as outlined in Section 2. However, this experiment remains beyond the scope of this paper.

1. The simple random walk model for the monthly inflation (*RW*). The nowcast is generated as $p_{c,t+1} = p_{c,t}$.
2. The random walk model for the monthly inflation à la the Atkeson and Ohanian (2001) specification adjusted by a seasonal factor (*AO^{SA}*). As the majority of the food components follows a strong seasonal pattern, in order to improve the benchmark we apply the typical averaging scheme to seasonally adjusted data. Next, we adjust the forecast by a Tramo-Seats seasonal factor as follows: $p_{c,t+1} = \frac{1}{12} \sum_{j=1}^{12} p_{c,t-j+1}^{SA} + \hat{\varepsilon}_{c,t+1}^{TS}$. Tramo-Seats specifications for components rely on the automatic selection (no fine tuning).
3. The best SARMA model based on in-sample fit (*BS^{IS}*). The specification is as follows: $p_{c,t} = \mu_c + \varepsilon_{c,t} + \sum_{i=1}^p \phi_{c,i} p_{c,t-i} + \sum_{i=1}^q \theta_{c,i} \varepsilon_{c,t-i} + \sum_{i=1}^P \Phi_{c,i} p_{c,t-i*12} + \sum_{i=1}^Q \Theta_{c,i} \varepsilon_{c,t-i*12}$, assuming that $\varepsilon_{c,t} \sim NIID(0, \sigma_c^2)$. In total, we consider 64 specifications by either including or excluding the AR, the MA, the seasonal AR and the seasonal MA parts. Conditional upon inclusion, we allow $p \in \{1, 2, 3\}$, $q \in \{1, 2, 3\}$, $P \in \{1\}$, $Q \in \{1\}$. The best model is chosen using the Bayesian Information Criterion (BIC). The model is iterated one-step ahead to obtain the nowcast of inflation.
4. The best SARMA model is based on out-of-sample accuracy (*BS^{OS}*). We use the same specification for the model as in *BS^{IS}*, but we choose the best model using the minimal RMSFE criterion calculated on a pseudo validation set.¹³ We use this model as the benchmark.
5. The pure, ex-post E-CPI nowcast (*EC^{EX}*). This approach accounts for all available online prices from a given month. We denote it with the superscript *EX* to indicate that this is, in fact, an *ex-post* estimate of the change in prices. Thus, the nowcast is generated as $p_{c,t+1} = o_{c,t+1}$.
6. The pure, real-time E-CPI nowcast (*EC^{RT}*). Since the publication of the CPI takes place around the 15th day of the month $t + 1$, in this approach, we exploit the timeliness of our online data. Hence, we calculate the change in online prices $o_{c,t}^*$ using the readily available data until the middle of the month to account for a similar information set with respect to the statistical office. Hence, the nowcast is generated as $p_{c,t+1} = o_{c,t+1}^*$, where * denotes that only the data from the first half of the month is used.

7. The best SARMAX model with online prices as the exogenous variable (*EC^{SX}*). The specification is the following: $p_{c,t} = \mu_c + \varepsilon_{c,t} + \sum_{i=0}^B \beta_{c,i} o_{c,t-i} + \sum_{i=1}^p \phi_{c,i} p_{c,t-i} + \sum_{i=1}^q \theta_{c,i} \varepsilon_{c,t-i} + \sum_{i=1}^P \Phi_{c,i} p_{c,t-i*12} + \sum_{i=1}^Q \Theta_{c,i} \varepsilon_{c,t-i*12}$ and $\varepsilon_{c,t} \sim NIID(0, \sigma_c^2)$. The model is iterated one-step ahead to obtain the nowcast. Given the fact, that this is a real-time exercise, we substitute $o_{c,t+1}$ with $o_{c,t+1}^*$ to account only for the data available up to the nowcast date. In total, we consider 256 specifications by accounting for various lags of B , $B \in \{0, 1, 2, 3\}$ and either excluding or including the AR, the MA, the seasonal AR and the seasonal MA parts (with lags as in the *BS^{IS}* case). The best model is chosen using the minimal RMSFE criterion calculated on a pseudo validation set (similarly as in the *BS^{OS}* case). We treat this model as our baseline approach.
8. The simple combination of the *EC^{SX}* models (*EC^{MC}*). We pool predictions from 256 individual models with equal weights to examine whether this leads to further gains in the nowcast accuracy, given the superiority of simple combinations forecasts upon more refined approaches (Genre et al., 2013; Timmermann, 2006). This is also motivated by the desire to reduce the probable model selection bias given a large number of possible specifications.¹⁴
9. The combination of *BS^{OS}* models with equal weights (*BS^{MC}*). We provide this as a benchmark to the previous model ensemble.
10. The judgemental forecast (*JD*). The nowcasts are prepared internally at Narodowy Bank Polski by an expert for in-house purposes. This paper reports errors only for the aggregate food and non-alcoholic beverages inflation.

We apply a recursive estimation strategy using the expanding window in the baseline scenario. We suspect that the more data on online prices we include, the larger the gains in the forecasting accuracy should be. This assumption is also in line with the practical approach towards forecasting. The estimation sample starts in January 1999 ($t = 1$) and ends in December 2016 ($t = T^*$).¹⁵ The evaluation period starts in January 2017 and ends in December 2020 ($t = T$). Hence, we evaluate our predictions using 48 point forecasts, a sizeable sample given the nature of our data.

¹³ For the recursive estimation strategy, whenever we choose the best model on a validation set, we introduce the following iterative process. In each vintage we (i) generate 24 training samples by subtracting $24-s+1$ observations with $s = 1, \dots, 24$, (ii) estimate all of the considered models and produce 24 pseudo out-of-sample nowcasts, (iii) calculate the RMSFE statistics using the subtracted data, (iv) choose the specification yielding the minimal RMSFE across all samples, (v) re-estimate the best model to produce the final nowcast using all available data in a given vintage. For the rolling window estimation, this process is slightly altered. In each vintage, we (i) choose in total the 84 most recent observations and generate 24 training samples with the rolling window of 60 observations. Next, we proceed with steps (ii)-(iv). As regards (v), we re-estimate the model on the 60 most recent observations to produce the nowcast.

¹⁴ We also investigate the accuracy of forecast combination with weights inversely proportional to the RMSFEs. We denote this model as *EC^{MCJ}* and study its nowcasting ability in Section 6.

¹⁵ Despite the early launch of our database, online data are available from December 2009. Whenever we use $o_{c,t}$ in model-based approaches in recursive estimation, we backcast the series using the inflation data lagged by 12 months. Though more refined frameworks towards imputing data are common (i.e., based on auxiliary regressions or state-space representations), we use this simple framework. This is motivated twofold. First, given the computational complexity of our nowcasting experiment, imputing the data for all inflation components in real-time using sophisticated models would prohibitively prolong the process of generating results. Second, our in-sample comparison shows that online prices precisely approximate the development in consumer prices as measured by the statistical office. Therefore, regression-based approaches would lead to similar results. We also indicate that we dispense with any data imputation for the rolling window approach.

Table 4
RMSFE of the competing models – expanding window estimation.

