NBER WORKING PAPER SERIES

USING SOCIAL MEDIA TO IDENTIFY THE EFFECTS OF CONGRESSIONAL VIEWPOINTS ON ASSET PRICES

Francesco Bianchi
Roberto Gomez Cram
Howard Kung

Working Paper 28749
http://www.nber.org/papers/w28749

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2021, Revised May 2022

Previously circulated as "Using Social Media to Identify the Effects of Congressional Partisanship on Asset Prices." We thank Fernando Stein, Lin Xie, and Sai Zhang for excellent research assistance. We are grateful for comments from Matthew Gentzkow, Bryan Kelly, Thilo Kind, Lubos Pastor, Pietro Veronesi, and many seminar and conference participants. This research was funded in part by the LBS AQR Asset Management Institute. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Using Social Media to Identify the Effects of Congressional Viewpoints on Asset Prices
Francesco Bianchi, Roberto Gomez Cram, and Howard Kung
NBER Working Paper No. 28749
April 2021, Revised May 2022
JEL No. D72,G14

ABSTRACT

This paper documents that individual politicians affect asset prices using a high-frequency identification approach. We exploit the regular flow of viewpoints contained in tweets from Congress members. Supportive (critical) tweets increase (decrease) the stock prices of the targeted firm in minutes around the tweet. The price response persists for several days, during which analysts revise their forecasts about the firm's cash flows. We link the tweets to legislation and find that surges in viewpoints within a bill predict roll call votes months before its signing. Overall, we show that congressional social media accounts are an important source of political news.

Francesco Bianchi
Social Sciences Building, 201B
Department of Economics
Duke University
Box 90097
Durham, NC 27708-0097
and CEPR
and also NBER
francesco.bianchi@jhu.edu

Roberto Gomez Cram
London Business School
Regent's Park
London NW1 4SA
United Kingdom
rgomezcram@london.edu

Howard Kung
London Business School
Regent's Park, Sussex Place
London NW1 4SA
United Kingdom
hkung@london.edu

1 Introduction

Social media platforms have become an increasingly important tool for politicians to communicate with the public. Both federal and state elected government officials actively maintain Twitter accounts by posting a large volume of content regularly. For example, the members of congress collectively average around 30,000 tweets per month, where 20 percent of those posts communicate important viewpoints on key economic agendas. Around 90 percent of these politicians tweet about a trending economic issue each month, with large surges in social media activity occurring around roll call votes. The tweets offer direct and accurate to the second snapshots of politician opinions that are widely available to the public. This paper examines the informational content of tweets by US congress members by studying their effect on asset prices and analyst expectations. We find that these tweets contain a large set of new and unique information about future legislation.

We scrape the official Twitter accounts of all members of the US Senate and House of Representatives. We then select tweets that explicitly convey an opinion about a specific company to facilitate identification, yielding a total of around ten thousand tweets and 500 unique firm mentions. Textual analysis is used to classify if the tone is critical or supportive in a continuous measure. Our benchmark analysis estimates the effect of the politician opinions on stock prices of the targeted firm in the minutes around each tweet. The identifying assumption is that no additional information affecting the stock price is systematically released over such a short time window.

We find that supportive (critical) tweets increase (decrease) the stock prices of the targeted firms in a statistically significant way. We therefore provide direct evidence that congress members affect asset prices through social media. Figure 1 provides an example of the effects of a critical viewpoint by plotting the minute-level stock price of Amazon.com Inc. in a forty-minute window surrounding a tweet posted by Senator Ron Wyden, the top Democrat on the Senate Finance Committee. The red dotted line highlights the exact time of the post. Sen. Ron Wyden posted on the 24th of October of 2019 at 12:01:53 EST: "@SenWarren and I are demanding the FTC investigate whether Amazon's reckless treatment of Americans' personal data broke the law" over concerns the

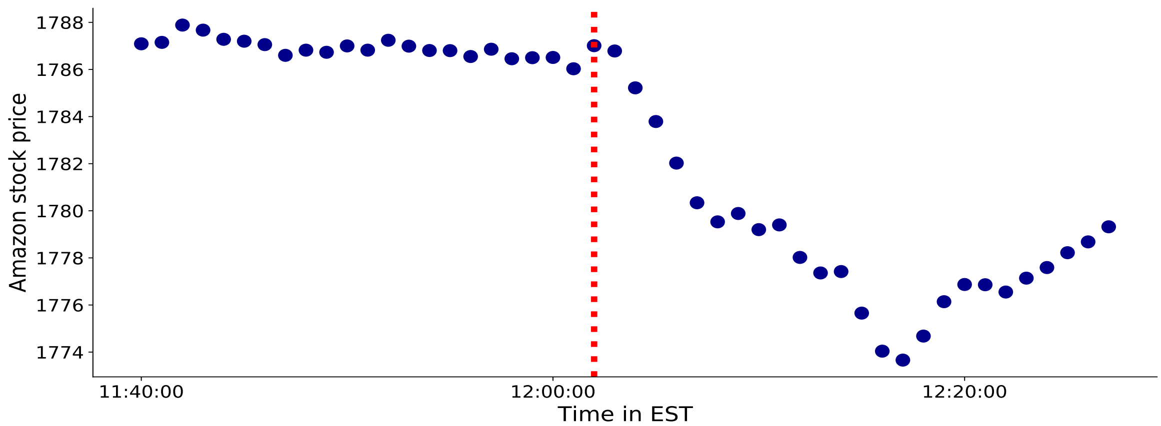

Fig. 1. *Notes*: This figure shows the minute-level stock price of Amazon.com Inc. in a forty-minute window surrounding a tweet posted by Senator Ron Wyden. Sen. Ron Wyden posted on the 24th of October of 2019 at 12:01:53 EST: "@SenWarren and I are demanding the FTC investigate whether Amazon's reckless treatment of Americans' personal data broke the law". The red dotted line highlights the exact time of the post.

company neglected security warnings regarding its cloud-computing system.[1] Amazon's stock price dropped by around 33 bps within minutes after the tweet.

While the high-frequency approach in our benchmark analysis allows for clean identification, the estimates may not quantify the full extent of the effect. To assess the economic relevance, additional analyses are conducted using daily stock price data. The event window in our benchmark estimation is extended from minutes to a daily frequency. With a daily event window, the estimated effect is six times larger, even when controlling for the market factor and several controls. The stock returns in fact drift upwards for the two days after the tweet but then flatten out with no subsequent reversals. There are also no perceptible price trends before the tweets. These patterns suggest that, although stock prices react within minutes of the tweet, it takes several days for the information contained in the politician viewpoints to be fully incorporated into stock prices.

The persistence in the price reaction is exploited to form a long-short portfolio strategy that uses the congressional viewpoints as trading signals. This exercise further quantifies the economic significance of our findings in an out of sample setting. The strategy buys (sells) the stock of the targeted firm if the politician is more supportive (critical). If,

[1]The Wall Street Journal covered this issue and added: "Lawmakers have grown increasingly concerned that these cloud-computing systems, which many companies, including banks, are using to replace traditional data centers, have security problems that are poorly understood by the financial sector." For the news click here.

in a day, a firm is targeted by multiple politicians, we compute a daily average tone measure and use this mean opinion as our trading signal. Value-weighted portfolios are constructed at the end of each day with daily rebalancing. The long-short portfolio strategy earns mean returns of 22 bps per day and is largely unaffected after controlling for standard risk factors. Overall, the strong stock price responses suggest that the tweets by US congress members contain new and relevant information that gets priced in over the next few days.

The stock price responses to the tweets are related to revisions in expectations about firm cash flows. We test if financial analysts revise their expectations about future firm cash flows in the days immediately following congressional tweets and if the tweets predict analysts' forecast errors. We find that the tone of the politicians' viewpoints aggregated at the day-firm level strongly predict subsequent sales and earnings surprises. Moreover, analysts revise their forecasts in line with the tone of the politician opinions during the one week period in which we document a price drift after the tweet. This evidence further highlights that the tweets with politician viewpoints potentially contain new information about firm fundamentals.

We next explore in detail the specific news that drives the price and expected cash flow revisions around the tweets. In particular, we focus on the subset of tweets targeting individual firms that are explicitly related to legislation. Proposed legislation is identified from hashtags included in the tweets. For example, Republican politicians used hashtags such as #TaxReform, #TaxCutsandJobsAct, and #TaxRelief to support the Tax Cuts and Jobs Act of 2017. Democrat politicians used hashtags such as #GOPTaxScam and #TrumpCut to oppose the bill. Given that legislation very rarely impacts a single firm, we find that the stock prices of firms in the same industry as the targeted firm are affected similarly in the tweets directly related to legislation. In contrast, the tweets not explicitly related to legislative news only have an effect on the targeted firm but not on the industry.

The informational content of the tweets can naturally vary depending on the source and the timing. For example, certain politicians may be more influential for a particular legislation and the initial viewpoints can generate larger revisions in expectations. We show that for the tweets relating to legislation, the estimated effects on stock prices are enhanced when the tweet is from a politician that chairs a committee or is more likely to vote independently, linked to a hashtag that appeared for the first time, and is associated with significant user-generated activity.

The timeline of politician viewpoints within a particular bill are examined. The selected tweets offer snapshots of the congressional viewpoints in real time that closely track the legislative process. Tweets targeting firms for a given bill pinpoint the exact informational content regarding the proposed policy change and affected members. This approach helps to identify the effect of a specific policy shock. As an example, we consider the Crapo bill, which was an important legislative initiative in our sample that generated significant social media activity by members of congress. This bill was favorable for the banking sector as it relaxed various restrictions imposed by the Dodd-Frank Act.

The timeline around this bill illustrates how surges in relevant news from politician accounts occurred months before the bill became public law. Significant spikes in Twitter activity are concentrated on the main legislative events, such as when it was received and when it was passed in the Senate and the House, respectively. We document strong price increases for firms in the banking industry in response to tweets from Republican (Democrat) politicians supporting (criticizing) the Crapo bill. The tone of the tweets strongly predict the actual voting behavior of the politicians, highlighting the credibility of the politicians' viewpoints expressed in social media. The appendix shows similar results for other important legislative initiatives such as the "Tax Cuts and Jobs Act." Overall, these results provide support for the idea that markets and analysts use the tweets to extract information about the likelihood or specific details of the proposed legislation.

Our findings that stock prices respond significantly to political news extracted from politician tweets provides direct support for the economic mechanisms featured in the models of Pástor and Veronesi (2012) and Pástor and Veronesi (2013). These papers show that political news shape investor beliefs about which government policy is likely to be adopted in the future. Asset prices therefore respond to the continual flow of political signals before a policy is implemented. Our paper extracts political signals directly from the social media accounts of a large panel of US Congress members. The effect of the signals on asset prices are then identified using a high-frequency approach.

An emerging literature examines the impact of politics on stock returns (e.g., Santa-Clara and Valkanov (2003), Belo, Gala, and Li (2013), Kelly, Pástor, and Veronesi (2016), Addoum and Kumar (2016), and Pástor and Veronesi (2020)). These papers focus on linkages between the party of the incumbent president and aggregate stock returns. Our paper complements this literature by providing granular evidence tying individual politician viewpoints to stock price responses. Cohen, Diether, and Malloy (2013) finds that the voting decisions of legislators have important information about stock returns. Relative to Cohen et al. (2013) who show that legislation after it gets passed affects asset prices (i.e., news about realized legislation), we instead obtain political signals about legislation that start months before voting (i.e., news about expected legislation). A novel aspect of our analysis is that we can observe regular snapshots of the political climate and legislative intent of Congress members even before the official proposal of the legislation. These signals forecast roll call votes and move the stock price of the targeted firm and the corresponding industry portfolio.

Our methodological approach connects to a literature using textual analysis to extract information that affects stock returns (e.g., Boudoukh, Feldman, Kogan, and Richardson (2013), Buehlmaier and Whited (2018), Chen, De, Hu, and Hwang (2014), Engle, Giglio, Kelly, Lee, and Stroebel (2020), Giglio, Maggiori, Rao, Stroebel, and Weber (2021), Hoberg and Moon (2019), Kelly, Manela, and Moreira (2019), Gentzkow, Kelly, and Taddy (2019a), Cookson, Engelberg, and Mullins (2021), and Arteaga-Garavito, Croce, Farroni, and Wolfskeil (2021)) and high-frequency identification in macroeconomics (e.g., Gürkaynak, Sack, and Swanson (2005), Bernanke and Kuttner (2005), Nakamura and Steinsson (2018), and Bianchi, Kind, and Kung (2019)).[2] Kuchler and Stroebel (2021) documents linkages between social media engagement and stock market investments. We build on these strands of literature by highlighting that congressional social media accounts are an important source of political news to market participants.

Our paper relates to the literature measuring political opinions from media platforms. Mullainathan and Shleifer (2005), Gentzkow and Shapiro (2010), and Martin and Yurukoglu (2017) analyze political polarization from media sources such as newspapers and cable news. Gentzkow, Shapiro, and Taddy (2019b) and Jensen, Naidu,

[2]See Gürkaynak and Wright (2013) for a survey on high-frequency event studies in macroeconomics.

Kaplan, Wilse-Samson, Gergen, Zuckerman, and Spirling (2012) measure partisanship in congressional speeches. Birney, Graetz, and Shapiro (2006) examine the role of public opinion on political legislation. We complement this literature by measuring congressional viewpoints about economic agendas at high frequencies directly from the social media accounts of individual members of congress and by studying the impact of the social media posts on stock prices.

2 Data

Our primary data source is the complete set of posts on Twitter (i.e., tweets) by members of the U.S. Senate and House of Representatives from January 2013 to December 2020. The prevalence of the Twitter platform as a congressional communication tool offers several advantages in our empirical analysis relative to traditional mediums of communication. In particular, tweets are in a standardized format and include a time stamp that is accurate to the second. Both of these features allow us to directly measure the viewpoints of individual politicians at high frequencies.

Every official, campaign, and personal account for each congressional member are obtained from their congressional websites. Congress members who did not list a Twitter account on their website or do not have a verified account are dropped. Information on institutional accounts that change hands between consecutive congressional terms are not collected so that each account can be linked to a single legislator. Around 85 percent of the congressional accounts are captured.

Politicians generate a large amount of social media content. Our dataset contains 2.5 million tweets from 740 different Twitter accounts. There are 30 million total words contained in theses posts with 77,000 unique words. In a median month, the median member of congress produces 42 tweets per month with a total of 1,200 likes and retweets per tweet.

A central part of our benchmark analysis is to exploit tweets by members of congress that explicitly convey an opinion about individual companies. Identifying firm mentions in a congressional social media post faces several challenges. First, politicians may mention the same company using different versions or variations of its name. For example, *Apple*

Inc appears as *Apple Computer Inc* in Compustat. However, politicians most likely will write either *Apple* or the twitter account of the company *@Apple* in their tweets. Alternatively, they could even target the company by mentioning the name of the CEO *Tim Cook* or the twitter account of the CEO *@tim_cook* instead of the company name itself. Second, a politician may mention a company name (or several variations of its name) in a tweet without expressing a view about the company being mentioned. For example, H.R. Keith Ellison (@keithellison) writes, "'Good morning! We are on *Apple* Podcasts!" 8 August 2019, 16:50:31 EST, Tweet. This problem is exacerbated if the company name has multiple meanings. For instance, Sen. Chris Coons (@ChrisCoons) "Heading to Bridgeville for *Apple* Scrapple! #netde" 11 October 2014, 16:19:04 EST, Tweet.[3] In both of these cases, members of congress mentioned the word *Apple*, although neither was expressing an opinion about the company. This task of identifying tweets of politicians that target specific companies would be greatly simplified if we just searched for a stock using its ticker. However, in our setting politicians seldom refer to a company by means of its ticker.

The aforementioned challenges are addressed as follows. To make the sample construction manageable, our search is restricted to stocks in the Russell 3000 index. This index contains the 3,000 largest U.S. traded stocks, comprising roughly 98% of the U.S. equity market index in terms of market capitalization. We then select tweets that contained either: (1) the full company legal name or parts thereof, after removing common words such as 'Inc' or 'Corp.'. For the company legal name we use both CONML and CONM in Compustat. (2) the official twitter account for each company, or (3) the Company CEO name or corresponding twitter account. This search generates a total of 24,032 matches. To avoid the problem that a company name might have multiple meanings, we manually checked all twenty four thousand matches to distinguish those that were erroneously classified and did not explicitly mention a company. We will use these subset of false-positive tweets for placebo tests.

