Banco de México

Working Papers

N° 2023-21

The Influence of Central Bank's Projections and Economic Narrative on Professional Forecasters' Expectations: Evidence from Mexico

Arturo Antón Sarabia
Banco de México

Santiago Bazdresch
Banco de México

Alejandra Lelo-de-Larrea
Banco de México

December 2023

Documento de Investigación
2023-21

Working Paper
2023-21

The Influence of Central Bank's Projections and Economic Narrative on Professional Forecasters' Expectations: Evidence from Mexico*

Arturo Antón Sarabia†
Banco de México

Santiago Bazdresch‡
Banco de México

Alejandra Lelo-de-Larrea§
Banco de México

Abstract: This paper evaluates the influence of central bank's projections and narrative signals provided in the summaries of its Inflation Report on the expectations of professional forecasters for inflation and GDP growth in the case of Mexico. We use the Latent Dirichlet Allocation model, a text-mining technique, to identify narrative signals. We show that both quantitative and qualitative information have an influence on inflation and GDP growth expectations. We also find that narrative signals related to monetary policy, observed inflation, aggregate demand, and inflation and employment projections stand out as the most relevant in accounting for changes in analysts' expectations. If the period of the COVID-19 pandemic is excluded, we still find that forecasters consider both types of information for their inflation expectations.
Keywords: Central bank projections, economic forecasting, machine learning, text mining
JEL Classification: E52, E58, C55

Resumen: Este documento evalúa la influencia de las proyecciones y de las señales narrativas incluidas en los resúmenes del Informe de Inflación del banco central sobre las expectativas de inflación y crecimiento del PIB de los pronosticadores profesionales en el caso de México. Usamos el modelo de Asignación Latente de Dirichlet, que es una técnica de minería de texto, para identificar las señales narrativas. Mostramos que la información cuantitativa y cualitativa influye en las expectativas de inflación y crecimiento del PIB. También encontramos que las señales narrativas sobre política monetaria, inflación observada, demanda agregada y proyecciones de inflación y empleo serían las que más toman en cuenta los analistas para sus expectativas. Si se excluye el periodo de la pandemia de COVID-19, hallamos que los pronosticadores siguen considerando ambos tipos de información en sus expectativas de inflación.
Palabras Clave: Proyecciones de bancos centrales, pronósticos económicos, aprendizaje de máquina, minería de textos

*We thank Paulina Anciola, Gustavo Hernández, Jorge Herrera, Luis Pech and Jorge Silva for data construction, interpretation and classification work. We also thank two anonymous Banco de México's referees for their detailed comments.

† Dirección General de Investigación Económica. Email: arturo.anton@banxico.org.mx.

‡ Dirección General de Investigación Económica. Email: cbazdreschb@banxico.org.mx.

§ Dirección General de Investigación Económica. Email: alejandra.lelo@banxico.org.mx.

1 Introduction

In most economies around the globe, central banks (CB) issue their own projections regarding the future evolution of the economy. In this capacity, as a public economic forecaster, CB provide economic agents with a view about the future. In addition, these forecasts serve to some extent as a scenario upon which monetary policy decisions are based, and thus contribute to the broad description of the CB's policies and stance. On the other hand, CB also generate qualitative information about the economic environment that may be deemed useful by economic agents. In an era in which information is so easily and widely distributed, it has become highly relevant for CB to have a clear sense of the impact that the publication of their own projections and qualitative views may have on the economy. In fact, an increasingly common monetary policy regime among CB in advanced and emerging economies is inflation targeting, in which agents' expectations about inflation, and generally about the CB's policy function, play a central role. In this context, it is especially important to improve our understanding of how CB's projections and "soft" information may be taken into consideration by private agents in their own expectations.

In this paper we examine the role of CB's projections and qualitative information in accounting for changes in professional analysts' expectations in Mexico. For that purpose, we use the information provided in the summaries of *Banco de México*'s (BM) key analytical publication, the Inflation Report (IR).[1] Although a wealth of information is compiled and processed by highly specialized staff at CB, identifying which information in their publications is beyond the one already known by economic analysts is a challenging task. Typically, experts are well-informed and it is possible that they might react modestly to the provision of information by a CB. Moreover, the IR in many cases is not accompanied by explicit policy decisions, so for the most part these projections are not "monetary policy news" in the sense that a monetary policy statement is. Therefore, in order to assess whether the information generated by the CB have an influence on public expectations, we focus on current- and next year

[1] For maximum clarity, we refer to CB's forward expected trajectories for the variables of interest respectively as "projections" and to those of the private sector as "forecasts" or "expectations".

projections for inflation and GDP growth. We also study the text of the IR summaries to extract their qualitative information (henceforth, their "narrative signals"). We restrict our analysis to the period 2008Q1-2021Q2. It starts in 2008Q1 because BM adopted the Overnight Interbank Interest Rate as the bank's main policy instrument in January 2008, and thus this date represents the time when the main elements of the current monetary policy strategy were adopted. It ends in 2021Q2 because that is another breaking point in central bank policies. At that point, BM began to publish its inflation projections in its monetary policy decision statement ahead of their publication in the IR. Thus, at that point the IR was no longer the main vehicle by which the central bank issues its projections.[2]

In a first stage, we test if the quantitative information provided by the CB through inflation and GDP growth projections have an influence on the corresponding forecasts elaborated by professionals. Specifically, at this stage we are interested in evaluating in what proportion the gap between CB projections and experts' forecasts for variable x at year T is incorporated by experts into their forecast revision. To account for changes in private agents' expectations, we rely on experts' forecasts for inflation and GDP growth right before and after the IR is published. The source of this information is the monthly survey of professional forecasters collected by BM. Given that private expectations may be affected by new economic information available within the period of two consecutive surveys, we control for data "surprises" in our empirical specification. We also include additional controls such as experts' forecasts for US GDP growth, the nominal exchange rate, and the short-run interest rate on government bonds.

In a second stage, we add the qualitative information included in the IR summaries to our previous specification, and evaluate to what extent it is taken into account by forecasters in their own expectations. To this end, we use the Latent Dirichlet Allocation (LDA) model of Blei et al. (2003), first introduced in the economics literature by Hansen et al. (2014). LDA is a Machine Learning algorithm in which words and paragraphs from a given text are related to topics based on the probability that each word from a paragraph belongs to a topic. In this

[2] In addition, the short time period does not allow for enough observations to fully account for this change.

paper, we implement the LDA model with 26 topics under two alternative approaches for the identification of narrative signals. Under the first approach, we select the paragraphs in the IR summary particularly associated with either inflation or GDP growth. Next, we use a dictionary of our own based on the work of Tobback et al. (2017) to estimate a tone index on these paragraphs and compute the mean tone and its change for each IR. The tone index is meant to represent whether the text conveys either higher or lower inflationary pressures, or stronger or weaker economic dynamism. In principle, text that expresses higher inflationary pressures might lead private forecasters to increase their expected inflation, while text referring to stronger economic dynamism might lead them to increase their expected level of output. To explore the impact of the tone as well as its change, we compute specifications with either the mean tone or its change as potential explanatory variables in an empirical model that also includes quantitative information. An advantage of this approach is that the econometric model may be estimated with standard methods. However, its weakness is that it may be ignoring the content in the IR not related to inflation and GDP growth, which may provide additional valuable information to market participants for shaping their expectations. Thus, under the second approach, we use all 26 topics identified in the *full* IR and apply the tone index to measure their mean tone and its change. We follow Hansen et al. (2019) to estimate "narrative shocks" and test if the whole text of the IR contains "news" that is distinct from the information provided by the CB projections. Given that we have 54 IR and 52 narrative shocks to estimate, we address this problem with the Elastic Net Regression of Zou and Hastie (2005). Finally, we identify the four most important topics and add them to the regression that only includes quantitative information to test if they have additional explanatory power for changing experts' expectations. Even though the second approach is relatively more complex, it has the advantage of letting all of the text provided in the IR have an impact, and is also more robust to changes in the econometric specification.

It is worth noting that the advantage of the LDA model over lexicon-based techniques is that the former is designed to find the best association between words and topics based on the principle that words and documents may be related to latent topics. As a result, assigning a predetermined context for words or sequences of words is not required, and thus the process lets "the data speak for itself", avoiding any possible researcher bias or need for personal judgement. To the best of our knowledge, this is the first paper that evaluates the role of CB's narrative signals for explaining changes in private forecasts through the lens of the LDA model.[3]

Our main results are as follows. First, we find that, in general, CB projections may have an influence on professional forecasters' expectations. Our estimates suggest that roughly 25% of the difference between CB projections and private sector forecasts for current-year inflation is translated into changes in private inflation expectations when the next survey is collected. For example, if the CB projection is currently 10 percentage points above the median value expected by private forecasters, for the next survey the specialists will raise their forecast by 2.5 percentage points on average, *ceteris paribus*. This number decreases to about 10% if next-year inflation forecasts are considered instead. For GDP projections, we estimate that about 39% of the difference between the CB and private forecasts for current-year GDP growth is incorporated by experts in their revised expectations. However, when dealing with next-year GDP growth forecasts, the impact of CB projections on expectations is not consistently significant. Second, our empirical findings suggest that the IR's qualitative information is also important in accounting for changes in agents' expectations in general. Under the first approach to account for qualitative information, we report that the tone related to inflationary pressures or its corresponding change have an influence on next-year inflation expectations but not current-year expectations. For economic activity, either its tone or its change seem to have an impact on next- or current-year GDP growth expectations, respectively. For the second approach, we find that narrative signals have an influence on experts' forecasts for current- and next-year inflation, as well as for next-year GDP growth. We also note that narrative topics related to monetary policy, observed inflation, aggregate demand, and inflation and formal employment projections stand out as the most relevant to be taken into account by experts for

[3] The literature review in section 2 describes how the LDA model has been used previously under other contexts in the economics literature. A detailed description of the model is presented in section 3.1.2.

their own expectations.

To evaluate if our results are sensitive to the economic environment, we perform similar exercises for a sample that excludes the COVID-19 period. While both samples exhibit similar standard deviations for current- and next-year inflation expectations and for next-year GDP growth, we observe striking differences in current-year GDP growth expectations. Specifically, its standard deviation is 40% larger in the full sample, suggesting that forecasters experienced a significantly higher uncertainty about their own short-run GDP growth expectations due to the pandemic.[4] Regarding inflation expectations, we report that the differences between CB projections and experts' forecasts in accounting for changes in current- and next-year inflation expectations are quantitatively similar to those obtained under the full sample. Furthermore, we do not observe changes on the relevance of narrative signals. In contrast, regarding GDP growth expectations, we find that the difference between the CB and forecasters for explaining changes in these expectations is larger in the sample that excludes the COVID-19 period. Moreover, in the shorter sample narrative signals have an explanatory power for current-year GDP growth expectations. Overall, these findings suggest that private forecasters may systematically incorporate CB's information into their inflation expectations, while its impact on GDP growth expectations may be sensitive to the economic context. In this paper we do not analyze in further detail the determinants behind the differentiated impact of CB's projections and narrative signals. However, the Results section provides food for thought regarding the potential reasons for these differences.

Our estimates about the relevance of the gap between CB and private sector forecasts in accounting for changes in professional forecasters' expectations can be compared directly with the findings of Pedersen (2015) for Chile and of de Mendonça and de Deus (2019) for Brazil, Poland and Mexico. In general, our empirical results are in line with, or imply smaller impacts than, those found by Pedersen (2015) and de Mendonça and de Deus (2019). What distinguishes our work from theirs is the analysis of the role of narrative signals that

[4]As discussed previously, our sample ends in 2021Q2. Therefore, it does not include most of the sharp (and mostly non-anticipated) raise in inflation experienced since the middle of 2021.

is provided in this paper. In this regard, empirical evidence suggests that CB's narrative signals are relevant for private forecasters once CB's quantitative information is controlled for (cf. Ullrich (2008), López Marmolejo (2013), Montes et al. (2016), Hubert (2017), Anderes et al. (2021) and Baranowski et al. (2021)). However, relative to the existing literature on the influence of qualitative information, in which the standard process is to rely on a predefined context for words or a sequence of words from the researcher, in this paper we show that it is possible to use non-supervised techniques for text analysis that do not require defining a specific context ex-ante.

The rest of the paper is structured as follows. First, Section 2 presents a short review of the literature on CB information provision and its impact on agents' expectations. Section 3 describes the elements of BM's IR summaries and the data from the survey of professional forecasters. We also detail the quantitative information extracted from the IR and the methodology implemented to transform the IR's narrative signals into numerical variables. Section 4 presents the econometric specification and the procedure to compute the tone related to both inflationary and economic activity pressures and to extract the narrative shocks. In Section 5 we present the results and identify the main topics that contribute to accounting for changes in professional forecasters' expectations. Finally, Section 6 presents our final remarks and conclusions.

2 Literature Review

Currently, there is a consensus about the importance of CB's information as a mechanism of transparency and accountability, and as a monetary policy tool (see, for example, Blinder et al. (2008), Woodford (2005), Hayo and Neuenkirch (2010), El-Shagi and Jung (2015), and Lustenberger and Rossi (2020)). In this regard, a series of benefits have been identified in the literature. First, CB information can help to improve the accuracy of market participants' forecasts (Hayo and Neuenkirch (2010), El-Shagi and Jung (2015), and Jung and Kuehl (2021)). Also, CB information can reduce financial markets volatility, especially during periods of high uncertainty (see Ehrmann and Fratzscher (2009) and Hayo et al. (2012)), and transparency can indeed mitigate uncertainty (Naszodi et al. (2016)). Lastly, some studies highlight that a solid communication strategy is of special importance for inflation targeters, since their policy actions are unambiguously related to inflation forecasts (see Fracasso et al. (2003) and Svensson (2010)). Nevertheless, the literature also points out the risks that can result from the provision of too much information. Österholm et al. (2008) note that more information can be in detriment of the functions and credibility of the CB. For instance, economic agents might interpret future actions of the CB as a promise; therefore, small deviations from these actions due to unexpected shocks may be enough to diminish these agents' trust in the institution. Finally, some studies suggest that a "noisy" communication can increase uncertainty in financial markets and affect macroeconomic forecasts (Coenen et al. (2017), Lustenberger and Rossi (2020)).

The specific question on the impact of an increase in CB transparency on the economy has also been an important focus of the literature, at least since the time of the rational expectations revolution, for example, in Barro (1976). Under the most basic paradigm, the provision of more information by the CB to the public reduces the overall level of uncertainty and thus, increases welfare. For example, Tarkka and Mayes (1999) show that the publication of unconditional inflation forecasts could lead to greater predictability and less output variability. Indeed, one assumption for the first welfare theorem to hold is the presence of complete information. However, other authors describe models in which the provision of more information by the CB does not improve welfare. For example, Morris and Shin (2002) show that, in the context of a principal that receives a noisy signal that serves to coordinate heterogeneous agents, more transparency leads to higher volatility in the average action, which is harmful to the principal. Gersbach (2003) and Cukierman (2001) also present models where CB information disclosure may not be beneficial because in order to stabilize output, the CB needs inflation to exceed expected inflation in the event of negative supply shocks, and thus in that setting it would be optimal to hide information about such shocks after expectations have been set.

An important part of the revolution in CB's communication and transparency has to do specifically with the publication of their economic and policy forecasts. These forecasts, elaborated using complex statistical methods and sometimes taking advantage of data that can only be accessed by the monetary policy authorities, are thought to provide a highly informed view of the economy. Moreover, these forecasts in some cases are built considering the institution's own expectations about its future policy and thus, they may be also important as a window into the likely future decisions of the CB. For all these reasons, economists and financial market analysts carefully follow the publication of CB's forecasts. However, its publication poses important questions for policymakers. For example, it is unclear whether there is a benefit for the economy or for policymakers in reducing the uncertainty around private sector forecasts.

As discussed by Hubert (2015), there are three plausible explanations about why CB forecasts may have an influence on private ones. First, as a result of possessing highly technical staff members, it is possible that CB forecasts may be "better" than the private sector's in terms of yielding lower forecast errors, especially for inflation (see Gamber and Smith (2009), El-Shagi and Jung (2015) and the references therein), which would rationalize private forecasters taking CB forecasts into account. Second, it is possible that due to its data collection responsibilities and infrastructure, there may exists asymmetric information between the CB and private forecasters. For example, the CB may have information relevant for forecasting inflation which is available in the payment system's data, but which is not available to the private sector, or additional information about the future state of the economy more generally (see, for example, Romer and Romer (2000), Nakamura and Steinsson (2018) and Hansen et al. (2019)). Third, CB forecasts may provide policy signals about the future stance of monetary policy. In any of these cases, private agents may adjust their forecasts to the extent that those of the CB have relevant information not available to them.

