BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel—2022-3

Last updated: January 11, 2022

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

by Wenting Song[1] and Samuel Stern[2]

[1]Canadian Economic Analysis Department, Bank of Canada
wsong@bankofcanada.ca

[2] University of Michigan
sternsa@umich.edu

DOI: https://doi.org/10.34989/swp-2022-3 | **ISSN** 1701-9397

Acknowledgements

We would like to thank John Leahy, Pablo Ottonello, Linda Tesar, and Toni Whited for generous advice and guidance. We also thank Sushant Acharya, Hassan Afrouzi, Martin Eichenbaum, Andrei Levchenko, Juan Herreno, Yulei Luo, Alistair Macauley, Albert Marcet, Michael McMahon, Laura Veldkamp, and the participants at numerous seminars and conferences for helpful comments and discussions.

Abstract

This paper provides direct evidence of the importance of firm attention to macro-economic dynamics. We construct a text-based measure of firm attention to macro-economic news and document firm attention that is polarized and countercyclical. Differences in attention lead to asymmetric responses to monetary policy: expansionary monetary shocks raise market values of attentive firms more than those of inattentive firms, and contractionary shocks lower values of attentive firms by less. We use the measure to calibrate a quantitative model of rationally inattentive firms with heterogeneous costs of information. 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.

Topics: Business fluctuations and cycles, Inflation and prices, Monetary policy
JEL codes: D83, E44, E52

1 Introduction

Public information often goes unused because attention is scarce. Rational inattention models pioneered by Sims (2003) and a broader set of incomplete-information models (Mankiw and Reis, 2002; Woodford, 2009)[1] consider firm managers who gather information to maximize value while facing cognitive costs of processing information. Inattention provides an intuitive microfoundation for monetary policy non-neutrality in which firm managers misinterpret nominal monetary policy as shocks to real demand. However, empirically assessing the importance of attention is challenging because neither a firm's allocation of attention nor information-processing costs are readily observable.

This paper is one of the first to provide direct evidence of the importance of firm attention to macroeconomic dynamics using a novel text-based measure of firm attention. We document countercyclical firm attention and uncover substantial heterogeneity in attention across firms. Moreover, our measure is consistent with the asymmetric prediction of inattention models that attentive firms exhibit higher profit semi-elasticities in response to expansionary monetary shocks and lower semi-elasticities following contractionary shocks. We then use this measure to calibrate information costs in a quantitative general equilibrium model with rationally inattentive firms, and show that firm inattention generates monetary non-neutrality. Together with our empirical evidence on countercyclical firm attention, this result suggests that aggregate attention to macroeconomic conditions is an important dimension of state-dependence in monetary policy.

To construct our attention measure, we compile a corpus based on approximately 200,000 annual SEC filings of US publicly-traded firms and search each document 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 document two 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 is countercyclical. Among the remaining firms

[1] Additional work includes Lucas (1972); Angeletos and La'O (2013); Gabaix (2019); Farhi and Werning (2019).

with time-varying attention, the number of firms that reference macroeconomic news rises notably during recessions.

Our main empirical result validates that our text-based methodology effectively measures attention by testing for an asymmetry in firm performance that is predicted by inattention models: Following a macroeconomic shock, firms with greater information-processing capacity should respond closer to the optimal response regardless of the shock's direction. Therefore, more attentive firms should exhibit higher profit elasticities in response to positive shocks and lower elasticities in response to negative shocks as they update prices more accurately than inattentive competitors. We test for this asymmetry using an event-study design that exploits high-frequency variation in firms' market values around FOMC announcements. This test requires combining our prevalence attention measure with daily CRSP stock prices, quarterly Compustat firm financials, and high-frequency monetary shocks (constructed as in Gürkaynak et al., 2005; Gorodnichenko and Weber, 2016; Nakamura and Steinsson, 2018).