Our selection criteria yields 11,602 tweets by members of congress that explicitly convey their opinions about specific companies. Figure 2 plots the total number of

[3]The Apple Scrapple Festival is held annually during the second weekend in October in Bridgeville, Delaware.

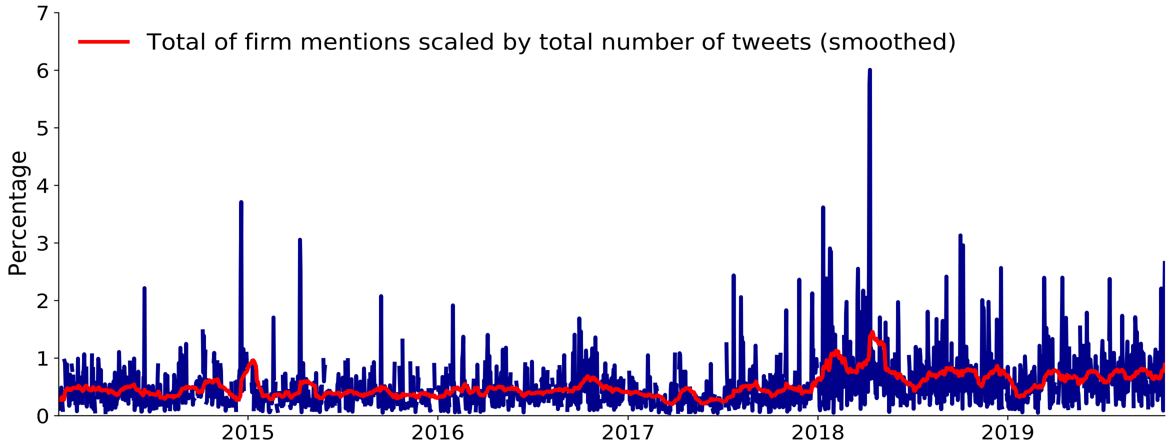

Fig. 2. *Notes*: This figure shows the total daily number of tweets that target an individual company. The count is scaled by the total number of tweets posted during each day. The red line depicts the smoothed series obtained by taking a 15-day moving average. The raw series is in blue. We use the complete set of tweets created by any member of the U.S. Senate and House of Representatives from the 113th Congress to the 116th Congress.

congressional tweets that express an opinion about a company in each day, scaled by the total number of congressional tweets posted during that day. The average value equals 0.5% which amounts to six daily tweets, since Congress as a whole posts on average 1,100 tweets per day. The figure also exhibits substantial amount of variation over time with a standard deviation of 0.46%.

After selecting the tweets that express opinions about specific companies, we proxy for politician viewpoints about companies using a relative tone measure that classifies the tweet as being supportive or critical. To systematically compute the tone, the lexicon developed by Loughran and McDonald (2011) is used. Loughran and McDonald (2011) create word dictionaries of negative and positive words that account for the nuances of finance jargon.[4] Using these dictionaries, we then count the number of positive and negative words for each tweet. We define the relative *Tone* measure as the difference between the positive and the negative word count scaled by the total number of words contained in the tweet.

Intuitively, the tweet will be in support of the company if the tone measure is positive (e.g., Sen. James Lankford (@SenatorLankford) *"The positive news just keeps on coming.*

[4]Loughran and McDonald (2011) generate a list of 2,337 words (353 words) that typically have negative (positive) implications in a financial sense. See `https://sraf.nd.edu/textual-analysis/` for details.

Wal-Mart now joining the growing list of companies w/ plans to increase wages for workers because of the #TaxCutsandJobsAct" 01 November 2018, 15:58:06 EST, Tweet., $Tone = 3.57\%$) and it will criticize the company if the tone measure is negative (e.g., Sen. Ron Wyden (@RonWyden) *"@SenWarren and I are demanding the FTC investigate whether Amazon's reckless treatment of Americans' personal data broke the law"* 24 October 2019, 12:01:53 EST, Tweet., $Tone = -11.76\%$). The average $Tone$ measure is -0.61% with a standard deviation of 5.7%. Figure A.1 in the Appendix shows the time series variation of this series.

We use the Loughran and McDonald dictionaries for our benchmark analysis because it is tractable, scales with ease, and has proved useful in signing text in many related contexts (e.g., Tetlock (2007), Tetlock, Saar-Tsechansky, and Macskassy (2008), Das and Chen (2007), Chen et al. (2014)). The appendix provides robustness results using two alternative tone measures. The first one uses the dictionaries proposed by García, Hu, and Rohrer (2020) which are based on a sophisticated machine learning algorithm. For the second one, we manually classify the tweets as positive, negative, or neutral. We find very similar results when using these alternative tones measures.

Since part of our analysis estimates the impact of politician viewpoints on firm valuations with a high-frequency identification approach, tick-by-tick data on stock prices from the NYSE Trade and Quote (TAQ) database is used. For this analysis, we keep all tweets posted during regular trading days from 9:30 to 16:00 Eastern Standard Time (EST) and collect the firm's ticker for all companies mentioned in our sample, which is then used to merge with the TAQ database. The raw series is cleaned following the procedures described in Brownlees and Gallo (2006).

To establish the economic significance of our findings, daily stock prices from the Center for Research in Security Prices (CRSP) is used. In some of our regressions, we control for firm-level characteristics, such as firm size, book-to-market, and assets. These variables are obtained from Compustat. Analyst-by-analyst sales and earnings per share forecasts are used from the Thomson Reuters unadjusted I/B/E/S Detail History File. Politician characteristics are collected using data on the state, party, chamber, congressional committee, rank therein, and record their entire history of legislative roll call votes. To determine the political ideology of each member, we use the DW-NOMINATE estimates of Carroll, Lewis, Lo, McCarty, Poole, and Rosenthal (2015). DW-NOMINATE scores provide a measure of legislators' ideological locations over time by examining their voting record.

It is plausible that other news about a company sparks a change in asset prices and a social media response from politicians. We use the company-level Bloomberg news sentiment measure as a control to address this concern. Bloomberg uses a supervised machine-learning algorithm to determine the sentiment measure of a news story (published in Bloomberg News, Web content, and select premium news wires) towards a given company. All news stories with sentiment scores for a given company are then aggregated at the daily frequency to produce a company-level sentiment score. The sentiment measure ranges from -1 to 1, where positive (negative) values denote a buy (sell) signal at the company level. The average company-level correlation between the politicians' tone and Bloomberg news sentiment measure is 0.07 (t-statistic = 1.7).

3 Congressional Tweets and Stock Prices

This section documents how politicians' viewpoints about a company conveyed in a tweet have a material impact on the stock price. This evidence suggests that these viewpoints contain valuable new information to market participants. Section 4 provides an economic interpretation of the effects documented in this section by relating tweets to legislative agendas.

3.1 High-frequency analysis

We begin with a high-frequency study of the impact of the selected congressional tweets on the stock prices of the companies for which explicit viewpoints are expressed in the tweet. The tweets targeting individual companies allow for a clean identification in measuring the effect of politician opinions about companies on stock prices. The goal in this section is to assess whether politician opinions about companies can systematically affect their valuations via their social media accounts. Specifically, we analyze the change in the price of the company mentioned in the tweet over a short window of time around

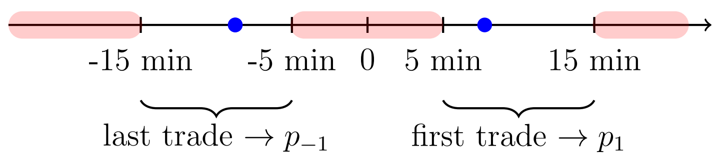

Fig. 3. *Notes*: Time zero denotes the time of the tweet. We are interested in the difference between the log-price of the stock after the tweet and the log-price of the stock before the tweet. We allow for a [-5 minutes,+5 minutes] window around the tweet to give time to markets to react to the new information. If there is no price change after 15 minutes, we conclude that the tweet did not have an impact on the price of the stock.

the tweet. The identifying assumption is that over such a short window of time there is no other relevant information being systematically released affecting the stock price.

Figure 3 provides a visual depiction of the event window used to identify the effect of the tweet on stock prices. Time zero denotes the time of the tweet. We are interested in the difference between the log price of the stock after the tweet and the log price of the stock before the tweet. We allow for a [-5 minutes,+5 minutes] window around the tweet to give time to markets to react to the new information. If there is no trade before and after 15 minutes, we conclude that the tweet did not have an impact on the price of the stock. As a robustness check, we also compute the price change for the tweets that mentioned, but did not express an opinion, about a company to show that such tweets have no material effect on prices. Hence, the total sample in our regressions below includes all tweets that contain a firm mention.

We then regress the log price change on the tone measure for each selected tweet:

$$\Delta p_{i,t} = a + b \cdot D_r \cdot \text{Tone}_{i,t} + \epsilon_{i,t}, \tag{1}$$

where $\Delta p_{i,t}$ denotes the log change in company i's stock price in a five minute window around the tweet, D_r is a dummy variable that equals one if the tweet is in fact relevant, and $\text{Tone}_{i,t}$ is the tone of the tweet referring to company i. The log price change is in basis points (bps) and $\text{Tone}_{i,t}$ is in percentages. The parameter of interest is b, which captures the average marginal response of stock prices in bps around each relevant tweet when the fraction of positive minus negative words is 1% higher. In all regressions below, we use stock fixed effects and cluster the standard errors at the stock-day level.

Two alternative specifications are also considered. In the first one, we include a level effect for the relevant tweets by interacting the dummy variable D_r with a constant as follows:

$$\Delta p_{i,t} = a + D_r \cdot (c + b \cdot \text{Tone}_{i,t}) + \epsilon_{i,t}. \tag{2}$$

In the second alternative specification, we also add the tone for the placebo tweets:

$$\Delta p_{i,t} = a + D_r \cdot (c + b \cdot \text{Tone}_{i,t}) + d \cdot (1 - D_r) \cdot \text{Tone}_{i,t} + \epsilon_{i,t}. \tag{3}$$

Table 1 contains the results. The first column reports the results for the baseline specification described in equation (1). A positive tone has a positive and statistically significant effect on the stock price as implied by the positive slope coefficient $\hat{b}$ equal to 0.36 (t-statistic = 2.8). Stock returns increase by around 2 bps when the tone measure increases by 1 standard deviation ($\approx 0.36 \times 6\%$). This estimate suggests that more supportive (critical) opinions about companies increase (decrease) stock valuations.

Column (2) shows that the point estimate of 0.36 is substantially unaffected by introducing the level effect for the relevant tweets. The level effect is equal to 0.9 bps ($= \hat{a} + \hat{c}$) but it is statistically insignificant. Nevertheless, the positive point estimate suggests that a tweet with a tone equal to zero has a positive effect on the stock price over the sample considered in our analysis. The total effect of a 1% increase in the tone measure equals to 1.3 bps ($= \hat{a} + \hat{b} + \hat{c}$). The positive intercept $\hat{c}$ suggests that a relative tone measure of zero seems to convey positive information. As such, analysis with a discrete measure of viewpoints below that classifies tweets into a 'supportive' and a 'critical' bin based on our relative tone measure considers different cutoffs besides zero.[5] Finally, Column (3) shows that when we introduce the effect of the tone for the placebo tweets in regression (3), we find that the coefficient $\hat{d}$ is close to zero and not statistically significant. Furthermore, the R^2 does not changes when controlling for the tone of the placebo tweets.

Table 2 extends the analysis by considering a discrete measure of tone. We run the

[5]The fact that the relevant tweets with a zero tone measure are associated with an increase of the price of a stock can be explained by the fact that the dictionary used to construct the tone has many more negative than positive words.

Table 1. **High-frequency stock prices responses to congressional viewpoints**

Coef	Variable	(1)	(2)	(3)
a	1	-0.192	-0.896	-0.902
		[-0.591]	[-1.882]	[-1.936]
b	$\text{Tone}_{i,t} \cdot D_{i,r}$	0.359	0.381	0.381
		[2.769]	[2.849]	[2.824]
c	$D_{i,r}$		1.838	1.838
			[1.377]	[1.381]
d	$\text{Tone}_{i,t} \cdot (1 - D_{i,r})$			-0.024
				[-0.231]
nObs		9,932	9,932	9,932
R2		0.118%	0.160%	0.160%

The independent variables are a combination of (a) the tone measure $\text{Tone}_{i,t}$; and (b) a dummy variable that equals one if the tweet is in fact relevant and zero otherwise $D_{i,r}$. In all regressions, the dependent variable is the log change in company i's stock price in a 10 minute window around the tweet $\Delta p_{i,t}$. In all regressions, we use stock fixed effects. Standard errors are double clustered at the stock-day level and t-statistics are in brackets. The tone measure is in percentage and the log price change is in basis points.

following regressions:

$$\Delta p_{i,t} = a + b^s \cdot Support_{i,t} + b^c \cdot Criticize_{i,t} + \epsilon_t, \tag{4}$$

where $\Delta p_{i,t}$ denotes the log change in company i's stock price in the same 10 minute event window around the tweet as the analysis above and the regressors are dummy variables indicating whether the politician tweet is supporting or criticising the targeted firm. In regression (1) the tweet is supportive (critical) of company i if the relative tone measure is positive (negative). Given that we showed that the effect is positive when the tone measure is zero, in regressions (2) and (3) we take a more conservative approach in selecting the cutoffs. The tweet is classified as supportive (critical) of company i if the tone measure is above its 75th percentile (below its 25th percentile) and its 90th percentile (below its 10th percentile), respectively.

The discrete specifications lead to stronger results when focusing on more stricter cutoffs. A positive tweet in the 75th percentile leads to a 7.8 bps increase, while a negative tweet in the 25th percentile leads to 5 bps decline. A positive tweet in the 90th percentile leads to an 11.9 bps increase, while a negative tweet in the 10th percentile leads to 7.1 bps decline. Notably, the t-statistics also increase as we move to tails of the

Table 2. **High-frequency stock price responses to a discrete measure of congressional viewpoints**

	Distribution of Tone		
Support equal to one if	> 0	$> 75\%$	$> 90\%$
Criticize equal to one if	< 0	$< 25\%$	$< 10\%$
	(1)	(2)	(3)
a	-0.275	-0.314	-0.316
	[-0.563]	[-0.775]	[-0.781]
Support	2.454	7.864	11.931
	[1.557]	[2.576]	[2.248]
Criticize	-2.054	-4.973	-7.141
	[-1.836]	[-2.658]	[-2.062]
nObs	9,932	9,932	9,932

This table reports coefficient estimates from the following Equation: $\Delta p_{i,t} = a + b^s \cdot Support_{i,t} + b^c \cdot Criticize_{i,t} + \epsilon_t$, where $\Delta p_{i,t}$ denotes the log change in company i's stock price in a 10 minute window around the tweet. The regressors are dummy variables indicating if the politician tweet is supporting or criticizing the targeted firm. In regression (1) the tweet supports (criticizes) company i if the tone measure is positive (negative). In regressions (2) and (3) the tweet supports (criticises) company i if the tone measure is above its 75th percentile (below its 25th percentile) and its 90th percentile (below its 10th percentile), respectively. In all regressions, we use stock fixed effects. Standard errors are double clustered at the stock-day level and t-statistics are in brackets. The log price change is in basis points.

Tone distribution. These results confirm the findings presented in our baseline analysis and show that more forceful tweets have proportionally larger effects. This is arguably due to the fact that tweets that are stronger in their tone are easier to interpret and likely to be associated with a strong political view, as opposed to a casual observation about a company.

3.2 Persistence of the effects

The previous subsection showed that the opinion of politicians about a company conveyed in a tweet has a statistically significant impact on its stock price in a short window around the tweet, with the sign of the impact depending on if the opinion is supportive or critical. In this subsection, we use daily data to ask whether these effects persist over time and whether they grow in magnitude. Expanding the event window allows us to better measure the economic significance of the effects. On the other hand, a wider event window gives us weaker identification and potential endogeneity concerns for which a variety of firm controls are used in the ensuing analysis.

