Extracting firms' short-term inflation expectations from survey comments using text analysis*

Jouchi Nakajima† Tatsushi Okuda‡ Hiroaki Yamagata‡
Shinnosuke Katsuki‡ Takeshi Shinohara‡

May 2022

Abstract

This paper proposes a monthly quantitative indicator of firms' inflation expectations, developed from the textual data of a nation-wide survey for firms in Japan. The text-based expectations are computed by extracting firms' views from survey comments, using a machine learning method. Empirical analyses show that the indicator tends to precede consumer price inflation by several months and that it is highly correlated with existing quarterly indicators of inflation expectations, implying that the text-based expectations could be a timely indicator of firms' inflation expectations. The analyses also indicates that the text-based expectations comove with both demand and cost variables while it also includes unique information for forecasting inflation rates.

JEL classification: C53, C55, E31, E37.

Key words: Inflation Expectations, Machine Learning, Text Analysis.

*The authors thank Kosuke Aoki, Ryo Jinnai, Seisaku Kameda, Kazushige Kamiyama, Takuji Kawamoto, Ichiro Muto, Takashi Nagahata, Teppei Nagano, Koji Nakamura, Koji Takahashi, and staff members of the Bank of Japan for their valuable comments. Financial support from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government through Grant-in-Aid for Scientific Research (No.20H00073) and the Hitotsubashi Institute for Advanced Study is gratefully acknowledged. The views expressed in this paper are those of the authors and do not necessarily reflect the official views of the Bank of Japan.

†Corresponding author. Hitotsubashi University. 2-1 Naka, Kunitachi-city, Tokyo 186-8603, Japan. E-mail: jouchi.nakajima@gmail.com

‡Bank of Japan.

1 Introduction

Inflation expectations are a key variable affecting macroeconomic outcomes and in recent years have attracted attention in both theoretical and empirical research. Some studies suggest that firms form their inflation expectations through a different mechanism than households and market participants (e.g., Kumar et al. (2015), Coibion, Gorodnichenko, and Kamdar (2018); Coibion, Gorodnichenko, and Kumar (2018)). On one hand, significant progress has been made in the study of households' and market participants' inflation expectations, reflecting the accumulation of related data. On the other hand, while firms are price setters and their inflation expectations are conventionally regarded as a critical variable that affects price developments, there has been relatively little progress in the study of firms' inflation expectations, partly reflecting the paucity of relevant data.

Against the backdrop, this paper proposes a monthly quantitative indicator of firms' inflation expectations, developed from the textual data of a nation-wide survey for firms in Japan. This study contributes to one strand of literature that uses textual data to create an indicator of inflation expectations. Guzman (2011) develops an indicator of U.S. inflation expectations using the number of Google search queries. Angelico et al. (2021) construct an indicator of Italian inflation expectations using textual data from Twitter. These indicators do not explicitly focus on firms' inflation expectations, but those of broader economic agents. An advantage of the current study is to explicitly focus on firms' inflation expectations by using the text data from the survey for firms.

The methodology to compute the indicator follows Otaka and Kan (2018), who demonstrate applications of machine learning methods to textual data of firms' comments in the Economy Watchers Survey (EWS, hereafter) conducted by the Cabinet Office of Japan. Among the applications, the authors develop an indicator, which is aimed to be a leading indicator for consumer price inflation. The firms' comments are classified into categories such as that implies inflation and that does deflation by a naïve Bayes classifier, which is a popular machine learning method (e.g., Murphy, 2012). The indicator is defined as the share of comments implying inflation minus the share of comments implying deflation. Otaka and Kan (2018) show that the indicator tends to precede the consumer price index (CPI) by several months, based on simple lag-lead correlations between the indicator and CPI inflation rates, and argue that the indicator appears to have potential for firms' inflation expectations.

This paper computes the indicator, defined as the *text-based inflation expectations*, up to the recent period, following the methodology of Otaka and Kan (2018), and formally examines the properties of the text-based inflation expectations, including the correlation with existing indicators of firms' inflation expectations, determinants of fluctuations in the text-based inflation expectations, the link with macroeconomic variables, and its usefulness in forecasting CPI inflation rates. Otaka and Kan (2018) originally propose the method to compute the indicator by applying the naïve Bayes classifier (e.g., Murphy, 2012) to the EWS textual data. The current paper uses the same method computing the indicator and uncovers the indicator's usefulness, superiority, and caveat in macroeconomic analysis.

Our empirical analyses show that the text-based inflation expectations tend to precede CPI inflation by several months and that it is highly correlated with existing quarterly indicators of firms' inflation expectations, implying that the text-based expectations could be a novel monthly indicator of firms' inflation expectations. The analyses also confirms that the text-based expectations comove with both demand and cost variables while it also includes unique information for forecasting inflation rates.

The text-based expectations have the following advantages. First, they are timely, since the EWS is a monthly survey and its results are released early in the following month. Second, the text-based expectations can be computed back to January 2000, meaning that they provide sufficiently long time series data to allow for quantitative analyses. Third, the text-based expectations reflect the views of EWS respondents, who hold jobs that enable them to closely watch developments in economic activity, in particular, of consumers. There is no such survey or statistics which gauge firms' inflation expectations in Japan. Furthermore, the method is generally applicable to a wide range of textual data in other countries.

The paper contributes to the literature on economic forecasting with text analysis which has recently grown rapidly. Several studies propose frameworks for forecasting business cycle indicators such as a GDP: Bybee et al. (2020), Shapiro et al. (2022), and Barbaglia et al. (2022) for the United States, Kalamara et al. (2020) for the United Kingdom, Barbaglia et al. (2021) for European countries, and Thorsrud (2020) for Norway. All of these studies use newspaper articles to derive textual information that is relevant predictors for the macroeconomic indicators. For forecasting the CPI inflation rates, Seabold and Coppola (2015), and Wei et al. (2017) exploit text analysis methods to construct useful indicators. Also, Goshima et al. (2021) develop a business cycle index based on the textual data of daily Japanese newspaper articles, utilizing a machine learning method with the EWS comments data exploited for training a text classification model. In these studies, the indexes obtained in the text analysis are used to forecast the inflation. In contrast, the current paper proposes the indicator of inflation expectations computed directly from the firms' comments based on the text analysis method. This work is also closely related to previous studies that create indicators of economic expectations from textual data (Sharpe et al., 2017; Ke et al., 2019).

The paper also contributes to the literature on measuring firms' inflation expectations. As Coibion et al. (2020) point out, the availability of surveys of firms is relatively limited. In the United States, the Atlanta Fed's Business Inflation Expectations survey is available monthly (Bryan et al., 2014). Yet, the time series of the data is relatively short in the sense that the start date of the survey is 2012 year. Livingston Survey, conducted by the Federal Reserve Bank of Philadelphia, provides a long time series of large firms' long-term inflation expectations, but the frequency is semi-annual. the Bank of Japan surveys Japanese firms' inflation expectations (Short-Term Economic Survey of Enterprises in Japan, known as *Tankan*; see Muto (2015)). However, the figures for inflation expectations are available only from 2014, when their questionnaire was added to the survey items, and the survey frequency is quarterly.[1] The methodology to compute the text-based expectations used in this study has the potential to create higher frequency indicators for firms' inflation expectations with more extended time series than existing ones.

