Red Nacional de Investigadores en Economía

Are professional forecasters inattentive to public discussions? The case of inflation in Argentina

J. Daniel Aromí (UBA-CONICET)

Martín Llada (UBA-CONICET)

DOCUMENTO DE TRABAJO N° 300

Enero de 2024

Citar como:
Aromí, J. Daniel y Martín Llada (2024). Are professional forecasters inattentive to public discussions? The case of inflation in Argentina. Documento de trabajo RedNIE N°300.

Are professional forecasters inattentive to public discussions? The case of inflation in Argentina*

J. Daniel Aromí
IIEP-Baires UBA-Conicet and FCE, Universidad Católica Argentina

Martín Llada
IIEP-Baires UBA-Conicet

September 10, 2022

Abstract

We evaluate whether professional forecasters incorporate valuable information from public discussions on social media. The study covers the case of inflation in Argentina for the period 2016-2022. We find solid evidence consistent with inattention. A simple indicator of attention to inflation on social media is shown to anticipate professional forecast errors. A one standard deviation increment in the indicator is followed by a rise of 0.4% in mean forecast errors in the subsequent month and by a cumulative increment of 0.7% over the next six months. Furthermore, social media content anticipates significant revisions in forecasts that target multiple months ahead inflation and calendar year inflation. These findings are different from previously documented forms of inattention. Consistent results are verified implementing out of sample forecasts and using content from an alternative social network. The study has implications for the use of professional forecasts in the context of policy-making and sheds new evidence on the nature of imperfect information in macroeconomics.

1 Introduction

Expectations play a central role in macroeconomic analysis. In particular, expectations are key for monetary policy-making (Bernanke 2007, Sims 2009, Adam 2007, Coibon

*We would like to thank the comments from seminar participants at Rednie, the Complex Systems Workshop (FCE UBA) and FCE UCA.

et al. 2020). In this context, professional forecasts constitute an important input that yields information on expectations and, more broadly, on macroeconomic conditions. The adequate use of this data requires a precise understanding of its properties. One traditional option is to assume full information rational expectations (FIRE). But a growing body of literature has provided evidence inconsistent with this traditional alternative. These contributions have documented information rigidities, underreaction to incoming information and overreaction or delayed over-extrapolation in professional forecasts (Coibon & Gorodnichenko 2015, Angeletos et al. 2021, Bordalo et al. 2020, Aromí 2018).

In this work, we contribute to this empirical literature evaluating a novel form of inattention. We assess whether professional forecasts incorporate information that can be extracted from public discussions. With this objective, in a preliminary step, we summarize the level of attention assigned to inflation on the microblogging platform Twitter. Then, we evaluate if, as expected under the inattention hypothesis, this indicator is able to anticipate forecast errors and forecast revisions. The study covers the case of inflation in Argentina for the period 2016-2022.

Our analyses provide strong support to the inattention hypothesis. We find that the index of attention to inflation on social media anticipates inflation forecast errors. According to the baseline specification, an increment of one standard deviation in the index anticipates an increment of 0.4% in the mean error of one-month-ahead forecasts. The estimated impulse-response function indicates that, six months after a one standard deviation shock to attention, mean cumulative forecast errors increase by 0.7%. Extended empirical models show that this form of inattention is different from the previously reported evidence on positively correlated forecast revisions (Coibon & Gorodnichenko 2015).

Moving beyond one month ahead forecast errors, we find that social media content also anticipates revisions in forecast. That is, further evidence on inattention results from inspecting changes in forecasts that target multiple months ahead inflation and, also, updates in forecasts that target calendar year inflation. Consistent with the inattention hypothesis, increments in attention to inflation on social media anticipate higher inflation forecast for this ample range of horizons.

We extend the analysis, and evaluate the robustness of the results, through out of sample forecasts. These exercises show that simple forecasting models that learn about the historic pattern of inattention can be used to anticipate forecast errors and forecast revisions in subsequent dates. This robustness exercise can be also interpreted as solid evidence on the persistence of inattention or, in other words, the absence of a fast learning process that eliminates the divergence from full information rational expectations.

For a proper interpretation of the results and to assess their relevance in alternative context, it is important to observe particular features of the Argentine economy. The case under study involves a macroeconomic regime characterized by fluctuating and elevated levels of inflation. In this context, inflation constitutes an important concern and is a common topic of public discussions. As a consequence, professional forecasters and the public have incentives to set a high level of attention to inflation hoping to reduce the elevated levels of uncertainty. Hence, according to this observation, inattention to public discussions could be conjectured to be unimportant. On the other hand, it must be noted that social media data can be costly to process and its value for forecasting purposes is, in principle, uncertain. Hence, significant levels of inattention to public discussions remains a plausible hypothesis. In other words, empirical analysis is needed to resolve the ambiguity resulting from these contrasting observations.

The empirical evaluations developed in the current study are motivated by an important theoretical literature that allows for deviations from FIRE that are explained by information stickiness, constrained perceptual channels or out of equilibrium strategic reasoning (Mackowiak & Wiederholt 2009, Mankiw & Reis 2002, Angeletos & La'O 2009, Farhi & Werning 2019). From a different but related perspective, our evaluations are also motivated by the understanding of inflation as an evolutionary process driven by forward looking, adaptive behavior (Heymann & Leijonhufvud 1995, De Grauwe & Ji 2019). According to this perspective, the value of social media data could be explained by its capacity to reflect information regarding the emergence collective beliefs and behavior.

This work contributes to an important set of empirical studies that document properties of macroeconomic forecasts (Coibon & Gorodnichenko 2015, Andrade & Le Bihan 2013, Fuhrer 2018, Angeletos et al. 2021, Bordalo et al. 2020). In particular, these contributions have documented evidence consistent with inattention to new information in the form of positive correlation of aggregate forecast errors or forecast revisions. In this study we document a novel pattern consistent with inattention: data on public discussions can be used to anticipate forecast errors and forecast revisions.

