International Journal of Forecasting 36 (2020) 232–247

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Forecasting inflation with online prices

Diego Aparicio [a], Manuel I. Bertolotto [b,*]

[a] IESE Business School, Spain
[b] Universidad de San Andrés, Department of Economics, 284 Vito Dumas, Victoria, Buenos Aires, 1644, Argentina

ARTICLE INFO

Keywords:
Inflation forecasting
Online prices
Survey of professional forecast
High frequency
Monthly forecast
Quarterly forecast

ABSTRACT

Are survey-based forecasts unbeatable? They are not. This paper uses online price indices to forecast the Consumer Price Index. We find that online price indices anticipate changes in official inflation trends more than one month in advance. Our baseline one-month forecast outperforms Bloomberg surveys of forecasters, which only predict the contemporaneous inflation rate. Our baseline specification also outperforms statistical benchmark forecasts for Australia, Canada, France, Germany, Greece, Ireland, Italy, the Netherlands, the United Kingdom, and the United States. Similarly, our quarterly forecast for the US inflation rate substantially outperforms the Survey of Professional Forecasters.

1. Introduction

In 2016, the word inflation was mentioned in over 250,000 news articles, which amounts to more than 28 articles per hour.[1] It is hardly questionable that improving our abilities to anticipate the inflation rate can benefit a broad spectrum of relevant decisions, from macroeconomic policies to hedging strategies, and to even financial decisions of individual households. For instance, inflation forecasts are key elements in the design of monetary policy (Bernanke, 2007; Yellen, 2017). New Keynesian models of optimal nominal short-term rate also depend on expected inflation (Clarida, Gali, & Gertler, 1999; Woodford, 2011). Even asset prices vary based on unexpected inflation rate news (Gurkaynak, Sack, & Swanson, 2004; Rigobon & Sack, 2004).[2] Our current projection abilities, however, are still limited and our forecast errors are too large.[3]

Survey-based inflation forecasts, which average predictions from multiple economics experts, have been considered the best inflation forecasting tools since Ang, Bekaert, and Wei (2007)'s work on the subject and were further reinforced by Faust and Wright (2013). Other forecasting methods were held back because they depended on regressions that used variables which accurately predicted the inflation rate in some stages of the economic cycle but not in others.[4] In this paper, we find that online prices collected from retailer websites are even more effective tools for predicting the future level of the Consumer Price Index (CPI) than previous models. Online price indices use data that is representative of retailer transactions and screen, on average, over 700,000 varieties of products on any given date. Additionally, these indices track consumer prices from mostly the same economic sectors as the traditional CPI, which allows them

* Corresponding author.
E-mail addresses: dapa@mit.edu (D. Aparicio), mbertolotto@udesa.edu.ar (M.I. Bertolotto).

1 Retrieved from Google News as of August 2018.

2 See, for example, the movements in the treasury yields described in Boesler (2016).

3 For instance, the mean absolute error of the one quarter inflation forecast annualized) from the median Survey of Professional Forecasters (SPF) in the United States is around 1.4 percent when the CPI fluctuates in the −1.4 to 2.2 percent interval 70 percent of the time (estimates using data from the first quarter of 2009 through the second quarter of 2016).

4 See Stock and Watson (2003) for a detailed explanation.

https://doi.org/10.1016/j.ijforecast.2019.04.018
0169-2070/

to closely move together with the CPI despite economic fluctuations.

We first forecast the CPI inflation rate one month in advance and subsequently calculate the two- and three-month-ahead forecasts. We find that parsimonious models using online price indices are more accurate than the survey of professional forecasters published by Bloomberg. In addition, these online models are more accurate than traditional benchmarks such as the AR(1), the random walk (RW), and the Phillips curve. These findings are robust to multiple specifications, and across all ten countries in our data, namely the United States (USA or US), United Kingdom (UK), Australia, Canada, France, Germany, Greece, Ireland, Italy, and the Netherlands. To our knowledge, this is the first study to show that online prices are a good predictor of changes in the CPI in multiple economies. Details about these series, as well as their advantages, are explained in Section 2.

These online indices are calculated at a daily frequency and published with a three-day lag. By the end of any month, we already know the monthly inflation rate in the online market, two weeks before the statistical office's publication of the CPI. These advantages in both timing and frequency boost the predictive value of the online series. However, we show that even after removing these advantages online indices still predict the CPI more accurately than the forecasting benchmarks. We provide preliminary explanations for why this anticipation might occur and discuss possible areas for future research.

Finally, we calculate a quarterly inflation forecast for the United States and evaluate its performance against the Survey of Professional Forecasters (SPF). We find that models using online price data substantially outperform this survey, which is considered one of the leading forecasting benchmarks in the literature.[5]

The paper is related to a vast body of literature about inflation forecasting. The methods employed in this field are diverse, ranging from VARs (Ang et al., 2007; Clements & Galvão, 2013; Marcellino, Stock, & Watson, 2003), dynamic factor models (Eickmeier & Ziegler, 2008; Forni, Hallin, Lippi, & Reichlin, 2003; Stock & Watson, 2006), artificial neural networks (Chen, Racine, & Swanson, 2001; Nakamura, 2005; Stock & Watson, 1999a), Bayesian methods (Wright, 2009), Phillips curve forecasts (Atkeson & Ohanian, 2001; Stock & Watson, 1999b), to survey-based forecasts (Ang et al., 2007; Bates & Granger, 1969; Cecchetti, Hooper, Kasman, Schoenholtz, & Watson, 2007; Croushore, 2010). The list is far from exhaustive and exemplifies how active and relevant the field remains. Comprehensive literature reviews on the subject can be found in Faust and Wright (2013) and Stock and Watson (2009).

Our work is also related to the rapidly growing field of literature using online prices to research diverse economic topics. The ability to collect thousands of product's prices at a fine granularity, from across the globe, and on a daily basis provides researchers unprecedentedly rich datasets to re-examine both macro- and microeconomic theories. Several recent papers have studied the relationship between online and offline markets, including Cavallo (2019), Gorodnichenko and Talavera (2017) and Lünnemann and Wintr (2011). However, to our knowledge, there has been no paper that uses online prices to forecast official inflation measurements. The closest examples are Bertolotto (2018) and Cavallo (2013) who show that online indices approximate the levels and trend dynamics of official inflation rates in multiple countries; and Aparicio and Cavallo (2018) and Cavallo and Rigobon (2016) who use online prices to measure country-level inflation rates. These result can also be understood as a validation of the notion that online-based price indices are comparable to the traditional CPI despite their methodological differences.

Lastly, the daily frequency of online indices studied in this paper relates to a strand of literature using other high-frequency market-based indicators of inflation expectations, as in Gürkaynak, Sack, and Wright (2010) or Haubrich, Pennacchi, and Ritchken (2012), and to models that combine different data frequencies such as Ghysels, Santa-Clara, and Valkanov (2004), Knotek and Zaman (2015), Modugno (2013) and Monteforte and Moretti (2013).

The paper is organized as follows. Section 2 describes the datasets used, and Section 3 argues that online indices should improve our forecasting accuracy. Section 4 describes the forecasting methodologies, and presents the monthly and quarterly out-of-sample inflation forecasts. Section 5 discusses potential reasons for why online prices can anticipate the traditional CPI. Section 6 concludes.

2. The data

This section describes the datasets used to evaluate the forecasting accuracy of online prices in Australia, Canada, France, Germany, Greece, Ireland, Italy, the Netherlands, the UK, and the US from July 2008 to September 2016.

We are interested in forecasting the monthly non-seasonally adjusted Consumer Price Indices, for all urban consumers, calculated by each country's national statistical office (NSO). In the case of Australia, where the Bureau of Statistics does not calculate a monthly CPI, we use their quarterly measurements instead.

We use online indices provided by *PriceStats*, a private company that spun off from the Billion Prices Project (BPP)[6] at MIT. These series have been designed to measure each country's aggregate inflation rate. Their daily frequency is an advantage over other measures used to forecast the CPI because it helps capture changes in inflation trends before the end of the month. A second advantage is their immediate release. Statistical offices usually publish their new CPI measurements for each month around the 15th day of the following month. On the other hand, the monthly inflation rate for the online price index is available immediately, giving it a 15-day

[5] See Ang et al. (2007) and Faust and Wright (2013) for further details.

[6] See http://www.thebillionpricesproject.com/ and Cavallo and Rigobon (2016) for additional details on the BPP and the scraping methodology.

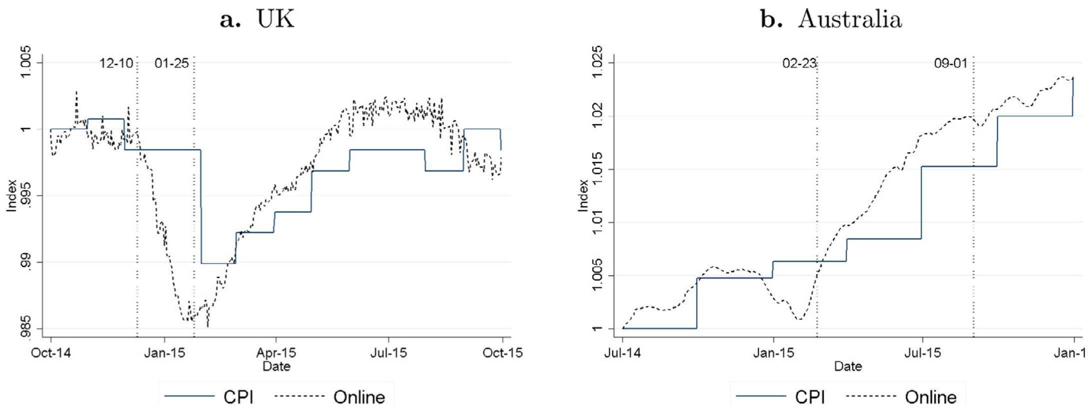

Fig. 1. Online index anticipation. *Note:* This figure provides a graphical illustration of the anticipatory effect of the online price indices. *Source:* Authors using online indices computed by *PriceStats* and the Consumer Price Indices, non-seasonally adjusted, all items, from the National Statistical Offices of the UK and Australia.

lead over the official release. These advantages are more pronounced in countries like Australia, where the official inflation rate is released once per quarter instead of once per month. Even though the online index only covers a fraction of product categories included in the statistical office's CPI, it includes prices from a larger number of goods and varieties within each category, and provides similar estimates to the aggregate CPIs (see Cavallo, 2013). Other differences between the online and offline price indices, such as the treatment of missing prices or product substitutions, are discussed in Cavallo (2019).

The methodology used to construct the online price indices can be summarized in three steps: data source selection, data collection and aggregation. The starting point is to select the retailers to sample. This decision is driven by *PriceStats* needs to get prices that are representative of retailer transactions. Therefore, the company focuses almost exclusively on the largest players in the market that sell both offline and online. These retailers concentrate the majority of the online transactions and actively manage their websites. On average, a retailer changes its prices two times per week.