The remainder of the paper is organized as follows. Section 2 explains the computational method for the text-based expectations. Section 3 examines the validity of the text-based expectations as a proxy for inflation expectations. Section 4 analyzes the causes of changes in the text-based expectations and the relationship between the text-based expectations and macroeconomic variables. Section 5 concludes.

[1]In terms of the lists of the regularly conducted surveys of firms' inflation expectations in other countries and regions, see Table 2 of Coibion et al. (2020). In addition to these surveys, surveys of firms' inflation expectations have been conducted ad-hoc. For example, Kumar et al. (2015), Coibion, Gorodnichenko, and Kumar (2018), and Coibion et al. (2021) conducted surveys of New Zealand firms. Coibion et al. (2019) and Andrade et al. (2022) surveyed Italian and French firms, respectively.

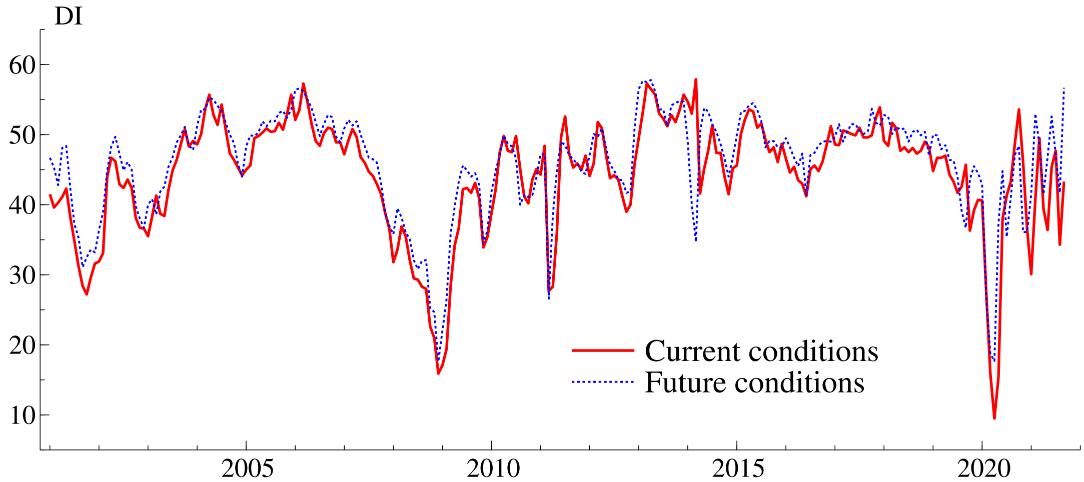


Figure 1: The diffusion index (DI) for current and future economic conditions in Economy Watchers Survey.

2 Methodology to compute the text-based expectations

2.1 The Economy Watchers Survey (EWS)

The EWS has been conducted monthly by the Cabinet Office since January 2000. The survey aims to grasp developments in Japan's economy in a timely manner. Each month, 2,050 people across Japan receive the survey and about 1,800 of them provide valid responses. Survey respondents consist of "economy watchers," that is, individuals holding jobs that enable them to closely watch developments in economic activity: for example, business managers and grocery clerks. In the survey, those engaged in household activity-related sectors account for about two thirds of respondents, while those working in corporate activity-related and employment-related sectors account for around 20 percent and 10 percent, respectively. This means that many of the survey respondents are engaged in industries that have a relatively close link with consumers.

The headline results from the EWS are the diffusion indices (DIs) for current and future economic conditions, presented in Figure 1. They are calculated using each respondent's assessment of current or future economic conditions on a scale comprising five categories ranging from, e.g., "better" to "worse." The DI for current economic condi-

Table 1: Examples of comments by EWS respondents on their assessment of economic conditions.

Assessment of economic conditions	Sector (Occupation)	Comment
Slightly better	Supermarket (Store manager)	While average sales per customer remain sluggish, the number of customers has been increasing.
Unchanged	Job placement office (Staff)	Despite a downward trend in job openings compared with the previous year, managers seem to struggle to fill vacancies and that there remains a sense of labor shortage in the nursing-care and construction sectors.

Note: Comments are authors' translations of the Japanese original.

tions has been regarded as a timely and useful indicator for assessing economic activity, as it shows some correlation with other macroeconomic indicators that capture economic developments (e.g., Bragoli, 2017).

The EWS is unique in that it collects not only respondents' assessment of economic conditions on a scale as just described but also their comments giving reasons for their assessment. On average, approximately 1,100 of the about 1,800 respondents in each survey provide comments on their economic assessments. These comments are organized and released as textual data on the Cabinet Office's website. Such data in the release for each survey consist of about 100,000 words in total. Examples of such comments are provided in Table 1. Note that the EWS comments presented in the table are our translations of those in the Japanese original. The respondents report how the economic conditions have changed based on findings from their business activity. There is no questionnaire specifically about the price developments, while respondents often refer to words regarding price developments in their answers to reasons for their assessment of economic conditions. The text-based expectations are derived from this big textual data and are computed by extracting and quantifying information regarding price developments from the data, using text analysis. The specific methods for computing the text-based expectations are presented in the next subsection.

2.2 Computing method for the text-based expectations

When respondents provide comments giving the reasons for their economic assessments in the EWS, they sometimes also refer to consumers' spending stance and to developments in prices such as commodity prices. The text-based expectations are designed to capture developments in the difference between the share of comments implying inflation and the share of comments implying deflation by classifying comments. Specifically, survey comments are classified into the following four categories:

(A) Comments implying inflation;

(B) Comments implying deflation;

(C) Comments implying zero inflation (neither inflation nor deflation);

(D) Comments not referring to price developments.

Manually screening the comments received in each survey to classify them into these four categories would require considerable time and effort. Moreover, such manual classification could result in incorporating the analysts' subjective views into the text-based expectations. Therefore, for computing the text-based expectations, comments are automatically classified into the four categories based on the words contained in each comment. This is done using the naïve Bayes classifier (e.g., Murphy, 2012).

Let $(W_1, \ldots, W_I)$ denote a set of all unique words that appear in comments, where I is the number of words. We define w_{hi} as the number of times the word W_i appears in the h-th comment in data, denoted by $s^{(h)}$. We treat the comment as a vector of the number of word's appearance, i.e., $s^{(h)} = (w_{h1}, \ldots, w_{hI})$.