Consistent with the theoretical prediction, expansionary monetary shocks raise stock returns of attentive firms by 2% more than those of their inattentive peers, whereas contractionary shocks lower returns of attentive firms by 6% less. The suboptimal responses to monetary shocks by inattentive firms are direct evidence of the cost of inattentive behavior. Moreover, the asymmetry invalidates some concerns about measuring firm attention with text analysis. Concern that filings contain macroeconomic buzzwords as a form of cheap talk to appease investors would imply a zero effect; concern that firms mention keywords solely as a function of exposure to monetary policy would imply symmetric responses to monetary shocks; and concern that stock returns vary with investor attention rather than firm attention would also fail to explain the asymmetric responses.

We then use our attention measure in a quantitative rational inattention model to study the aggregate implications of the heterogeneity in firm attention. Firms with heterogeneous information costs optimally trade off between the precision of their signals of aggregate demand and the cost of acquiring and processing information. Information-processing costs and the distribution of firm attention are calibrated using our text-based attention measure. Consistent with our empirical findings, attentive firms in the calibrated model have higher semi-elasticities to expansionary monetary shocks and lower semi-elasticities to contractionary shocks. We incorporate the empirical countercyclicality of firm attention to show that the efficacy of monetary policy declines as the fraction of attentive firms increases and more firms set prices closer to the optimum. This new interpretation of attention-dependent monetary policy implies that central banks should expect the effects of policy to be weaker when an aggregate shock has already drawn firm attention to macroeconomic policy.

Related literature Our paper contributes to four strands of literature. First, we contribute to the empirical literature on macroeconomic expectations by developing an ongoing, broad-based measure of firm attention that extends back to the mid-1990s. Recent literature has highlighted the importance of expectations for macroeconomic policy.[2] and consequently the need for empirical measures[3] Existing research has successfully measured attention in lab experiments (Reutskaja et al., 2011), field experiments (Bartoš et al., 2016; Fuster et al., 2018), and for individual consumers (McCaulay, 2020). Our methodology complements those measures as well as survey-based evidence on firm expectations by Tanaka et al. (2019), Coibion et al. (2018), Afrouzi (2020), and Candia et al. (2021), and enables researchers to explore questions that lie outside the coverage of existing surveys.

Second, our findings on firm inattention lend empirical support to a broad body of theoretical work on incomplete information as a source of monetary non-neutrality (Sims, 2003; Mankiw and Reis, 2002; Woodford, 2009). Microfoundations proposed in rational inattention and sticky information models are successful in explaining firm pricing (Mackowiak and Wiederholt, 2009; Afrouzi and Yang, 2021), business cycles (Maćkowiak and Wiederholt, 2015), asset prices (Van Nieuwerburgh and Veldkamp, 2009), discrete choices (Matějka and McKay, 2015; Caplin et al., 2019), and reconciling micro and macro evidence (Auclert et al., 2020). However, the lack of measurement on firm attention makes it challenging to assess the empirical importance of these microfoundations. Our results estimate a substantial cost of information frictions in the US data, providing direct support for these theories.

Our findings on the relationship between countercyclical attention and monetary policy efficacy relate to existing literature on state dependencies of monetary policy. Tenreyro and

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

[3]Mackowiak et al. (2021) and Gabaix (2019) provide comprehensive surveys of existing measure of attention.

Thwaites (2016) estimates non-linear responses in monetary policy, which are weaker in recessions than in expansions. Vavra (2014), McKay and Wieland (2019), and Ottonello and Winberry (2020) consider volatility, durable consumption, and default risk as other channels through which state dependency arises. This paper suggests that attention may be an important source of state dependency of monetary policy.