In the second step, web-scraping software parses a retailer's public HTML[7] code and collects details for each product, such as the price, description, brand, and size. The retailer assigns a unique identifier (ID) to each good, which is constant over time. This ID is also collected and allows the algorithm to build a panel dataset, with millions of product-level prices per day. Online prices include the VAT tax and exclude transportation costs to match as closely as possible the price used in the traditional CPI. By the same token, the indices exclude out of stock items. In the third step, the price changes from multiple retailers are aggregated into a unique country price index using CPI weights. While the aggregation methodology is similar to the methodology used by the NSOs, we dedicate the last section of this paper to explain the main differences that lead online prices to anticipate their offline counterparts.

Some of the models addressed in this paper include two additional datasets. The first relates to offline gasoline and diesel prices from weekly surveys conducted in each country. These prices include excise taxes and sales/VAT taxes, and are published with a one-week delay, a timing advantage that is hard to find in other sectors included in the CPI. We use these surveys because Knotek and Zaman (2015) shows that they outperform the most advanced forecasts in the literature, even using parsimonious regressions. Second, we include a Phillips Curve model in the paper using the seasonally adjusted unemployment rate from IHS Economics.

We compare the accuracy of the monthly online forecast against the survey of forecasters provided by Bloomberg. Additionally, we compare the quarterly forecast accuracy against the Survey of Professional Forecasters published by the Federal Reserve Bank of Philadelphia.

Throughout the paper, we define the inflation rate as follows,

$$p_t = 100 * \ln \frac{P_t}{P_{t-1}} \tag{1}$$

where P_t and P_{t-1} denote the price index in period t and the previous period $t-1$.

3. Predictive ability of online series

This section suggests that online price indices are useful indicators for forecasting the CPI. First, we visually compare the CPI and their online counterparts. Second, we show the anticipatory effect of the online series by calculating the time required for a shock to the online index to be reflected in the CPI. Third, we show that the online series explain a significant proportion of the CPI's variability even after controlling for other macro variables.

Fig. 1 illustrates how online prices anticipate major changes in the trend of the official inflation rate. The online UK index started falling on December 10, 2014, as can be seen in graph 1.a. One month later it had already dropped by 1.2 percent; however, on January 13th the December release from the NSO showed no change to the CPI. When the January reading was released on February 17th, the index finally registered a 0.9 percent drop,

[7] HTML stands for Hyper Text Markup Language.

giving the online index a nearly two-month lead against the official CPI. The online price index stopped falling in January and started increasing once again, but the CPI did not show this change in trend until the estimates for February were published on March 24th. Fig. 1.b shows similar anticipation in Australia. The online index showed an uptick in its inflation rate in May of 2015, but the CPI did not reflect this increase until the second quarter's inflation rate was published on July 22nd. Similar anticipation patterns can be seen in most of the online series, which are displayed in Appendix A.1.[8]

We next compute impulse response functions (IRFs), which show the effect of a one percentage point shock in the online inflation rate on the CPI's inflation rate, holding all other variables constant. To this end, we calculate a vector autoregression (VAR) for each country, where the online data is taken as the exogenous variable. This is equivalent to estimating an autoregressive distributed lag (ADL). We complement the analysis with 95 percent confidence bands computed by block bootstrapping[9] and Hall (1992)'s percentile-interval method.[10] The model is defined as follows,

$$p_t = \alpha + \sum_{i=1}^{i=n} \beta_{t-i}\, p_{t-i} + \sum_{i=0}^{i=n} \gamma_{t-i}\, o_{t-i} + \epsilon_t \tag{2}$$

where Eq. (2) includes $i = 1, 2, \ldots, n$ lags of offline, p_{t-i}, and online, o_{t-i}, aggregate monthly inflation rates, plus the contemporaneous value of online inflation rate (o_t). We run the model in varying lag lengths before choosing $n = 6$ lags according to the Akaike's final prediction error (FPE) and Lütkepohl (2005)'s lag order selection criteria (HQIC).[11]

The IRFs in Fig. 2 show that shocks to the online indices are slowly incorporated into the CPI, suggesting an anticipatory effect of online prices. All country IRFs are statistically significant, and the CPI takes one-to-five months to react to a shock in online inflation.[12] Moreover, the results hold when we remove the contemporaneous observation of the online price indices, suggesting that their predictive content goes beyond their immediate release to the public.

The IRF analysis does not control for other economic variables. Therefore, we further evaluate the explanatory power of the online series with offline fuel prices, which have been found relevant predictors of the CPI in recent times (e.g., Knotek & Zaman, 2015). To this end, we sequentially partial-out from the CPI the effects of its lagged values, of fuel, and of the online series, and then observe the R^2 of these regressions. That is, first, the CPI is regressed against its lags. Then, the residuals of that regression are explained by fuel prices. Finally, the residuals of that second procedure are regressed against the online index. An R^2 that is higher than zero in the last regression is taken as evidence that the online index is a valuable source of information to predict the CPI's inflation rate. The R^2 in the last regression describes the share of the remaining variance (after partialling-out the offline and fuel prices) explained by the online price index.

Notice that this decomposition runs strongly against the online index since we are allowing the CPI inflation to be fully explained by its own lagged values and also by fuel. In other words, the explanatory power derived from online prices is at the lowest bound of their information content. Bring (1996) shows that the more correlated the variables are, the lower the importance assigned to the last partialed-out regressor. Although not shown in this paper, by reversing the order of the regressions we find that the CPI and fuel explain little compared to the online series.

The horizontal lines on Fig. 3 show the in-sample decomposition for the entire time series, using six lags in each of the regressions, for Australia and USA. The time-varying lines depict the decomposition in a rolling window of 24 months,[13] also using six lags. The rest of the countries are plotted in Appendix A.2.

In all countries, the online aggregate price index accounts for 4 to 20 percent of the CPI's variability in the full sample. In many countries, the fuel series increase their contribution through time, which is expected given the recent context of very low inflation rates and large shocks in fuel prices.

4. Forecasting inflation using online prices

This section presents the forecast exercises. We start by describing the methodology. Section 4.2 shows the results from one- to three-month-ahead inflation forecasts, and Section 4.3 shows the quarterly predictions of the US inflation index.

[8] Ireland's online index disserves a special mention since it shows the largest discrepancy with the CPI. The official series show a marked seasonality, where prices increase significantly in March and April, and September and October, and decrease in February and August. Such seasonality is mostly due to its dynamic apparel sector, which varies prices according to the season. While the online index captures the deflationary months, it has failed to capture the full extent of the seasonal price increases. While our dataset is not disaggregated enough to explain this phenomenon fully, we believe that apparel retailers in Ireland change the ID of each good after the season is over. As a consequence, the online index captures the large price decreases when the goods become on sale and ultimately on clearance, but fail to capture the increase when the same or a similar model is introduced into the market at a higher price the next season.

[9] See Hongyi Li and Maddala (1996) for details about the methodology.

[10] Residual-based bootstrapping based on Benkwitz, Lütkepohl, and Wolters (2001) yields tighter bands, but may not control for the serial correlation of the series.

[11] We run this exercise in STATA. We estimate the coefficients in Eq. (2) using the "var" module, specifying that the online aggregate monthly inflation rate is exogenous. After that, the IRFs are calculated using the "irf" command. We account for the serial correlation of the series calculating the confidence bands by block bootstrapping. The number of bootstrap iterations is 200, but the results do not change running the exercise with 500 iterations.

[12] We complement the IRF analysis by running a Granger causality test (see Granger (1969) and Lütkepohl (2005)). The idea behind this test is that if the online series leads their offline counterpart, the former should improve the prediction of the latter. More formally, we estimate the coefficients in Eq. (2) and test the null hypothesis that the coefficients on o_{t-i} are jointly zero. The test rejects the hypothesis that online prices do not Granger-cause the offline series with a 1 percent significance level for every country in the sample.

[13] Results remain qualitatively similar using alternative lag and window lengths.

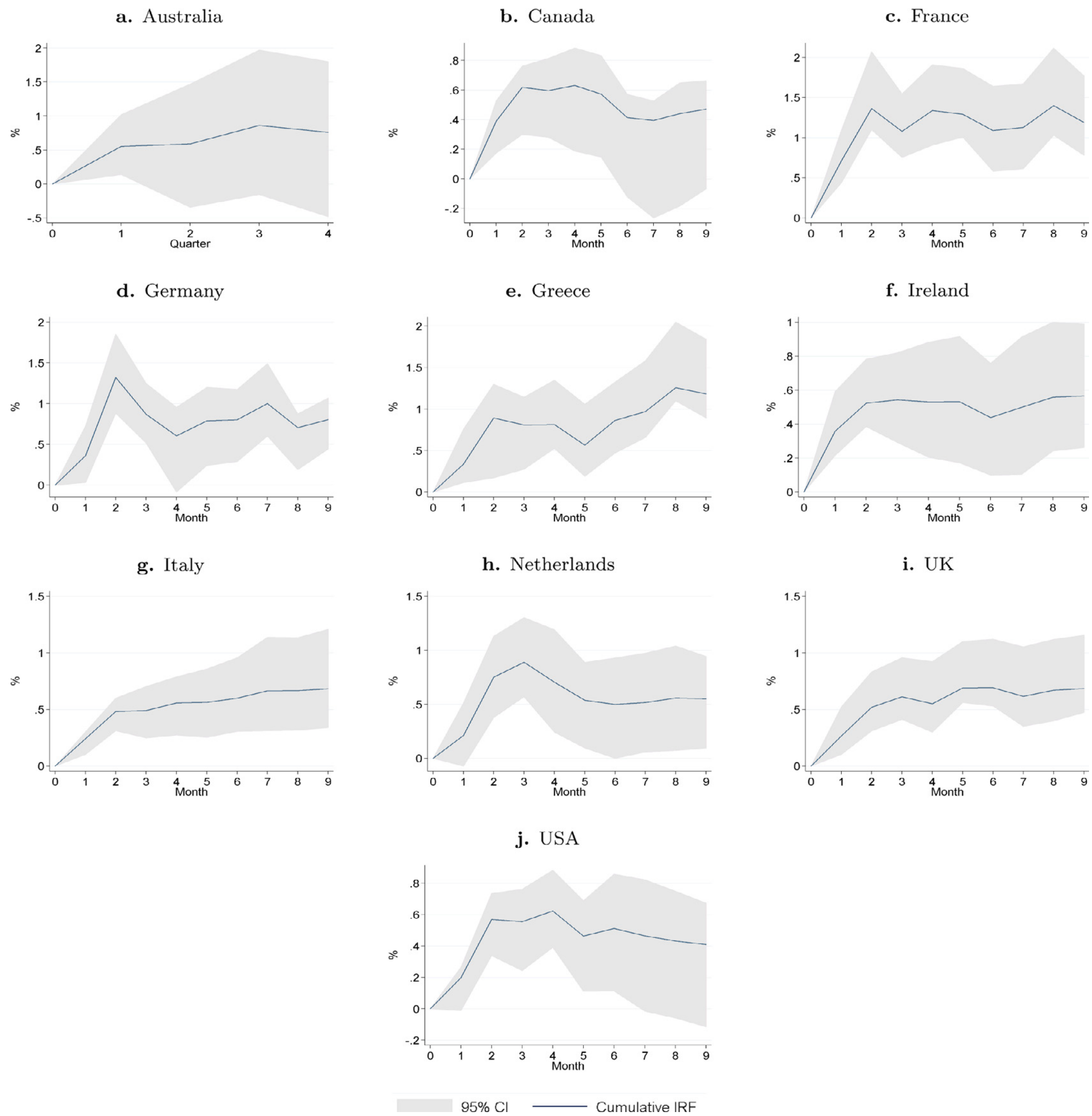

Fig. 2. Impulse responses - One percent shock to the Online price index. *Note:* The solid line in each graph is the impulse response functions of a one percent shock to the online price index on the consumer price index for a particular country. The shaded areas represent the 95 percent confidence intervals of the impulse response, calculated by block bootstrapping. All series are non-seasonally adjusted.
Source: Authors.

