BANK OF CANADA
BANQUE DU CANADA

Staff Discussion Paper/Document d'analyse du personnel—2023-13

Last updated: June 29, 2023

What Can Earnings Calls Tell Us About the Output Gap and Inflation in Canada?

by Marc-André Gosselin and Temel Taskin

Canadian Economic Analysis Department
Bank of Canada
mgosselin@bankofcanada.ca, ttaskin@bankofcanada.ca

DOI: https://doi.org/10.34989/sdp-2023-13 | **ISSN** 1914-0568

Acknowledgements

We owe special thanks to Troy David for excellent research assistance. We would also like to thank Fabio Canova, Michael Ehrmann, Justin Damien Guinette, Christopher Hajzler, Tolga Ozden, Nicolas Petrosky-Nadeau, Adam Hale Shapiro and Alexander Ueberfeldt for their comments and feedback.

Abstract

We construct new indicators of the imbalance between demand and supply for the Canadian economy by using natural language processing techniques to analyze earnings calls of publicly listed firms. The results show that the text-based indicators are highly correlated with official inflation data and estimates of the output gap and improve the accuracy of inflation forecasts. This suggests that these indicators could help central banks foresee inflationary pressures in the economy. Our examination of other topics in earnings calls, such as supply chain disruptions and capacity constraints, points to the potential benefits of using textual data to quickly draw insights on a range of relevant topics. We conclude that text-based measures of economic slack should be included in central banks' monitoring and forecasting toolkits.

Topics: Central bank research; Domestic demand and components; Econometric and statistical methods; Inflation and prices; Potential output

JEL codes: C1, C3, E3, E5

Résumé

Nous élaborons de nouveaux indicateurs de déséquilibre entre la demande et l'offre dans l'économie canadienne au moyen de techniques de traitement automatique du langage naturel pour analyser les transcriptions de présentations des résultats financiers de sociétés cotées en bourse. Les résultats de l'étude montrent que les indicateurs fondés sur des données textuelles présentent une forte corrélation avec les données officielles sur l'inflation et les estimations de l'écart de production, et améliorent l'exactitude des prévisions d'inflation. Cela laisse entendre que ces indicateurs pourraient aider les banques centrales à prévoir les pressions inflationnistes dans l'économie. Notre examen d'autres sujets mentionnés lors des présentations des résultats financiers, tels que les perturbations des chaînes d'approvisionnement et les contraintes de capacité, indique des avantages possibles de l'utilisation de données textuelles pour tirer rapidement des conclusions sur divers sujets pertinents. Nous concluons que les mesures de capacités excédentaires basées sur des données textuelles devraient être ajoutées aux outils de suivi et de prévision des banques centrales.

Sujets : Demande intérieure et composantes; Inflation et prix; Méthodes économétriques et statistiques; Production potentielle; Recherches menées par les banques centrales
Codes JEL : C1, C3, E3, E5

1. Introduction

Monitoring aggregate demand and supply conditions in the economy is a key element in conducting monetary policy because it helps determine the degree of economic slack and associated inflationary pressures. Slack is typically measured by the output gap, which is the difference between the actual level of output and an estimated trend that captures the economy's productive capacity. This approach has important caveats since the estimated trend depends on the methodology used (e.g., filters, unobserved components, growth accounting) and is subject to significant revisions as new data come in (see Cheung, Frymire and Pichette 2021; Champagne, Poulin-Bellisle and Sekkel 2018; Marcellino and Musso 2011). These caveats are shown to be an important factor in evaluating and conducting monetary policy (Coibion, Gorodnichenko and Ulate 2018; Kozicki 2004). Moreover, standard measures of economic slack fall short of explaining movements in inflation since the start of the COVID-19 pandemic, when aggregate demand and supply have been subject to large shocks (Faucher et al. 2022; Ruch and Taskin 2022). This paper departs from standard methods to measure the output gap and presents an alternative indicator of economic slack using transcripts of earnings calls, which are available almost immediately after the calls take place. The proposed methodology provides more timely and direct measures of aggregate demand and supply conditions based on firm-level information. It also captures inflationary pressures since the start of the COVID-19 pandemic when conventional measures of slack have been largely uninformative about inflation.

Earnings calls are an important channel of communication between market participants and the management of publicly traded companies. They provide insights on companies' own outlooks and views around broader financial and economic developments. Therefore, public statements from corporate management teams could be used to draw insights on broad economic activity. In this paper, we extract information from earnings calls for Canadian companies that are publicly listed, covering the period between the first quarter of 2013 and the fourth quarter of 2022. To do so, we borrow methods from natural language processing (NLP).

Our methodology relies on identifying mentions of demand and supply in the transcripts of earnings calls and measuring sentiment around those concepts, similar to the approaches of Baker, Bloom and Davis (2016) and Hassan et al. (2020). This allows us to conduct a systematic quantitative analysis of a large volume of textual data that would be impossible to do by reading and interpreting each transcript.

We then assess the usefulness of these indicators in a forecasting environment. Results show that our text-based estimate of economic slack tends to be highly correlated with the Bank of

Canada's estimate of the output gap in before the COVID-19 pandemic. However, these measures stop moving in tandem beginning in 2020. Specifically, the Bank's measure of the output gap indicates a persistent state of excess supply since the start of the COVID-19 pandemic. In contrast, our novel estimate points to a sharp transition from excess supply to excess demand that coincides with the emergence of disruptions in global supply chains. This result is important because it shows that our text-based indicator can help provide more accurate forecasts of inflation. In addition, our estimate of slack is available to policy-makers in real time, whereas traditional measures of the output gap are available with a delay given the publication lag of two months for Canada's gross domestic product (GDP). Moreover, our examination of other topics in earnings calls, such as supply chain disruptions and capacity constraints, shows the potential benefits of using textual data to draw timely insights on a broad range of relevant topics. We conclude that text-based measures of economic slack should be included in central banks' toolkits for monitoring and forecasting inflation.

This paper contributes to two major branches of research. First, our newly constructed indicators of demand and supply draw from the expanding literature about applying NLP techniques to digital texts with economic content. For instance, Hassan et al. (2020) use transcripts of earnings calls to estimate the impact of Brexit on publicly listed firms in the United Kingdom and across the world. Baker, Bloom and Davis (2016) construct an index for political uncertainty using newspaper articles and present evidence on the impact on economic activity. Bybee et al. (2020) document the topics covered in economic news articles and show that real activity is significantly correlated with the weight of topics that reflect the state of the business cycle in the economy. Similarly, Manela and Moreira (2017) provide evidence of the link between economic disasters and news-based uncertainty using the textual information on front pages of *The Wall Street Journal* between 1890 and 2007. Shapiro, Sudhof and Wilson (2022) build a news-based measure of sentiment and show that it predicts movements in survey-based measures of consumer sentiment. Chen and Houle (2023) generate high-frequency and up-to-date indicators to monitor news media coverage of supply and labour shortages in Canada. Angelico et al. (2022) and Larsen, Thorsrud and Zhulanova (2021) use text-based methods to construct consumers' inflation expectations.[1] We contribute to this strand of literature by constructing a new proxy for the output gap by computing demand and supply sentiments using text from earning calls and then linking those measures to other output gap estimates and inflation data.

The second area of research that we contribute to is the empirical literature quantifying economic slack and its role in forecasting inflation. This literature extensively covers estimating

[1] See Gentzkow, Kelly and Taddy (2019) for a recent survey of research about using text as data.

the output gap and discusses its real-time properties (for more, see Orphanides and van Norden 2002; Marcellino and Musso 2011). A number of follow-up studies argue that conventional measures of the output gap are unreliable in real time because they commonly use trend estimation methods that are subject to large revisions at the end of the sample (see Cayen and van Norden 2005; Kamada 2005; Cusinato, Minella and da Silva Pôrto Júnior 2013, among others). We contribute to this literature by evaluating how well our proxy for the output gap captures inflationary pressures. Our method uses direct discussions of demand and supply and doesn't depend on estimating unobserved trends in actual output data, making our approach robust to the end-point bias commonly cited with traditional techniques. Moreover, earnings calls cover the recent past, current state and near future of a company and the broader economy, providing a better picture of short-term demand and supply conditions than the trend estimation, which likely overlooks short-term supply-side developments.

The rest of the paper is organized as follows. Section 2 describes the textual dataset, elaborates on our proxy for the output gap and compares it with the relevant data for the Canadian economy. Section 3 reports the use of the textual indicator in inflation forecasts and evaluates its performance. Section 4 concludes.

2. Constructing demand, supply and output gap indicators using textual data

In this section we present our textual dataset. We also describe the indicators of demand and supply that we extract and compare them with the relevant Canadian economic data.