Finally, our paper relates to a broader and emerging literature that brings natural language processing techniques to economics. The seminal work of Loughran and McDonald (2011) applies the "bag of words" method to firm filings and develops word lists specific to economic and financial texts. Recent work uses textual analysis to study financial constraints (Buehlmaier and Whited, 2018), central bank communication (Hansen et al., 2018), firm-level political risk (Hassan et al., 2019), inflation expectation formation (Larsen et al., 2021), and uncertainty (Handley and Li, 2020). We contribute to this literature by constructing a set of keyword dictionaries based on macroeconomic news releases that correspond to nine macroeconomic topics. While this paper focuses on attention to monetary policy, our method for measuring attention and its effects can be generalized to the other macroeconomic topics.

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 and exposure to macro shocks and derive the predicted asymmetry; 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; and Section 6 concludes.

2 Textual Measure of Attention

This section presents our measure of firm attention and documents several stylized facts about how attention varies between firms and has evolved over our sample period.

2.1 SEC filings

To measure firm attention, we employ the universe of annual 10-K filings with the US 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) require a more extensive discussion of business conditions and audited financial statements, while the quarterly filings (Form 10-Q) are usually less descriptive and only require 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 require 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 appear 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, inflation, 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 mentions any keyword related to a given topic k in period t:

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

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:

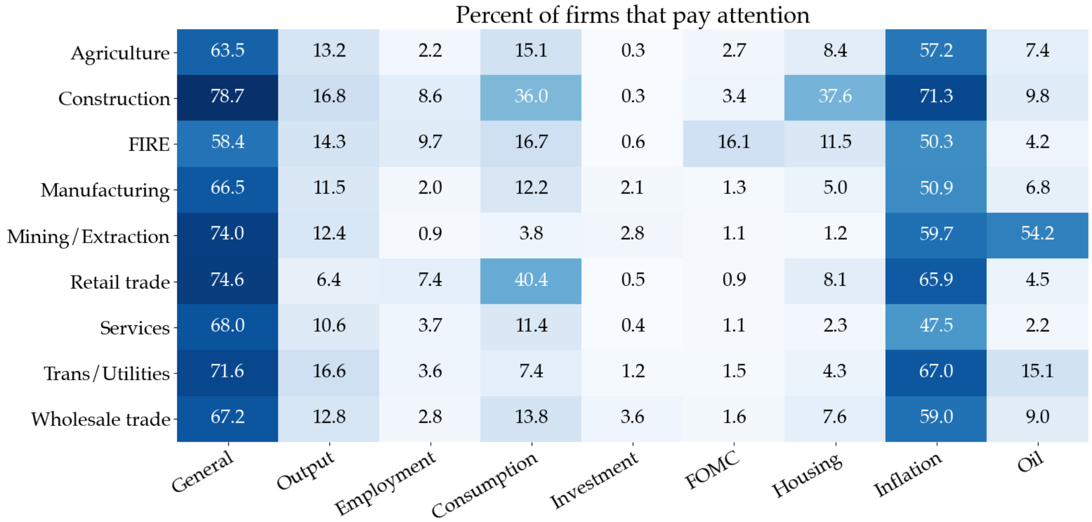
$$s_{it}^k = \frac{\text{Total topic } k \text{ words}_{it}}{\text{Total words}_{it}} \qquad \text{(intensity)}$$

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.

Sense check of the textual measure As a preliminary sense check of the textual measure, Table A.2 in the Appendix reports the summary statistics of firm characteristics by attention. Attentive firms, whose prevalence attention to the general topic is nonzero in any year in the sample period, tend to be larger, older, and slightly less levered than their inattentive counterparts.

We then investigate the cross-industry variation in attention. Figure 1 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

Figure 1: 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.