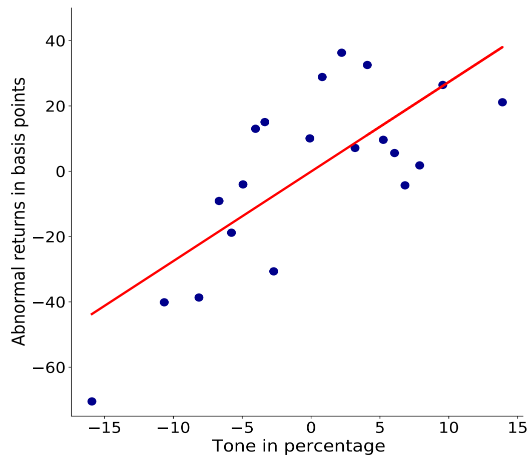

Fig. 4. *Notes*: This figure shows the daily abnormal returns as a function of the tone measure. We sort the tone measure into 20 bins and report the average tone measure and the average abnormal daily return for each of the 20 sorted bins. The abnormal daily return is defined as the company i's return minus the return on the aggregate market. The tone measure is in percentage, and daily abnormal returns are in basis points. The red line denotes the regression fit line.

We focus on daily abnormal returns, defined as:

$$r^{ab}_{i,t,t+1} = r_{i,t,t+1} - r_{m,t,t+1},$$

where $r_{i,t,t+1}$ denotes the daily return on stock i the day after the tweet and $r_{m,t,t+1}$ corresponds to the daily return on the aggregate market during the same day. We construct the daily measure of tone for day t, $Tone_{i,t}$, taking the average of the tone of all tweets about company i on day t.

Figure 4 shows the relation between the daily abnormal returns and the daily tone measure. We sort the tone measure into 20 bins and report the average tone measure and the average abnormal daily return for each of the 20 sorted bins. The red line denotes the regression fit line. We observe a clearly positive relation between the two variables, in line with the high frequency analysis presented above.

We formally analyze the relation between abnormal returns and tone by estimating the following regression:

$$r^{ab}_{i,t,t+1} = a + b \cdot D_r \cdot \text{Tone}_{i,t} + \delta \cdot Controls_{i,t} + \epsilon_{i,t}, \tag{5}$$

where we use a full panel of a firm's stock returns (i.e., days with and without politician tweets targeting a company i). The variable $r^{ab}_{i,t,t+1}$ denotes the company i's abnormal return from day t to day $t+1$, $\text{Tone}_{i,t}$ is the average tone measure for all tweets that target company i on day t. As firm *Controls* we include the average Bloomberg news sentiment measure for company i from $t-3$ to t, $\text{SentimentNews}_{i,t-3,t}$. This variable controls for any news released in media outlets that may affect company i. We also control for the stock i abnormal cumulative return from $t-3$ to $t-1$, $r^{ab}_{i,t-3,t-1}$, and measures of firm size, book-to-market, and assets. Finally, we also include firm fixed effects. Standard errors are double clustered at the stock and day level and t-statistics are in brackets. Abnormal returns are in basis points and the tone measure is in percentages.

Table 3 formalizes the positive relation between abnormal returns and tone by reporting the estimated coefficients for the regression. As shown in column (1), the estimated coefficient on $\text{Tone}_{i,t}$ equals 2.1 (t-statistic = 2.75), suggesting that daily abnormal returns are 2.1 bps higher when the tone measure is 1% higher. Column (2) shows that the coefficient estimate on the tone for the placebo tweets is close to zero and not statistically significant, consistent with our high-frequency results. Column (3) shows that our results are robust to including variables reflecting news sentiment in media outlets towards a given company. Columns (4) and (5) show that the coefficient for the tone of the relevant tweets is substantially unaffected when controlling for the lagged stock i abnormal cumulative returns and measures of firm size, book-to-market, and assets. Thus, the results are robust when moving to daily data and the effects increase in magnitude.

We then ask whether the effect persists over time by extending the holding period. We run the following regression:

$$r^{ab}_{i,t,t+h} = a_h + b_h \cdot D_r \cdot \text{Tone}_{i,t} + \delta \cdot Controls_{i,t} + \varepsilon_{i,t} \tag{6}$$

where $r^{ab}_{i,t,t+h}$ denotes the company i's cumulative return in excess of the return on the aggregate market from t to $t+h$. Thus, the slope coefficient b_h measures the change in abnormal returns (in basis points) over an horizon of h days if the tone measure increases by 1 percent. We consider the following horizons $h = -4$ days, $\cdots, 8$ days. We use the same *Controls* as regression (5) and Table 3.

Table 3. **Daily asset prices responses to congressional viewpoints**

	(1)	(2)	(3)	(4)	(5)
Tone	2.164	2.164	2.069	2.012	1.998
	[2.753]	[2.753]	[2.681]	[2.500]	[2.534]
Tone placebo		0.106	0.084	0.035	0.050
		[0.230]	[0.181]	[0.077]	[0.124]
$\text{SentimentNews}_{i,t-3,t}$			0.231	0.248	0.219
			[8.272]	[8.434]	[8.150]
$r^{ab}_{i,t-3,t-1}$				-0.011	-0.012
				[-0.455]	[-0.479]
log(Size)					6.128
					[2.554]
log(BM)					-4.599
					[-3.007]
log(Assets)					-5.980
					[-2.717]
nObs	784,259	784,259	784,259	754,749	754,749

This table reports coefficient estimates from the following Equation: $r^{ab}_{i,t,t+1} = a + b \cdot D_{r,t} \cdot \text{Tone}_{i,t} + \delta \cdot Controls_{i,t} + \epsilon_{i,t}$, where we use a full panel of a firm's stock returns (i.e., days with and without politician tweets targeting a company i). $r^{ab}_{i,t,t+1}$ denotes the company i's abnormal return from day t to day $t+1$. Abnormal returns are the company's return minus the return on the aggregate market. $\text{Tone}_{i,t}$ is the average tone measure for all tweets that target company i on day t. $D_{r,t}$ is a dummy variable that equals one if tweet on day t is in fact relevant. As $Controls_{i,t}$ we include the average Bloomberg news sentiment measure for company i from $t-3$ to t, $\text{SentimentNews}_{i,t-3,t}$, the stock i abnormal cumulative return from $t-3$ to $t-1$, and measures of firm size, book-to-market, and assets. In all regressions, we use stock fixed effects. Standard errors are double clustered at the stock-day level and t-statistics are in brackets. Abnormal returns are in basis points and the tone measure is in percentage.

Figure 5 reports the results. The dots denote the point estimates for the slope coefficients (b_h), while the vertical lines correspond to 95 percent confidence intervals. Standard errors are double clustered at the firm and day level. Returns significantly drift upward for the next two days and then flatten out. We do not observe a pre-announcement drift in the days before the social media post, with all horizons not statistically significant, except for the -3 horizon. On the contrary, after the tweet the effects are statistically significant at all horizons. Based on these results, we can conclude that the effects of the tweets are persistent and increase in magnitude in the days immediately following a relevant tweet.

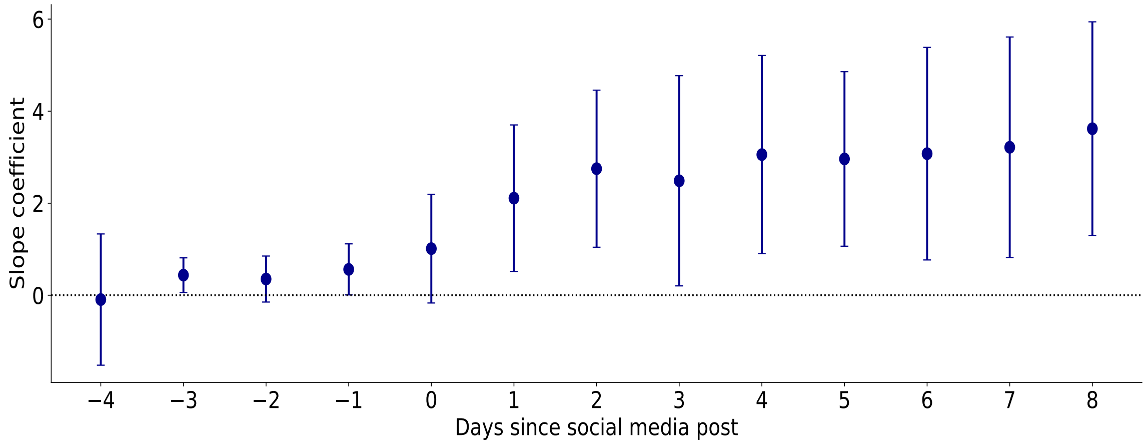

Fig. 5. *Notes*: This figure shows the coefficient estimate b_h from the following specification: $r^{ab}_{i,t,t+h} = a_h + b_h \cdot D_r \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \varepsilon_{i,t}$, where $r^{ab}_{i,t,t+h}$ denotes the company i's cumulative return in excess of the return on the aggregate market from t to $t+h$. We consider the following horizons $h = -4$ days, $\cdots$, 8 days. $\text{Tone}_{i,t}$ is the average tone measure for all tweets that target company i on day t. D_r is a dummy variable that equals one if tweet is in fact relevant. We use the same *Controls* as regression (4) in Table 3. The slope coefficient b_h measures the change in abnormal returns (in basis points) if the tone measure increases by 1 percent. The vertical lines denote 95 percent confidence intervals. We use stock fixed effects and standard errors are double clustered at the firm-day level.

3.3 Trading strategy

We documented that opinions about companies from politician tweets can move stock prices. While stock prices react within minutes of the tweet, the central finding in Figure 5 is that it takes several days for the information to be fully incorporated into asset prices. This slow diffusion of information suggests the possibility of building a trading strategy to exploit the response of markets to the news. The performance of such a strategy can provide another way of quantifying the economic relevance of our findings. Moreover, the trading strategy can be seen as an out-of-sample exercise on the return predictability documented above, since we use information up-to time t to predict $t+1$ returns.

This section presents a feasible and simple strategy that trades on politician viewpoints inferred from the tweets. The $Tone_{i,t}$ measure on firm i is used as a trading signal to buy or sell stock i. We compare the value of $Tone_{i,t}$ with all $\{Tone_{j,\tilde{t}}\}_{j,\tilde{t}<t}$ measures computed for all companies in the previous 12 months. If $Tone_{i,t}$ is above the 75th percentile, the tweet is classified as being supportive of firm i and we buy the stock. If $Tone_{i,t}$ is below the 25th percentile, the tweet is classified as being critical of firm i and we short sell the stock. We then form Long and Short portfolios by value-weighting all stocks that received a signal to buy and sell during a day, respectively. The portfolios

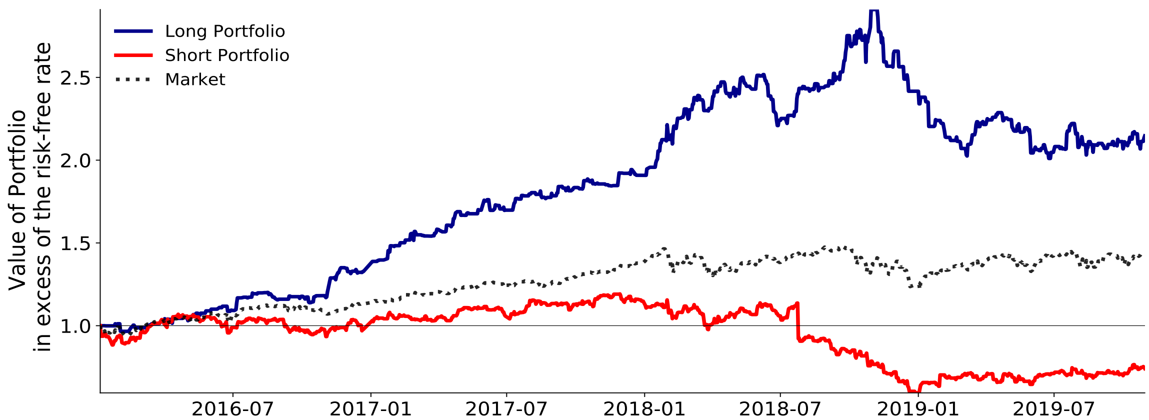

Fig. 6. *Notes*: The figure shows the cumulative daily returns for one-dollar investments in the long portfolio (blue line), the short portfolio (red line), and the aggregate market (black dashed line). All returns are in excess of the risk-free rate. The long (short) portfolio buys stock i if the corresponding tone measure is above (below) the 75 (25) percentile of the tone distribution computed over the previous 12 months. We value-weight the stocks to form portfolios whenever several firms are assigned to the same leg in the same day.

are rebalanced daily. Appendix B.2 presents robustness with respect to the choice of the 25/75 percentiles and with respect to the 12 month look-back measurement period.

We assess the performance of the trading strategy using a variety of profitability measures. To begin, Figure 6 shows the cumulative daily returns in excess of the risk-free rate for investments in the Long, Short, and Market portfolios. A dollar fully invested in the Long portfolio in January 2016 accumulates to around \$2.04 (in excess of risk-free investments) at the end of 2019, but decreases to \$0.74 if invested in the Short portfolio. For comparison, a dollar fully invested in the market grows to about 1.44 dollars over the same sample period.

Next, we compute average daily abnormal returns and report the results in Panel A of Table 4. The Long portfolio return in excess of the market averages 13.7 bps per day (t -statistic = 2.7), and the Short portfolio excess return averages -6.8 bps per day (t -statistic = -1.74). The difference between the Long and Short portfolios is statistically significant. The Long minus Short portfolio that buys each day the Long portfolio and sells the Short portfolios averages a daily excess return of 22 bps with a t-statistic of 3.3.[6] Notably, the standard deviations are remarkably similar across the three portfolios

[6]The mean excess return of the Long - Short portfolio differs slightly from the mean excess return of the Long portfolio minus the mean excess return of the Short portfolio because the former requires that for each date both long and short returns exists.

Table 4. **Performance Evaluation: Mean returns and factor alphas**

	Future returns (day $t+1$)		
	Short	Long	Long - Short
	(1)	(2)	(3)
Mean abnormal return in bps	-6.821	13.725	22.084
	[-1.747]	[2.760]	[3.309]
Std.Dev. (%)	1.228	1.258	1.523
Panel B: Factor alphas in bps			
CAPM alpha	-2.412	15.217	17.235
	[-0.437]	[2.458]	[2.554]
Three-factor alpha	-2.336	15.193	17.148
	[-0.429]	[2.457]	[2.536]
Four-factor alpha	-2.302	15.248	17.488
	[-0.423]	[2.458]	[2.590]

This table reports daily mean returns and daily factor alphas for three different portfolios. The long (short) portfolio buys stock i if the corresponding tone measure is above (below) the 75 (25) percentile of the tone distribution computed over the previous 12 months. We value-weight the stocks to form portfolios whenever several firms are assigned to the same leg in the same day. The "Long-Short" portfolio buys the long portfolio and sells the short portfolio each day. The alphas denote the intercepts form time series regression of the portfolio excess returns on factor alphas. The four factors are the aggregate market excess return, the size factor, the value factor, and the momentum factor. Standard errors adjust for heteroskedasticity and autocorrelation. We report the t-statistics in brackets.

as shown in the bottom row of Panel A.

The strategy also earns large risk-adjusted returns computed using various factor models. These models contain a linear combination of the market factor, the size factor, the value factor, and the momentum factor. The risk-adjusted return or alpha is equal to the intercept from a time-series regression of the portfolio return on the factors. The alpha of this strategy possibly reflects a slow diffusion of information contained in these congressional tweets, consistent with the notion that the stock price responses take a few days to incorporate the viewpoints, as presented in Figure 5.

Panel B of Table 4 reports the results with the alphas quoted in basis points to facilitate interpretation. The alphas are both statistically and economically significant. For example, the Long-Short portfolio consistently earns a daily alpha of 17 bps (t-statistic = 2.5), irrespective of the factor model considered. This point estimate translates into a monthly risk-adjusted return of 3.74%.

3.4 Real effects

To provide insights into the economic underpinnings of the stock price responses documented in the sections above, we next look at financial analysts' forecasts about future cash flows. Specifically, we analyze if analysts revise their expectations in the days immediately following congressional tweets and if the tweets predict analysts' forecast errors. The first exercise allows us to assess whether professionals believe that the tweets contain valuable information about the performance of a company. The second exercise allows to assess whether this is in fact the case. Assuming that the analysts produce a forecast that already incorporates the available information, a positive surprise in earnings following a tweet with a positive (negative) tone indicates that the tweet anticipated positive (negative) developments for the targeted company.