The classifier is a supervised learning algorithm, utilizing the Bayes theorem, to classify the vector of quantities $s^{(h)}$ to one of J categories, denoted by $(c_1, \ldots, c_J)$. Using the Bayes theorem, we obtain a conditional probability of the comment $s^{(h)}$ belongs to the category c_j as

$$\mathrm{P}(c_j|s^{(h)}) \propto \mathrm{P}(c_j) \cdot \mathrm{P}(s^{(h)}|c_j). \tag{1}$$

In the classifier, we assume a "naïve" assumption of conditional independence, that is,

$$\mathrm{P}(w_{hi}|c_j, w_{h1}, \ldots, w_{h,i-1}, w_{h,i+1}, \ldots, w_{hI}) = \mathrm{P}(w_{hi}|c_j),$$

for all $i = 1, \ldots, I$, and $j = 1, \ldots, J$. Then, equation (1) is simplified as

$$\mathrm{P}(c_j|s^{(h)}) \;\propto\; \mathrm{P}(c_j) \cdot \prod_{i=1}^{I} \mathrm{P}(w_{hi}|c_j).$$

Define $p_j = \mathrm{P}(c_j)$, and $q_{ij} = \mathrm{P}(w_{hi}|c_j)$. A classification rule is given by

$$j^*(s^{(h)}) \;=\; \arg\max_j p_i \prod_{i=1}^{I} q_{ij},$$

where $j^*(s^{(h)})$ denotes the category index for the comment $s^{(h)}$.

We use the maximum a posteriori (MAP) estimator to estimate $p \equiv (p_1, \ldots, p_J)$ and $q \equiv (q_1, \ldots, q_J)$, where $q_j \equiv (q_{1j}, \ldots, q_{Ij})$. We assume that p follows a Dirichlet distribution, $\mathrm{Dir}(\alpha_p)$, with a vector of hyperparameters, $\alpha_p = (\alpha_{p1}, \ldots, \alpha_{pJ})$; and $q_j \sim \mathrm{Dir}(\alpha_{qj})$, with $\alpha_{qj} = (\alpha_{qj1}, \ldots, \alpha_{qjI})$. The category of a comment is assumed to follow the multinomial distribution, $\mathrm{Mult}(1, p)$. We also assume that the number of words in the comment $s^{(h)}$, denoted by N_h, is independent of the category. The number of each word which appears in the comment follows the multinomial distribution, $\mathrm{Mult}(N_h, q_j)$, conditional on that the comment belongs to the category c_j.

For a training dataset which consists of K comments, define $s^{(k)}$ and $c^{(k)}$ denote the k-th comment and its category, where $s^{(k)} \equiv (w_{k1}, \ldots, w_{kI})$. The MAP estimator is given by

$$\hat{p}_j \;=\; \frac{M_j + \alpha_{pj} - 1}{\sum_{j=1}^{J}(M_j + \alpha_{pj} - 1)}, \quad \hat{q}_{ij} \;=\; \frac{m_{ij} + \alpha_{qij} - 1}{\sum_{i=1}^{I}(m_{ij} + \alpha_{qij} - 1)},$$

where $m_{ij} = \sum_{k=1}^{K} w_{ki} \cdot I[c^{(k)} = c_j]$, $M_j = \sum_{i=1}^{I} m_{ij}$, and $I[\,\cdot\,]$ denotes an indicator function that takes one when the argument is true, and zero otherwise. In the following analysis, we set $\alpha_{pj} = \alpha_{qij} = 2$, for all i and j to make a prior information as uninformative as possible.

To construct the training data, we randomly select K = 1,500 comments from the EWS during the period 2001–2017. Each of these comments is then manually read and classified into one of the $J = 4$ categories. This procedure leaves some room for analysts' subjective opinions to be incorporated into the index. To minimize such potential bias to the greatest extent possible, comments in which respondents' views on price developments appear to be ambiguous are classified as category D. Further, the classification process is conducted by two annotators so that the classification reflects average opinions.

The text-based expectations are defined as the share of comments implying inflation (type A) minus the share of comments implying deflation (type B). Specifically, for each survey (month), we compute

$$X_t = \frac{n_t(A) - n_t(B)}{n_t(A) + n_t(B) + n_t(C)},$$

where X_t denotes the text-based expectations, and $n_t(c)$ denotes the number of comments which belongs to the category c in the survey (month) t. In the remainder of this paper, we use the series of text-based expectations normalized by the mean and standard deviation for the period 2000–2019.

We use all the comments in the dataset to construct the text-based expectations. As mentioned earlier, not all the comments mention the price development. An alternative approach can be to isolate the more relevant comment to forecast price development, for example, by focusing on specific terms related to the price or inflation. This approach may target our purpose and perform better than the current approach. However, in the EWS, we did not find appropriate stop words to isolate the comments on price development generally, and such selection of comments may be arbitrary. For these reasons, we end up using all the available comments. Moreover, note that the share of the type-D comments is, on average, over half of all comments, while it varies according to economic conditions. This finding implies that, to some extent, our approach isolates the comments which are irrelevant to price developments.[2]

While the text-based expectations are computed separately for comments on current economic conditions and on future economic conditions, the following analyses use the text-based expectations based on comments on current economic conditions. We also conducted the analyses using the text-based expectations based on comments on future economic conditions and found that results do not change significantly. In the following analysis, we use the text-based expectations from 2001 onward as the number of responses

[2]On average, the shares of type-A and type-B comments are around 15 and 30 percent, respectively. The number of type-C comments in each wave is a few.

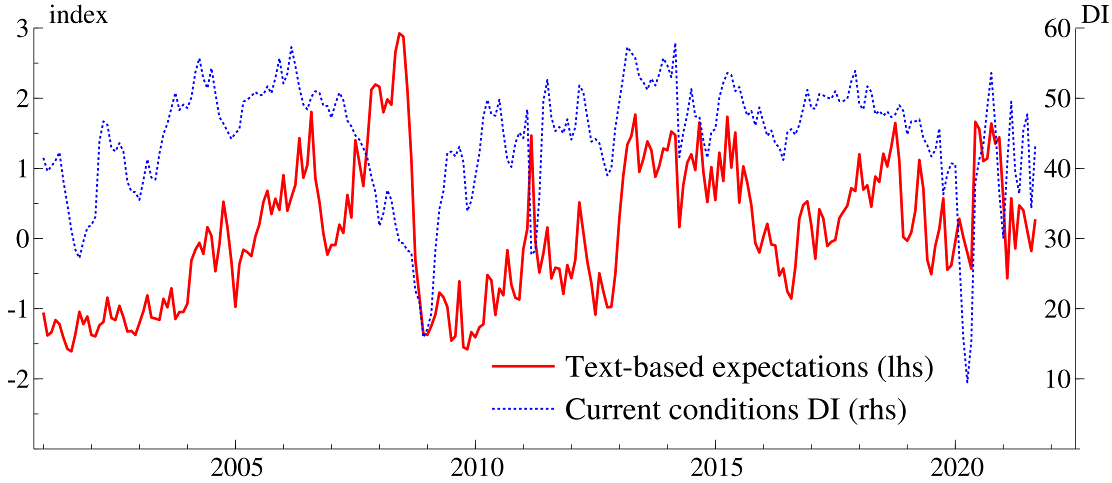


Figure 2: The text-based expectations and current economic conditions DI.

in 2000, the year the survey was started, is considerably smaller than from 2001 onward.