Regardless of the relative importance of the aforementioned explanations, the literature reports that private agents typically adjust their economic projections after the publication of CB forecasts. Fujiwara (2005), Hubert (2014), Hubert (2015), Pedersen (2015), de Mendonça and de Deus (2019), Jain and Sutherland (2020) and Hattori et al. (2021), among others, find that CB inflation projections may have an influence on private inflation forecasts after controlling for other factors that may also affect private agents' expectations. As described later, our results regarding projections mostly confirm what these papers have found in different contexts. Interestingly, the evidence is that CB forecasts for other variables such as GDP growth and interest rates on government bonds also have an impact on private forecasts (see, for example, Fujiwara (2005), Hubert (2014), Pedersen (2015), de Mendonça and de Deus (2019), and Jain and Sutherland (2020)). Overall, there is supportive evidence that quantitative information in the form of CB projections may lead private agents to review their own forecasts.

Different studies have also examined if qualitative information generated by the CB is also able to have an impact on private sector's forecasts. This type of information is provided through different channels, such as press releases, minutes, IR, and press conferences, and may be delivered simultaneously with "hard" information such as forecasts and monetary policy decisions. Empirical evidence suggests that CB's qualitative information may have an influence on private forecasters' expectations even after controlling for the quantitative content of CB's projections and other relevant variables (see Ullrich (2008), López Marmolejo (2013), Montes et al. (2016), Hubert (2017), Gardner et al. (2021), Anderes et al. (2021), and Baranowski et al. (2021), among others). In these papers, lexicon-based methods are used to quantify the CB tone and content for a given text.[5] Nevertheless, Hansen et al. (2018) acknowledges that accounting for context under lexicon-based methods may be difficult. As discussed by Bholat et al. (2015), these methods are deductive approaches in which a general theory is tested based on a predefined list of words. While this approach takes advantage of its simplicity and applicability, its weakness relies on ignoring words not included in the list that may be informative to test the theory of interest.

In order to assess the usefulness of an alternative that may overcome these difficulties, we

[5]For a discussion on lexicon-based methods, see Algaba et al. (2020).

apply the Latent Dirichlet Allocation (LDA) model in our analysis. A major advantage of this method over lexicon-based techniques is that LDA is a non-supervised algorithm in which words and documents may be related to multiple latent topics. As a result, the algorithm is designed to find the best association between words and topics, in contrast with lexicon-based methods where a pre-defined word lists or groups is required. Therefore, LDA overcomes the difficulties of assigning a proper context to words or sequences of words. This method has been implemented previously to study different aspects of CB communication. For example, the effects of qualitative information provided in monetary policy statements on both market and real economic variables (Hansen and McMahon, 2016); the consequences of transparency on monetary policy makers' deliberations as reflected in the transcripts of monetary policy meetings (Hansen et al., 2018); and the role of both quantitative and qualitative information of inflation reports on market interest rates (Hansen et al., 2019). To the best or our knowledge, this paper is the first to examine the influence of qualitative CB projections on professional forecasters' expectations through the lens of the LDA model.

3 Data

In this section we briefly describe the IR published by BM (known as the "Quarterly Report") and its projections on inflation and GDP growth, as well as the information from BM's survey of professional forecasters and from other data sources. Our sample period covers from 2008Q1 to 2021Q2, with the exception of next-year GDP growth for which CB projections are only available since 2009Q2.[6]

[6]The sample period starts in the first quarter of 2008 because BM adopted the Overnight Interbank Interest Rate as its monetary policy instrument in January, 2008. Before that period, the CB conducted its monetary policy through the cumulative balance of commercial bank reserves with the CB, an instrument known as the "corto". A detailed explanation of the "corto" is available at: https://www.banxico.org.mx/monetary-policy/monetary-policy-implementatio.html

3.1 Inflation Report

The IR is typically one of the main elements of a CB projection strategy. It usually provides the data, model estimations, analysis and forecasts used as inputs for the monetary policy decision, the motivation behind the latest decisions, and the evaluation of risks associated with the economic outlook (Fracasso et al., 2003).

BM publishes its IR on a quarterly basis since 2000 with the following contents: i) a review of both economic and financial outlooks for the foreign and domestic economies; ii) the observed evolution of inflation up to the corresponding quarter; iii) a discussion about the factors affecting inflation; iv) an analysis of the monetary policy decisions taken during the period; and v) the publication of its projections and its assessment on the balance of risks for economic activity and inflation. Additionally, the IR includes informative "Boxes" that analyze a specific topic of interest given the economic context of the corresponding quarter. Along with the IR, BM also publishes a summary of the IR and a deck of slides used during the IR presentation to the public, and gives a press conference following the IR publication. More recently, starting with the 2021Q2 IR, BM also publishes infographics and a visual summary of each IR.[7]

To analyze the impact of BM's projections on analysts' expectations, we use the quantitative and qualitative information provided in the English version of 54 IR summaries published during the period 2008Q1 - 2021Q2.[8] Quantitative information includes BM's numerical projections of headline inflation and GDP growth for the next eight quarters, while qualitative information is provided in the plain text of the summaries. The following sections describe the process for extracting both types of information. It is worth mentioning that the IR for a given quarter is published around 6-8 weeks after the end of that quarter; thus, the IR for a given

[7]All the materials are available at BM's website in Spanish. Some translations to English are available. Translations of the IR full text are available up to the July-September 2018 IR, translations of the presentation are available since the July-September 2010 IR, and translations of the summary are available for each IR. These documents can be found at: https://www.banxico.org.mx/publications-and-press/quarterly-reports/quarterly-reports-prices-banc.html

[8]We select the English versions of the IR summaries given that the text analysis requires the use of dictionaries that are available in English.

quarter is published in the following quarter. For example, the January-March 2021 IR was published on June 2^{nd}, 2021.

3.1.1 Quantitative Information: Banco de México's Forecast

BM publishes projections for several economic variables in each IR. Projection variables include headline and core inflation, GDP growth, job growth in the formal sector, trade balance and the current account. As previously mentioned, we focus on headline inflation and GDP growth projections.

BM has used different ways to present its projections for annual headline inflation across time. For the period 2008Q1-2010Q4, they were reported in terms of an interval for the current quarter and for each of the next seven quarters. For the next seven years (2011Q1-2017Q4), the information was provided through fan charts. Starting with the IR published in the first quarter of 2018, projections are presented simultaneously as point estimates and fan charts for the current and the following seven quarters. We focus on end-year inflation projections for the current and next year. For those cases where an interval is reported, we use the midpoint of the interval. Additionally, for the 2011Q1-2017Q4 period, we compute the inflation projection series as the average of the estimates from five BM's specialists based on their readings from the fan charts.[9]

For GDP growth, projections are published in terms of the annual average growth for the current year and for each of the next two years. In general, the information is provided as an interval. In just a few cases, a point estimate is reported. For our purposes, we use the midpoint of the interval. The information for current-year GDP growth is available for the entire period of analysis. However, for next-year GDP growth the information is available starting 2009Q2.

[9] In general, we find small variations in the estimates reported by these specialists. Specifically, the difference between the highest and lowest estimates for current- and next-year inflation averaged 13 and 18 basis points, respectively. For this reason, we believe that the average of the estimates should not be affected significantly if the number of specialists is increased. A natural question is whether its effect may be affected by the publication format. Unfortunately, the sample size and frequency limit us to adequately perform this analysis. For example, given that there are only four IR per year, we would only have 28 observations for the longest period available (2011Q1-2017Q4).

Projections for the next two years are discarded because they are only reported on a yearly basis.

3.1.2 Qualitative Information: Banco de Mexico's IR summaries

To measure the impact of qualitative information on analysts' expectations, we use the text of IR summaries. Following the methodology of Hansen et al. (2019), we build a high-dimensional set of variables related to the narrative of the IR which, in principle, can be treated as "news" to market participants. This narrative information provides the CB's view about the economic outlook and its risks, as well as the projections for inflation, GDP growth and other variables. In this sense, the CB can signal its view on economic uncertainty. As Hansen et al. (2019) point out, narrative information is much richer than the information provided by numeric forecasts alone, and thus, potentially allows us to capture different signals that the CB sends to market participants.

To collect narrative data, we follow a three-step process. First, we remove headers, page numbers, graphs and references to the boxes and tables. Then, each IR summary is divided into paragraphs. For the 54 IR in the sample, we have a total of 1,089 paragraphs. We then pre-process the text in each paragraph by removing all non-alphabetic terms and *stopwords*.[10] We also convert the text to lower case and transform sequences of terms of special interest into a single term. For example, we replace "economic activity" with "economicactivity". After this, we stem the remaining terms into their linguistic root.[11] As a result, we get a total of 61,136 terms and 2,498 unique terms.

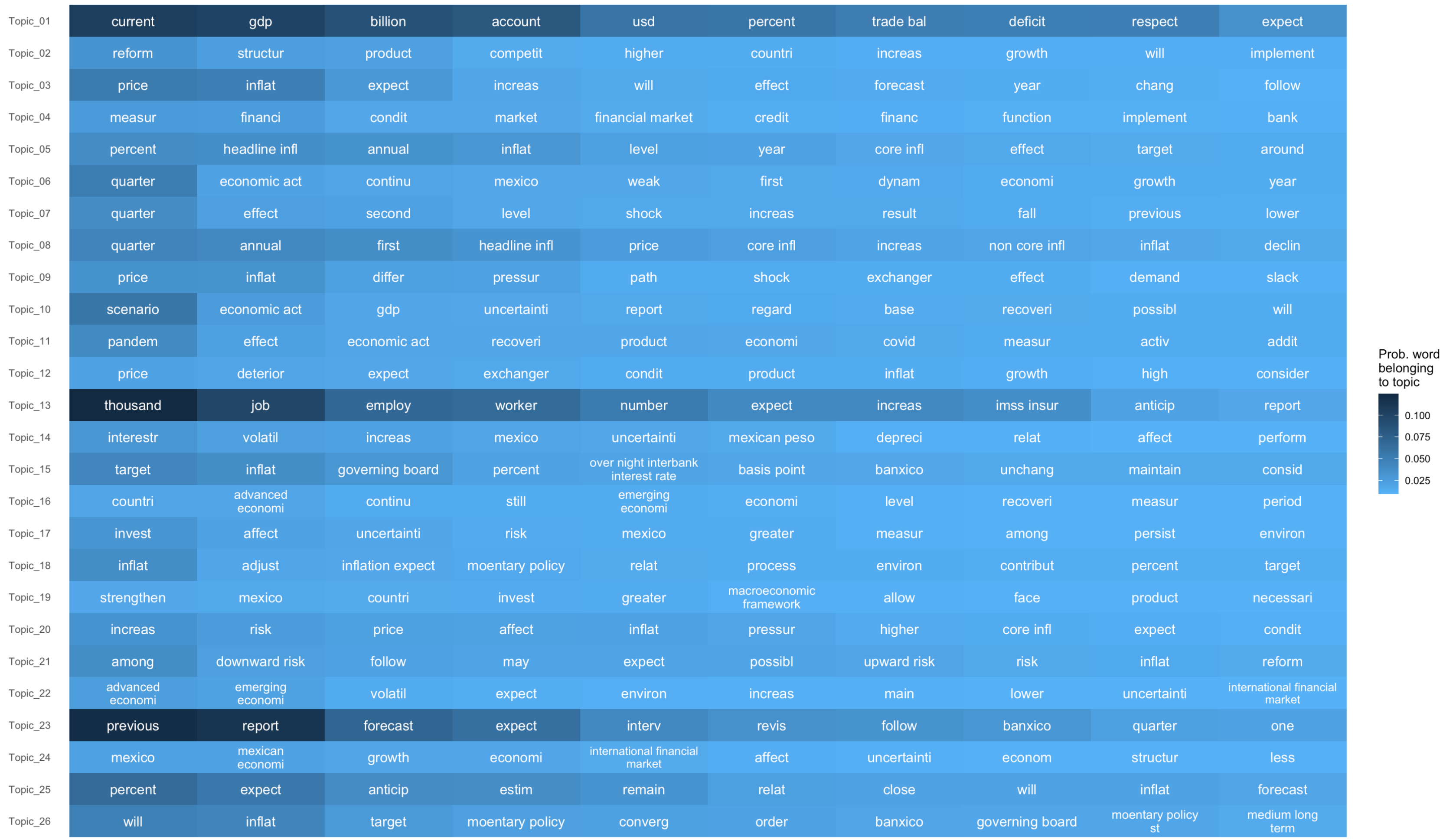
Second, we apply the LDA probabilistic topic model in order to reduce the dimensionality of the narrative information. For each document d, the LDA model estimates a distribution over topics, θ_d, in which each topic k is itself a "probability vector" over the V unique terms in the dataset.[12] Accordingly, let θ_d^k represent topic k's "share" of document d. We exploit

[10] *Stopwords* refer to common words that are meaningless or uninformative, like 'the", "and", and "then".
[11] To stem the terms in the IR we use the Porter's stemming algorithm.
[12] See Blei et al. (2003) and Hansen et al. (2019) for further details.

the variation in the IR by estimating the model at the paragraph level, i.e. each paragraph is considered a separate document. We evaluate a model with a total of $K = 26$ topics over $D = 1,089$ documents and $V = 2,498$ terms.[13] Then, to obtain a topic distribution at the IR level instead of the paragraph level, we obtain the mean values of all the $\theta_{k,t}$ over all the paragraphs of a specific IR, where $\theta_{k,t}$ is the "weight" of topic k for the IR published at time t. Figure 1 shows the 10 terms with the highest probability of being associated with each of the 26 topics in the model. As can be observed, the estimated topics can be interpreted as focusing on identifiable subjects. For example, topics 1, 3, 10, and 13 represent projections for the current account, inflation, economic growth and formal employment, respectively, while topics 5, 11, 15 and 24 are related to inflation convergence, the COVID-19 pandemic, monetary policy and domestic economic activity, respectively. We also compute the change in the mean topic distribution from one IR to the next, $\eta_{k,t} = \theta_{k,t} - \theta_{k,t-1}$, in order to account for changes in the topic's relevance over time. For example, $\eta_{1,2020Q1}$ shows how much more/less relevant is topic 1 in the first quarter of 2020 relative to the fourth quarter of 2019. Therefore, we have a 52-dimensional representation of the information provided in the IR summaries.

[13] To select the number of topics in the model, four specialized readers independently assigned one or more topics to each of the paragraphs of 12 randomly selected IR over the sample period. This procedure resulted in a total of 26 different topics.

Figure 1: Top 10 Terms Within Topics Ranked by Probability



Third, similar to Hansen et al. (2019) and given that LDA does not provide a directional interpretation of each topic, we calculate a measure of the tone used for each topic in every IR. To this end, we first calculate the tone of each paragraph based on our Tone Index (TI), which is created using dictionaries of our own.[14] Next, we assign each paragraph to have at most four topics to account for the possibility that a single paragraph may refer to more than one topic.[15] Finally, we obtain the tone of each topic for each IR, $\theta_{k,t}^{HL}$, by calculating the mean TI for the assigned topics and its corresponding first difference, $\eta_{k,t}^{HL} = \theta_{k,t}^{HL} - \theta_{k,t-1}^{HL}$.[16]

3.2 Survey of Professional Forecasters

To measure the experts' expectations on headline inflation and GDP growth, we use the Expectations Survey of Private Sector Forecasters collected by BM. This is a monthly survey over several economic variables which is publicly available since 1999. Typically, the survey starts in the middle of the month and ends between three and six days before the end of the month. The exception is December, in which information is collected within the first week of the month. On average, there are 33 respondents per survey. An advantage of using this data rather than other long-running surveys from private institutions is that we can observe the date at which BM receives the answers of the forecasters (more on this below). The survey reports the values for the mean, median, standard deviation, first and third quartile, as well as the lowest and highest values. For our purposes, we consider the median value of the responses in order to better account for the presence of outliers.

For headline inflation, the survey reports expectations for the end of the current year and each of the next three years. Given that BM's projections only cover two years, we use the expectations for the current and next year only. For GDP growth, the survey collects expectations on the annual average growth for the current year and each of the next three years.