2.1. Earnings calls data

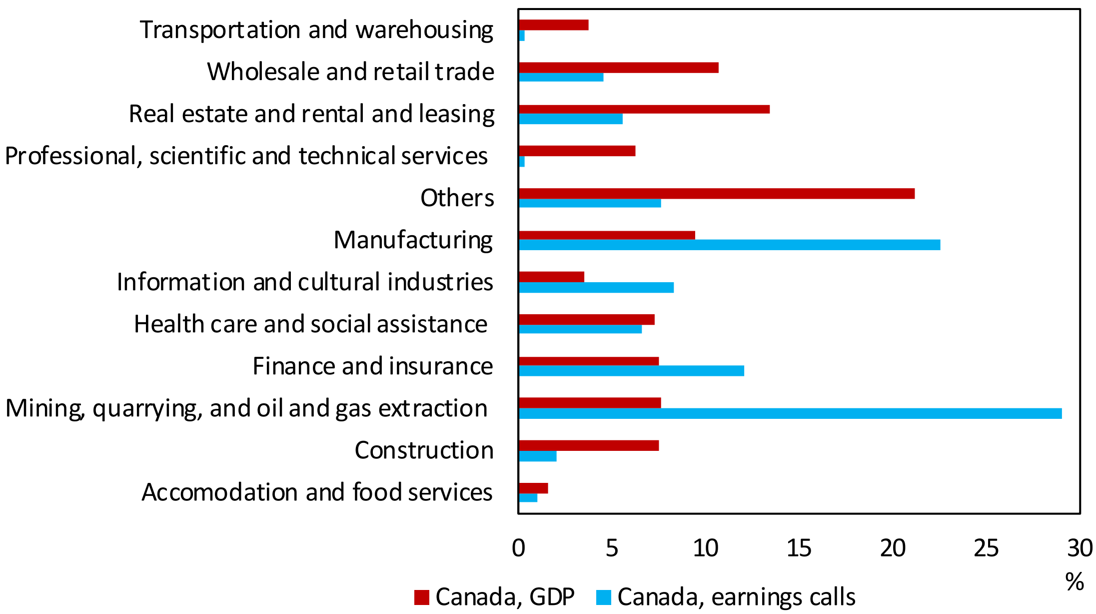
An earnings call is a quarterly conference between the executives of a public company, market participants and the media to communicate the company's financial performance and discuss broader, yet relevant, economic and financial developments. We use transcripts of these calls to assess their value when it comes to monitoring the macroeconomy. We focus on companies that are headquartered in Canada since we are interested in Canadian economic conditions.[2]

We obtained transcripts of earning calls from Seeking Alpha Ltd, a financial data provider. We focus on the period starting in 2013 because the number of observations per quarter for Canadian companies are usually too small in earlier periods. Our dataset contains 9,221 transcripts for companies headquartered in Canada between the first quarter of 2013 and the

[2] Companies listed in the US stock market conduct earnings calls, but these calls cover a broad range of countries identified by the location of their headquarters. One can use other methods to identify Canada-relevant discussions within the entire set of earnings calls, for instance, by extracting the text around mentions of Canada.

fourth quarter of 2022.[3] **Chart 1** provides a sectoral breakdown of firms in the transcript data compared with value-added shares of GDP, by industry, from Statistics Canada. We can see that industry shares tend to be similar, although the manufacturing and mining, quarrying and oil and gas sectors have a greater representation in the earnings call data.

Chart 1: Sectoral breakdown by transcript data and share of gross domestic product



Sources: Seeking Alpha Ltd., Statistics Canada and Bank of Canada Last observation: 2022Q2

Note: Sectors are identified by the North American Industry Classification System (NAICS). "Others" includes administrative and support; waste management and remediation services; agriculture, forestry, fishing and hunting; arts, entertainment and recreation; educational services; management of companies and enterprises; public administration, utilities; and other services.

We clean the raw data using NLP techniques to make the information suitable for quantitative analysis. NLP is a field of machine learning that focuses on the interaction between computers and human language. The methods developed in this field aim to improve computers' ability to quantify textual data. The standard steps involved in preparing text for quantitative analysis include tokenization and the removal of stop words (Gentzkow, Kelly and Taddy 2019).[4] Tokenizing splits sentences into individual words, known as tokens, based on text delimiters such as spaces and commas. This converts text to a machine-readable format and is an

[3] The complete dataset contains about 190,000 earnings calls starting from 2006 for companies headquartered in 80 different countries. Examples of text from earnings calls are available **in Table A-1** in the Appendix.

[4] Stop words are frequently used functional words such as "the," "are," "and," "or," "in" and "as." We use a common NLP library of stop words to identify these words in the text. For further details, see the Python Software Foundation.

important step in preparing data to be input into models. Finally, we remove stop words from the tokenized text. The cleaning process is important not only to reduce the total number of unique terms in each document but also to arrive at a word count for each document that is comparable across earnings calls.

2.2. Quantifying demand and supply conditions

In this section, we describe the text mining techniques we use to construct our measures of demand and supply sentiment and then present the models we use to assess the usefulness of these indicators in making inflation forecasts.

We follow Hassan et al. (2020) in measuring sentiment variables in the preprocessed transcripts of earnings calls. Demand or supply sentiment on a given call is obtained by aggregating sentiment scores around every mention of each word. Sentiment is measured by the frequency of positive-toned terms minus negative-toned terms within the r-terms range of the mention, divided by the total number of words on a given call. More specifically, the sentiment score is calculated as follows:

$$Sentiment_{it}^{D,S} = \frac{1}{|B_{it}|} \sum_{b \in B_{it}} \{\mathbf{1}^{\boldsymbol{D,S}}(b) \times [\sum_{c \in C^r(b)} S(c)]\}, \quad (1)$$

where B_{it} denotes the entire list of words in the call of firm i at time t, and $\mathbf{1}^{\boldsymbol{D,S}}(.)$ is an indicator function that takes the value of 1 if the input word is in the demand (supply) word list and 0 otherwise. $C^r(b)$ denotes the set of words in the r-terms range of word b (before and after), and the function $S(.)$ is defined as follows:[5]

$$S(c) = \begin{cases} +1 \; if \; c \in S^+ \\ -1 \; if \; c \in S^- \\ 0 \; otherwise, \end{cases}$$

in which S^+ and S^- represent the lists of positive- and negative-toned words, respectively.[6]

[5] Here, the r-terms range refers to the list of r-words before and after the mention of a specific term. We set r equal to 10 in the baseline calculations. Repeating the same exercise by setting r equal to 20 or 30 results in broadly similar results.

[6] Some examples of positive-toned words: good, strong, benefit, improved and favourable. Examples of negative-toned words: concerns, challenges, weak, disruption and adverse. Examples are from the Loughran and McDonald (2016) dictionary.

The positive- and negative-toned words are identified using the Loughran and McDonald (2016) sentiment dictionary. This dictionary contains economics- and finance-related sentiment texts that allows us to identify the most relevant words for our purposes.

We standardize the raw sentiment series for demand and supply (by subtracting the sample average and dividing by the sample standard deviation) to allow for comparisons. We measure demand/supply imbalance by calculating the difference between these standardized sentiment scores of demand and supply.

2.3. Benchmarking textual indicators against hard data

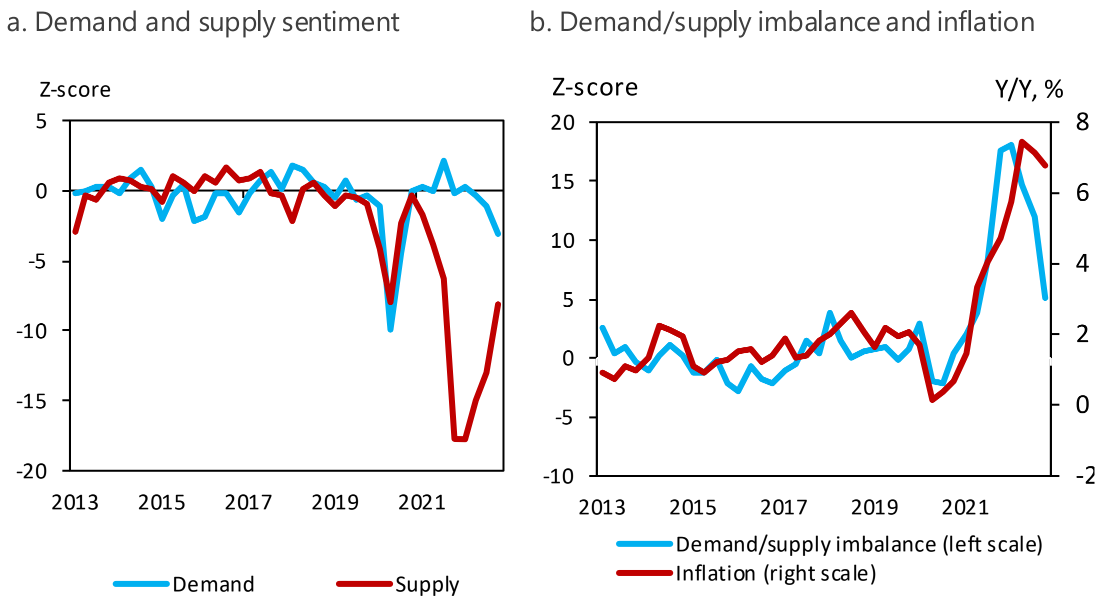
In this section, we describe the constructed indicators and compare them with relevant data for the Canadian economy.