From an applied perspective, this paper is related to analyses of macroeconomic policy-making and its relationship with indicators of expectations. Expectations can be linked to policymaking in two manners. First, expectations constitute an important input that informs monetary policymaking (Batini & Haldane 1999, Levin et al. 2003, Orphanides 2004). Additionally, policymakers communicate decisions to influence expectations which are conjectured to determine behavior and outcomes (Melosi 2017, Miranda-Agrippino & Ricco 2021). The evidence reported in this paper suggests that policymakers trying to learn about expectations and the state of the macroeconomy need to go beyond survey forecast. In particular, policymaking can benefit from novel sources of information such as public discussion data on social media.

In terms of its methodology, this work is part of a group of studies that have used text as data to produce valuable indicators in macroeconomic settings and financial settings (Tetlock 2007, Baker et al. 2016, Aromí 2020, Gardner et al. forthcoming). In particular, Twitter content has been used to produce valuable indicators of inflation and inflation expectations (Angelico et al. 2022, Aromi & Llada 2020). Extending this literature, this paper shows that professional forecasters fail to incorporate the information that can be extracted processing text on social media.

This work is organized as follows. In the following section data and methodology are presented. Professional forecasts are evaluated in section 3. An analysis using data of public discussions on an alternative social media is presented in section 4. Next, an out-of-sample forecast exercise is detailed. The last section presents concluding remarks.

2 Data and methodology

Professional forecasts correspond to those collected by the survey conducted by the Argentine Central Bank (Banco Central de la República Argentina).[1] After the normalization of official economic statistics in 2016, the first release of the survey corresponds to June 2016. In the first exercises implemented in this paper we use median monthly inflation forecasts for one-month-ahead forecasts. In the extended analysis, we also consider median forecasts of monthly inflation for up to 6 months ahead and median forecasts of calendar year inflation for up to two years ahead. The last observation used in this study corresponds to the March 2022 release.

[1] Relevamiento de Expectativas de Mercado (REM): http://www.bcra.gob.ar.

Starting in June 2016 inflation data correspond to INDEC´s Consumer Price Index[2]. Inflation data for earlier periods are used in preliminary analyses of the information content of social media indicators. In this case, due to concerns regarding data quality, inflation data for these earlier dates correspond to that provided by the statistics agency of the City of Buenos Aires (9/2012-5/2016) and the consulting firm Buenos Aires City (1/2012-8/2012).

Social media content, Twitter messages, correspond to Tweets from users that self-reported “Argentina” in the free-text section of their user profile. The tweets were collected from two sources. One source is the ”sample” stream that provides 1% of tweets in realtime.[3] This source was complemented with tweets extracted through Twitter API´s ”user timeline” endpoint.[4] The database comprises 240 million tweets.

We summarize social media content implementing a very simple and transparent strategy. First, for each month, we count the number of mentions of the word inflation (“inflación” or “inflacion” or “inflacionario” in Spanish). Then, the value of the index in that month equals the ratio of mentions of inflation to the number of words (tokens) in tweets of the corresponding month.[5] This index can be interpreted as an indicator of attention allocated to inflation. Considering likely structural breaks on social media content, we also consider an alternative specification in which the original index is adjusted by historical values observed in the previous 12-month period. More formally, the original specification of the index is given by:

$$I_t = \frac{\#\, mentions\, inflation}{\#\, words} * 1000 \tag{1}$$

The alternative specification is given by:

$$\hat{I}_t = I_t - \sum_{k=1}^{12} I_{t-k}/12 \tag{2}$$

[2]https://www.indec.gob.ar/indec/web/Nivel4-Tema-3-5-31

[3]This information was accessed through Internet Archive's “Twitter Stream Grab”, https://archive.org/details/twitterstream.

[4]More specifically, Argentine users identified inspecting the first source (“sample” stream tweets) were randomly selected and their tweets were collected using “user timeline” endpoint.

[5]We also considered an alternative specification in which words that can be linked to inflation concerns are incorporated in the analysis. More specifically, we computed the sum of the original index and the time series corresponding to the frequency of the following words: ”dólar” (dollar), ”precio” (price) and ”salario” (wage). The results under this alternative specification are very similar to the results observed under the original methodology.

Table 1 shows descriptive statistics. During the sample period, monthly inflation averaged 2.7% with a standard deviation of 0.9%. According to the index of social media content, inflation is mentioned on average, 0.32 times per ten thousand words. As expected from the elevated levels of inflation in Argentina, this number is high when compared to selected countries. More precisely, the mean value of the indicator of social media attention to inflation in Argentina is about 4 times the average level observed in a sample of Brazilian tweets and 10 times the average level observed in a sample of tweets corresponding to Chile.

Additionally, the last two rows of Table 1 show descriptive statistics for one-month-ahead forecasts and one-month-ahead forecast errors. During the sample period, the average one-month-ahead forecast of the inflation rate was 2.7% and the standard deviation was 0.9%. Forecast errors were of a similar scale as indicated by a standard deviation of 0.8%, a maximum value of 2.6% and a minimum of -1.8%.

Figure 1 shows the evolution of the three main indicators used in this study. This figure shows the co-movement between the three indicators. More specifically, substantial increments in the attention index coincide with episodes of important increments in inflation forecasts and the inflation rate. Coincident spikes can be detected in the second half of 2018, the year 2019 and the instances of inflation acceleration that took place by the end of the sample period. Additionally, it is worth noting that the coefficient of correlation between inflation and the attention index is 0.54 while the coefficient of correlation between inflation and inflation forecasts is 0.62.

3 Forecast errors and public discussions on social media

In the first exercise we show that the indicator of attention to inflation on social media contains information regarding future inflation. More specifically, we estimate a simple autoregressive model that is extended to incorporate an indicator of social media attention to inflation. The model is given by:

$$\Delta cpi_{t+1} = \alpha_0 + \alpha_1 \Delta cpi_t + \beta I_t + u_{t+1} \tag{3}$$

where Δcpi_t is the monthly inflation rate, I_t is an index of social media content and u_{t+1} is the error term. Columns 2 and 3 in table 2 show the estimated coefficients for different specifications of the index. According to these estimated models, a one standard deviation increment in the corresponding social media indicator is followed by an increment in mean inflation that is between 0.3% and 0.4%.