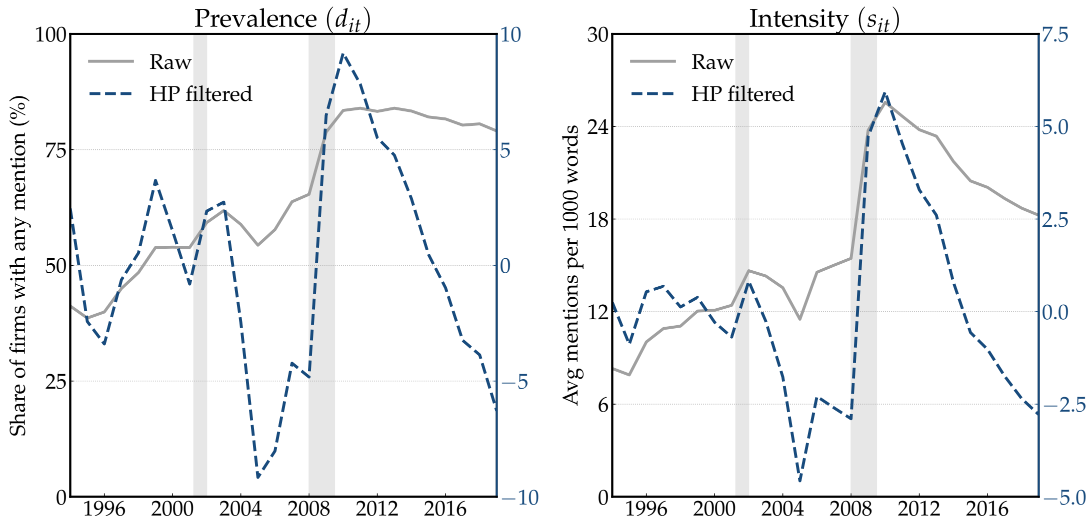
results should be interpreted across industry rather than across topic.

For each macro topic, attention is highest in industries for which profits are most sensitive to the topic a priori. For example, Mining, Oil, and Gas (NAICS 21) have 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 serve as sense 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 attention to that topic, and some industries appear to have greater overall interest in aggregate economic conditions.

Figure 2: Time series of attention to "economic conditions"



Notes: Time series of firm attention to the keyword "economic conditions." The 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 with smoothing factor 400. Shares are reported in percent.

2.3 Stylized facts about firm attention

We now apply our prevalence and intensity measures to document two stylized facts about time and firm variation in attention: firm attention in the US is countercyclical and polarized. We then investigate firm characteristics that drive attention.

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 2 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 with smoothing factor 400.

The share of firms that mentions "economic conditions" increases over the sample period, 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 keywords "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.

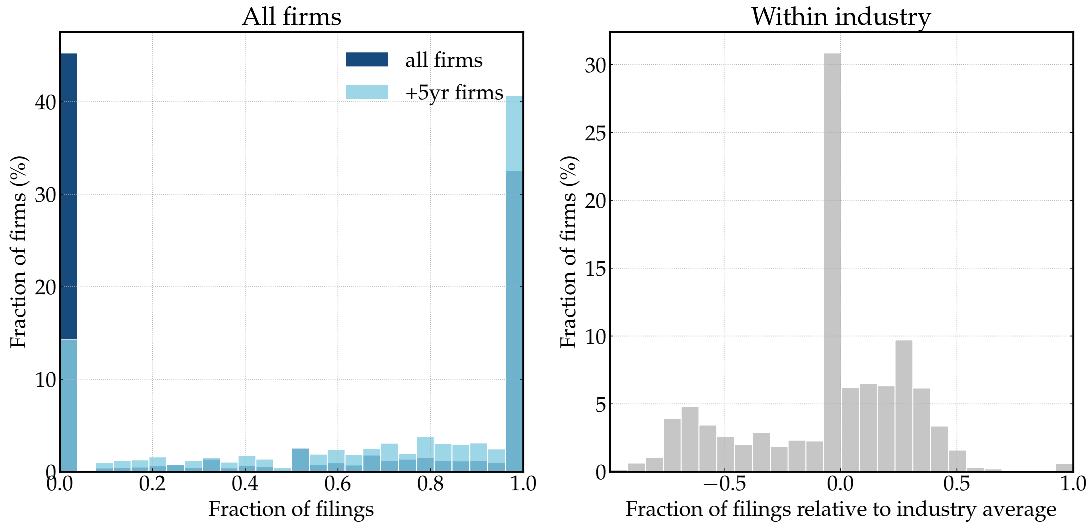
Countercyclical attention exhibited in Figure 2 is consistent with predictions in Mackowiak and Wiederholt (2009), which models firms that allocate attention between aggregate and idiosyncratic conditions. Their model predicts that firms will pay more attention to aggregate conditions in downturns if those conditions become more uncertain. This result is also consistent with Chiang (2021), which develops a generalized information structure where agents pay greater attention to uncertain aggregate conditions when expecting a bad economic state, which subsequently generates countercyclical attention and uncertainty.

