Firm Inattention and the Efficacy of Monetary Policy: A Text-Based Approach

Wenting Song[*] Sam Stern[†]
University of Michigan University of Michigan

November 6, 2020

Click here for the latest version

Abstract

This paper studies the empirical importance of information frictions on monetary non-neutrality. We construct a text-based measure of firm attention to macroeconomic news and document firm attention that is countercyclical and polarized. Differences in attention lead to asymmetric responses to monetary policy: expansionary monetary shocks raise stock returns of attentive firms more than those of inattentive firms, and contractionary shocks lower returns of attentive firms by less. We interpret the findings using a quantitative model of rationally inattentive firms and calibrate parameters for information frictions using our text-based measure. In the model, firms invest in attention endogenously and face heterogeneous information costs. Less attentive firms adjust prices slowly in response to monetary innovations, which yields non-neutrality. As average attention varies over the business cycle, so does the efficacy of monetary policy.

JEL: D83, E44, E52

Keywords: Rational inattention, monetary policy, natural language processing

[*]Department of Economics, University of Michigan, wtsong@umich.edu. We would like to thank John Leahy and Pablo Ottonello for advice, guidance, encouragement and support. We also thank Alberto Arredondo-Chavez, Dominick Bartelme, Martin Eichenbaum, Xing Guo, Andrei Levchenko, Frank Jiaze Li, Alistair Macauley, Michael McMahon, Linda Tesar, Aditi Thapar, Andrew Usher, Toni Whited, and seminar participants at Michigan for helpful comments and discussions. We acknowledge financial support from the University of Michigan through MITRE Research Award, Rackham Research Grant, and Library Data Grant.

[†]Department of Economics, University of Michigan, sternsa@umich.edu.

1 Introduction

Public information often goes unused because attention is scarce. Rational inattention models pioneered by Sims (2003) consider agents that allocate attention to state variables while facing cognitive costs of processing information. Paying attention to certain state variables reduces uncertainty about those states and allows agents to set choice variables closer to their optimal values. Firm managers face similar trade-offs as they allocate attention to maximize firm values. Empirically identifying attention in this framework is challenging because neither a firm's allocation of attention nor information-processing costs are readily observable.

In this paper we introduce a novel measure of firms' attention to macroeconomic news using textual analysis on SEC filings. To construct the attention measure, we search through almost 200,000 annual filings of US publicly-traded firms for macroeconomic keywords. We define two measures of attention: "prevalence", whether firm managers discuss macro conditions at all, and "intensity", the frequency at which managers discuss macro conditions.

We begin by documenting four stylized facts about firm attention. First, firm attention is polarized. The majority of firms in our sample either mention macroeconomic conditions in every filing or in none of their filings. Second, attention to macro news is countercylical. Among the remaining firms with time-varying attention, the number of firms that mentioned macroeconomic news rose notably during the recessions in 2001 and 2007. Third, firms in a given industry are most attentive to macroeconomic topics that are important to their industry a priori. Construction companies follow housing starts; mining, oil, and gas companies follow oil prices; and retail trade companies follow consumer confidence. Finally, attention rises with firm age and size.

Our main empirical results demonstrate an unusual asymmetry that we derive from rational inattention models: expansionary monetary shocks raise stock returns of attentive firms more than those of their inattentive peers, and contractionary shocks lower returns of attentive firms by less. This asymmetry appears uniquely consistent with a theory of attention under imperfect information and rules out alternative interpretations of our attention measure, such as a measure of a firm's profit exposure to macroeconomic conditions, that should yield symmetric responses to positive or negative monetary shocks. To arrive at this result, we combine our attention measure with CRSP stock return data, quarterly Compustat firm financials, and high-frequency monetary shocks constructed as in Gürkaynak et al. (2005). Our main empirical model regresses daily firm stock price changes on firm attention interacted with both positive and negative high-frequency monetary shocks. We control for firm characteristics and include industry fixed effects to study otherwise similar firms within narrowly defined industry groups, and cluster standard errors by FOMC announcement to allow for correlated errors at each announcement.

Motivated by the empirical finding of heterogeneous firm attention and its effects on monetary transmission, we construct a quantitative model of rational inattention in which firms trade off the precision of their signals of aggregate demand against a cost of acquiring and processing information. Firms invest in attention endogenously and are heterogeneous in their information cost. We use empirical moments to calibrate key parameters of the model governing the severity of information frictions, including the fraction of attentive firms and the heterogeneous costs of information. In the calibrated model, attentive firms have higher semi-elasticities to expansionary monetary shocks and lower semi-elasticities to contractionary shocks, consistent with the data.

We then apply use this model to study aggregate implications of attention on monetary policy. Empirically, firm attention rises during recessions and displays a countercyclical pattern. We quantify the effects of average attention on the efficacy of monetary policy by varying the fraction of attentive firms in the model. As the fraction of attentive firms increases, more firms set prices closer to the optimum, monetary non-neutrality weakens, and efficacy of monetary policy declines. This new interpretation of attention-dependent monetary policy has an important implication for policy implementation: central banks should expect the effects of their actions to be weaker when responding to aggregate shocks that have already raised average firm attention to macroeconomic policy.

Related Literature Our paper contributes to four strands of literature. First, we contribute to the literature on information frictions by quantifying the empirical severity of information frictions. Our results shed light on the empirical importance of both sticky information models (Reis, 2006; Mankiw and Reis, 2002), which are successful in matching dynamics of monetary policy and help to reconcile micro and macro evidence (Auclert et al., 2020), and rational inattention models (Sims, 2003, 2010), which microfound the information-acquisition process and provide useful applications in explaining firm pricing (Woodford, 2009; Mackowiak and Wiederholt, 2009) and asset pricing (Van Nieuwerburgh and Veldkamp, 2009), and discrete choices (Matějka and McKay, 2015; Caplin et al., 2019). In particular, Afrouzi and Yang (2019) model firms' price setting under rational inattention and are successful in matching stylized facts on firm responses to monetary policy. Contributing to the literature, our paper provides evidence of firms displaying inattentive behavior, and we show empirically that firm inattention is important for monetary non-neutrality.

Second, we contribute to the empirical literature on macroeconomic expectations by constructing a direct measure of firm attention. Recent literature has highlighted the importance of expectations for macroeconomic policy[1]. Existing measures of attention come in roughly three categories[2]: deviations from model optimality (McCaulay, 2020), lab evidence such as eye-tracking (Reutskaja et al., 2011), and survey evidence. The first two measures quantify inattention in individuals but are not available for firms. Tanaka et al. (2019) conduct survey on the expectations of Japanese firms and Coibion et al. (2018) and Afrouzi (2020) survey the expectations of New Zealand firms. To our knowledge, similar survey evidence for US firms does not exist. Our methodology is able to provide a direct measure of attention for a large panel of firms on an ongoing basis.

Third, we relate to the literature of state dependency of monetary policy. Tenreyro and Thwaites (2016) estimate non-linear responses in monetary policy which are weaker in recessions than in expansions. Vavra (2014), McKay and Wieland (2019) and Ottonello and Winberry (2018) model channels from which the state dependency arises. We measure the importance of attention which leads to a new source of state dependency of monetary policy.

Finally, our paper relates methodologically to a broader and emerging literature that applies natural language processing techniques to economics. The seminal work of Loughran and McDonald (2011) applies the "bag of words" method from textual analysis to firm filings and develops word lists specific to economic and financial texts. Recent works have

[1]See, for example, Coibion and Gorodnichenko (2015); Coibion et al. (2020); Malmendier and Nagel (2016)

[2]See Gabaix (2019) for a comprehensive survey.

used textual analysis to measure financial constraints (Buehlmaier and Whited, 2018), central bank communication (Hansen et al., 2018), and firm-level political risk (Hassan et al., 2016) and uncertainty (Handley and Li, 2020). We contribute to the literature by constructing a dictionary of macroeconomic keywords with detailed categories based on releases of macroeconomic series.

Road map The rest of the paper proceeds as follows: in Section 2 we describe our methodology for measuring attention and present evidence of the stylized facts listed above; in Section 3 we present a theoretical framework that incorporates attention to FOMC announcements; in Section 4 we outline an empirical strategy for testing the effects of attention on expected returns and present our results; in Section 5 we construct a quantitative model of rational inattention and conduct policy counterfactuals; in Section 6 we discuss limitations of our measures and mitigation. Section 7 concludes.

2 Textual Measure of Attention

This section presents our measure of firm attention to macroeconomic news for the universe of US publicly-traded firms between 1994 and 2019. We then document several stylized facts about firm attention before conducting the main empirical analysis in Section 4.

2.1 SEC filings

To measure firm attention, we employ the universe of annual 10-K filings with the U.S. Securities and Exchange Commission (SEC) between 1994 and 2019. Under Regulation S-K, all public companies are required to disclose financial statements and business conditions in these filings. The annual filings (Form 10-K) requires a more extensive discussion of business conditions and audited financial statements, while the quarterly filings (Form 10-Q) is usually less descriptive and only requires unaudited financial statements. Our sample contains 201,751 unique annual 10-K filings by 35,655 firms. Table 1 shows the summary statistics on the 10-K filings. The average length of 10-Ks is 30,647 words with 2,433 unique words.

Table 1: Summary Statistics on 10-K filings

	N	Mean	Median	SD	Min	Max
Total word count	201,751	30,647	26,133	23,031	152	199,520
excl. stopwords	201,751	18,912	16,128	14,232	98	164,734
Unique word count	201,751	2,433	2,496	1,039	74	7,937
excl. stopwords	201,751	2,337	2,395	1,026	68	7,822

Discussion of economic conditions in an SEC filing typically appears in two contexts: recent or future firm performance and the risk factors that shareholders face by investing in the company. The former context usually appears in Item 7 of 10-K and 10-Q filings, which requires managers to discuss and analyze the firm's financial conditions and results of operations. This section is written as a narrative and can vary in length across firms (for instance, Item 7 of Alphabet's 2020 10-K filing is 17 pages long). Economic conditions in the context of risk factors commonly appears in Items 1A and 7A, which detail general firm risks and near-term market risks, respectively.

2.2 Methodology

Textual measure of firm attention To construct our main measures of firm attention to macroeconomic news, we employ dictionary-based frequency counts in natural language processing. We identify instances in which firms discuss the following nine macroeconomic topics: general economic conditions, output, labor market, consumption, investment, monetary policy, housing, and oil. Each topic is matched with a keyword dictionary that consists of names of major macroeconomic releases from Econoday (the data provider behind Bloomberg's economic calendar) as well as words and phrases that commonly appear in popular articles on each topic. Any words or phrases that might apply to both aggregate- and firm-specific conditions are removed to avoid misidentification. For example, the phrase "interest rates" is excluded from the monetary policy dictionary because firms may mention interest rates in the context of their own liabilities. The dictionary of topics and associated keywords appears in Table A.1.

We then construct two measures of attention based on these keywords. Attention *prevalence*, d_{it}^k, indicates whether a firm i mentioned any keyword related to a given topic k in period t:

$$d_{it}^k = \mathbb{1}(\text{Total topic-k words}_{it} > 0) \tag{prevalence}$$

Attention *intensity*, s_{it}^k, records the rate at which keywords are mentioned as a share of total words in the filing. We interpret this measure as the average intensity with which firms pay attention to economic conditions:

$$s_{it}^k = \frac{\text{Total topic k words}_{it}}{\text{Total words}_{it}} \tag{intensity}$$

Total word count is generated by following the parsing strategy in Loughran and McDonald (2011). First, a text is stripped of all numbers and "stop words" such as articles. The text is then mapped onto a dictionary of words constructed by extending 2of12inf, a commonly-used collection of English words, to include additional words in 10-K documents.

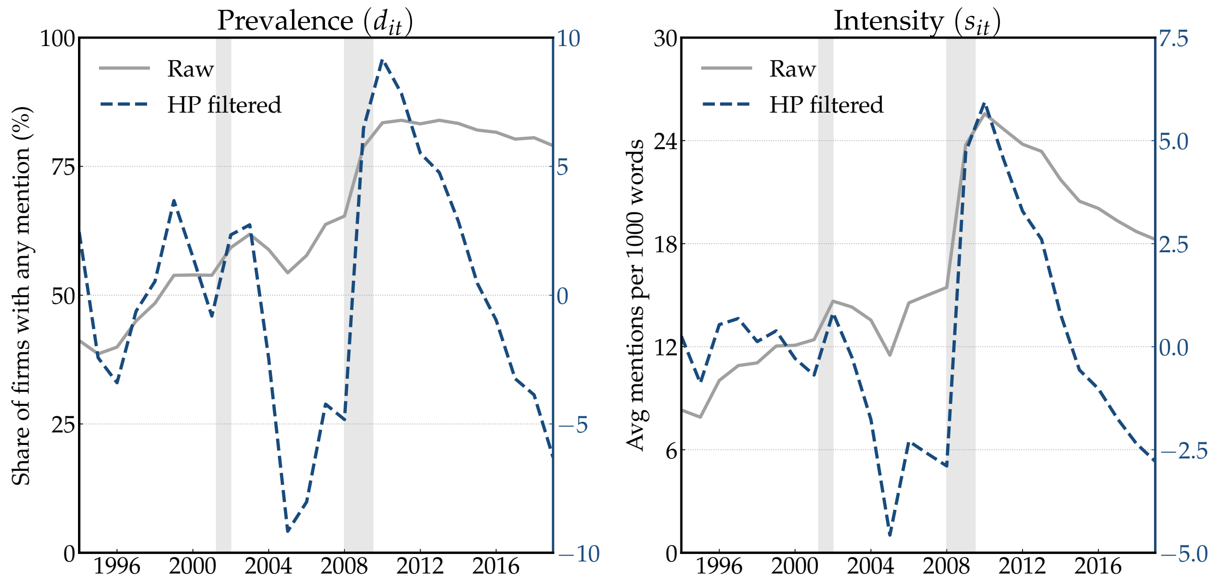
2.3 Stylized facts about firm attention

We first apply our prevalence and intensity measures to document four stylized facts about time and firm variation in attention. The first two facts presented below, on countercyclical and industry-specific attention, are consistent with existing theory on rational inattention and serve as sanity checks of the measures. The last two facts, on the polarization of attention and characteristics of attentive firms, extend our understanding of firm attention. We interpret these results in the context of existing literature and discuss how they might discipline future models of rational inattention.