Chart 2 plots the evolution of our sentiment indicators of demand and supply over the sample period. While both series show relatively small variations over 2013–19, large swings start in 2020. These movements capture the commonly reported pandemic narrative: the contemporaneous collapse and rebound of demand and supply in 2020. Following this cycle was a sharp decline in supply sentiment associated with disruptions in global supply chains and increasing shortages of labour, which took place against the backdrop of resilient consumer demand for goods in 2021. Supply improved and demand conditions eased in 2022.[7]

Taking the difference between these two series yields our indicator of the demand/supply imbalance (DSI). Movements in our measure of demand and supply are broadly consistent with the evolution of inflation over the sample period (**Chart 2**, panel b). The decline in inflation during the initial phase of the pandemic and the sharp increase over the past two years are well reflected in the DSI series. The peak correlation between DSI and inflation is .90 and occurs with a one-quarter lag. Therefore, the fall in our indicator of the demand-supply imbalance at the end of the sample indicates an impending period of disinflation.

[7] The increased frequency of discussions in earning calls around supply chain disruptions and capacity constraints corroborate these narratives. **Chart A-1** and **Chart A-2**, for instance, show that frequency of mentions in earnings calls of supply disruptions and capacity constraint were roughly 30 and 7 standard deviations higher than their long-term averages, respectively. Results on other indicators are also available in the **Appendix**.

Chart 2: Demand and supply sentiments are indicative of inflation



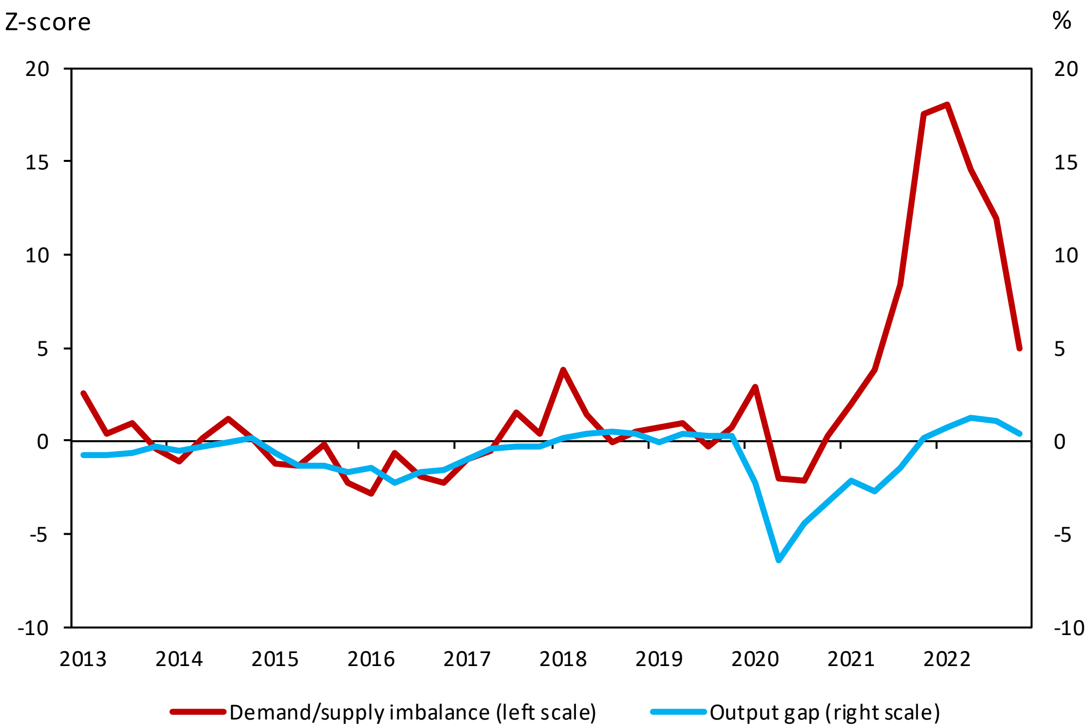
Sources: Seeking Alpha Ltd. and Bank of Canada

Last observation: 2022Q4

Chart 3 compares our indicator of the imbalance between demand and supply sentiment with the final estimate of the output gap from the Bank's *Monetary Policy Report* (MPR).[8] The two series move in tandem between 2013 and 2019, but a divergence after the pandemic started, pointing to different implications for inflation dynamics.

[8] **Chart 3** plots our demand/supply imbalance against the final output gap in the MPR for illustrative purposes, yet we use both final and real-time output gap series in the forecast exercises. Final output gap refers to the latest available vintage whereas the real-time output gap refers to the historical vintages. Accordingly, the real-time series incorporates information up to the corresponding vintage date. For details, see "Indicators of capacity and inflation pressures for Canada" on the Bank's website.

Chart 3: Demand/supply imbalance and the output gap move in tandem until the pandemic



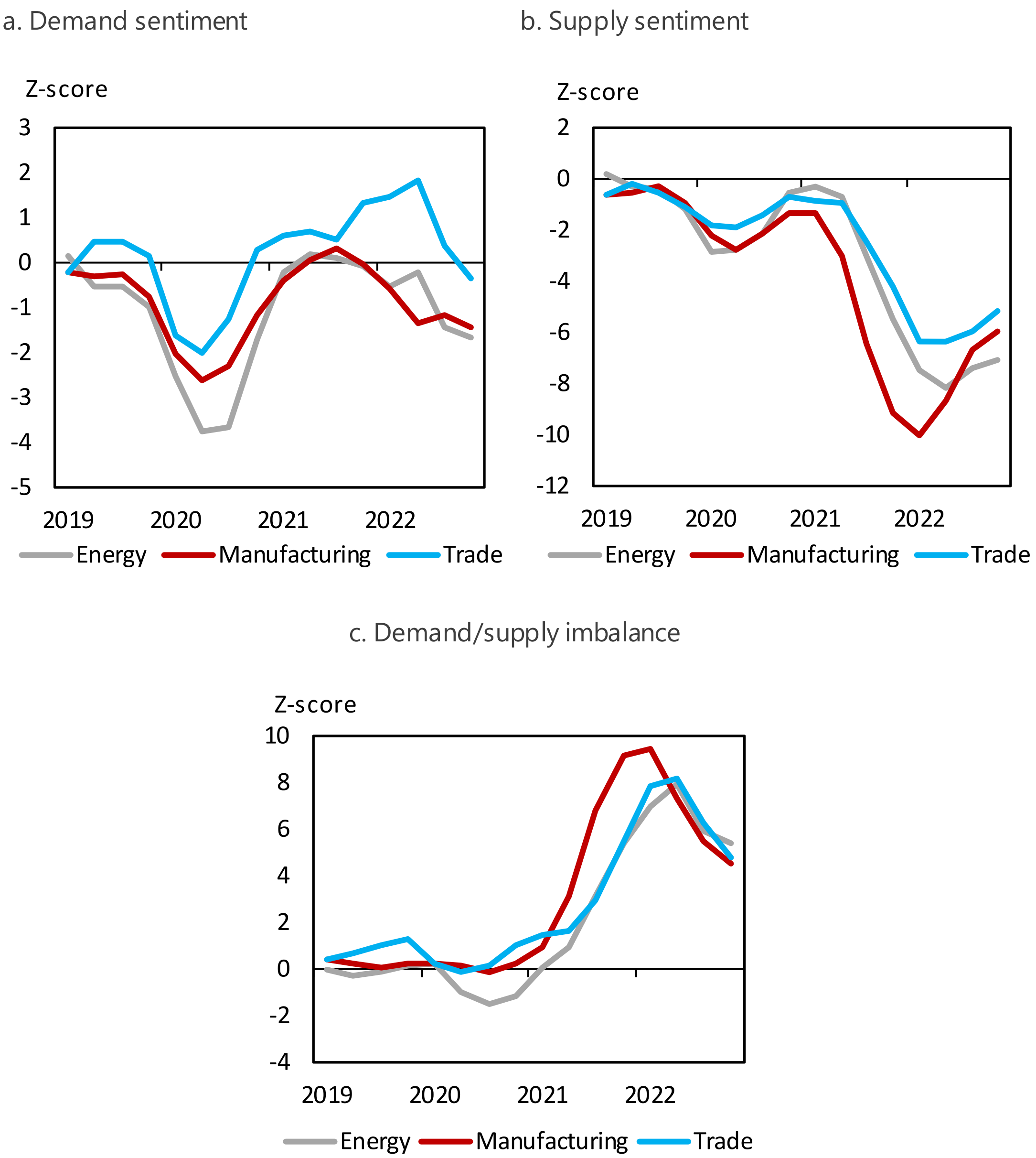
Sources: Seeking Alpha Ltd. and Bank of Canada

Last observation: 2022Q4

2.4. Sector-level indicators

To provide more details about our results, we assess how the imbalance between demand and supply manifested in selected sectors that represent relatively larger shares of our sample over the past three years. Since sectoral series tend to be more volatile, we focus on moving averages for illustrative purposes. A sectoral breakdown of sentiment scores illustrates a broad-based decline in demand sentiment during the initial phase of the pandemic and a wide-ranging decline in supply sentiment afterward (**Chart 4**).

Chart 4: Movements in demand/supply imbalance during the pandemic are broad-based



Sources: Seeking Alpha Ltd. and Bank of Canada Last observation: 2022Q4

Note: Trade includes wholesale trade and retail trade. Energy includes mining, quarrying, and oil and gas extraction.

In the next section, we explain how we use the textual indicators in inflation forecasts.

3. Using textual indicators in inflation forecasts

This section outlines the formal approach we take, given the promising descriptive results presented in the previous section, to test the usefulness of information in earnings calls for predicting inflation. We extend commonly used inflation forecast models with our textual indicators. Then, we compare the root mean squared errors (RMSEs) of out-of-sample forecasts of various models to assess the marginal contribution of the information in earnings calls.