3 The text-based expectations as a proxy for short-term inflation expectations

3.1 The text-based expectations

Figure 2 plots the series of text-based expectations in addition to the current economic conditions DI from the EWS. It is clear that developments in the text-based expectations differ from those in the current economic conditions DI. For example, around 2007–2008, when commodity prices were surging, the current economic conditions DI declined due to concerns over a decrease in profits, whereas the text-based expectations rose, clearly reflecting the rise in raw materials prices. Around 2008–2009, the text-based expectations fell substantially in tandem with the current economic conditions DI amid the significant decline in demand both at home and abroad due to the impact of the global financial crisis. These findings suggest that developments in the text-based expectations are significantly affected by changes in demand due to the business cycle and also by cost factors such as commodity price changes.

We investigate the relationship between the text-based expectations and CPI inflation. Figure 3 plots the text-based expectations and the year-on-year rate of change in the

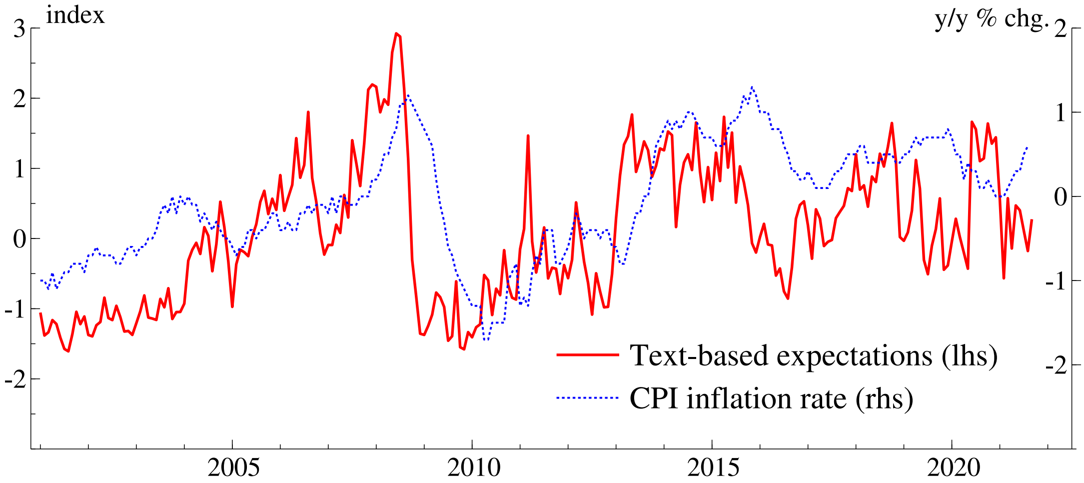


Figure 3: The text-based expectations and CPI inflation rate (all items less fresh food and energy). The CPI inflation rate excludes temporary factors such as mobile phone charges and the effects of the consumption tax hikes.

consumer price index (CPI, all items less fresh food and energy). As can be seen in the figure, developments in the text-based expectations appear to somewhat precede those in the CPI inflation rate. This visual impression is confirmed when we estimate simple lead-lag correlation coefficients between the two variables for the period through the end of 2019. The correlation coefficient between the text-based expectations and the seasonally adjusted quarter-on-quarter rate of change in the CPI is largest, taking a value of 0.543, when the text-based expectations lead the CPI inflation rate by one month. We also find that the correlation coefficient between the text-based expectations and the year-on-year rate of change in the CPI is largest, taking a value of 0.765, when the text-based expectations leads the CPI inflation rate by seven months. These correlation coefficients are surprisingly high, which suggests that the text-based expectations could be useful as a proxy for firms' inflation expectations.

Behind these interesting findings, the naïve Bayes classifier plays a key role in reflecting the textual information about the near-future price development. The driver of increasing (decreasing) text-based expectations is relatively increasing the share of comments implying inflation (deflation), defined as type-A (type-B) comments in the previous section. The estimated "score" of the word i for the comment type j, $\hat{q}_{ij}$ measures a marginal increase in the likelihood that the comment belongs to the type j when the word i is used in the comment. We find that terms such as "rise", "high", "exceed", "price increase",

Table 2: Correlation between the text-based expectations and existing indicators of firms' inflation expectations.

(a) The text-based expectations and *Tankan* DI

	Actual result	Forecast
DI for output prices	0.826	0.860
DI for input prices	0.893	0.923
DI for supply and demand conditions	0.703	0.711

(b) The text-based expectations and firms' inflation outlook (*Tankan*)

	1-year ahead	3-year ahead	5-year ahead
Outlook for output prices	0.809	0.782	0.716
Outlook for general prices	0.687	0.645	0.629

Note: The sample period is (a) 2001/Q1–2019/Q4, and (b) 2014/Q1–2019/Q4.

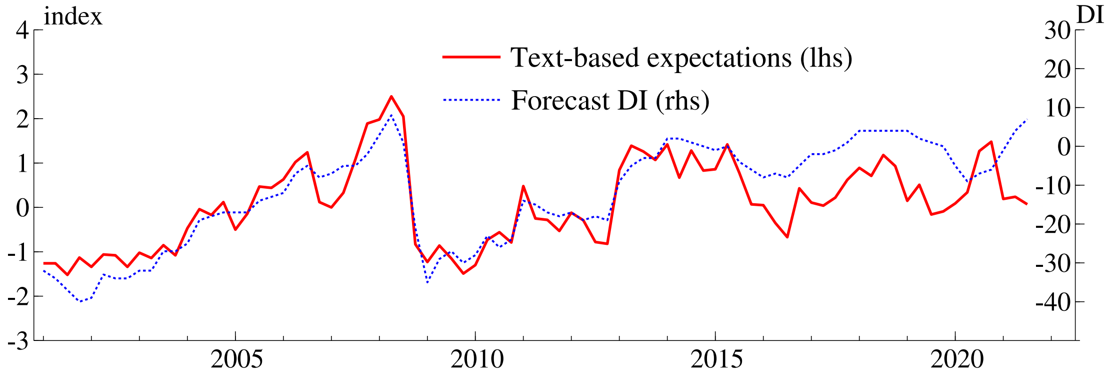
and "surge" have a relatively higher score for type A, compared to type B. In contrast, the terms such as "decline", "cheap", "price cut", "sluggish", and "sale" have a relatively higher score for type B, compared to type A. This estimation result is so intuitive that we see that the naïve Bayes classifier works properly in our data and context.