COICOP	Group	Frameworks with online prices			Traditional benchmarks				Forecast combinations		
		EC^{SX}	EC^{EX}	EC^{RT}	RW	AO^{SA}	BS^{IS}	BS^{OS}	BS^{MC}	EC^{MC}	EC^{MCI}
01.	Food and non-alcoholic beverages	0.338	1.071	1.082	1.784***	1.528***	1.446***	1.524***	1.621***	1.181**	1.176**
01.1.1	Bread and cereals	0.264	1.295**	1.338**	1.189**	1.153*	0.987	0.991	0.975	0.955	0.949
01.1.2	Meat	0.747	0.825	0.952	1.483***	1.285**	1.153***	1.151**	1.166***	1.062	1.061
01.1.3	Fish and seafood	0.319	2.070***	2.242***	1.621***	0.983	1.093	0.991	1.080	1.017	1.019
01.1.4	Milk, cheese and eggs	0.536	0.892	0.840	1.630	1.489**	1.400	1.349	1.411*	1.065	1.063
01.1.5	Oils and fats	1.237	0.818**	1.072	1.400***	1.155	1.083**	1.030	1.121**	1.071*	1.068*
01.1.6	Fruits	2.085	1.126	1.285	1.460***	1.224**	1.103*	1.136*	1.227**	1.088	1.070
01.1.7	Vegetables	1.737	1.011	0.976	2.432***	1.707***	1.655***	1.686***	1.734***	0.957	0.954
01.1.8	Sugar, jam, honey, chocolate and confectionery	0.421	1.272***	1.536***	1.573***	1.235**	1.050**	1.061	1.115	1.002	1.000
01.1.9	Food products n.e.c.	0.260	1.474**	1.783**	2.130***	1.101	1.088**	1.122	1.222**	1.066	1.061
01.2	Non-alcoholic beverages	0.234	1.845***	1.743***	1.940***	1.103	1.080	1.104	1.205**	1.048	1.046

Note: Table 4 reports the outcomes of the nowcasting exercise performed using estimation on an expanding window. Results are reported for food and non-alcoholic beverages inflation and its main subaggregates. Nowcast errors are calculated for the whole evaluation period, i.e., January 2017 - December 2020. RMSFE statistics for EC^{SX} are reported in levels. For the competing models, the RMSFE statistics are reported as ratios – a value above 1 (below 1) indicates that the competing model produces on average less (more) accurate nowcasts than EC^{SX} . Asterisks correspond to the outcomes of the two-sided Diebold and Mariano (1995) test, where *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level. For model definitions, the reader is referred to Section 4, whereas for the outcomes for all food and non-alcoholic beverages inflation components, the reader is referred to section S2 in the Supplementary Material (available online).

To assess the accuracy of the nowcasts, we use two common criteria: the mean forecasting error (MFE) and the root mean squared forecasting error (RMSFE):

$$MFE_c = (T - T^*)^{-1} \sum_{t=T^*+1}^T (p_{c,t} - \hat{p}_{c,t}) \tag{1}$$

$$RMSFE_c = \sqrt{(T - T^*)^{-1} \sum_{t=T^*+1}^T (p_{c,t} - \hat{p}_{c,t})^2}$$

where $p_{c,t}$ is the realization of inflation and $\hat{p}_{c,t}$ is the nowcast for the considered inflation component c . We compare the nowcasting accuracy of the models with the Diebold and Mariano (1995) test. Specifically, we assess the statistical significance of the differences in predictive accuracy between the baseline model EC^{SX} vis-à-vis a competing model by employing a standard, quadratic loss function. While we are aware that tests evaluating the predictive accuracy of models require a fixed window length and a rolling estimation strategy, we have decided to use an expanding window to account for the growing amount of online data and to study their usefulness from a practical approach.¹⁶ However, as a robustness check, we also perform the estimation on the rolling window.

5. Baseline results

We start the discussion of the results by presenting the outcomes of our nowcasting experiment performed using the recursive estimation scheme. In Table 4 we provide

RMSFE statistics for EC^{SX} in levels (column 1).¹⁷ For the remaining models, RMSFE statistics are reported as ratios, with the values above 1 indicating the lower average accuracy of the competing model. Statistical differences in the nowcast accuracy are examined with the use of the two-sided Diebold and Mariano (1995) test. This test checks the differences in the predictive accuracy between the EC^{SX} vis-à-vis a competing model. Table 5 summarizes the MFE statistics reported in levels.

Our results indicate that using online prices leads to a superior predictive ability in the nowcast. For the food inflation aggregate, the RMSFE statistic falls by 29% when using the pure, real-time estimate of the change in online prices with respect to the benchmark model. Moreover, employing the information on online prices into model frameworks leads to further improved accuracy. The EC^{SX} model outperforms the benchmark by 34.4%. Changes in the nowcast accuracy are statistically significant at $\alpha = 0.01$. Employing online prices for the nowcast also beats in a statistically significant manner the traditional benchmarks, the rigorous Atkeson and Ohanian (2001) model corrected for the seasonality patterns in the series, whose nowcast accuracy is the same as BS^{OS} . Moreover, the nowcasts from the EC^{SX} model are characterized by a smaller bias in comparison to other approaches. We also evidence that over the whole evaluation period nowcasts from the EC^{SX} model are more accurate than those prepared internally using judgment. However, the difference in the accuracy is significant only at $\alpha = 0.10$ (Table 6).

At a disaggregated level, the conclusions become more diverse. The EC^{SX} model provides more accurate nowcasts relative to BS^{OS} in terms of the RMSFE statistic for 52 out of 84 low-level food inflation components (62% cases).

¹⁶ A quick survey of the forecasting literature indicates that DM-type tests are frequently used to compare the predictive accuracy of models estimated with the use of an expanding window (e.g. Aparicio & Bertolotto, 2020; Kolasa & Rubaszek, 2015; Monteforte & Moretti, 2013; Szafranek, 2019; Tallman & Zaman, 2017, 2020).

¹⁷ In the paper, we report the results for food and non-alcoholic beverages inflation and its main subcomponents. Outcomes for all food and non-alcoholic beverages inflation components are provided in the Supplementary Material, available online (Section S2).

Table 5
MFE of the competing models – expanding window estimation.

COICOP	Group	Frameworks with online prices			Traditional benchmarks				Forecast combinations		
		EC^{SX}	EC^{EX}	EC^{RT}	RW	AO^{SA}	BS^{IS}	BS^{OS}	BS^{MC}	EC^{MC}	EC^{MCI}
01.	Food and non-alcoholic beverages	0.018	-0.106	-0.093	0.042	0.118	0.109	0.111	0.134	0.025	0.022
01.1.1	Bread and cereals	-0.031	-0.203	-0.191	-0.042	-0.001	-0.044	-0.024	-0.027	-0.011	-0.032
01.1.2	Meat	-0.052	-0.031	-0.034	-0.006	0.074	-0.016	-0.028	-0.011	-0.012	-0.012
01.1.3	Fish and seafood	0.005	-0.044	-0.042	-0.094	0.025	0.019	0.026	0.018	0.021	0.022
01.1.4	Milk, cheese and eggs	-0.061	-0.057	-0.052	0.006	0.035	-0.022	-0.018	-0.011	-0.040	-0.042
01.1.5	Oils and fats	0.054	-0.031	-0.009	-0.166	0.148	0.032	0.041	0.046	-0.001	0.000
01.1.6	Fruits	0.207	-0.332	-0.254	-0.007	0.299	0.358	0.312	0.290	0.142	0.152
01.1.7	Vegetables	0.226	-0.325	-0.259	0.755	0.606	0.827	0.826	1.006	0.174	0.176
01.1.8	Sugar, jam, honey, chocolate and confectionery	0.030	-0.032	-0.021	-0.123	0.023	0.044	0.082	0.078	0.034	0.033
01.1.9	Food products n.e.c.	-0.007	-0.028	-0.019	-0.089	0.011	0.005	-0.005	0.028	0.013	0.013
01.2	Non-alcoholic beverages	-0.023	-0.024	-0.021	-0.070	-0.009	-0.025	0.001	-0.009	-0.020	-0.020

Note: Table 5 reports the outcomes of the nowcasting exercise performed using recursive estimation on an expanding window. Results are reported for food and non-alcoholic beverages inflation and its main subaggregates. Nowcast errors are calculated for the whole evaluation period, i.e., January 2017 - December 2020. All MFE statistics are reported in levels. For model definitions, the reader is referred to Section 4, whereas for the outcomes for all food and non-alcoholic beverages inflation components, the reader is referred to section S2 in the Supplementary Material (available online).