Countercyclical attention to economic conditions Both the share of firms that mention macro keywords and the intensity with which firms mention macro keywords vary countercyclically over the business cycle. To illustrate this, we plot the time series related to the keyword "economic conditions". Figure 1 plots the share of firms that mention the keyword. The left panel reports the prevalence measure, and the right panel reports the intensity measure. Both panels also show the cyclical components of the HP-filtered series in both panels.

The share of firms that mention "economic conditions" increased over the sample period,

Figure 1: Time series of attention to "economic conditions"



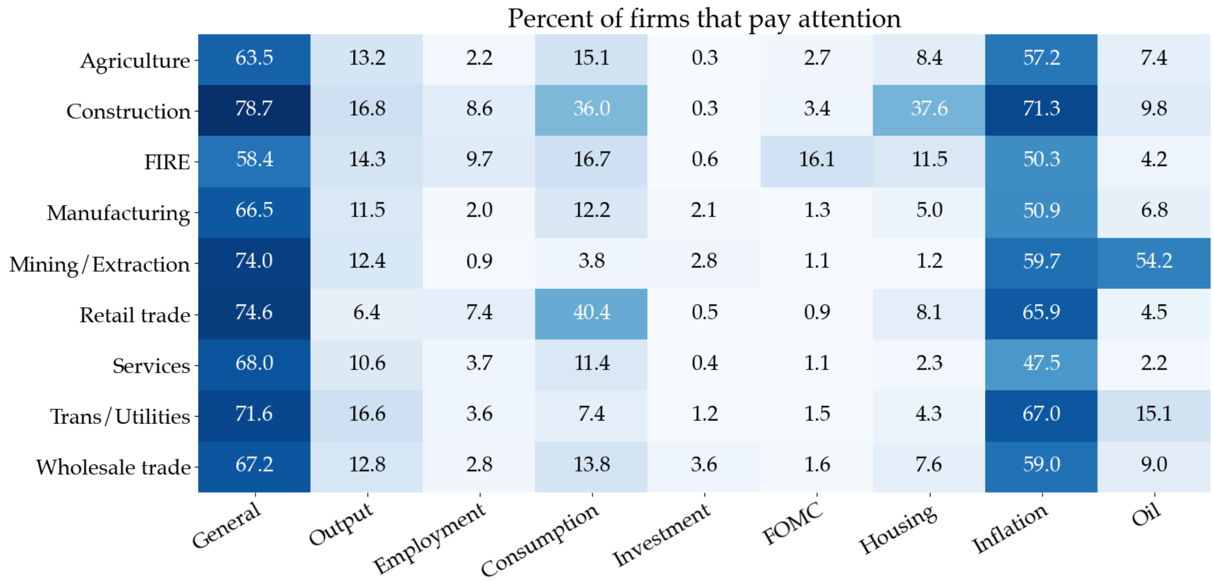
Notes: Time series of firm attention to the keyword "economic conditions". Left panel plots the prevalence measure and reports the share of firms that mention the keyword. The right panel plots the intensity measure and reports the average mentions of the keyword per 1,000 words. "Raw" refers to the unfiltered series and "HP filtered" refers to the cyclical components of the HP-filtered series. Shares are reported in percent.

with faster growth during recessions. The share of firms jumped by about 15 percentage points during the Great Recession and has moderated to approximately 80% in subsequent years.

The intensity related to the keyword "economic conditions" across all filings displays a stronger cyclical trend than the share of firms mentioning output. The share of words increases more during recessions and falls faster during recoveries compared to the share of firms mentioning output.

Countercyclical attention exhibited in Figure 1 is consistent with predictions in Mackowiak and Wiederholt (2009) about the allocation of attention between aggregate and idiosyncratic state variables. The authors argue that firms will allocate more attention to aggregate conditions than idiosyncratic conditions when aggregate conditions become more variable or relatively more important. If uncertainty is countercyclical, as macroeconomic conditions worsen and uncertainty rises, firms must pay greater attention to economic conditions when planning operations and forecasting performance.

Figure 2: Firm attention by industry



Notes: Heat map of the fraction of firms in an industry that pay attention to each macroeconomic topic. Industry is defined as 2-digit NAICS. Darker color represents a higher fraction of firms that pay attention.

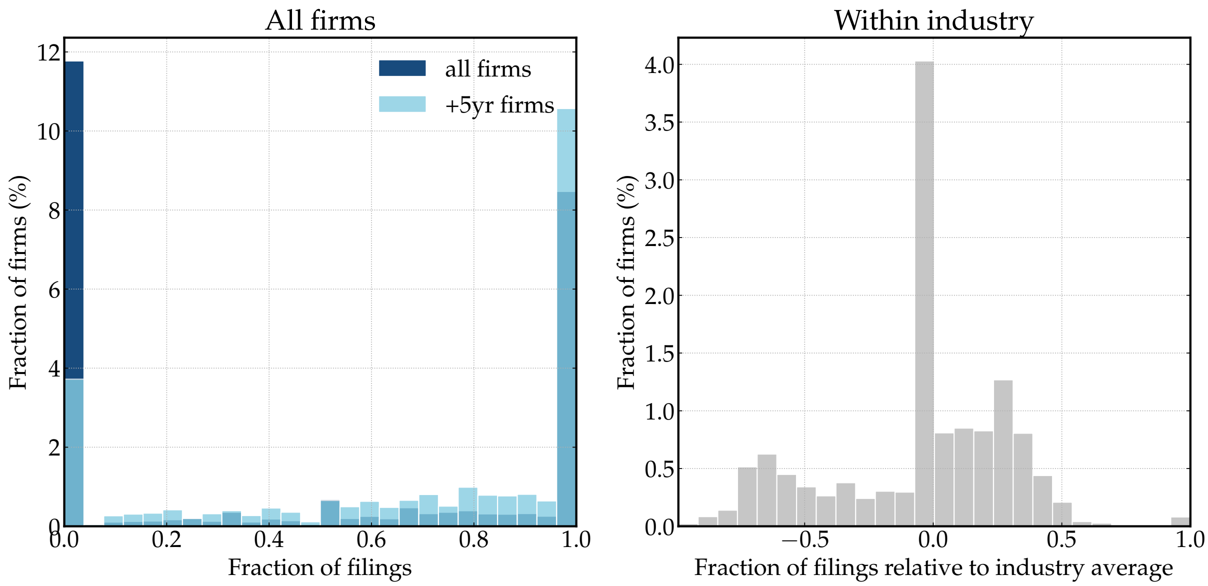
Cross-industry variation in attention Figure 2 reports the share of firms that pays attention to each topic by industry. Industry is measured using 2-digit NAICS from Compustat. The quality of our attention measure varies by topic so these results should be interpreted across industry rather than across topic.

For each macro topic, attention is highest in industries for which profits are most sensitive to the topic. For example, Mining, Oil, and Gas (NAICS 21) has the highest share of firms that pay attention to news about oil prices; Retail trade (NAICS 44-45) pays the greatest attention to news about consumption; and finance (NAICS 52) pays the greatest attention to news about FOMC meetings.

Furthermore, some industries appear to pay greater overall attention than others. Finance ranks among the most attentive industries to employment, FOMC, output, and interest rates, while agriculture (NAICS 11) and Professional, Scientific, and Technical Services (NAICS 54) appear least attentive overall.

The two features of cross-industry variation described above are fairly unsurprising and should be considered as sanity checks of our attention measure. Put simply, industries whose profitability depends more on a certain macro topic have a higher share of firms that pay

Figure 3: Share of filings that mention "economic conditions"



Notes: Histogram of share of filings that mention "economic conditions". The left panel shows the histogram of the average fraction of filings that mention the keyword "economic conditions" over the sample period of 1994-2019. Dark blue bars correspond to the distribution of all firms, and light blue bars correspond to firms appearing for at least 5 years in the sample. A value of 0 corresponds to a firm that has never mention the keyword in any of its filings. The right panel shows the histogram of the time series averages of the residuals of firm attention to "economic conditions" after regressing on industry fixed effects (2-digit NAICS). A value of 0 corresponds to a firm at industry average. Shares of firms on the vertical axes are reported in percent.

attention to that topic, and some industries appear to have greater overall sensitivity to aggregate economic conditions.

Polarization in firm attention The left panel of Figure 3 plots the histogram of firms by average attention over the sample period. The number of bins matches the number of annual observations in our sample and can be doubly interpreted as the number or fraction of filings in which firms pay attention. A firm with a value of 0 for the fraction of filings on the horizontal axis has never mentioned "economic conditions" over the sample period, whereas a firm with a value of 1 has mentioned that phrase in every filing over the sample period. Most notably, firms are concentrated at either never mentioning a macroeconomic keyword in their filings or mentioning a macroeconomic keyword in every filing. Despite the countercyclical variation found above, it appears that most variation in attention occurs across firms and that attention is largely invariant over time.

To test whether this polarization is driven by firms with few filings we replicate the histogram using a restricted sample of firms with at least five years of filings. Although this restriction greatly reduces the number of firms that never pay attention to macroeconomic news in our sample, the polarization between always- and never-attentive firms remains.

We also test whether polarized attention is attributable to industry patterns in attention. The right panel of Figure 3 demeans firm attention by industry to isolate within-industry heterogeneity. This panel depicts a large degree of variation in firm attention even after accounting for industry averages. Aside from a high concentration of attention at the industry average, demeaned attention appears bimodally dispersed around average.

The concentration at the industry average raises concern about the text-based measure: Does the frequency of macroeconomic keywords in 10-K filings capture firm attention to macroeconomic news or firm exposure to aggregate conditions? It is entirely plausible that a firm does not discuss the macroeconomy because its profits are not sensitive to aggregate fluctuations. Our main empirical analysis in Section 4 will focus on disentangling the attention channel from the exposure channel. We validate the text-based measures as capturing attention through the empirical findings of asymmetric responses to aggregate disturbances. If firms discuss macro news more because they are more exposed, then "attentive" firms would profit more when there is a positive shock and lose more when there is a negative shock, generating symmetric responses to monetary shocks. On the other hand, if the text-based measures correctly capture attention, then attentive firms would outperform inattentive firms in response to both positive and negative shocks, resulting in asymmetric responses. The theoretical framework in Section 3 discusses the mechanism in details.

Heterogeneous attention to publicly available news about U.S. output provides the clearest evidence that firms are limited in their capacity to process available information. The profitability of all publicly traded firms in our sample is arguably exposed to variation in U.S. economic conditions, and we should expect firms with unlimited information-processing bandwidth to incorporate this news into their decision making. Evidence of heterogeneity is to the contrary and provides new insights into how firms allocate attention differently.

Table 2: Firm characteristics and attention

	N	Mean	Median	SD
Inattentive				
Total assets (Millions)	33,277	2,873.36	104.02	35,004.36
Age	33,796	7.78	7.00	4.98
Leverage	32,955	0.35	0.17	0.69
Attentive				
Total assets (Millions)	102,493	7,311.57	538.12	65,274.94
Age	103,312	11.57	10.00	7.37
Leverage	101,981	0.30	0.20	0.46
Total				
Total assets (Millions)	135,770	6,223.78	370.50	59,333.37
Age	137,108	10.64	9.00	7.05
Leverage	134,936	0.31	0.19	0.53

Notes: Firm characteristics by firm attention type. In this table, a firm is attentive if its prevalence attention to the general topic is positive in any year in the sample period. Leverage is measured as total debt over asset.

Firm characteristics and attention We now examine firm characteristics that drive attention. Table 2 shows the summary statistics of firm characteristics by attention. A firm is attentive if its prevalence attention to the general topic is nonzero in any year in the sample period. Firm size is measured by the log of total assets, age is measured as the number of years since the firm first appeared in our sample, and leverage is defined as the ratio of total debt to market equity. We observe little right-censoring in our age measure because the Compustat sample starts 13 years before the attention sample begins.

Attentive firms tend to be larger and older than their inattentive counterparts, with an average asset size of $7.3 billion for attentive firms as compared to $2.9 billion for inattentive firms and an average age of 12 years for attentive firms compared to 8 years for inattentive firms. The leverage ratio for attentive is slightly lower at 0.30 as compared to 0.35 for inattentive firms

To better understand firm characteristics that are associated with attention, we estimate the marginal effects of firm size and age on the likelihood that firms pay attention to each macro news topic using a probit model. We include industry fixed effects by 4-digit NAICS to control for cross-industry variation in attention and firm covariates.

Table A.2 in the Appendix displays the results of the probit regressions. Both larger and older firms are more likely to pay attention to each macroeconomic series. A one percent increase in total assets is associated with a 2.2 basis point increase in the probability that a firm pays attention to GDP news, and an additional year of age is associated with a 0.4 basis point increase in the probability of attention.