Using textual information in forecasts is promising not only because of the strong correlation between the output gap and inflation shown in the previous section but also because the information in the earnings calls is timely. The data needed to estimate the output gap are usually available with substantial lags, unlike the information in earnings calls, which is immediately available. The information in earnings calls can therefore improve our assessment of the current conditions of supply and demand, and therefore the output gap.

Our estimation and forecast periods cover the first quarter of 2013 to the fourth quarter of 2019 and the first quarter of 2020 to the first quarter of 2023, respectively. Given that historical text-based data are limited, we divide the sample period as before and after the start of the COVID-19 pandemic to assess the usefulness of our indicators during the recent unusual times. More specifically, we estimate the models between the first quarter of 2013—the earliest available data point for sentiment series—and the fourth quarter of 2019 and generate a forecast for the first quarter of 2020. We repeat the assessment by rolling the forecast window forward.

The next section discusses the empirical and forecasting results based on our text-based demand and supply indicators for the Canadian economy.

3.1. Forecast models

We compare the following models to assess how much our text-based demand and supply indicators improve the accuracy of inflation forecasts:

- **A benchmark univariate autoregressive *(AR)* model**: $\pi_t = \rho_0 + \sum_{i=1}^{k} \rho_i \pi_{t-i} + \epsilon_t$. This model is used as the benchmark case when comparing the performance of alternative models. AR models generate out-of-sample forecasts for inflation that are hard to improve upon (see Faust and Wright 2009; Champagne, Poulin-Bellisle and Sekkel 2018).[9]

[9] Given the small sample we have due to the availability of textual data, our benchmark case focuses on AR(1) processes where $k = 1$.

- **A generic Phillips curve model with real-time output gap ($\boldsymbol{PC^{RT}}$)**: $\pi_t = \rho_0 + \sum_{i=1}^{k} \rho_i \pi_{t-i} + \gamma x_t^{RT} + \epsilon_t$. The variable x_t^{RT} represents real-time estimates of the output gap as found in the Bank's MPR. We then use an AR(1) process for the output gap for the out-of-sample dynamic simulations that generate the forecasts for π_{t+h}, where h is the horizon of the forecast in quarters.

- **Previous model with final output gap ($\boldsymbol{PC^F}$)**: $\pi_t = \rho_0 + \sum_{i=1}^{k} \rho_i \pi_{t-i} + \gamma x_t^F + \epsilon_t$, where x_t^F is the final estimate of the output gap estimate in the MPR. This final estimate is then used for the out-of-sample dynamic simulations that generate the forecasts for π_{t+h}, where again h is the horizon of the forecast in quarters.[10]

- **A model with textual output gap indicator ($\boldsymbol{DS}$)**: $\pi_t = \rho_0 + \sum_{i=1}^{k} \rho_i \pi_{t-i} + \gamma z_t + \epsilon_t$. The variable z_t represents the demand and supply proxy of the output gap from the transcripts of earnings calls that we construct in this paper. We then use an AR(1) process for the proxy of the imbalance between demand and supply for the out-of-sample dynamic simulations that generate the forecasts for π_{t+h}, where h is the horizon of the forecast in quarters.

3.2. Forecast performance

Table 1 shows the performance results from the baseline forecast. The RMSEs of the three alternative models for the nowcast of the current quarter (T0) and for the following one-, two-, three- and four-quarters-ahead horizons are reported relative to the benchmark model ($\boldsymbol{AR}$). Several interesting results emerge from the forecast exercises. The model with our demand-supply indicator ($\boldsymbol{DS}$) outperforms the baseline model and the two reference Phillips curve models. This result provides more evidence of the usefulness of the information in earnings calls beyond the descriptive correlations presented in the section 2. The RMSE improvements are larger in the short-term forecasts but remain substantial over the entire forecast horizon.

Another interesting result is that the model with the final output gap ($\boldsymbol{PC^F}$) performs better than the model with the real-time output gap ($\boldsymbol{PC^{RT}}$), in line with results covering earlier periods and various countries in the literature (e.g., Orphanides and van Norden 2002). The $\boldsymbol{PC^{RT}}$ model underperforms the AR model, confirming the earlier results from Champagne, Poulin-Bellisle and Sekkel (2018) for Canada, and Faust and Wright (2013) for the United States. Comparing $\boldsymbol{PC^F}$ with the $\boldsymbol{AR}$ returns mixed results, with these models performing better than the other over different time horizons, yet with no statistical significance.

[10] See footnote 3 for a brief description of real-time and final output gap estimates.

Table 1: Inflation forecasts, Y/Y percentage change

RMSE relative to $\boldsymbol{AR}$	(1) $\boldsymbol{PC^{RT}}$	(2) $\boldsymbol{PC^{F}}$	(3) $\boldsymbol{DS}$
T0	1.117	1.077	0.777**
T1	1.155	1.068	0.718
T2	1.143**	1.026	0.681*
T3	1.134***	0.992	0.962
T4	1.132***	0.962	0.934

Note: *, * and *** represent statistical significance at the 10%, 5% and 1% levels, respectively. Statistical significance is determined using the Diebold and Mariano (1995) test. Columns (1), (2) and (3) show root mean squared errors (RMSEs) of the corresponding model forecasts relative to the baseline model. $\boldsymbol{AR}$: autoregressive. $\boldsymbol{PC^{RT}}$: Phillips curve with the official real-time output gap. $\boldsymbol{PC^{F}}$: Phillips curve with the official final output gap. $\boldsymbol{DS}$: model with demand and supply imbalance indicator. Official output gap series reflect figures in the Bank of Canada's *Monetary Policy Report*. See text in section 3.2 for more details.

3.3. Robustness

As described above, we use an AR(1) process in the benchmark case for the out-of-sample dynamic simulations of $x_t^{RT}$and z_t over the forecast horizons. Moreover, in the benchmark case, we assume quarterly inflation at time *t* is unknown and nowcasted. We assess the robustness of our results in two ways. First, we use alternative models to simulate dynamic out-of-sample values for $x_t^{RT}$and z_t. Second, we relax the assumption that quarterly information is unknown at quarter *t* and adopt alternative information structures.

The following are the alternative models we use to generate out-of-sample dynamic simulations:

- **Random walk** ($\boldsymbol{RW}$): This model simply takes x_{t-1} as the forecast for the x_{t+h}, where h represents the forecast horizon.
- **An alternative AR model for each forecast horizon** ($\boldsymbol{ALT}$): $x_t = \rho_0^h + \rho_1^h x_{t-h} + \epsilon_t^h$, for each forecast horizon h. This model uses the most recent inflation value in all forecast horizons with different coefficients estimated in the equations described above.[11]

Next, we relax the assumption about the available information in quarter t.[12] The benchmark case assumes that π_t was unknown during quarter *t* and nowcasted using information up to

[11] See Gosselin and Tkacz (2001) for a detailed discussion of this model.

[12] We thank Fabio Canova for suggesting these alternative information structures in the forecast exercises.

$t-1$. In fact, at the end of quarter t, inflation figures for the first two months of the quarter are observed. In the following alternatives, we use this information and proceed accordingly with the forecasts:

- Shift the monthly consumer price index (CPI) series back one period ($\boldsymbol{SHIFT}$): $P_t^{SHIFT} = (\frac{P_{m1}(t)+P_{m2}(t)+P_{m3}(t-1)}{3})$. In this case, actual inflation at quarter t is assumed to be equal to the inflation rate calculated by the shifted price series, P_t^{SHIFT}.
- Omit the third month ($\boldsymbol{2M}$): $P_t^{2M} = (\frac{P_{m1}(t)+P_{m2}(t)}{2})$. In this case, actual inflation at quarter t is assumed to equal the average of the first two months of the quarter.
- Assume the price level is known and equals the actual value in quarter t ($\boldsymbol{KNOWN}$).

The nowcast is omitted in these alternative information structures because inflation at quarter t is assumed to equal the described alternative. Inflation is forecast for horizons 1 to 4 as described in the models ($\boldsymbol{PC^{RT}}$, $\boldsymbol{PC^{F}}$ and $\boldsymbol{DS}$).

Table 2 presents the results with alternative assumptions. The robustness exercises return broadly similar results compared with the ones under the baseline assumptions. If the variables on the right-hand side are treated consistently, the improvement in forecast accuracy holds with our newly constructed variable, exemplifying the usefulness of information from earnings calls.