Table 6
The comparison of nowcast errors for the EC^{SX} and JD approaches.

	Recursive estimation Whole evaluation period		Rolling estimation		Recursive estimation COVID-19 period	
	EC^{SX}	JD	EC^{SX}	JD	EC^{SX}	JD
	MFE	0.018	0.021	0.019	0.021	0.146
RMSFE	0.338	1.144*	0.427	0.905	0.340	0.965

Note: Table 6 reports the MFE and RMSFE statistics for the EC^{SX} and JD . Results are reported for food and non-alcoholic beverages inflation aggregate. The whole evaluation period refers to the sample January 2017 - December 2020, whereas the COVID-19 period refers to the sample January 2020 - December 2020. The MFE statistics are reported in levels. The RMSFE statistics for the EC^{SX} model are reported in levels, while for the JD approach, they are reported as a ratio – a value above 1 (below 1) indicates that the competing approach produces on average less (more) accurate nowcasts than EC^{SX} . Asterisks correspond to the outcomes of the two-sided Diebold and Mariano (1995) test, where *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level. For model definitions, the reader is referred to Section 4.

Table 7
The comparison between the nowcast errors obtained in the rolling and expanding window estimations.

		EC^{SX}	EC^{MC}	EC^{MCI}	BS^{IS}	BS^{OS}	BS^{MC}
RMSFE	Rolling window	0.427	0.428	0.412	0.581	0.589	0.565
	Expanding window	0.338	0.399	0.397	0.489	0.515	0.548
	Ratio	26%	7%	4%	19%	14%	3%
MFE	Rolling	0.019	0.080	0.019	0.041	0.017	0.025
	Expanding	0.018	0.025	0.022	0.109	0.111	0.134
	Difference	0.001	0.054	-0.003	-0.068	-0.093	-0.109

Note: Table 7 presents the comparison between the RMSFE and MFE errors obtained in the rolling and expanding window estimations. Results are reported for food and non-alcoholic beverages inflation aggregate. Nowcast errors are calculated for the whole evaluation period, i.e., January 2017 - December 2020. For the RMSFE statistics, the ratio denotes the percent change in the RMSFE when the expanding windows estimation is switched to the rolling windows. For the MFE, a simple difference is reported. For model definitions, the reader is referred to Section 4.

Moreover, in 19 cases, these differences are statistically significant at least at $\alpha = 0.1$. At the subaggregate level EC^{SX} beats the benchmark model in eight out of ten cases in terms of the RMSFE statistic (in four cases these differences are statistically significant). Thus, the results indicate that while the nowcasts from EC^{SX} are not consistently better at the elementary group level, they enable us to considerably improve the estimates for the subaggregates and the food inflation aggregate.

To illustrate the advantage online prices have upon our benchmark, we provide a simple bar chart in Fig. 4. On Panel A, we use a red bar to indicate a group for which the pure, real-time estimate of the change in online prices EC^{RT} provides a better nowcast than our benchmark BS^{OS} in terms of the RMSFE statistics. In turn, on Panel B we use a red bar to indicate these groups, for which the nowcast from the EC^{SX} model is more accurate in comparison to the nowcast from BS^{OS} . It is visible that the use of pure

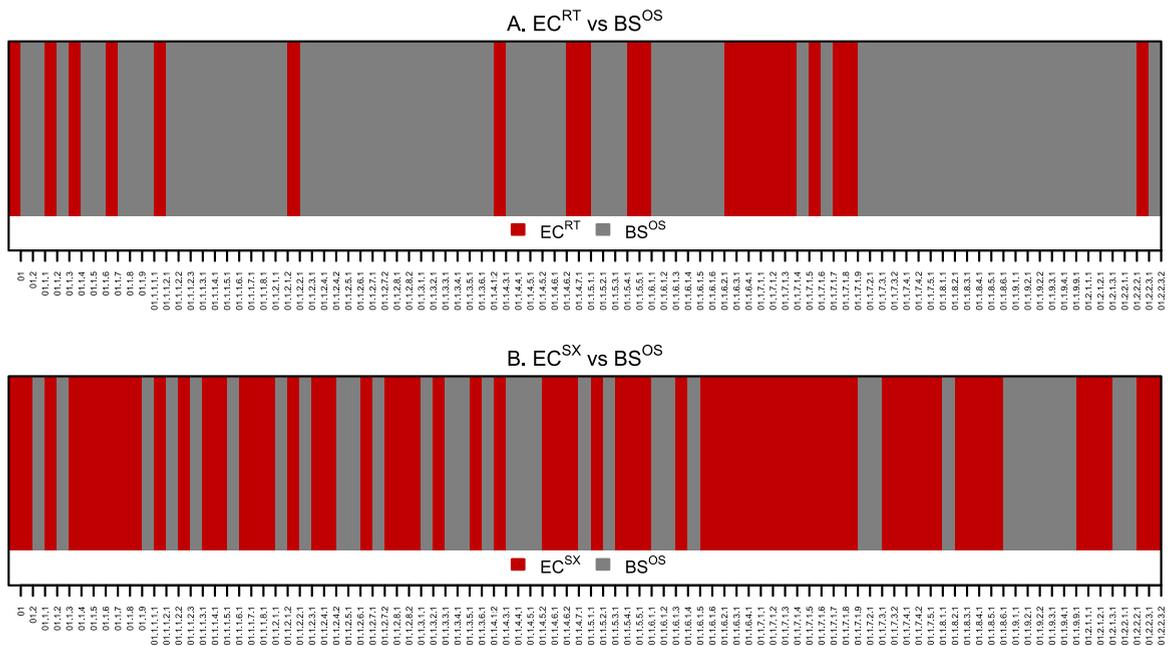


Fig. 4. The choice of the best model across all COICOP groups. Note: Fig. 4 illustrates the best model across all considered inflation components when the nowcast from EC^{RT} is compared against BS^{OS} (panel A) or the nowcast from EC^{SX} is compared against BS^{OS} (panel B). The grey bar indicates that the BS^{OS} framework yields a lower RMSFE statistic, whereas the red bar denotes that either the EC^{RT} model (panel A) or the EC^{SX} model (panel B) produces a lower RMSFE. Forecast errors are calculated for the whole evaluation period, i.e., January 2017–December 2020. For model definitions, the reader is referred to Section 4. (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this article).