During the sample period there were multiple instances of large devaluations that were followed by accelerations of the inflation rate. Given this feature, we also consider an extended forecasting model in which the rate of devaluation is incorporated as a predictor. The estimated models, summarized in columns 4 through 6 of table 2, provide similar conclusions. These findings is reported in more detail in Aromi & Llada (2020).

Having established that the index provides relevant information regarding future levels of inflation, the next task involves evaluating the extent to which professional forecasters are able to incorporate this information in a timely manner. This evaluation is carried out through simple tests of forecast conditional bias. If forecasters are inattentive to social media data, the indicator that summarizes this data could be used to anticipate forecast errors or forecasts revisions. A standard evaluation of professional forecasters is implemented through an empirical model in which mean forecast errors are a function of lagged social media content. Let $F_t cpi_{t+1}$ represent the one-month-ahead forecast at time t of the inflation rate at time $t+1$. The associated forecast error is given by $fe_{t|t+1} = \Delta cpi_{t+1} - F_t cpi_{t+1}$. Then, the evaluation model is given by:

$$fe_{t|t+1} = \alpha + \beta I_t + u_{t+1} \tag{4}$$

Where $fe_{t|t+1}$ is the forecast error of the median one-month-ahead forecast that is released in period t, I_t is an index of social media content and u_{t+1} is the error term. If forecasts incorporate the information on social media, the mean forecast error is independent of the indicator I_t. Table 3 shows the fitted models. According to the estimated coefficients, an increment in the value of the index capturing social media attention is associated with larger mean forecast errors in the following month. Interestingly, the estimated coefficient for the index is between 0.003 and 0.004. These values are similar to those corresponding coefficients reported for the case of inflation forecasts presented in the previous table. This is consistent with professional forecasters that, to a large extent, ignore the information provided by social media content.[6]

It could be conjectured that the variation in attention is particularly informative in

[6] In an extended analysis, we evaluate if the results change when we consider that professional forecasts are reported a few days before each month ends. We found that there are no significant changes in the results when the indicator of Twitter content is computed excluding these days that follow the date in which forecasts are reported.

the case of increments in attention. To evaluate this conjecture, column 3 in table 3 reports the estimated coefficient for a simple form of nonlinearity. Under this specification, mean forecast errors are responsive to attention only when it is higher than the values observed, on average, during in the previous 12 months. The estimated coefficient and adjusted R^2 do not suggest that there are gains to be made under this alternative specification.[7]

To evaluate if the finding reported above is closely related to information rigidities that are manifested in the form of correlated revisions (Coibon & Gorodnichenko 2015), the empirical model is extended to include lagged forecast revisions. Let $F_{t-k}cpi_t$ be the k-month-ahead forecast at time $t-k$ of the inflation rate at time t. More specifically, we incorporate a new predictor that is given by: $rev_{t|t+1} = F_t cpi_{t+1} - F_{t-1} cpi_{t+1}$. The corresponding estimated coefficients are reported in columns 5 through 7. According to the reported figures, there are no strong indications of the presence of information rigidity in the form of correlated forecast revisions. On the other hand, the evidence on inattention to social media content remains, for the most part, unaltered in this extended model. In other words, the evidence points to a new form on imperfect information in macroeconomics: inattention to public discussions.

A richer description of the association between forecast errors and social media content is generated estimating a simple VAR model and computing the associated impulse response functions. The VAR implementation is a parsimonious one lag specification where the impulse response functions are computed using a Cholesky decomposition in which it is assumed that a shock in the index has a contemporaneous impact on forecast errors but forecast errors have no impact on social media attention. This is a plausible assumption since inflation is reported reported more than 10 days after the corresponding month has ended.[8]

Figure 2 shows the cumulative response of one-month-ahead forecast errors to a shock in the inattention index. According to the fitted values, starting with the shock in month 1, cumulative forecast errors increase in a persistent manner and, after six months, the cumulative response is 0.7%. In this way, social media content is shown to provide information goes beyond that observed in the subsequent month. In the

[7]In another evaluation, we classified messages as future oriented or backward-looking using sets of keywords. The results did not point to any gain in information content resulting from this strategy.

[8]Qualitatively similar results are observed under the alternative ordering in the Cholesky decomposition.

next subsection, additional evaluations of forecast associated to longer horizons and the revisions of those forecasts are presented.

3.1 Longer forecast horizons and forecast revisions

Professional forecasts collected by the Argentine Central Bank include monthly inflation forecasts for horizons that range from one up to six months ahead. In this way, moving beyond the anticipation of forecast errors, we can evaluate whether social media information provides information regarding the revision of multiple months ahead forecasts. In addition, the collected forecast include expected inflation for the current calendar year and the next two calendar years. Exploiting this rich set of point forecast, we report a complementary characterizations of professional forecasts that extends the previous analysis to the relevant case of longer horizons.

Considering revisions in monthly forecasts first, let $F_{t-k}cpi_t$ represent the k-month-ahead forecast that is released at the end of month $t-k$ and targets the inflation rate in month t. Then, in month $t-k$, fixed event forecast revisions are given by: $rev_{t-k|t} = F_{t-k}cpi_t - F_{t-k-1}cpi_t$ for $k \in \{0, 5\}$. In the case of $k = 0$, we set $F_t cpi_t = cpi_t$ and, in this way, the particular revision, $rev_{t-1|t}$, coincides with the one-month-ahead forecast error. The analysis considers the case of the variation in expected inflation for six months windows. That is, we add the six revisions realized once a new edition of REM is released ($\sum_{h=0}^{5} rev_{t|t+h}$) and evaluate whether this change in expectations can be anticipated using social media data. We also consider the case in which the revision in the immediately following month excluded, that is, we evaluate the ability to anticipate $\sum_{h=1}^{5} rev_{t|t+h}$. As in the previous analysis, under the null hypothesis of fully informed rational expectations, the mean revision cannot be anticipated using information available at the time the original forecasts were released.