4.1. Forecasting models

We forecast the one-month ahead CPI inflation rate as follows,

$$E_{t-1}\, p_t = \hat{\alpha} + \sum_{i=1}^{p} \hat{\beta}_{t-i}\, p_{t-i} + \sum_{i=0}^{p} \hat{\theta}_{t-i} f_{t-i} + \sum_{i=0}^{p} \hat{\gamma}_{t-i}\, o_{t-i} + \sum_{i=0}^{p} \hat{\eta}_{t-i}\, of_{t-i} \tag{3}$$

where p_t, f_t, o_t and of_t are the CPI, offline fuel, online aggregate, and online fuel inflation rates in period t. We take advantage of the daily frequency in online prices by forecasting the inflation rate in t using data through the 15th of month $t-1$, one month in advance of t's CPI release. For example, to forecast the official inflation rate for February, we use online prices through February 15th, and the official CPI rate on January 31st, which is released on February 15th. In other words, we build the forecast with all data that is available on February 15th.[14] The indices have been seasonally adjusted with monthly dummy variables, except for Australia which uses quarterly dummy variables.

We allow the data to speak for itself by calculating a forecast for each combination of regressors in Eq. (3). For example, the first model estimates $\hat{\beta}_{t-1}$ and sets $\hat{\theta}_{t-i} = \hat{\gamma}_{t-i} = \hat{\eta}_{t-i} = 0$. A second model estimates $\hat{\beta}_{t-1}$ and $\hat{\beta}_{t-2}$, and sets all other parameters to zero. We calculate

[14] Our results hold using data until the end of each month.

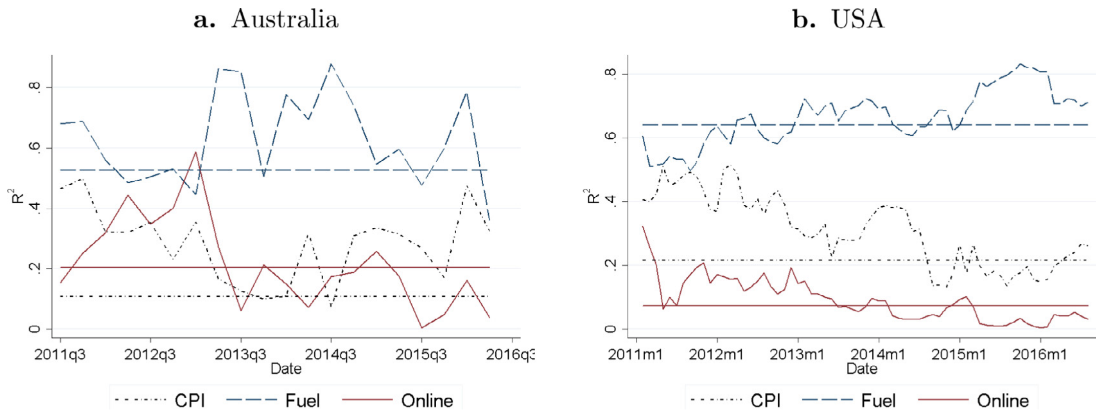

Fig. 3. R^2 Decomposition. *Note:* This figure shows the R^2 of three sequential regressions. The horizontal lines represent full sample regressions, while the moving lines show the R^2 of a 24-month rolling-window estimation. The dotted-and-dashed line represents the first regression, CPI against its lags. The residuals are regressed against fuel, and the resulting R^2 is represented with the dashed lines. Finally, the residuals of that second procedure are regressed against the online index, and the resulting R^2 is depicted with the solid line. All series are non-seasonally adjusted.
Source: Authors.

every combination of values for the regressors so that the last model estimates all parameters where no coefficients are set to zero. Going forward, we define offline models as those that include past CPI inflation rates, offline fuel data, or both. Also, we define online models as those that use any source of online data, such as aggregate inflation indices, fuel prices, or both.

Our baseline specification is an equal-weighted average of forecasts using online data.[15] We choose this method because pooled forecasts have been found to produce more accurate results than ex-ante best individual forecasts, as noted by Faust and Wright (2013), Stock and Watson (1999a), and Timmermann (2006). The popularity of surveys of professional forecasters provided by the Federal Reserve Bank of Philadelphia and by Bloomberg further confirm the accuracy of pooled forecasts.

The baseline specification competes with five benchmarks. The first is an average of offline forecasts following Model (3). This is the analogous specification to our baseline, so any performance improvement against this benchmark exemplifies the value of online price indices as useful predictors of the CPI's inflation rate. The second comparison is against the survey of 1-month ahead forecast published by Bloomberg, which is one of the most well-known forecasts in the market. Finally, we calculate three models commonly used in the literature: an AR(1), a Phillips Curve, and a Random Walk. Details on their methodology are found in Appendix A.3.

Two features need to be defined in our forecasting exercises. *(i)* The in-sample or training window length w. This defines the amount of data used to estimate the parameters of the model. We set w to 24, 36, and 48 months, and take into account all single models based on these windows in the averages.[16,17] *(ii)* Whether the estimation window is fixed or increases over time. An increasing window-length uses all available information through month t before forecasting $t+1$, whereas a fixed window uses the same number of data points on a rolling basis. Our baseline model and its offline equivalent take into account both specifications.[18]

Additionally, we predict the CPI level two and three months ahead. For example, we use all data available until February 15 to predict the inflation rate of March and April. These forecasts can be estimated directly or indirectly. Indirect models predict the one-month ahead inflation rate first and carry over this forecast to the second month, and recursively iterate until the *ith*-month ahead. Direct models, in contrast, forecast the *ith*-month of interest without intermediate regressions. This paper focuses on direct forecasts only.[19]

4.2. Forecast using monthly observations

Our baseline specification identifies the best performing individual models out of all combinations of regressors from Eq. (3). In particular, we determine the top 100 best performers based on their individual root mean square error (RMSE). Therefore, we calculate each individual model and determine the top 100 best performers based on their RMSE. For any given month, we average the predictions of those models. The resulting forecast is compared against the CPI analogously to the comparisons between the CPI and any single model estimated in Eq. (3).

15 There are presumably more efficient weighting schemes, e.g. Bayesian averaging, shrinkage methods, inverse MSE weighting. See Timmermann (2006) for a literature review on forecast combination.

16 The window length in Australia is set to 8, 12, and 16 quarters.

17 A training window of 36 or 48 months yields very similar results. A 24-month window produces slightly higher root mean square errors for all specifications, but results are qualitatively unaltered excluding this window from the analysis.

18 The results remain unchanged removing the increasing or fixed window length from the analysis.

19 Indirect forecasts impose significant modeling structure to the online and fuel inflation data. While an AR(p) can iterate forward its one-month ahead forecast until the *ith*-month, models with online or fuel prices need to forecast the CPI and next month's online and fuel inflation rate. As indicated by Marcellino, Stock, and Watson (2006), assuming too much structure tends to amplify forecast errors when models are incorrectly specified.

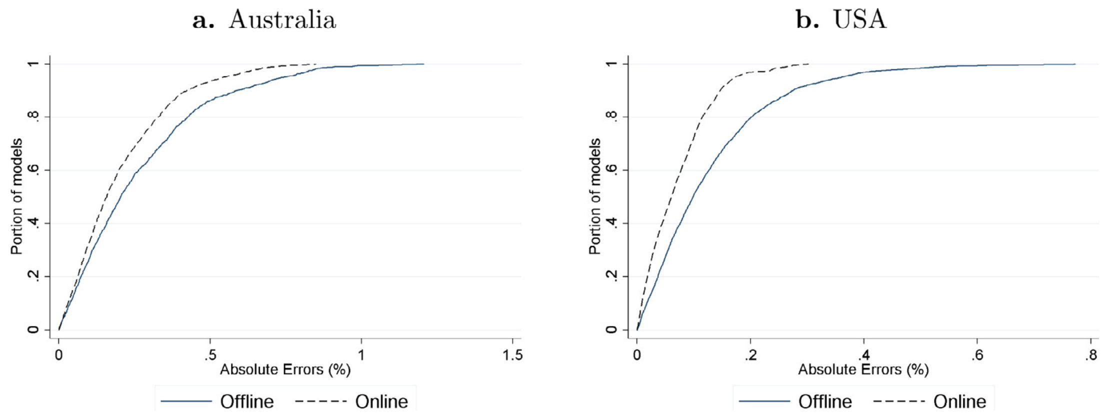

Fig. 4. Cumulative distribution of absolute value of errors. *Note:* This figure shows the cumulative distribution functions of the root mean square errors of the models included in the pooled forecast. The solid line is the distribution based on models using offline data, and the dashed one includes models using online indices.
Source: Authors.

Table 1 shows the RMSE of our baseline and competing models. The results are summarized as follows. First, the 1-month ahead forecast using online series outperforms the rest of the benchmarks in the sample. The online model is on average 16 percent more accurate than the offline benchmark, and 17 percent more accurate than the survey of forecasters published by Bloomberg. We find that the forecasting accuracy using online series is statistically different from the rest of the models using Diebold and Mariano (1995)'s test with a small sample adjustment from Harvey, Leybourne, and Newbold (1997). However, the time span of our series is short, so in some cases, we may not find significant differences due to the low power of the test. We further highlight the contribution of the online indices using Stock and Watson (1999b)'s test.[20]

Second, the online price indices prove to be a useful predictor of the CPI's inflation rate even over longer time horizons. The two-period-ahead RMSE is, on average, 8 percent smaller than its offline benchmark, and the three-period-ahead forecast registers a 4 percent improvement over the offline benchmark. The decrease in forecasting advantage at longer time horizons is consistent with the IRFs in Section 3, where the largest impact of an online shock to the CPI occurs in the first two periods.