[14] See Appendix A for more details. We use the same dictionaries regardless of the identified topic.

[15] For each paragraph we ordered the probability vector θ_d and select topic k if $\theta_d^k \geq 0.25$. Note that it is highly unlikely that four topics are assigned to a given paragraph.

[16] Under this procedure, it might be possible that no paragraphs are assigned to a specific topic and, therefore, the tone cannot be calculated. In these cases, the TI is classified as "neutral".

We only use the expectations for the current and next year to match the periodicity of the CB projections.

In our econometric specification we use forecasts on other variables of interest available in the survey that may also impact experts' expectations on headline inflation and GDP growth. We consider inflation expectations for the current month and GDP growth expectations for the current quarter to construct a measure of "surprises", namely, the difference between the observed and expected data. We also use the expectation series for the current and next year on the nominal exchange rate US dollar/Mexican peso, the 28-day nominal interest rate on government bonds, and the US GDP growth rate.

We use the surveys that are collected prior to and immediately after the publication of the IR to observe how private forecasters may change their expectations in response to *new* information provided by the CB. In a few cases, forecasters surveyed prior to the publication of the report send their answers after the report is published. To avoid the possibility that they may have adjusted their expectations after the publication of the IR, we eliminate the observations received in the day the report is published or thereafter. A similar situation is observed for the survey collected immediately after the publication of the report. Sometimes a few forecasters send their answers before the report is published. Given that we are interested in evaluating how experts may adjust their expectations in response to the publication of the report, we eliminate the answers received before or the day the report is published.[17]

3.3 Other Data Sources

We collect information on monthly headline inflation and quarterly GDP growth rates from the National Institute of Statistics and Geography (INEGI for its acronym in Spanish). As mentioned above, they are used to measure "surprises", i.e. the gap between the data observed

[17] Under these criteria, we identify 17 IR in which there is an overlap between the date of publication of the report and the date forecasters send their answers. On average, we subtract 4 respondents in these 17 reports. Given that the mean number of respondents per survey is 33, in such cases we eliminate 12% of the observations on average.

and the data expected by experts. It is worth mentioning that there are two types of GDP data for a given quarter in Mexico: a timely and an official version. Timely data is published about 30 days after the end of the corresponding quarter while official data is released about 30 days afterwards. The release of timely GDP data started in 2015Q3. Given that our sample starts in either 2008Q1 or 2009Q2, we rely on official GDP data for measuring "surprises".

4 Methodology

We estimate the impact of CB projections on the expectations of professional forecasters in the spirit of Hansen et al. (2019). The authors make a distinction between projections and narrative signals in the Bank of England's IR to seize its impact on market interest rates. In our case, the quantitative elements are provided by BM's projections while the qualitative content is extracted from the IR text through the text mining methods described earlier. In general terms, the identification proceeds as follows. We first estimate a regression model where we only consider the quantitative projections. At this stage, we check for the robustness of our results and select our preferred specification. Next, as a first approach to measure the impact of IR's qualitative information, we use LDA and our TI index to calculate the tone of the paragraphs specifically related to inflation and GDP growth. This tone measure is added to the econometric model selected previously. As an alternative approach, we apply an econometric analysis to the topic-by-topic tone indexes estimated using LDA as described in section 3.1.2 to construct "narrative shocks" using the text of the entire IR. The four most important narrative shocks are added to the selected econometric model with "hard" information only to measure the additional explanatory power of narrative signals.

4.1 Econometric Specification

In this section we describe the econometric model used for the first part of the estimation. Let $F_t(x_h)$ and $F_t(X_h)$ represent the experts' and CB's forecast for variable x at time t for horizon h, respectively. Based on Pedersen (2015) and Baranowski et al. (2021), we consider the following specification for the forecast revisions by experts:

$$\Delta F_{t+1}(x_h) = \beta_0 + \beta_1 [F_t(X_h) - F_t(x_h)] + \phi Z_t + \varepsilon_t, \tag{1}$$

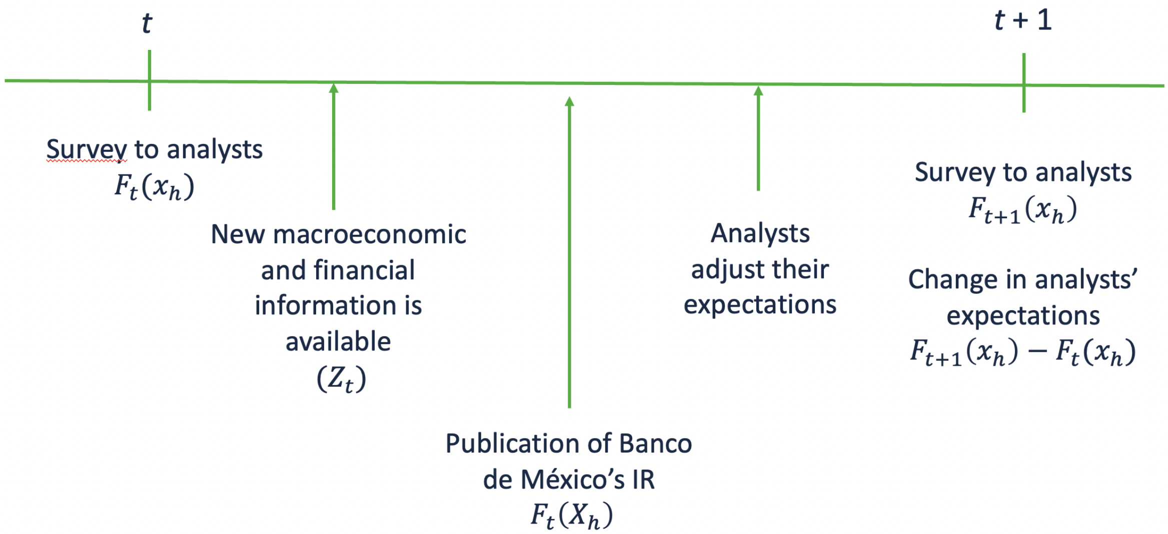
where $\Delta F_{t+1}(x_h) \equiv F_{t+1}(x_h) - F_t(x_h)$ is the change in the experts' forecast for variable x between t and $t+1$. In equation (1), β_0 is a constant term, β_1 is the coefficient that measures how the difference between the CB and experts' forecast may impact next period's expectation for variable x, ϕ is a vector of coefficients, Z_t is a vector of other variables that may affect forecast revisions, and ε_t is the error term. The term x in equation (1) may refer either to headline inflation (π) or GDP growth (y), given our interest in evaluating the impact of CB projections on these two variables. Similarly, horizon h indicates whether variable x is forecast for year T or $T+1$.

Figure 2 illustrates the timing of events behind the estimation of equation (1). At month t forecast $F_t(x_h)$ is gathered from experts. This exercise is repeated at month $t+1$. Between t and $t+1$, forecasters are exposed to new information, including the CB's forecast, $F_t(X_h)$, and public macroeconomic and financial data included in vector Z_t. This new information may lead experts to revise their previous forecast from $F_t(x_h)$ to $F_{t+1}(x_h)$. Accordingly, the survey at $t+1$ would reflect the updated forecast $F_{t+1}(x_h)$.

Expression (1) implies that if the CB's projection is identical to that of experts, the later will not adjust their expectations in the next period due to the publication of the CB's projection at time t. In such a case, any change in the forecast should be attributed to other variables in Z_t. In the data, we observe that $F_t(X_h) \neq F_t(x_h)$ in general. Therefore, we should expect that $\beta_1 \neq 0$ if CB's projections cause experts to revise their expectations.

Variables in vector Z_t depend on the left-hand side term $\Delta F_{t+1}(x_h)$ under consideration. For current-year regressions, we include the difference between the monthly inflation rate at time t and the corresponding experts' forecast, $\pi_t - F_t(\pi_t)$, and the difference between the actual quarterly growth rate at time t and the corresponding experts' forecast, $y_t - F_t(y_t)$, to account

Figure 2: The Timing of Events



Source: Own elaboration.

for the possibility of data "surprises". To incorporate the notion of information rigidities, as in Coibion and Gorodnichenko (2015), a one-period lag in the change of experts' forecast for variable x, $\Delta F_t(x_h)$, is included. We also control for the one-period lag in CB's projections $F_{t-1}(X_h)$. For regressions involving inflation expectations, we follow Pedersen (2015) and include the difference between the experts' inflation forecast for the next two years and the inflation target, $F_t(\pi_{T+2}) - \pi^*$, as a proxy of CB's credibility. Given that BM's annual inflation target is 3%, we set π^*=3. For regressions involving GDP growth expectations, we take into account the change in expectations for US GDP growth $\Delta F_{t+1}(y_h^{US})$. For next-year regressions, we account for the possibility that changes in current-year forecasts for x_T may affect next-year forecasts for x_{T+1}. For all specifications, we include the change in expectations for the nominal exchange rate US dollar/Mexican peso, $\Delta F_{t+1}(e_h)$, and for the 28-day nominal interest rate on government bonds, $\Delta F_{t+1}(i_h)$ as additional controls, both collected from the survey of professional forecasters.

4.2 Tone of Inflationary and GDP Growth Pressures

As a first approach for the identification of narrative signals, we take advantage of the topic distribution obtained from the LDA model (see Section 3.1.2) to extract the tone and the change in the tone of those topics associated with inflation or GDP growth. In a first stage, we assign up to four topics to each paragraph of the IR.[18] Next, we extract all the terms related to inflation (GDP growth) and create a list of all the topics associated with the selected inflation (GDP growth) terms.[19] In a third stage, we select all the IR paragraphs that were assigned a topic belonging to the inflation (GDP growth) topic list. Then, we calculate the TI of these paragraphs and compute the mean tone index for each IR. Finally, we augment model (1) with the mean tone indicator to analyze if BM's qualitative information on inflation (GDP growth) has an impact on professional forecasters' expectations.

4.3 Narrative Shocks

A potential problem with the previous approach is that inflation and/or GDP growth may not be the only topics that provide additional information to market participants to form their inflation and GDP growth expectations. Therefore, as an alternative approach we incorporate *all* the qualitative information from the IR summary using the topic distributions from the LDA model to estimate narrative shocks. Following Hansen et al. (2019), we eliminate the variation in the narrative variables ($\theta_{k,t}^{HL}$ and $\eta_{k,t}^{HL}$) that can be thought of as repeating the information in the numerical forecasts to ensure that narrative shocks are solely "news", i.e. information different to the one already available in the quantitative information provided. To this end, we fit the following models:

$$\theta_{k,t}^{HL} = \alpha_{0,k}^{\theta} + {\alpha_{1,k}^{\theta}}^{T}[F_t(X_h) - F_t(x_h)] + \nu_{k,t}^{\theta}, \tag{2}$$

$$\eta_{k,t}^{HL} = \alpha_{0,k}^{\eta} + {\alpha_{1,k}^{\eta}}^{T}[F_t(X_h) - F_t(x_h)] + \nu_{k,t}^{\eta}. \tag{3}$$

[18]See footnote 15.

[19]We implement a Boolean search to extract all the terms containing the word "inflat" ("growth" or "gdp") and discard all those terms with a per-topic per-word probability less than 0.0015 (0.003) to avoid including irrelevant terms in the list.

Therefore, the residuals $\hat{v}^{\theta}_{k,t}$ and $\hat{v}^{\eta}_{k,t}$ represent the variation in the tone of topic k and its change from the previous IR that is not explained by the economic forecasts published in the IR, i.e. these residuals are the narrative shocks.

To investigate if the IR's qualitative information is statistically significant, we need to determine whether the narrative shocks obtained from equations (2) and (3) explain the residuals from equation (1). Namely, we want to address how much of the variability in experts' forecasts not explained by quantitative information is due to the narrative shocks. We have a total of 54 observations and 52 narrative shocks and, therefore, OLS is not feasible for our setting. Similar to Hansen et al. (2019), we address this problem through regularization by estimating an elastic net regression. Specifically, we solve the following optimization problem:

$$\min_{\gamma} \sum_{t} \left(\hat{\varepsilon}_t - \gamma^T \hat{v}_t\right)^2 + \lambda \left[\alpha ||\gamma||_1 + (1-\alpha)||\gamma||_2^2\right], \tag{4}$$

where $\hat{v}_t = [\hat{v}^{\theta}_{1,t}, \hat{v}^{\theta}_{2,t}, \cdots, \hat{v}^{\theta}_{k,t}, \hat{v}^{\eta}_{1,t}, \hat{v}^{\eta}_{2,t}, \cdots, \hat{v}^{\eta}_{k,t}]$ and $||\gamma||_p = (\sum_{i=1}^{N} |\gamma_i|^p)^{1/p}$ is the standard ℓ^p norm. The first term in equation (4) is the objective function of an OLS regression for the change in the expert's forecast residuals, $\hat{\varepsilon}$, on the narrative shocks. The second expression is a penalty term on regression coefficients γ. This penalty term is a weighted average between a Ridge regression ($\alpha = 0$) and a LASSO regression ($\alpha = 1$). We set $\alpha = 0.99$ to induce a high degree of sparsity (like LASSO) but allowing for some flexibility given the correlations of the narrative shocks.

Prior to the estimation of equation (4), we need to choose the value of the penalty parameter λ. The most common approach is to select the value of λ using cross validation for the out-of-sample predictive performance. Given the sample size, we use a leave-one-out cross validation process and set a grid search for the possible value of λ to vary from 0.00001 to 0.002 with increments of 0.00001.[20] Under this procedure, the percentage of selected narrative shocks for each specific variable x will be an indicator of how important is the qualitative information provided in the IR for explaining changes in experts' forecasts. As reported in

[20]The best model was selected using the RMSE criteria.

section 5.3, we find that more than 50% of the narrative shocks are selected in each case.

4.4 Augmenting the econometric specification with narrative shocks

If we find that narrative signals are an important source of news, we still need to identify which of them contribute the most to explain changes in analysts' expectations and if they differ among the different forecasts under study.

To this end, we implement a bootstrap procedure in the spirit of Hansen et al. (2019) to select the four most important narrative signals in each case. For each of the 500 simulations, we first draw a bootstrap sample with replacement from the original data.[21] Second, using this new sample we re-estimate the model in equation (1) and obtain the new narrative signals by re-estimating equations (2) and (3). Third, using the same cross validation procedure, we evaluate the elastic net regression in (4) and record whether each narrative shock is selected or not. Lastly, we compute the percentage of times that each narrative shock is selected across all the bootstrap draws. This is an indicator of how important is each IR's topic (or its difference) for explaining changes in analysts' expectations. We choose the four most selected topics for each forecast.

Finally, to confirm that the most important narrative shocks actually help to explain some of the variation in experts' forecast revisions, we augment model (1) with the four most selected narrative shocks for each type of forecast and calculate the change in the adjusted R^2. Ideally, we would expect an increase in the value of the adjusted R^2 when adding the IR's qualitative information to the model.

[21]There is not a rule of thumb regarding the appropriate number of simulations in a Bootstrap procedure. Given the computational cost for each simulation, we decided to follow Hansen et al. (2019) and implemented 500 simulations. We believe this number is suitable to give robust results.

5 Results

In this section we report our main results. First, we discuss the impact of CB's projections on experts' forecast revisions. For each of the four sets of forecasts, we select our preferred specification. Next, we show the impact of qualitative information under our first approach, namely, when the tone and the change in the tone of inflationary and GDP growth pressures are added to the model with the quantitative projections. Finally, we estimate the additional impact of the narrative shocks under our second approach for the identification of narrative signals, namely, when the whole text of the IR is accounted for. We also present a robustness analysis of our findings to evaluate if CB projections are perceived differently in a period of lower uncertainty.