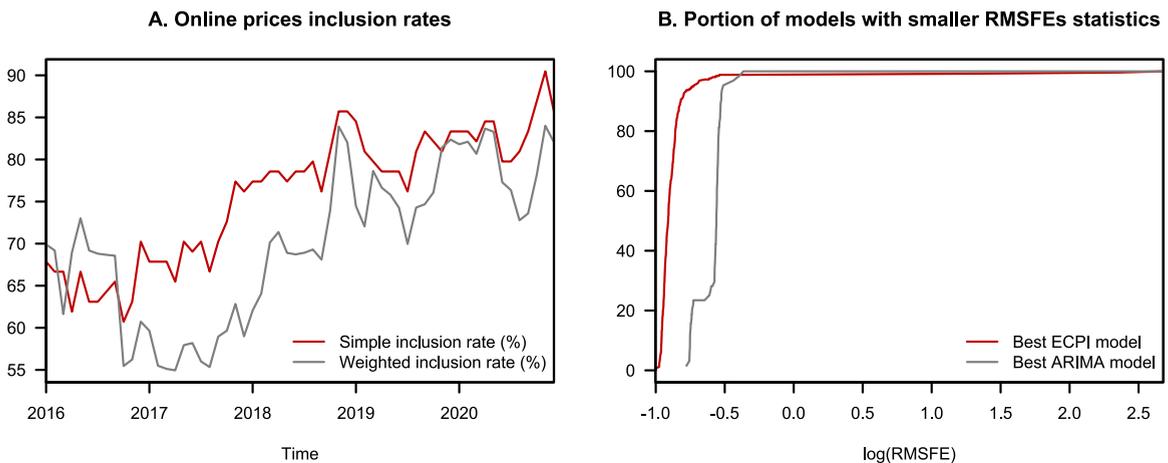


Fig. 5. Comparison of the nowcasting accuracy across components and models. Note: Fig. 5 presents the percentage of elementary groups for which forecasts with the E-CPI are more accurate than benchmarks in terms of the RMSFE statistics in time (panel A) and the cumulative portion of models with smaller RMSFE statistics (panel B). (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this article).

online estimates results in a better nowcast for the aggregate food inflation, but only for several components at the disaggregated level. However, once the online prices are incorporated into the model approach, the inferiority of the traditional model becomes more pervasive.

Our findings at the disaggregated level indicate that nowcasts based on scraped data are especially accurate for groups characterized with high price volatility (particularly unprocessed food, fruits, and vegetables). We

attribute this phenomenon to the fact that retailers are forced to change prices in a coordinated manner in the presence of a common shock on a competitive market. In turn, when there are moderate price swings, individual differences between retailers become more important (such as different suppliers, contracts, or the pricing policy). This is particularly vivid for the vegetable group (01.1.7), for which the benchmark generates 70% less accurate nowcasts with respect to the baseline model.

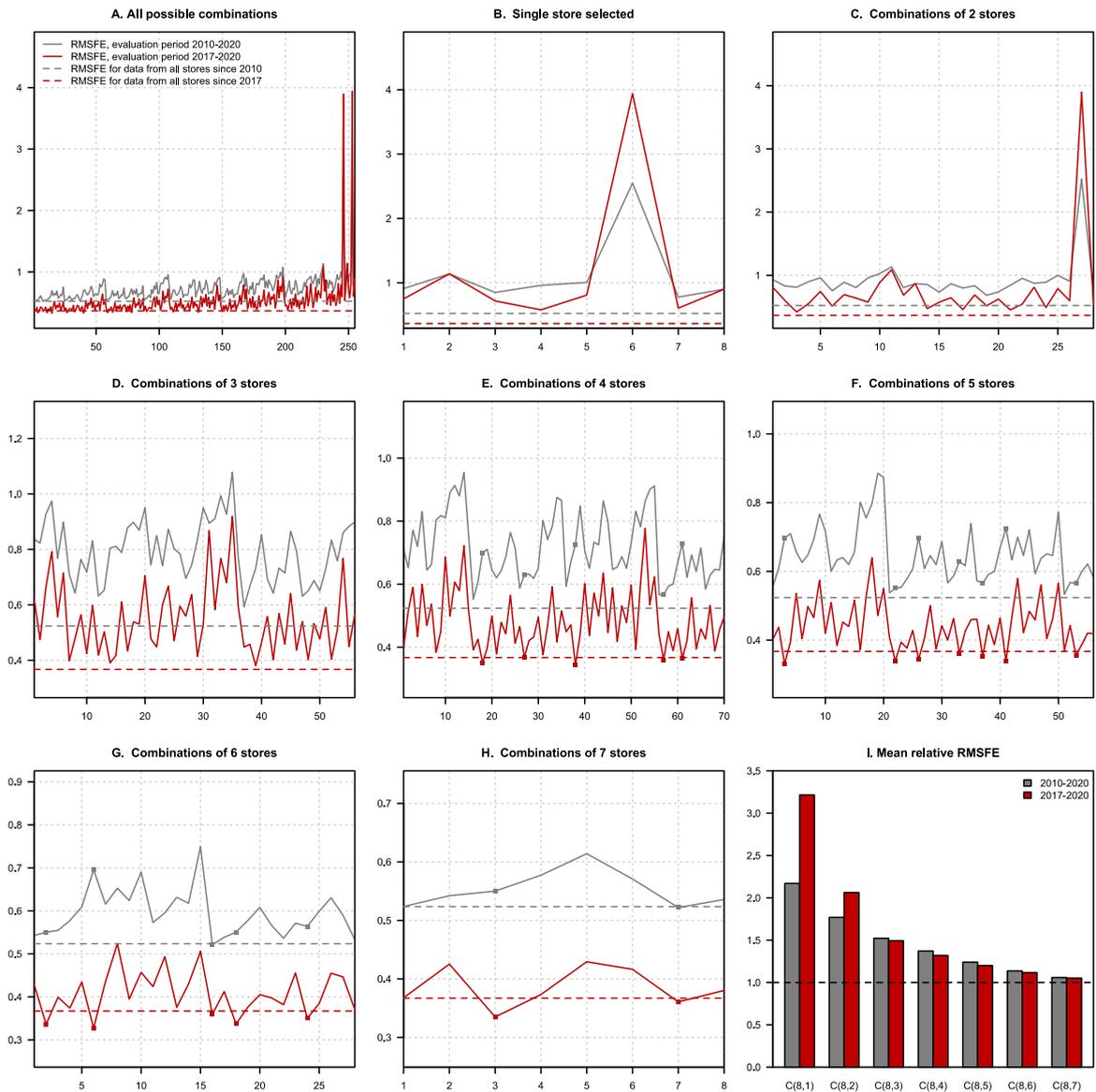


Fig. 6. The impact of sample composition on nowcast accuracy. Note: Fig. 6 presents the impact of the sample composition on the nowcast accuracy of the EC^{RT} model, measured with the RMSFE. The results are presented for all combinations (Panel A) and combinations of k stores (Panels B-H) for two evaluation periods: 2010–2020 and 2017–2020. Panel I shows the average relative RMSFE (with respect to the total sample) across different combinations of stores, indicated as $C(n, k)$, where $n = 8$ is the total number of stores and k is the number of stores selected. (For interpretation of the references to color in this figure legend, the reader is referred to the online version of this article).