3 Illustrative Framework

Motivated by the evidence that firms are heterogeneous in their attention to macroeconomic news, we set out to study how firm attention affects monetary transmission. Before doing so, we address a key identification challenge: how often a firm mentions macroeconomic keywords can be driven by its exposure to macroeconomic conditions, rather than attention. To confront the identification challenge, we lay out a stylized model in which firms are heterogeneous in both attention and exposure. For the two sources of heterogeneity, the model yields contrasting predictions for stock return responses to monetary shocks, which we then exploit to guide our regression specifications. The model environment is kept minimal to illustrate the key mechanisms for attention and exposure. In the quantitative model in Section 5, we expand the model settings to incorporate more realistic assumptions.

Environment Time is static. Consider a firm whose profits, $\pi(s, a)$, depend on an aggregate state variable, s, and a firm action, a. Assume that $\pi(s, a)$ is twice continuously differentiable, a single-peaked function of a, and maximized at $a^* = s$. For concreteness, we think of a as the price a monopolistic competitive firm sets and s as the exogenous optimal price determined by factors outside of a firm's control, as in Woodford (2009).

Firm profits can be approximated under a second-order log approximation around the non-stochastic steady state as[3]:

$$\hat{\pi}(\hat{s}, \hat{a}) = \pi_s(\bar{s}, \bar{a})\bar{s}\hat{s} + \frac{1}{2}\left(\pi_{ss}(\bar{s}, \bar{a})\bar{s}^2 - \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\right)\hat{s}^2 + \frac{1}{2}\pi_{aa}(\bar{s}, \bar{a})\bar{a}^2(\hat{a} - \hat{s})^2 \tag{1}$$

[3]Under this approximation, $\pi_a(s, a)$ drops out because of the first-order condition and assumption that $a^* = s$ at the optimum. Appendix A.5 contains detailed derivations of the approximation.

where $\bar{s}$ and $\bar{a}$ denote the steady-state values, $\hat{\pi}$, $\hat{s}$ and $\hat{a}$ denote the log deviations from the steady state, and $\pi_s \equiv \frac{\partial}{\partial s}\pi(s,a)$, $\pi_{aa} \equiv \frac{\partial^2}{\partial a^2}\pi(s,a)$ and $\pi_{ss} \equiv \frac{\partial^2}{\partial s^2}\pi(s,a)$.

Lastly, assume that firm profits are increasing in s, $\pi_s > 0$, and that the second-order condition for a stable equilibrium holds, $\pi_{aa} < 0$.

Attention and Exposure We can now define attention and exposure in the model. Intuitively, a firm is more exposed to aggregate conditions if its profits are more elastic with respect to aggregate shocks. On the other hand, a firm is more attentive if its action responds more to shocks. Definitions 1 and 2 formalize the ideas.

***Definition* 1** (attention)**.** *Let a firm's action be a function of the state: $\hat{a} = f(\hat{s})$, with $f(0) = 0$ and $0 < f'(\hat{s}) \leq 1$. Firm i is attentive to macroeconomic conditions if $f_i'(\hat{s}) = 1$, and firm j is inattentive to macroeconomic conditions if $0 < f_j'(\hat{s}) < 1$.*

An attentive firm reacts one-for-one with innovations to the aggregate state, whereas an inattentive firm responds less than one-for-one. The simplified definition of inattention is consistent with that in rational inattention models such as Sims (2003) which yields a steady-state Kalman gain between 0 and 1.

***Definition* 2** (exposure)**.** *Firm i is more exposed to macroeconomic conditions than firm j if $\pi_s^i(s,a) > \pi_s^j(s,a)$.*

Differences in attention and exposure We now derive model predictions for heterogeneity in attention and exposure that guide the empirical analysis in the next section.

We first construct stock returns, which is the dependent variable in our empirical analysis. As in Gorodnichenko and Weber (2016), a firm's stock price is equal to its firm value, which in the simple static setting equals its profits:

$$v = \pi(s,a)$$

The *realized equity returns*, measuring the log changes in a firm's values before and after an aggregate shock is realized, are given by:

$$r = \hat{v} - \mathbb{E}_{-1}\,\hat{v} \tag{2}$$

where $\hat{v}$ denotes the log deviations of firm values from the steady state.

Proposition 1 highlights the asymmetric responses of stock returns to positive and negative aggregate shocks that emerge from the attention channel of the model. In contrast, the model predicts symmetric return responses from the exposure channel.

***Proposition* 1.** *The return elasticity with respect to aggregate shocks for the exposure and the attention channels can be characterized as below:*

(i) ***Exposure****: If firm i is more exposed to macroeconomic conditions than firm j, then holding all else equal the return elasticity of firm i with respect to the aggregate shock is higher than the return elasticity of firm j for all shocks:*

$$\frac{\partial r_i}{\partial \hat{s}} > \frac{\partial r_j}{\partial \hat{s}} \quad \forall \hat{s}$$

(ii) ***Attention****: Suppose firm i is attentive to macroeconomic conditions and firm j is inattentive. Then, holding all else equal, for positive (expansionary) shocks, the return elasticity with respect to the aggregate shock for the attentive firm i is higher than the return elasticity of the inattentive firm j. For negative (contractionary) shocks, the return elasticity for the attentive firm i is lower than for the inattentive firm j. For zero shocks, the return elasticities for attentive and inattentive firms equal:*

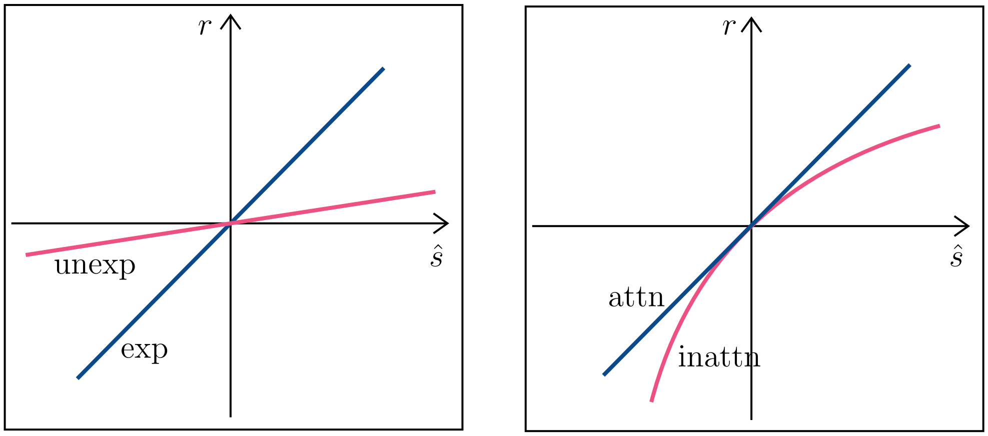
$$\begin{cases} \frac{\partial r_i}{\partial \hat{s}} > \frac{\partial r_j}{\partial \hat{s}} & \text{if } \hat{s} > 0 \\ \frac{\partial r_i}{\partial \hat{s}} = \frac{\partial r_j}{\partial \hat{s}} & \text{if } \hat{s} = 0 \\ \frac{\partial r_i}{\partial \hat{s}} < \frac{\partial r_j}{\partial \hat{s}} & \text{if } \hat{s} < 0 \end{cases}$$

Proof. See Appendix A.6 ■

Figure 4 illustrates the predictions from Proposition 1. In Panel (a), firms are heterogeneous in their exposures to aggregate shocks. Firms that are more exposed to aggregate

Figure 4: Model predictions for exposure vs attention

(a) Heterogeneity in exposure **(b)** Heterogeneity in attention



Notes: Illustration of model predictions of return elasticity with respect to aggregate shocks. Verticle axes represent conditional realized return, and horizontal axes represent the magnitude of shocks. Left panel shows return elasticity for firms that are exposed to macro conditions (*exp*) and firms that are unexposed (*unexp*). Right panel shows return elasticity for attentive firms (*attn*) and inattentive firms (*inattn*). Exposure and attention are as defined in the main text.

shocks see bigger fluctuations in their stock returns following an aggregate shock. Their stock returns rise by more in response to a positive shock and drop by more to a negative shock, compared to those of their less sensitive peers. Higher exposure leads to higher return elasticity to aggregate fluctuations, regardless of the sign of the shock.

Panel (b) illustrates the mechanism of attention. Attentive firms are better at tracking the state variable, so their stock returns outperform those of inattentive firms after any aggregate disturbance. In response to a positive shock, stock returns of both attentive and inattentive firms rise, but returns of attentive firms rise more. In contrast, in response to a negative shock, returns of both types of firms decrease, but returns of attentive firms drop by less.

The asymmetry in responses to aggregate shocks is a unique feature of the attention channel, which allows us to distinguish between the effects of firm attention to macro news from firm exposure to macro news.

In the next section, we exploit this predicted asymmetry to test whether our text-based measure is identifying exposure or attention, and estimate the cost of inattention to assess its qualitative significance. The mechanism highlighted through the stylized inattention model is general to any aggregate shocks, including TFP and fiscal policy shocks. We focus on implementing the empirical analysis with monetary policy shocks since it is one of the most well-identified macroeconomic shocks[4].

4 Empirical Analysis

Given our attention measures and theoretical predictions, we set out to test the hypothesis that attentive firms respond to monetary policy shocks better than inattentive firms. We use a high-frequency identification strategy to isolate plausibly exogenous shocks to monetary policy from FOMC announcements and look at changes in stock prices of attentive and inattentive firms within a similarly narrow window around such announcements. The advantage of using stock prices as the outcome variable is that asset prices quickly reflect changes in expected future profits. More direct measures of firm responses such as price adjustments, investment, and hiring decisions are only observed over longer time horizons. The weak power of high-frequency monetary policy shocks combined with confounding factors that influence firms' choices would likely prevent us from confidently estimating the effects of monetary policy shocks on these firm choice variables.

Once we identify the effects of monetary policy shocks on stock returns, we must correctly identify the differences in stock price responses that are attributable to firm attention. We exploit the asymmetry prediction from Section 3 to disentangle the effects of attention from exposure. Furthermore, to best isolate the effects of attention from a firm's exposure to monetary policy, we include controls for firm size, age, leverage, and industry measured by 4-digit NAICS in all our baseline specifications. We assume firms within a narrowly defined industry that have similar size, age, and financial structure have similar exposure to monetary policy shocks and that residual variation in stock prices can be attributed to firm attention rather than cross-firm variation in the exposure to monetary policy.

[4] Ramey (2016) provides a comprehensive survey on the efforts on identifying monetary shocks.

4.1 Data

Monetary policy shocks are constructed using the high-frequency identification strategy developed by Cook and Hahn (1989) and used recently in Gorodnichenko and Weber (2016), Nakamura and Steinsson (2018), and Ottonello and Winberry (2018). These shocks are measured as the change in the fed funds futures rate within a one-hour window surrounding FOMC announcements. Any changes within such a narrow window can be attributed to unanticipated changes to monetary policy as it is unlikely that other shocks occurred within the same window.

Monthly fed funds futures contracts clear at the average daily effective fed funds rate over the delivery month, so rate changes are weighted by the number of days in the month that are affected by the monetary policy shock. Following notation in Gorodnichenko and Weber (2016), the final shock series is defined as,

$$v_t = \frac{D}{D-\tau}(ff^0_{t+\Delta t^+} - ff^0_{t-\Delta t^-}), \tag{3}$$

where t is the time of the FOMC announcement, $ff^0_{t+\Delta t^+}$ and $ff^0_{t-\Delta t^-}$ are the fed funds futures rates 30 minutes before and after the announcement, D is the number of days in the month of the announcement, and τ is the date of the announcement. We use the series published by Gorodnichenko and Weber (2016) and Nakamura and Steinsson (2018) for monetary shocks from 1994 to 2014.

Firm outcome and control variables are constructed using CRSP and Compustat data available through Wharton Research Data Services (WRDS). Daily stock returns are measured as the open-to-close change in stock prices on the day of an FOMC announcement. Returns are winsorized at 1% to truncate outliers in daily stock movements. Firm size, age, and industry controls are constructed as described in Section 2.3.

Firm attention is measured using the *prevlance* measure, d_{it}, described in Section 2. To better suit a high-frequency methodology, firm attention at the time of an FOMC announcement is identified using the firm's most recent public filing rather than the filing that applied to the same quarter as the FOMC announcement. This modification excludes the possibility that firms are identified as attentive to an FOMC announcement that happened earlier in the quarter.

4.2 Methodology

The key feature of Figure 4(b) that we test is whether the stock returns of inattentive firm fall by more following negative (contractionary) shocks and rise by less following positive (expansionary) shocks. The average interaction effect between monetary shocks and firm attention is ambiguous alone because the sign of the effect depends on the direction of the shock: the interaction coefficient should be negative for negative monetary shocks and positive for positive monetary shocks. We separately estimate the slope of the interaction between monetary shocks and firm attention for positive and negative shocks, and then test whether these two coefficients are statistically different.

For a firm i in industry j on day t, our baseline model takes the form,

$$r_{it} = \delta_j + \beta_{v+} v_t \mathbb{1}_{v_t>0} + \beta_{v-} v_t \mathbb{1}_{v_t<0} + \beta_{dv+} d_{it} v_t \mathbb{1}_{v_t>0} + \beta_{dv-} d_{it} v_t \mathbb{1}_{v_t<0} + \boldsymbol{\beta}'_{\boldsymbol{X}} \mathbf{X}_t + \varepsilon_{it}, \qquad (4)$$

where d_{it} is the attention prevalence, v_t is the monetary policy shock, $\mathbb{1}_{v>0}$ indicates positive monetary policy shocks, $\mathbb{1}_{v<0}$ indicates negative monetary policy shocks, and $\mathbf{X}_t$ is a set of controls including the indicator variable for positive shocks and quarterly firm controls for size, age, and leverage. $\mathbf{X}_t$ also includes interaction terms of monetary shocks with industry fixed effects and firm controls to capture the possibility of firm characteristics drive the differential responses to monetary shocks. Standard errors are clustered by FOMC announcement to allow for correlated errors across firms at each FOMC announcement.