Polarization in firm attention Heterogeneous attention to publicly available news about US 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 US economic conditions, and we should expect firms with unlimited information-processing bandwidth to incorporate this news into their decision-making.[4] Evidence of heterogeneity is to the contrary and provides new insights into how firms allocate attention differently.

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 never mentions "economic conditions" over the sample period, whereas a firm with a value of 1 mentions that phrase in every filing. Most notably, firms are concentrated at each extreme: 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

[4]See, for example, Jaimovich and Rebelo (2009).

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



Notes: Histograms of the share of filings by a firm 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. 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. Shares of firms on the vertical axes are reported in percent.

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 attention even after accounting for industry averages. Aside from a high concentration of attention at the industry average, demeaned attention also appears bimodally dispersed.

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 insensitive to aggregate fluctuations. Our main empirical analysis in Section 4 will focus on disentangling our hypothesized attention channel from this alternative exposure channel. We test our hypothesis by separately estimating the response of stock prices to positive and negative macro shocks. If firms discuss macro news more often because they are more exposed to aggregate fluctuations, then "attentive" firms would profit more from a positive shock and lose more from a negative shock, generating symmetric relative responses to macro shocks. On the other hand, if the text-based measures indeed capture attention, then attentive firms would outperform inattentive competitors regardless of the direction of the shock, resulting in asymmetric relative responses. The theoretical framework in Section 3 discusses the mechanism in detail.

Firm characteristics and attention We next study the cross-sectional and time-series relationship between firm characteristics and attention by estimating a series of annual, univariate regression models of the following form:

$$\text{Cross-sectional variation:} \quad d_{it} = \delta_t + \delta_j + \beta \cdot x_{it} + \varepsilon_{it} \tag{1}$$

$$\text{Time-series variation:} \quad d_{it} = \delta_i + \beta \cdot x_{it} + \varepsilon_{it}, \tag{2}$$

where x_{it} is either firm size, age, or leverage[5] and d_{it} is our prevalence measure of attention. Equation (1) includes a time fixed effect, δ_t, and a sector fixed effect, δ_j, at 4-digit NAICS level, to focus our analysis on the cross-section. Equation (2) includes a firm fixed effect, δ_i, to study variation in attention over a firm's life cycle.

Data on firm characteristics are from `Compustat`. Size is measured as the log of total assets, age as the years since a firm's first appearance in the Compustat sample, and leverage as the debt-to-asset ratio. All firm covariates are standardized across all observations so that the unit is one standard deviation.

Panel (a) in Table 2 shows that in the cross section in a given year, firm attention to general economic news increases with size and age, and decreases with leverage. Panel (b) in Table 2 shows that as a firm grows older and larger, it is more likely to become attentive to macro news.

In addition to firm characteristics, firm managers' limited cognitive bandwidth is a po-

[5] Existing literature has found each of these characteristics to be relevant for the transmission of macroeconomic policy (Gertler and Gilchrist, 1994; Cloyne et al., 2018; Ottonello and Winberry, 2020).