The quarterly frequency of the Australian forecast deserves a special mention because the online series have a larger time advantage in this country than they do in others. Therefore, we can calculate a 1-period-ahead forecast three months before the official release. For example, Australia releases the October-December inflation rate around January 25th. On this date, we forecast the January-March inflation rate, which is released on around April 25th. Countries such as Greece and the Netherlands show a higher RMSE than countries like Canada. These European economies show a high amount of month-to-month volatility, mostly due to seasonality in the clothing sector, which is not shown in the online series. Provided with a longer time span of testable data, the authors

Table 1
Root mean square error of monthly forecasts.
Source: Authors and 1-month-ahead survey of professional forecasters published by Bloomberg.

	Online	Offline	Survey	AR(1)	Phillips	RW
1 month ahead						
Australia	0.164	0.230	0.258	0.299	0.286	0.361
Canada	0.071	0.096	0.135	0.186	0.185	0.278
France	0.137	0.134	0.127	0.185	0.174	0.219
Germany	0.188	0.181	0.092	0.262	0.256	0.351
Greece	0.370	0.419	.	0.476	0.454	0.618
Ireland	0.129	0.141	.	0.176	0.254	0.217
Italy	0.097	0.121	1.253	0.176	0.143	0.202
Netherlands	0.144	0.179	.	0.284	0.276	0.306
UK	0.106	0.149	0.140	0.189	0.148	0.181
USA	0.081	0.085	0.096	0.200	0.285	0.244
2 month ahead						
Australia	0.221	0.282	.	0.368	0.326	0.329
Canada	0.161	0.168	.	0.185	0.190	0.281
France	0.169	0.167	.	0.185	0.183	0.212
Germany	0.236	0.238	.	0.246	0.241	0.315
Greece	0.389	0.442	.	0.454	0.437	0.537
Ireland	0.159	0.160	.	0.176	0.248	0.218
Italy	0.135	0.135	.	0.181	0.134	0.193
Netherlands	0.172	0.227	.	0.293	0.276	0.314
UK	0.151	0.164	.	0.210	0.171	0.176
USA	0.193	0.195	.	0.210	0.253	0.310
3 month ahead						
Australia	0.276	0.296	.	0.350	0.406	0.442
Canada	0.155	0.172	.	0.180	0.181	0.288
France	0.165	0.156	.	0.189	0.166	0.226
Germany	0.235	0.225	.	0.247	0.243	0.323
Greece	0.378	0.459	.	0.472	0.450	0.747
Ireland	0.158	0.162	.	0.181	0.226	0.219
Italy	0.129	0.131	.	0.169	0.132	0.175
Netherlands	0.256	0.285	.	0.301	0.283	0.379
UK	0.167	0.175	.	0.221	0.219	0.198
USA	0.212	0.203	.	0.213	0.211	0.309

Notes: Root mean square errors are expressed in non-annualized monthly percentage points. Diebold and Mariano (1995)'s and Stock and Watson (1999b)'s significance tests are shown in Appendix A.4.

[20] The test regresses $p_t = \lambda f_{t-1}^{online} + (1-\lambda) f_{t-1}^{benchmark} + e_t$, where f_{t-1} denotes the forecast of the CPI's inflation rate, p_t. The regression suggests that the online index is a valuable source of information to forecast the inflation rate when λ is significantly higher than zero.

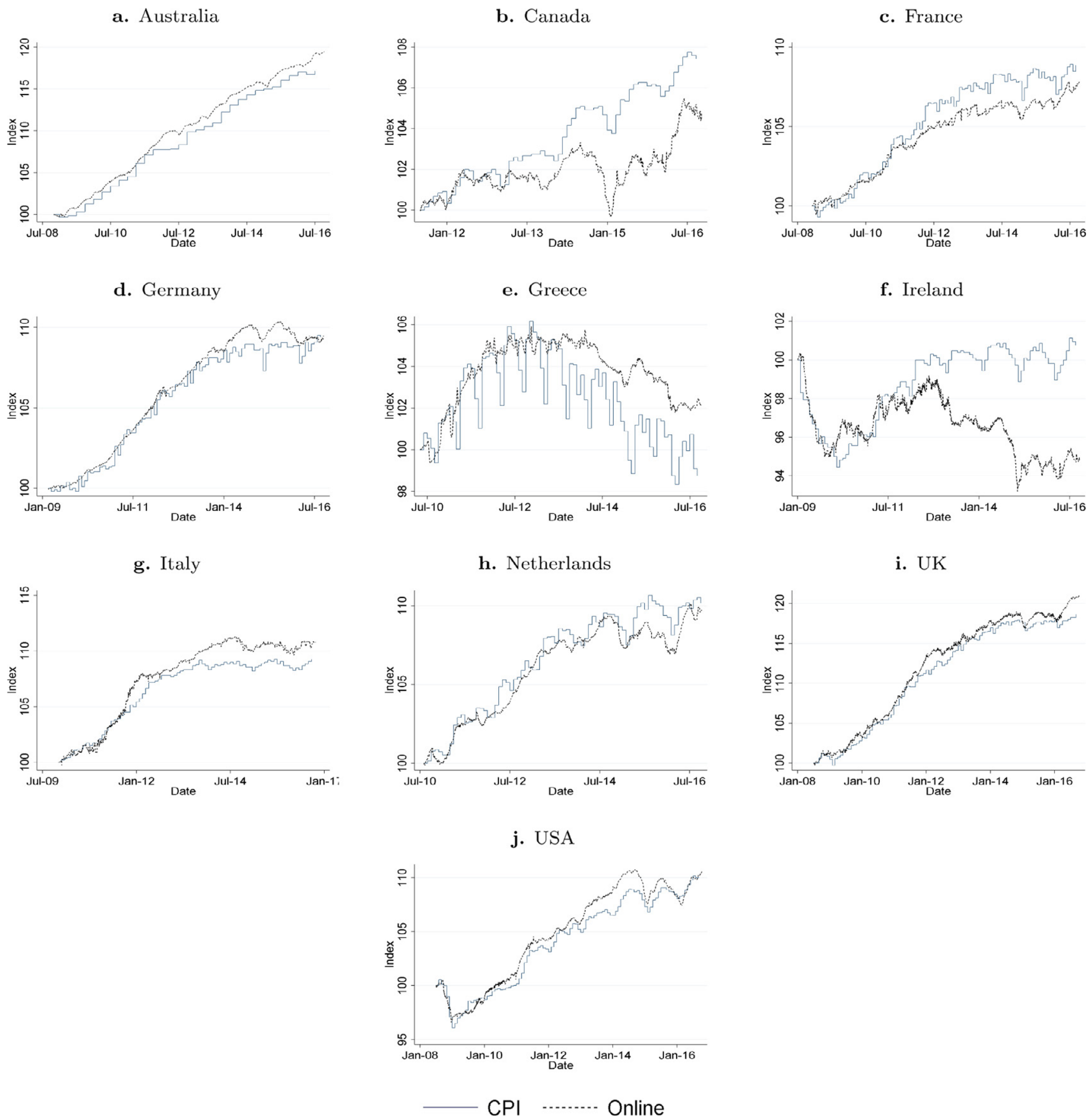

Fig. A.1. Online aggregate inflation indices. *Note:* This figure showcases the online price index next to the Conumer Price Index for each country. *Source:* Authors using online indices computed by *PriceStats* and the Consumer Price Indices, non-seasonally adjusted, all items, from the National Statistical Offices.

would add explanatory variables that correlate with the seasonality of clothing to address this phenomenon. We view the construction of more disaggregated price indices as a promising area of future research.

The baseline specification could show a smaller RMSE because either a few models perform extremely well, compensating for others that perform poorly, or because most of the models perform slightly better than those that exclude online indices. We prefer the second scenario because we are more likely to make small forecasting mistakes independently of the models included in the average. We calculate a cumulative distribution of the absolute forecasting errors of the individual models included in the baseline and offline benchmark. Fig. 4 shows that online prices reduce the likelihood of making large forecast errors in Australia and the US. The same result holds for the rest of the countries, plotted in Appendix A.5. Furthermore, the distribution of offline forecasts is stochastically dominated by the online forecasts for 8 out of 10 economies.[21]

We find that the baseline specification is particularly accurate when the offline benchmark cannot pick up the CPI surprises. For each country, we keep the 50 percent of the observations where the offline benchmark makes the largest forecast errors (in absolute value). For this subset of data, we estimate the forecasting accuracy of

21 The results hold when we include 20, 50, or 200 models in the cumulative distributions. It also holds when we restrict the distribution to positive or negative errors.

Table 2
Root mean square error of quarterly forecast for USA.
Source: Authors and the Survey of Professional Forecasters published by the Federal Reserve Bank of Philadelphia.

	Online	Offline	SPF Mean	SPF Median	AR(1)	Phillips	RW
RMSE	0.429	0.806	1.677	1.785	1.198	1.968	1.575
D-M		0.002	0.000	0.000	0.008	0.000	0.000
S-W		1.165	1.251	1.236	1.016	1.086	1.048
		(0.224)	(0.215)	(0.185)	(0.123)	(0.103)	(0.084)

Notes: Root mean square errors are expressed in non-annualized monthly percentage points. D-M shows the p-values of the null hypothesis of the online model presenting similar predictive ability than each alternative model. The test is based on Diebold and Mariano (1995), using a small sample adjustment from Harvey et al. (1997). S-W represents the λ coefficient of Stock and Watson (1999b)'s test. This test suggests that the online index is a valuable source of information to forecast the inflation rate when λ is significantly higher than zero. λ's standard errors are in parenthesis.

the online versus the offline forecasts. The baseline specification is statistically more accurate than the offline benchmark for every country except France, Germany, and the US.[22] For example, in March 2015 the offline benchmark predicted the inflation rate in the UK would accelerate, reaching 0.32 percent. For the same period, the online model predicted a deceleration, with an estimate of 0.14 percent. When the CPI came out, the inflation rate was 0.15 percent.

The conclusions of this paper do not change when our baseline and main offline benchmark average the top 20, 50, or 200 models. Moreover, Appendix A.6 shows that online prices are a useful predictor of the CPI even after indiscriminately combining every single model from Eq. (3).

4.3. Forecast using quarterly observations

This section presents a quarterly inflation forecast for the US and compares its results to the Survey of Professional Forecasters released by the Federal Reserve Bank of Philadelphia. We use the SPF survey because it is considered a leading source of medium-term forecast in the US, and it is regularly monitored by the Federal Reserve System.[23] Evidence of this survey's accuracy is discussed in Ang et al. (2007), Croushore (2010), and Faust and Wright (2013).

The quarterly forecast is constructed compounding the one, two, and three-month ahead forecasts from a given quarter of the baseline specification and then annualizing the result. Similarly to Section 4.2, the online and offline models are a pooled forecast of the 100 best models, but we only include forecasts with an increasing window, starting at 36 months.[24] All series have been seasonally adjusted using dummy variables as in the previous section.