The coefficients of interest are β_{dv+} and β_{dv-}. The theoretical framework in Section 3 hypothesizes β_{dv+} to be positive and β_{dv-} to be negative, implying attentive firms should outperform inattentive firms in response to both expansionary and contractionary monetary shocks. To formally test the hypothesis, we conduct a Wald Test with the null hypothesis $H_0 : \beta_{dv+} = \beta_{dv-}$.

Table 3: Baseline results

	(1)	(2)	(3)	(4)
Shock	4.55*	4.55*		
	(2.64)	(2.65)		
Attention		-0.01	-0.07	-0.03
		(0.05)	(0.05)	(0.05)
Shock × Attn		1.07		
		(0.64)		
Shock × $\mathbb{1}_{v_t>0}$			4.93*	6.54**
			(2.70)	(2.61)
Shock × $\mathbb{1}_{v_t<0}$			-3.57	-0.95
			(4.39)	(4.45)
Shock × Attn ×$\mathbb{1}_{v_t>0}$			2.02***	1.55**
			(0.64)	(0.64)
Shock × Attn ×$\mathbb{1}_{v_t<0}$			-5.87**	-5.77*
			(2.87)	(2.89)
Observations	575667	575667	575667	432458
R^2	0.022	0.022	0.026	0.027
Clustered SE	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes
4-digit NAICS FE	yes	yes	yes	yes
excl. ZLB	no	no	no	yes
Wald Test p-value			0.013	0.025

Notes: Results from variants of estimating the baseline specification

$$r_{it} = \delta_j + \beta_{v+} v_t \mathbb{1}_{v_t>0} + \beta_{v-} v_t \mathbb{1}_{v_t<0} + \beta_{dv+} d_{it} v_t \mathbb{1}_{v_t>0} + \beta_{dv-} d_{it} v_t \mathbb{1}_{v_t<0} + \boldsymbol{\beta}_X \mathbf{X}_t + \varepsilon_{it}$$

where δ_j is an industry fixed effect, v_t is the monetary shock, D_{it} is the prevalence attention measure, and $\mathbf{X}_t$ contains the indicator variable for positive shocks $\mathbf{1}_{v_t>0}$ and firm level controls of size, age and leverage. We also include firm controls and industry fixed effects interacted with the monetary shocks. Standard errors are clustered at the shock level. We have normalized the sign of teh monetary shock ν_t so that a positive shock is expansionary (corresponding to a decrease in interest rates). Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

4.3 Results

Our baseline results are reported in Table 3. In the first column, we estimate the effect of high-frequency monetary shocks without our attention measures and find a that a 25 basis point unanticipated increase in the fed funds rate is associated with about a one percent increase in stock prices. This results is consistent with existing estimates from Gorodnichenko and Weber (2016) and Nakamura and Steinsson (2018). The second column introduces the unconditional interaction between monetary shocks and firm attention. We find that attentive firms experience slightly higher stock returns than their inattentive counterparts but our estimate is not statistically distinguishable from zero. This result is consistent with the framework outlined in Section 3, which remains agnostic as to the average interaction over the entire range of monetary shocks.

The main results from Equation (4) are presented in the third column. We test whether attention leads to differential responses to positive and negative monetary shocks. We compare firms within a narrow industry by including 4-digit NAICS industry fixed effects and firm-level controls for size, age and leverage, both standalone and interacted with the shocks. Consistent with predictions from rational inattention models, attentive firms appear to experience larger increases in stock returns following expansionary monetary shocks and smaller decreases in stock returns following contractionary monetary shocks. The coefficients are statistically different from zero, and the Wald Test of whether these coefficients are equivalent is rejected at 5% significance.

Finally, fourth column ends the sample in 2007 to exclude the zero lower bound period following the Great Recession. Results are both qualitatively and quantitatively similar as in the full sample, suggesting our findings are not driven by anomalies from the financial crisis or the zero lower bound periods.

The asymmetric responses to positive and negative shocks are uniquely consistent with heterogeneous responses predicted by a model of incomplete attention. If the attention measure used above misidentified attention as the sensitivity of firms' profits to macroeconomic conditions, then supposedly attentive firms would perform better in response to a positive shock and suffer larger losses in response to a negative shocks. In this case we would expect to see a positive and significant effect from the interaction term between shock and attention (β_{dv}) in the second column. This is inconsistent with the results in Table 3. Yet another alternative hypothesis that explains why firms may mention FOMC meetings is that firms attribute their own poor performance to broader economic forces. If firms mention more macroeconomic keywords when they are underperforming, then we would expect attentive firms to underperform in response to a negative monetary shocks, corresponding to a positive coefficient for β_{dv-} in the third column, which is also at odds with our empirical findings.

5 Quantitative Model

Motivated by the empirical importance of inattention on firm performance, we now construct a general-equilibrium model with rationally-inattentive firms. Key parameters of the model are calibrated using the attention measure and empirical moments from the sections above. Using the quantitative model, we explore the effects of inattention on the efficacy of monetary policy.

5.1 Model environment

The model mechanism is an extension to the stylized model outlined in Section 3. Time is discrete and infinite. The economy consists of households, firms and the central bank. Households and the central bank have full information about the economy, while firms face information frictions. We start with a standard general equilibrium model with rationally inattentive firms as in Mackowiak and Wiederholt (2009) and Afrouzi and Yang (2019). Attention is modeled with the Shannon mutual information following Sims (2003, 2010). Then we incorporate heterogeneous costs of information and connect model objects to the data to calibrate parameters for information frictions.

Household A representative household maximizes its life-time utility,

$$\max_{C_{it},N_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t (\log C_t - \psi N_t), \tag{5}$$

where N_t denotes the labor supply and ψ parameterizes the disutility of labor. Consumption C_t is aggregated over each good type i with a CES aggregator,

$$C_t = \left(\int_0^1 C_{it}^{\frac{\varepsilon_p - 1}{\varepsilon_p}} dj \right)^{\frac{\varepsilon_p}{\varepsilon_p - 1}}, \tag{6}$$

where ε_p is the elasticity of substitution. In addition to the wage income, households have access to a one-period bond B_t with the interest rate ι_t and receives a lump-sum transfer D_t

from the government. The household budget constraint is given by:

$$\int_0^1 P_{it} C_{it} di + B_t \leq W_t N_t + (1 + \iota_t) B_{t-1} + D_t \tag{7}$$

Central Bank The central bank targets aggregate money supply similar to Caplin and Spulber (1987) and Gertler and Leahy (2008). As a result, the nominal aggregate demand follows an autoregressive process:

$$\Delta \log Q_t = \rho \Delta \log Q_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2) \tag{8}$$

Firms Firms are owned by a risk-neutral agent and have production technology that is linear in labor:

$$Y_{it} = N_{it}$$

The functional form of a firm's information flow is specified with Shannon's mutual information:

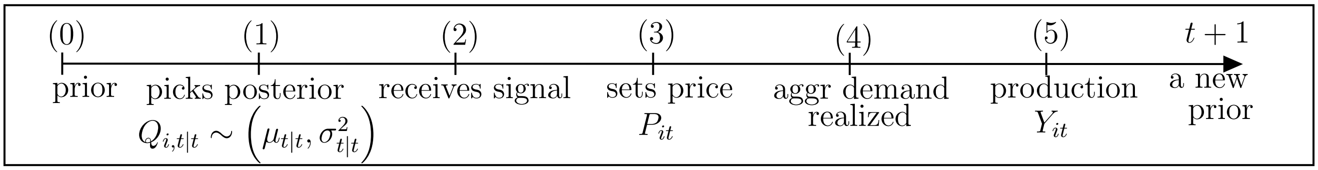
$$\mathcal{I}(\tilde{Q}_{i,t|t-1}, \tilde{Q}_{i,t|t}) = \frac{1}{2} \log \frac{\sigma^2_{i,t|t-1}}{\sigma^2_{i,t|t}} \tag{9}$$

which is decreasing in the posterior variance, so that more precise posteriors are more expansive. The marginal cost of information per nat, $2\omega_i$ is heterogeneous across firms and can be either high or low:

$$\omega_i \in \{\omega_H, \omega_L\}$$

Figure 5 shows a firm's timeline. It enters a period with a prior on the aggregate demand. Then it chooses the posterior distribution. Since the Shannon mutual information in (9) does not depend on the posterior mean, it is optimal for a firm to center the posterior distribution around the true mean. So the firm's information choice is only of the posterior variance $\sigma^2_{t|t}$. Based on the chosen posterior distribution, the firm receives a signal on the aggregate demand and sets its price P_{it} based on the posterior belief. Then, the aggregate demand is realized,

Figure 5: Firm's Timeline



the firm produces and enters the next period with a new prior.

A firm's value function is given by

$$V(\sigma^2_{t|t-1}) = \max_{P_{jt},\sigma^2_{t|t}} \mathbb{E}_t \Big[\underbrace{\frac{Y_{jt}}{P_t}\left(P_{jt} - MC_t\right)}_{\text{flow op. profits}} - \underbrace{2\omega_i \mathcal{I}(\tilde{Q}_{i,t|t-1}, \tilde{Q}_{i,t|t})}_{\text{info costs}} + \beta \underbrace{\tilde{V}(\sigma^2_{t+1|t})}_{\text{cont. value}} \Big| \, \sigma^2_{t|t} \Big], \tag{10}$$

which consists of flow operational profits that are maximized when firms successfully track the aggregate demand, information costs that depend on firms' information acquisition choices, and a continuation value. The expectation operator of a firm is based on its time-t information set. The problem of a firm's manager in each period is to maximize the firm value by jointly setting prices and investing in attention.

Firms optimize subject to the following constraints:

$$Y_{jt} = \left(\frac{P_{jt}}{P_t}\right)^{-\varepsilon_p} C_t \tag{demand}$$

$$\sigma^2_{i,t+1|t} = \rho\sigma^2_{i,t|t} + \sigma^2_{\nu} \tag{law of motion for prior}$$

$$0 \leq \sigma^2_{i,t|t} \leq \sigma^2_{i,t|t-1} \tag{no forgetting}$$

The demand function comes from the household's problem, and the law of motion for a firm's prior belief is derived from the central bank's monetary rule. The no-forgetting constraint prohibits firms from discarding previously-acquired information to make room for new information, ensuring the Shannon information costs are non-negative.

Equilibrium The equilibrium consists of the household allocation, $\{C_t, \{C_{it}\}_{i\in[0,1]}, N_t\}_t$, firms allocations, $\{\sigma^2_{i,t|t}, P_{it}, Y_{it}\}_t$, and a set of prices $\{P_t, W_t\}_t$ such that:

(i) Given prices and the firms' choices, the household optimizes (5);

(ii) Given an initial prior $\sigma^2_{i,0|-1}$, prices and the households' choices, firms optimize (10);

(iii) Monetary policy follows (8);

(iv) All markets clear.

Model Solution Following Mackowiak and Wiederholt (2009) and Afrouzi and Yang (2019), we approximate firm's flow profits with second order log approximations around the full-information steady state.[5] This approximation yields an imperfect-information firm value, $\tilde{v}$. We decompose a firm's total value under log approximation, v, into a full-information value, v^*, representing the firm's value under optimal pricing with full information, and the imperfect information value, $\tilde{v}$, representing the loss in firm value from imperfect information.

The firm's imperfect information problem is solved numerically using the algorithm for dynamic rational inattention problems developed by Afrouzi and Yang (2019).

5.2 Calibration

Calibration features two sets of parameters: standard parameters unrelated to information frictions, which are set exogenously, and parameters related to information frictions. Importantly, we calibrate parameters related to information frictions to match the stylized facts on attention and the empirical elasticities estimated in the empirical analysis.

Standard parameters The top panel of Table 4 shows the calibration for predetermined parameters. The model period is a quarter, so the discount rate is set as $\beta = 0.95^{1/4}$. The monetary shock process is calibrated using quarterly US nominal output between 1994 and 2019. To match our empirical specification, which compares firms within a sector, we restrict our attention to nominal output in the manufacturing sector. The persistence of the shock is calibrated to $\rho = 0.89$ and the standard deviation is calibrated to $\sigma_\nu = 0.063$. The elasticity

[5] Log-quadratic approximation is a common simplifying assumption in rational inattention models to address the curse of dimensionality that arises from firms having the joint distribution of prices and nominal aggregate demand as the state variable. Sims (2003) shows that the optimal distribution under Gaussian priors and quadratic payoffs is also Gaussian, so log-quadratic approximation of the profit function greatly reduces the dimensionality of the problem.

Table 4: Calibration

Parameter	Description	Value
Standard parameters		
β	discount rate	$0.95^{1/4}$
ρ	shock persistence	0.89
σ_ν	shock std. dev.	0.063
ε_p	elasticity of substitution	11
ψ	disutility of labor	0.91
Information-friction parameters		
θ	fraction of attentive firms	65%
ω_L	cost of information	30
ω_H	cost of information	47

of substitution is set to $\varepsilon_p = 11$, implying a steady-state markup of 10%, and the disutility of labor is set to $\psi = 0.91$ to offset the steady-state distortions from monopolistic competition.