Beyond monthly forecasts, we also consider the revision of calendar year forecasts. We let $rev_{t|y0}$ represent month t revision of current calendar year forecast and $rev_{t|yk}$ represent month t revision of expected inflation k calendar years ahead. For example, $rev_{2020/07,1}$ represents the forecast revision of inflation for year 2021 that is realized at the end of July 2020. This is given by the difference between the one-year-ahead forecast released in July 2020 and June 2020.

As suggested in previous discussions, the most convenient way to aggregate social media information is unknown and needs to be evaluated empirically. In line with this observation, the empirical analysis is implemented allowing for different specifications of the metric of social media inflation attention. We consider the inflation attention index (I_t) and the variation in the attention index with respect to the average value in the previous 12 months ($\hat{I}_t$). Considering that these information associated to this longer terms forecasts might arrive at a different rate, we also propose a lower frequency metric of change in the level of attention. More specifically, the new specification is equal to the difference between the average value of the index in the most recent semester versus the average value in the preceding five year window. Formally, this metric is given by: $\hat{I}_t^{lf} = \sum_{h=0}^{5} I_{t-h}/6 - \sum_{h=6}^{65} I_{t-h}/60$.

As in the previous analysis, a simple model is proposed to evaluate if social media information is adequately incorporated in inflation forecasts. In this empirical model, the expected revision is dependent on the value of the lagged metric of social media attention. Formally, the univariate model is given by:

$$Revision_{t+1} = \alpha + \beta Attention_t + u_{t+1} \tag{5}$$

where $Revision_{t+1}$ is one of the proposed metrics of forecast revisions that is realized in month $t+1$ and $Attention_t$ is one of the proposed metrics of social media attention to inflation computed for month t and u_{t+1} is a noise term. The coefficient of interest is β. Under the null hypothesis of efficient forecasts, the value of this coefficient is zero. That is, lagged social media content cannot be used to anticipate forecast revisions.

Table 4 reports, for each specification of the social media content indicator and for the alternative forecast horizons, the estimated value of the coefficient of interest, $\hat{\beta}$. In line with the previously reported results, the estimations point to a positive association between social media content and subsequent forecast revisions. According to these estimations, increments in attention to inflation on social media content is followed by an upward adjustment in monthly and calendar year forecasts. For example, in the case of cumulative six-months-ahead forecasting windows, the estimated increment is between 0.6% and 0.8%. Larger estimated coefficients can be observed in the case of revisions of yearly inflation forecasts.

Beyond these evaluations of forecast efficiency, the dynamic association between social media inflation attention can be described in more comprehensive manner estimating VAR models and computing impulse-response functions. As in the previous section, we estimate a parsimonious bi-variate model with one lag. In the Cholesky decomposition it is assumed that forecast revisions have no contemporaneous impact on social media attention. This is a plausible assumption since forecasts are reported by the end of each month.

The estimated impulse response functions show that social media content provides significant information regarding the trajectory of forecast revisions. A shock to the inflation attention index is followed, on average, by a significant and persistent upward shift in forecasts. Figure 3 shows the case of cumulative six-month-ahead forecasts, six months after the shock, the mean cumulative revision increases by 1.72%. In the case of cumulative three-year-ahead forecasts revision, the mean increment after six months of the shock is estimated at 2.67% (see, Figure 4). These estimated cumulative responses provide complementary evidence that strengthens the previously reported results regarding inattention in inflation forecasts.

3.2 Evidence for alternative social media

In this study we analyze professional forecasts inattention to public discussions. Twitter content was used to construct a metric that estimates attention to inflation in public discussions. One relevant aspect is whether the findings are robust to changes in the source of data used to construct the indicator of public discussions. That is, can the findings reported above be reproduced using alternative data of emerging public attitudes?

To answer this question we implement a similar exercise for the case of an alternative social network. We analyze the content of "Reddit Argentina"[9], a discussion forum with 333 thousand members at the time of this writing. This is a small number compared to the estimated 5 millions of monthly active users of Twitter in Argentina[10]. Hence, while this evaluation is informative, a weaker association is to be expected.

We used Pushshift Reddit API[11] to collect the full texts of submissions and comments corresponding to the discussion forum. The resulting corpus comprises 293 million tokens. Attention to inflation is proxied as the monthly ratio of mentions of inflation to the total number of tokens. As in the case of Twitter data, we consider multiple specifications of the indicator of social media content. Also, as in the previous section, expert forecast are evaluated using the standard linear model in which forecast errors or

[9]https://www.reddit.com/r/argentina/.
[10]https://www.statista.com/statistics/558273/number-of-twitter-users-in-argentina/
[11]https://github.com/pushshift/api

revisions are a function of lagged indicators of social media content.

As a first exploratory exercise, we verify that over the period 6/2016-4/2021 the correlation between Reddit and Twitter indices is high: 0.70. Then, consistent with the previously reported findings, table 5 shows that, according to the estimated models, Reddit content provides information regarding subsequent forecast errors and forecast revisions. It must be noted that the estimated coefficients are smaller. Also, in multiple cases, the null hypothesis of a zero coefficient cannot be rejected. The most plausible explanation for this difference is the small number of active users and the associated smaller dataset in the case of this alternative social network. That is, beyond this weaker results, this complementary evaluation suggests that the previously reported evidence on inattention is robust to changes in the source of public discussions data.

4 Out-of-sample Forecast

To provide further insights on inattention to public discussions we implement out-of-sample forecast exercises in which models are trained recursively with past information. In this way, an alternative, more robust evaluation of expert forecast inattention is produced. In this exercise, the performance of two models of forecast errors or forecast revisions are compared. Under the first model, consistent with full information rational expectations, expected forecast errors and expected forecast revisions are zero. In other words, according to this baseline model, forecast errors cannot be anticipated. In contrast, the second model allows for inattention, that is, mean forecast errors and mean forecast revisions are a function of an indicator of social media attention to inflation. According to this second empirical approach forecast error (or forecast revision) is modeled as:

$$Fore_{t+1} = \beta I_t + u_{t+1} \tag{6}$$

Where $Fore_{t+1}$ is the corresponding forecast error or forecast revision and, is in the previous analysis, I_t is one of the three different specifications of the indicator of inflation attention. ($\{I_t, \hat{I}_t$ or $\hat{I}_t^{sf}\}$. Each specification of this model is trained recursively with data corresponding to an expanding window. The first model is estimated to predict forecast errors or forecast revisions observed in February 2020 using data corresponding to the training window 2016/07-2019/12. Then, in each subsequent recursion of the prediction exercise, the target and the upper bound of the training window are moved forward one period.