Figure 7 illustrates how the forecast revisions and surprises for sales of firm i are computed. To construct our sample of forecast errors and revisions, we collect analyst-by-analyst sale forecasts from the Thomson Reuters I/B/E/S Detail History File and the actual figures from the I/B/E/S actuals file database. Let FQ_j denote the fiscal quarter j, $Actual_j$ is the corresponding sales realization, and time zero depicts the day of the tweet. We keep all analyst forecasts for the current (i.e., $j = 0$) and next fiscal quarter (i.e., $j = 1$). To compute the analyst consensus for fiscal quarter j before the day of the tweet (i.e., F^j_{-1}), we calculate the median of all forecasts submitted 45 days to one day prior to the tweet. To compute the analyst consensus after the day of the tweet (i.e., F^j_1), we compute the median value for all forecasts submitted between day zero and day eight.

Suppose an analyst makes several estimates for the same firm and event window over the same fiscal quarter. In that case, we take the forecast closest to the day of the tweet to ensure that we consider the most recent estimate. We define forecast revisions as the change in the analyst consensus (i.e., $FR^j = F^j_1 - F^j_{-1}$) before and after the day of the tweet. Forecast errors are computed as the difference between the firm's actual sales and the analyst consensus before the day of the tweet (i.e., $FE^j = Actual_j - F^j_{-1}$). Finally, we scale forecast revisions and surprises by the stock price before the previous fiscal quarter announcement.

The reason for using a relatively short window of eight days after the tweet is that

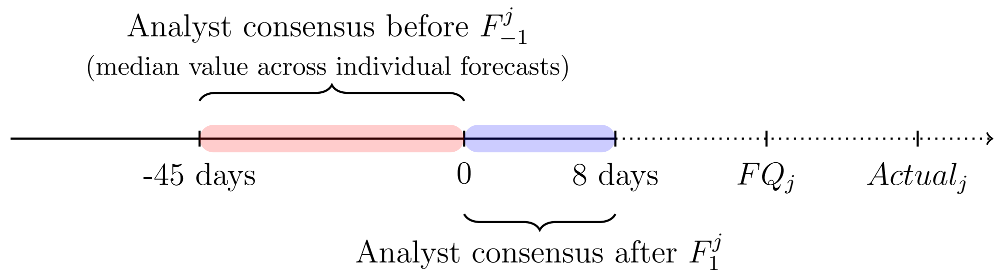

Fig. 7. *Notes*: This figure illustrates how the forecast revisions and surprises for firm i are computed. Time zero denotes the day of the tweet. The analyst consensus for fiscal quarter j before the tweet (i.e., F^j_{-1}) is equal to the median of all forecasts submitted 45 days to one day prior to the tweet. The analyst consensus after the tweet (i.e., F^j_1), equals the median value for all forecasts submitted between day zero and day eight. Forecast revisions are defined as the difference between the analyst consensus (i.e., $FR_j = F^j_1 - F^j_{-1}$).Forecast errors are given by the difference between the firm's actual sales and the analyst consensus before (i.e., $FE_j = Actual_j - F^j_{-1}$).

it takes about a week for the information to be fully reflected in asset prices (see, for instance, Figure 5). Hence, this short window attempts to capture analysts forecasts issued precisely during this learning period. The idea behind the relatively long window of 45 days before the tweet is that we want to use as many forecasts as possible when computing the analyst consensus before the tweet. However, taking a longer window raises the concern that some material information for the fiscal quarter j could have been released during this more extended sample period. Therefore, we add meaningful controls in our regression tests, such as the firm-level cumulative return from day $t - 45$ to day $t - 1$. The Appendix B.3 shows that the results are robust with respect to the choice of the 45 day window.

Figure 8 illustrates the strong positive relation between the tone of the tweets and the analysts' forecast revisions (left panel) and between the tone of the tweets and the analysts' forecast errors (right panel). On the horizontal axis, we report the tone for the tweets after sorting them in 10 different bins. The vertical axis of the left panel is the average price-scaled sales forecast revision. The vertical axis of the right panel is the average price-scaled sales forecast error. All variables are in percentages and the red lines denote the regression fit lines. Thus, a positive tone leads to a positive revision in expectations and positive forecast errors.

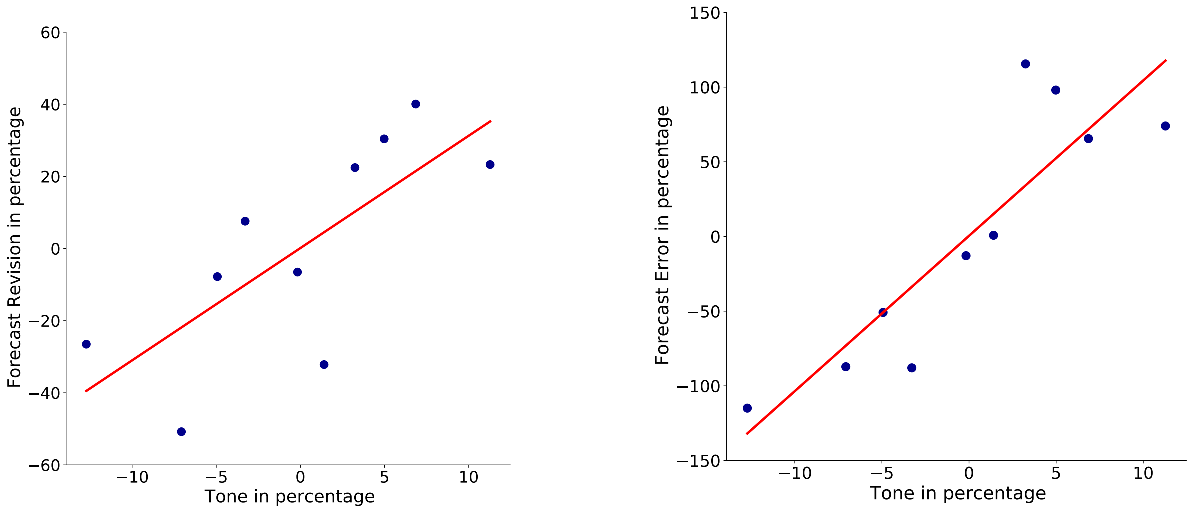

Fig. 8. *Notes*: We sort the tone measure into 10 bins. The left panel reports the average tone measure and the average price-scaled sales forecast revision for each of the 10 sorted bins. The right panel reports the average tone measure and the average price-scaled sales forecast error for each of the 10 sorted bins. Forecast revision is the difference between the consensus sales forecast after and before the tweet. Sales forecast error is the difference between reported quarterly sales and the consensus sales forecast before the tweet. All variables are in percentage. The red lines denote the regression fit lines.

We formalize this finding by running the following regressions:

$$FR_{i,t}^{j} = a + b \cdot D_r \cdot \text{Tone}_{i,t} + c \cdot (1 - D_r) \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}, \tag{7}$$

and

$$FE_{i,t}^{j} = a + b \cdot D_r \cdot \text{Tone}_{i,t} + c \cdot (1 - D_r) \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}, \tag{8}$$

where the dependent variables $FR_{i,t}^{j}$ and $FE_{i,t}^{j}$ denote the price-scaled sales forecast revision and price-scaled sales forecast errors for firm i and fiscal quarter j, respectively. The controls include the average Bloomberg news sentiment measure for company i from $t-45$ to t, the stock i abnormal cumulative return from $t-45$ to $t-1$ and measures of firm size and assets. We also include year-month and industry fixed effects. Standard errors are double clustered at the industry and year-month level and t-statistics are in brackets. $FR_{i,t}^{j}$, $FE_{i,t}^{j}$, and $\text{Tone}_{i,t}$ are in percentage.

Table 5 reports coefficient estimates. In column (1), we estimate equation (7) without considering any controls. We find that $\hat{b} = 2.16$ (t-statistic $= 2.4$). This implies that a one standard deviation higher tone measure is associated with a 12.3 higher analyst sales forecast (relative to prices) in the days immediately following a tweet. For comparison,

Table 5. **Real effects of congressional viewpoints**

			Forecast errors		
	Forecast revisions i.e., $F_1^j - F_{-1}^j$		Before i.e., $Actual_j - F_{-1}^j$		After i.e., $Actual_j - F_1^j$
	(1)	(2)	(3)	(4)	(5)
Tone	2.16	1.58	3.48	3.28	3.29
	[2.40]	[1.85]	[3.20]	[2.64]	[1.76]
Tone placebo		0.79		-2.27	-5.85
		[0.54]		[-1.29]	[-1.76]
$\text{SentimentNews}_{i,t-45,t}$		2.10		4.59	9.28
		[2.88]		[5.21]	[6.18]
$r^{ab}_{i,t-45,t}$		0.03		0.03	0.02
		[6.49]		[6.27]	[2.37]
log(Size)		108.85		395.82	522.93
		[5.17]		[16.14]	[10.98]
log(Assets)		-96.33		-541.85	-640.76
		[-5.26]		[-24.81]	[-15.79]
Obs	18,981	17,384	44,742	41,838	17,384

This table reports coefficient estimates from the following equation $y^j_{i,t} = a + b \cdot D_r \cdot \text{Tone}_{i,t} + c \cdot (1 - D_r) \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}$. In Columns (1) and (2) the dependent variable is the price-scaled sales forecast revision $FR^j_{i,t}$. In Columns (3) and (4) the dependent variable is the price-scaled sales forecast error $FE^j_{i,t}$. In Column (5) the dependent variable is the price-scaled sales forecast errors computed with respect to the consensus analysts' forecasts produced over the 8 days after the tweet. Figure 7 provides the details on how we compute these measures. The regressor of interest is $\text{Tone}_{i,t}$ which denotes the average tone measure for all tweets that target company i on day t. As controls we include the average Bloomberg news sentiment measure for company i from $t - 45$ to t, the stock i abnormal cumulative return from $t - 45$ to t, and measures of firm size and assets. We also include year-month and industry fixed effects. Standard errors are double clustered at the industry and year-month level and t-statistics are in brackets.

scaled analyst revisions have a standard deviation of 475 in our sample (the mean is slightly negative and equals -18.5). Column (2) shows that the estimated coefficient does not change once we control for previous cumulative returns, size, assets, and the year-month and industry fixed effects. Next we test whether forecast surprises tend to be positive following a tweet with a positive tone. Columns (4) and (5) shows that this is indeed the case. The tone of the politicians tweet positively forecasts errors. The estimated coefficient in the full specification equals 2.53 (t-statistic = 2.2). Notably, when considering the placebo tweets, we find that the tone is not statistically significant and with no consistent sign between forecast revisions and forecast errors.

Next, we examine whether analysts correctly revised their expectations in the right amount following the politician tweet. To do so, we estimate the following equation:

$$FE^{j}_{i,t}(after) = a + b \cdot D_r \cdot \text{Tone}_{i,t} + c \cdot (1 - D_r) \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}, \qquad (9)$$

where the dependent variable $FE^{j}_{i,t}(after)$ denotes the price-scaled sales forecast errors computed with respect to the consensus analysts' forecasts produced over the 8 days after the tweet (i.e., $FE(after)^{j}_{i,t} = Actual_j - F^{j}_{1}$). Column (6) of Table 5 reports the results. Interestingly, we still find that the estimated coefficient is positive, even if smaller and not statistically significant. This result suggests that although analyst forecasts move in the right direction following the tweets, the revisions are not large enough to fully eliminate the forecast errors. This underreaction is consistent with a slow dissemination of the new information embedded in the tweet. The next section examines the informational content of the tweets and provides an economic interpretation.

Appendix B.3 shows that the results presented in this subsection for sales also hold for earnings per share (see Table B.5): A positive tone is associated with positive revisions in forecasts and actual earnings. Summarizing, when a politician supports (criticizes) a company, both sales and earnings for the company increase (decrease) and analysts revise their forecasts accordingly. Thus, politician opinions about firms contained in the tweets have predictive power for the actual path of firm cash flows of the targeted firms and analysts revise their expectations accordingly.

4 News About Legislation

This section investigates why politician opinions about companies expressed in tweets impact their stock prices. We explore the possibility that some of these tweets contain new and important information about future legislation related to the targeted company.

4.1 Selecting the news

We first select the subset of tweets targeting individual firms that are explicitly related to legislation. Proposed legislation is identified from hashtags included in the tweets. The

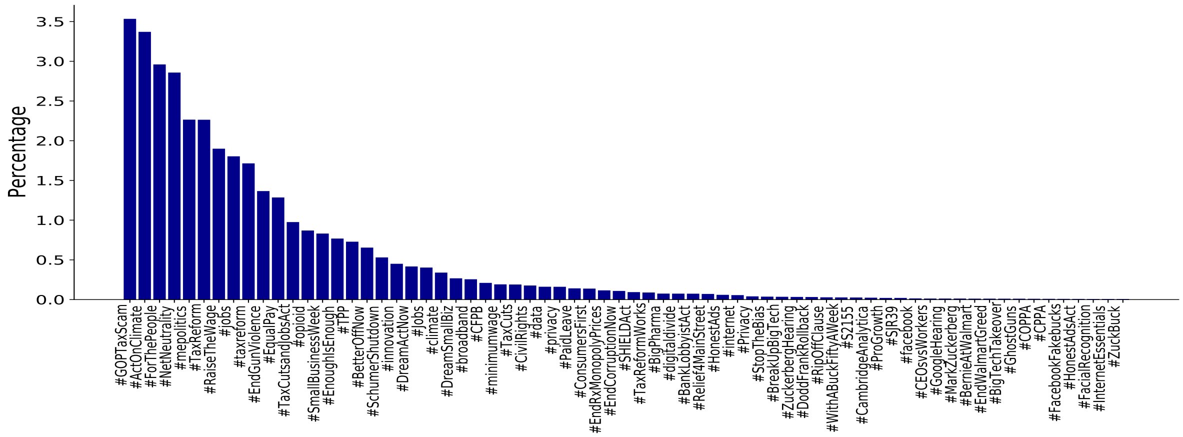

Fig. 9. *Notes*: This figure shows the policy relevant hashtags obtained from all tweets that target a firm. The vertical bars indicate the percentage of times each specific hashtags is mentioned scaled by all tweets that contain a hashtag.

tweets are first sorted based on their hashtag and then the hashtags related to legislation are selected. The use of hashtags provides a clean way of identifying the topic of the post without requiring any estimation.

Figure 9 shows the policy relevant hashtags obtained from all tweets that target a firm and also contain a hashtag. On the horizontal axis we display the hashtag, while the vertical bars capture how many times the corresponding hashtag is mentioned as a percentage of all tweets that target a company and contain a hashtag. The hashtags align closely with future legislation. Below we highlight a few representative examples.

Republican legislators used hashtags such as #TaxReform, #TaxCutsandJobsAct, and #TaxRelief to support the Tax Cuts and Jobs Act of 2017.[7] Democrats, on the other hand, strongly opposed this legislation through the use of opposing hashtags such as #GOPTaxScam and #TrumpCuts. The 2017 tax reform made the largest changes to the US tax code in over thirty years. Other important examples are the following four hashtags: #S2155, #DoddFrankRollback, #BankLobbyistAct, and #Relief4MainStreet. Lawmakers used these four hashtags to express their viewpoints about the Crapo Bill. The bill reversed some of the reforms of the Dodd-Frank Act. In the next section and in Appendix B.4, we study these two bills in detail and identify the effect of congressional viewpoints on asset prices within a specific policy reform.

[7]The short title "Tax Cuts and Jobs Act" was not approved by Senate and to comply with Senate rules, the official title of the bill was changed to "An Act to provide for reconciliation pursuant to titles II and V of the concurrent resolution on the budget for fiscal year 2018."

Other notable hashtags are #BigPharma and #EndRxMonopolyPrice, which were used by Democrats to criticize firms in the Pharmaceutical sector and to push legislative proposals aiming to reduce the industry prices and profits. In addition, Democrats in the U.S. Senate and some Republican senators used the hashtag #HonestAds in support of a bill called the "Honest Ads Act" that regulates political ads and help to prevent foreign interference in U.S. elections. Moreover, #BreakUpBigTec reflect bipartisan consensus on antitrust lawsuits.