5.1 The Influence of Quantitative Information

Table 1 shows our estimates for the impact of projections on the change in analysts' current-year inflation expectations, i.e. $\Delta F_{t+1}(\pi_T)$. The first column of results only considers the difference in inflation expectations between the CB and professional forecasters, $F_t(\Pi_T) - F_t(\pi_T)$, and the surprise in the published data on inflation as explanatory variables. Both are significant at standard levels. In particular, the results suggest that around 25% of the difference in current-year inflation forecasts is translated into changes in inflation expectations, and that inflation surprises generate an upward revision in inflation forecasts. The next column adds the one-period lag in the experts' forecast revision as an explanatory variable. The corresponding coefficient of 0.51 is statistically significant, suggesting the presence of information rigidities (cf. Coibion and Gorodnichenko, 2015).[22] These results are robust to the addition of explanatory variables such as the one-period lag in the CB's inflation projection and the surprise in GDP growth data. The last column presents the estimates when all variables

[22] Capistrán and López-Moctezuma (2014) report that participants in the survey of professional forecasters conducted by BM update their revisions to both inflation and GDP growth smoothly when faced with new information. Using data from the same survey, Capistrán and López-Moctezuma (2010) find that forecasts for short-run inflation and GDP growth do not use information from macroeconomic variables efficiently.

are simultaneously considered.[23] From these exercises we conclude that approximately one-quarter of the difference in inflation expectations between the CB and professional forecasters at time t is translated into changes in current-year inflation expectations at time $t+1$.

The results for the change in next-year inflation expectations are reported in Table 2. The first column of results illustrates that 12.5% of the gap between the CB and experts' forecasts is translated into changes in experts' expectations for next year's inflation. The results also indicate that if experts adjust their current-year inflation forecasts upwards, they are also likely to adjust their next-year expectations in the same direction. The next columns report that the proxy for CB's credibility and changes in expectations for the nominal interest rate and the exchange rate are statistically significant. For each of these regressions, the null that the difference in inflation forecasts cannot impact next-year inflation expectations is rejected at standard significance levels.[24]

Table 3 shows the regressions corresponding to the change in current-year growth expectations, i.e. $\Delta F_{t+1}(y_T)$. Similar to previous cases, the gap in growth forecasts between the CB and experts, $F_t(Y_T) - F_t(y_T)$, is statistically significant under all specifications. On average, 39% of such difference carries to changes in growth forecasts. On the other hand, the surprise in GDP growth leads experts to revise their forecasts upwards. Interestingly, the change in current-year U.S. growth expectations is translated roughly one-to-one to changes in growth forecasts. Finally, the significance of the one-period lag in growth forecasts is not robust, suggesting the absence of information rigidities in this case.[25]

[23] We also estimated the baseline model by including our proxy for CB's credibility and the change in expectations for the nominal exchange rate and for the 28-day nominal interest rate on government bonds as potential explanatory variables. However, none of these variables were statistically significant and thus they are not reported here. This may suggest that short-run surprises are relatively more important for explaining changes in forecasters' expectations. Pedersen (2015) finds a positive but small impact of variations in the exchange rate on changes in current-year inflation expectations while de Mendonça and de Deus (2019) find no impact. Pedersen (2015), de Mendonça and de Deus (2019) and Jain and Sutherland (2020) examine the role of the monetary policy rate (rather than the interest rate on government bonds) on experts' inflation expectations.

[24] We did not find evidence of information rigidities for next-year inflation expectations under alternative specifications.

[25] We included inflation "surprises" and revised expectations for the nominal exchange rate and for the 28-day nominal interest rate on government bonds as additional controls. Nonetheless, they were not statistically significant.

Table 1: Influence of IR's Quantitative Information on Current-Year Inflation Expectations

Regressor	Model 1	Model 2	Model 3	Model 4	Model 5
$F_t(\Pi_T) - F_t(\pi_T)$	0.246**	0.267***	0.244***	0.252***	0.226***
	(0.107)	(0.074)	(0.064)	(0.049)	(0.048)
$\pi_t - F_t(\pi_t)$	1.303***	0.803***	0.771**	0.836***	0.802***
	(0.309)	(0.247)	(0.3)	(0.243)	(0.251)
$y_t - F_t(y_t)$				-0.017**	-0.019**
				(0.007)	(0.009)
$\Delta F_t(\pi_T)$		0.511***	0.504***	0.516***	0.509***
		(0.145)	(0.167)	(0.166)	(0.183)
$F_{t-1}(\Pi_T)$			0.026		0.030*
			(0.021)		(0.016)
Constant	0.049**	0.036**	-0.064	0.034***	-0.080
	(0.022)	(0.017)	(0.077)	(0.009)	(0.063)
R^2	0.364	0.520	0.530	0.540	0.552
Adj. R^2	0.339	0.491	0.492	0.502	0.506
Obs.	54	54	54	54	54

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: *** (**) [*] significant at 1% (5%) [10%] level (two-sided test). The models are estimated with OLS and robust Newey-West (HAC) standard errors. A total of seven models were estimated. We only present the five models with the highest adjusted R^2 values. The rest of the estimations are available upon request. The sample has 54 observations corresponding to the period 2008Q1-2021Q2. Inflation expectations are for the end of year T.

Table 2: Influence of IR's Quantitative Information on Next-Year Inflation Expectations

Regressor	Model 1	Model 2	Model 3	Model 4	Model 5
$F_t(\Pi_{T+1}) - F_t(\pi_{T+1})$	0.125**	0.129**	0.117***	0.098**	0.101***
	(0.059)	(0.058)	(0.041)	(0.038)	(0.035)
$\Delta F_{t+1}(\pi_T)$	0.165***	0.157***	0.112***	0.119**	0.105**
	(0.051)	(0.045)	(0.04)	(0.051)	(0.041)
$\Delta F_{t+1}(e_{T+1})$				0.053*	0.057*
				(0.03)	(0.029)
$\Delta F_{t+1}(i_{T+1})$			0.119*	0.086*	0.092**
			(0.071)	(0.044)	(0.042)
$F_{t+1}(\Pi_{T+2}) - \pi^*$		0.156*	0.186**		0.203**
		(0.090)	(0.088)		(0.100)
Constant	0.063*	-0.019	-0.035	0.052**	-0.055
	(0.033)	(0.046)	(0.041)	(0.024)	(0.048)
R^2	0.247	0.287	0.425	0.439	0.505
Adj. R^2	0.218	0.244	0.378	0.393	0.454
Obs.	54	54	54	54	54

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 1. A total of eight models were estimated. Inflation expectations are for the end of year $T+1$.

Table 3: Influence of IR's Quantitative Information on Current-Year GDP Growth Expectations

Regressor	Model 1	Model 2	Model 3	Model 4	Model 5
$F_t(Y_T) - F_t(y_T)$	0.381***	0.318***	0.442***	0.440***	0.373***
	(0.071)	(0.065)	(0.112)	(0.116)	(0.113)
$y_t - F_t(y_t)$	0.222***	0.251***	0.089**	0.072*	0.125***
	(0.054)	(0.071)	(0.036)	(0.042)	(0.032)
$\Delta F_{t+1}(y_T^{US})$			0.975***	0.979***	0.983***
			(0.153)	(0.159)	(0.173)
$\Delta F_t(y_T)$				0.135*	-0.037
				(0.08)	(0.104)
$F_{t-1}(Y_T)$		0.049***			0.055***
		(0.013)			(0.014)
Constant	-0.149**	-0.241***	-0.128*	-0.117*	-0.234***
	(0.068)	(0.071)	(0.07)	(0.068)	(0.059)
R^2	0.257	0.294	0.744	0.747	0.789
Adj. R^2	0.228	0.251	0.729	0.726	0.767
Obs.	54	54	54	54	54

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 1. A total of eight models were estimated. GDP growth expectations are for year T.

To complete our analysis, Table 4 presents the estimates involving changes in next-year growth forecasts. Although in some of the models 15-20% of the gap in growth forecasts between the CB and experts is incorporated by private analysts, in contrast with previous findings this result is not robust to alternative specifications. The results also indicate that an upward change in current-year growth expectations lead experts to adjust downwards their next-year forecasts, suggesting that analysts consider that short-run deviations of GDP growth from their forecasts are likely to be transitory. In addition, changes in next-year's U.S. growth expectations have some explanatory power to account for changes in domestic growth forecasts but to a lower extent compared to the results in Table 3.[26]

Our findings can be compared to previous results. Hubert (2015) reports that the *level* of CB inflation projections for the current year can impact the *level* of experts' inflation forecast for the same year, and that the corresponding coefficient decreases in magnitude for next-year

[26]Our alternative estimates also included the one-period lag change in next-year growth forecast as an explanatory variable. However, for each specification this variable was not significant at standard levels, implying that information rigidities are not relevant for this case.

Table 4: Influence of IR's Quantitative Information on Next-Year GDP Growth Expectations

Regressor	Model 1	Model 2	Model 3	Model 4	Model 5
$F_t(Y_{T+1}) - F_t(y_{T+1})$	0.209	0.183*	0.173*	0.151	0.145
	(0.144)	(0.101)	(0.1)	(0.094)	(0.109)
$\Delta F_{t+1}(y_T)$	-0.054*	-0.099***	-0.131***	-0.097***	-0.121***
	(0.03)	(0.025)	(0.032)	(0.023)	(0.031)
$\Delta F_{t+1}(e_{T+1})$		-0.179***	-0.225***	-0.192***	-0.227***
		(0.036)	(0.049)	(0.042)	(0.053)
$\Delta F_{t+1}(i_{T+1})$			0.132		0.102
			(0.096)		(0.075)
$\Delta F_{t+1}(y^{US}_{T+1})$				0.479**	0.445**
				(0.209)	(0.205)
Constant	-0.039	-0.031	-0.028	-0.026	-0.024
	(0.036)	(0.03)	(0.03)	(0.028)	(0.025)
R^2	0.124	0.338	0.369	0.420	0.438
Adj. R^2	0.085	0.294	0.311	0.367	0.372
Obs.	49	49	49	49	49

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 1. A total of eight models were estimated. The sample has 48 observations corresponding to the period 2009Q2-2021Q2. GDP growth expectations are for year $T+1$.

inflation forecasts. Clearly, our specification is different to Hubert (2015). However, it is worth noting that the coefficient associated with the CB projection in our case (either for inflation or GDP growth) is also smaller for next-year forecasts. Our results can be compared more directly with those of Pedersen (2015) and de Mendonça and de Deus (2019).[27] For current-year inflation, these authors report values for coefficient β_1 in equation (1) between 0.33 and 0.73, which are higher than our estimates of 0.25. For next-year inflation, our estimates for β_1 are between those reported by de Mendonça and de Deus (2019) and Pedersen (2015). For current-year GDP growth, our findings for β_1 are above those from these authors with the exception of Brazil, where a coefficient value of 0.98 is reported. Finally, de Mendonça and

[27] Pedersen (2015) studies the impact of CB projections on private expectations for Chile and de Mendonça and de Deus (2019) perform a similar exercise for three emerging countries: Brazil, Poland and Mexico. However, there are important distinctions between that paper and ours regarding the estimates for Mexico. A first difference is the period of study (2001Q1-2016Q4 in the former case), in which two different monetary policy instruments were adopted. We limited our sample to cover only the period in which the Overnight Interbank Interest Rate is used as monetary policy instrument. A second difference is that de Mendonça and de Deus (2019) use the quarterly average of the monthly forecasts by experts for variable x_h, which is arguably an imprecise measure for our purposes. We believe taking the monthly forecasts (as we do) as the relevant data is more appropriate for evaluating changes in experts' forecasts. Finally, de Mendonça and de Deus (2019) do not examine the relevance of CB's narrative signals.

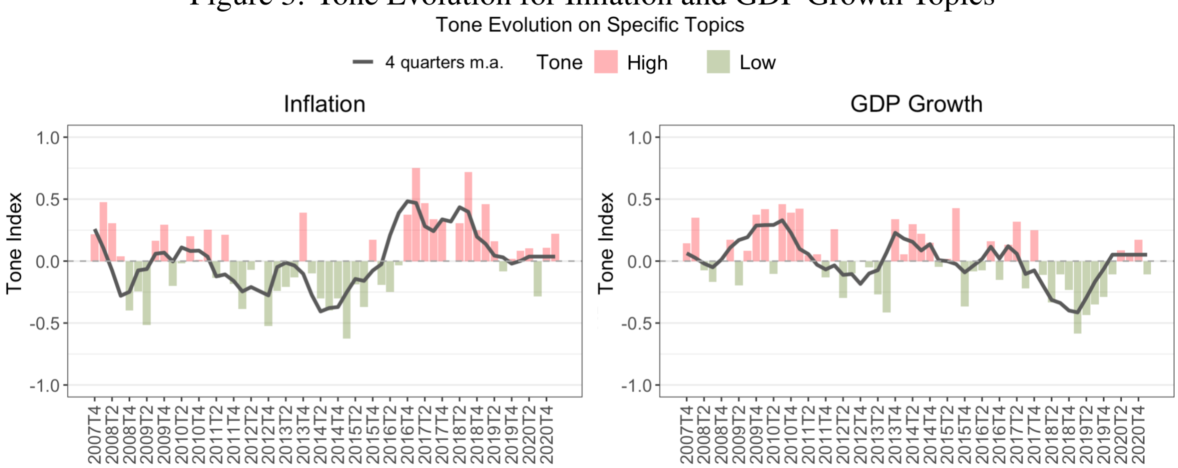
de Deus (2019) obtain estimates between 0.42 and 0.88 for regressions involving next-year GDP growth forecasts, which are well above our findings. Overall, the estimates for β_1 reported in Tables 1 - 4 are roughly in line or somewhat smaller those found elsewhere.

From these exercises, we conclude that CB's projections provide information that experts deem to be valuable for updating their own forecasts, with the exception of those for next-year GDP growth.

5.2 The Influence of Qualitative Information - Tone of Inflationary and GDP Growth Pressures

Figure 3 shows the evolution of the Tone index for those paragraphs exclusively associated with inflation and GDP growth. As observed, in recent years inflation topics have had a higher tone. The opposite is true for the tone of GDP growth.[28] However, the tone of both inflation- and growth-associated paragraphs has been close to a neutral level since the beginning of the pandemic.

Figure 3: Tone Evolution for Inflation and GDP Growth Topics



[28]It is worth noting that the same dictionaries are used for both inflation and GDP growth. In the case of GDP growth, a higher tone is associated with a stronger economic activity, while a lower tone is associated with a weaker economic activity.

Table 5 shows that neither the tone nor the change in the tone of the paragraphs associated with inflation have an explanatory power on the change in forecasters' expectations for current-year inflation. In contrast, Table 6 shows that information on inflation is important when addressing changes in next-year inflation expectations. Approximately 9.2% (14.2%) of the tone (change in tone) expressed by the CB on inflation topics is translated into changes in experts' expectations for next-year's inflation. This result may reflect the relevance of CB projections on inflation for the anchoring of medium-term inflation expectations (but less so for short-run expectations).[29] Furthermore, IR's quantitative information is also significant: again, around 10% of the gap between the CB and professional forecasters is translated into revisions for next-year inflation expectations. Results for the control variables remain quantitatively and qualitatively similar to those reported under Model 5.

For current-year GDP growth expectations, Table 7 shows that the tone of the paragraphs related to GDP growth is not relevant when explaining current-year GDP growth expectations. However, changes in the narrative of this topic provide additional information to forecasters. Lastly, Table 8 shows that the opposite holds for next-year GDP growth expectations: approximately 14.4% of the tone expressed by the CB on GDP growth topics is translated into changes in experts' expectations, while the change in the narrative is not statistically significant.[30] Additionally, it is worth noting that the gap between the CB and experts' forecasts remains statistically non-significant.

These results suggest that CB's narrative signals provided in specific paragraphs of the IR might have some explanatory power on experts' forecasts. In the following section, we take full advantage of the LDA model, which makes the text analysis less dependent on the researcher's

[29]The results shown in Tables 10 and 11 also support this view. In Table 10, topics related to financial markets, aggregate demand and economic activity are important for explaining current-year inflation expectations. In contrast, Table 11 illustrates that narrative shocks associated with inflation are relevant for next-year inflation expectations.

[30]We view the tone and its change as two alternatives for the identification of CB narrative signals. What these results indicate is that forecasters are attentive to CB's discussion on GDP growth. As illustrated by Tables 12 and 13, the alternative approach to account for qualitative information also reports that either levels or changes in narrative shocks related to economic activity may have an influence on GDP growth's expectations.