For each model specification and each prediction target, performance is evaluated computing the root-mean-square prediction error (RMSE). For the models that allow for inattention, this measure of accuracy is also expressed as a fraction of the RMSE of the baseline model. Considering the uncertainty surrounding the optimal specification of the indicator of social media content, the performance of forecast combinations are also computed. More specifically, forecasts of the different models that allow for inattention are averaged with equal weights and, then, the performance of the resulting forecast combination is compared to that resulting from the baseline model.

Table 6 shows the results of these out-of-sample forecast exercises. The evidence indicates that historical patterns consistent with inattention can be used to learn about systematic errors observed in future periods. This conclusion can be verified checking both the case of monthly forecasts and calendar year forecasts. The last row of the table shows that forecast combinations allow in most of the cases substantial gains in forecast accuracy. In summary, these out-of-sample forecast exercises provide further support to the idea that social media content provides valuable information regarding future forecast errors and forecast revisions. This finding is inconsistent with full information rational expectations and indicates that professional forecasters are inattentive to public discussions.

5 Conclusions

This work evaluates professional inflation forecasts for the case of Argentina. The analysis shows that forecasters fail to incorporate information on public discussions that can be summarized from social media. The results are economically significant and are corroborated under multiple robustness tests. In the baseline evaluation, a one standard increment in the social media indicator anticipates an increment of 0.4% in average one-month-ahead forecast error. Furthermore, similar evidence is found in the case of multiple months ahead forecast and calendar year forecasts. Impulse response functions point to a substantial and persistent dynamic association. This form of inattention is different from the previously reported evidence on positively correlated forecast revisions.

Given the prominent role given to expectations and professional forecasts, this evidence is clearly relevant for the analysis and design of monetary policy. This novel form of inattention delivers a new warning regarding the breadth of imperfect information in macroeconomic contexts. At the same time, the reported regularities points to a new form of data that can inform policymakers about the state of the macroeconomy: decentralized public discussions on social media.

There are several directions in which this analysis can be extended. In the current work social media data is aggregated using a very simple and transparent strategy. While there are benefits from this approach, it leaves space for extension in which alternative data processing strategies are evaluated. For example, the network structure in social media could be exploited to evaluate if there exist communities in which public discussions are particularly informative. Complementarily, natural language processing models could be implemented to provide richer characterizations of public discussions.

Finally, having in mind the special characteristics of the Argentine macroeconomic regime during the sample period, one natural question that can be tackled in future work is to evaluate under which conditions the regularities reported in this work are manifested by professional forecasts corresponding to other economies. In a related manner, it would be of interest to verify whether the inattention manifested by professional forecasters in Argentina persists in subsequent years.

References:

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

Angeletos, G. M., Huo, Z., & Sastry, K. A. (2021). Imperfect macroeconomic expectations: Evidence and theory. NBER Macroeconomics Annual, 35(1), 1-86.

Angeletos, G. M., & La'O, J. (2009). Incomplete information, higher-order beliefs and price inertia. Journal of Monetary Economics, 56, S19-S37.

Aromí, J. D. (2018). GDP growth forecasts and information flows: Is there evidence of overreactions?. International Finance, 21(2), 122-139.

Aromi, J. D. (2020). Linking words in economic discourse: Implications for macroeconomic forecasts. International Journal of Forecasting, 36(4), 1517-1530.

Aromi, J. D. & Llada, M. (2020). Forecasting Inflation with Twitter. Reunión Anual de la AAEP.

Andrade, P., & Le Bihan, H. (2013). Inattentive professional forecasters. Journal of Monetary Economics, 60(8), 967-982.

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

Batini, N., & Haldane, A. (1999). Forward-looking rules for monetary policy. In Monetary policy rules (pp. 157-202). University of Chicago Press.

Bernanke, B. (2007). "Inflation Expectations and Inflation Forecasting." Speech to the Monetary Economics Workshop of the NBER Summer Institute, Cambridge, Massachusetts, 10 July.

Bordalo, P., Gennaioli, N., Ma, Y., & Shleifer, A. (2020). Overreaction in macroeconomic expectations. American Economic Review, 110(9), 2748-82.

Bomberger, W. A. (1996). Disagreement as a Measure of Uncertainty. Journal of Money, Credit and Banking, 28(3), 381-392.

Coibion, O., & Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. American Economic Review, 105(8), 2644-78.

De Grauwe, P., & Ji, Y. (2019). Inflation targets and the zero lower bound in a behavioural macroeconomic model. Economica, 86(342), 262-299.

Fuhrer, J. C. (2018). Intrinsic expectations persistence: evidence from professional and household survey expectations. Available at SSRN 3296152.

Coibion, O., Gorodnichenko, Y., Kumar, S., & Pedemonte, M. (2020). Inflation expectations as a policy tool?. Journal of International Economics, 124, 103297.

Farhi, E., & Werning, I. (2019). Monetary policy, bounded rationality, and incomplete markets. American Economic Review, 109(11), 3887-3928.

Gardner, B., Scotti, C., & Vega, C. (2021). Words speak as loudly as actions: Central bank communication and the response of equity prices to macroeconomic announcements. Journal of Econometrics.

Heymann, D., & Leijonhufvud, A. (1995). High Inflation: The Arne Ryde Memorial Lectures. OUP Catalogue.

Lahiri, K., & Sheng, X. (2010). Measuring forecast uncertainty by disagreement: The missing link. Journal of Applied Econometrics, 25(4), 514-538.