22 The number of observations used in this exercise is extremely low, so we may not find significant differences due to the low power of the test.

23 See, for example, Bernanke (2007) and Yellen (2017).

24 Results are similar using a 24 or 48-month window as well as using a fixed time window.

Table 2 presents the out-of-sample RMSE for the quarterly forecasting exercise. The forecasts using online indices substantially outperform the quarterly survey of professional forecasters and the other benchmarks included in this paper. Despite the small sample size, both Diebold and Mariano (1995)[25] and Harvey et al. (1997) tests suggest that the results are significant at the 1 percent level.

5. Why do online prices anticipate the CPI inflation?

The paper so far has shown that the online price indices accurately predict the CPI's values several months in advance. This is a reasonable result since both indices track prices from the same sectors of the economy, and therefore should be highly correlated. To address this point, Cavallo (2017) compares offline and online prices and shows that the mean absolute size of price changes is very similar in several countries. Some of the papers in Section 2 expand further on this topic.

Interestingly, forecasts using online data are highly accurate even after removing their timing advantage, suggesting that pricing dynamics in the online market differ from those in the offline market. Indeed, Cavallo (2017) finds a 30 percent price difference between online and offline retailers, which could arise from unsynchronized pricing. These differences in reaction times between the online and offline markets present an area of future research, but we mention four factors that likely contribute to the anticipatory feature of online prices.

First, retailers usually incur a cost to change listed prices, and high costs can become a barrier to adjusting prices frequently (Kehoe & Midrigan, 2015; Levy, Bergen, Dutta, & Venable, 1997; Nakamura & Zerom, 2010). In a comparison of offline and online retailers, Brynjolfsson and Smith (2000) finds evidence that online stores change prices in smaller amounts, suggesting lower menu costs and less price rigidity. However, Cavallo (2019) documents that small price changes are not as common as previously reported. The matter is far from settled and the frequency of price changes may, in fact, be changing over time as online purchases become more popular and technology improvements make price changes

25 The test uses a small sample adjustment from Harvey et al. (1997).

Table A.1
Root mean square error of monthly forecasts - *One month ahead*.
Source: Authors and 1-month-ahead survey of professional forecasters published by Bloomberg.

	Online	Offline	Survey	AR(1)	Phillips	RW
Australia	0.164	0.230	0.258	0.299	0.286	0.361
D-M		0.019	0.288	0.001	0.002	0.005
S-W		0.813	0.666	0.868	0.922	0.901
		(0.243)	(0.003)	(0.180)	(0.188)	(0.134)
Canada	0.071	0.096	0.135	0.186	0.185	0.278
D-M		0.000	0.000	0.003	0.002	0.005
S-W		1.034	1.093	0.870	0.872	0.911
		(0.418)	(0.001)	(0.107)	(0.109)	(0.074)
France	0.137	0.134	0.127	0.185	0.174	0.219
D-M		0.295	0.190	0.002	0.002	0.006
S-W		0.663	0.255	0.999	0.986	1.135
		(0.623)	(0.095)	(0.233)	(0.227)	(0.141)
Germany	0.188	0.181	0.092	0.262	0.256	0.351
D-M		0.254	0.000	0.032	0.020	0.000
S-W		−0.276	0.125	1.007	1.041	1.097
		(0.696)	(0.161)	(0.245)	(0.255)	(0.117)
Greece	0.370	0.419	.	0.476	0.454	0.618
D-M		0.000	.	0.002	0.008	0.008
S-W		1.754	.	1.574	1.317	1.199
		(0.700)	.	(0.610)	(0.587)	(0.215)
Ireland	0.129	0.141	.	0.176	0.254	0.217
D-M		0.053	.	0.002	0.001	0.001
S-W		1.139	.	0.958	0.992	0.948
		(0.379)	.	(0.180)	(0.161)	(0.120)
Italy	0.097	0.121	1.253	0.176	0.143	0.202
D-M		0.013	0.000	0.000	0.005	0.001
S-W		0.985	0.985	1.160	0.909	1.189
		(0.325)	(0.000)	(0.207)	(0.210)	(0.130)
Netherlands	0.144	0.179	.	0.284	0.276	0.306
D-M		0.117	.	0.005	0.004	0.003
S-W		1.040	.	1.143	1.165	1.032
		(0.376)	.	(0.203)	(0.200)	(0.144)
UK	0.106	0.149	0.140	0.189	0.148	0.181
D-M		0.014	0.045	0.000	0.001	0.000
S-W		0.908	0.773	0.858	0.860	1.141
		(0.214)	(0.000)	(0.142)	(0.121)	(0.121)
USA	0.081	0.085	0.096	0.200	0.285	0.244
D-M		0.141	0.510	0.002	0.000	0.001
S-W		0.684	0.581	1.001	0.936	1.045
		(0.310)	(0.000)	(0.071)	(0.065)	(0.059)

Notes: The table shows the root mean square errors (RMSE) of the baseline and our main benchmark when the top 100 single models in Eq. (3) are included in the pooled forecast. RMSEs are expressed in non-annualized monthly percentage points. D-M shows the p-values of the null hypothesis of the online model presenting similar predictive ability than each alternative model. The test is based on Diebold and Mariano (1995), using a small sample adjustment from Harvey et al. (1997). S-W represents the λ coefficient of Stock and Watson (1999b)'s test. This test suggests that the online index is a valuable source of information to forecast the inflation rate when λ is significantly higher than zero. λ's standard errors are in parenthesis.

Table A.2
Root mean square error of monthly forecasts - *Two month ahead*.
Source: Authors.

	Online	Offline	Survey	AR(1)	Phillips	RW
Australia	0.221	0.282	.	0.368	0.326	0.329
D-M		0.000	.	0.001	0.015	0.185
S-W		1.162	.	1.355	1.097	0.955
		(0.546)	.	(0.329)	(0.345)	(0.232)
Canada	0.161	0.168	.	0.185	0.190	0.281
D-M		0.049	.	0.027	0.031	0.008
S-W		2.933	.	1.870	1.325	0.960
		(2.036)	.	(1.118)	(0.835)	(0.234)
France	0.169	0.167	.	0.185	0.183	0.212
D-M		0.298	.	0.078	0.121	0.060
S-W		0.148	.	0.591	0.464	0.950
		(0.483)	.	(0.421)	(0.524)	(0.190)
Germany	0.236	0.238	.	0.246	0.241	0.315
D-M		0.476	.	0.206	0.293	0.021
S-W		0.802	.	0.510	0.174	1.030
		(1.388)	.	(0.509)	(0.516)	(0.196)
Greece	0.389	0.442	.	0.454	0.437	0.537
D-M		0.007	.	0.020	0.010	0.001
S-W		1.124	.	1.119	1.117	0.881
		(0.458)	.	(0.510)	(0.500)	(0.218)
Ireland	0.159	0.160	.	0.176	0.248	0.218
D-M		0.271	.	0.083	0.004	0.005
S-W		0.946	.	0.842	0.940	0.992
		(0.941)	.	(0.362)	(0.244)	(0.185)
Italy	0.135	0.135	.	0.181	0.134	0.193
D-M		0.434	.	0.023	0.397	0.035
S-W		0.747	.	0.866	0.631	1.027
		(0.393)	.	(0.322)	(0.413)	(0.177)
Netherlands	0.172	0.227	.	0.293	0.276	0.314
D-M		0.116	.	0.014	0.016	0.001
S-W		0.782	.	1.089	1.117	0.903
		(0.383)	.	(0.268)	(0.266)	(0.165)
UK	0.151	0.164	.	0.210	0.171	0.176
D-M		0.288	.	0.003	0.028	0.025
S-W		1.101	.	0.675	0.931	0.963
		(0.262)	.	(0.251)	(0.219)	(0.186)
USA	0.193	0.195	.	0.210	0.253	0.310
D-M		0.082	.	0.014	0.007	0.000
S-W		0.843	.	0.790	0.794	1.048
		(0.391)	.	(0.270)	(0.275)	(0.133)

Notes: The table shows the root mean square errors (RMSE) of the baseline and our main benchmark when the top 100 single models in Eq. (3) are included in the pooled forecast. RMSEs are expressed in non-annualized monthly percentage points. D-M shows the p-values of the null hypothesis of the online model presenting similar predictive ability than each alternative model. The test is based on Diebold and Mariano (1995), using a small sample adjustment from Harvey et al. (1997). S-W represents the λ coefficient of Stock and Watson (1999b)'s test. This test suggests that the online index is a valuable source of information to forecast the inflation rate when λ is significantly higher than zero. λ's standard errors are in parenthesis.

less expensive. However, it is still plausible that online prices are less sticky than their offline counterparts, so price changes should happen sooner online than in the brick-and-mortar stores.

Second, when a product's price is not available, the surveying agent at the NSO directly compares the price from the closest alternative product against the previous price of the original good.[26] If the agent finds a similar product but does not consider it comparable, a hedonic regression removes the price differential associated with the quality discrepancy between the products. When the alternative product is not comparable, the price for that period is assumed to change by the average price change of comparable items. These indirect substitutions may lead to delays in the reaction time of the CPI to recent price updates. In contrast, as of the publishing date of this paper, the *PriceStats*' online indices calculate price changes for identical items only.

Third, price changes may be reported late when the "different day pricing" methodology is applied.[27] When

[26] See International Labour Office (2004) for additional details on standard NSOs methods on CPI's.

[27] See Bureau of Labor Statistics (2015) for details.