Finally, we argue that the quality of models with on-line data is expected to increase over time. We denote the E-CPI inclusion rate as the percentage share in the elementary groups in which the model that includes the online data features a lower RMSFE than the best benchmark model. This rate is calculated every month, and it indicates whether utilizing online prices is preferred over time in a real-time forecasting experiment. This indicator is consistently above the 50% threshold and it tends to increase over time (Fig. 5, panel A), both when the ordinary share and the weighted share with CPI basket weights are considered. The simple inclusion rate evolved in the 60%–70% range in 2016 and most of 2017, increases

to over 80% in 2020, and ultimately reaches 90% in late 2020. The weighted inclusion rate increases over time but is more volatile as it is affected by transitory factors. This evidence further indicates the usefulness of online prices in inflation nowcasts.

6. Sensitivity analysis

In this section, we test the robustness of our results by performing three sensitivity analyses. Firstly, we check whether using the rolling scheme of estimation instead of the expanding window affects the quality of the nowcasts.

Secondly, we evaluate the accuracy of forecasts combinations. Finally, we inspect how the nowcasting errors evolved during the outbreak of the COVID-19 pandemic.

Results obtained on a rolling sample. In this robustness check, we apply a rolling window estimation scheme. Specifically, we use a short span of data by setting our window equal to 60 months. We report the nowcast error statistics in the same manner as for the baseline results in Tables S3 and S4 in the Supplementary Material, available online. Moreover, we compare the RMSFE and MFE statistics between the recursive and rolling scheme in Table 7.

For the food inflation aggregate, nowcasts from the EC^{SX} model still outperform traditional approaches by a significant margin. The increase in the accuracy with respect to the benchmark approach amounts to around 28% (similar gains are observed in comparison to other frameworks). Simultaneously, when working on a short sample, the nowcast based on a pure change in online prices delivers a further gain in the nowcast accuracy, albeit not statistically significant. We observe a large reduction in the RMSFE statistics for unprocessed food components at the disaggregated level. MFE statistics show that the EC^{SX} approach is negligibly biased.

In comparison to the results obtained in the recursive window experiment, our findings for the rolling scheme show that this framework leads to the deterioration in the nowcast quality in terms of the RMSFE statistic – both the online and the traditional frameworks are affected (Table 7). Specifically, the RMSFE statistics tend to be higher, from 4% to 26% for models that incorporate online prices and from 3% to 19% for traditional benchmarks. A possible explanation is that shorter time series are insufficient to estimate seasonality patterns precisely. They also may fail to extract additional predictive information from online price dynamics. However, the increase in the predictive error is less pronounced in the combination frameworks. This provides anecdotal evidence that specification selection bias is present, but it tends to manifest instead for the rolling rather than the expanding window approach. The rolling sample offers a smaller bias in terms of the MFE statistics for traditional benchmarks. As regards the comparison between the judgemental approach and the EC^{SX} model estimated on a rolling window, we show that the accuracy between these two approaches is comparable in the statistical sense, but the JD approach is slightly more accurate in terms of the RMSFE statistic (Table 6).

Considering these results, using the rolling window seems to be a futile approach to improve nowcasts as the smaller information set of online prices is a limiting factor, while the nowcast bias is already negligible. Therefore, we recommend using recursive estimation on samples for macroeconomic practitioners if possible.

Forecast combinations. In forecasting practice, combinations of competing models are expected to provide a reasonable forecast performance while minimizing the risk of selecting the underperforming approach (Timmermann, 2006). In this step, we inspect whether combining nowcasts from different models provide an additional gain in the nowcast accuracy. Thus, we focus on the

predictive performance of EC^{MC} and EC^{MCI} , which are nowcast combinations using equal weights and weights inversely proportional to RMSFEs obtained on the validation set, respectively. Forecast accuracy statistics for these approaches are presented in the last three columns in Table 4 in the case of RMSFE statistics and Table 5 in the case of MFE measures.

Surprisingly, for the recursive sample our results indicate that RMSFEs of forecast combinations are higher for most of the inflation components than the best performing individual model (selected in real-time). For the food inflation aggregate, both nowcast combinations are equally accurate and statistically outperformed by the EC^{SX} model. For the main aggregate, the RMSFEs of EC^{MC} and EC^{MCI} are 18% higher in comparison to EC^{SX} . Instead, for food inflation subaggregates, the deterioration in the nowcasts quality is smaller and does not exceed 8%. In this case, we observe that the aggregation across inflation components negatively impacts the overall accuracy of the aggregate inflation measure.

We conclude that in our case, the nowcast combination is inferior to the best individual model selected on a validation set. This result can be interpreted as follows. Given the already significant predictive power of online prices EC^{RT} , one needs to use an aggressive strategy to improve the nowcast further. We observe that specifications that accurately model both the seasonality of the official CPI and the impact of the online price dynamics are rare. However, these results should be interpreted with caution because, in a highly disaggregated framework, the performance of competing models is prone to be affected by outliers originating from seasonality or market shocks. While our results obtained for the recursive estimation provide an incentive to risk an individual model approach, our general view on forecast combination usefulness is rather inclusive as the rolling window experiment shows that the nowcast accuracy is comparable.

The usefulness of online prices in nowcasting food inflation during the COVID-19 pandemic. Lastly, we investigate the usefulness of online prices following the outburst of the COVID-19 pandemic in December 2019. To this end, we report the RMSFE and MFE statistics again, but only for the period January 2020 – December 2020 (Tables S5 and S6 in the Supplementary Material, available online). Given that this evaluation period accounts for only 12 nowcasts, we dispense with statistical testing of the differences in the predictive accuracy.

Our results indicate that the EC^{SX} model accounting for online prices produces, on average, much more accurate nowcasts than the benchmark model BS^{OS} for the food inflation aggregate. In relative terms, the gain in the nowcast accuracy increases to almost 47% from a 34% increase obtained when considering the whole evaluation period. For this aggregate, it is also visible that other approaches are characterized by much lower accuracy in terms of the RMSFE statistic (this measure increases from 30% to 100% across competing approaches). The bias of almost all models increases, but for frameworks relying on online prices, it remains very low. At the disaggregated level, we observe again that the relative gain in the nowcast accuracy occurs for groups characterized by

high volatility throughout the year (mostly unprocessed food, i.e., meat products, eggs, mostly domestic fruits, and vegetables). Nowcasting with online prices returns predictions of comparable quality with the judgemental framework (Table 6). It turns out that during the COVID-19 period, the accuracy of both these approaches has been essentially the same – both in terms of the bias and the dispersion measured by the RMSFE. Our reading of these results is that online prices have provided important information on the consumer price formation during COVID-19. However, given no statistical testing, we treat these outcomes only as anecdotal evidence and leave the evaluation of the usefulness of online prices following the outbreak of the COVID-19 to the reader.

7. The effect of data quantity and data quality on nowcast accuracy

In this section, we investigate the link between data quality, data quantity, and the nowcast accuracy. First, we study the impact of store composition on nowcast accuracy since we do not have any prior knowledge of the usefulness of data sources (stores). Second, we scrutinize the effect of data quality on our nowcasts by choosing different product classification approaches.