Levin, A., Wieland, V., Williams, J. C. (2003). The performance of forecast-based monetary policy rules under model uncertainty. American Economic Review, 93(3), 622-645.

Mackowiak, B., & Wiederholt, M. (2009). Optimal sticky prices under rational inattention. American Economic Review, 99(3), 769-803.

Mankiw, N. G., & Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the New Keynesian Phillips curve. The Quarterly Journal of Economics, 117(4), 1295-1328.

Melosi, L. (2017). Signalling effects of monetary policy. The Review of Economic Studies, 84(2), 853-884.

Miranda-Agrippino, S., & Ricco, G. (2021). The transmission of monetary policy shocks. American Economic Journal: Macroeconomics, 13(3), 74-107.

Orphanides, A. (2004). Monetary policy rules, macroeconomic stability, and inflation: A view from the trenches. Journal of Money, Credit and Banking, 151-175.

Rossi, B., & Sekhposyan, T. (2016). Forecast rationality tests in the presence of instabilities, with applications to Federal Reserve and survey forecasts. Journal of Applied Econometrics, 31(3), 507-532.

Sims, C.. (2009). Inflation Expectations, Uncertainty and Monetary Policy BIS Working Paper No. 275.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. The Journal of finance, 62(3), 1139-1168.

Zhang, F. (2017). Rational inattention in uncertain business cycles. Journal of Money, Credit and Banking, 49(1), 215-253.

Table 1: Descriptive statistics

Sample period is 02/2012-03/2022. Data frequency is monthly. Δcpi_t: rate of monthly change in the Consumer Price Index. I_t is the index of attention to inflation, $F_{t-1}cpi_t$: one-month-ahead forecast at time $t-1$ of the inflation rate at time t. $fe_{t-1|t}$: one-month-ahead inflation forecast error.

Variable	Obs.	Mean	St. Dev.	Min	Q1	Q3	Max
cpi_t	122	0.02707	0.01199	0.00202	0.01888	0.03419	0.06729
I_t	122	0.03159	0.01779	0.00763	0.01810	0.03929	0.09623
$F_{t-1}cpi_t$	69	0.02730	0.00943	0.01300	0.01850	0.03400	0.05820
$fe_{t-1\|t}$	69	0.00213	0.00808	-0.01798	-0.00198	0.00639	0.02609

Table 2: Inflation forecast models

Sample period is 2012-2022. cpi_t: monthly inflation rate (CPI). er_t: monthly rate of devaluation. I_t : ratio between total mentions of the term “inflation” and total number of words. $\hat{I}_t$: variation in the inflation attention index with respect to the average value in the previous 12-month period. Standardized coefficients. Standard errors are estimated following Newey & West (1987, 1994).*p<0.1; **p<0.05; ***p<0.01

	(1)	(2)	(3)	(4)	(5)	(6)
cpi_t	0.008***	0.006***	0.007***	0.007***	0.005***	0.006***
	(11.6)	(6.2)	(10.4)	(7.4)	(5.7)	(7.7)
I_t		0.004***			0.004***	
		(3.8)			(3.9)	
$\hat{I}_t$			0.003***			0.003***
			(3.5)			(3.4)
er_t				0.003***	0.001***	0.001***
				(3.2)	(4.9)	(3.6)
Constant	0.027***	0.027***	0.027***	0.026***	0.027***	0.027***
	(38.9)	(36.5)	(34.9)	(36.1)	(37.9)	(36.1)
Observations	121	121	121	121	121	121
Adjusted R^2	0.434	0.520	0.488	0.465	0.549	0.511

Table 3: Evaluation of forecast errors

Estimated coefficients for different specifications of model 4. Sample period is 07/2016-03/2022. $rev_{t|t+1}$: revision at time t of the one-month-ahead forecast of the inflation rate. I_t : inflation attention index. $\hat{I}_t$: variation in the inflation attention index with respect to the average value in the previous 12-month period. Standardized coefficients. $\hat{I}_t^+ = max[0, \hat{I}_t]$: positive variation of the attention index with respect to the value in the previous 12-month period. Standardized coefficients. Standard errors are estimated following Newey & West (1987, 1994).*p<0.1; **p<0.05; ***p<0.01

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$rev_{t\|t+1}$	-	-	-	0.0023	0.0014	0.0012	0.0016
				(1.3)	(0.8)	(1.0)	(1.1)
I_t	0.0032***	-	-	-	0.0027**	-	-
	(2.8)				(2.3)		
$\hat{I}_t$	-	0.0039***	-	-	-	0.0036***	-
		(3.5)				(3.5)	
$\hat{I}_t^+$	-	-	0.0032***	-	-	-	0.0027**
			(2.7)				(2.6)
Constant	0.0021*	0.0024**	0.0021*	0.0022*	0.0021**	0.0022**	0.0022**
	(1.8)	(2.4)	(1.9)	(1.9)	(2.1)	(2.5)	(2.3)
Adjusted R^2	0.146	0.227	0.141	0.067	0.155	0.233	0.160

Table 4: Forecast revisions and social media attention

The table reports the estimated coefficient, $\hat{\beta}$, for model 5. Sample period is 07/2016-03/2022. I_t: inflation attention index. $\hat{I}_t$: variation in the inflation attention index with respect to the average value in the previous 12-month period. $\hat{I}_t^{lf}$: difference between the average value of inflation attention index in the most recent semester and the average value in the preceding five year window. $rev_{t|t+1}$: one-month-ahead forecast error. $\sum_{h=1}^{5} rev_{t|t+h}$: sum of forecast revisions for horizons that range from 1 up to 5 months ahead. $\sum_{h=0}^{5} rev_{t|t+h}$: sum of six forecast revisions for horizons that range from 1 up to 5 months ahead. $rev_{t|y0}$: revision in current year inflation forecast. $\sum_{k=1}^{2} rev_{t|yk}$: sum of forecast revisions for one-year-ahead and two-years-ahead horizons. $\sum_{k=0}^{2} rev_{t|yk}$: sum of revisions of forecasts for current year and next two years. Standard errors are estimated following Newey & West (1987, 1994).*p<0.1; **p<0.05; ***p<0.01