Table 2: Robustness tests using alternative assumptions

RMSE, relative to $\boldsymbol{AR}$	Baseline $\boldsymbol{PC^{RT}}$ (1)	$\boldsymbol{PC^{F}}$ (2)	$\boldsymbol{DS}$ (3)	ALT $\boldsymbol{PC^{RT}}$ (4)	$\boldsymbol{DS}$ (5)	RW $\boldsymbol{PC^{RT}}$ (6)	$\boldsymbol{DS}$ (7)
Baseline							
T0	1.117	1.077	0.777**	1.117	0.777	1.117	0.777
T1	1.155	1.068	0.718	1.155	0.718**	1.167	0.615*
T2	1.143**	1.026	0.681*	1.133**	0.631	1.166**	0.510
T3	1.134***	0.992	0.962	1.106***	0.823	1.170***	0.652
T4	1.132***	0.962	0.934	1.075***	0.979	1.188***	0.812
$\boldsymbol{P_t^{SHIFT}}$							
T1	1.280*	1.184	0.782**	1.280*	0.782**	1.298*	0.673*
T2	1.218**	1.084	0.745	1.206**	0.704	1.246**	0.574
T3	1.174***	1.022	1.005	1.144***	0.869	1.213***	0.687
T4	1.146***	0.973	0.947*	1.088***	0.993	1.201***	0.825*
$\boldsymbol{P_t^{2M}}$							
T1	1.331*	1.229	0.784**	1.331*	0.784**	1.350*	0.666*
T2	1.244**	1.105	0.754	1.231**	0.709	1.272**	0.577
T3	1.186***	1.032	1.012	1.156***	0.873	1.226***	0.691
T4	1.151***	0.978*	0.950*	1.093***	0.996	1.207***	0.827*
$\boldsymbol{P_t^{KNOWN}}$							
T1	1.374*	1.262	0.805	1.374*	0.805**	1.395*	0.665*
T2	1.259**	1.115	0.769	1.246**	0.723	1.288**	0.582*
T3	1.192***	1.037	1.018	1.162***	0.877	1.232***	0.693
T4	1.155***	0.981*	0.952	1.097***	0.998	1.211***	0.829*

Note: *, * and *** represent statistical significance at the 10%, 5% and 1% levels, respectively. Statistical significance is determined using the Diebold and Mariano (1995) test. Columns (1), (2) and (3) show root mean squared errors (RMSEs) of the corresponding model forecasts relative to the baseline model. $\boldsymbol{AR}$: Autoregressive. $\boldsymbol{PC^{RT}}$: Phillips curve with the official real-time output gap. $\boldsymbol{PC^{F}}$: Phillips curve with the official final output gap. $\boldsymbol{DS}$: model with demand and supply imbalance indicator. Official output gap series reflect figures in the Bank of Canada's *Monetary Policy Report*. $\boldsymbol{ALT}$ and $\boldsymbol{RW}$ represent alternative processes for the out-of-sample dynamic simulations. $\boldsymbol{SHIFT}$, $\boldsymbol{2M}$ and $\boldsymbol{KNOWN}$ represent alternative assumptions about the information content at time *t*. See section 3.3 for more details.

3.4. Inflation shocks and the demand/supply imbalance

Finally, we examine the correlation between the DSI and the inflation shocks recovered from the Bank's two macroeconomic models: the Large Empirical and Semi-structural Model (Gervais and Gosselin 2014) and the Terms-of-trade Economic Model (Corrigan et al. 2021). In doing so, we regress the DSI on the models' shock series for the entire sample period of the first quarter of 2013 to the fourth quarter of 2022 to assess whether our series captures some of the variation in inflation that the Bank's macroeconomic models do not explain.

Table 3 reports the estimated coefficients of the DSI in the inflation shock regressions. It is statistically significant in both regressions, and the regressions capture substantial fractions of the variation in inflation shocks, pointing to the potential benefits of using our series in large-scale macroeconomic models used to forecast inflation.

Table 3: Demand and supply imbalance and inflation shocks

	LENS inflation shocks	TOTEM inflation shocks
Demand/supply imbalance	0.00019*	0.00184**
	(0.000)	(0.001)
Constant	-0.00027	-0.00007
	(0.000)	(0.002)
Observations	40	40
R-squared	0.14	0.38

Note: Robust standard errors are in parentheses. *, * and *** represent statistical significance at the 10%, 5% and 1% levels, respectively. LENS: Large Empirical and Semi-structural Model, TOTEM: Terms-of-trade Economic Model.

4. Conclusion

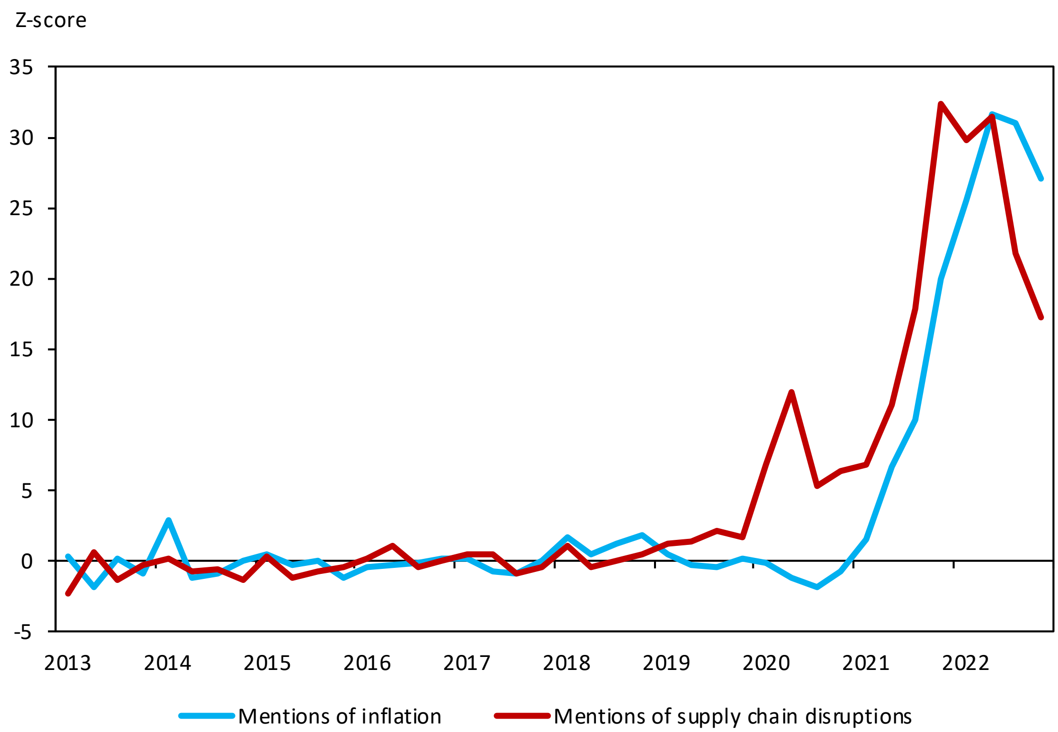
The output gap, a key concept for monetary policy, refers to the difference between the level of actual output and an estimate of the economy's potential output. Traditional methods of estimating potential output are subject to large uncertainty and considerable revisions as new data are received. This represents an important drawback for analyzing and forecasting the economy in real time (Champagne, Poulin-Bellisle and Sekkel 2018; Marcellino and Musso 2011). These measures also did not anticipate the pickup in inflation during the pandemic. This paper proposes an alternative approach to calculate the output gap by analyzing the transcripts of earnings calls. We find that our new measures improve forecasts of inflation during the pandemic.

Central banks could benefit from incorporating our text-based measures into their forecasting models. The findings indicate that the new measures could identify inflationary pressures in the economy because the text-based indicators of demand/supply imbalance are substantially linked with inflation data and official estimates of the output gap. Additionally, the accuracy of inflation forecasts improves when using these new variables. Moreover, our examination of earnings calls around other topics such as supply chain disruptions and capacity constraints points to the potential benefits of using textual data to draw timely insights on a range of relevant topics (**Charts A-1** to **A-4**). A more detailed analysis of these issues is left for future research.

Appendix

Transcripts of earning calls are obtained from Seeking Alpha Ltd., a financial data provider. The complete dataset contains about 190,000 earnings calls starting from 2006 for companies headquartered in 80 different countries. The dataset covers all major sectors. Earnings calls from Canada cover roughly 4% of the sample.

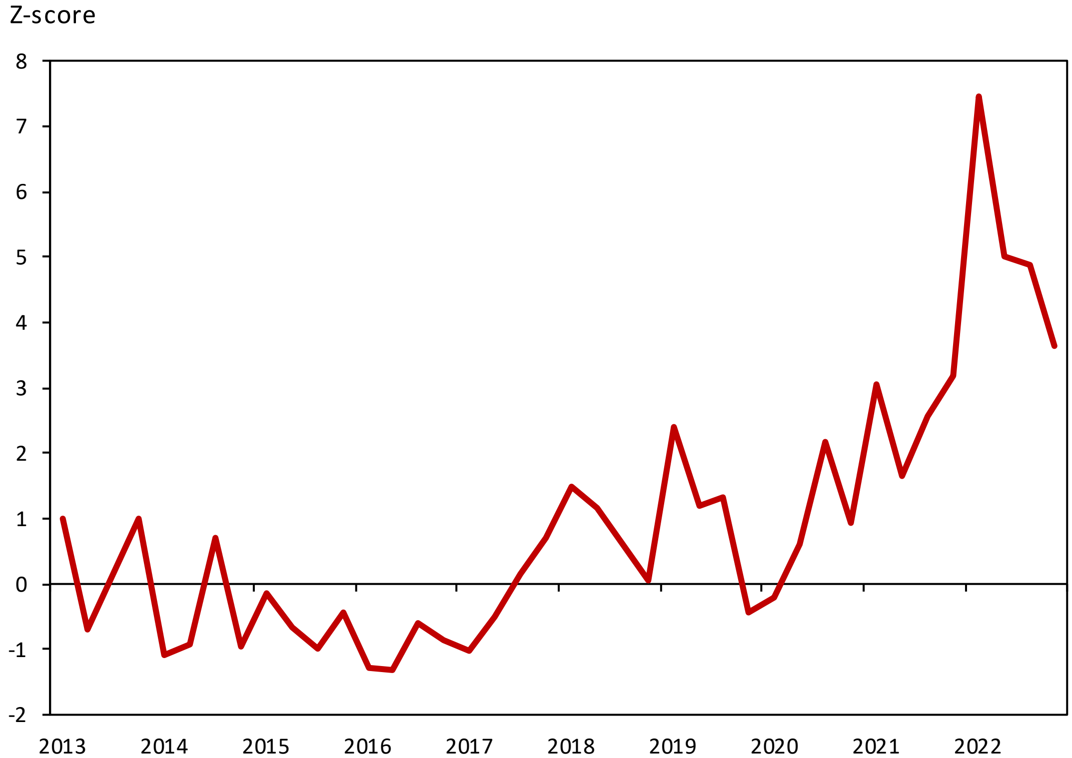
Chart A-1: Mentions of inflation and of supply chain disruptions, Canadian earnings calls