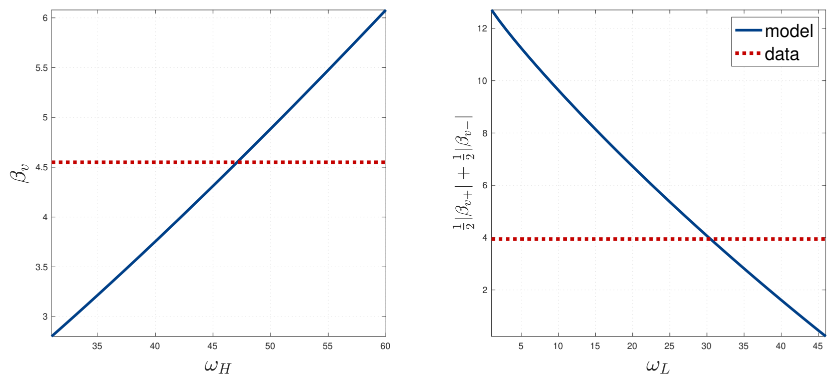
Information-friction parameters The bottom panel of Table 4 contains calibrations for parameters $(\theta, \omega_L, \omega_H)$. To calibrate these important parameters governing the degree of information frictions in the model, we use our text-based measure of attention and the empirical moments from Section 4.

The fraction of attentive firms is set to $\theta = 65\%$ to match the average fraction of firms that have paid attention to the keyword "economic conditions" over the sample period. Attention to economic conditions conveys firm attention of the aggregate demand, which is direct counterpart with the model state variable firms are tracking.

To calibrate the costs of attention, ω_L and ω_H, we target regression coefficients in Table 3 by running the same regressions with simulated model data. We first define model objects that match those observed in the data. Stock returns in the model are defined as the log change in a firm's value function in Equation (10), $r_{it} = \log V_{it} - \log \mathbb{E}_{t-1}(V_{it})$. We define attention in the model to be the Shannon mutual information. Since our main empirical specification uses the prevalence attention measure, we define a corresponding attention indicator, d_{it}, to equal 1 when a firm's attention is above the cross-sectional mean in a given period and 0 otherwise. Finally, we use ν_t as the monetary shocks. We simulate the model for a panel of 100 firms and for 1000 quarters, discarding the first 100 quarters as burn-in.

The cost of information for inattentive firms, ω_H, is calibrated to target $\hat{\beta}_v$ in Column (2)

Figure 6: Sensitivity of simulated moments to costs of information



Notes: Simulated moments for a range of costs of information parameters. We simulate models for a panel of 100 firms and for 1000 periods with 100 periods burn-ins. Simulated moments are generated with regressions discussed in the text.

of Table 3, which measures the average response of stock returns to monetary policy. With simulated data, we run the following regression:

$$r_{it} = c + \beta_v \nu_t + \beta_d d_{it} + \beta_{dv} d_{it} \nu_{it} + \varepsilon_{it},$$

and set ω_H so that the simulated β_v matches the empirical moment $\hat{\beta}_v$. The left panel of Figure 6 shows how ω_H is identified. We simulate the model for a range of values of ω_H. As the costs of information for attentive firms ω_H increases, the average response to monetary policy β_v increases monotonically.

For a given ω_H, we then set the cost of information for attentive firms, ω_L, to match $\hat{\beta}_{dv+}$ and $\hat{\beta}_{dv-}$ in Column (3) of Table 3, which measure the heterogeneous return semi-elasticity to monetary policy . The distance between ω_H and ω_L reflects the relative cost of information for inattentive firms compared to attentive firms. We run the regression with simulated data:

$$r_{it} = c + \beta_1 \mathbb{1}_{v>0} + \beta_{v+} v_t \mathbb{1}_{v>0} + \beta_{v-} v_t \mathbb{1}_{v<0} + \beta_d d_{it} + \beta_{dv+} d_{it} \nu_{it} \mathbb{1}_{v>0} + \beta_{dv-} d_{it} \nu_{it} \mathbb{1}_{v<0} + \varepsilon_{it}$$

In particular, the elasticity from Column (3) we target is we target is $\frac{1}{2}|\hat{\beta}_{dv+}| + \frac{1}{2}|\hat{\beta}_{dv-}|$, which measures the relative stock return losses of firms that do not pay attention. The right panel of Figure 6 shows how ω_L is identified. Given a value of ω_H, we simulate the model for a range of ω_L. As ω_L increases and the gap between ω_H and ω_L narrows, the simulated elasticity monotonically decreases, implying lowering heterogeneity between attentive and inattentive firms. Figure A.1 in the appendix shows how simulated β_{dv+} and β_{dv-} change individually as we vary ω_L. β_{dv+} is positive and β_{dv-} is negative, suggesting the stock returns of attentive firms outperform those of their inattentive peers for both positive and negative monetary shocks, consistent with our empirical findings. As ω_L increases and the gap between the information costs for attentive and inattentive firms narrow, β_{dv+} decreases and β_{dv-} increases, implying a lower degree of heterogeneity between attentive and inattentive firms.

The costs of information parameters are calibrated to $\omega_L = 30$ and $\omega_H = 47$. To our knowledge, the only existing study that quantitatively calibrates firm cost of attention is Afrouzi (2020), which studies the rational inattention problem of New Zealand firms under strategic complementarity and calibrates $\omega = 0.3$ using firm beliefs reported in New Zealand surveys. Our calibration differs both in our sample of US firms and in our approach of using the equity prices and their conditional responses to monetary policy shocks.

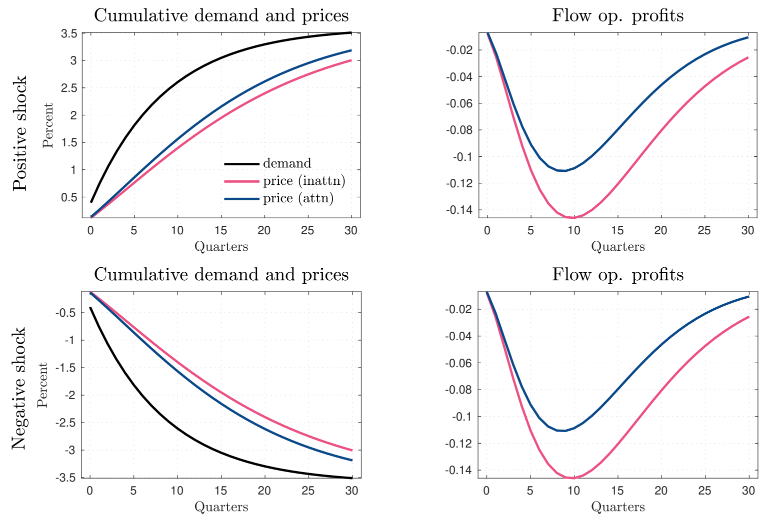
The calibration implies significant information costs for firms, which might seem surprising considering macroeconomic series are freely available. However, as plant-level evidence by Zbaracki et al. (2004) suggests, information costs involve not only information gathering costs but also information processing costs and communication costs. More recently, Abis and Veldkamp (2020) estimate the data production function which takes labor and capital inputs to process unstructured data into structure data and analyze data to produce knowledge. It requires significant manpower and expertise to process, summarize and forecast macroeconomic series into sufficient statistics that aids a firm's investment, production and pricing decisions, as highlighted in Reis (2006). The parameters of information costs in our model capture costs associated with processing information in addition to acquiring it.

5.3 Model dynamics

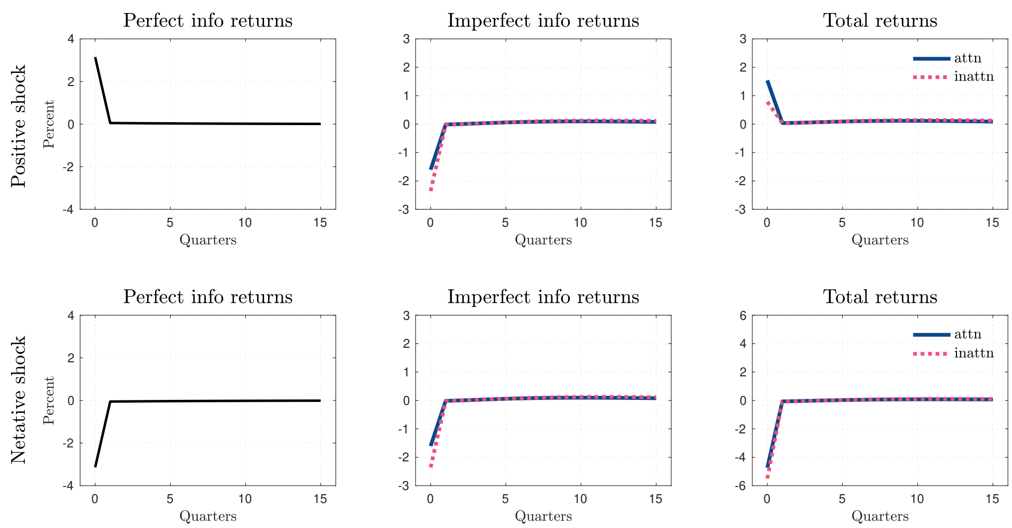
With our quantified model, we now study how firm inattention results in monetary non-neutrality. Figure 7 shows the impulse responses to expansionary and contractionary monetary shocks of one standard deviation. Inattentive firms are shown in red, and attentive firms are shown in blue. Panel (a) shows the responses of firm's prices and flow operating profits. As the nominal aggregate demand rises, firms' prices respond sluggishly, reflecting firms'

Figure 7: Firm impulse responses to monetary shocks

(a) Firm prices and operating profits



(b) Conditional realized returns



Notes: Firm impulse responses to a one standard deviation positive (expansionary) monetary shock and negative (contractionary) shock. Impulse responses are in percent deviations from the perfect-information steady state. “demand” refers the nominal aggregate demand. “attn” refers to the impulse responses of attentive firms, “inattn” refers to the impulse responses of inattentive firms.

partial incorporation of noisy signals about demand. Attentive firms are able to better track the aggregate demand than inattentive firms. Since we approximate firm profits around the full-information steady state, any deviation from the full-information benchmark results in a loss. The inattentive firms experience greater operational loss because they have less precise information about the aggregate demand. Inattentive firms also pay higher information costs despite acquiring less information, because they face a higher marginal cost of information. Costs of information are constant and do not result in change in returns.

Panel (b) shows the responses of stock returns. In response to an expansionary monetary shock, full-information equity returns of both attentive and inattentive firms increase, since firms are monopolistically competitive. Compared to the full-information returns, imperfect-information returns of attentive firms drop less than those of inattentive firms, because they track the optimal price more closely. In total, stock returns of both firms rise in response to an expansionary shock, but returns of attentive firms rise by more. In contrast, in response to a contractionary monetary shock, returns of both types of firms decrease, but returns of attentive firms drop by less.

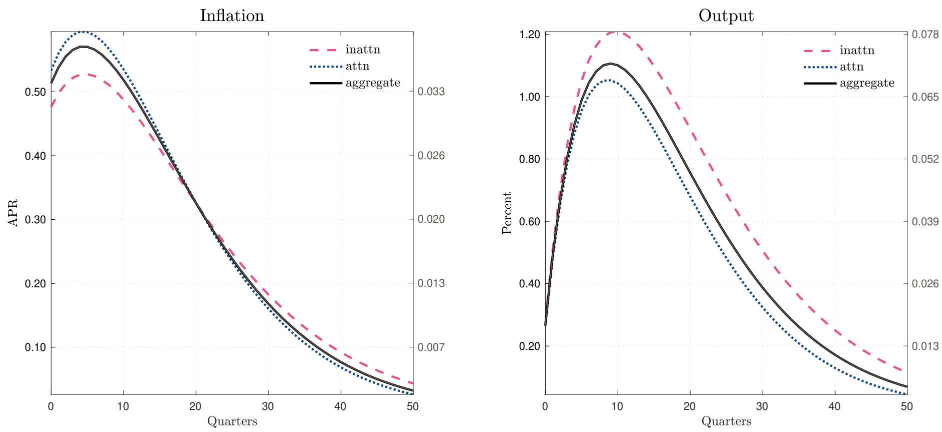
In Figure 8, we study the aggregate responses of output and inflation to a one standard deviation expansionary monetary shock by aggregating attentive and inattentive firms.

In responses to an equivalent of 25 basis point expansionary monetary policy shock[6], inflation increases by a peak value of an annualized 0.04% and output increases by a peak value of 0.07%. As a benchmark, Christiano et al. (2005) estimate the peak effect of monetary policy shocks to be an annualized 0.2% for inflation and 0.5% for output. With information as the only source of friction, the model generates about one seventh of the output responses.

The attentive firms, plotted in blue dotted lines, track nominal shocks more successfully than their peers and raise their prices faster. Inattentive firms, plotted in red dashed lines, ultimately set lower prices are must produce more than attentive firms.

[6]Our model considers monetary policy shock to the nominal aggregate demand and Christiano et al. (2005) consider shocks to the interest rate. In Appendix A.7 we estimate the passthrough of interest rate on the nominal aggregate demand with manufacturing output data.

Figure 8: Aggregate responses to expansionary monetary shock



Notes: Impulse responses of inflation and output. The right scales show the impulse responses to a one standard deviation expansionary monetary shock, and the right scales show the impulse responses to an equivalent of 25 basis point expansionary monetary policy shock. Impulse responses are in percent deviations from the perfect-information steady state. "attn" refers to the impulse responses of attentive firms, "inattn" refers to the impulse responses of inattentive firms, and "aggregate" refers to the aggregate impulse responses.

5.4 Inattention and the efficacy of monetary policy

In the rational inattention model, monetary non-neutrality increases with the fraction of inattentive firms and cost of information acquisition. Section 2 documents that firm attention evolves countercyclically over the business cycle. In Figure 1, firm attention rose during both the 2001 recession and the Great Recession.