3.2 The link between the text-based expectations and existing indicators of firms' inflation expectations

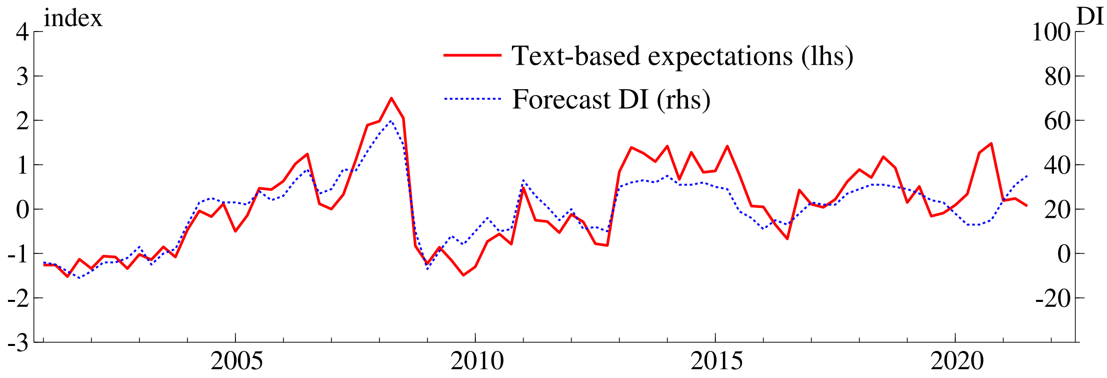
To examine the validity of the text-based expectations as a proxy for firms' inflation expectations, we compare the series with existing lower frequency indicators for firms' inflation expectations. Figure 4 plots the text-based expectations and the forecast (one quarter ahead) DIs of *Tankan*.[3] Because the DI series is quarterly, we convert the text-

[3]The *Tankan* survey asks roughly 10,000 Japanese firms to report their assessments of questionnaire items, including the current ("actual result") and future output prices, input prices, and supply and demand conditions. Firms' answers are qualitative in that their answers are chosen from three possible responses. Specifically, the question and the candidate responses for output price questionnaire are as follows: *Please choose the option which best describes the current conditions, excluding seasonal factors:*

(a) DI for output prices



(b) DI for input prices



(c) DI for supply and demand conditions

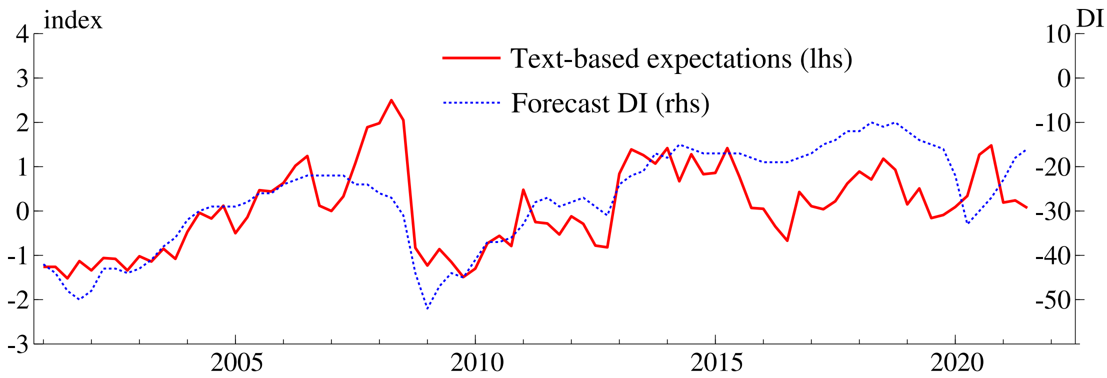


Figure 4: The text-based expectations and *Tankan* forecast DIs.

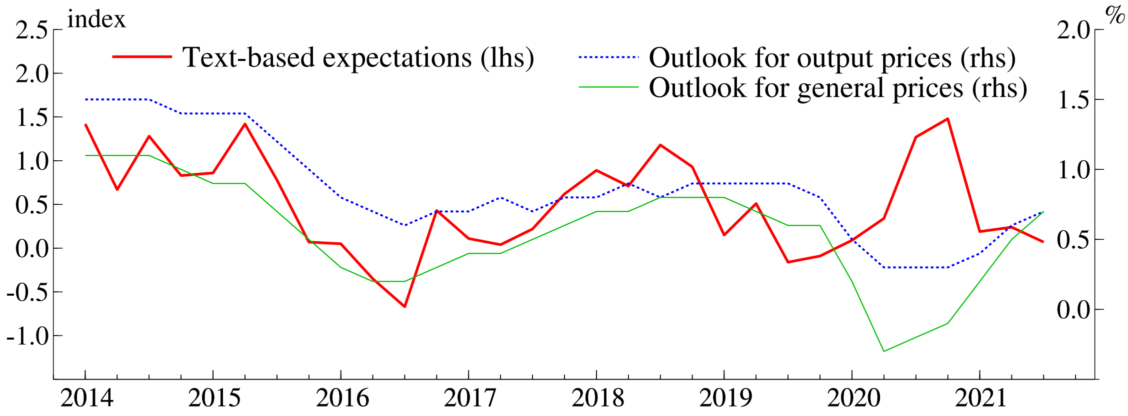


Figure 5: The text-based expectations and 1-year-ahead inflation outlook of enterprises (*Tankan*).

based expectations to a quarterly series by taking a three-month average. We find that the text-based expectations has a high correlation with the DIs for output prices, for input prices, and for domestic supply and demand conditions. This suggests that the text-based expectations reflect firms' expectations of their prices, demand conditions and input costs. Table 2(a) shows that the correlation coefficients between the text-based expectations and the *Tankan* DIs are high, and that the coefficients are all, albeit slightly, higher for the *forecast* DI than for the *actual* DI. This suggests that the text-based expectations provide a useful proxy for short-term inflation expectations that captures firms' price-setting stance for the period ahead rather than their current price-setting stance. This is also consistent with the earlier finding that the text-based expectations somewhat lead the inflation rate.

Figure 5 plots the text-based expectations and firms' one-year-ahead outlook for general prices reported in *Tankan*, which shows a close relationship.[4] Table 2(b) reports the correlation coefficient between the text-based expectations and the inflation outlook at one-year, three-year, and five-year inflation expectations in *Tankan*. We find that the one-year inflation expectations exhibit a higher correlation coefficient with the text-based expectations than the three-year and five-year. This evidence suggests that the text-

Rise, Unchanged, or Fall. Each answer is assigned a score value of 1, 0, and −1, respectively. The Bank of Japan then calculates the aggregated series, called Diffusion Indexes (DIs), by taking simple averages of the samples.

[4]The *Tankan* survey started to ask firms to report their inflation outlook in 2014. The item is qualitative in the sense that firms choose their answers from ten possible answers (e.g., around 0 percent, around 1 percent, and so on) and the Bank of Japan calculate the averages of them.

based expectations are closely related to the firms' inflation expectations, in particular, for several months ahead.