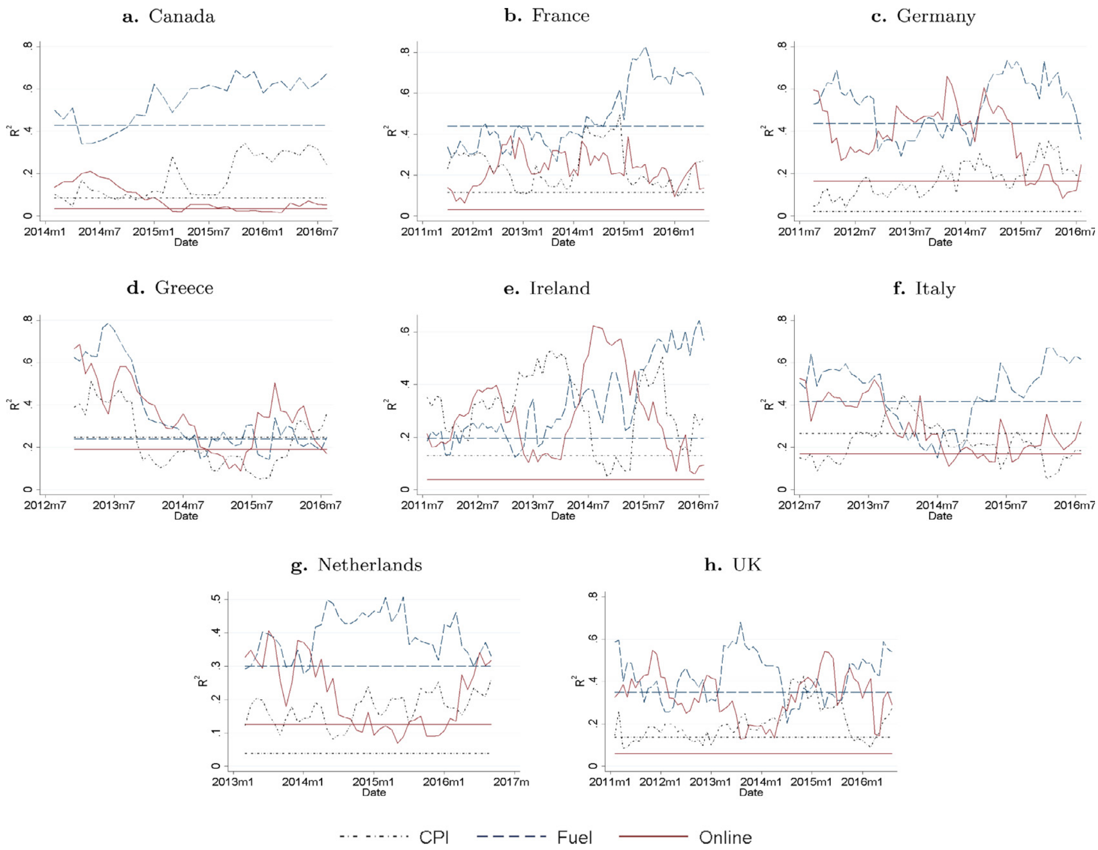

Fig. A.2. R^2 Decomposition. *Note:* This figure shows the R^2 of three sequential regressions. The horizontal lines represent full sample regressions, while the moving lines show the R^2 of a 24-month rolling-window estimation. The dotted-and-dashed line represents the first regression, CPI against its lags. The residuals are regressed against fuel, and the resulting R^2 is represented with the dashed lines. Finally, the residuals of that second procedure are regressed against the online index, and the resulting R^2 is depicted with the solid line. All series are non-seasonally adjusted. *Source:* Authors.

a product is not available at the time of collection, its previous price is eligible as a substitute, provided that the item was available for sale in the last seven days. As a result, the CPI may record price spikes one month after they actually occur.

Finally, online anticipation can be amplified when NSOs collect prices on a bimonthly frequency. For example, the Bureau of Labor Statistics in the US collects prices monthly for all items in the three largest publication areas, but only collects prices bimonthly in the remaining regions.[28] This infrequent data collection schedule may further delay the recording of price changes in the CPI.[29]

6. Conclusions

Our work introduces online price indices as a useful predictor of the CPI's inflation rate for many economies and at multiple horizons. We use parsimonious models that do not exploit the high frequency of the online series. However, when online prices are included, these models outperform the most common benchmarks in the literature as well as two leading surveys of professional forecasters. We argue that these are reasonable results since the CPI and the online price indices collect prices from the same sectors of the economy and are therefore closely correlated.

We discussed how inflation values from online series are available earlier than the CPI's, and that therefore policymakers and industry practitioners alike can leverage them to include more recent information into their models. Our analysis shows that even after removing this timing advantage online prices still improve forecasting accuracy, which suggests that those prices tend to move before offline prices. Furthermore, we outlined how some methodological procedures employed to calculate a CPI may delay the recording of certain price changes, leading to slower reaction times for offline series.

Our analysis suggests several areas that would benefit from further research. First, online prices should be included in forecasting models that take advantage of their high frequency, such as those explained by Modugno

28 See Bureau of Labor Statistics (2015) for details.

29 See also Hausman and Leibtag (2009) for a discussion on the gradual introduction of new items and additional sources of bias in the CPI.

Table A.3
Root mean square error of monthly forecasts - *Three month ahead*.
Source: Authors.

	Online	Offline	Survey	AR(1)	Phillips	RW
Australia	0.276	0.296	.	0.350	0.406	0.442
D-M		0.085	.	0.063	0.001	0.002
S-W		2.374	.	1.179	1.737	1.303
		(0.943)	.	(0.609)	(0.548)	(0.213)
Canada	0.155	0.172	.	0.180	0.181	0.288
D-M		0.038	.	0.007	0.007	0.001
S-W		1.378	.	2.120	2.124	1.221
		(1.030)	.	(0.898)	(0.875)	(0.284)
France	0.165	0.156	.	0.189	0.166	0.226
D-M		0.394	.	0.002	0.285	0.002
S-W		0.517	.	0.665	0.702	1.021
		(0.404)	.	(0.414)	(0.399)	(0.171)
a Germany	0.235	0.225	.	0.247	0.243	0.323
D-M		0.339	.	0.139	0.089	0.001
S-W		−0.359	.	0.148	−0.232	1.117
		(0.884)	.	(0.615)	(1.004)	(0.203)
Greece	0.378	0.459	.	0.472	0.450	0.747
D-M		0.018	.	0.000	0.002	0.000
S-W		1.460	.	1.994	1.729	1.126
		(0.480)	.	(0.644)	(0.553)	(0.162)
Ireland	0.158	0.162	.	0.181	0.226	0.219
D-M		0.130	.	0.046	0.028	0.032
S-W		3.272	.	0.954	1.087	0.983
		(1.227)	.	(0.346)	(0.248)	(0.175)
Italy	0.129	0.131	.	0.169	0.132	0.175
D-M		0.449	.	0.005	0.384	0.012
S-W		0.924	.	1.007	0.932	0.975
		(0.476)	.	(0.464)	(0.447)	(0.198)
Netherlands	0.256	0.285	.	0.301	0.283	0.379
D-M		0.122	.	0.057	0.081	0.002
S-W		1.104	.	0.917	0.906	0.889
		(0.774)	.	(0.653)	(0.565)	(0.220)
UK	0.167	0.175	.	0.221	0.219	0.198
D-M		0.363	.	0.000	0.040	0.033
S-W		1.286	.	0.689	1.125	1.087
		(0.299)	.	(0.623)	(0.240)	(0.195)
USA	0.212	0.203	.	0.213	0.211	0.309
D-M		0.338	.	0.103	0.388	0.001
S-W		2.094	.	−0.193	1.605	1.126
		(1.914)	.	(2.590)	(0.990)	(0.163)

Notes: The table shows the root mean square errors (RMSE) of the baseline and our main benchmark when the top 100 single models in Eq. (3) are included in the pooled forecast. RMSEs are expressed in non-annualized monthly percentage points. D-M shows the p-values of the null hypothesis of the online model presenting similar predictive ability than each alternative model. The test is based on Diebold and Mariano (1995), using a small sample adjustment from Harvey et al. (1997). S-W represents the λ coefficient of Stock and Watson (1999b)'s test. This test suggests that the online index is a valuable source of information to forecast the inflation rate when λ is significantly higher than zero. λ's standard errors are in parenthesis.

Table A.4
Root mean square error of monthly forecasts - *One month ahead. All-model specification.*
Source: Authors and 1-month-ahead survey of professional forecasters published by Bloomberg.

	Online	Offline	Survey	AR(1)	Phillips	RW
Australia	0.234	0.272	0.258	0.299	0.286	0.361
D-M		0.077	0.628	0.022	0.017	0.018
S-W		0.659	0.446	0.703	0.769	0.779
		(0.342)	(0.082)	(0.252)	(0.278)	(0.172)
Canada	0.093	0.134	0.135	0.186	0.185	0.278
D-M		0.034	0.000	0.005	0.004	0.007
S-W		1.061	1.338	0.962	0.962	0.978
		(0.340)	(0.011)	(0.177)	(0.179)	(0.115)
France	0.151	0.158	0.127	0.185	0.174	0.219
D-M		0.051	0.069	0.006	0.003	0.007
S-W		0.791	0.169	1.101	1.093	1.236
		(0.623)	(0.248)	(0.343)	(0.334)	(0.172)
Germany	0.209	0.217	0.092	0.262	0.256	0.351
D-M		0.236	0.000	0.064	0.045	0.000
S-W		0.465	0.113	0.814	0.921	1.303
		(0.619)	(0.143)	(0.314)	(0.371)	(0.154)
Greece	0.439	0.449	.	0.476	0.454	0.618
D-M		0.068	.	0.131	0.491	0.036
S-W		1.418	.	0.329	0.121	1.067
		(2.193)	.	(0.883)	(0.833)	(0.259)
Ireland	0.151	0.164	.	0.176	0.254	0.217
D-M		0.134	.	0.013	0.006	0.001
S-W		0.945	.	0.938	0.976	0.997
		(0.373)	.	(0.257)	(0.215)	(0.160)
Italy	0.117	0.140	1.253	0.176	0.143	0.202
D-M		0.000	0.000	0.000	0.063	0.002
S-W		1.604	0.985	1.448	0.929	1.366
		(0.553)	(0.000)	(0.301)	(0.297)	(0.161)
Netherlands	0.198	0.216	.	0.284	0.276	0.306
D-M		0.052	.	0.002	0.001	0.001
S-W		0.726	.	1.024	1.050	0.964
		(0.623)	.	(0.335)	(0.327)	(0.213)
UK	0.131	0.166	0.140	0.189	0.148	0.181
D-M		0.005	0.720	0.001	0.097	0.002
S-W		1.328	0.785	1.217	1.089	1.366
		(0.316)	(0.000)	(0.222)	(0.172)	(0.140)
USA	0.092	0.125	0.096	0.200	0.285	0.244
D-M		0.015	0.933	0.002	0.000	0.001
S-W		1.063	0.486	1.105	1.001	1.118
		(0.191)	(0.001)	(0.090)	(0.082)	(0.071)

Notes: The table shows the root mean square errors (RMSE) of the baseline and our main benchmark when all the single models in Eq. (3) are included in the pooled forecast. RMSEs are expressed in non-annualized monthly percentage points. D-M shows the p-values of the null hypothesis of the online model presenting similar predictive ability than each alternative model. The test is based on Diebold and Mariano (1995), using a small sample adjustment from Harvey et al. (1997). S-W represents the λ coefficient of Stock and Watson (1999b)'s test. This test suggests that the online index is a valuable source of information to forecast the inflation rate when λ is significantly higher than zero. λ's standard errors are in parenthesis.

(2013) and Knotek and Zaman (2015). Second, forecasts that include online sector indices might better capture changes in price trends in volatile sectors, such as clothing. Currently, the main limitation is the short time span of testable data available for these indices, but as time goes by, the inclusion of sectoral data will become a feasible exercise. Third, new research avenues should investigate the reasons why online prices anticipate the CPI, even after removing their frequency and immediate-release advantages.

Acknowledgments

We are very grateful to Alberto Cavallo, Jerry Hausman, Anna Mikusheva, Emi Nakamura, Roberto Rigobon, Pete Salas, and Michael Metcalfe for many helpful discussions. We also thank *PriceStats* for granting us access to the data. Manuel Bertolotto is the Chief Operating Officer of *PriceStats*, but the opinions expressed in this paper are those of the authors and do not necessarily reflect those of *PriceStats*. The authors have no other financial or material

Table A.5
Root mean square error of monthly forecasts - *Two month ahead. All-model specification.*
Source: Authors.