The countercyclicality of overall attention leads to an important insight about the efficacy of monetary policy: when the Federal Reserve cuts rates in response to a realized recession, monetary policy is less powerful because more firms are already paying attention. With a higher fraction of attentive firms, information frictions are less severe, monetary policy is closer to neutral, and monetary stimulus have smaller effects on output. In contrast, monetary policies aimed at preemptively fending off a recession is more powerful, because a smaller fraction of firms are paying attention. As such, a preemptive monetary stimulus has a stronger effect on the real output.

To illustrate the quantitative scope of the effect, we exogenously vary the fraction of

Table 5: Attention and monetary non-neutrality

	Least attentive	Baseline	Most attentive
Fraction of attentive firms (θ)	56%	65%	73%
Average output response (%)	0.1016	0.0992	0.0971

Notes: Dependence of output responses on the fraction of attentive firms in the economy. Average output responses are calculated over 50 periods. Calibration for the least and most attentive economy is described in the main text.

attentive firms in the model and measure the average responses to a one standard deviation expansionary monetary shock. We start with the baseline calibration for the fraction of attentive firms, $\theta^{\texttt{baseline}} = 65\%$, which is the time series average of the prevalence measure of firm attention to aggregate demand between 1994 and 2019. Then, we decompose the time series of attention into the trend and cyclical components with the HP filter:

$$d_t = \tau_t + \zeta_t + u_t$$

where τ_t, ζ_t and u_t denote the trend, cyclical and error components of the attention measure d_t, respectively. The frequency is annual, and smoothing parameter for the HP filter is set to 400. We then add the minimum (maximum) of the cyclical component to the baseline calibration to form the most (least) attentive calibration of the model:

$$\theta^{\texttt{least attn}} = \theta^{\texttt{baseline}} + \min(\zeta_t)$$
$$\theta^{\texttt{most attn}} = \theta^{\texttt{baseline}} + \max(\zeta_t)$$

where $\min(\zeta_t)$ and $\max(\zeta_t)$ correspond to the minimum and and maximum of the HP-filtered prevalence measure in the left panel of Figure 1. Therefore, $\theta^{\texttt{least attn}} = 56\%$ and $\theta^{\texttt{most attn}} = 73\%$.

Then we study how aggregate responses to monetary policies change as we vary the fraction of attentive firms in the economy. Table 5 shows the average responses of output relative to the steady state over 50 periods. Compared to the least attentive calibration, the average output response to monetary policy is 5% weaker in the most attentive calibration. This suggests if the Federal Reserve cuts rates in the depth of a crisis period such as the COVID-

19 pandemic when all firms are paying attention to macroeconomic policies, its monetary stimulus will be 5% weaker than if it cuts rates in a preemptive fashion to lean against the wind. The results are consistent with studies on the state dependency of monetary policy, which finds US monetary policy to be weaker in recessions than in expansions (Tenreyro and Thwaites, 2016).

6 Limitations and Robustness

Before we conclude, we discuss several potential sources of limitations of our measures and their mitigation.

Context of attention First, our measure of attention is limited to whether a firm mentions macroeconomic news, but mentioning the news does not necessarily mean firms translate them into optimal policy responses. We consider this translation as the primary mechanism in our framework and seek to compare firms that vary in how much they process new information. To do so, we assume that firms mentioning a macroeconomic topic are also processing news related to that topic more than firms that do not mention the topic.

To assess whether the translation from words to action is plausible, we conduct two additional exercises, frequency search within granular sections of 10-K and topic modelling via Latent Dirichlet Allocation (LDA), to gauge the context in which firms discuss the macroeconomy. What firms discuss in conjunction with the macroeconomy helps us understand if and how macro variables enter into firms' policy functions. Detailed methodology of the additional robustness are discussed in Sections A.2 (itemized search) and A.3 (LDA) in the appendix.

Figures A.2 show the frequency search for granular sections in 10-K filines, and Figures A.3 and A.4 show topics discussed around macro keywords. Results from the additional robustness exercises show that firms pay attention to macro news to assess the impact on their business operations and risks, consistent assumptions that firms mentioning a macroeconomic topic do so in order to incorporate the news into their decision making.

Lexical similarity A possible explanation for firm persistence in attention is that firms recycle the boilerplate language year over year. To rule out the case, we measure the Jaccard similarity between filings of the same firm, defined as the share of unique non-stop words that appear between current year's and last year's 10-K filings. Section A.4 in the appendix contains details for the methodology. The Jaccard similarity score provides a measure of lexical similarity, and a high lexical similarity suggests the possibility of boilerplate languages.

Figure A.5 reports the Jaccard score for each section of 10-K, showing that the sections Business (Item 1) and Management's Discussion (Item 7) have the most distinct languages across filings. We then restrict our empirical analysis to only using attention measures from these two 10-K sections with low lexical similarity. Results in Table A.3 are both qualitatively and quantitatively similar as the baseline results.

False negatives Another potential limitation is that our measure misses any firms that follow macroeconomic news despite not mentioning it in their public filings with the SEC. In practice, the bias would attenuate our results, which makes our estimate to be a lower bound of information frictions.

Despite the limitations, we consider the measure to be informative as it provides a measure of firm attention which has been difficult to quantify. SEC filings are also the main channel through which firms communicate with shareholders who own firms, and therefore we consider the assumption that mentioning the macroeconomic topics translates into attention to be plausible.

7 Conclusion

This paper presents a new measure of firm attention to macroeconomic news. We validate the measure by testing for and finding an asymmetric effect of rational inattention on monetary policy transmission. We show that firms that pay attention to FOMC news have larger increases in stock returns after positive monetary shocks and smaller decreases in stock returns after negative monetary shocks. We document stylized facts about attention that can be used to discipline future research in rational inattention. To interpret the findings, we construct a quantitative model with rationally-inattentive firms and calibrate the model with empirical moments from the text-based attention measure. Inattention plays an important role in driving monetary non-neutrality, which leads to a new source of state dependency of monetary policy. We show how the countercyclical nature of firm attention to macroeconomic news reduces the efficacy of monetary policies that are aimed at counteracting downturns.

References

Abis, Simona and Laura Veldkamp, "The Changing Economics of Knowledge Production," *Available at SSRN 3570130*, 2020.

Afrouzi, Hassan, "Strategic Inattention, Inflation Dynamics, and the Non-neutrality of Money," 2020.

_ **and Choongryul Yang**, "Dynamic Rational Inattention and the Phillips Curve," *Available at SSRN 3465793*, 2019.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub, "Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model," Technical Report, National Bureau of Economic Research 2020.

Blei, David, Andrew Ng, and Michael Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 2003, *3* (Jan), 993–1022.

Buehlmaier, Matthias and Toni Whited, "Are Financial Constraints Priced? Evidence from Textual Analysis," *The Review of Financial Studies*, 2018, *31* (7), 2693–2728.

Caplin, Andrew and Daniel Spulber, "Menu Costs and the Neutrality of Money," *The Quarterly Journal of Economics*, 1987, *102* (4), 703–725.

_ **, Mark Dean, and John Leahy**, "Rational Inattention, Optimal Consideration Sets, and Stochastic Choice," *The Review of Economic Studies*, 2019, *86* (3), 1061–1094.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans, "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 2005, *113* (1), 1–45.

Coibion, Olivier and Yuriy Gorodnichenko, "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, 2015, *105* (8), 2644–78.

_ , _ , and Saten Kumar, “How Do Firms Form Their Expectations? New Survey Evidence,” *American Economic Review*, 2018, *108* (9), 2671–2713.

_ , _ , and Tiziano Ropele, “Inflation Expectations and Firm Decisions: New Causal Evidence,” *The Quarterly Journal of Economics*, 2020, *135* (1), 165–219.

Cook, Timothy and Thomas Hahn, “The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970s,” *Journal of Monetary Economics*, 1989, *24* (3), 331–351.

Gabaix, Xavier, “Behavioral Inattention,” in “Handbook of Behavioral Economics: Applications and Foundations 1,” Vol. 2, Elsevier, 2019, pp. 261–343.

Gertler, Mark and John Leahy, “A Phillips Curve with an Ss Foundation,” *Journal of Political Economy*, 2008, *116* (3), 533–572.

Gorodnichenko, Yuriy and Michael Weber, “Are Sticky Prices Costly? Evidence from the Stock Market,” *American Economic Review*, 2016, *106* (1), 165–99.

Gürkaynak, Refet S, Brian Sack, and Eric T Swansonc, “Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements,” *International Journal of Central Banking*, 2005.

Handley, Kyle and J Frank Li, “Measuring the Effects of Firm Uncertainty on Economic Activity: New Evidence from One Million Documents,” Technical Report, National Bureau of Economic Research 2020.

Hansen, Stephen, Michael McMahon, and Andrea Prat, “Transparency and Deliberation within the FOMC: A Computational Linguistics Approach,” *The Quarterly Journal of Economics*, 2018, *133* (2), 801–870.

Hassan, Tarek, Stephan Hollander, Ahmed Tahoun, and Laurence van Lent, “Aggregate and Idiosyncratic Political Risk: Measurement and Effects,” *Available at SSRN 2838644*, 2016.

Jordà, Òscar, "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 2005, *95* (1), 161–182.

Loughran, Tim and Bill McDonald, "When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *The Journal of Finance*, 2011, *66* (1), 35–65.

Mackowiak, Bartosz and Mirko Wiederholt, "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 2009, *99* (3), 769–803.

Malmendier, Ulrike and Stefan Nagel, "Learning from Inflation Experiences," *The Quarterly Journal of Economics*, 2016, *131* (1), 53–87.

Mankiw, Gregory and Ricardo Reis, "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *The Quarterly Journal of Economics*, 2002, *117* (4), 1295–1328.

Matějka, Filip and Alisdair McKay, "Rational Inattention to Discrete Dhoices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 2015, *105* (1), 272–98.

McCaulay, Alistair, "Cyclical Attention to Saving," Technical Report, Working Paper 2020.

McKay, Alisdair and Johannes Wieland, "Lumpy Durable Consumption Demand and the Limited Ammunition of Monetary Policy," Technical Report, National Bureau of Economic Research 2019.

Nakamura, Emi and Jón Steinsson, "High Frequency Identification of Monetary Nonneutrality: The Information Effect," *Quarterly Journal of Economics*, 2018.

Nieuwerburgh, Stijn Van and Laura Veldkamp, "Information Immobility and the Home Bias Puzzle," *The Journal of Finance*, 2009, *64* (3), 1187–1215.

Ottonello, Pablo and Thomas Winberry, "Financial Heterogeneity and the Investment Channel of Monetary Policy," Technical Report, National Bureau of Economic Research 2018.

Ramey, Valerie, "Macroeconomic Shocks and Their Propagation," in "Handbook of Macroeconomics," Vol. 2, Elsevier, 2016, pp. 71–162.

Reis, Ricardo, "Inattentive Producers," *The Review of Economic Studies*, 2006, *73* (3), 793–821.

Reutskaja, Elena, Rosemarie Nagel, Colin Camerer, and Antonio Rangel, "Search Dynamics in Consumer Choice under Time Pressure: An Eye-tracking Study," *American Economic Review*, 2011, *101* (2), 900–926.

Sims, Christopher, "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

_ , "Rational Inattention and Monetary Economics," in "Handbook of Monetary Economics," Vol. 3, Elsevier, 2010, pp. 155–181.

Tanaka, Mari, Nicholas Bloom, Joel David, and Maiko Koga, "Firm Performance and Macro Forecast Accuracy," *Journal of Monetary Economics*, 2019.

Tenreyro, Silvana and Gregory Thwaites, "Pushing on a String: US Monetary Policy is Less Powerful in Recessions," *American Economic Journal: Macroeconomics*, 2016, *8* (4), 43–74.

Vavra, Joseph, "Inflation Dynamics and Time-varying Volatility: New evidence and an Ss Interpretation," *The Quarterly Journal of Economics*, 2014, *129* (1), 215–258.

Woodford, Michael, "Information-constrained State-dependent Pricing," *Journal of Monetary Economics*, 2009, *56*, S100–S124.

Zbaracki, Mark, Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen, "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," *Review of Economics and Statistics*, 2004, *86* (2), 514–533.

A Appendices

A.1 Additional Tables and Figures

Table A.1: Macroeconomic topics and keywords

Topic	**Keywords**
General	economic conditions
Output	GDP, economic growth, macroeconomic condition, construction spending, national activity, recession
Employment	unemployment, JOLTS, labor market, jobless claims, jobs report, non-farm payroll, ADP employment report, empoyment cost index
Consumption	consumer confidence, consumer credit, consumer sentiment, durable goods, personal income, retail sales
Investment	business inventories, manufacturing survey, factory orders, business outlook survey, manufacturing index, industrial production, business optimism, wholesale trade
FOMC	FOMC, monetary policy, quantitative easing
Housing	home sales, home prices, housing starts, housing market
Inflation	price index, price level, consumer price index, CPI, PMI, PPI, inflation, inflationary, disinflation, disinflationary, hyperinflation, hyperinflationary
Oil	oil prices, oil supply, oil demand

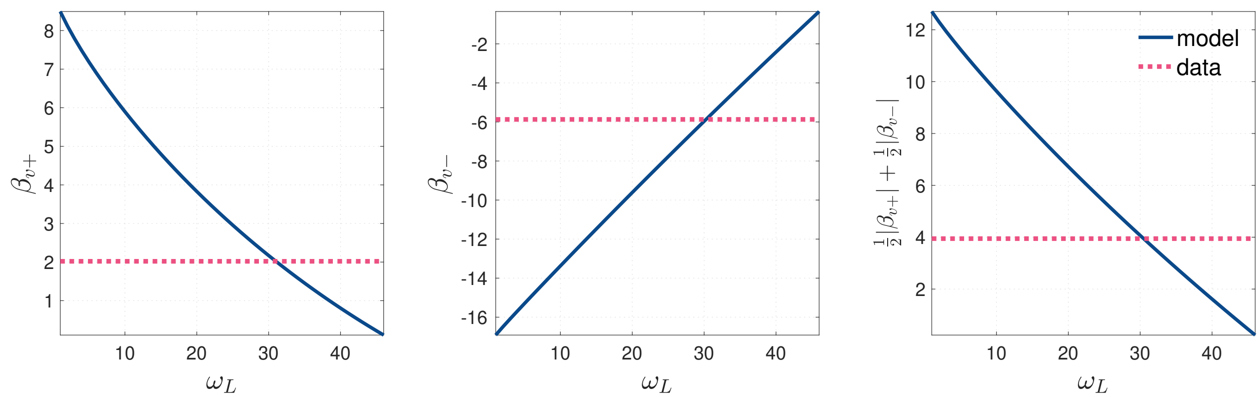
Notes: Dictionary of keywords used in constructed text-based attention measures. Keywords are based on names of macroeconomic releases from EconodayPlus, complemented with macroeconomic words and phrases from popular press.