4 Further exploration of text-based expectations

We examine whether the text-based expectations provide unique information such that it complements information on the output gap and other conventional macroeconomic variables for forecasting of inflation. Specifically, we conduct regression analyses using the h-quarter-ahead inflation rate as the dependent variable, for $h = 1, \ldots, 4$. In this analysis, we first estimate a regression equation using the exchange rate and the output gap as independent variables. We then add the text-based expectations as an independent variable to the equation and examine how the regression results change as a result. The inflation rate is measured in terms of the year-on-year rate of change in the CPI (all items less fresh food and energy, excluding the effects of the consumption tax hikes), while for the exchange rate the year-on-year rate of change in the nominal effective exchange rate is used. The output gap is estimated by the Bank of Japan. The estimation period for the regression is from 2001/Q1 to 2019/Q4. We limit the sample period through to the end of 2019 because the output gap estimate may have a great uncertainty in 2020 onward due to a large decline in the output during the COVID-19 pandemic.

The regression results are shown in Table 3. In the specification without the text-based expectations, the coefficients on the output gap and the exchange rate are statistically significant for all the horizons. When the text-based expectations are added, these coefficients remain statistically significant, and importantly, the coefficient on the text-based expectations is significant for the one-quarter to three-quarter horizons. We further find that the explanatory power in terms of the adjusted R-squared is higher when the text-based expectations are included. These results suggest that the text-based expectations capture additional information relevant for changes in the inflation rate not captured by the output gap and the exchange rate.

Next, we employ a vector autoregression (VAR) model to examine the relationship between the text-based expectations and macroeconomic variables including the inflation rate. In this estimation, we use four variables: the nominal effective exchange rate (quarter-on-quarter change); the output gap; the text-based expectations; and the CPI (all items less fresh food and energy; seasonally adjusted quarter-on-quarter change). The

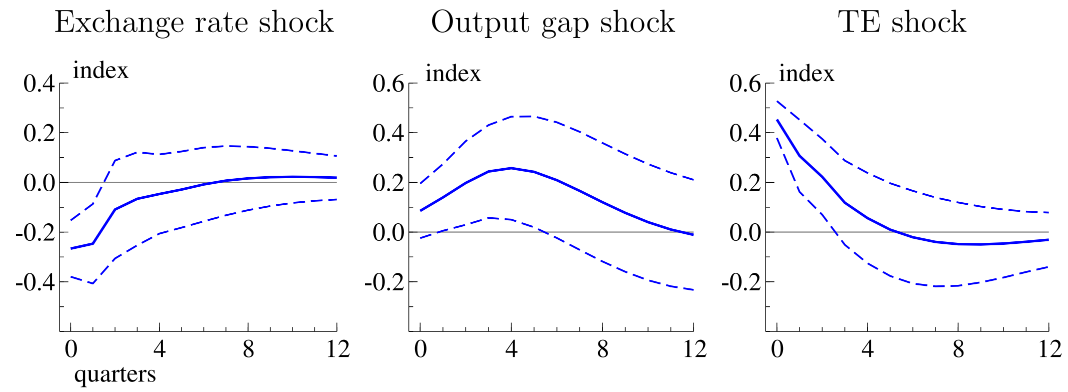
Table 3: Regression results: specifications including and excluding the text-based expectations (TE).

	CPI, 1-quarter ahead		CPI, 2-quarter ahead	
	w/o TE	with TE	w/o TE	with TE
Constant	0.050 (0.023)**	0.013 (0.026)	0.087 (0.046)*	0.019 (0.053)
CPI (current)	0.858 (0.037)***	0.808 (0.051)***	0.663 (0.070)***	0.568 (0.115)***
Output gap	0.068 (0.023)***	0.037 (0.021)*	0.127 (0.040)***	0.067 (0.036)*
Exchange rate	−0.009 (0.003)***	−0.006 (0.003)**	−0.017 (0.006)**	−0.012 (0.004)***
TE		0.114 (0.052)**		0.206 (0.096)**
Std.Err.	0.203	0.192	0.288	0.261
Adjusted R^2	0.906	0.916	0.810	0.844

	CPI, 3-quarter ahead		CPI, 4-quarter ahead	
	w/o TE	with TE	w/o TE	with TE
Constant	0.100 (0.052)*	0.040 (0.066)	0.094 (0.059)	0.055 (0.061)
CPI (current)	0.477 (0.093)***	0.388 (0.112)***	0.322 (0.117)***	0.260 (0.112)**
Output gap	0.159 (0.043)***	0.106 (0.041)**	0.169 (0.036)***	0.135 (0.048)***
Exchange rate	−0.024 (0.008)***	−0.020 (0.006)***	−0.027 (0.009)***	−0.025 (0.010)**
TE		0.181 (0.070)**		0.115 (0.079)
Std.Err.	0.346	0.330	0.407	0.404
Adjusted R^2	0.726	0.765	0.621	0.628

Note: The CPI (current) refers to the year-on-year rate of change in the CPI (all items less fresh food and energy, excluding the effects of the consumption tax hikes and policies concerning the provision of free education) at a current quarter. The exchange rate refers to the year-on-year rate of change in the nominal effective exchange rate. The estimation period is from 2001/Q1 to 2019/Q4. Figures in parentheses are heteroskedasticity- and autocorrelation-consistent (HAC) standard errors. ***, **, and * denote statistical significance at the 1, 5, and 10 percent levels, respectively.

(a) Impulse response of the text-based expectations (TE)



(b) Impulse response of the CPI

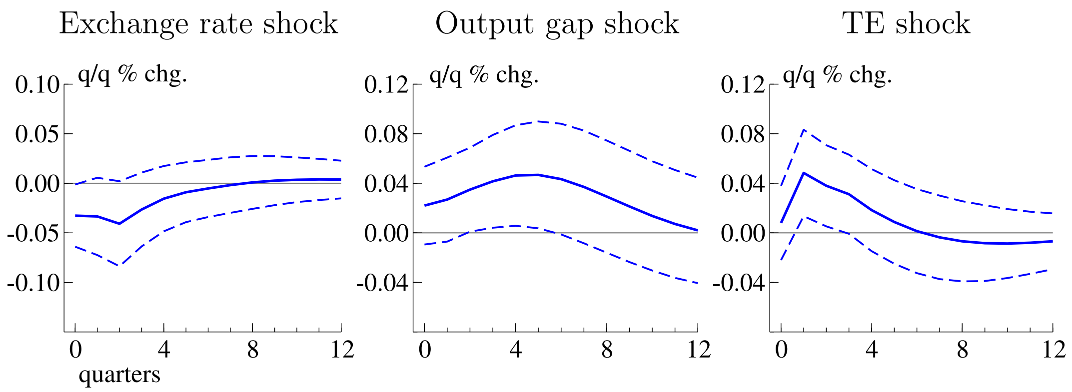


Figure 6: Impulse responses from VAR model with the exchange rate. The size of the shock is one standard deviation. The dashed lines indicate the 95 percent confidence intervals.

estimation period is from the 2001/Q1 to the 2019/Q4. Based on the Akaike information criterion (AIC), the lag length is set to two quarters.