Sources: Seeking Alpha Ltd. and Bank of Canada Last observation: 2022Q4

Note: To calculate mentions per call, we identify the number of mentions of a given topic in a call by identifying the keywords related to that topic, add the number of mentions of each keyword and divide by the total number of words in the call. Aggregate time series are computed as a simple average across calls in each quarter. The following keyword lists are used as the baseline case in the chart: supply chain disruption: supply chain, supply disruption, supply bottleneck; inflation: inflation. The inflation keyword list is extended with CPI, and the supply chain disruption keyword list is extended with logistic constraint, logistic disruption, supply shortage, component shortage, chip shortage, semiconductor shortage, input shortage, container shortage, driver shortage, part shortage, product shortage, inventory shortage and ship shortage. The evolution of the time series remains broadly same.

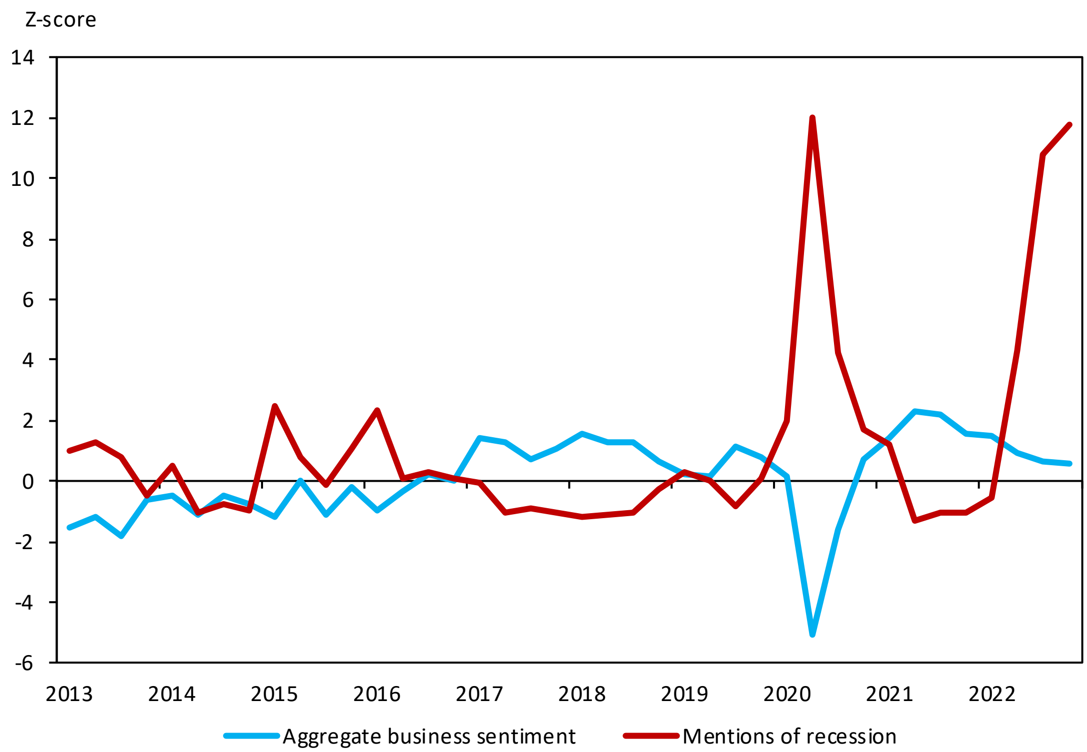
Chart A-2: Mentions of capacity constraints, Canadian earnings calls



Sources: Seeking Alpha Ltd. and Bank of Canada Last observation: 2022Q4

Note: To calculate mentions per call, we identify the number of mentions of a given topic in a call by identifying the keywords related to that topic, add the number of mentions of each keyword and divide by the total number of words in the call. Aggregate time series are computed as a simple average across calls in each quarter. The following keywords are used as the baseline case in the chart: capacity constraint and production constraint supply constraint.

Chart A-3: Aggregate business sentiment and mentions of a recession, Canadian earnings calls

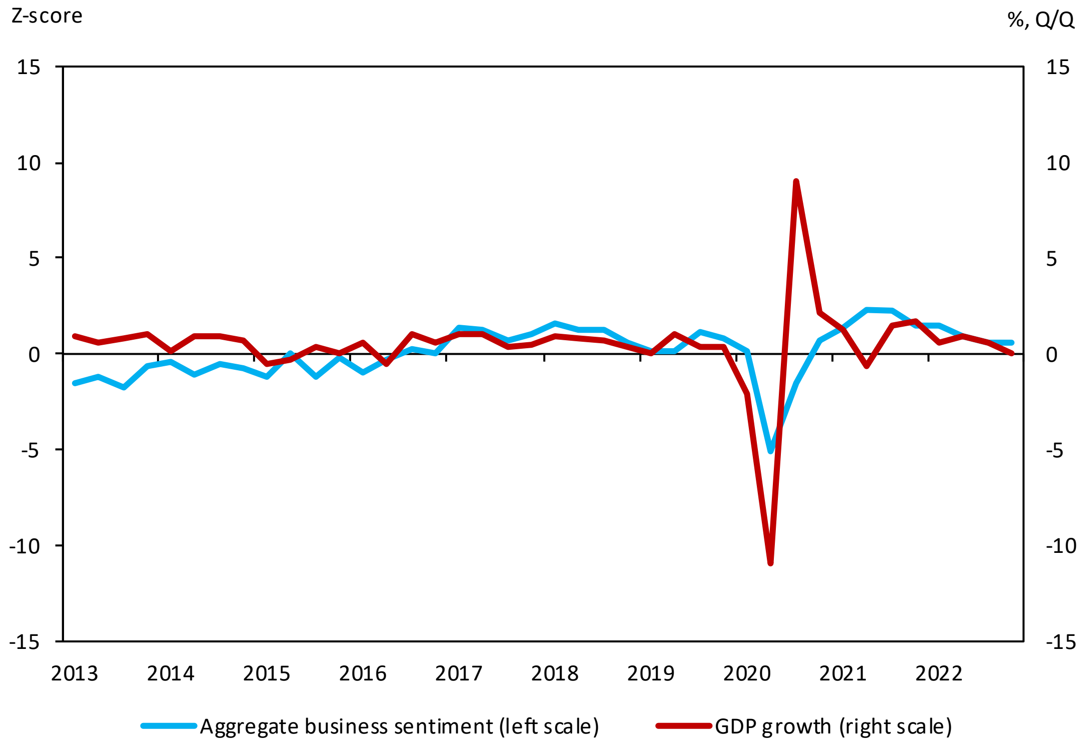


Sources: Seeking Alpha Ltd. and Bank of Canada

Last observation: 2022Q4

Note: To calculate mentions per call, we identify the number of mentions of a given topic in a call by identifying the keywords related to that topic, add the number of mentions of each keyword and divide by the total number of words in the call. Aggregate time series are computed as a simple average across calls in each quarter. The following keywords are used as the baseline case in the chart: recession: recession and recessionary. The recession keyword list is extended with economic uncertainty, economic crisis, economic downturn, financial crisis, weak economy, economic slowdown and economic turmoil. The evolution of the time series remains broadly same. To calculate aggregate sentiment, we identify the total number of positive and negative tone words—using the sentiment dictionary in Laughran and McDonalds (2016)—in a call and calculate the balance divided by the total number of words in the call. Aggregate time series are computed as a simple average across calls in each quarter.