Table 5: Influence of the Tone of Inflation on Current-Year Inflation Expectations

Regressor	Model 5	Augmented Model with Tone	Augmented Model with Change in Tone
$F_t(\Pi_T) - F_t(\pi_T)$	0.226***	0.229***	0.228***
	(0.048)	(0.035)	(0.042)
$\pi_t - F_t(\pi_t)$	0.802***	0.793***	0.808***
	(0.251)	(0.264)	(0.279)
$y_t - F_t(y_t)$	-0.019**	-0.019**	-0.019**
	(0.009)	(0.008)	(0.007)
$\Delta F_t(\pi_T)$	0.509***	0.508***	0.506**
	(0.183)	(0.189)	(0.192)
$F_{t-1}(\Pi_T)$	0.030*	0.029	0.032**
	(0.016)	(0.019)	(0.015)
Constant	-0.080	-0.077	-0.089
	(0.063)	(0.076)	(0.058)
Tone		0.015	
		(0.09)	
Change in Tone			0.050
			(0.226)
Adj. R^2	0.506	0.496	0.506
ΔAdj. R^2		-0.01	0

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: *** (**) [*] significant at 1% (5%) [10%] level (two-sided test). The models are estimated with OLS and robust Newey-West (HAC) standard errors. A total of seven models were estimated. We only present the model with the highest adjusted R^2 value for the baseline scenario (quantitative information only). This model is augmented with the tone of the paragraphs related to inflation and the change in the tone of those paragraphs. The rest of the estimations are available upon request. The sample has 54 observations corresponding to the period 2008Q1-2021Q2. Changes in expectations are calculated for the end of year T.

point of view, by applying this technique to the complete text in the IR summaries.

5.3 The Influence of Qualitative Information - Narrative Shocks

As described in section 4.3, once we select our preferred model, first we calculate the narrative shocks and estimate an elastic net regression choosing the penalty parameter λ using a leave-one-out cross validation for the out-of-sample predictive performance of the models.[31] Accordingly, Table 9 presents the number of narrative shocks whose coefficient is not drawn to zero; in other words, it shows the number of narrative shocks that are relevant for each

[31] Table B.1 in Appendix B shows the selected λ value for each forecast.

Table 6: Influence of the Tone of Inflation on Next-Year Inflation Expectations

Regressor	Model 5	Augmented Model with Tone	Augmented Model with Change in Tone
$F_t(\Pi_{T+1}) - F_t(\pi_{T+1})$	0.101***	0.117***	0.092**
	(0.035)	(0.034)	(0.036)
$\Delta F_{t+1}(\pi_T)$	0.105**	0.097**	0.117***
	(0.041)	(0.037)	(0.039)
$\Delta F_{t+1}(e_{T+1})$	0.057*	0.060**	0.054**
	(0.029)	(0.026)	(0.026)
$\Delta F_{t+1}(i_{T+1})$	0.092**	0.076**	0.086**
	(0.042)	(0.037)	(0.04)
$F_{t+1}(\Pi_{T+2}) - \pi^*$	0.203**	0.202**	0.211**
	(0.100)	(0.099)	(0.096)
Constant	-0.055	-0.049	-0.062
	(0.048)	(0.049)	(0.049)
Tone		0.092**	
		(0.043)	
Change in Tone			0.142*
			(0.076)
Adj. R^2	0.454	0.496	0.471
Δ Adj. R^2		0.042	0.017

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 5. A total of eight models were estimated. Changes in expectations are calculated for the end of year $T+1$.

forecast. As can be seen, under this procedure more than 67% of the narrative shocks are selected for all forecasts. This is suggestive evidence about the importance of qualitative information provided in IR summaries for explaining changes in analysts' expectations.

Since narrative shocks seem to be an important source of news, we implement a bootstrap procedure to select those that contribute the most to explain changes in analysts' expectations for each forecast.[32] Figure 4 graphically represents the four key topics when examining current-year and next-year inflation forecasts. These word clouds intend to give the reader a quick understanding of how the different topics identified by the model can be easily distinguished. As can be seen in the top panel, analysts mainly focus in the development of financial markets (topic 14), aggregate demand (topic 6), domestic economic activity (topic 24) and changes in

[32]Table B.2 in Appendix B presents the number of bootstrapping iterations (out of a total of 500) in which each narrative shock was selected when estimating the elastic net regression model.

Table 7: Influence of the Tone of GDP Growth on Current-Year GDP Growth Expectations

Regressor	Model 5	Augmented Model with Tone	Augmented Model with Change in Tone
$F_t(Y_T) - F_t(y_T)$	0.373***	0.387***	0.401***
	(0.113)	(0.128)	(0.117)
$y_t - F_t(y_t)$	0.125***	0.135***	0.122***
	(0.032)	(0.034)	(0.032)
$\Delta F_{t+1}(y_T^{US})$	0.983***	0.980***	0.969***
	(0.173)	(0.185)	(0.171)
$\Delta F_t(y_T)$	-0.037	-0.053	-0.076
	(0.104)	(0.121)	(0.11)
$F_{t-1}(Y_T)$	0.055***	0.056***	0.050***
	(0.014)	(0.016)	(0.013)
Constant	-0.234***	-0.236***	-0.230***
	(0.059)	(0.064)	(0.055)
Tone		-0.214	
		(0.188)	
Change in Tone			-0.884*
			(0.495)
Adj. R^2	0.767	0.765	0.772
ΔAdj. R^2		-0.002	0.005

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México*.

Notes: See notes in Table 5. A total of eight models were estimated. This model is augmented with the tone of the paragraphs related to GDP growth and the change in the tone of those paragraphs. Changes in expectations are calculated for year T.

inflation convergence (Δ topic 5) to set their current-year inflation forecasts. The bottom panel shows the key topics for analysts when determining their next-year inflation forecasts. Not surprisingly, topic 3, which is related to CB's inflation projections, appears in both levels and differences. This suggests that the IR's narrative around the CB's projection, as well as how this narrative changes from one IR to another, gives additional information to market analysts. The other two key topics are related to observed inflation (topic 8) and monetary policy (topic 15).

Similarly, Figure 5 presents the four key topics related to current-year and next-year GDP growth forecasts. As can be seen in the top panel, when analysts adjust their current-year GDP growth forecasts, they mainly focus on how the CB expresses risks to inflation (topic 20), the change in the narrative around the investment environment (Δ topic 19), aggregate demand (topic 6), and the CB's projection for formal employment (topic 13). Additionally, the bottom panel presents the key topics for next-year GDP growth forecast. Market analysts pay attention to how the CB's narrative on formal employment projections changed from the

Table 8: Influence of the Tone of GDP Growth on Next-Year GDP Growth Expectations

Regressor	Model 5	Augmented Model with Tone	Augmented Model with Change in Tone
$F_t(Y_{T+1}) - F_t(y_{T+1})$	0.145	0.164	0.150
	(0.109)	(0.109)	(0.091)
$\Delta F_{t+1}(y_T)$	-0.121***	-0.129***	-0.113***
	(0.031)	(0.029)	(0.026)
$\Delta F_{t+1}(e_{T+1})$	-0.227***	-0.230***	-0.227***
	(0.053)	(0.053)	(0.053)
$\Delta F_{t+1}(i_{T+1})$	0.102	0.111	0.097
	(0.075)	(0.073)	(0.07)
$\Delta F_{t+1}(y^{US}_{T+1})$	0.445**	0.454**	0.443**
	(0.205)	(0.205)	(0.189)
Constant	-0.024	-0.026	-0.023
	(0.025)	(0.025)	(0.027)
Tone		0.144*	
		(0.079)	
Change in Tone			0.283
			(0.252)
Adj. R^2	0.372	0.380	0.376
ΔAdj. R^2		0.008	0.004

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México*.

Notes: See notes in Table 5. A total of eight models were estimated. The sample has 48 observations corresponding to the period 2009Q2-2021Q2. This model is augmented with the tone of the paragraphs related to GDP growth and the change in the tone of those paragraphs. Changes in expectations are calculated for year $T+1$.

Table 9: Selected Narrative Shocks for each Forecast

Forecast	# Selected	% Selected
Current-year Inflation	35	67.3
Next-year Inflation	51	98.0
Current-year GDP Growth	43	82.6
Next-year GDP Growth	36	69.2

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México*.

Figure 4: Key Topics for Current-Year and Next-Year Inflation Forecasts

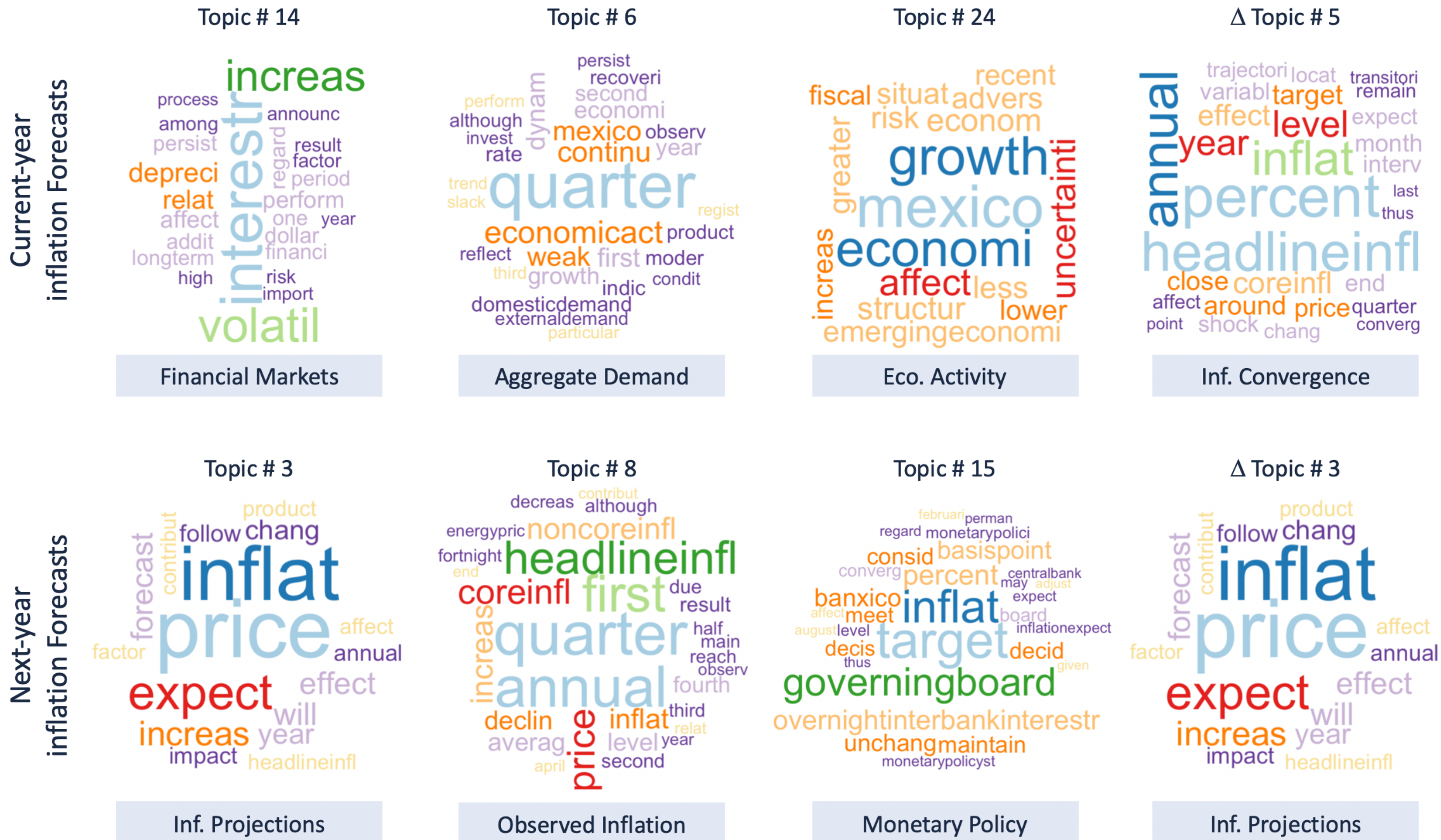
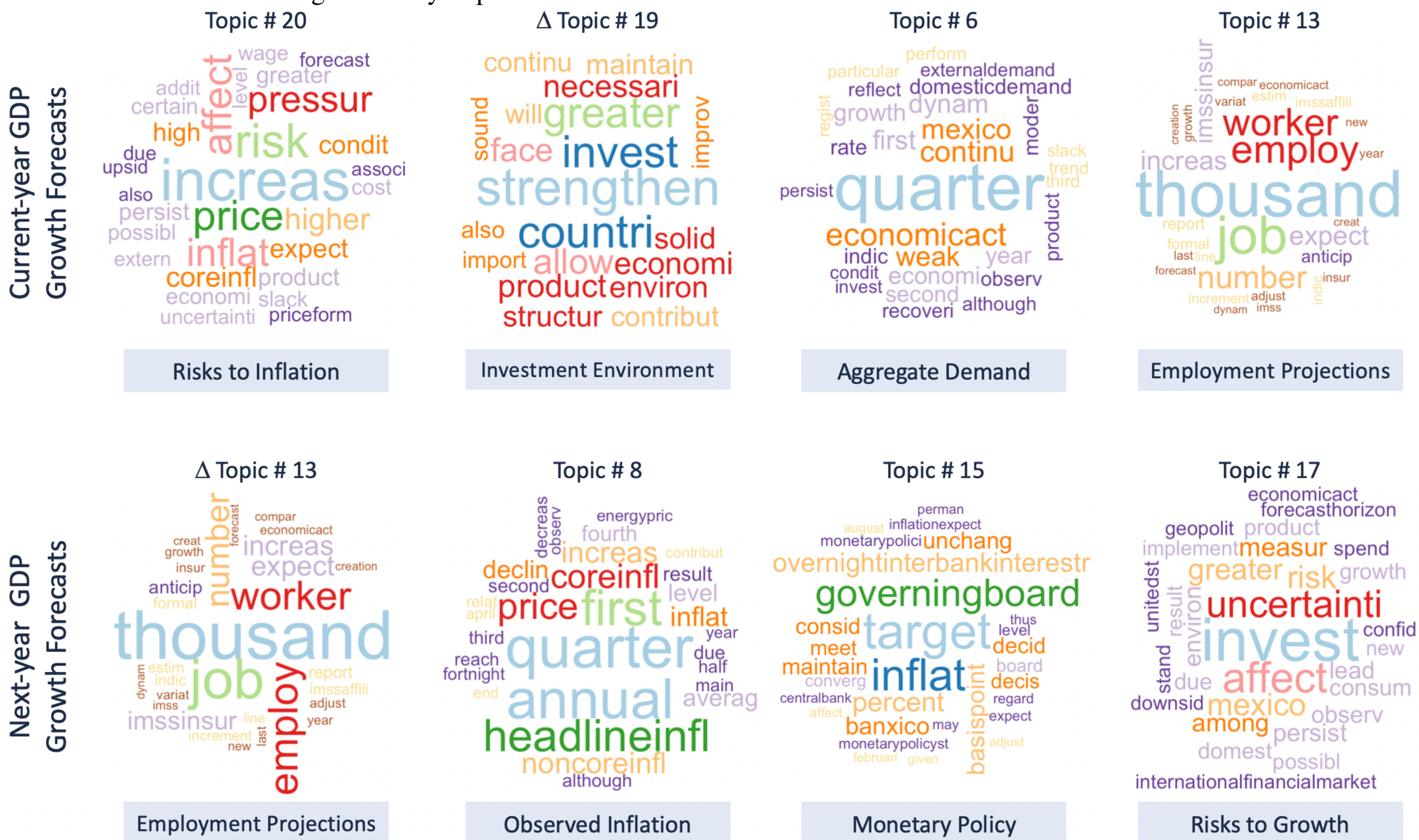


Figure 5: Key Topics for Current-Year and Next-Year GDP Growth Forecasts



previous IR (Δ topic 13), observed inflation (topic 8), monetary policy (topic 15), and risks to growth (topic 17).

We augment model (1) with the top four key topics for each forecast to verify that IR's text gives additional information to professional forecasters when adjusting their current and next-year inflation and GDP growth forecasts.[33] Starting with current-year inflation expectations, Table 10 presents estimates for the baseline model including the most important narrative shocks. Under this specification, around 26% of the difference between the CB and analysts' forecasts is translated into changes in inflation expectations (compared to the 22.6% under the baseline model). However, some of the explanatory variables loose significance, like GDP surprises and the lag in the CB's inflation projection. This change might be due to the addition of the narrative shocks related to economic activity and aggregate demand, as well as those related to the change in the narrative of inflation convergence. If the CB explains in detail the foreseen evolution of the economy and the reasons behind its expectations on inflation, it is likely that the CB's quantitative information provided by GDP and inflation forecasts becomes less informative for professional forecasters. Notably, the variability explained by the model increases by 10.5 percentage points when the IR's qualitative information is taken into account.