	$rev_{t\|t+1}$	$\sum_{h=1}^{5} rev_{t\|t+h}$	$\sum_{h=0}^{5} rev_{t\|t+h}$	$rev_{t\|y0}$	$\sum_{k=1}^{2} rev_{t\|yk}$	$\sum_{k=0}^{2} rev_{t\|yk}$
I_t	0.0032***	0.0031***	0.0064***	0.0067***	0.0079**	0.0137***
	(2.8)	(2.7)	(3.6)	(2.8)	(2.8)	(3.5)
Adj. R^2	0.146	0.038	0.102	0.051	0.053	0.048
$\hat{I}_t$	0.0039***	0.0020	0.0060***	0.0049	0.0023	0.0059
	(3.5)	(1.4)	(2.7)	(1.4)	(0.5)	(0.9)
Adj. R^2	0.227	0.007	0.088	0.019	-0.010	-0.005
$\hat{I}_t^{lf}$	0.0024	0.0052***	0.0076***	0.0111***	0.0086***	0.0201***
	(1.6)	(3.3)	(4.3)	(6.3)	(2.8)	(5.9)
Adj. R^2	0.072	0.131	0.152	0.163	0.069	0.124

Table 5: Forecast revisions and alternative social media attention

The table reports the estimated coefficient, $\hat{\beta}$, for model 5. Sample period is 07/2016-03/2022. I_t: inflation attention index. $\hat{I}_t$ variation in the inflation attention index with respect to the average value in the previous 12-month period. $\hat{I}_t^{lf}$: difference between the average value of inflation attention index in the most recent semester and the average value in the preceding five year window. $rev_{t|t+1}$: one-month-ahead forecast error. $\sum_{h=1}^{5} rev_{t|t+h}$: sum of five forecast revisions. $\sum_{h=0}^{5} rev_{t|t+h}$: sum of six forecast revision. $rev_{t|y0}$: one-year-ahead forecast error. $\sum_{k=1}^{2} rev_{t|yk}$: sum of two revisions of calendar year forecast $\sum_{k=0}^{2} rev_{t|yk}$: sum of three revisions of calendar year forecast. Standard errors are estimated following Newey & West (1987, 1994).*p<0.1; **p<0.05; ***p<0.01

	$rev_{t\|t+1}$	$\sum_{h=1}^{5} rev_{t\|t+h}$	$\sum_{h=0}^{5} rev_{t\|t+h}$	$rev_{t\|y0}$	$\sum_{k=1}^{2} rev_{t\|yk}$	$\sum_{k=0}^{2} rev_{t\|yk}$
I_t	0.0024*	0.0021*	0.0044**	0.0034	0.0019	0.0038
	(1.7)	(1.8)	(2.1)	(1.4)	(0.5)	(0.6)
Adj. R^2	0.072	0.008	0.042	0.001	-0.012	-0.013
$\hat{I}_t$	0.0027**	0.0010	0.0037**	0.0014	-0.0030	-0.0031
	(2.4)	(1.1)	(2.3)	(0.6)	(0.6)	(0.5)
Adj. R^2	0.101	-0.010	0.025	-0.013	-0.007	-0.014
$\hat{I}_t^{lf}$	0.0005	0.0034	0.0039	0.0071*	0.0085**	0.0167***
	(0.4)	(1.7)	(1.4)	(2.2)	(2.5)	(2.9)
Adj. R^2	-0.012	0.044	0.028	0.045	0.061	0.065

Table 6: Out of sample forecast of forecast errors and forecast revisions

The table reports the RMSE of forecast model estimated following the equation 6. Sample period is 07/2016-03/2022. I_t: inflation attention index. $\hat{I}_t$ difference of the attention index with respect to average values in the previous 12 months. $\hat{I}_t^{lf}$: difference between the average value of inflation attention index in the most recent semester and the average value in the preceding five years window. $rev_{t|t+1}$: one-month-ahead forecast error. $\sum_{h=1}^{5} rev_{t|t+h}$: sum of five forecast revisions. $\sum_{h=0}^{5} rev_{t|t+h}$: sum of six forecast revision. $rev_{t|y0}$: one-year-ahead forecast error. $\sum_{k=1}^{2} rev_{t|yk}$: sum of two revisions of calendar year forecast $\sum_{k=0}^{2} rev_{t|yk}$: sum of three revisions of calendar year forecast. Forecast combination are implemented through simple averages.

	$rev_{t\|t+1}$	$\sum_{h=1}^{5} rev_{t\|t+h}$	$\sum_{h=0}^{5} rev_{t\|t+h}$	$rev_{t\|y0}$	$\sum_{k=1}^{2} rev_{t\|yk}$	$\sum_{k=0}^{2} rev_{t\|yk}$
RMSE-REM	0.0078	0.0051	0.0107	0.0171	0.0376	0.0478
I_t	0.0075	0.0086	0.0128	0.0217	0.0267	0.0411
(ratio)	0.95	1.68	1.193	1.27	0.71	0.86
$\hat{I}_t$	0.0067	0.0059	0.0099	0.0199	0.0393	0.0521
(ratio)	0.86	1.15	0.92	1.168	1.05	1.09
$\hat{I}_t^{lf}$	0.0072	0.0061	0.0092	0.0177	324	0.0402
(ratio)	0.91	1.18	0.86	0.96	0.86	0.84
Combination	0.0066	0.0057	0.0090	0.0177	0.0311	0.0404
(ratio)	0.85	1.12	0.84	1.03	0.83	0.84