Store selection and nowcast accuracy. While the big datasets derived from the web are perceived to be very useful in empirical analysis, in the statistical strain of literature researchers highlight that there are also some limitations of big data, including the selectivity issues (Beręsewicz et al., 2018). Online datasets are typically very large, but they are usually non-probabilistic, resulting in potential biases and representativeness issues that may hinder the inference based on such data (Bethlehem, 2010). In the case of web scraping, a selection bias may result from the self-selection of the retailers that offer their products on the web – some retailers are just not present online. Moreover, even if all the major domestic retailers are represented in the online sample, the bias may emerge when the researcher decides which data sources to tap, given the limited resources. Therefore, the question of how the sample composition affects the nowcasting accuracy arises.

We address this issue empirically. Given the fact that we are unable to assess the usefulness of a particular store ex-ante, we calculate nowcasts using EC^{RT} approach on datasets that comprise every possible combination of retailers. Hence, for the total number of 8 stores we construct 255 k -combinations of subsamples (k among 8, where $k = 1, 2, \dots, 8$). We do not perform any estimation; we use our standard evaluation period (January 2017 - December 2020) and our entire sample (January 2010 - December 2020).

We report aggregate results in Table 8 and plot them on Fig. 6. Our simulation reveals that on average larger samples offer higher nowcast accuracy as the average relative RMSFE decreases with the increase in the number of stores selected in the sample and the total number of products included. Moreover, the marginal gain tends to fall as the number of stores in the sample increases.

Table 8
The impact of sample composition on nowcast accuracy.

k	Average relative RMSFE		N	\bar{q}_s
	2017–2020	2010–2020		
C(8,1)	3.21	2.17	8	20.0
C(8,2)	2.06	1.77	28	39.9
C(8,3)	1.49	1.52	56	59.9
C(8,4)	1.32	1.37	70	79.9
C(8,5)	1.20	1.24	56	99.8
C(8,6)	1.12	1.14	28	119.8
C(8,7)	1.05	1.06	8	139.7
C(8,8)	1.00	1.00	1	159.7

Note: $C(n, k)$ denotes the combination of stores taken into account where $n = 8$ is the total number of stores and k is the number of stores selected. Average relative RMSFE is calculated with respect to the total sample ($n = k = 8$) and its value above 1 indicates that the k -combination provides a higher RMSFE than the total sample. N denotes the number of combinations in the group, whereas \bar{q}_s is the average number of products in the group (in thousand).

The mean relative RMSFE does not fall below one, which indicates that using as much data as possible is preferred.

Detailed results across all possible combinations plotted in Fig. 6 show that there is a very limited number of combinations providing slightly better results in comparison to the full sample when the 2017–2020 evaluation period is considered (Panels B–H, marked with a red square). Moreover, when we check their performance on our entire sample of data (2010–2020, marked with a grey square), we notice that they fail to provide a generous value added. We also show that the nowcast accuracy tends to be very low, especially when a small subset of stores is considered. We conclude that data quantity coupled with multiple data sources affect the nowcast accuracy. Next, we investigate whether more data is always better.

Does product classification affect nowcast accuracy? To

answer this question, we reproduce the entire nowcasting experiment, but we switch off our exact classification of products. Instead, we use an approximate, unsupervised approach based on dictionaries of keyword stems.¹⁸ Subsequently, we use official weights to aggregate the inflation components. Thus, we illustrate whether an approximate but fully automatic approach is sufficient to provide high quality nowcasts similar to those produced with precise human annotation. We provide results for the food inflation aggregate in a similar manner as in the baseline approach in Table 9.

Our results indicate that dispensing with this data curation step severely affects the quality of the nowcasts. We show that when using an approximate approach, the nowcast based on the change in the online prices index EC^{RT} is 40% less accurate than our baseline model EC^{SX} , a difference statistically significant at $\alpha = 0.01$. Moreover, in this framework, we observe a significant increase in

¹⁸ Section S1 in the Supplementary Material (available online) provides details on the approaches towards product classification.

Table 9
RMSFE and MFE of the competing models for food and non-alcoholic beverages inflation.

		RMSFE	MFE
Frameworks with online prices	EC^{SX}	0.338	0.018
	EC^{EX}	1.071	-0.106
	EC^{RT}	1.082	-0.093
Approximate classification	EC^{SX}	1.299**	0.036
	EC^{RT}	1.400***	-0.117
Traditional benchmarks	RW	1.784***	0.042
	AO^{SA}	1.528***	0.118
	BS^{IS}	1.446***	0.109
	BS^{OS}	1.524***	0.111
Forecast combinations	BS^{MC}	1.621***	0.134
	EC^{MC}	1.181**	0.025
	EC^{MCI}	1.176**	0.022

Note: Table 9 reports the outcomes of the nowcasting exercise performed using estimation on an expanding window. Forecast errors are calculated for the whole evaluation period, i.e., January 2017 - December 2020. RMSFE statistics for EC^{SX} are reported in levels. For the competing models, the RMSFE statistics are reported as ratios – a value above 1 (below 1) indicates that the competing model produces on average less (more) accurate nowcasts than EC^{SX} . Approximate classification refers to the robustness check, where products are classified into respective groups using the approximate approach based on dictionaries of word stems. Asterisks correspond to the outcomes of the two-sided Diebold and Mariano (1995) test, where *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level. For model definitions, the reader is referred to Section 4.

bias. In turn, in the model approach EC^{SX} the loss in the accuracy – when approximate classification is considered – falls slightly, but still using exact classification outperforms the approximate approach by a large, significant margin (around 30%). Interestingly, despite the imperfect classification of products, this framework is already able to beat traditional benchmarks, including the Atkeson and Ohanian (2001) model in terms of the RMSFE accuracy while maintaining a low bias.

We conclude that proper data curation is of paramount importance for the quality of nowcasts. Our reading of these results is that both the quantity and quality of data determine the success of the nowcast, especially when the problem complexity rises.¹⁹ Put differently, we argue that more data is not always better.

8. Conclusions

Online prices have recently gained a lot of interest in the empirical literature. The spread of new data collection methods such as web scraping offers a unique opportunity to utilize highly disaggregated and timely data in economic policy analysis and forecasting. Simultaneously, we have observed a rapid increase in the e-commerce market. This development has been strikingly vivid after the outbreak of the COVID-19 pandemic, when traditional stores closed due to the introduction of severe

confinement measures for several weeks and e-commerce experienced a boom. Both these phenomena have confronted researchers and policymakers with the availability of new sources of information ready to be tapped.

In this paper, we assess the ability of online prices in the nowcasting of food inflation. To this end, we use a massive dataset of online prices collected at Narodowy Bank Polski from the websites of online retailers in Poland since the beginning of 2009. We provide evidence for a broad spectrum of CPI components, focusing on the food inflation aggregate, its 84 components and ten sub-aggregates and ten subaggregates. We contribute to the rapidly growing literature on the usefulness of online prices across four margins.

First, we show that online prices can markedly improve nowcasts of food inflation. Our results show that including the information on online prices in the model leads to a substantial and statistically significant reduction in the nowcasting errors with respect to popular benchmarks. We report that the food nowcast errors from our recursively optimized model are around 34% smaller in comparison both to the benchmark, which does not account for online data, and the rigorous Atkeson and Ohanian (2001) nowcast corrected for seasonality patterns. Moreover, the bias of the nowcast remains negligible. For macroeconomic practitioners considering launching similar projects, we show that even without a model-based approach, the researcher can improve the accuracy of the nowcast by around 30% by utilizing real-time estimates of price changes based on online data. Therefore, we argue that the advantages of using online prices can be exploited very quickly, already after several months of collecting data via web scraping. Our work provides complementary evidence to Aparicio and Bertolotto (2020) that scraping data is also beneficial for emerging economies where the expansion of the e-commerce market is still underway.