We identify shocks using Cholesky decomposition, with the variables ordered as above. The variables are ordered from the most exogenous to the least exogenous one. This reflects our assumptions regarding the nature of the quarterly shock to each variable. Specifically, we assume that, during the same quarter, (i) an exchange rate shock may affect all the other variables, (ii) an output gap shock may influence firms' inflation expectations (the text-based expectations) and the actual inflation rate (the CPI), and (iii) a text-based expectations-shock may have an impact on the CPI. A shock to the CPI here is assumed to have no impact on the text-based expectations during the same

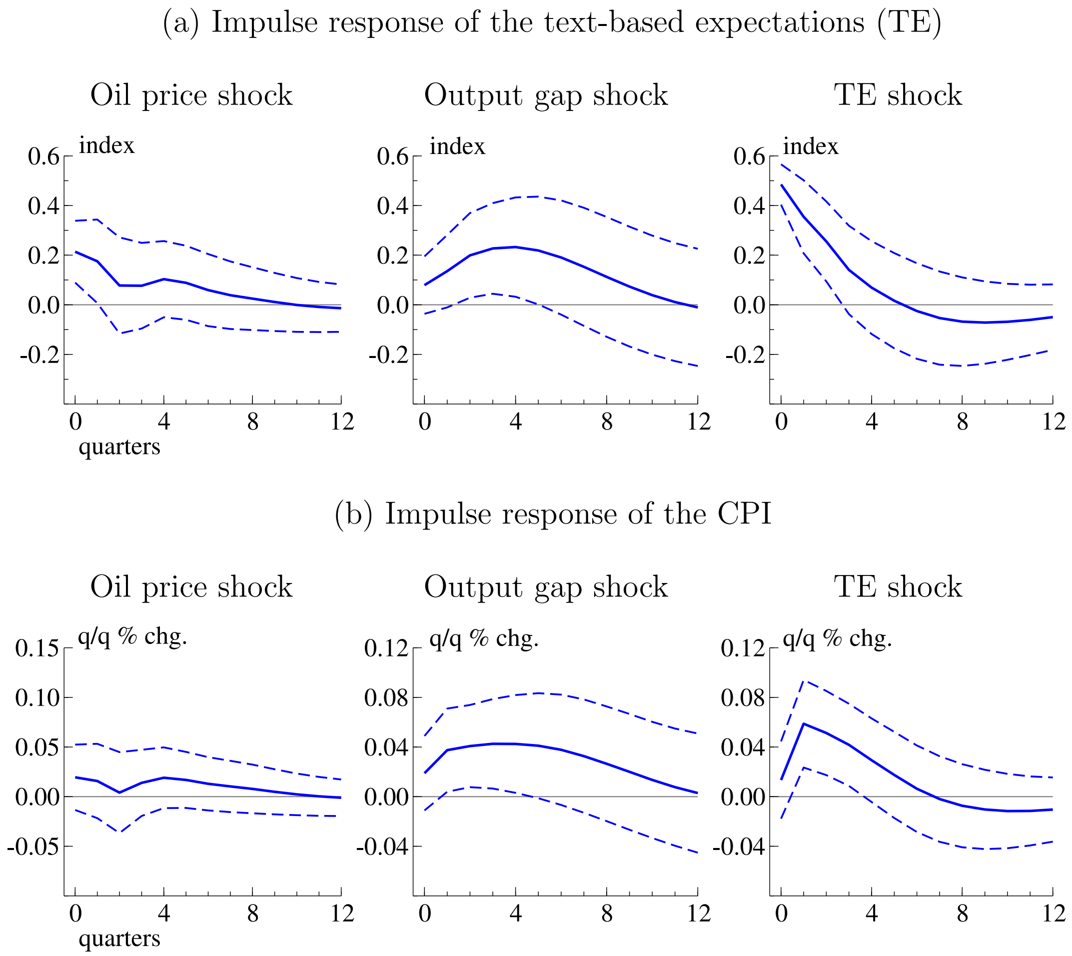


Figure 7: Impulse responses from VAR model with the oil price. The size of the shock is one standard deviation. The dashed lines indicate the 95 percent confidence intervals.

quarter. It should be noted that even if we change the order of the text-based expectations and the CPI by assuming that a CPI shock in this ordering may affect the text-based expectations during the same quarter, we obtain qualitatively the same impulse responses of the variables as presented below.

Figure 6 plots the impulse responses of the VAR model. Figure 6(a) indicates that the responses of the text-based expectations to an exchange rate shock and an output gap shock are statistically significant. The text-based expectations reacts to an exchange rate shock almost contemporaneously and to an output gap shock with a lag of about 3–4 quarters. This implies that the text-based expectations is closely related to the macroeconomic variables which affect the inflation rate.

Turning to Figure 6(b), we further find that the response of the inflation rate to a text-based expectations-specific shock is statistically significant. It is noteworthy that the inflation rate reacts to a text-based expectations-specific shock with a lag of about 1–2 quarters, indicating that the text-based expectations tend to lead the inflation rate. These results suggest that the text-based expectations contain unique information regarding future changes in the inflation rate not captured by the exchange rate and the output gap.

To check the robustness of the impulse response, we also estimate the VAR model with the exchange rate replaced by the oil price, which is another important variable which affects the consumer price. We use the quarter-on-quarter change in the Dubai crude oil spot price. Figure 7 plots estimated impulse responses from the VAR model. The impulse response of the text-based expectations to an oil price shock is statistically significant, which indicates that the text-based expectations include information about the oil price development. For the rest of the impulse responses, the result remains the same qualitatively as the VAR with the exchange rate.

Table 4 reports a result of variance decomposition from the VAR models. For the VAR with the exchange rate, about 15 percent and 35 percent of the variance of the text-based expectations are attributable to the exchange rate and the output gap, respectively, at the 2-year horizon. In the VAR model with the oil price, about 10 percent of the text-based expectations' variation is explained by the oil price. In the variance of the CPI, it is notable that about 10–20 percent is explained by the text-based expectations-specific variation.

Turning back to the regression model, we now test the predictive power of the text-based expectations for the inflation rate in terms of out-of-sample forecasting. Specifically, we conduct one-quarter to four-quarter-ahead forecasting of the inflation rate for each quarter from 2012/Q1 to 2019/Q4. We start the test by estimating the regression equation using the data for the period through the 2011/Q4 and then predict the inflation rate for 2012/Q1–Q4. Next, we estimate the regression equation again using the data for the period through 2012/Q1 and then forecast the inflation rate for 2012/Q2–2013/Q1. By repeating this out-of-sample forecasting for each quarter, we obtain the predicted inflation rates for the period through 2019/Q4. We measure the accuracy of these out-of-sample forecasts by calculating the root-mean-squared error (RMSE) between the forecasts and

Table 4: Variance decomposition based on the VAR models including the text-based expectations (TE). The figures are in percent.