Chart A-4: Aggregate business sentiment and growth in Canadian gross domestic product

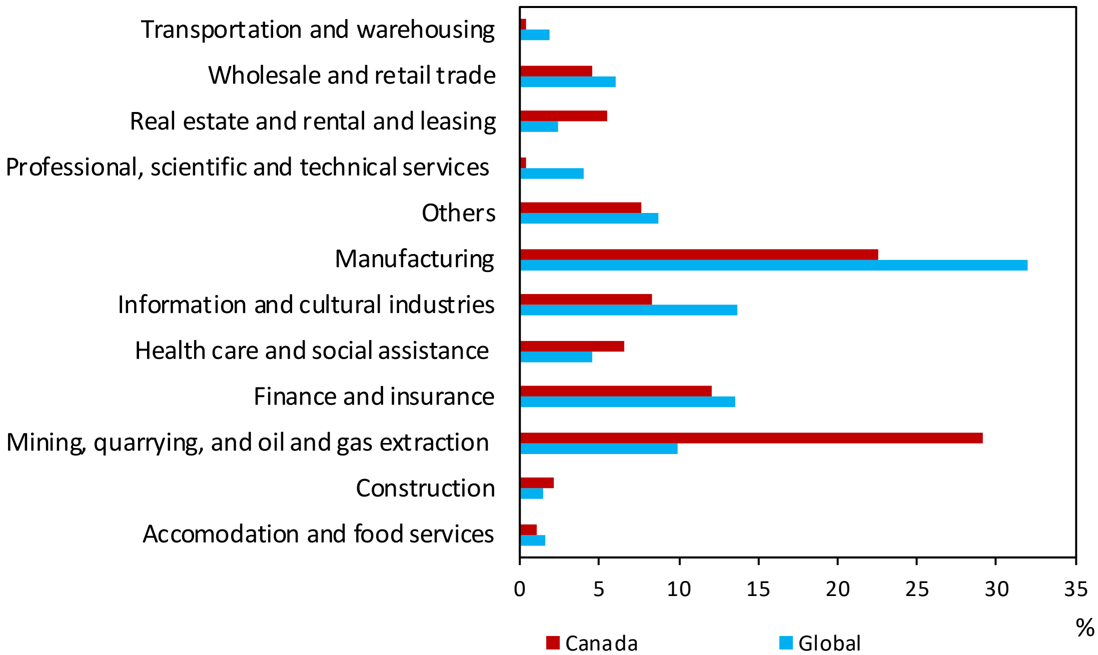


Sources: Seeking Alpha Ltd., Statistics Canada and Bank of Canada calculations

Last observation: 2022Q4

Note: To calculate aggregate sentiment in each call, we identify the total number of positive and negative tone words—using the sentiment dictionary in Laughran and McDonalds (2016)—and calculate the balance divided by the total number of words in the call. Aggregate time series are computed as a simple average across calls in each quarter.

Chart A-5: Sectoral breakdown in Canada and global samples of earnings calls

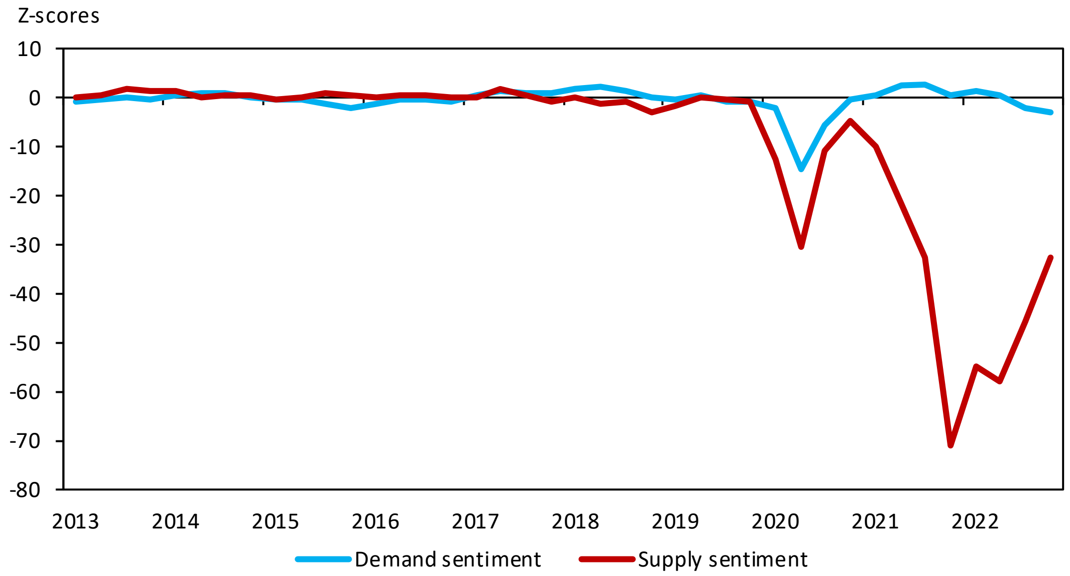


Sources: Seeking Alpha Ltd. and Bank of Canada

Last observation: 2022Q2

Note: Sectors are identified by the North American Industry Classification System (NAICS). "Others" includes administrative and support; waste management and remediation services; agriculture, forestry, fishing and hunting; arts, entertainment and recreation; educational services; management of companies and enterprises; public administration; utilities; and other services.

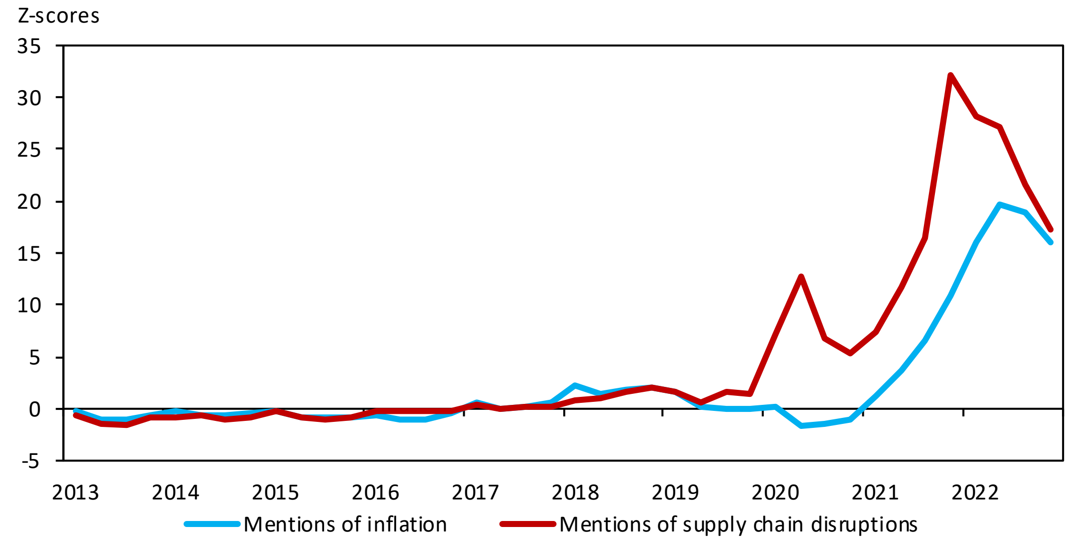
Chart A-6: Demand and supply sentiment in global earnings calls



Sources: Seeking Alpha Ltd. and Bank of Canada Last observation: 2022Q4

Note: See section 2 for the details about the methodology for calculating sentiment scores.

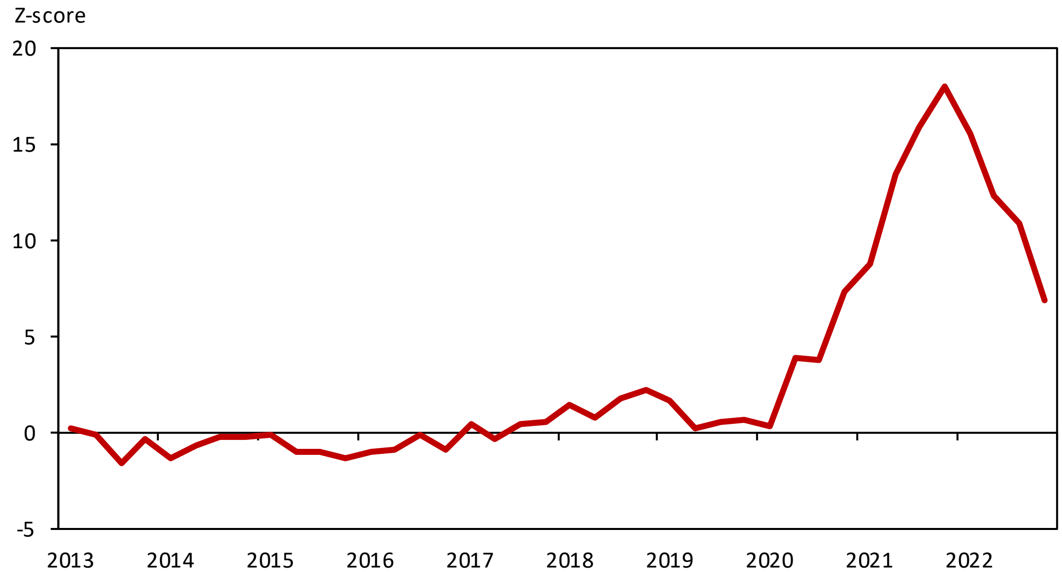
Chart A-7: Mentions of inflation and supply chain disruptions in global earnings calls



Sources: Seeking Alpha Ltd. and Bank of Canada Last observation: 2022Q4

Note: See the note to **Chart A-1** for the methodology.