4.2 Industry effects

While our results in Section 3 focus on the stock price responses of the targeted firm, very rarely does proposed legislation only impact an individual firm (e.g., Cohen et al. (2013)). As such, this subsection examines if the selected tweets also impacts the stock prices of the other companies in the industry in the same direction.

To assess whether a tweet commenting on firm i also affects the other firms in the same industry, we form industry portfolios excluding firm i and compute the daily abnormal returns following the tweet:

$$r^{ab}_{ind-i,t,t+1} = r_{ind-i,t,t+1} - r_{m,t,t+1},$$

where $ind - i$ denotes the industry of stock i, $r^{ab}_{ind-i,t,t+1}$ is the value-weighted abnormal return on industry $ind - i$ excluding the return on company i, $r_{ind-i,i,t,t+1}$ denotes the daily value-weighted return on the industry of stock i (excluding stock i), and $r_{m,t,t+1}$ measures the daily return on the aggregate market. Concerning the industry classification, we assign each company to an industry based on its four-digit SIC code following the 48 industry definitions provided on Kenneth French's website.

We then run the following regression:

$$r^{ab}_{ind-i,t,t+1} = a + b \cdot D_r \cdot \text{Tone}_{i,t} + c \cdot (1 - D_r) \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t},$$

where $\text{Tone}_{i,t}$ is the average tone measure for all tweets that target company i on day t. As controls we include the average Bloomberg news sentiment measure in industry $ind-i$

Table 6. **Daily industry price responses to congressional viewpoints**

	(1)	(2)	(3)
Tone	0.272	0.300	0.072
	[1.835]	[2.360]	[0.973]
Tone-placebo	0.322	0.244	0.271
	[0.476]	[0.419]	[0.416]
Tone–#legislation			1.650
			[2.096]
Industry SentimentNews$_{ind,t-3,t}$		0.374	0.305
		[2.420]	[2.213]
Industry $r^{ab}_{ind-i,t-3,t-1}$		-0.084	-0.015
		[-1.091]	[-0.906]
Industry Size		0.000	0.000
		[3.368]	[2.209]
Industry Assets		-0.001	-0.001
		[-2.221]	[-1.409]
Industry Book-to-Market		-31.564	-27.830
		[-1.027]	[-0.902]
nObs	14,923	15,172	14,635

Denote with $ind-i$ the industry of stock i. This table reports coefficient estimates from the following Equation: $r^{ab}_{ind-i,t,t+1} = a + b \cdot \text{Tone}_{i,t} + c \cdot D_{\#legislation,t} \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}$, where $r^{ab}_{i,t,t+1}$ denotes the value-weighted abnormal return on industry $ind-i$ excluding the return on company i. Abnormal returns are the $ind-i$ return minus the return on the aggregate market. Tone$_{i,t}$ is the average tone measure for all tweets that target company i on day t. $D_{\#legislation,t}$ is a dummy variable that takes the value of one if the tweet contains one of policy relevant hashtags shown in Figure 9. As controls we include the average Bloomberg news sentiment measure in industry $ind-i$ from $t-3$ to t, the $ind-i$ abnormal cumulative return from $t-3$ to $t-1$, and measures of industry-level average size, book-to-market, and assets. We also include year-month and industry fixed effects. Standard errors are double clustered at the industry and year-month level and t-statistics are in brackets. Abnormal returns are in basis points and the tone measure is in percentage.

from $t-3$ to t, the $ind-i$ abnormal cumulative return from $t-3$ to $t-1$, and measures of industry size, book-to-market, and assets. We also include year-month and industry fixed effects. Standard errors are double clustered at the industry and year-month level and t-statistics are in brackets. Abnormal returns are in bps and the tone measure is quoted in percentages.

Table 6 reports the results. We find that a positive tone in a tweet about firm i is associated with an increase in the abnormal returns of the industry to which the firm belongs. The effect becomes larger and strongly statistically significant when introducing the controls. In terms of magnitudes, the effect is quantitatively smaller than the stock price response of the targeted firm in each tweet (see Table 3). A possible explanation for the weaker industry response is due to the noise induced by the industry classification. As before, the tone of the placebo tweets has a non-statistically significant effect on industry returns.

These results indicate that the selected tweets contain material information beyond the targeted firm that it is also relevant to the industry. These industry effects are consistent with an interpretation that the tweets contain news about future legislation and not just isolated news affecting an individual firm.

To test whether the industry effects are driven by news about legislation, we run the following regression:

$$r^{ab}_{ind-i,t,t+1} = a + b \cdot \text{Tone}_{i,t} + c \cdot D_{\#legislation,t} \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}, \tag{10}$$

where $D_{\#legislation,t}$ is a dummy variable that takes the value of one if the tweet contains hashtags related to legislation as shown in Figure 9. The coefficient c measures the stock price response to the subset of tweets targeting companies *and* that explicitly reference specific legislation.

The third column of Table 6 shows that the tone of the tweets related to legislation has a positive and statistically significant effect, while the coefficient on the tone alone becomes smaller and not statistically significant (but still positive). This result implies that the industry effects are primarily concentrated in the tweets that contain news about legislation.

4.3 Distinguishing the news content

The regression below introduces an additional regressor to the previous specification that allows us to identify the tweets that generate the largest effects:

$$r^{ab}_{ind-i,t,t+1} = a+b\cdot\text{Tone}_{i,t}+c\cdot D_{\#policy,t}\cdot\text{Tone}_{i,t}+d\cdot D_{X,t}\cdot D_{\#policy,t}\cdot\text{Tone}_{i,t}+\delta\cdot Controls+\epsilon_{i,t} \tag{11}$$

where $D_{X,t}$ is a dummy variable that takes the value of one if the tweet is relevant based on a criterion "X". We consider four possible criteria: (1) The politician chairs a committee; (2) the hashtag appeared for the first time; (3) the tweet is newsworthy,

Table 7. **Daily industry responses to congressional viewpoints related to policy**

	Criterion "X"			
	Politician chairs a Committee (1)	First day of # (2)	Newsworthy (3)	DW-nominate dissidents (4)
Tone	0.100	0.100	0.091	0.089
	[1.250]	[1.257]	[1.058]	[1.053]
Tone–#legislation	1.569	1.532	1.488	1.490
	[1.928]	[2.026]	[2.597]	[1.970]
Tone–#legislation & X	1.723	8.589	1.459	6.519
	[0.666]	[1.802]	[0.536]	[1.651]
Tone placebo	0.312	0.313	0.245	0.245
	[0.460]	[0.461]	[0.389]	[0.388]
Controls	Yes	Yes	Yes	Yes
nObs	14635	14635	13901	13901

This table reports coefficient estimates from the following Equation: $r^{ab}_{ind-i,t,t+1} = a + b \cdot \text{Tone}_{i,t} + c \cdot D_{\#policy,t} \cdot \text{Tone}_{i,t} + d \cdot D_{X,t} \cdot D_{\#policy,t} \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}$, where $r^{ab}_{ind-i,t,t+1}$ are the value-weighted abnormal returns for the industry to which firm i belongs, excluding firm i, $D_{\#policy,t}$ is a dummy variable that takes the value of one if the tweet contains one of policy relevant hashtags shown in Figure 9, and $\text{Tone}_{i,t}$ is the relevant tone measure. $D_{X,t}$ is a dummy variable that takes the value of one if the tweet is relevant based on a criterion "X". We consider four possible criteria: (1) The politician chairs a committee; (2) the hashtag appeared for the first time; (3) the tweet is newsworthy, meaning that the sum of likes, comments, and retweets is above the 90th percentile for the politician; (4) the politician is not very partisan, meaning that he/she is between the 40th and 60th percentile (i.e., middle 20%) in terms of the DW-NOMINATE scores. *Controls* includes the following variables: the average Bloomberg news sentiment measure in industry $ind - i$ from $t-3$ to t, the industry cumulative abnormal return from $t-3$ to $t-1$, $r^{ab}_{i,t-3,t-1}$, industry-level average size, book-to-market, and assets. Abnormal returns are in basis points and the tone measure is in percentage.

meaning that the sum of likes, comments, and retweets is above the 90th percentile for the politician; (4) the politician does not have a strong partisan lean, meaning that the congress member is between the 40th and 60th percentile (i.e., middle 20%) in terms of the DW-NOMINATE scores. Columns (1) through (4) in Table 7 report the results.

We find that each of the four criteria magnify the effects of the tweets. The effect is an order of magnitude larger (with larger t-statistics) when the tone is associated with a tweet in which a hashtag associated with legislation appeared for the first time and if the politician who tweeted does not have a strong partisan lean. These last results deserve some further discussion. Column (2) shows that when the hashtag about legislation appears for the first time, the estimated effect is 8.5 higher, with a t-statistic of 1.8 on the difference. The fact that the effect is greatly magnified when a hashtag appears for the first time can be explained as an information effect. More information is likely to be contained in these first tweets because they bring attention to the legislation at hand. Subsequent tweets about the legislation have smaller stock price effects than the initial one.

Column (4) shows that the viewpoints of politicians exhibiting weaker partisanship have stronger stock price effects. The point estimate equals 6.5 (t-statistic = 1.65). The fact that the opinions of politicians who are less partisan have a larger effect is arguably due to the fact that these politicians are more likely to be the marginal voter on closely contested bills. Thus, the opinion of these politicians potentially have a large impact on markets' expectations about the future legislative environment.

4.4 The timeline of news within a policy reform

The timeline of a policy reform is examined in real time through the congressional social media accounts. The tweets offer frequent snapshots of politician opinions that closely track the legislative process. We now focus on a subset of our tweets targeting firms and that are related to a specific bill to nail down the informational content regarding the proposed policy and affected members. This approach helps to identify the effect of a specific policy shock. The timeline of tweets around this bill illustrates how our data sample captures surges in relevant news months before the bill became public law.

An important legislative initiative that generated a large volume of congressional social media activity and intense media scrutiny is the Crapo Bill, officially called Economic Growth, Regulatory Relief, and Consumer Protection Act (S.2115), was signed into law in 2018. The bill takes the name from Mike Crapo (R-ID), the United States Senator who sponsored it and the chair of the Senate Banking Committee at the time the bill was passed. The bill passed the Senate vote by a margin of 67 to 31 in March 2018, passed the House by a Yea-Nay vote of 258 - 159, and was signed by former President Donald Trump in May 2018.

The bill was perceived as favorable for the banking sector because it was relaxing several restrictions introduced after the 2008/9 financial crisis. The bill raised the asset threshold for banks to be considered too big to fail from $50 billion to $250 billion. The

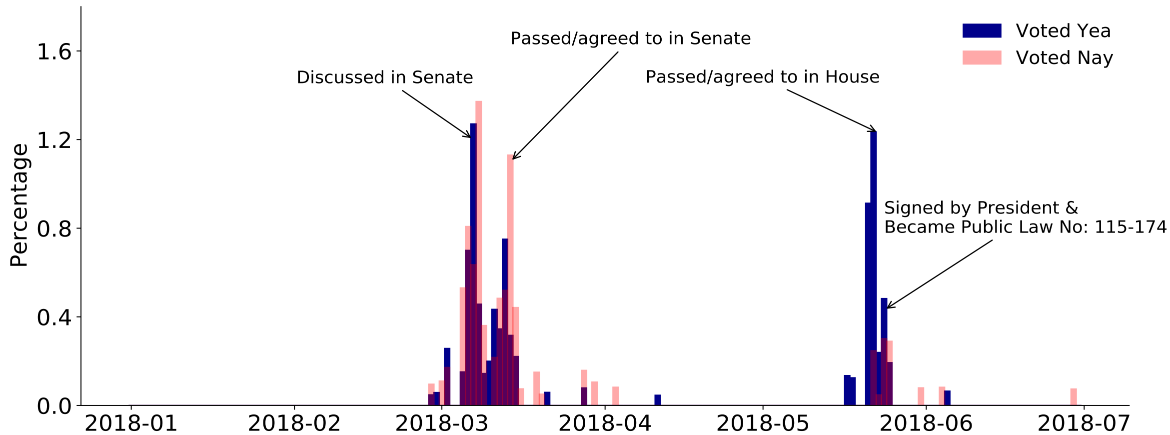

Fig. 10. *Notes*: This figure shows the percentage of congressional posts that explicitly reference the Crapo Bill (officially called Economic Growth, Regulatory Relief, and Consumer Protection Act (S.2115)) as identified by the following hashtags #S2155, #DoddFrankRollback, #BankLobbyistAct, and #Relief4MainStreet. We further divide the tweets made by members of congress who voted Yea (blue bars) or Nay (red bars). The bill passed the Senate with an amendment by Yea-Nay vote of 67 - 31. The bill passed the House by Yea-Nay vote of 258 - 159.

bill also eliminated the Volcker Rule for small banks with less than $10 billion in assets. The Volcker rule takes the name from the ex-Fed Chairman Paul Volcker who proposed it in the aftermath of the 2008/9 financial crisis to restrict commercial banks from engaging in proprietary trading backed by deposits.

Figure 10 shows the percentage of congressional tweets that explicitly reference the Crapo Bill given by the following hashtags: #S2155, #DoddFrankRollback, #BankLobbyistAct, and #Relief4MainStreet. We distinguish between the tweets made by members of congress who voted Yea (blue bars) or Nay (red bars). The number of tweets related to the bill clearly increases on the main legislative events, suggesting that the Congress members use the tweets to communicate their views on the bill.

As a next step, we analyze the relation between the tone used in the tweets and subsequent votes. For each Congress member, we compute the average tone in the tweets referring to the Crapo Bill. We then sort the Congress members based on the tone measure into bins and compute the average tone measured over the corresponding bin. We then track how the Congress members in each bin voted and compute the percentage of yea votes for the corresponding bin. We then plot the percentage of yea votes as a function of the average tone in Figure 11. We find that the tone is a strong predictor of how the Congress members vote, underscoring the credibility of the viewpoints contained

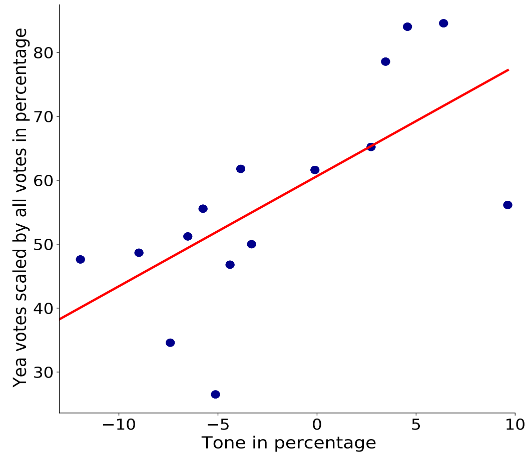

Fig. 11. *Notes*: This figure shows the percentage of yea votes as function of the tone measure. For each Congress member, we compute the average tone in the tweets referring to the Crapo Bill. We then sort the resulting averages based on their tone measure into bins and compute the average tone measure over the corresponding bin. We then track how the Congress members in each bin voted and compute the percentage of yea votes for the corresponding bin. The tone measure is computed over all tweets that explicitly reference the Crapo Bill (officially called Economic Growth, Regulatory Relief, and Consumer Protection Act (S.2115)) as identified by the following hashtags #S2155, #DoddFrankRollback, #BankLobbyistAct, and #Relief4MainStreet. The bill passed the Senate with an amendment by Yea-Nay vote of 67 - 31. The bill passed the House by Yea-Nay vote of 258 - 159. All variables are in percentage. The red line denotes the regression fit line.

in the politician tweets. These results provides support for the idea that markets and analysts use the tweets to extract information about the likelihood or specific details of the proposed legislation. Figure A.2 in the Online Appendix shows that the tone used in the tweets also aligns with the ideological position of the corresponding Congress member as measured by the DW-NOMINATE score.