Table 2: Firm characteristics and attention

(a) Cross-sectional variation

	(1)	(2)	(3)
Size	0.0858***		
	(0.0013)		
Age		0.0055***	
		(0.0016)	
Leverage			-0.0096***
			(0.0010)
Observations	131885	131421	131384
R^2	0.265	0.243	0.243
Time-industry FE	yes	yes	yes
Firm FE	no	no	no

(b) Time-series variation

	(1)	(2)	(3)
Size	0.2345***		
	(0.0037)		
Age		0.2100***	
		(0.0016)	
Leverage			-0.0015
			(0.0011)
Observations	131896	131431	131396
R^2	0.553	0.600	0.538
Time-industry FE	no	no	no
Firm FE	yes	yes	yes

Notes: Panel (a) reports the estimated coefficient β from $d_{it} = \delta_t + \delta_i + \beta \cdot x_{it} + \varepsilon_{it}$, and Panel (b) reports the estimated coefficient β from $d_{it} = \delta_i + \beta \cdot x_{it} + \varepsilon_{it}$, where x_{it} is the firm size, age, or leverage, d_{it} is the prevalence attention to general economic news, δ_i is a firm fixed effect, δ_j is an industry fixed effect (4-digit NAICS), and δ_t is a time fixed effect.

tential source of information frictions. In Appendix Table A.3, we proxy for management quality with data on board members' education levels from `BoardEx`. Consistent with theory, firms whose board members have attained higher levels of education are more attentive to macro news.

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 the transmission of macroeconomic policy. Before doing so, we address a key identification challenge: whether our text-based attention measures identify differences in firm attention to macroeconomic conditions, conditional on firm characteristics, rather than differences in exposure to those conditions. To confront this 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 minimal to highlight the key mechanisms for attention and exposure. In Section 5, we expand the model environment 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 that a monopolistically competitive firm sets and s as the exogenous optimal price determined by factors outside of that 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[6]

$$\hat{\pi}(\hat{s},\hat{a}) = \pi_s(\bar{s},\bar{a})\bar{s}\hat{s} + \frac{1}{2}\left(\pi_{ss}(\bar{s},\bar{a})\bar{s}^2 - \pi_{aa}(\bar{s},\bar{a})\bar{a}^2\right)\hat{s}^2 + \frac{1}{2}\pi_{aa}(\bar{s},\bar{a})\bar{a}^2(\hat{a}-\hat{s})^2, \tag{3}$$

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. A firm is more exposed to aggregate conditions if its profits are more sensitive to aggregate shocks, while a firm is more attentive if its actions are more sensitive to shocks. Definitions 1 and 2 formalize these 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.

[6]Under this approximation, $\pi_a(s,a)$ drops out because of the first-order condition and assumption that $a^* = s$ at the optimum. Appendix D.1 contains detailed derivations of the approximation.

***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 to come.

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:

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

Realized equity returns, measuring the log change in a firm's value around an aggregate shock, are given by:

$$r = \hat{v} - \hat{v}_{-1}, \tag{4}$$

where $\hat{v} \equiv \log V - \log \bar{V}$ denotes the log deviation of firm value from the steady state, and $\hat{v}_{-1} \equiv \log \mathbb{E}_{-1} V - \log \bar{V}$ denotes the log deviation of firm value before the shock is realized.

Proposition 1 highlights the asymmetric responses of stock returns to positive and negative aggregate shocks that result from the attention channel and the symmetric responses from the exposure channel.