Similarly, Table 11 reports the results for the change in next-year inflation expectations when adding qualitative information. The coefficient for the gap between the CB and experts' forecast remains quantitatively similar and significant. As for the explanatory variables, our proxy for CB's credibility is no longer statistically significant as we add the qualitative information related to inflation projections and observed inflation. In line with current-year

[33] We only select the four most important narrative shocks given the size of our sample period and the number of control variables. We also estimate the model adding the two and the six most important narrative shocks and find that our results are robust to this specification. When adding only two narrative shocks for the four types of forecasts: i) the adjusted R^2 value increases in all cases, however to a lesser extent than the baseline model with four narrative shocks (except for current-year GDP growth forecasts); and ii) similar topics are selected in the bootstrapping procedure. When adding six narrative shocks: i) the adjusted R^2 value increases in all cases to a greater extent than the baseline model with four narrative shocks (except for next-year inflation forecasts); and ii) similar topics are selected in the bootstrapping procedure.

Table 10: Influence of IR's Qualitative Information on Current-Year Inflation Expectations

Regressor	Model 5	Augmented Model
$F_t(\Pi_T) - F_t(\pi_T)$	0.226***	0.259***
	(0.048)	(0.054)
$\pi_t - F_t(\pi_t)$	0.802***	0.700**
	(0.251)	(0.293)
$y_t - F_t(y_t)$	-0.019**	-0.014
	(0.009)	(0.009)
$\Delta F_t(\pi_T)$	0.509***	0.374**
	(0.183)	(0.144)
$F_{t-1}(\Pi_T)$	0.030*	0.028
	(0.016)	(0.019)
Constant	-0.080	-0.101
	(0.063)	(0.072)
Top 4 Narrative Shocks		T14: Financial Markets T6: Aggregate Demand T24: Economic Activity ΔT5: Inflation Convergence
Adj. R^2	0.506	0.611
ΔAdj. R^2		0.105

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: *** (**) [*] significant at 1% (5%) [10%] level (two-sided test). The models are estimated with OLS and robust Newey-West (HAC) standard errors. A total of seven models were estimated. We only present the model with the highest adjusted R^2 value for the baseline scenario (quantitative information only) and augment it with the top 4 narrative shocks selected from a bootstrap procedure. The rest of the estimations are available upon request. The sample has 54 observations corresponding to the period 2008Q1-2021Q2. Changes in expectations are calculated for the end of year T.

inflation expectations, the variability explained by the model increases by 12.8 percentage points when adding the top four narrative shocks.

As for current-year GDP growth forecasts, Table 12 shows that IR's qualitative information seems not to improve the baseline estimation. Accordingly, narrative information does not help to increase the variability already explained by the quantitative information provided by the CB and other variables. In contrast, Table 13 shows that narrative shocks are useful for explaining changes on next-year GDP growth expectations because the value of the adjusted R^2 increases by 13.7 percentage points. Notably, the coefficients associated with the quantitative

Table 11: Influence of IR's Qualitative Information on Next-Year Inflation Expectations

Regressor	Model 5	Augmented Model
$F_t(\Pi_{T+1}) - F_t(\pi_{T+1})$	0.101***	0.100***
	(0.035)	(0.034)
$\Delta F_{t+1}(\pi_T)$	0.105**	0.150***
	(0.041)	(0.047)
$\Delta F_{t+1}(e_{T+1})$	0.057*	0.049**
	(0.029)	(0.022)
$\Delta F_{t+1}(i_{T+1})$	0.092**	0.081***
	(0.042)	(0.026)
$F_{t+1}(\Pi_{T+2}) - \pi^*$	0.203**	0.168
	(0.100)	(0.105)
Constant	-0.055	-0.033
	(0.048)	(0.054)
Top 4 Narrative Shocks		T3: Inflation Projections T8: Observed Inflation T15: Monetary Policy ΔT3: Inflation Projections
Adj. R^2	0.454	0.582
Δ Adj. R^2		0.128

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 10. A total of eight models were estimated. Changes in expectations are calculated for the end of year $T+1$.

variables are relatively robust to the inclusion of narrative shocks.

5.4 Results Excluding the Period of the COVID-19 Pandemic

In this section we seek to investigate if CB projections become more relevant for professional forecasters during unusual periods of economic uncertainty. Particularly, we are interested in understanding the role of CB projections during the recent COVID-19 sanitary crisis. Ideally, we would re-estimate our model using 2020Q1 - 2021Q2 as our sample period. However, given that the IR is published on a quarterly basis, we would only have six IR available. Instead, we remove the COVID-19 crisis from the sample, re-estimate the models and compare our results with those in sections 5.1 and 5.3.[34] Therefore, we reduce the sample period to

[34] We also run regressions using only the tone and change in tone of the inflation (GDP growth) paragraphs excluding the COVID-19 period. We do not report these results because they are not as robust as those using all

Table 12: Influence of IR's Qualitative Information on Current-Year GDP Growth Expectations

Regressor	Model 5	Augmented Model
$F_t(Y_T) - F_t(y_T)$	0.373***	0.381**
	(0.113)	(0.160)
$y_t - F_t(y_t)$	0.125***	0.145***
	(0.032)	(0.042)
$\Delta F_{t+1}(y_T^{US})$	0.983***	0.944***
	(0.173)	(0.218)
$\Delta F_t(y_T)$	-0.037	-0.070
	(0.104)	(0.147)
$F_{t-1}(Y_T)$	0.055***	0.057***
	(0.014)	(0.019)
Constant	-0.234***	-0.295***
	(0.059)	(0.099)
Top 4 Narrative Shocks		T20: Risks to Inflation ΔT19: Investment Environment T6: Aggregate Demand T13: Employment Projections
Adj. R^2	0.767	0.767
ΔAdj. R^2		0.000

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 10. A total of eight models were estimated. Changes in expectations are calculated for year T.

2008Q1-2019Q3.

Even though we are only excluding seven observations, GDP growth volatility is strikingly different in both samples: the standard deviation of GDP growth under the full sample is 1.8 times larger than the standard deviation under the sample without COVID-19. Perhaps not surprisingly, the economic uncertainty brought by the pandemic was reflected in a higher volatility for current-year GDP growth forecasts. Specifically, its standard deviation in the sample without COVID-19 increases by almost 40% when the full sample is considered.[35] Therefore, professional forecasters were indeed subject to a higher short-run uncertainty for GDP growth during the pandemic. By comparing the previous results with those excluding

topics. These results are available upon request.

[35] Noticeably, the standard deviation for next-year GDP growth and inflation forecasts are similar under both samples. Also, the standard deviation of headline inflation is similar in both samples.

Table 13: Influence of IR's Qualitative Information on Next-Year GDP Growth Expectations

Regressor	Model 5	Augmented Model
$F_t(Y_{T+1}) - F_t(y_{T+1})$	0.145	0.150
	(0.109)	(0.104)
$\Delta F_{t+1}(y_T)$	-0.121***	-0.100***
	(0.031)	(0.020)
$\Delta F_{t+1}(e_{T+1})$	-0.227***	-0.204***
	(0.053)	(0.032)
$\Delta F_{t+1}(i_{T+1})$	0.102	0.056
	(0.075)	(0.053)
$\Delta F_{t+1}(y^{US}_{T+1})$	0.445**	0.379**
	(0.205)	(0.179)
Constant	-0.024	-0.029
	(0.025)	(0.021)
Top 4 Narrative Shocks		ΔT13: Employment Projections T8: Observed Inflation T15: Monetary Policy T17: Risks to Growth
Adj. R^2	0.372	0.509
ΔAdj. R^2		0.137

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México*.

Notes: See notes in Table 10. A total of eight models were estimated. The sample has 48 observations corresponding to the period 2009Q2-2021Q2. Changes in expectations are calculated for year $T+1$.

the COVID-19 period, we can evaluate if CB projections are relatively more important under higher uncertainty.

Again, we follow the estimation steps detailed in Section 4. Accordingly, Tables C.1 - C.4 in Appendix C present the results for each of the variables of interest when only projections are included. Figure C.1 in Appendix C also shows the new topic distribution for the sample ending in 2019Q3. It is worth noting that, since the sample period changes, the IR corpus used to estimate the LDA model is different. Therefore, even though we estimate the same number of topics, their specification may vary. Moreover, some topics might be different from those in the previous section. For example, Topic 2 is now related to formal employment forecasts, while previously this topic was identified as number 13. Similarly, for the entire sample Topic 11 was associated with the COVID-19 crisis while, naturally, there is no such topic for the reduced period. As before, for each of the four forecasts we select our preferred specification, which corresponds to Model 5 in all cases except for next-year GDP growth forecasts, for which Model 4 is selected. We also choose the four most important narrative shocks under each specification as reported in Appendix C.

Table 14 presents estimates for the change in current-year inflation expectations for the sample excluding the COVID-19 period. Under this specification, the difference between the CB and analysts' forecasts that is translated into changes in inflation expectations decreases from 26% for the entire sample (see Table 10) to almost 22% for the reduced sample. This result suggests that forecasters might be slightly more sensitive to the CB's inflation projections in an environment of larger volatility in GDP growth. Additionally, the sign and significance of the other explanatory variables remains the same. As before, augmenting the model with qualitative information increases the value of the adjusted R^2 by a similar magnitude, suggesting that qualitative information is equally important in both periods. When comparing the most important narrative shocks with those under the full sample, we note that topics related to financial markets, aggregate demand and inflation convergence "loose" importance, while those associated with formal employment projections, risks to the economic activity and risk to inflation are the most informative for professional forecasters.

Table 15 reports the results for the change in next-year inflation expectations for the reduced sample. Again, 10% of the gap between the CB and professional forecasters is translated into revisions of next-year inflation expectations. Results for the additional explanatory variables are quantitatively and statistically similar, except for the nominal interest rate on government bonds, which is no longer significant (see also Table 11). Again, augmenting the model with qualitative information increases the value of the adjusted R^2 by a similar magnitude. As for the most important narrative shocks, the most informative topics for professional forecasters are now related to the monetary policy decision, risks to inflation, long-run growth perspectives and growth projections. After comparing the shocks in Table 11 with those in Table 15, we observe that the CB's narrative about inflation projections were particularly

Table 14: Influence of IR's Qualitative Information on Current-Year Inflation Expectations - Sample Without COVID Crisis

Regressor	Model 5	Augmented Model
$F_t(\Pi_T) - F_t(\pi_T)$	0.204**	0.217**
	(0.091)	(0.088)
$\pi_t - F_t(\pi_t)$	0.774**	0.725*
	(0.327)	(0.374)
$y_t - F_t(y_t)$	-0.018	-0.011
	(0.014)	(0.012)
$\Delta F_t(\pi_T)$	0.354**	0.401**
	(0.143)	(0.155)
$F_{t-1}(\Pi_T)$	0.041	0.043
	(0.027)	(0.027)
Constant	-0.129	-0.137
	(0.11)	(0.111)
Top 4 Narrative Shocks		ΔT02: Employment Projections T22: Risks to Eco. Act. ΔT13: Risks to Inflation T15: Economic Activity
Adj. R^2	0.482	0.564
Δ Adj. R^2.		0.082

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México*.

Notes: See notes in Table 10. The sample has 47 observations corresponding to the period 2008Q1-2019Q3.

relevant for professional forecasters during the period covering the COVID-19 pandemic. If this period is excluded, forecasters seem to be more attentive to narratives related to inflation risks.

As for current-year GDP growth forecasts, Table 16 shows that the difference between the CB and analysts' forecasts that is translated into changes in growth expectations almost doubles compared to the results for the entire sample from 38 to 66%. Simultaneously, the change in expectations for US GDP growth is no longer significant. Given that the restricted sample exhibits less uncertainty for GDP growth, the increase from 38 to 66% indicates that forecasters are more sensitive to CB projections when uncertainty is relatively low. But if uncertainty increases, the importance of CB projections diminishes at the expense of expectations for US GDP growth. Accordingly, CB's projections seem to be less influential if the uncertainty

Table 15: Influence of IR's Qualitative Information on Next-Year Inflation Expectations - Sample Without COVID Crisis

Regressor	Model 5	Augmented Model
$F_t(\Pi_{T+1}) - F_t(\pi_{T+1})$	0.093**	0.104**
	(0.038)	(0.039)
$\Delta F_{t+1}(\pi_T)$	0.127**	0.147**
	(0.059)	(0.057)
$\Delta F_{t+1}(e_{T+1})$	0.072*	0.081***
	(0.039)	(0.026)
$\Delta F_{t+1}(i_{T+1})$	0.080*	0.034
	(0.045)	(0.05)
$F_{t+1}(\Pi_{T+2}) - \pi^*$	0.205*	0.122
	(0.111)	(0.091)
Constant	-0.061	-0.009
	(0.053)	(0.045)
Top 4 Narrative Shocks		ΔT20: Monetary Policy Decision T13: Risks to Inflation ΔT12: Long-run Growth Perspectives T17: Growth Projections
Adj. R^2	0.451	0.582
ΔAdj. R^2		0.131

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 11. The sample has 47 observations corresponding to the period 2008Q1-2019Q3.

surrounding GDP growth is higher. Moreover, the most relevant topics for professional forecasters in the reduced sample are related to forward guidance, observed inflation, and financial markets. Presumably, this result may be driven by the 2008-2009 financial crisis. In contrast with the full sample, qualitative information now has additional explanatory power in accounting for changes in analysts' expectations, suggesting that it becomes more relevant when uncertainty is low.

Lastly, Table 17 shows the estimated results for next-year GDP growth forecasts. In sharp contrast with the findings for the complete sample period, now 26% of the difference between CB and analysts' forecast account for changes in expectations. Similar to the case of current-year forecasts, this result seems to reflect the higher impact of CB's projections when economic uncertainty is relatively low. Formal employment projections and monetary policy are still

Table 16: Influence of IR's Qualitative Information on Current-Year GDP Growth Expectations - Sample Without COVID Crisis

Regressor	Model 5	Augmented Model
$F_t(Y_T) - F_t(y_T)$	0.524***	0.658***
	(0.154)	(0.152)
$y_t - F_t(y_t)$	0.122***	0.100***
	(0.038)	(0.031)
$\Delta F_{t+1}(y_T^{US})$	0.359**	0.195
	(0.176)	(0.145)
$\Delta F_t(y_T)$	0.087	-0.013
	(0.099)	(0.116)
$F_{t-1}(Y_T)$	0.031*	0.035**
	(0.018)	(0.016)
Constant	-0.141**	-0.139**
	(0.059)	(0.055)
Top 4 Narrative Shocks		T23: Forward guidance ΔT25: Risks to Financial Markets ΔT07: Observed Inflation T10: Financial Markets
Adj. R^2	0.535	0.652
Δ Adj. R^2		0.117

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 12. The sample has 47 observations corresponding to the period 2008Q1-2019Q3.

important topics, while observed inflation and risks to growth are replaced by investment and risks to inflation.

From these exercises, we conclude that the way in which experts process both quantitative and qualitative information from the CB about inflation and GDP growth is different and may be sensitive to the economic environment. For inflation, the share of the gap between CB's projections and experts' forecasts that is incorporated by experts into their forecast revision is robust across samples. Furthermore, narrative signals are important for changing forecasters' expectations in both periods. These results might be due to the fact that the standard deviation of inflation and inflation forecasts for current- and next-year are very similar in both periods, suggesting that there were no changes in the inflationary environment. In contrast, the relative importance of the differences between the CB and forecasters in accounting for changes in

Table 17: Influence of IR's Qualitative Information on Next-Year GDP Growth Expectations - Sample Without COVID Crisis

Regressor	Model 4	Augmented Model
$F_t(Y_{T+1}) - F_t(y_{T+1})$	0.228***	0.263***
	(0.082)	(0.070)
$\Delta F_{t+1}(y_T)$	-0.079	-0.091
	(0.084)	(0.073)
$\Delta F_{t+1}(e_{T+1})$	-0.159***	-0.154***
	(0.023)	(0.017)
$\Delta F_{t+1}(y^{US}_{T+1})$	0.444***	0.308**
	(0.157)	(0.145)
Constante	-0.057***	-0.046***
	(0.017)	(0.015)
Top 4 Narrative Shocks		T08: Investment T20: Monetary Policy Decision ΔT02: Employment Projections T01: Risks to Inflation
Adj. R^2	0.533	0.611
Δ Adj. R^2		0.078

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México*.