The following regression quantifies the relevance of the politician viewpoints about the bill:

$$r^{ab}_{banking-i,t,t+1} = a + b \cdot \text{Tone}_{i,t} + c \cdot D_{\#Crapo,t} \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t},$$

where $r^{ab}_{banking-i,t,t+1}$ is the abnormal return for the banking sector excluding the firm i explicitly mentioned in the tweet, $D_{\#Crapo,t}$ is a dummy variable that equals one when the tweet contains one of the hashtags that refer to the Crapo bill, and, to save on notation, $\text{Tone}_{i,t}$ is the tone of a *relevant* tweet mentioning firm i. We include as controls the average Bloomberg news sentiment measure in the banking industry from $t-3$ to t, the three-day holding period abnormal return for the banking industry $r^{ab}_{banking-i,t-3,t-1}$

Table 8. **Daily banking price responses to congressional viewpoints**

	(1)	(2)
Tone	-0.305	-0.293
	[-0.574]	[-0.525]
Tone #policyRelevant	17.893	17.855
	[2.390]	[2.381]
Tone placebo		-0.471
		[-0.738]
Banking $\text{SentimentNews}_{ind,t-3,t}$		0.154
		[1.365]
$r^{ab}_{ind-i,t-3,t-1}$		0.024
		[0.529]
Banking Avg. Size		-0.000
		[-0.301]
Constant	8.510	12.813
	[2.856]	[0.798]
nObs	1,006	981
R2	0.588	1.319

This table reports coefficient estimates from the following Equation: $r^{ab}_{banking-i,t,t+1} = a + b \cdot \text{Tone}_{i,t} + c \cdot D_{\#Crapo,t} \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}$, where $r^{ab}_{banking-i,t,t+1}$ are abnormal returns for the banking sector excluding the firm i explicitly mentioned in the tweet, $\text{Tone}_{i,t}$ is the tone of a relevant tweet mentioning firm i, and $D_{\#Crapo,t}$ is a dummy variable that equals one when the tweet contains one of the hashtags that refer to the Crapo bill: #S2155, #DoddFrankRollback, #BankLobbyistAct, and #Relief4MainStreet. We include as controls the average Bloomberg news sentiment measure in the banking industry from $t-3$ to t, the three-day holding period abnormal return for the banking industry $r^{ab}_{banking-i,t-3,t-1}$, and the banking average size. Standard errors adjust for heteroskedasticity and t-statistics are in brackets. Abnormal returns are in basis points and the tone measure is in percentage.

and the banking average size. The parameter of interest is c, which measures the effect of the politician viewpoint about the Crapo bill on the banking industry abnormal return.

The first column of Table 8 shows that the effect of the politician viewpoints on stock prices is concentrated on the tweets that explicitly mention the Crapo bill. A one percent increase in the tone implies an additional 0.2% (t-statistic = 2.4) in abnormal daily returns. Notably, the effect of all other tweets is just below zero, but not statistically significant. The second column of Table 8 shows that these results are virtually the same once we add the controls.

The Appendix B.4 further show how politicians systematically time their tweets around key legislative events, such as the "Tax Cuts and Jobs Act" (see Figure B.3). More generally, in the Appendix we document that a high volume of tweets related to legislation occur on the corresponding roll call days (see Figure B.4). To this end, we identify tweets that contain at least one keyword associated with the following six categories: (1) government budget; (2) homeland security; (3) energy, commerce, and labor; (4) regulation; (5) corporations; and (6) healthcare. Then, for each roll call we use their corresponding title or brief descriptions as compiled through the electronic voting machines and identify those roll calls that include at least one of these topic keywords. Finally, we select tweets by members of congress that contain the relevant economic keywords that were posted on the same day as the economic roll calls. See Appendix B.5 for details.

5 Conclusion

This paper extracts political opinions from individual US Congressional social media accounts. Politician tweets that support (criticize) a specific firm increase (decrease) the stock price within minutes around the tweet. The price response persists for several days but then flatten out with no subsequent reversals. A trading strategy that exploits the slow diffusion of political news earns sizable abnormal returns, highlighting the economic significance of the new information contained in the tweets. Financial analysts revise cash flow forecasts consistently with the price response in the days immediately following the tweet. The forecast revisions suggest that the politician viewpoints contain relevant news about firm fundamentals.

A subset of the tweets targeting firms are explicitly linked to legislation, yielding information about proposed policy reforms. We find that the stock prices of firms in the same industry as the targeted firm are impacted similarly by these tweets about legislation. Tweets not related to legislation only have an effect on the targeted firm but not on the industry. We also exploit the cross-sectional variation across politicians and tweets to identify the tweets that generate the largest impact on stock prices. Politicians who are more likely to be marginal voters have larger effects. Politician opinions within a specific bill exhibits surges in relevant news that predict roll call votes months before the signing of the bill. Overall, we show that congressional social media accounts are an important source of political news.

References

Addoum, J. M., Kumar, A., 2016. Political sentiment and predictable returns. The Review of Financial Studies 29, 3471–3518.

Arteaga-Garavito, M. J., Croce, M., Farroni, P., Wolfskeil, I., 2021. When the markets get co.v.i.d: Contagion, viruses, and information diffusion. Working paper, Bocconi University.

Auerbach, A. J., 2018. Measuring the effects of corporate tax cuts. Journal of Economic Perspectives 32, 97–120.

Belo, F., Gala, V. D., Li, J., 2013. Government spending, political cycles, and the cross section of stock returns. Journal of Financial Economics 107, 305–324.

Bernanke, B. S., Kuttner, K. N., 2005. What explains the stock market's reaction to federal reserve policy? The Journal of finance 60, 1221–1257.

Bianchi, F., Kind, T., Kung, H., 2019. Threats to central bank independence: High-frequency identification with twitter. Tech. rep., National Bureau of Economic Research.

Birney, M., Graetz, M. J., Shapiro, I., 2006. Public opinion and the push to repeal the estate tax. National Tax Journal pp. 439–461.

Boudoukh, J., Feldman, R., Kogan, S., Richardson, M., 2013. Which news moves stock prices? a textual analysis. Tech. rep., National Bureau of Economic Research.

Brownlees, C. T., Gallo, G. M., 2006. Financial econometric analysis at ultra-high frequency: Data handling concerns. Computational Statistics & Data Analysis 51, 2232–2245.

Buehlmaier, M. M., Whited, T. M., 2018. Are financial constraints priced? evidence from textual analysis. The Review of Financial Studies 31, 2693–2728.

Carroll, R., Lewis, J., Lo, J., McCarty, N., Poole, K., Rosenthal, H., 2015. 'common space'dw-nominate scores with bootstrapped standard errors. September 2, 2015.

Chen, H., De, P., Hu, Y. J., Hwang, B.-H., 2014. Wisdom of crowds: The value of stock opinions transmitted through social media. The Review of Financial Studies 27, 1367–1403.

Cohen, L., Diether, K., Malloy, C., 2013. Legislating stock prices. Journal of Financial Economics 110, 574–595.

Cookson, J. A., Engelberg, J., Mullins, W., 2021. Echo chambers. Available at SSRN 3603107 .

Das, S. R., Chen, M. Y., 2007. Yahoo! for amazon: Sentiment extraction from small talk on the web. Management science 53, 1375–1388.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., Stroebel, J., 2020. Hedging climate change news. The Review of Financial Studies 33, 1184–1216.

García, D., Hu, X., Rohrer, M., 2020. The color of finance words. Available at SSRN 3630898 .

Gentzkow, M., Kelly, B., Taddy, M., 2019a. Text as data. Journal of Economic Literature 57, 535–74.

Gentzkow, M., Shapiro, J. M., 2010. What drives media slant? evidence from us daily newspapers. Econometrica 78, 35–71.

Gentzkow, M., Shapiro, J. M., Taddy, M., 2019b. Measuring group differences in high-dimensional choices: method and application to congressional speech. Econometrica 87, 1307–1340.

Giglio, S., Maggiori, M., Rao, K., Stroebel, J., Weber, A., 2021. Climate change and long-run discount rates: Evidence from real estate. The Review of Financial Studies 34, 3527–3571.

Gürkaynak, R. S., Sack, B., Swanson, E., 2005. Do actions speak louder than words? the response of asset prices to monetary policy actions and statements-ijcb-may 2005. International Journal of Central Banking .

Gürkaynak, R. S., Wright, J. H., 2013. Identification and inference using event studies. The Manchester School 81, 48–65.

Hoberg, G., Moon, S. K., 2019. The offshoring return premium. Management Science 65, 2876–2899.

Jensen, J., Naidu, S., Kaplan, E., Wilse-Samson, L., Gergen, D., Zuckerman, M., Spirling, A., 2012. Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech [with comments and discussion]. Brookings Papers on Economic Activity pp. 1–81.

Kelly, B., Pástor, L., Veronesi, P., 2016. The price of political uncertainty: Theory and evidence from the option market. The Journal of Finance 71, 2417–2480.

Kelly, B. T., Manela, A., Moreira, A., 2019. Text selection. Tech. rep., National Bureau of Economic Research.

Kuchler, T., Stroebel, J., 2021. Social finance. Annual Review of Financial Economics 13.

Lewis, J. B., Poole, K., Rosenthal, H., Boche, A., Rudkin, A., Sonnet, L., 2018. Voteview: Congressional roll-call votes database. See https://voteview.com/(accessed 27 January 2020) .

Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. The Journal of Finance 66, 35–65.

Martin, G. J., Yurukoglu, A., 2017. Bias in cable news: Persuasion and polarization. American Economic Review 107, 2565–99.

Mullainathan, S., Shleifer, A., 2005. The market for news. American economic review 95, 1031–1053.

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

Pástor, L., Veronesi, P., 2012. Uncertainty about government policy and stock prices. The journal of Finance 67, 1219–1264.

Pástor, L., Veronesi, P., 2013. Political uncertainty and risk premia. Journal of financial Economics 110, 520–545.

Pástor, L., Veronesi, P., 2020. Political cycles and stock returns. Journal of Political Economy 128, 4011–4045.

Rohde, D. W., 1991. Parties and leaders in the postreform House. University of Chicago Press.

Santa-Clara, P., Valkanov, R., 2003. The presidential puzzle: Political cycles and the stock market. The Journal of Finance 58, 1841–1872.

Snyder Jr, J. M., Groseclose, T., 2000. Estimating party influence in congressional roll-call voting. American Journal of Political Science pp. 193–211.

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

Tetlock, P. C., Saar-Tsechansky, M., Macskassy, S., 2008. More than words: Quantifying language to measure firms' fundamentals. The journal of finance 63, 1437–1467.

Wagner, A. F., Zeckhauser, R. J., Ziegler, A., 2018. Unequal rewards to firms: Stock market responses to the trump election and the 2017 corporate tax reform. In: *AEA Papers and Proceedings*, vol. 108, pp. 590–96.

Appendix A Additional Figures

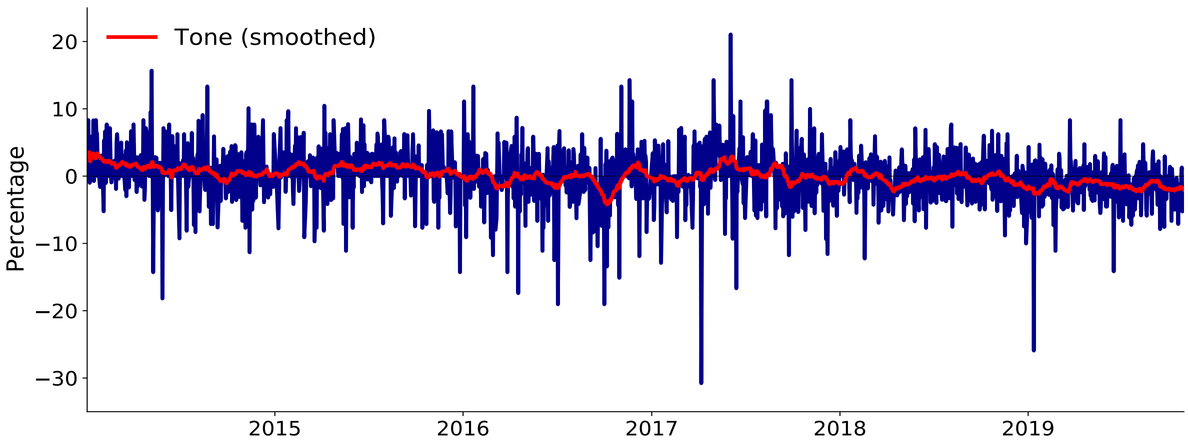

Fig. A.1. *Notes*: This figure shows the daily tone measure. To systematically compute the tone, lexicon developed by Loughran and McDonald (2011) is used. Using these dictionaries, we then count the number of positive and negative words that each tweet has. We define the $Tone$ measure as the difference between the positive and the negative word count scaled by the total number of words contained in the tweet.

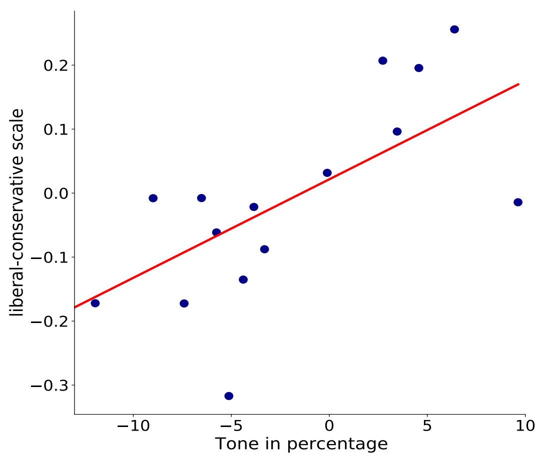

Fig. A.2. *Notes*: This figure shows the DW-nominate score as function of the tone measure. We sort the tone measure into bins and report the average tone measure and the DW-nominate score. The tone measure is computed over all tweets that explicitly reference the Crapo Bill (officially called Economic Growth, Regulatory Relief, and Consumer Protection Act (S.2115)) as identified by the following hashtags #S2155, #DoddFrankRollback, #BankLobbyistAct, and #Relief4MainStreet. The bill passed the Senate with an amendment by Yea-Nay vote of 67 - 31. The bill passed the House by Yea-Nay vote of 258 - 159. The DW-NOMINATE scores provide a measure of legislators' ideological locations over time by estimating how similar their voting records are. All variables are in percentage. The red line denotes the regression fit line.

Appendix B Additional Empirical Results

B.1 Alternative tone measures

In the main text, we use the lexicon developed by Loughran and McDonald (2011) to systematically compute the tone measure of the relevant tweets. In this appendix, we provide robustness results using two alternative tone measures. The first one uses the dictionaries proposed by García et al. (2020) which are based on a sophisticated machine learning algorithm. The main output from this algorithm is a set of n-grams that characterize the sentiment of a text as positive and negative. We use the positive/negative n-gram dictionaries (hereafter ML dictionaries) to sign the tweets. Specifically, we count the number of positive and negative n-gram for each tweet and scale this word count by the total number of words contained in the tweet. We use the positive/negative unigrams, and bigrams reported in the appendix in García et al. (2020). For the second tone measure, we manually classify the tweets as positive, negative, or neutral.

We first regress the high-frequency log price change on the first tone measure (ML tone measure). The regression specifications are described in the main text in equations (1), (2), and (3). Table B.1 reports the results. We document very similar results. A positive tone has a positive and statistically significant effect on the stock price as implied by the positive slope coefficient $\hat{b}$ equal to 0.39 (t-statistic = 2.5). Consistent with the results in the main text, we also find a zero effect for the placebo tweets. Note that the estimated coefficient $\hat{d}$ is not statistically significant.

We then use a discrete cutoff for the ML tone measure. We run specification (4) in the main text and present results in Table B.2. Consistent with the results reported in the main text, we find that the discrete specifications lead to stronger results when focusing on more stricter cutoffs. Notably, the t-statistics also tend to increase as we move to tails of the ML tone distribution. Column (4) reports results for when we manually classified the tweets. The estimated coefficients are larger (in absolute value) and the t-statistics also increase.

Table B.1. **High-frequency stock prices responses to congressional viewpoints: Using an alternative tone measure**

Coef	Variable	(1)	(2)	(3)
a	1	-0.410	-1.535	-1.229
		[-1.111]	[-1.058]	[-0.958]
b	$\text{Tone}_{i,t} \cdot D_{i,r}$	0.399	0.380	0.379
		[2.526]	[2.639]	[2.637]
c	$D_{i,r}$		3.037	2.713
			[0.795]	[0.746]
d	$\text{Tone}_{i,t} \cdot (1 - D_{i,r})$			-0.234
				[-0.724]
nObs		9,932	9,932	9,932
R2		0.184%	0.297%	0.401%

The independent variables are a combination of (a) the tone measure $\text{Tone}_{i,t}$; and (b) a dummy variable that equals one if the tweet is in fact relevant and zero otherwise $D_{i,r}$. To compute the $\text{Tone}_{i,t}$ we use the LM dictionaries of García et al. (2020). In all regressions, the dependent variable is the log change in company i's stock price in a 10 minute window around the tweet $\Delta p_{i,t}$. In all regressions, we use stock fixed effects. Standard errors are double clustered at the stock-day level and t-statistics are in brackets. The tone measure is in percentage and the log price change is in basis points.