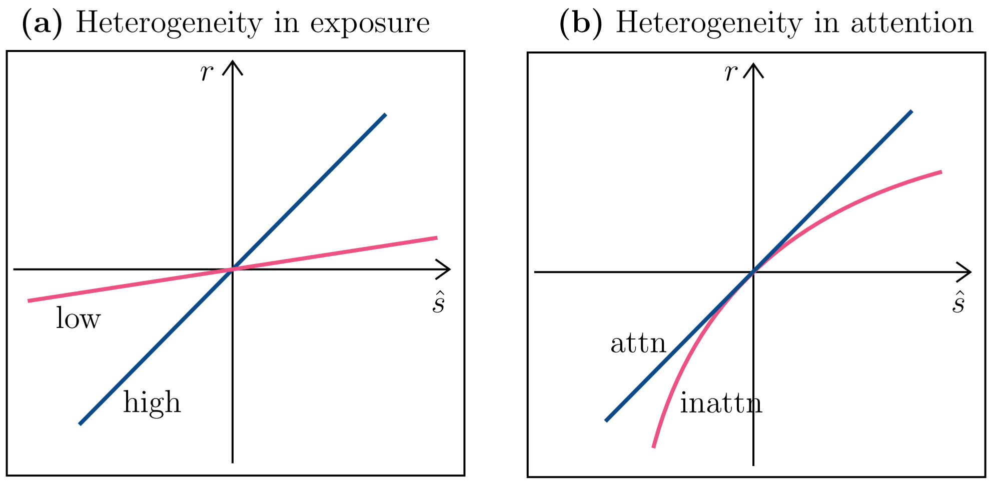
***Proposition* 1.** *The return elasticity with respect to aggregate shocks for the exposure and the attention channels can be characterized as below:*

(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, the return elasticity of a positive (expansion-*

Figure 4: Model predictions for exposure vs 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 highly exposed to macro conditions (*high*) and firms that are unexposed (*low*). Right panel shows return elasticity for attentive firms (*attn*) and inattentive firms (*inattn*). Exposure and attention are as defined in the main text.

ary) shock is higher for the attentive firm i than that 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:

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

Proof. See Appendix D.2 ■

Figure 4 illustrates the predictions from Proposition 1. In Panel (a), firms are heterogeneous in their exposures to aggregate shocks, and those with high exposure exhibit higher return elasticities to aggregate shocks 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 response to a negative shock, returns of both types of firms decrease, but returns of attentive firms drop by less.

This asymmetry in return elasticities is a unique feature of the attention channel and allows us to distinguish between the effects of firm attention and exposure to macro news. In the next section, we use this predicted asymmetry to show that our text-based measure correctly identifies firm attention, and then estimate the cost of inattention based on the difference in return elasticities for positive and negative shocks.

4 Empirical Analysis

Given our attention measures and theoretical predictions, we set out to test the hypothesis that attentive firms respond to macro shocks better than inattentive firms. We use a high-frequency identification strategy that isolates plausibly exogenous shocks to monetary policy from FOMC announcements and compares changes in stock prices of attentive and inattentive firms within a similarly narrow window around these announcements. We implement our empirical analysis with monetary policy shocks since they are familiar and well-identified[7] though the mechanism highlighted in our stylized inattention model is general and can be applied to other aggregate shocks with the corresponding attention measure.

Stock prices are a particularly informative outcome variable because they are forward-looking and quickly reflect changes in expected future profits. By focusing on the high-frequency windows of stock price movements, we are able to separate effects of monetary surprises from other confounding factors. More direct measures of firm responses, such as investment and hiring decisions, are only observed over longer time horizons and are confounded by other factors that influence firms' choices.

To best isolate the effects of attention, our baseline specification controls for firm size, age, leverage, and industry measured by 4-digit NAICS. The underlying identifying assumption is that firms have similar exposure to monetary policy shocks within a narrowly defined industry after conditioning on firm characteristics and financial structure. Residual variation in stock prices can then be attributed to firm attention rather than cross-firm variation in

[7] Ramey (2016) provides a comprehensive survey on the efforts to identify monetary shocks.

the exposure to monetary policy.

4.1 Data

Monetary policy shocks are constructed using the high-frequency identification strategy developed in Cook and Hahn (1989) and Gürkaynak et al. (2005), and used recently in Gorodnichenko and Weber (2016), Nakamura and Steinsson (2018), and Ottonello and Winberry (2020). 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,

$$\nu_t = \frac{D}{D-\tau}(ff^0_{t+\Delta t^+} - ff^0_{t-\Delta t^-}), \tag{5}$$

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 15 minutes before and 45 minutes 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 in Gorodnichenko and Weber (2016) and Nakamura and Steinsson (2018) for monetary shocks from 1994 to 2014. For easier interpretation of our empirical results, we normalize the sign of the monetary shock so that a positive shock is expansionary (corresponding to a decrease in interest rates).