Table A.2: Marginal effects of firm characteristics on likelihood of attention

	General	FOMC	GDP	Emp	Cons	Inflation	Inv	Oil	Housing
Age	0.014***	0.001***	0.004***	0.003***	0.006***	0.008***	0.001***	0.002***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log(assets)	0.024***	0.006***	0.022***	0.008***	0.017***	0.027***	0.002***	0.010***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Observations	135132	116257	133664	128704	134064	135001	105692	128815	128425
4-digit NAICS FE	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: Results of the probit regressions on the marginal effects of firm characteristics on the likelihood of firms paying attention to a macroeconomic topic. The dependent variables are the prevalence attention measures. 4-digit NAICS Industry fixed effects are included.

Figure A.1: Sensitivity of simulated moments to ω_L



Notes: Calibration plots showing simulated moments for a range of costs of information parameters (ω_L). We simulate models for a panel of 100 firms and for 1000 periods with 100 periods burn-ins. Simulated moments are generated with regressions discussed in the text in Section 5:

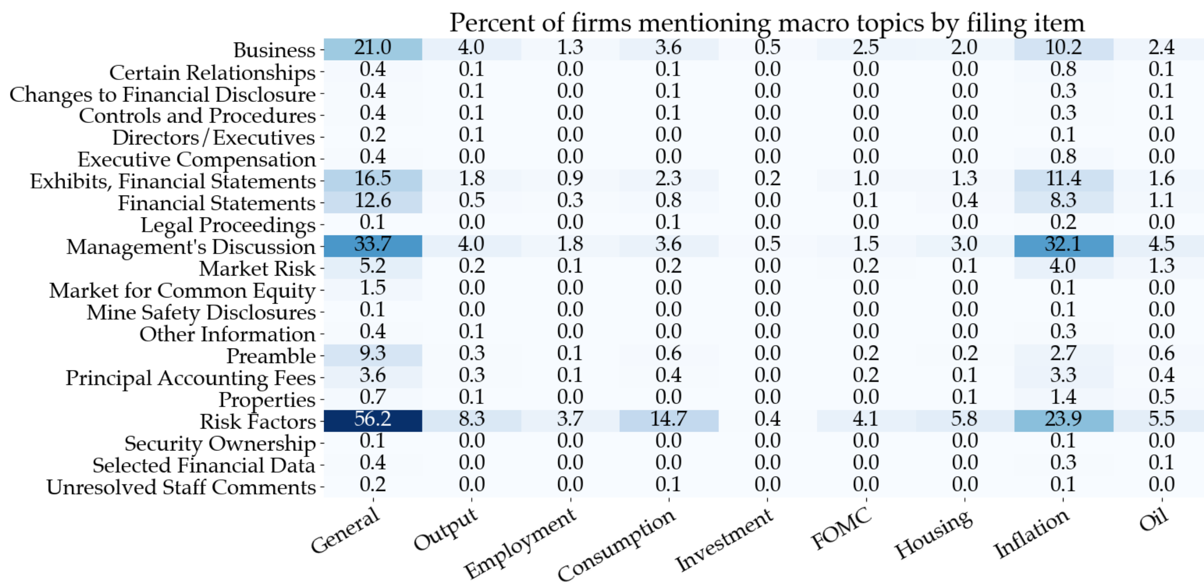
$$r_{it} = c + \beta_1 \mathbb{1}_{v>0} + \beta_{v+} v_t \mathbb{1}_{v>0} + \beta_{v-} v_t \mathbb{1}_{v<0} + \beta_d d_{it} + \beta_{dv+} d_{it} \nu_{it} \mathbb{1}_{v>0} + \beta_{dv-} d_{it} \nu_{it} \mathbb{1}_{v<0} + \varepsilon_{it}$$

The left panel shows the sensitivity of simulated β_{v+} to the calibration of ω_L; the middle panel shows the sensitivity of β_{v-}; the right panel shows the sensitivity of $\frac{1}{2}|\beta_{v+}| + \frac{1}{2}|\beta_v|$, which we use to calibrate ω_L to match the empirical moment in the data.

A.2 Itemized Frequency Search

10-K filings have standard formats and are organized in sections. We perform refined frequency counts for each of the section, or "items", to see where attention is concentrated in. Results of frequency counts of macroeconomic keywords by filing item are shown in Figure A.2 in the Appendix. Discussions of the macroeconomy are concentrated in Description of Business (Item 1), Risk Factors (Item 1A) and Management Discussion and Analysis of Financial Condition and Results of Operations (Item 7A).

Figure A.2: Firm attention by filing items



Notes: Heat map of firm attention by filing items. Each row represents a section ("item") of 10-K, and each column represents a macroeconomic topic. Darkness represents a higher fraction of firms that pay attention to a macroeconomic topic in an item.

A.3 Latent Dirichlet Allocation

To enable automated context detection, we use the Latent Dirichlet Allocation (LDA) model to uncover topics firms tend to discuss in conjunction with macro news. LDA (Blei et al., 2003) is an unsupervised learning algorithm aimed at grouping words in documents into meaningful topics. We apply LDA to texts in earning filings within 20 words surrounding a macroeconomic keyword and set the number of topics to be 10.

Following Hansen et al. (2018), we pre-process texts of 10-K filings for LDA as follows: we remove numbers and words that are only one character. Then we lemmatize to combine different word forms (for example, "operated" and "operates" are lemmatized to "operate"). The advantage of lemmatizing over stemming is that the resulting LDA outputs are more friendly to interpret. Our corpus include words and bigrams which appear for at least 20 times. We filter out words that occur in less than 20 documents or more than 50% of the documents. Then we transform the texts through bag-of-words representation.

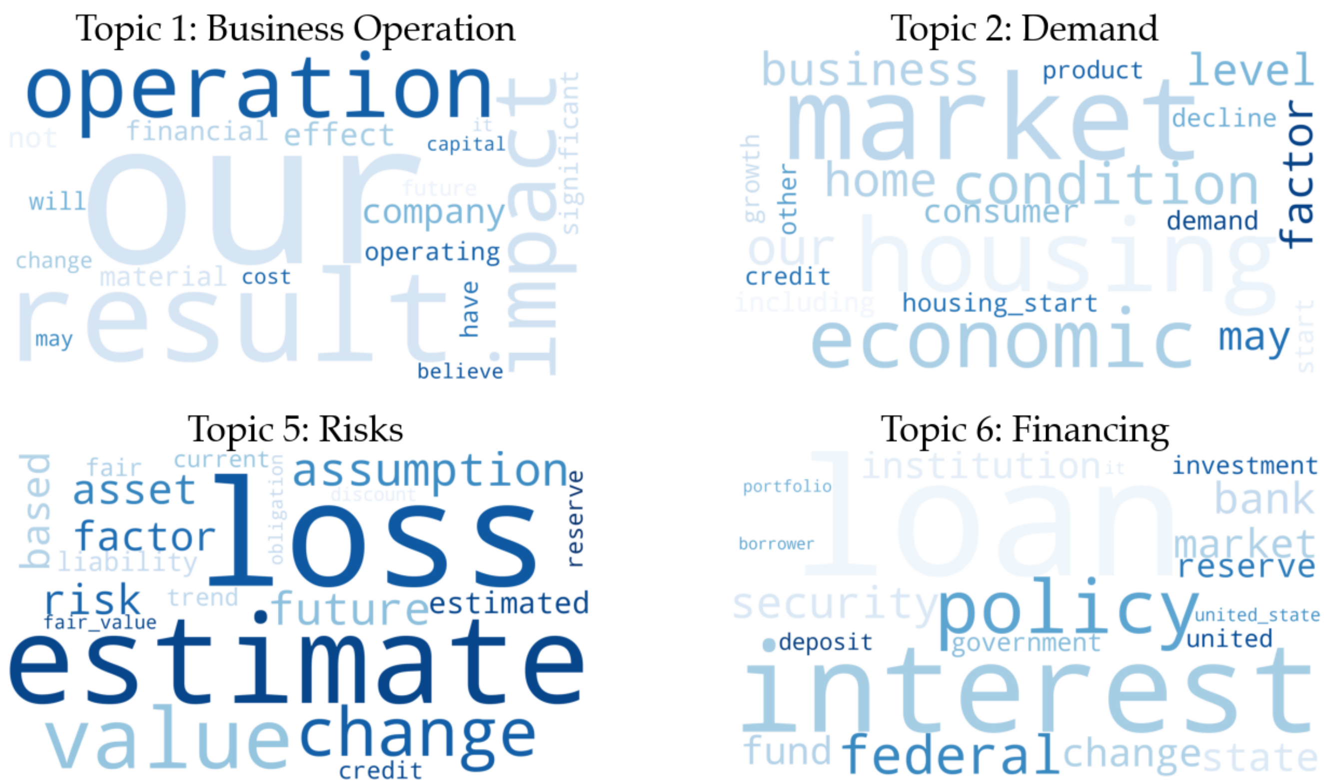
We model topics surrounding each of the nine macro categories for the attention measure, as well as an aggregate category containing keywords from all categories. Figures A.3 and A.4 visualize the LDA output surrounding keywords in all categories.Figure A.3 shows the heat map of LDA outputs. Each row represent a topic clustered by LDA, and the darkness of the cell within a topic represent the likelihood of a word to appear in the topic. Figure A.4 highlights the word cloud of selected topics in A.3.

Although LDA output does not label topics, it is natural to characterize some of the topics. Topic 1 relates to business operations, as firms discuss how macro conditions feed into into their daily operations; Topic 2 relates to demand, as firms track and gauge the aggregate demand; Topic 6 relate to financing costs, as firms pay attention to how monetary policy affect their financial costs, investment decisions, and portfolio holdings; Topic 10 relates to labor costs, as firms assess the tightness of the labor market. Rest of the topics relate to housing, currency, and risk factors.

Figure A.3: LDA output for texts surrounding all macro keywords

topic 1	our	result	operation	impact	company	effect	not	material	financial	significant	operating	will	have	change	may	future	cost	believe	capital	it
topic 2	market	housing	economic	condition	our	home	business	level	factor	may	consumer	including	demand	start	growth	decline	housing start	other	product	credit
topic 3	increase	lease	real	estate	real estate	index	year	property	annual	price	based	rent	consumer	term	adjustment	operating	rental	payment	building	expense
topic 4	currency	foreign	fluctuation	foreign currency	risk	exchange	dollar	country	political	international	change	tax	may	law	exposure	including	u	other	government	china
topic 5	loss	estimate	value	change	assumption	future	asset	risk	factor	based	estimated	liability	fair	trend	credit	reserve	current	fair value	obligation	discount
topic 6	loan	interest	policy	federal	security	bank	market	state	change	fund	institution	reserve	government	investment	united	deposit	united state	it	portfolio	borrower
topic 7	asset	return	statement	financial	plan	consolidated	interest	note	longterm	historical	expected	hedge	liability	performance	investment	data	pension	dollar	due	relative
topic 8	price	sale	million	year	increase	cost	due	increased	december	production	higher	primarily	net	compared	ended	volume	approximately	offset	fiscal	oil
topic 9	cost	company	service	contract	certain	adjusted	unit	of	our	agreement	cpi	be	equipment	customer	labor	health	facility	benefit	existing	to
topic 10	cash	claim	flow	cash flow	benefit	employee	stock	salary	share	shipment	legislative	senior	common	holding	vehicle	indexed	mac	restaurant	five	plan

Figure A.4: LDA output for texts surrounding all macro keywords: Selected topics



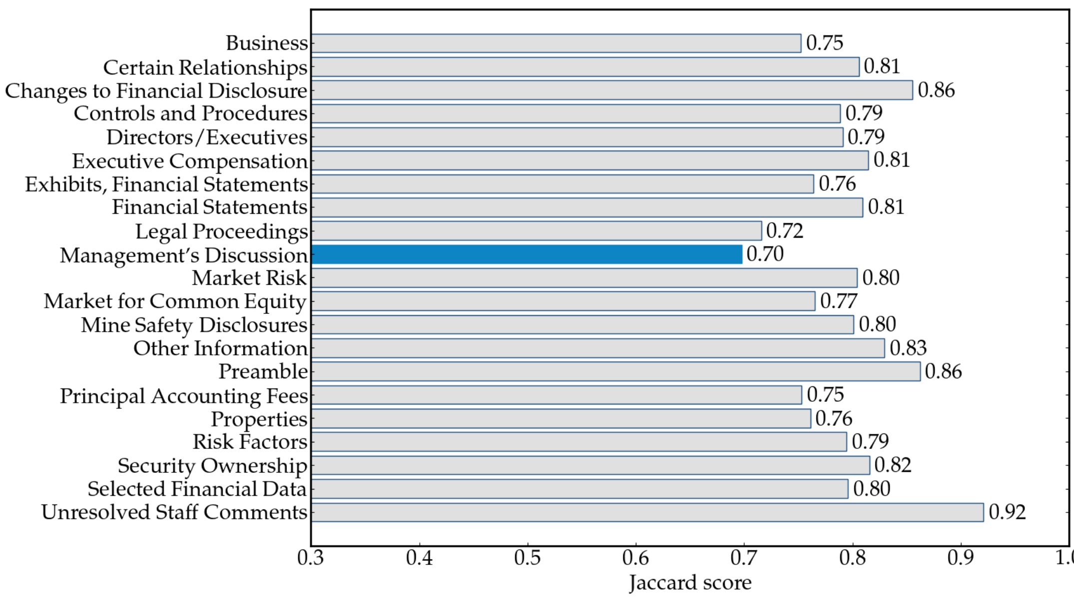
A.4 Lexical Similarity

Our measure of lexical similarity is a Jaccard score, $J(y_{it}, y_{it-1})$, which measures the share of unique non-stop words that appear between the current year's 10-K (y_i) compared to the previous year's 10-K (y_{it-1}).

$$J(y_i, y_{it-1}) = \frac{|y_i \cap y_{it-1}|}{|y_i \cup y_{it-1}|}$$

The Jaccard score is bounded by the unit interval, and is decreasing with the "uniqueness" of the text. Figure A.5 reports the average Jaccard score for each section of 10-K filings.