Table B.2. **High-frequency stock price responses to a discrete measure of congressional viewpoints: Using an alternative tone measure**

	Distribution of ML Tone			Manual classification
Support equal to one if	> 0	$> 75\%$	$> 90\%$	
Criticize equal to one if	< 0	$< 25\%$	$< 10\%$	
	(1)	(2)	(3)	(4)
a	-0.798	-0.453	-0.339	[-0.901]
	[-0.953]	[-1.067]	[-1.023]	[-1.753]
Support	4.512	8.729	7.544	11.562
	[1.665]	[2.272]	[3.581]	[3.839]
Criticize	-0.685	-2.489	-3.943	-6.198
	[-0.442]	[-2.143]	[-1.980]	[-3.574]
nObs	9,932	9,932	9,932	9,932

This table reports coefficient estimates from the following Equation: $\Delta p_{i,t} = a + b^s \cdot Support_{i,t} + b^c \cdot Criticize_{i,t} + \epsilon_t$, where $\Delta p_{i,t}$ denotes the log change in company i's stock price in a 10 minute window around the tweet. The regressors are dummy variables indicating if the politician tweet is supporting or criticizing the targeted firm. In regression (1) the tweet supports (criticizes) company i if the tone measure is positive (negative). In regressions (2) and (3) the tweet supports (criticises) company i if the tone measure is above its 75th percentile (below its 25th percentile) and its 90th percentile (below its 10th percentile), respectively. In regressions (1), (2), and (3) we use the LM dictionaries of García et al. (2020) to compute the tone measure. In regression (4), we manually classify the tweets as positive, negative, or neutral. In all regressions, we use stock fixed effects. Standard errors are double clustered at the stock-day level and t-statistics are in brackets. The log price change is in basis points.

B.2 Trading Strategy

Table B.3. **Factor alphas for different look-back measurement periods**

	Look-back measurement periods in months								
	3	9	12	15	18	21	24	27	30
CAPM alpha	5.08	15.41	17.23	17.14	14.02	16.83	20.03	18.19	23.94
	[0.64]	[2.04]	[2.55]	[2.46]	[2.21]	[2.56]	[2.55]	[1.97]	[2.59]
Three-factor alpha	4.76	14.88	17.14	16.76	14.09	16.84	19.52	17.95	23.07
	[0.62]	[2.03]	[2.53]	[2.48]	[2.21]	[2.55]	[2.56]	[2.05]	[2.66]
Four-factor alpha	4.94	17.48	16.76	17.23	14.48	17.20	19.76	17.58	22.74
	[0.64]	[2.59]	[2.33]	[2.56]	[2.28]	[2.61]	[2.61]	[2.02]	[2.65]

This table reports daily factor alphas for the Long - Short portfolio. The long (short) portfolio buys stock i if the corresponding tone measure is above (below) the 75 (25) percentile of the tone distribution computed using different look-back measurement periods. Each column represents a different look-back measurement period. We value-weight the stocks to form portfolios whenever several firms are assigned to the same leg in the same day. The "Long-Short" portfolio buys the long portfolio and sells the short portfolio each day. The alphas denote the intercepts form time series regression of the portfolio excess returns on factor alphas. The four factors are the aggregate market excess return, the size factor, the value factor, and the momentum factor. Standard errors adjust for heteroskedasticity and autocorrelation. We report the t-statistics in brackets.

Table B.4. **Factor alphas for different thresholds in the trading signal**

	Distribution of Tone		
Signal to buy	> 0	$> 75\%$	$> 90\%$
Signal to sell	< 0	$< 25\%$	$< 10\%$
	(1)	(2)	(3)
CAPM alpha	15.87	17.23	25.73
	[2.26]	[2.554]	[2.97]
Three-factor alpha	15.92	17.148	24.92
	[2.37]	[2.536]	[3.01]
Four-factor alpha	15.77	17.488	24.45
	[2.35]	[2.590]	[3.01]

This table reports daily factor alphas for the Long - Short portfolio. The long (short) portfolio buys stock i if the corresponding tone measure is above (below) the certain percentile of the tone distribution computed over the previous 12 months. We value-weight the stocks to form portfolios whenever several firms are assigned to the same leg in the same day. The "Long-Short" portfolio buys the long portfolio and sells the short portfolio each day. The alphas denote the intercepts form time series regression of the portfolio excess returns on factor alphas. The four factors are the aggregate market excess return, the size factor, the value factor, and the momentum factor. Standard errors adjust for heteroskedasticity and autocorrelation. We report the t-statistics in brackets.

Table B.5. **Real effects of congressional viewpoints: Earnings per share**

	Forecast revisions i.e., $F_1^j - F_{-1}^j$		Forecast errors i.e., $Actual_j - F_{-1}^j$	
	(1)	(2)	(3)	(4)
Tone	0.0001	0.0003	0.0018	0.0019
	[1.5703[	[1.2914]	[3.8669]	[3.9055]
Tone placebo		0.0001		0.0004
		[0.1717]		[0.4950]
$r^{ab}_{i,t-45,t-1}$		0.0000		0.0000
		[10.0086]		[5.0006]
log(Size)		0.0081		0.3052
		[1.3046]		[23.6111]
log(Assets)		0.0090		-0.3104
		[1.5702]		[-26.4367]
Obs	22736	21264	51971	48664

This table reports coefficient estimates from the following equation $y^j_{i,t} = a + b \cdot D_r \cdot \text{Tone}_{i,t} + c \cdot (1 - D_r) \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}$. In Columns (1) and (2) the dependent variable is the price-scaled earnings per share forecast revision $FR^j_{i,t}$. In Columns (3) and (4) the dependent variable is the price-scaled earnings per share forecast error $FE^j_{i,t}$. Figure 7 in the main text provides the details on how we compute these measures. The regressor of interest is $\text{Tone}_{i,t}$ which denotes the average tone measure for all tweets that target company i on day t. As $Controls$ we include the stock i abnormal cumulative return from $t-45$ to $t-1$, and measures of firm size and assets. We also include year-month and industry fixed effects. Standard errors are double clustered at the industry and year-month level and t-statistics are in brackets.

B.3 Real Effects

Table B.6. **Real effects of congressional viewpoints: Sales**

	Forecast errors i.e., $Actual_j - F^j_{-1}$ Days used to compute F^j_{-1}		
	15 days (1)	30 days (2)	60 days (3)
Tone	4.3732	4.3574	3.0355
	[2.5057]	[3.8443]	[2.8165]
Tone. placebo	0.6064	-1.3119	-0.6603
	[0.2263]	[-0.8573]	[-0.5016]
Controls	Yes	Yes	Yes
R2	1.8657	1.4523	1.4010
Obs	27889	44831	49176

This table reports coefficient estimates from the following equation $FE^j_{i,t} = a + b \cdot D_r \cdot \text{Tone}_{i,t} + c \cdot (1 - D_r) \cdot \text{Tone}_{i,t} + \delta \cdot Controls + \epsilon_{i,t}$.The dependent variable is the price-scaled earnings per share forecast error $FE^j_{i,t}$. Each column presents results for a different window used to compute the analysis consensus before the tweet. Column (1) uses 15 days. Column (2) uses 30 days. Column (3) uses 60 days. In the main text we use as a benchmark 45 days. Figure 7 in the main text provides the details on how we compute this measure. The regressor of interest is $\text{Tone}_{i,t}$ which denotes the average tone measure for all tweets that target company i on day t. As *Controls* we include the stock i abnormal cumulative return over the same sample period used to compute the consensus forecast, and measures of firm size and assets. We also include year-month and industry fixed effects. Standard errors are double clustered at the industry and year-month level and t-statistics are in brackets.

B.4 Illustrative Example: Tax Cuts and Jobs Act

In this section we focus on the Public Law 115–97, commonly referred to as the Tax Cuts and Jobs Act (henceforth TCJA) or the 2017 tax reform.[8] This tax reform is used as an other illustrative case of how members of congress can influence expectations about policy by communicating their stance in real time through their social media accounts. The 2017 tax reform is one of the most important bills in our sample since it made the largest changes to the US tax code in over thirty years. For example, the tax reform had an important impact on firms given that the major changes were to the US corporate tax system, including a reduction in the federal corporate tax rate from 35 percent to 21 percent.[9] This reform moved swiftly through the legislative process taking less than three months from the release of a nine-page "Unified Framework for Tax Reform" on September 27, 2017 to a nearly 200 page final bill signed into law by President Trump on December 22, 2017. The bill was extensively revised as it was rushed by Republicans through the House and Senate generating substantial uncertainty both in the actual content of the bill and on whether it would pass (Wagner, Zeckhauser, and Ziegler, 2018).[10] The uncertainty remained until the passing of the bill. Importantly, members of congress actively used their social media accounts to communicate their stance about these changes in real time. For instance, Democrat Senator Dick Durbin (@SenatorDurbin) posted "Trying to review the #GOPTaxScam but they are making hand-written changes to brand new text as we speak – can anyone else read this?" [attached a screenshot of a page of the bill with the changes], 1 Dec, 2017, 23:25:27 EST, Tweet.

B.4.1 The legislative process of the Tax Cuts and Jobs Act

The legislative process of the Tax Cuts and Jobs Act directly is captured from the Twitter accounts of the members of congress. The main advantage of this approach is that congressional social media posts provide representative snapshots of their viewpoints at high frequencies. Importantly, we show that these partisan viewpoints about policy generate significant revisions in market expectations.

To identify the legislative process of the Tax Cuts and Jobs Act, we start by building an index of tax related terms based on our panel of congressional social media posts. All posts that contain a tax related word or hashtag are counted, such as "tax", "taxation", and "#TrumpCuts" in a given day. This raw count is then scaled by the total number of congressional tweets posted in the same day. Figure B.3 depicts the resulting index.

[8] The short title "Tax Cuts and Jobs Act" was not approved by Senate and to comply with Senate rules, the official title of the bill was changed to "An Act to provide for reconciliation pursuant to titles II and V of the concurrent resolution on the budget for fiscal year 2018."

[9] Auerbach (2018) provides a detail explanation of the main changes of the the Tax Cuts and Jobs Act. The full set of changes are in `https://www.congress.gov/bill/115th-congress/house-bill/1`.

[10] It is important to note that the Republican party did not have 60 or more votes in the Senate to pass the bill over a Democratic filibuster. However, the Republicans passed the tax cuts via a procedural maneuver known as budget reconciliation. This fast-track process allows the bill to be passed by majority vote as long as the bill does not increase the deficit in the next decade. For details on this see Alex Tausanovitch & Sam Berger, Center for American Progress, TheImpact of the Filibuster on Federal Policymaking (Dec. 5, 2019), available at `https://www.americanprogress.org/issues/democracy/reports/2019/12/05/478199/impact-filibuster-federalpolicymaking/`.

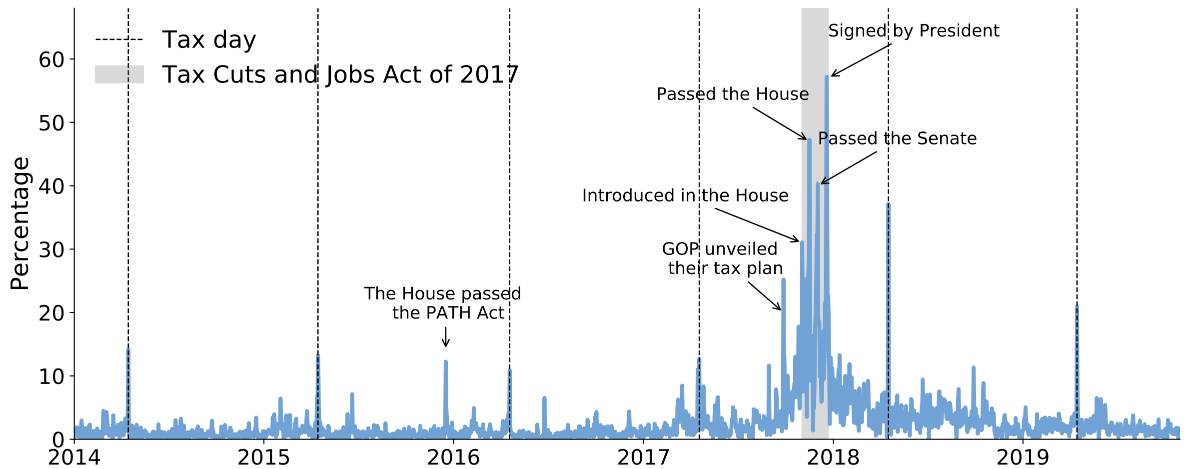

Fig. B.3. *Notes*: This figure shows the percentage of congressional post containing tax related word or hashtag, such as "tax", "taxation", and "#TrumpCuts" in a given day. The gray shaded area in the figure highlights the time the bill spent in Congress till it became Public Law. We use the complete set of tweets created by any member of the U.S. Senate and House of Representatives from the 113th Congress to the 116th Congress.

The gray shaded area in the figure highlights the time the bill spent in Congress until it became Public Law.

The figure shows significant spikes on the main legislative events of the tax reform. Republicans unveiled their tax plan on September 27. On November 2, the House Ways and Means Committee introduced the bill, which was then passed on November 16. On the same day, the Senate Finance Committee passed a draft of the bill, which was subsequently passed by the Senate in the early hours of December 2. After reconciling the difference between the House and Senate bills, the final version of the bill passed each chamber in a mostly party line vote. Finally, the President signed the bill into law on December 22. The figure also displays other events such as the Tax day (dotted vertical lines) which usually falls on the 15 of April of each year. In the U.S., the Tax Day denotes the due date on which individual income tax returns should be submitted to the federal government.

B.4.2 Corporate taxes and the aggregate market

Next, we use a high-identification approach and focus on movements in asset prices in a short time window around the tweets that communicate the congressional viewpoints about the 2017 tax reform. The key finding is that these viewpoints provide real-time updates to market participants about the stance of each party to the continually changing provisions made to the bill.

We proceed in two steps. In the first step, we select posts that explicitly express a view about the tax reform, and classify them as being supportive or critical of the tax framework. Members of congress primarily used four hashtags (#TaxReform, #TaxCutsandJobsAct, #TrumpCuts, #GOPTaxScam) to engage with the public and express their viewpoints about the GOP's tax bill. Therefore, to select the tax reform tweets posts are selected that contain any of these four hashtags. All of these tweets that do not directly express a view about the reform are dropped. We then classify the remaining tweets. To do so, for each politician we cross-referenced the use of hashtags with the *ex*-post voting behavior. It turns out that 92% and 98% of the tweets containing the hashtags #TaxReform and #TaxCutsandJobsAct, respectively, were posted by a legislator who voted in support of the bill. Conversely, 100% of the tweets that contained #TrumpCuts and #GOPTaxScam came from a member of congress who voted against the bill.[11] Therefore, to avoid subjectivity in the classification procedure, we classify a tweet being in support of the bill if it contains #TaxReform or #TaxCutsandJobsAct, and opposing the bill if it contains either #TrumpCuts or #GOPTaxScam.

In the second step, we measure the average effect of the politician tweet that is either in support or opposes the tax reform on the aggregate stock market. Specifically, for the aggregate stock market we take the intraday prices of the exchange-traded fund that tracks the S&P 500 stock market index (henceforth, SPY) which is obtained from the TAQ database. To clean the raw tick-by-tick series, the same procedure as in Section 2 is followed and a 90-second window around the tweet is used. Similarly, only the tweets posted during normal NYSE trading hours are used, which begin at 9:30 a.m. EST and end at 4 p.m EST.

Table B.7 presents the results. All estimated coefficients are in basis points. In regression 1 we estimate the following model of aggregate stock prices:

$$\Delta p_t = a + b^{s/c} \cdot t_{i,t}^{s/c} + \epsilon_t, \tag{B.1}$$

where Δp_t denotes the 90-second log aggregate stock price change and $t_{i,t}^{s/c}$ is a dummy variable that equals one if politician i tweeted at time t that was supportive or critical of the tax reform. The intercept a captures the average effect of the policy-related economic topics. When we do not condition on the direction of the tweet, the estimated coefficient $b^{s/c}$ is economically small and statistically insignificant (t-statistic = 0.59).