	Online	Offline	Survey	AR(1)	Phillips	RW
Australia	0.361	0.353	.	0.368	0.326	0.329
D-M		0.328	.	0.384	0.316	0.403
S-W		0.204	.	0.629	0.382	0.604
		(0.909)	.	(0.540)	(0.433)	(0.294)
Canada	0.192	0.175	.	0.185	0.190	0.281
D-M		0.066	.	0.069	0.166	0.020
S-W		−0.539	.	0.028	0.270	0.792
		(1.169)	.	(0.876)	(0.680)	(0.243)
France	0.177	0.175	.	0.185	0.183	0.212
D-M		0.177	.	0.130	0.450	0.115
S-W		−0.482	.	0.612	0.343	1.073
		(1.171)	.	(0.700)	(0.626)	(0.218)
Germany	0.256	0.239	.	0.246	0.241	0.315
D-M		0.031	.	0.165	0.138	0.032
S-W		−0.990	.	−0.275	−0.430	0.981
		(0.625)	.	(0.545)	(0.472)	(0.225)
Greece	0.451	0.454	.	0.454	0.437	0.537
D-M		0.442	.	0.337	0.469	0.003
S-W		0.451	.	0.268	0.323	0.766
		(1.008)	.	(0.678)	(0.643)	(0.268)
Ireland	0.195	0.178	.	0.176	0.248	0.218
D-M		0.001	.	0.008	0.036	0.074
S-W		−1.509	.	−0.085	0.556	0.748
		(0.697)	.	(0.404)	(0.358)	(0.212)
Italy	0.148	0.146	.	0.181	0.134	0.193
D-M		0.330	.	0.006	0.055	0.049
S-W		−0.052	.	0.850	0.113	1.069
		(1.023)	.	(0.533)	(0.564)	(0.214)
Netherlands	0.232	0.259	.	0.293	0.276	0.314
D-M		0.074	.	0.002	0.001	0.012
S-W		0.953	.	1.244	1.295	0.850
		(0.773)	.	(0.518)	(0.509)	(0.225)
UK	0.177	0.191	.	0.210	0.171	0.176
D-M		0.031	.	0.006	0.404	0.487
S-W		0.935	.	0.716	0.940	0.909
		(0.504)	.	(0.400)	(0.248)	(0.190)
USA	0.209	0.211	.	0.210	0.253	0.310
D-M		0.360	.	0.286	0.050	0.001
S-W		0.388	.	0.594	0.585	1.068
		(0.630)	.	(0.373)	(0.381)	(0.154)

Notes: The table shows the root mean square errors (RMSE) of the baseline and our main benchmark when all the single models in Eq. (3) are included in the pooled forecast. RMSEs are expressed in non-annualized monthly percentage points. D-M shows the p-values of the null hypothesis of the online model presenting similar predictive ability than each alternative model. The test is based on Diebold and Mariano (1995), using a small sample adjustment from Harvey et al. (1997). S-W represents the λ coefficient of Stock and Watson (1999b)'s test. This test suggests that the online index is a valuable source of information to forecast the inflation rate when λ is significantly higher than zero. λ's standard errors are in parenthesis.

Table A.6
Root mean square error of monthly forecasts - *Three month ahead. All-model specification.*
Source: Authors.

	Online	Offline	Survey	AR(1)	Phillips	RW
Australia	0.522	0.399	.	0.350	0.406	0.442
D-M		0.095	.	0.079	0.234	0.304
S-W		−0.147	.	−0.171	0.206	0.650
		(0.560)	.	(0.354)	(0.294)	(0.267)
Canada	0.199	0.191	.	0.180	0.181	0.288
D-M		0.306	.	0.268	0.299	0.002
S-W		0.131	.	0.207	0.244	1.056
		(1.272)	.	(0.952)	(0.973)	(0.382)
France	0.179	0.178	.	0.189	0.166	0.226
D-M		0.430	.	0.025	0.014	0.018
S-W		−0.517	.	0.419	0.512	1.034
		(0.844)	.	(0.564)	(0.565)	(0.185)
Germany	0.252	0.247	.	0.247	0.243	0.323
D-M		0.091	.	0.471	0.071	0.017
S-W		−0.603	.	−0.362	−0.794	0.994
		(0.732)	.	(0.647)	(0.831)	(0.197)
Greece	0.440	0.478	.	0.472	0.450	0.747
D-M		0.103	.	0.114	0.091	0.000
S-W		1.336	.	0.712	0.955	1.274
		(0.833)	.	(0.741)	(0.781)	(0.215)
Ireland	0.212	0.186	.	0.181	0.226	0.219
D-M		0.009	.	0.045	0.220	0.385
S-W		−2.365	.	−0.595	0.132	0.619
		(0.636)	.	(0.394)	(0.490)	(0.227)
Italy	0.151	0.153	.	0.169	0.132	0.175
D-M		0.318	.	0.112	0.021	0.069
S-W		0.023	.	0.154	0.023	0.785
		(0.920)	.	(0.485)	(0.538)	(0.219)
Netherlands	0.296	0.292	.	0.301	0.283	0.379
D-M		0.369	.	0.308	0.253	0.008
S-W		−0.118	.	0.118	0.445	1.000
		(1.813)	.	(1.121)	(1.056)	(0.306)
UK	0.196	0.200	.	0.221	0.219	0.198
D-M		0.049	.	0.002	0.251	0.458
S-W		1.088	.	0.625	1.037	1.089
		(0.692)	.	(0.538)	(0.224)	(0.196)
USA	0.229	0.221	.	0.213	0.211	0.309
D-M		0.001	.	0.022	0.203	0.004
S-W		−2.083	.	−0.854	0.679	0.980
		(1.201)	.	(0.642)	(0.804)	(0.167)

Notes: The table shows the root mean square errors (RMSE) of the baseline and our main benchmark when all the single models in Eq. (3) are included in the pooled forecast. RMSEs are expressed in non-annualized monthly percentage points. D-M shows the p-values of the null hypothesis of the online model presenting similar predictive ability than each alternative model. The test is based on Diebold and Mariano (1995), using a small sample adjustment from Harvey et al. (1997). S-W represents the λ coefficient of Stock and Watson (1999b)'s test. This test suggests that the online index is a valuable source of information to forecast the inflation rate when λ is significantly higher than zero. λ's standard errors are in parenthesis.

interest related to this project to disclose. All errors are our own.

Appendix A

A.1. Online price indices

See Fig. A.1.

A.2. R^2 Decomposition

See Fig. A.2.

A.3. The models in detail

This section explains the methodological details of the AR(p), Phillips curve, and Random-Walk models used in the paper.

A.3.1. AR(p)

We set the first-order univariate autoregressive model to be the benchmark as it is a simple model but remains hard to outperform in the literature (see Faust & Wright, 2013; Stock & Watson, 2003 for examples). However, we have also considered an autoregressive model with n

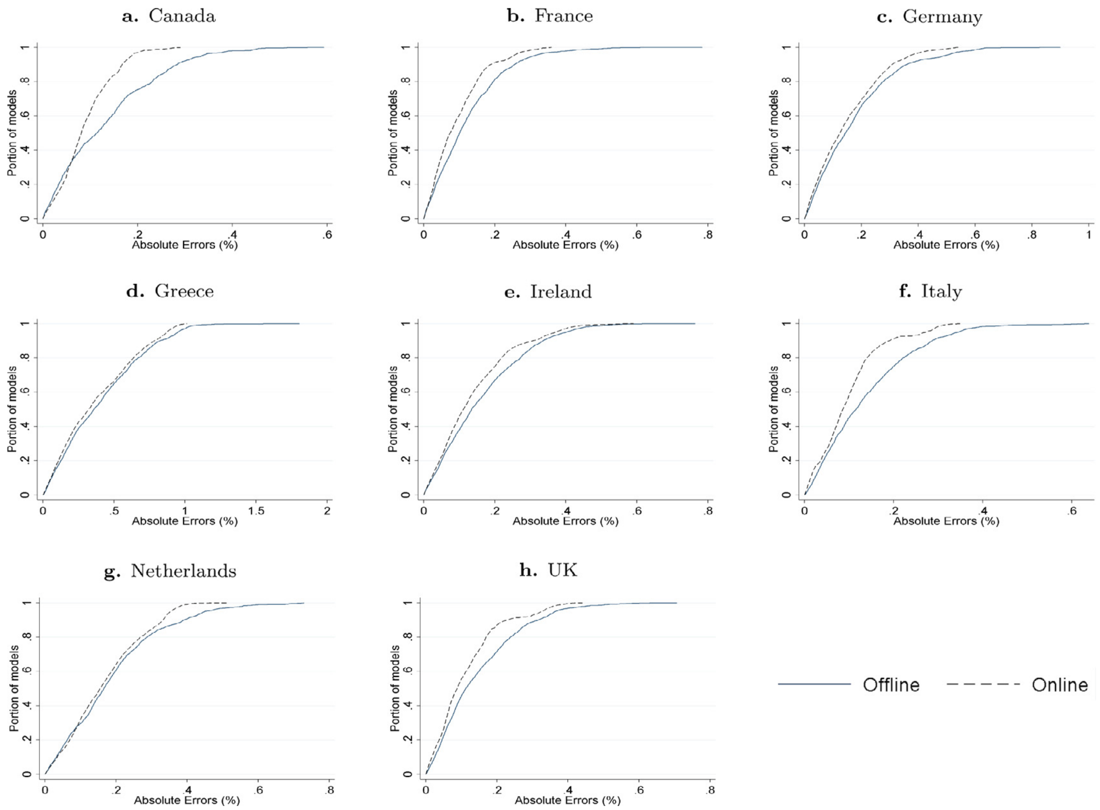

Fig. A.3. Cumulative distribution of absolute value of errors. *Note:* This figure depicts the cumulative distribution functions of the root mean square errors of the individual models included in the online pooled forecast and the offline benchmark. The solid line is the distribution based on models using offline data, and the dashed one includes models using online indices.
Source: Authors.

equal to 2, 3, and 4 lags such that,

$$p_t = a + \phi_1 p_{t-1} + \phi_2 p_{t-2} + \cdots + \phi_n p_{t-n} + \epsilon_t \tag{A.1}$$

Models with $n > 1$ do not show any advantage over $n = 1$.

A.3.2. Phillips curve

The Phillips-curve model uses n CPI lags and the last month's unemployment rate, u_{t-1}. Thus,

$$p_t = a + \phi_1 p_{t-1} + \phi_2 p_{t-2} + \cdots + \phi_p p_{t-n} + b\, u_{t-1} + \epsilon_t \tag{A.2}$$

The conclusions of this paper do not change using the seasonally-adjusted or non-seasonally adjusted unemployment rate, so we only report results using the seasonally adjusted values.