Second, we demonstrate that both data quantity and data quality matter for the precision of the nowcast. We show that larger samples of online prices offer, on average, higher nowcast accuracy. Hence, macroeconomic practitioners should tap multiple data sources as relying on prices from a few stores may be insufficient to provide accurate nowcasts. However, we also stress the importance of data quality. Specifically, we show that proper handling of scraped data is essential for the nowcasting accuracy. Dispensing with exact product classification and a subsequent expenditure weighing in line with the methodology provided by the statistical office is detrimental to the nowcasting accuracy. Thus, we recommend a tedious approach towards classifying products in line with the consumption basket and representative goods for each component of the CPI. While obtaining scraped data is relatively fast and inexpensive, properly selecting goods to approximate the consumption basket necessitates machine-assisted human classification, essential for accurate nowcasts.

Third, we provide mixed evidence on the usefulness of nowcast combinations. In our recursive estimation combining nowcasts from multiple models provides an inferior nowcast with respect to our baseline approach.

¹⁹ This observation for large datasets is well described in statistics. For instance, Meng (2018) shows that estimation errors may be decomposed into three (multiplicative) components: (i) data quality, measuring the sign and degree of the selection bias, (ii) data quantity, and (iii) problem complexity, measured with standard deviation.

However, when we perform a rolling experiment on a short sample, these two approaches become comparable in terms of the RMSFE statistics.

Finally, we argue that online prices have proven useful after the outbreak of the COVID-19 pandemic. Our results indicate that during 2020 the accuracy of our baseline model has increased with respect to the benchmark in terms of the RMSFE statistic and remained comparable with the judgemental approach. While we perceive this instead as anecdotal evidence, we feel safe to conclude that following the COVID-19 outbreak, online prices have proved helpful in nowcasting changes in inflation deviating from their seasonal trends.

The paper provides interesting insights into the forecasting process using online prices at the highly disaggregated level and valuable practical information regarding implementing such an approach. However, it does not discuss several issues related to our future research. First, we do not consider online data for other components of the CPI basket. This is because data collection, selection, and curation processes for other goods are far more exhaustive, and we have very limited human resources. We currently extend our database to include prices of other goods and services. Second, after the outbreak of the COVID-19 pandemic, we have begun to scrutinize the availability of products. While the literature already provides some approaches towards measuring the quantity sold (Gorodnichenko et al., 2018a), this remains an open research question. Third, in practical (business) approaches, product classification may be done using fully automatic machine learning frameworks. Finally, we plan to study the price characteristics to learn more about their distinct micro features.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank two anonymous referees and the Editor for their valuable comments, which substantially improved the quality of the paper. The authors are grateful for the remarks shared by the members of the ESCB Research Network PRISMA (Price-setting Microdata Analysis) and also the participants of the 2021 International Symposium on Forecasting, the 2021 Conference on Non-traditional Data, Machine Learning and Natural Language Processing in Macroeconomics, the 9th Summer Workshop of Narodowy Bank Polski and the 2017 Narodowy Bank Polski Workshop on Forecasting. Earlier drafts of the paper benefited from comments by Grzegorz Szafranski. The authors would also like to thank Aneta Błażejowska for help in maintaining the database daily. The data supporting the study findings are collected within the project 'E-CPI' by Narodowy Bank Polski. The dissemination of the dataset is limited. Computations have been performed in SAS, Matlab, Python, and R. The

views expressed in this paper are those of their authors and do not necessarily correspond to the opinions of Narodowy Bank Polski. All remaining errors are the authors' sole responsibility.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.ijforecast.2022.02.007>.

References

- Ang, A., Bekaert, G., & Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, 54(4), 1163–1212.
- Aparicio, D., & Bertolotto, M. I. (2020). Forecasting inflation with online prices. *International Journal of Forecasting*, 36(2), 232–247.
- Aparicio, D., Metzman, Z., & Rigobon, R. (2021). The pricing strategies of online grocery retailers. In *Working Paper Series*, (28639), National Bureau of Economic Research.
- Atkeson, A., & Ohanian, L. E. (2001). Are phillips curves useful for forecasting inflation? *Quarterly Review*, 25(Win), 2–11.
- Belmonte, M. A., Koop, G., & Korobilis, D. (2014). Hierarchical shrinkage in time-varying parameter models. *Journal of Forecasting*, 33(1), 80–94.
- Bereśiewicz, M., Lehtonen, R., Reis, F., Di Consiglio, L., & Karlberg, M. (2018). An overview of methods for treating selectivity in big data sources. Publication Office of the European Union, Luxembourg.
- Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, 78(2), 161–188.
- Binner, J. M., Elger, C. T., Nilsson, B., & Tepper, J. A. (2006). Predictable non-linearities in U.S. inflation. *Economics Letters*, 93(3), 323–328.
- Błażejowska, A., & Szafranek, K. (2020). Growth of eCommerce in Poland and new challenges. <https://www.obserwatorfinansowy.pl/in-english/business/growth-of-e-commerce-in-poland-and-new-challenges/>. (Accessed: 12 December 2020).
- Cavallo, A. (2013). Online and official price indexes: Measuring Argentina's inflation. *Journal of Monetary Economics*, 60(2), 152–165.
- Cavallo, A. (2017). Are online and offline prices similar? Evidence from large multi-channel retailers. *American Economic Review*, 107(1), 283–303.
- Cavallo, A. (2018). Scraped data and sticky prices. *The Review of Economics and Statistics*, 100(1), 105–119.
- Cavallo, A., & Rigobon, R. (2016). The billion prices project: Using online prices for measurement and research. *Journal of Economic Perspectives*, 30(2), 151–178.
- Diebold, F., & Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263.
- Duarte, C., & Rua, A. (2007). Forecasting inflation through a bottom-up approach: How bottom is bottom? *Economic Modelling*, 24(6), 941–953.
- Edge, R. M., & Gurkaynak, R. S. (2010). How useful are estimated DSGE model forecasts for central bankers? *Brookings Papers on Economic Activity*, 41(2 (Fall)), 209–259.
- Eickmeier, S., & Ziegler, C. (2008). How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting*, 27(3), 237–265.
- Eurostat (2020a). Guidance note for NSIs on the replacement and imputations techniques to be used for a harmonised approach across EU member states. <https://ec.europa.eu/eurostat/web/hicp/methodology>. (Accessed: 26 August 2021).
- Eurostat (2020b). Science, technology, digital society. Individuals - internet use. <https://ec.europa.eu/eurostat/data/database>. (Accessed: 10 December 2020).
- Eurostat (2021). Information on imputations made related to Covid-19. <https://ec.europa.eu/eurostat/web/hicp/methodology>. (Accessed: 26 August 2021).
- Faust, J., & Wright, J. H. (2013). Chapter 1 - Forecasting inflation. In G. Elliott, & A. Timmermann (Eds.), *Handbook of Economic Forecasting: vol. 2, Part A, Handbook of Economic Forecasting* (pp. 2–56). Elsevier.

- Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2003). Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, 50(6), 1243–1255.
- Genre, V., Kenny, G., Meyler, A., & Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting*, 29(1), 108–121.
- Goolsbee, A. D., & Klenow, P. J. (2018). Internet rising, prices falling: Measuring inflation in a world of E-commerce. *AEA Papers and Proceedings*, 108, 488–492.
- Gordnichenko, Y., Sheremirov, V., & Talavera, O. (2018a). Price setting in online markets: Does IT click? *Journal of the European Economic Association*, 16(6), 1764–1811.
- Gordnichenko, Y., Sheremirov, V., & Talavera, O. (2018b). The responses of internet retail prices to aggregate shocks: A high-frequency approach. *Economics Letters*, 164(C), 124–127.
- Gordnichenko, Y., & Talavera, O. (2017). Price setting in online markets: Basic facts, international comparisons, and cross-border integration. *American Economic Review*, 107(1), 249–282.
- Groen, J. J., Paap, R., & Ravazzolo, F. (2013). Real-time inflation forecasting in a changing world. *Journal of Business & Economic Statistics*, 31(1), 29–44.
- Inoue, A., & Kilian, L. (2008). How useful is bagging in forecasting economic time series? A case study of U.S. consumer price inflation. *Journal of the American Statistical Association*, 103(482), 511–522.
- Jaravel, X., & O'Connell, M. (2020). Real-time price indices: Inflation spike and falling product variety during the great lockdown. *Journal of Public Economics*, 191, Article 104270.
- Jaworski, K. (2021). Measuring food inflation during the COVID-19 pandemic in real time using online data: a case study of Poland. *British Food Journal*, 123(13), 260–280.
- Jochmann, M., Koop, G., & Strachan, R. W. (2010). Bayesian forecasting using stochastic search variable selection in a VAR subject to breaks. *International Journal of Forecasting*, 26(2), 326–347, Special Issue: Bayesian Forecasting in Economics.
- Kabukcuoğlu, A., & Martínez-García, E. (2018). Inflation as a global phenomenon—Some implications for inflation modeling and forecasting. *Journal of Economic Dynamics & Control*, 87, 46–73.
- Kichian, M., & Rumler, F. (2014). Forecasting Canadian inflation: A semi-structural NKPC approach. *Economic Modelling*, 43, 183–191.
- Knotek, E. S., II, & Zaman, S. (2017). Nowcasting U.S. headline and core inflation. *Journal of Money, Credit and Banking*, 49(5), 931–968.
- Kolasa, M., & Rubaszek, M. (2015). Forecasting using DSGE models with financial frictions. *International Journal of Forecasting*, 31(1), 1–19.
- Koop, G., & Korobilis, D. (2012). Forecasting inflation using dynamic model averaging*. *International Economic Review*, 53(3), 867–886.
- Korobilis, D. (2017). Quantile regression forecasts of inflation under model uncertainty. *International Journal of Forecasting*, 33(1), 11–20.
- Lamboray, C. (2021). Index compilation techniques for scanner data - an overview. https://unece.org/sites/default/files/2021-05/Session_1_Eurostat_Paper.pdf. (Accessed: 26 August 2021).
- Lünnemann, P., & Wintr, L. (2011). Price stickiness in the US and europe revisited: Evidence from internet prices. *Oxford Bulletin of Economics and Statistics*, 73(5), 593–621.
- Macias, P., & Stelmasiak, D. (2019). Food inflation nowcasting with web scraped data. NBP Working Paper No. 302.
- Mandalinci, Z. (2017). Forecasting inflation in emerging markets: An evaluation of alternative models. *International Journal of Forecasting*, 33(4), 1082–1104.
- McAdam, P., & McNelis, P. (2005). Forecasting inflation with thick models and neural networks. *Economic Modelling*, 22(5), 848–867.
- McKnight, S., Mihailov, A., & Rumler, F. (2020). Inflation forecasting using the new keynesian phillips curve with a time-varying trend. *Economic Modelling*, 87, 383–393.
- Medeiros, M. C., Vasconcelos, G. F. R., Veiga, A., & Zilberman, E. (2019). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics*, 1–22.
- Meng, X.-L. (2018). Statistical paradises and paradoxes in big data (1): Law of large populations, big data paradox, and the 2016 US presidential election. *The Annals of Applied Statistics*, 12(2), 685–726.
- Modugno, M. (2013). Now-casting inflation using high frequency data. *International Journal of Forecasting*, 29(4), 664–675.
- Monteforte, L., & Moretti, G. (2013). Real-time forecasts of inflation: The role of financial variables. *Journal of Forecasting*, 32(1), 51–61.
- Nakamura, E. (2005). Inflation forecasting using a neural network. *Economics Letters*, 86(3), 373–378.
- ONS (2020). Web scraping policy. <https://www.ons.gov.uk/aboutus/transparencyandgovernance/datastrategy/datapolicies/webscrapingpolicy>. (Accessed: 10 December 2020).
- Powell, B., Nason, G., Elliott, D., Mayhew, M., Davies, J., & Winton, J. (2018). Tracking and modelling prices using web-scraped price microdata: towards automated daily consumer price index forecasting. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 181(3), 737–756.
- Rapach, D. E., & Strauss, J. K. (2010). Bagging or combining (or both)? An analysis based on forecasting U.S. employment growth. *Econometric Reviews*, 29(5–6), 511–533.
- Rumler, F., & Valderrama, M. T. (2010). Comparing the New Keynesian Phillips Curve with time series models to forecast inflation. *The North American Journal of Economics and Finance*, 21(2), 126–144, Special Issue: 50 Years of the Phillips Curve.
- Statista (2021). eCommerce Poland. <https://www.statista.com/outlook/243/146/ecommerce/poland>. (Accessed: 26 August 2021).
- Stelmasiak, D., & Szafranski, G. (2016). Forecasting the polish inflation using Bayesian VAR models with seasonality. *Central European Journal of Economic Modelling and Econometrics*, 8(1), 21–42.
- Stock, J. H., & Watson, M. W. (1999). Forecasting inflation. *Journal of Monetary Economics*, 44(2), 293–335.
- Stock, J. H., & Watson, M. W. (2006). Chapter 10 forecasting with many predictors. In G. Elliott, C. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting: vol. 1*, (pp. 515–554). Elsevier.
- Stock, J. H., & Watson, M. W. (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking*, 39, 3–33.
- Szafranek, K. (2019). Bagged neural networks for forecasting Polish (low) inflation. *International Journal of Forecasting*, 35(3), 1042–1059, Forecasting issues in developing economies.
- Szafranek, K., & Halka, A. (2019). Determinants of low inflation in an emerging, small open economy through the lens of aggregated and disaggregated approach. *Emerging Markets Finance and Trade*, 55(13), 3094–3111.
- Tallman, E. W., & Zaman, S. (2017). Forecasting inflation: Phillips curve effects on services price measures. *International Journal of Forecasting*, 33(2), 442–457.
- Tallman, E. W., & Zaman, S. (2020). Combining survey long-run forecasts and nowcasts with BVAR forecasts using relative entropy. *International Journal of Forecasting*, 36(2), 373–398.
- Timmermann, A. (2006). Forecast combinations. In G. Elliott, C. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting: vol. 1*, *Handbook of economic forecasting* (pp. 135–196). Elsevier, Chapter 4.
- Villani, M. (2009). Steady-state priors for vector autoregressions. *Journal of Applied Econometrics*, 24(4), 630–650.