Regression 2 in Table B.7 conditions on the direction of the tweet:

$$\Delta p_t = a + b^s \cdot t_{i,t}^s + b^c \cdot t_{i,t}^c + \epsilon_t. \tag{B.2}$$

$t_{i,t}^s$ and $t_{i,t}^c$ are dummy variables indicating if the politician tweet is supportive or critical of the bill. A striking change is evident in the estimated coefficients. Social media post that are supportive of the tax reform increase on average valuations on the aggregate stock market index. It is 0.37 bps higher, with a t-statistic of 2.11 on the difference. Conversely, tweets critical of the reform decrease valuations. The effect is somewhat smaller of around -0.22, , but not statistically significant (t-statistic = -1.4).

Next, we add one additional element to the regression to assess the extent to which the informational content of the congressional tweets differ on crucial days of the legislative process:

$$\Delta p_t = a + \left(b^s + b^{s-imp} \cdot I_{imp,t}\right) \cdot t_{i,t}^s + \left(b^c + b^{c-imp} \cdot I_{imp,t}\right) \cdot t_{i,t}^c + \epsilon_t, \tag{B.3}$$

[11] We did not use directly the party of the politician to classify the tweets because it was not entirely a party-line vote. Although, no Democrat supported the bill, there were 13 Republicans who voted against it.

Table B.7. **Effect of politicians on returns: Tax related tweets**

Coefficient	Variable	Estimated coefficients [t-statistics]				
		(1)	(2)	(3)	(4)	(5)
a	1	0.0164	0.0164	0.0164	0.0164	0.0164
		[14.448]	[14.541]	[14.522]	[14.587]	[14.573]
$b^{s/c}$	$t^{s/c}_{i,t}$	0.0746				
		[0.5922]				
b^{s}	$t^{s}_{i,t}$		0.3729	0.3387	0.2120	0.1585
			[2.1108]	[1.8069]	[1.0051]	[0.7080]
b^{c}	$t^{c}_{i,t}$		-0.2237	-0.1567	-0.1889	-0.1234
			[-1.4157]	[-0.8312]	[-1.0589]	[-0.5916]
b^{s-imp}	$I_{imp,t} \cdot t^{s}_{i,t}$			0.6251		0.7613
				[1.6560]		[1.9182]
b^{c-imp}	$I_{imp,t} \cdot t^{c}_{i,t}$			-0.4745		-0.4703
				[-2.0457]		[-2.0313]
b^{s-inf}	$I_{inf,t} \cdot t^{s}_{i,t}$				0.5866	0.6297
					[1.9035]	[2.0241]
b^{c-inf}	$I_{inf,t} \cdot t^{s}_{i,t}$				-0.2876	-0.2795
					[-1.0641]	[-0.9904]
nObs		357,815	357,815	357,815	357,815	357,815

This table reports the effect of tweets that explicitly express a view about the 2017 tax reform on the aggregate stock market. In all regressions, the dependent variable variable is the 90-second log stock price of the SPY ETF. The independent variables are a constant, $t^{s}_{i,t}$ and $t^{c}_{i,t}$ are dummy variables indicating if the politician tweet is in support or opposes the bill, an indicator for important days $I_{imp,t}$, which equals one if the tweet was posted on an important day in the legislative process, and an indicator, $I_{inf,t}$, that equals one if the tweet was posted by a politician that co-sponsored the bill or if she/he belonged either to the Committee on Ways and Means or to the Senate Finance Committee. We report t-statistics in brackets and all estimated coefficients are in basis points.

where $I_{imp,t}$ is an indicator that equals one if the tweet was posted on an important day in the legislative process. The important days in the legislative process of the 2017 tax reform are: (1) Republicans unveil their tax plan on the 27 of September of 2017; (2) The bill was introduced in the house (11/02/2017); (3) passed/agreed to in House (11/16/2017); (4) Passed/agreed to in Senate (12/02/2017); (4) Resolving differences between the House and Senate (12/20/2017); and (5) Signed into law by the President (12/22/2017).

Column (3) of Table B.7 presents the results. The estimates b^{s-imp} of 0.62 (t-statistic= 1.65) and b^{c-imp} of -0.47 (t-statistic= -2.04) show that a disproportionate amount of news about the provisions of the tax reform are revealed during these important days. During these days, the effect of supporting viewpoints is 0.98 bps ($= a + b^{s} + b^{s-imp}$), while the effect of opposing viewpoints is -0.61 ($= a + b^{c} + b^{c-imp}$).

Regression 4 modifies the previous specification to evaluate whether more influential politicians also have a larger market impact:

$$\Delta p_t = a + \left(b^s + b^{s-inf} \cdot I_{inf,t}\right) \cdot t^s_{i,t} + \left(b^c + b^{c-inf} \cdot I_{inf,t}\right) \cdot t^c_{i,t} + \epsilon_t, \tag{B.4}$$

where $I_{inf,t}$ is an indicator that equals one if the tweet was posted by a politician that co-sponsored the bill or if they belonged either to the Committee on Ways and Means or to the Senate Finance Committee, which are the government bodies in the House and Senate, respectively, in charge of making provisions to the tax reform. Column (4) highlights that positive viewpoints of more influential politicians have an average effect of 0.81 bps. In contrast, the point estimate effect of negative viewpoints is - 0.46 bps, but not statistically significant

B.5 Policy-related economic roll call votes and social media post.

In this appendix, we relate tweets to congressional votes. We document a high volume of tweets related to proposed legislation occurring on the corresponding roll call days.

To link tweets to congressional votes, we proceed in three stages. First, we capture the subset of social media posts by members of congress that contain their views about policy-related economic topics. To this end, we begin by selecting posts that contain at least one keyword associated with the following six categories: (1) government budget; (2) homeland security; (3) energy, commerce, and labor; (4) regulation; (5) corporations; and (6) healthcare. Appendix B.5.1 reports the complete set of keywords included in each topic. Broadly, these posts capture directly the viewpoints of politicians and cover a wide range of topics, including opinions about tax policy (e.g., Rep Paul Ryan (SpeakerRyan) "An army of lobbyists will come to protect special interest provisions & derail #TaxReform. But conservatives cannot shrink from this moment." 12 Oct 2017, 14:05:44 EST, Tweet.) to stances about specific companies (e.g., Dem Elizabeth Warren (@ewarren) "Amazon has too much power. Under my plan to #BreakUpBigTech , entrepreneurs would have a fighting chance to compete against tech giants." 25 June 2019, 15:43:43, Tweet.)

In the second stage, the roll call votes taken during the 113th-116th Congresses are linked to the policy-related economic topics. To do this, for each roll call we use their corresponding title or brief descriptions as compiled through the electronic voting machines[12] and identify roll calls that include at least one of the topic keywords described in Appendix B.5.1. Broadly, these roll calls capture key bills and motions related to tax policy, budget resolutions, the national debt limit, and social welfare policy. Interestingly, these roll calls address core issues that divide parties as manifested in the ex-post vote outcomes.

A common measure of the partisan divide based on roll call voting is the *party difference measure* (e.g., Rohde, 1991; Snyder Jr and Groseclose, 2000). The party difference for a given roll call j is

$$\text{RollCall}_j = \left| \frac{1}{\#D} \sum_{i \in D} v_{ij} - \frac{1}{\#R} \sum_{i \in R} v_{ij} \right| \times 100\% \quad \text{with} \quad v_{ij} = \begin{cases} 1, & \text{if legislator } i \text{ votes yea} \\ 0, & \text{if legislator } i \text{ votes nay} \end{cases}$$

where D and R denote the set of Democratic and Republican legislators, respectively. A value close to one implies that the vote on roll call j was lopsided across party lines. Whereas a value close to zero implies that the vote was bipartisan, with a similar fraction of yea votes in the two parties. The left panel of Figure B.4 plots the party difference measures for roll calls linked to the policy-related economic topics (blue line) and any other roll call (gray dashed line). The series is smoothed by taking the average of the party difference measures in a given quarter. The basic finding is that policy-related economic roll-call votes are more contested, where the average difference is 20 percentage points.

Next, we select tweets by members of congress that contain the relevant economic keywords and were posted on the same day as the roll calls recorded in the first stage. The

[12] This data can be download in a suitable format for analysis from Lewis, Poole, Rosenthal, Boche, Rudkin, and Sonnet (2018).

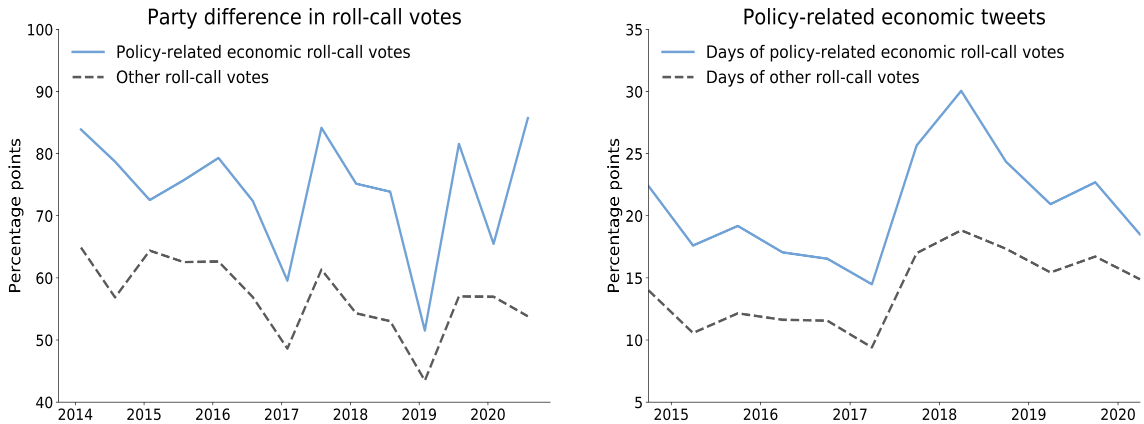

Fig. B.4. *Notes*: The left panel shows the party difference measure based on roll calls. This measure is defined as the absolute value of the difference between the percentage of Republicans who voted yea and the percentage of Democrats who voted yea. We depict this measure for roll calls related to policy-economic topics (blue line) and other type of roll calls (gray dashed line). We smooth the series by taking the average of the party difference measures in a given quarter. The right panel shows the total number of policy-related tweets as percentage of the total number of congressional tweets. We further distinguish between tweets posted during days in which a policy-related economic roll call was voted (blue line) or not (gray dashed line). Finally, we smooth both series by aggregating the daily time series to quarterly series. We use the complete set of tweets created by any member of the U.S. Senate and House of Representatives from the 113th Congress to the 116th Congress.

right panel of Figure B.4 plots the total number of policy-related tweets as percentage of the total number of congressional tweets. We further distinguish between tweets posted during days in which a policy-related economic roll call was voted (blue line) or not (gray dashed line). Finally, both series are smoothed by aggregating the daily time series to quarterly series. The figure shows that there is a high volume of tweets related to proposed legislation occurring on the corresponding roll call days. This increase in tweets is also large in magnitude. For instance, over the entire sample, members of congress as a whole increase their policy-related social media content by about 10 percentage points. Overall, this spike in tweet intensity is consistent with the findings in the illustrative examples documented in Section 4.4 in the main text and in Appendix B.4.

B.5.1 Words used in the policy-related economic topics and roll calls

Next we present the keywords that we used for each category. To assign the hashtags to each category we proceeded in two steps. First, we search for the most common hashtags used by all members of congress. Second, we manually classify each one of these common hashtags into one category.

- **Government budget:** taxes, tax, taxation, taxed, tariff, tariffs, government spending, federal budget, budget battle, balanced budget, defense spending, military spending, entitlement spending, fiscal stimulus, budget deficit, federal debt, national debt, debt ceiling, fiscal footing, government deficits, balance the budget, fiscal stimulus, debt ceiling, debt limit, welfare, food stamps, AFDC, tanf, oasdi, earned income tax credit, EITC, public assistance, head start program, entitlement program, wic program, government subsidies, deficit, budget, national debt, federal debt, sovereign debt, government policy, public policy, government spending, government expenditures, entitlement spending, entitlement expenditures, unemployment insurance, unemployment benefits, disability insurance, disability benefits, welfare reform, fiscal stimulus, Fiscal stimulus, fiscal policies, fiscal policy, fiscal reform

  - **Related hashtags:** #TaxDay, #taxreform, #TaxReform, #TaxCutsandJobsAct, #TrumpCuts, #GOPTaxScam, #IRS,#TrumpShutdown, #TrumpBudget, #SchumerShutdown, #EndTheShutdown, #budget, #DontShutDownOurSecurity, #StandWithPP.

- **Homeland security:** national security, war, military conflict, terrorism, terror, defense spending, military spending, police action, armed forces, base closure, military procurement, saber rattling, naval blockade, military embargo, no-fly zone, military invasion, national security, military invasion, military conflict, military embargo, military procurement, war, armed forces, police action, base closure, saber rattling, naval blockade, no-fly zone, defense spending, defense expenditures, military spending, military expenditures

  - **Related hashtags:** #NDAA, #SecDef, #ISIL, #ISIS, #Benghazi, #Ukraine, #NorthKorea, #Syria, #Iran, #IranDeal.

- **Energy, commerce, and labor:** carbon tax cap, pollution controls, environmental restrictions, clean air act, clean water act, energy policy, drill restrict, import tariffs, import duty, import barrier, government subsidies, government subsidy, wto, world trade organization, trade treaty, trade agreement, trade policy, trade act, doha round, uruguay round, gatt, dumping, trade policy, trade act, trade agreement, trade treaty, duty, duties, import tariff, import barrier, minimum wage, minimum wage, union rights, card check, national labor rel. board, nlrb, collective bargaining, right to work, closed shop, worker compensation, maximum hours, wages and hours, advanced notice requirement, affirmative action, overtime requirements, at-will employment, equal employment opportunity, eeo, osha, immigration, unemployment insurance, unemployment benefits,

  - **Related hashtags:** #JobsReport, #jobs, #RaiseTheWage, #climatechange, #ClimateChange, #energy, #GreenNewDeal, #ParisAgreement, #ActOnClimate, #EPA, #KeystoneXL, #China, #Russia, #USMCA, #TPP.

- **Regulation:** banking supervision, bank supervision, glass-steagall, tarp, thrift supervision, dodd-frank, financial reform, commodity futures trading commission, cftc, house financial services committee, basel, capital requirement, Volcker rule, bank stress test, securities and exchange commission, sec, deposit insurance, fdic, fslic, ots, occ, firrea, truth in lending, union rights, card check, collective bargaining law, national labor relations board, nlrb, minimum wage, living wage, right to work, closed shop, wages and hours, workers compensation, prescription drugs, drug policy, food and drug admin, FDA, Gramm- Rudman, Bank supervision, thrift supervision, malpractice reform, constitutional re- form, financial reform, medical insurance reform, welfare reform, tort reform, constitutional amendment, Glass-Steagall, Dodd-Frank, housing financial services committee, capital requirement, security exchange commission, sec, deposit insurance, fdic, fslic, ots, occ, firrea, truth in lending, monometallist, bimetallist, silver coinage, gold coinage, alcohol prohibition, liquor prohibition, bill

  - **Related hashtags:** #EqualityAct, #RenewUI, #VAWA, #FarmBill, #NetNeutrality, #EqualPay, #SCOTUS, #EnoughIsEnough, #EndGunViolence, #gunviolence, #DisarmHate, #NoBillNoBreak, #SOTU.

- **Healthcare:** health care, Medicaid, Medicare, health insurance, malpractice tort reform, malpractice reform, prescription drugs, drug policy, food and drug administration, FDA, medical malpractice, prescription drug act, medical insurance reform, medical liability, affordable care act, Obamacare, health care, social security, Medicare, Medicaid, affordable care act, nutritional assistant program, health insurance, health benefits, medical insurance reform, constitutional reform

  - **Related hashtags:** #Trumpcare, #TrumpCare, #MedicareForAll, #ProtectOurCare, #AHCA, #GetCovered, #ACA, #healthcare, #Medicaid, #Obamacare, #SocialSecurity, #ObamaCare, #GrahamCassidy, #PayMoreForLess, #CHIP