Notes: See notes in Table 13. The sample has 42 observations corresponding to the period 2009Q2-2019Q3.

GDP growth expectations varies sharply in both periods. Our results suggest that forecasters are less sensitive to CB's projections if the uncertainty about GDP growth among forecasters is higher. Moreover, we find that narrative signals do not have an influence on current-year GDP growth expectations if uncertainty is high.

It is worth noting that there may be other reasons why some of these projections may have a stronger influence on expectations than others. This might be the result of many factors, such as i) the differences in the timeliness of information regarding inflation (bimonthly) vs GDP (quarterly with an eight-week lag and with monthly partial updates), ii) the size and persistence of shocks to each variable (inflationary shocks may have a stronger persistence than those affecting GDP growth, but overall GDP growth is much more volatile than inflation), iii) the specific degree of economic uncertainty at any particular time, and iv) the analysts' perception that a particular CB projection may not only have a purely informative content but also an expectations' management element. Indeed, in theory, there is a role for publicly-funded institutions such as CB to function as coordinators of the public's expectations.

All the above considerations are likely to have different relevance when comparing inflation vs GDP projections, and also when comparing projections at different lags. Therefore, the analysis of the circumstances —such as high inflation vs. low inflation, underwhelming growth vs. higher-than-expected growth, the degree of recent forecasting success, among others— that lead analysts to incorporate more of the CB's projections into their own, remains an important question to answer in future work.

6 Conclusions

In this paper we have evaluated the influence of CB's quantitative and qualitative information, as laid out in BM's IR summaries, on the expectations of professional forecasters. Our findings suggest that the information provided by CB's projections is embodied by experts in their forecasts on inflation and GDP growth. For inflation, we report that roughly 25% of the difference between CB and private forecasts is translated into changes for current-year experts' forecasts, and that only 10% of these differences are incorporated by experts in their next-year forecasts. For GDP growth, experts assimilate about 39% of the difference between CB and private forecasts for the current year, but the evidence is not robust for next-year forecasts. As previously mentioned, these estimates are similar or smaller than those reported elsewhere in the literature.

To measure the importance of narrative signals, we have proposed two alternative approaches that take advantage of the LDA model. Under the first approach, we report that the tone used by the CB when addressing inflation in the IR summary and the change in such tone in the next IR summary are important in accounting for variations in experts' forecasts for next-year inflation but not for current-year inflation. Also, we observe that the tone used to identify economic activity pressures and its change may have an influence on experts' GDP growth forecasts.

When the full text of the IR summary is taken into account under the second approach, we find that narrative signals may have an influence on forecasters' expectations for current- and next-year inflation, and next-year GDP growth. We also identify that narrative topics related to monetary policy, inflation data, aggregate demand, and inflation and formal employment projections are the most relevant to experts. The findings from these two approaches suggest that CB's qualitative information may be relevant for experts' forecasts, as reported by Ullrich (2008), López Marmolejo (2013), Montes et al. (2016), Hubert (2017), Anderes et al. (2021) and Baranowski et al. (2021). As explained before, what distinguishes our paper from others is the use of the LDA model for the identification of narrative signals. In this sense, the paper confirms the possibility of using automatic text classification methods to examine whether such text has some information content for forecasters. Finally, we conduct a similar exercise for a sample excluding the COVID-19 period to evaluate if our results are robust to the economic environment. We find that forecasters systematically account for both types of information when revising their inflation expectations. However, we report that the relative importance of quantitative and qualitative information for GDP growth expectations may be sensitive to the short-run output uncertainty faced by forecasters.

There are various directions for future research using the text mining techniques described in this paper. First, it would be interesting to estimate the influence of narrative signals on financial markets. This may be particularly relevant as non-conventional monetary policy tools such as forward guidance have gained prominence in recent years. In this regard, it would be useful to analyze which elements of the text are the most relevant in accounting for movements in the yield curve at different maturities. Another possibility is to evaluate if CB's qualitative information may have an influence on the dispersion of forecasters' expectations. More generally, analyzing other characteristics of CB's narrative signals, such as their readability, their evolution over time, and their impact during periods of crisis, are all interesting avenues to pursue in the future.

Bibliography

Algaba, A., Ardia, D., Bluteau, K., Borms, S., and Boudt, K. (2020). Econometrics meets sentiment: An overview of methodology and applications. *Journal of Economic Surveys*, 34(3):512–547.

Anderes, M., Rathke, A., Streicher, S., and Sturm, J. E. (2021). The role of ECB communication in guiding markets. *Public Choice*, 186(3):351–383.

Baranowski, P., Doryń, W., Łyziak, T., and Stanisławska, E. (2021). Words and deeds in managing expectations: Empirical evidence from an inflation targeting economy. *Economic Modelling*, 95:49–67.

Barro, R. J. (1976). Rational expectations and the role of monetary policy. *Journal of Monetary Economics*, 2(1):1–32.

Bholat, D., Hansen, S., Santos, P., and Schonhardt-Bailey, C. (2015). Text mining for central banks. *London: Bank of England, Centre for Central Banking Studies*.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet allocation. *The Journal of Machine Learning Research*, 3:993–1022.

Blinder, A. S., Ehrmann, M., Fratzscher, M., De Haan, J., and Jansen, D. J. (2008). Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature*, 46(4):910–45.

Capistrán, C. and López-Moctezuma, G. (2010). Macroeconomic expectations of professional forecasters: An evaluation of short-run forecasts in Mexico (in Spanish). *El Trimestre Económico*, 77(306):275–312.

Capistrán, C. and López-Moctezuma, G. (2014). Forecast revisions of Mexican inflation and GDP growth. *International Journal of Forecasting*, 30(2):177–191.

Coenen, G., Ehrmann, M., Gaballo, G., Hoffmann, P., Nakov, A., Nardelli, S., Persson, E., and Strasser, G. (2017). Communication of monetary policy in unconventional times. ECB Working Paper no. 2080.

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

Cukierman, A. (2001). Accountability, credibility, transparency and stabilization policy in the eurosystem. In *The Impact of EMU on Europe and the Developing Countries*, pages 40–75. Oxford University Press.

de Mendonça, H. F. and de Deus, J. D. B. V. (2019). Central bank forecasts and private expectations: An empirical assessment from three emerging economies. *Economic Modelling*, 83:234–244.

Ehrmann, M. and Fratzscher, M. (2009). Explaining monetary policy in press conferences. *International Journal of Central Banking*, 5(2):42–84.

El-Shagi, M. and Jung, A. (2015). Have minutes helped markets to predict the MPC's monetary policy decisions? *European Journal of Political Economy*, 39:222–234.

Fracasso, A., Genberg, H., and Wyplosz, C. (2003). How do central banks write? *Geneva Reports on the World Economy*, 2.

Fujiwara, I. (2005). Is the central bank's publication of economic forecasts influential? *Economics Letters*, 89(3):255–261.

Gamber, E. N. and Smith, J. K. (2009). Are the Fed's inflation forecasts still superior to the private sector's? *Journal of Macroeconomics*, 31(2):240–251.

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

Gersbach, H. (2003). On the negative social value of central banks' knowledge transparency. *Economics of Governance*, 4(2):91–102.

Hansen, S. and McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99:S114–S133.

Hansen, S., McMahon, M., and Prat, A. (2014). Transparency and deliberation within the FOMC: A computational linguistics approach. *CEPR Discussion Paper 9994*.

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870.

Hansen, S., McMahon, M., and Tong, M. (2019). The long-run information effect of central bank communication. *Journal of Monetary Economics*, 108:185–202.

Hattori, M., Kong, S., Packer, F., and Sekine, T. (2021). The impact of regime change on the influence of the central bank's inflation forecasts: Evidence from Japan's shift to inflation targeting. *International Journal of Central Banking*, 17(4):257–290.

Hayo, B., Kutan, A. M., and Neuenkirch, M. (2012). Communication matters: US monetary policy and commodity price volatility. *Economics Letters*, 117(1):247–249.

Hayo, B. and Neuenkirch, M. (2010). Do Federal Reserve communications help predict federal funds target rate decisions? *Journal of Macroeconomics*, 32(4):1014–1024.

Hubert, P. (2014). FOMC forecasts as a focal point for private expectations. *Journal of Money, Credit and Banking*, 46(7):1381–1420.

Hubert, P. (2015). Do central bank forecasts influence private agents? Forecasting performance versus signals. *Journal of Money, Credit and Banking*, 47(4):771–789.

Hubert, P. (2017). Qualitative and quantitative central bank communication and inflation expectations. *The BE Journal of Macroeconomics*, 17(1):1–41.

Jain, M. and Sutherland, C. S. (2020). How do central bank projections and forward guidance influence private-sector forecasts? *International Journal of Central Banking*, 16(5):179–218.

Jung, A. and Kuehl, P. (2021). Can central bank communication help to stabilise inflation expectations? *Scottish Journal of Political Economy*, 68(3):298–321.

López Marmolejo, A. (2013). Deciphering Banco de Mexico's language (in Spanish). *El Trimestre Económico*, 80(318):345–370.

Lustenberger, T. and Rossi, E. (2020). Does central bank transparency and communication affect financial and macroeconomic forecasts? *International Journal of Central Banking*, 16(2):153–201.

Montes, G., Oliveira, L., Curi, A., and Nicolay, R. (2016). Effects of transparency, monetary policy signalling and clarity of central bank communication on disagreement about inflation expectations. *Applied Economics*, 48(7):590–607.

Morris, S. and Shin, H. S. (2002). Social value of public information. *American Economic Review*, 92(5):1521–1534.

Nakamura, E. and Steinsson, J. (2018). High-frequency identification of monetary non-neutrality: The information effect. *The Quarterly Journal of Economics*, 133(3):1283–1330.

Naszodi, A., Csavas, C., and Felcser, D. (2016). Which aspects of central bank transparency matter? A comprehensive analysis of the effect of transparency on survey forecasts. *International Journal of Central Banking*, 12(4):147–192.

Österholm, P., Dale, S., and Orphanides, A. (2008). Imperfect central bank communication-information versus distraction. *IMF Working Papers*, pages 1–31.

Pedersen, M. (2015). What affects the predictions of private forecasters? The role of central bank forecasts in Chile. *International Journal of Forecasting*, 31(4):1043–1055.

Romer, C. D. and Romer, D. H. (2000). Federal Reserve information and the behavior of interest rates. *American Economic Review*, 90(3):429–457.

Svensson, L. E. (2010). Inflation targeting. In *Handbook of Monetary Economics*, volume 3, pages 1237–1302. Elsevier.

Tarkka, J. and Mayes, D. G. (1999). The value of publishing official central bank forecasts. *Bank of Finland Research Discussion Paper 22*.

Tobback, E., Nardelli, S., and Martens, D. (2017). Between hawks and doves: Measuring central bank communication. ECB Working Paper no. 2085.

Ullrich, K. (2008). Inflation expectations of experts and ECB communication. *The North American Journal of Economics and Finance*, 19(1):93–108.

Woodford, M. (2005). Central bank communication and policy effectiveness. Technical report, National Bureau of Economic Research.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (statistical methodology)*, 67(2):301–320.

A Tone Index

For the purpose of this paper, the tone of a message signals either the severity of inflationary pressures or the strength of economic dynamism. Specifically, a higher tone value is associated with higher inflationary pressures or stronger economic dynamism. Therefore, the words chosen for specific topics allow us to measure the tone of the overall CB message.

We follow the approach of Tobback et al. (2017) and use semantic orientation to propose a Tone Index (TI) using dictionaries of our own. Our dictionary is the result of both monetary policy knowledge and understanding of the relevant vocabulary used in the monetary policy documents of advanced and emerging CB. It applies to any text related to monetary policy and has been reviewed by several monetary policy experts to avoid excluding important terms.

To compute the TI, we count the number of occurrences of words in each document associated with higher and lower inflationary pressures, or stronger and weaker economic dynamism based on our dictionaries (see Table A.1). First we split each document into sentences and then each sentence is tokenized into words. After this, we count the number of words in each sentence that match with our dictionaries that are not preceded by a negation word. When a word is followed or preceded by a negation word, we reverse the sentiment of that word. Finally, the TI is computed as

$$TI = \frac{I_H - I_L}{I_H + I_L} \tag{5}$$

where $I_H = \sum_{i=1}^{n} H_i$ and $I_L = \sum_{i=1}^{n} L_i$ are, respectively, the sum of occurrences of words associated with higher and lower inflationary pressures, or stronger and weaker economic dynamism over the n sentences of the document. If a text only has I_H words, then $TI = I_H/I_H = 1$, and if a text only includes I_L words, then $TI = -I_L/I_L = -1$. When a text has the same number of occurrences of I_H and I_L words, we have $TI = (I_H - I_L)/(I_H + I_L) = 0$ and therefore, we say that the document has a neutral tone. Note that when neither I_H nor I_L words are found, the text is considered as being neutral and therefore $TI = 0$.

Table A.1 presents the main words in the dictionaries used to construct the Tone Index. All terms derived from these words (like plurals and conjunctions) are also taken into account when calculating the Tone Index for each document.

Table A.1: Pressure Index Dictionaries

High pressure	**Low pressure**	**Negation**
augment	abate	not
boost	accommodate	not expected to
bump up	contain	unlikely to
climb	cut	fail to
elevate	damp	no reason to
expand	decelerate	
go up	decline	
hawkish	decrease	
head up	depress	
high	deteriorate	
hike	diminish	
improve	disappoint	
increase	dovish	
lift	down	
move up	downward	
pick up	drop	
put up	ease	
raise	go down	
rebound	head down	
rise	loose	
solid	low	
strong	moderated	
strength	move down	
tight	put down	
upward	reduce	
	shave	
	slash	
	slice	
	slow	
	subdue	
	underutil	
	weak	

Notes:
1/ This table presents the main words in the dictionaries used to construct the Tone Index.
2/ All terms derived from these words are also included in the dictionaries.
3/ The selection of these words is based on Tobback et al. (2017) dictionaries which have been strengthened by the authors of this paper. They are the result of both monetary policy knowledge and understanding of the relevant vocabulary in central banks' monetary policy documents.

B Bootstrapping

Table B.1: Selection of λ Values

Forecast	λ
Current-year Inflation	0.002
Next-year Inflation	0.00001
Current-year GDP Growth	0.002
Next-year GDP Growth	0.002

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México*.

Notes: This table reports the values of λ in equation (4) for each of the four dependent variables considered in the study. The values are selected through a leave-one-out cross validation process using a grid search where $\lambda \in \{0.00001, 0.00002, \cdots, 0.002\}$. We choose the λ value for the model with the smallest RMSE.

Table B.2: Bootstrapping: Topic Selection

Narrative Shock	Current-year Inflation	Next-year Inflation	Current-year GDP Growth	Next-year GDP Growth	Narrative Shock	Current-year Inflation	Next-year Inflation	Current-year GDP Growth	Next-year GDP Growth
T.1	497	496	497	4	Δ T.1	6	6	498	5
T.2	497	497	1	496	Δ T.2	498	5	498	5
T.3	498	499	5	496	Δ T.3	497	499	499	4
T.4	498	4	497	5	Δ T.4	4	498	497	4
T.5	1	498	1	3	Δ T.5	499	498	497	495
T.6	499	6	499	3	Δ T.6	497	497	498	496
T.7	498	5	497	493	Δ T.7	6	497	3	498
T.8	497	499	498	499	Δ T.8	499	495	498	497
T.9	498	6	498	2	Δ T.9	496	496	4	4
T.10	496	498	495	5	Δ T.10	4	498	498	497
T.11	6	1	497	5	Δ T.11	495	5	495	2
T.12	497	2	4	3	Δ T.12	499	496	496	496
T.13	497	5	499	1	Δ T.13	6	3	499	500
T.14	500	498	498	494	Δ T.14	499	499	498	497
T.15	5	499	5	499	Δ T.15	7	499	496	497
T.16	496	497	497	495	Δ T.16	5	4	499	497
T.17	498	498	497	498	Δ T.17	499	498	5	4
T.18	4	6	497	498	Δ T.18	497	4	497	4
T.19	497	4	498	3	Δ T.19	4	499	500	6
T.20	4	496	500	5	Δ T.20	4	7	496	495
T.21	497	4	6	498	Δ T.21	497	498	497	4
T.22	498	497	496	498	Δ T.22	498	498	3	496
T.23	498	498	498	1	Δ T.23	6	5	4	5
T.24	499	495	498	495	Δ T.24	497	498	496	497
T.25	497	497	4	498	Δ T.25	4	496	498	496
T.26	7	6	7	4	Δ T.26	499	499	6	498

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de Mexico*.