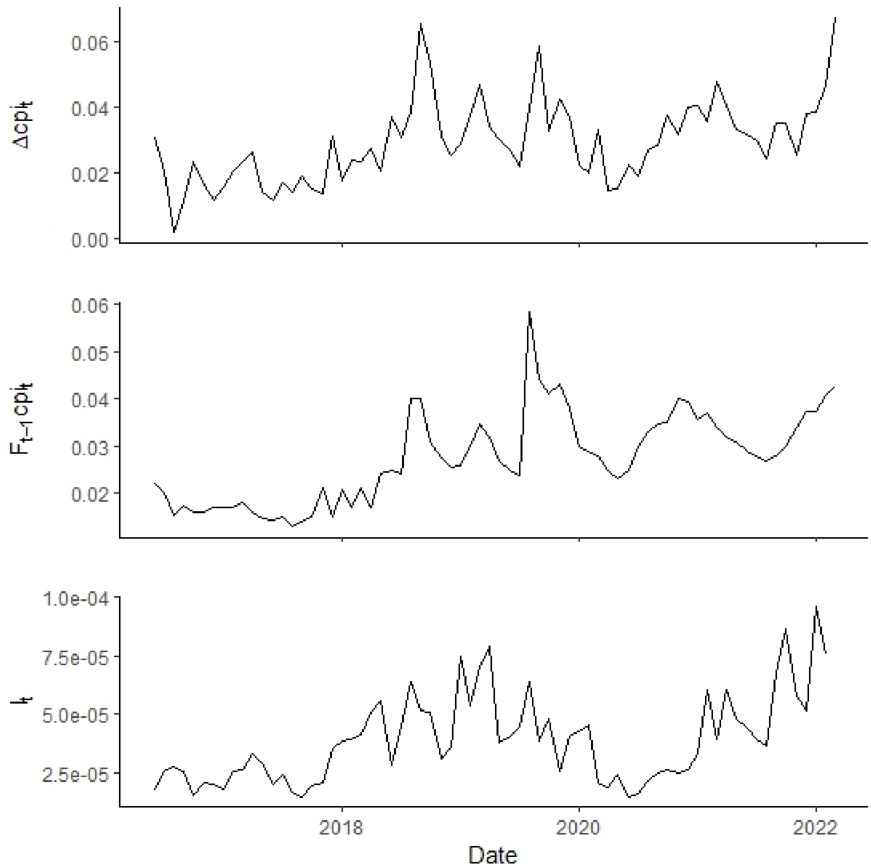

Figure 1: Inflation (Δcpi_t), forecast ($F_{t-1}cpi_t$) and Twitter attention index (I_t)

.

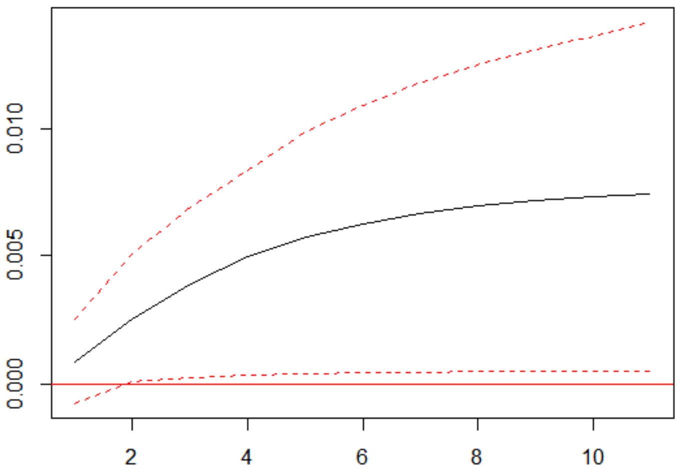

Figure 2: Impulse response function: cumulative response of one month ahead forecast errors ($fe_{t|t+1}$) to a shock on social media inflation attention (I_t).
Note: 95% bootstrap confidence intervals.

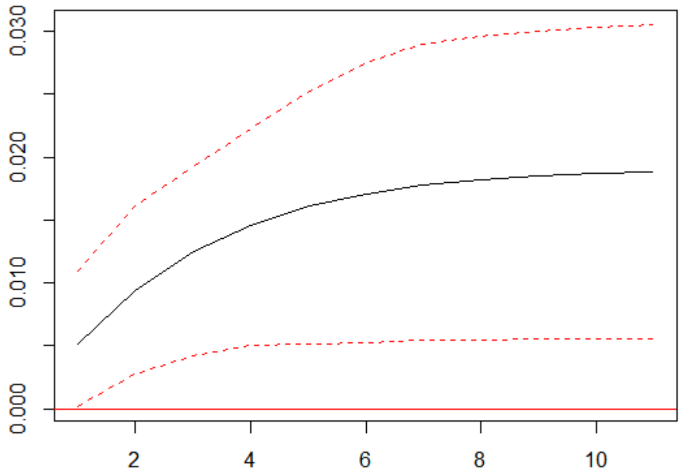

Figure 3: Impulse response function: Cumulative response of forecast revisions for inflation over the next six months ($\sum_{h=0}^{5} rev_{t|t+h}$) to a shock on social media inflation attention (I_t).

Note: 95% bootstrap confidence intervals.

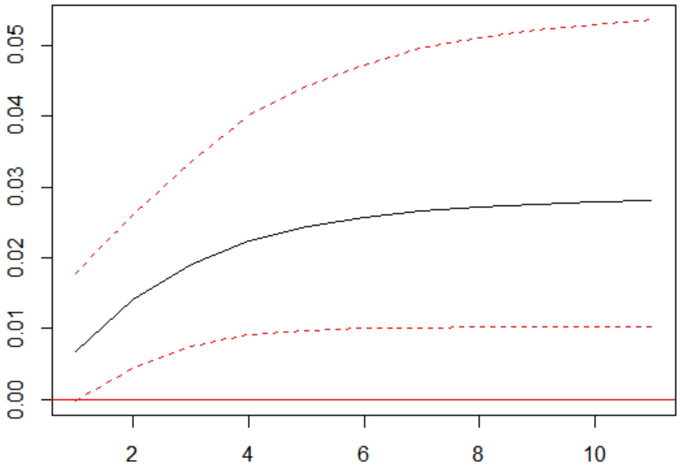

Figure 4: Impulse response function: Cumulative response of forecast revisions for inflation over the next three calendar years ($\sum_{k=0}^{2} rev_{t|yk}$) to a shock on social media inflation attention (I_t).

Note: 95% bootstrap confidence intervals.