Figure A.5: Lexical similarity by section of 10-K filings



Notes: Average Jaccard scores for sections in 10-K filings. The Jaccard score is bounded by the unit interval. A high Jaccard score represents high lexical similarity between filings. The Management's Discussion section has the lowest level of lexical similarity in all 10-K sections.

We then restrict the attention measures to keywords mentioned in low Jaccard score sections: Business (Item 1) and Managment's Discussion (Item 7). We exclude Legal Proceedings (Item 3) that has a low Jaccard score to avoid false positives from legal languages. Regression results with attention restricted to low lexical similarity 10-K sections are reported in Table A.3.

Table A.3: Restricting attention to low lexical similarity 10-K sections

	(1)	(2)	(3)	(4)
Shock	4.13	4.13		
	(2.53)	(2.53)		
Attention		-0.03	-0.08*	-0.05
		(0.03)	(0.04)	(0.05)
Shock × Attn		0.02		
		(0.47)		
Shock × $\mathbb{1}_{v_t>0}$			4.55*	6.21**
			(2.62)	(2.53)
Shock × $\mathbb{1}_{v_t<0}$			-4.16	-1.45
			(4.29)	(4.36)
Shock × Attn × $\mathbb{1}_{v_t>0}$			0.79	0.50
			(0.50)	(0.47)
Shock × Attn × $\mathbb{1}_{v_t<0}$			-5.24**	-4.95**
			(2.10)	(2.09)
Observations	546596	546596	546596	409889
R^2	0.023	0.023	0.026	0.027
Clustered SE	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes
4-digit NAICS FE	yes	yes	yes	yes
excl. ZLB	no	no	no	yes
Wald Test p-value			0.010	0.020

Notes: Results from variants of estimating the baseline specification, restricting to 10-K items that discuss firm operations (Items 1 and 7):

$$r_{it} = \delta_j + \beta_{v+} v_t \mathbb{1}_{v_t>0} + \beta_{v-} v_t \mathbb{1}_{v_t<0} + \beta_{dv+} d_{it} v_t \mathbb{1}_{v_t>0} + \beta_{dv-} d_{it} v_t \mathbb{1}_{v_t<0} + \boldsymbol{\beta}_X \mathbf{X}_t + \varepsilon_{it}$$

where δ_j is an industry fixed effect, v_t is the monetary shock, D_{it} is the prevalence attention measure, and $\mathbf{X}_t$ contains the indicator variable for positive shocks $\mathbf{1}_{v_t>0}$ and firm level controls of size, age and leverage. We also include firm controls and industry fixed effects interacted with the monetary shocks. Standard errors are clustered at the shock level. We have normalized the sign of teh monetary shock ν_t so that a positive shock is expansionary (corresponding to a decrease in interest rates). Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

A.5 Approximation of Firm Profits in the Stylized Model

Under second-order approximation around the non-stochastic steady state, the log approximation of a firm's profits, denoted by $\hat{\pi}(s_t, a_t)$, is given by:

$$\begin{aligned}
\hat{\pi}(s_t, a_t) &= \pi(\bar{s}, \bar{a}) + \pi_s(\bar{s}, \bar{a})\bar{s}\hat{s}_t + \pi_a(\bar{s}, \bar{a})\bar{a}\hat{a}_t + \frac{1}{2}\pi_{ss}(\bar{s}, \bar{a})\bar{s}^2\hat{s}_t^2 + \frac{1}{2}\pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\hat{a}_t^2 + \pi_{sa}(\bar{s}, \bar{a})\bar{s}\bar{a}\hat{s}_t\hat{a}_t \\
&= \pi(\bar{s}, \bar{a}) + \pi_s(\bar{s}, \bar{a})\bar{s}\hat{s}_t + \frac{1}{2}\pi_{ss}(\bar{s}, \bar{a})\bar{s}^2\hat{s}_t^2 + \frac{1}{2}\pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\hat{a}_t^2 - \pi_{aa}(\bar{s}, \bar{a})\bar{a}\bar{s}\hat{a}_t\hat{s}_t \\
&= \pi(\bar{s}, \bar{a}) + \pi_s(\bar{s}, \bar{a})\bar{s}\hat{s}_t + \frac{1}{2}\left(\pi_{ss}(\bar{s}, \bar{a})\bar{s}^2 - \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\right)\hat{s}_t^2 + \frac{1}{2}\pi_{aa}(\bar{s}, \bar{a})\bar{a}^2(\hat{a}_t - \hat{s}_t)^2
\end{aligned}$$

In the second line, $\pi_a(\bar{s}, \bar{a}) = 0$ because of optimal choice. In addition, the assumption that $a = s$ under full information yields $\pi_a(a, a) = 0\ \forall a$, which implies $\pi_{sa}(\bar{s}, \bar{a}) = -\pi_{aa}(\bar{s}, \bar{a})$. The third line added and subtracted $\frac{1}{2}\pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\hat{s}_t^2$ to complete squares and used the fact that $\bar{a} = \bar{s}$ in the steady state. The resulting expression is equation (1).

A.6 Proof of Proposition 1

Proof. We consider the responses of returns to an aggregate shock ε. Holding all else equal, that is, $\pi_{ss}^k(s,a) = \pi_{ss}(s,a)$ and $\pi_{aa}^k(s,a) = \pi_{aa}(s,a)$ for all firms k, we can show the following for heterogeneity in exposure and in attention.

(i) **Exposure**: Let firms be heterogeneous in exposure and homogeneous in attention. Specifically, suppose firm i is more exposed to macro conditions than firm j, that is, $\pi_s^i > \pi_s^j > 0$. We consider how heterogeneity in exposure affects return elasticity for cases in which both firms are attentive and both are inattentive.

  (a) Case 1 (both firms attentive): When firms are both attentive, $\hat{a}_t = \hat{s}_t$. Then by equation (1) we can derive the return elasticity with respect to the aggregate shock to be:

$$\frac{\partial r_k}{\partial \varepsilon} = \frac{\partial \hat{\pi}_k}{\partial \varepsilon} = \pi_s^k(\bar{s},\bar{a})\bar{s} + \left(\pi_{ss}(\bar{s},\bar{a})\bar{s}^2 - \pi_{aa}(\bar{s},\bar{a})\bar{a}^2\right)\varepsilon \quad \text{for firm } k = i, j.$$

  Therefore, the return elasticity for firms i is larger for the return elasticity for firm j for all magnitudes of shocks

$$\frac{\partial r_i}{\partial \varepsilon} - \frac{\partial r_j}{\partial \varepsilon} = \pi_s^i(\bar{s},\bar{a})\bar{s} - \pi_s^j(\bar{s},\bar{a})\bar{s} > 0$$

  because $\pi_s^i > \pi_s^j > 0$.

  (b) Case 2 (both firms inattentive): When both firms are inattentive, the return elasticity with respect to the shock can be expressed as:

$$\frac{\partial r_k}{\partial \varepsilon} = \frac{\partial \hat{\pi}_k}{\partial \varepsilon} = \pi_s^k(\bar{s},\bar{a})\bar{s} + \left(\pi_{ss}(\bar{s},\bar{a})\bar{s}^2 - \pi_{aa}(\bar{s},\bar{a})\bar{a}^2\right)\varepsilon$$
$$+ \pi_{aa}(\bar{s},\bar{a})\bar{a}^2(f_k(\varepsilon) - \varepsilon)(f_k'(\varepsilon) - 1) \quad \text{for firm } k = i, j.$$

  Since firms are only heterogeneous in exposure, the second and third term in the above expression for return elasticity is the same for both firms. Therefore:

$$\frac{\partial r_i}{\partial \varepsilon} - \frac{\partial r_j}{\partial \varepsilon} = \pi_s^i(\bar{s}, \bar{a})\bar{s} - \pi_s^j(\bar{s}, \bar{a})\bar{s} > 0$$

which is also independent of the magnitude of ε.

(ii) **Attention**: Now instead let firms be heterogeneous in attention and homogeneous in exposure, so the attentive firm i has $f_i'(\varepsilon) = 1$, the inattentive firm j has $f_j'(\varepsilon) < 1$, and both firms have $\pi_s^i = \pi_s^j$. The return elasticity for attentive and inattentive firms can be expressed as:

$$\frac{\partial r_i}{\partial \varepsilon} = \pi_s(\bar{s}, \bar{a})\bar{s} + \left(\pi_{ss}(\bar{s}, \bar{a})\bar{s}^2 - \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\right)\varepsilon \tag{11}$$

$$\frac{\partial r_j}{\partial \varepsilon} = \pi_s(\bar{s}, \bar{a})\bar{s} + \left(\pi_{ss}(\bar{s}, \bar{a})\bar{s}^2 - \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\right)\varepsilon + \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2(f_j(\varepsilon) - \varepsilon)(f_j'(\varepsilon) - 1) \tag{12}$$

since firms are homogenous in exposure: $\pi_s^i = \pi_s^j = \pi_s$. The relative magnitude of return elasticities between attentive and inattentive firms depends on the sign of the shock ε. Specifically, we consider three cases.

(a) Zero shock ($\varepsilon = 0$): Since $f(0) = 0$, (11) and (12) lead to:

$$\frac{\partial r_i}{\partial \varepsilon} = \pi_s(\bar{s}, \bar{a})\bar{s} = \frac{\partial r_j}{\partial \varepsilon}$$

(b) Positive shock ($\varepsilon > 0$): Since $\varepsilon_t > f_j(\varepsilon_t) > 0$,

$$\frac{\partial r_j}{\partial \varepsilon} - \frac{\partial r_i}{\partial \varepsilon} = \underbrace{\pi_{aa}(\bar{s}, \bar{a})\bar{a}^2}_{<0}\underbrace{(f_j(\varepsilon) - \varepsilon)}_{<0}\underbrace{(f_j'(\varepsilon) - 1)}_{<0} < 0$$

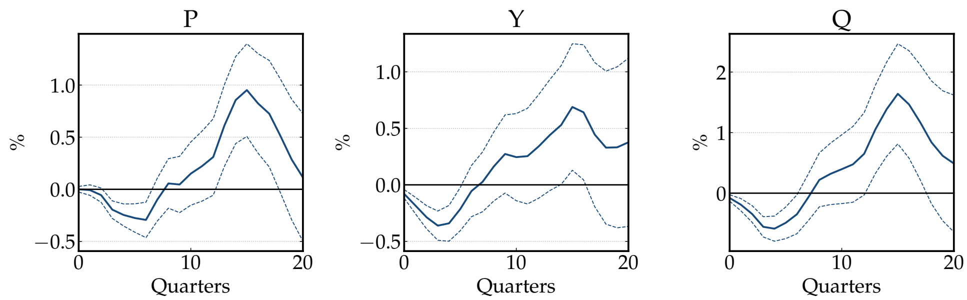
(c) Negative shock ($\varepsilon < 0$) Since $\varepsilon_t < f_j(\varepsilon_t) < 0$,

$$\frac{\partial r_j}{\partial \varepsilon} - \frac{\partial r_i}{\partial \varepsilon} = \underbrace{\pi_{aa}(\bar{s}, \bar{a})\bar{a}^2}_{<0}\underbrace{(f_j(\varepsilon) - \varepsilon)}_{>0}\underbrace{(f_j'(\varepsilon) - 1)}_{<0} > 0$$

■

A.7 Passthrough Regressions

Figure A.6: Passthrough of rates to nominal demand



The passthrough of nominal interest rate change to nominal demand change is estimated with local projections (Jordà, 2005). We estimate the following model for horizons $h = 1, 2, \cdots, 20$:

$$\Delta_h y_{t-1,t+h} = \alpha_h + \beta_h \varepsilon_t^{\mathtt{i}} + u_{th}$$

where y is the variable of interest, and $\varepsilon_t^{\mathtt{i}}$ is a shock to the nominal interest rate. Path of β_h informs the cumulative changes in the dependent variable in response to the interest rate shock.

The dependent variables are U.S. manufacturing output over the sample period of 1994 to 2019. We estimate the responses of manufacturing prices, real output and nominal output. We time aggregate high-frequency monetary policy shocks to quarterly to match the frequency of dependent variables. Figure A.6 shows the results of the local projection. A one percentage point expansionary shock to the interest rate leads to about 1.6 percent peak increase in nominal demand.