Notes: This table shows the number of times each topic is selected for each of the four dependent variables under the bootstrapping procedure. For more details, see Section 4.4.

C Results Without the COVID-19 Pandemic

This Appendix presents additional results for the sample excluding the period of the COVID-19 pandemic. The OLS regressions that include only projections are reported in Tables C.1 - C.4. In general, the sign and magnitude of the coefficients are similar to those obtained under the full sample for the regressions related to current- and next-year inflation. For regressions on current-year growth (Table C.3), the coefficients for the difference between CB and experts' forecasts are significantly larger than the corresponding estimates under the full sample. In contrast, the coefficients for the change in US GDP growth are now significantly smaller. For next-year growth regressions, the coefficients for the difference between CB and experts' forecasts are significant under all specifications, while the change in the forecast for current-year growth is no longer significant.

Table C.1: Influence of IR's Quantitative Information on Current-Year Inflation Expectations - Sample Without COVID Crisis

Regressor	Model 1	Model 2	Model 3	Model 4	Model 5
$F_t(\Pi_T) - F_t(\pi_T)$	0.220*	0.242**	0.205*	0.242***	0.204**
	(0.122)	(0.092)	(0.106)	(0.089)	(0.091)
$\pi_t - F_t(\pi_t)$	1.166***	0.805**	0.728**	0.851**	0.774**
	(0.294)	(0.328)	(0.275)	(0.367)	(0.327)
$y_t - F_t(y_t)$				-0.017	-0.018
				(0.014)	(0.014)
$\Delta F_t(\pi_T)$		0.392**	0.369***	0.378**	0.354**
		(0.146)	(0.136)	(0.141)	(0.143)
$F_{t-1}(\Pi_T)$			0.040		0.041
			(0.03)		(0.027)
Constant	0.040*	0.035	-0.123	0.032	-0.129
	(0.022)	(0.022)	(0.12)	(0.02)	(0.11)
R^2	0.383	0.489	0.519	0.508	0.539
Adj. R^2	0.354	0.454	0.473	0.461	0.482
Obs.	47	47	47	47	47

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 1. The sample has 47 observations corresponding to the period 2008Q1-2019Q3.

Table C.2: Influence of IR's Quantitative Information on Next-Year Inflation Expectations - Sample Without COVID Crisis

Regressor	Model 1	Model 2	Model 3	Model 4	Model 5
$F_t(\Pi_{T+1}) - F_t(\pi_{T+1})$	0.121**	0.126**	0.114***	0.087**	0.093**
	(0.059)	(0.058)	(0.038)	(0.042)	(0.038)
$\Delta F_{t+1}(\pi_T)$	0.171**	0.159**	0.121**	0.146*	0.127**
	(0.072)	(0.063)	(0.05)	(0.073)	(0.059)
$\Delta F_{t+1}(e_{T+1})$				0.070	0.072*
				(0.042)	(0.039)
$\Delta F_{t+1}(i_{T+1})$			0.139*	0.071	0.080*
			(0.082)	(0.049)	(0.045)
$F_{t+1}(\Pi_{T+2}) - \pi^*$		0.159*	0.197**		0.205*
		(0.093)	(0.089)		(0.111)
Constant	0.061*	-0.022	-0.043	0.047*	-0.061
	(0.033)	(0.047)	(0.041)	(0.027)	(0.053)
R^2	0.229	0.272	0.437	0.441	0.511
Adj. R^2	0.194	0.221	0.384	0.388	0.451
Obs.	47	47	47	47	47

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México*.

Notes: See notes in Table 2. The sample has 47 observations corresponding to the period 2008Q1-2019Q3.

Table C.3: Influence of IR's Quantitative Information on Current-Year GDP Growth Expectations - Sample Without COVID Crisis

Regressor	Model 1	Model 2	Model 3	Model 4	Model 5
$F_t(Y_T) - F_t(y_T)$	0.635***	0.600***	0.558***	0.563***	0.524***
	(0.144)	(0.140)	(0.142)	(0.196)	(0.154)
$y_t - F_t(y_t)$	0.133***	0.139***	0.126***	0.106***	0.122***
	(0.03)	(0.036)	(0.029)	(0.023)	(0.038)
$\Delta F_{t+1}(y_T^{US})$			0.362**	0.304**	0.359**
			(0.147)	(0.146)	(0.176)
$\Delta F_t(y_T)$				0.195***	0.087
				(0.066)	(0.099)
$F_{t-1}(Y_T)$		0.035*			0.031*
		(0.018)			(0.018)
Constant	-0.084***	-0.170***	-0.072**	-0.058**	-0.141**
	(0.031)	(0.056)	(0.03)	(0.024)	(0.059)
R^2	0.501	0.546	0.532	0.558	0.586
Adj. R^2	0.479	0.515	0.499	0.515	0.535
Obs.	47	47	47	47	47

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México*.

Notes: See notes in Table 3. The sample has 47 observations corresponding to the period 2008Q1-2019Q3.

Table C.4: Influence of IR's Quantitative Information on Next-Year GDP Growth Expectations - Sample Without COVID Crisis

Regressor	Model 1	Model 2	Model 3	Model 4	Model 5
$F_t(Y_{T+1}) - F_t(y_{T+1})$	0.307***	0.231**	0.217**	0.228***	0.220**
	(0.113)	(0.087)	(0.084)	(0.082)	(0.081)
$\Delta F_{t+1}(y_T)$	-0.108	-0.076	-0.092	-0.079	-0.089
	(0.095)	(0.095)	(0.098)	(0.084)	(0.085)
$\Delta F_{t+1}(e_{T+1})$		-0.163***	-0.201***	-0.159***	-0.184***
		(0.023)	(0.048)	(0.023)	(0.042)
$\Delta F_{t+1}(i_{T+1})$			0.102		0.064
			(0.116)		(0.077)
$\Delta F_{t+1}(y^{US}_{T+1})$				0.444***	0.403***
				(0.157)	(0.128)
Constant	-0.077***	-0.062***	-0.057***	-0.057***	-0.054***
	(0.021)	(0.016)	(0.017)	(0.017)	(0.019)
R^2	0.278	0.507	0.533	0.578	0.588
Adj. R^2	0.241	0.468	0.482	0.533	0.531
Obs.	42	42	42	42	42

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Notes: See notes in Table 4. The sample has 42 observations corresponding to the period 2009Q2-2019Q3.

Table C.5 shows that more than 53% of the narrative shocks are selected (and thus relevant) for all forecasts. Compared to the period with the COVID crisis, fewer narrative shocks are selected for next-year inflation forecasts and for current-year GDP growth forecasts.

Table C.5: Selected Narrative Shocks for each Forecast

Forecast	# Selected	% Selected
Current-year Inflation	39	75
Next-year Inflation	28	53.8
Current-year GDP Growth	36	69.2
Next-year GDP Growth	37	71.1

Source: Own calculations with data from the Expectations Survey of Private Sector Forecasters and IR summaries of *Banco de México.*

Figure C.1: Top 10 Terms Within Topics Ranked by Probability - Sample Without COVID Crisis

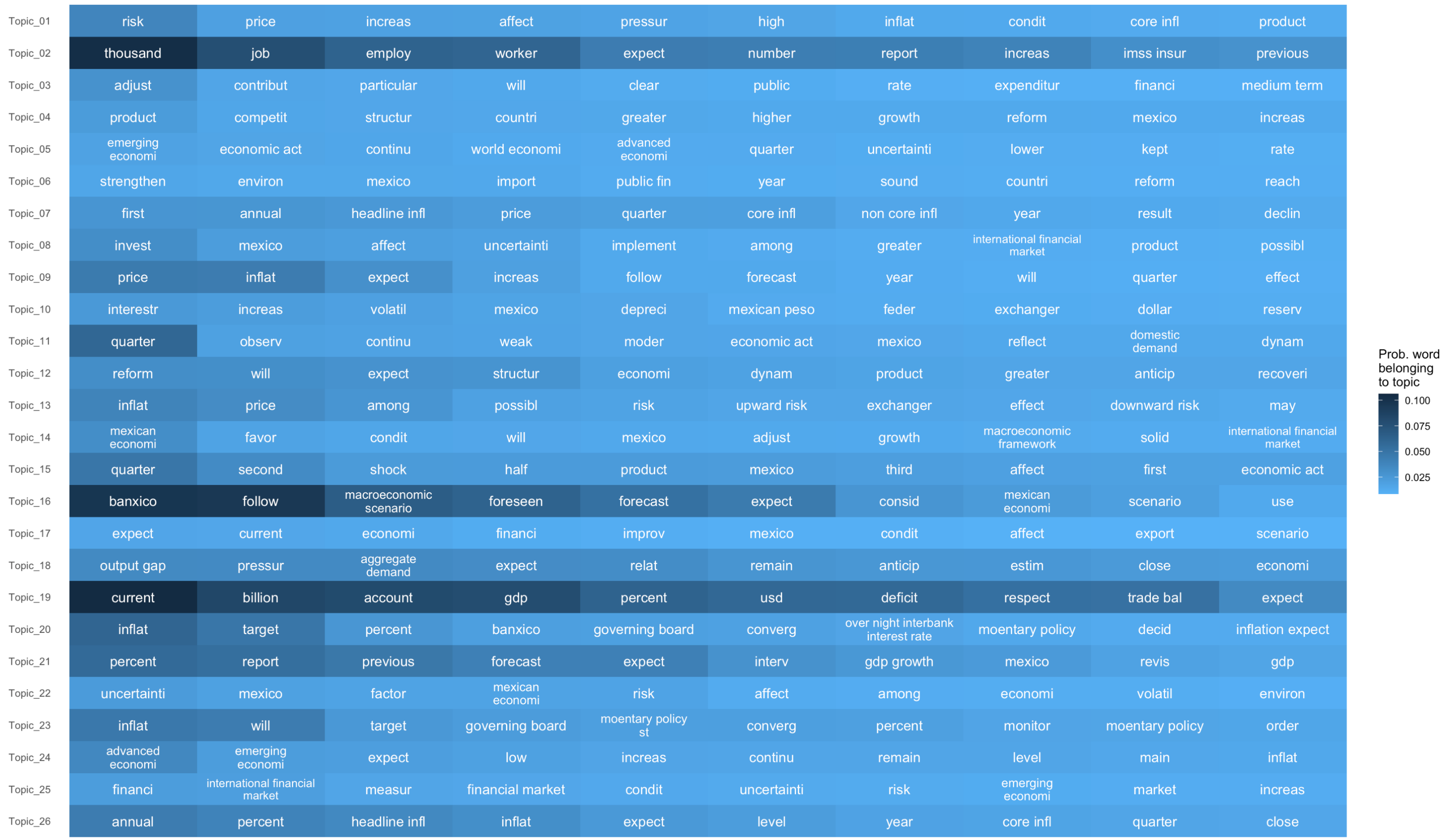


Figure C.1 shows the new topic distribution for the sample excluding the COVID-19 crisis. As mentioned in the main text, this topic distribution is different from the one obtained for the whole sample even though the number of topics remains the same.

The top panel in Figure C.2 represents the four key topics for analysts when examining current-year inflation forecasts: the change in formal employment projections (Δ Topic 2), risks to economic activity (Topic 22), the change in risks to inflation (Δ Topic 13) and economic activity (Topic 15). Analogously, the bottom panel shows that the change in the monetary policy decision (Δ Topic 20), risks to inflation (Topic 13), the long-run growth perspectives (Δ Topic 12) and growth projections (Topic 17) are important topics for assessing next-year inflation forecasts.

Similarly, Figure C.3 shows the four key topics when analyzing current and next-year GDP growth forecasts for the reduced sample. For current-year GDP growth forecasts, the top panel shows that forward guidance (Topic 23), the change in risks to financial markets (Δ Topic 25), the change in observed inflation (Δ Topic 7) and financial markets (Topic 10) are the most relevant topics for market analysts. For next-year GDP growth forecasts, the bottom panel illustrate that the main topics are those related to investment (Topic 8), the monetary policy decision (Topic 20), the change in formal employment projections (Δ Topic 2) and risks to inflation (Topic 1).

Figure C.2: Key Topics for Current-Year and Next-Year Inflation Forecasts - Sample Without COVID Crisis

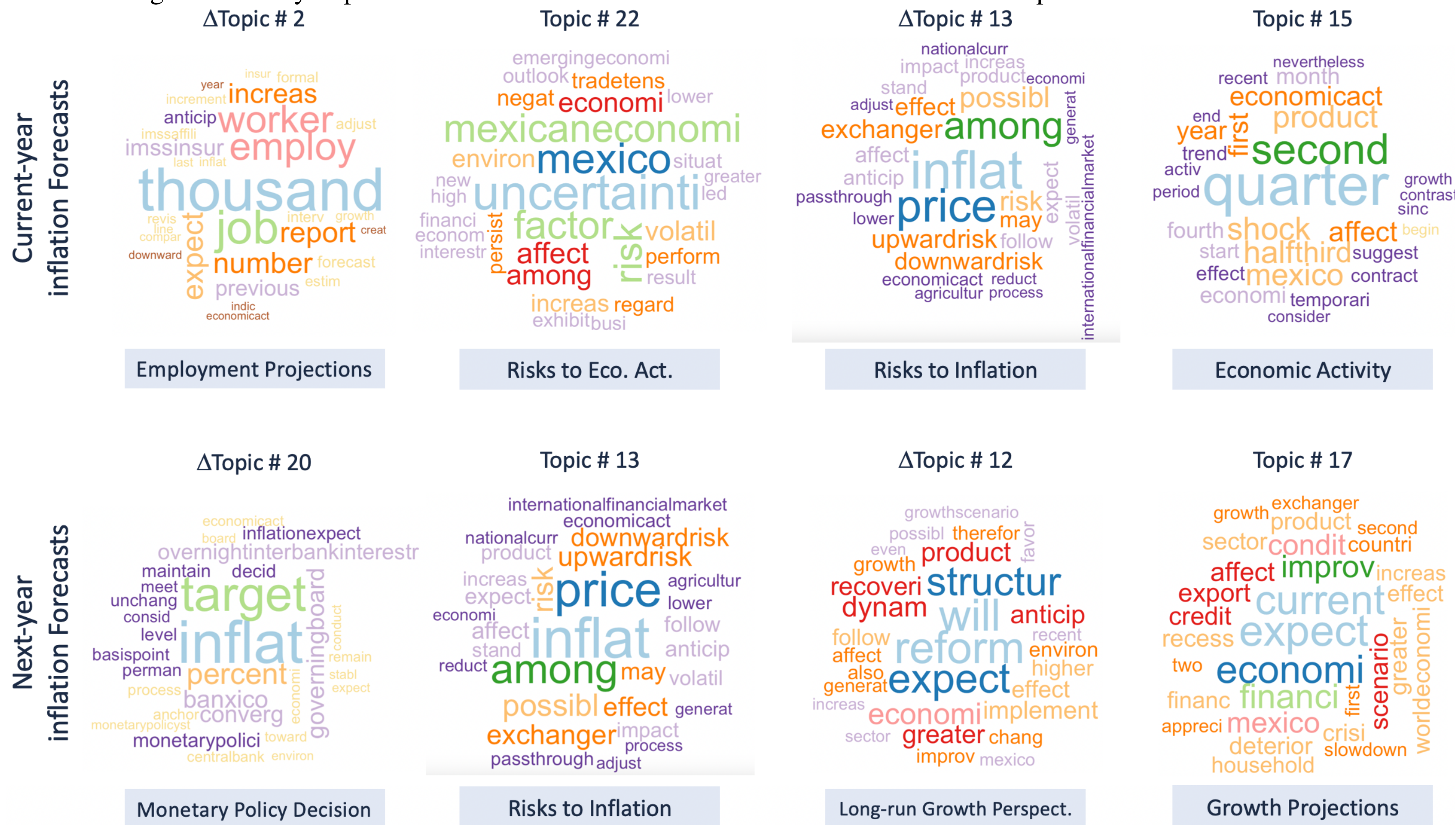


Figure C.3: Key Topics for Current-Year and Next-Year GDP Growth Forecasts - Sample Without COVID Crisis