A.3.3. Random walk

Similar to Ang et al. (2007), Atkeson and Ohanian (2001), Stock and Watson (2001) and Stock and Watson (2007), we report out-of-sample forecasts from a Random-Walk that averages the last four months of inflation. The model is defined as,

$$p_t = \frac{1}{n} \sum_{s=1}^{n} p_{t-s} \tag{A.3}$$

The results in this paper remain unchanged setting n to 1, 2, or 3.

A.4. Statistical significance of monthly forecast results

See Tables A.1–A.3.

A.5. Cumulative distribution of absolute value of errors

See Fig. A.3.

A.6. Sensitivity of the forecast average

The tables in this section show the RMSE of the baseline and our main benchmark when all the single models in Eq. (3) are included in the pooled forecast. This specification, therefore, averages all models from the 24, 36, and 48 window length, as well as fixed and increasing time window, and single models that do not take into account the CPI lags or the fuel survey.

In other words, the pooled forecasts assume there was no ex-ante analysis to distinguish between good and bad performing models. This is a disadvantageous assumption since most analysts would first restrict the set of models, avoiding cases, for example, where the CPI is not included in the regressions.

Nevertheless, the results on the table suggest that the online series are a useful predictor of the CPI. For example, the 1-month ahead forecast is, on average, 15 percent more accurate than the main offline benchmark, and 5 percent more accurate than the survey of professional forecasters published by Bloomberg.

Tables A.4, A.5, and A.6 show the one-month, two-month, and three-month ahead forecasts, respectively.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ijforecast.2019.04.018. The supplementary material includes the scripts and dataset needed to replicate our main forecasting analysis.

References

Ang, Andrew, Bekaert, Geert, & Wei, Min (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, *54*(4), 1163–1212.

Aparicio, Diego, & Cavallo, Alberto (2018). Targeted price controls on supermarket products. *NBER w24275*.

Atkeson, Andrew, & Ohanian, Lee E. (2001). Are phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review*, *25*(1), 2–11.

Bates, John M., & Granger, Clive W. J. (1969). The combination of forecasts. *Operations Research Quarterly*, *20*, 451–468.

Benkwitz, Alexander, Lütkepohl, Helmut, & Wolters, Jürgen (2001). Comparison of bootstrap confidence intervals for impulse responses of german monetary systems. *Macroeconomic Dynamics*, *5*(01), 81–100.

Bernanke, Ben S. (2007). Inflation expectations and inflation forecasting. In *Speech at the monetary economics workshop of the national bureau of economic research summer institute*. Cambridge, Mass., July 2007.

Bertolotto, Manuel (2018). Online Price Index with Automated Product Replacement: The Closest-Match Approach. Available at SSRN.

Boesler, Matthew for Bloomberg (2016). The best illustration of Feds inflation relations with markets. March 16, 2016.

Bring, Johan (1996). A geometric approach to compare variables in a regression model. *The American Statistician*, *50*(1), 57–62.

Brynjolfsson, Erik, & Smith, Michael D. (2000). Frictionless commerce? A comparison of Internet and conventional retailers. *Management Science*, *46*(4), 563–585.

Bureau of Labor Statistics (2015). The consumer price index. Handbook of Methods, Chapter 17: Updated June 2015.

Cavallo, Alberto (2013). Online and official price indexes: measuring Argentina's inflation. *Journal of Monetary Economics*, *60*(2), 152–165.

Cavallo, Alberto (2017). Are online and offline prices similar? Evidence from large multi-channel retailers. *The American Economic Review*, *107*(1), 283–303.

Cavallo, Alberto (2019). Scraped data and sticky prices. *Review of Economics and Statistics*, *100*(1), 105–119.

Cavallo, Alberto, & Rigobon, Roberto (2016). The billion prices project: Using online prices for measurement and research. *The Journal of Economic Perspectives*, *30*(2), 151–178.

Cecchetti, Stephen G., Hooper, Peter, Kasman, Bruce C., Schoenholtz, Kermit L., & Watson, Mark W. (2007). Understanding the evolving inflation process. *US Monetary Policy Forum*, *8*.

Chen, Xiaohong, Racine, Jeffrey, & Swanson, Norman R. (2001). Semiparametric ARX neural-network models with an application to forecasting inflation. *IEEE Transactions on Neural Networks*, *12*(4), 674–683.

Clarida, Richard, Gali, Jordi, & Gertler, Mark (1999). The science of monetary policy: a new keynesian perspective. *Journal of Economic Literature*, *37*, 1661–1707.

Clements, Michael P., & Galvão, Ana Beatriz (2013). Forecasting with vector autoregressive models of data vintages: US output growth and inflation. *International Journal of Forecasting*, *29*(4), 698–714.

Croushore, Dean (2010). An evaluation of inflation forecasts from surveys using real-time data. *The BE Journal of Macroeconomics*, *10*(1).

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

Eickmeier, Sandra, & Ziegler, Christina (2008). How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach. *Journal of Forecasting*, *27*(3), 237–265.

Faust, Jon, & Wright, Jonathan H. (2013). Forecasting inflation. *Handbook of Economic Forecasting*, *2*(Part A), 3–56.

Forni, Mario, Hallin, Marc, Lippi, Marco, & Reichlin, Lucrezia (2003). Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, *50*(6), 1243–1255.

Ghysels, Eric, Santa-Clara, Pedro, & Valkanov, Rossen (2004). The MIDAS touch: Mixed data sampling regression models. *Finance*.

Gorodnichenko, Yuriy, & Talavera, Oleksandr (2017). Price setting in online markets: Basic facts, international comparisons, and cross-border integration. *The American Economic Review*, *107*(1), 249–282.

Granger, Clive W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*.

Gurkaynak, Refet S., Sack, Brian P., & Swanson, Eric T. (2004). Do actions speak louder than words? The response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*, *1*, 55–94.

Gürkaynak, Refet S., Sack, Brian, & Wright, Jonathan H. (2010). The TIPS yield curve and inflation compensation. *American Economic Journal: Macroeconomics*, 70–92.

Hall, Peter (1992). *The bootstrap and Edgeworth expansion*. Springer.

Harvey, David, Leybourne, Stephen, & Newbold, Paul (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting*, *13*(2), 281–291.

Haubrich, Joseph, Pennacchi, George, & Ritchken, Peter (2012). Inflation expectations, real rates, and risk premia: evidence from inflation swaps. *Review of Financial Studies*, hhs003.

Hausman, Jerry, & Leibtag, Ephraim (2009). CPI bias from supercenters: does the BLS know that Wal-Mart exists? In *Price index concepts and measurement* (pp. 203–231). University of Chicago Press.

Hongyi Li, G. S., & Maddala (1996). Bootstrapping time series models. *Econometric Reviews*, *15*(2), 115–158.

International Labour Office (2004). Consumer price index manual: theory and practice. Geneva: 2004.

Kehoe, Patrick, & Midrigan, Virgiliu (2015). Prices are sticky after all. *Journal of Monetary Economics*, *75*, 35–53.

Knotek, Edward S., & Zaman, Saeed (2015). Nowcasting US Headline and Core Inflation. Federal Reserve Bank of Cleveland, Working Paper 14-03R.

Levy, Daniel, Bergen, Mark, Dutta, Shantanu, & Venable, Robert (1997). The magnitude of menu costs: direct evidence from large US supermarket chains. *Quarterly Journal of Economics*, 791–825.

Lünnemann, Patrick, & Wintr, Ladislav (2011). Price stickiness in the US and europe revisited: Evidence from internet prices. *Oxford Bulletin of Economics and Statistics*, *73*(5), 593–621.

Lütkepohl, Helmut (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Marcellino, Massimiliano, Stock, James H., & Watson, Mark W. (2003). Macroeconomic forecasting in the euro area: Country specific versus area-wide information. *European Economic Review*, *47*(1), 1–18.

Marcellino, Massimiliano, Stock, James H., & Watson, Mark W. (2006). A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics*, *135*(1), 499–526.

Modugno, Michele (2013). Now-casting inflation using high frequency data. *International Journal of Forecasting*, *29*(4), 664–675.

Monteforte, Libero, & Moretti, Gianluca (2013). Real-time forecasts of inflation: The role of financial variables. *Journal of Forecasting*, *32*(1), 51–61.

Nakamura, Emi (2005). Inflation forecasting using a neural network. *Economics Letters*, *86*(3), 373–378.

Nakamura, Emi, & Zerom, Dawit (2010). Accounting for incomplete pass-through. *Review of Economic Studies*, *77*(3), 1192–1230.

Rigobon, Roberto, & Sack, Brian (2004). The impact of monetary policy on asset prices. *Journal of Monetary Economics*, *51*(8), 1553–1575.

Stock, James H., & Watson, Mark W. (1999a). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In R. Engle, & H. White (Eds.), *Cointegration, causality and forecasting: A festschrift for clive W.J. Granger* (pp. 1–44). Oxford: Oxford University Press.

Stock, James H., & Watson, Mark W. (1999b). Forecasting inflation. *Journal of Monetary Economics*, *44*(2), 293–335.
Stock, James H., & Watson, Mark W. (2001). Vector autoregressions. *Journal of Economic perspectives*, 101–115.
Stock, James H., & Watson, Mark W. (2003). Forecasting output and inflation: the role of asset prices. *Journal of Economic Literature*, *28*, 204–430.
Stock, James H., & Watson, Mark W. (2006). Forecasting with many predictors. *Handbook of Economic Forecasting*, *1*, 515–554.
Stock, James H., & Watson, Mark W. (2007). Why has US inflation become harder to forecast? *Journal of Money, Credit and Banking*, *39*(s1), 3–33.
Stock, James H., & Watson, Mark W. (2009). Phillips curve inflation forecasts. In Jeffrey Fuhrer, Yolanda Kodrzycki, Jane Little, & Giovanni Olivei (Eds.), *Understanding inflation and the implications for monetary policy* (pp. 99–202). Cambridge: MIT Press.
Timmermann, Allan (2006). Forecast combinations. *Handbook of Economic Forecasting*, *1*, 135–196.
Woodford, Michael (2011). *Interest and prices: foundations of a theory of monetary policy*. Princeton University Press.
Wright, Jonathan H. (2009). Forecasting US inflation by Bayesian model averaging. *Journal of Forecasting*, *28*, 131–144.
Yellen, Janet L. (2017). Inflation, uncertainty, and monetary policy. *Business Economics*, *52*(4), 194–207.

Diego Aparicio is an Assistant Professor at the IESE Business School.

Manuel I. Bertolotto is the Chief Operating Officer at PriceStats, where he leads the economics, software engineering, and data sourcing teams. His research interests include the development and monitoring of real-time financial indicators using novel data sources and distributed systems, the continuous improvement of inflation and purchasing power parity indicators, and forecasting. Manuel holds a Ph.D. in Economics from Universidad de San Andrés.