Chart A-8: Mentions per call of capacity constraints in global earnings calls

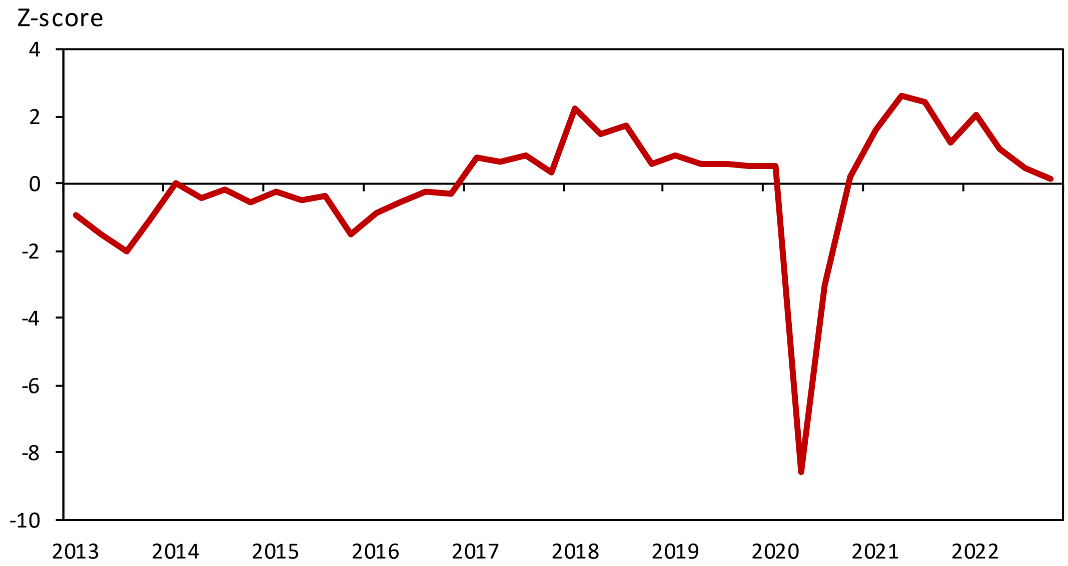


Sources: Seeking Alpha Ltd. and Bank of Canada

Last observation: 2022Q4

Note: See the note to **Chart A-2** for the methodology.

Chart A-9: Aggregate business sentiment in global earnings calls



Sources: Seeking Alpha Ltd. and Bank of Canada

Last observation: 2022Q4

Note: See the note to **Chart A-4** for the methodology.

Tables

This section presents sample excerpts from earnings call transcripts. Bolded terms reflect some of the keywords related to recently highlighted topics.

Table A-1: Sample excerpts from earnings call transcripts

Company	Date	Sample excerpt
Canadian Tire Corporation, Limited	**May 12, 2022**	"We continue to see **healthy demand** signals... There's no question we continue to operate in an environment where **inflation is real**, and global **supply chains** continue to be **challenged**." "As we look forward, here, we've had a late start to spring, but through what we're looking at, the **demand signals** are still pretty **strong**." "As for cargo, a **high demand** for cargo, especially in the Pacific market combined with our new freighter flying has led to a **strong** performance in this area."
Alimentation Couche-Tard Inc.	**June 29, 2022**	"This is stated by the impact of the **labour shortage** and a need to improve employee retention and increase of marketing initiatives and other discretionary expenses..." "...if we can get over time down that's as we faced in the industry face **staffing shortages** last year. We expect that to moderate—if labour trends continue." "We are pleased to report really a remarkable year...nonetheless the **pandemic**, a war in Europe, **supply chain**, **staffing challenges** and now inflation." "While **supply chain issues** have been a challenge in some items, our expanded supplier relationships and duplicate supply sourcing have enabled us to improve our in-stock positions versus prior quarters."
Bombardier Inc.	**August 6, 2020**	"Let me start this morning by recognizing at first how **difficult** the past quarter has been for all of us. I certainly did not expect my first quarter back to be so **challenging** with the **COVID-19 pandemic** affecting nearly every aspect of our operation, our end markets, and our financial performance." "We also took the **difficult** step of announcing a **significant workforce adjustment** as we needed to realign our production rate to the current **COVID impacted** market condition..."
Rogers Communications Inc	**April 22, 2020**	"We continue to see **positive demand** for our **internet offerings**, particularly as in-home bandwidth and reliability takes precedence in a work from home environment..." "In the last two weeks, we've been sort of seeing a **60% year-on-year increase** in **internet demand** from customers in terms of households in general, and the network has fared very well."

Sources: Seeking Alpha Ltd. and Bank of Canada

References

Angelico, C., J. Marcucci, M. Miccoli and F. Quarta. 2022. "Can We Measure Inflation Expectations Using Twitter?" *Journal of Econometrics* 228 (2): 259–277.

Baker, S. R., N. Bloom and S. J. Davis. 2016. "Measuring Economic Policy Uncertainty." *Quarterly Journal of Economics* 131 (4): 1593–1636.

Bybee, L., B. T. Kelly, A. Manela and D. Xiu. 2020. "The Structure of Economic News." National Bureau of Economic Research Working Paper No. 26648.

Cayen, J.-P. and S. van Norden. 2005. "The Reliability of Canadian Output-Gap Estimates." *North American Journal of Economics and Finance* 16 (3): 373–393.

Champagne, J., G. Poulin-Bellisle and R. Sekkel. 2018. "The Real-Time Properties of the Bank of Canada's Staff Output Gap Estimates." *Journal of Money, Credit and Banking* 50 (6): 1167–1188.

Chen, L. and S. Houle. 2023. "Turning Words into Numbers: Measuring News Media Coverage of Shortages." Bank of Canada Staff Discussion Paper No. 2023-8.

Cheung, C., L. Frymire and L. Pichette. 2021. "Can the Business Outlook Survey Help Improve Estimates of the Canadian Output Gap?" Bank of Canada Staff Discussion Paper No. 2020-14.

Coibion, O., Y. Gorodnichenko and M. Ulate. 2018. "The Cyclical Sensitivity in Estimates of Potential Output." Brookings Papers on Economic Activity, Fall 2018: 343–441.

Corrigan, P., H. Desgagnés, J. Dorich, V. Lepetyuk, W. Miyamoto and Y. Zhang. 2021. "ToTEM III: The Bank of Canada's Main DSGE Model for Projection and Policy Analysis." Bank of Canada Technical Report No. 119.

Cusinato, R. T., A. Minella and S. da Silva Pôrto Júnior. 2013. "Output Gap in Brazil: A Real-time Data Analysis." *Empirical Economics* 44: 1113–1127.

Diebold, F. X. and R. S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics* 13 (3): 134–144.

Faucher, G., C. Hajzler, M. Kuncl, D. Matveev, Y. Park and T. Taskin. 2022. "Potential Output and the Neutral Rate in Canada: 2022 Assessment." Bank of Canada Staff Analytical Note No. 2022-3.

Faust, J. and J. H. Wright. 2009. "Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset." *Journal of Business & Economic Statistics* 27 (4): 468–479.

Faust, J. and J. H. Wright. 2013. "Forecasting Inflation." In *Handbook of Economic Forecasting, Vol. 2*, edited by G. Elliott and A. Timmermann, 2–56. Elsevier B.V.

Gentzkow, M., B. Kelly and M. Taddy. 2019. “Text as Data.” *Journal of Economic Literature* 57 (3): 535–574.

Gervais, O. and M.-A. Gosselin. 2014. “Analyzing and Forecasting the Canadian Economy Through the LENS Model.” Bank of Canada Technical Report No. 102.

Gosselin, M.-A. and G. Tkacz. 2001. “Evaluating Factor Models: An Application to Forecasting Inflation in Canada.” Bank of Canada Staff Working Paper No. 2001-18.

Hassan, T. A., S. Hollander, L. van Lent and A. Tahoun. 2020. “The Global Impact of Brexit Uncertainty.” National Bureau of Economic Research Working Paper No. 26609.

Kamada, K. 2005. “Real-Time Estimation of the Output Gap in Japan and Its Usefulness for Inflation Forecasting and Policymaking.” *North American Journal of Economics and Finance* 16 (3): 309–332.

Kozicki, S. 2004. “How Do Data Revisions Affect the Evaluation and Conduct of Monetary Policy?” Federal Reserve Bank of Kansas City *Economic Review* 89 (Q1): 5–38.

Larsen, V. H., L. A. Thorsrud and J. Zhulanova. 2021. “News-driven Inflation Expectations and Information Rigidities.” *Journal of Monetary Economics* 117: 507–520.

Loughran, T. and B. McDonald. 2016. “Textual Analysis in Accounting and Finance: A Survey.” *Journal of Accounting Research* 54 (4): 1187–1230.

Manela, A. and A. Moreira. 2017. “News Implied Volatility and Disaster Concerns.” *Journal of Financial Economics* 123 (1): 137–162.

Marcellino, M. and A. Musso. 2011 “The Reliability of Real-Time Estimates of the Euro Area Output Gap.” *Economic Modelling* 28 (4): 1842–1856.

Orphanides, A. and S. van Norden. 2002. “The Unreliability of Output-Gap Estimates in Real Time.” *Review of Economics and Statistics* 84 (4): 569–583.

Ruch, F. and T. Taskin. 2022. “Demand and Supply Shocks: Evidence from Corporate Earning Calls.” World Bank Policy Research Working Paper No. 9922.

Shapiro, A. H., M. Sudhof and D. J. Wilson. 2022. “Measuring News Sentiment.” *Journal of Econometrics* 228 (2): 221–243.