(a) VAR with exchange rate

Horizon	Exchange rate	Output gap	TE specific	CPI inflation
Decomposition of text-based expectations				
1 year	19.2	24.6	47.0	9.2
2 years	15.5	34.4	38.0	12.1
3 years	15.4	34.6	38.0	12.1
Decomposition of CPI				
1 year	13.6	18.0	14.7	53.6
2 years	11.3	29.1	12.2	47.3
3 years	11.1	29.9	12.5	46.5

(b) VAR with oil price

Horizon	Oil price	Output gap	TE specific	CPI inflation
Decomposition of text-based expectations				
1 year	12.6	21.6	57.4	8.3
2 years	11.6	29.9	47.3	11.3
3 years	11.3	29.8	47.6	11.2
Decomposition of CPI				
1 year	3.4	20.0	25.2	51.3
2 years	4.3	27.9	21.7	46.1
3 years	4.2	28.5	22.1	45.2

Table 5: Out-of-sample predictive performance of the text-based expectations (TE): root mean squared errors (RMSEs, in percent) in forecasting the CPI (the year-on-year rate of change, all items less fresh food and energy) at one to four quarters ahead. The forecasting period is 2012/Q1–2019/Q4.

Horizon (quarters)	1	2	3	4
Not including TE	0.184	0.277	0.318	0.401
Including TE	0.160	0.216	0.280	0.383

the actual inflation rates.

Table 5 shows the results of the out-of-sample forecasting exercise, which indicate that the RMSE of the specification including the text-based expectations is about 10 percent smaller than that of the specification without the text-based expectations for one-quarter horizon, and that this gain reaches about 20 percent for the two-quarter-ahead forecast. The gain declines to about 5 percent for four-quarter horizon. This result indicates that the text-based expectations have the unique information in forecasting at one-quarter to three-quarter horizons.

In sum, the analyses reveal that the text-based expectations provide additional information on changes in consumer prices over the next several months not captured by such macroeconomic variables as exchange rates and the output gap. Computed from comments by respondents to the survey, the text-based expectations appear to be a useful proxy for firms' short-term inflation expectations.

5 Concluding remarks

This paper proposes a quantitative indicator of firms' inflation expectations computed from comments provided by respondents to the firms' survey. Our analyses suggest that the text-based expectations comove with both demand and cost factors and also includes unique information for forecasting CPI inflation. There are several alternative approaches for analyzing textual data. For instance, unsupervised learning methods such as topic models and multinomial logistic regression may apply to the EWS data. A comparison or an extension of our method with a different approach is left for future work.

References

Andrade, P., O. Coibion, E. Gautier, and Y. Gorodnichenko (2022). No firm is an island? How industry conditions shape firms' expectations. *Journal of Monetary Economics 125*, 40–56.

Angelico, C., J. Marcucci, M. Miccoli, and F. Quarta (2021). Can we measure inflation expectations using Twitter? Bank of Italy Working Papers, No. 1318.

Barbaglia, L., S. Consoli, and S. Manzan (2021). Forecasting GDP in Europe with textual data. Working Paper, SSRN 3898680.

Barbaglia, L., S. Consoli, and S. Manzan (2022). Forecasting with economic news. *Journal of Business and Economic Statistics*, forthcoming.

Bragoli, D. (2017). Now-casting the Japanese economy. *International Journal of Forecasting 33*(2), 390–402.

Bryan, M., B. Meyer, and N. Parker (2014). The inflation expectations of firms: What do they look like, are they accurate, and do they matter? Federal Reserve Bank of Atlanta Working Paper, No. 2014-27.

Bybee, L., B. T. Kelly, A. Manela, and D. Xiu (2020). The structure of economic news. NBER Working Paper Series, No. 26648.

Coibion, O., Y. Gorodnichenko, and R. Kamdar (2018). The formation of expectations, inflation, and the Phillips curve. *Journal of Economic Literature 56*(4), 1447–1491.

Coibion, O., Y. Gorodnichenko, and S. Kumar (2018). How do firms form their expectations? New survey evidence. *American Economic Review 108*(9), 2671–2713.

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

Coibion, O., Y. Gorodnichenko, S. Kumar, and J. Ryngaert (2021). Do you know that I know that you know...? Higher-order beliefs in survey data. *The Quarterly Journal of Economics 136*(3), 1387–1446.

Coibion, O., Y. Gorodnichenko, and T. Ropele (2019). Inflation expectations and firm decisions: New causal evidence. *The Quarterly Journal of Economics 135*(1), 165–219.

Goshima, K., H. Ishijima, M. Shintani, and H. Yamamoto (2021). Forecasting Japanese inflation with a news-based leading indicator of economic activities. *Studies in Nonlinear Dynamics & Econometrics 25*(4), 111–113.

Guzman, G. (2011). Internet search behavior as an economic forecasting tool: The case of inflation expectations. *Journal of Economic and Social Measurement 36*(3), 119–167.

Kalamara, E., A. Turrell, C. Redl, G. Kapetanios, and S. Kapadia (2020). Making text count: Economic forecasting using newspaper text. Staff Working Paper No. 865, Bank of England.

Ke, Z. T., B. T. Kelly, and D. Xiu (2019). Predicting returns with text data. NBER Working Paper Series, No. 26186.

Kumar, S., H. Afrouzi, O. Coibion, and Y. Gorodnichenko (2015). Inflation targeting does not anchor inflation expectations: Evidence from firms in New Zealand. *Brookings Papers on Economic Activity 2015*(2), 151–225.

Murphy, K. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.

Muto, T. (2015). Results of a survey on inflation outlook of firms in the Bank of Japan's "Short-term economic survey of enterprises in Japan". In *Indicators to support monetary and financial stability analysis: data sources and statistical methodologies*, Volume 39. Bank for International Settlements.

Otaka, K. and K. Kan (2018). Economic analysis using machine learning: Text mining of the Economy Watchers Survey. Bank of Japan Working Paper Series, No. 18-J-8 (in Japanese).

Seabold, S. and A. Coppola (2015). Nowcasting prices using Google trends: An application to Central America. Policy Research Working Paper, No. 7398, World Bank.

Shapiro, A. H., M. Sudhof, and D. J. Wilson (2022). Measuring news sentiment. *Journal of Econometrics 228*, 221–243.

Sharpe, S. A., N. R. Sinha, and C. Hollrah (2017). What's the story? A new perspective on the value of economic forecasts. Finance and Economics Discussion Series 2017-107. Washington: Board of Governors of the Federal Reserve System.

Thorsrud, L. A. (2020). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics 38*, 393–409.

Wei, Y., X. Zhang, and S. Wang (2017). Can search data help forecast inflation? Evidence from a 13-country panel. *Proceedings* 2017 IEEE International Conference on Big Data.