Firm outcome and control variables are constructed using `CRSP` and `Compustat` data. Daily stock returns are measured as the open-to-close change in stock prices on the day of an FOMC announcement. Firm size, age, and industry controls are constructed as described in Section 2.3.

Firm attention is measured using the *prevalence* 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 annual filing rather than the filing in the same year as the FOMC announcement. This modification precludes the possibility that firms are identified as attentive to an FOMC announcement that inspired their attention.

4.2 Methodology

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

$$\begin{aligned} r_{it} = \delta_j + \delta_j^{'}\nu_t + \beta_d d_{it} + \beta_1 \mathbb{1}_{\nu_t>0} + \beta_{\nu_+}\nu_t \mathbb{1}_{\nu_t>0} + \beta_{\nu_-} v_t \mathbb{1}_{\nu_t<0} \\ + \beta_{d\nu_+}(d_{it}\nu_t \mathbb{1}_{\nu_t>0}) + \beta_{d\nu_-}(d_{it}\nu_t \mathbb{1}_{\nu_t<0}) + \Gamma_1' X_t + \Gamma_2' X_t \nu_t + \varepsilon_{it}, \end{aligned} \tag{6}$$

where d_{it} is the attention prevalence, v_t is the monetary policy shock, $\mathbb{1}_{\nu>0}$ indicates positive monetary policy shocks, $\mathbb{1}_{\nu<0}$ indicates negative monetary policy shocks, and X_t is a vector of controls including the indicator variable for positive shocks and quarterly firm controls for size, age, and leverage. We also control for the interaction of monetary shocks with industry fixed effects and with firm controls, to capture the effects of firm characteristics on 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 $\beta_{d\nu_+}$ and $\beta_{d\nu_-}$. The theoretical framework in Section 3 hypothesizes $\beta_{d\nu_+}$ to be positive and $\beta_{d\nu_-}$ 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 $H_0 : \beta_{d\nu_+} = \beta_{d\nu_-}$.

4.3 Empirical 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 that a 25 basis point expansionary monetary shock is associated with about a 1% increase in stock prices.

This result 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 (6) are presented in the third column. We test whether attention leads to differential responses to positive and negative monetary 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, the fourth column ends the sample in 2007 to exclude the zero lower bound period following the Great Recession. This excludes periods of forward guidance and unconventional monetary policy and allows us to focus on conventional monetary transmission. Results are both qualitatively and quantitatively similar as in the full sample, suggesting our findings are not driven by anomalies from the financial crisis, the zero lower bound period.

The *asymmetric* responses to positive and negative shocks are consistent with heterogeneous responses predicted by a model of inattention and rule out alternative interpretations of the textual measure that predict symmetric responses. The first alternative interpretation, discussed in detail in Section 3, is that the textual measure misidentifies firms' profit exposure to macroeconomic conditions as attention. In this case, symmetric responses to positive and negative monetary shocks would yield a positive and significant effect from the interaction term between shock and attention (β_{dv}) in the second column, which is inconsistent with our empirical findings. A second alternative hypothesis is that firms attribute poor performance to broader economic forces and are more likely to mention FOMC meetings when underperforming. We would then expect attentive firms to underperform in response to negative monetary shocks, corresponding to a positive coefficient for β_{dv-} in the third

Table 3: Baseline results

		(1) Average	(2) Exposure	(3) Attention	(4) excl. ZLB
β_ν	Shock	5.61***	4.55*		
		(1.21)	(2.65)		
β_d	Attention		-0.01	-0.07	-0.03
			(0.05)	(0.06)	(0.06)
$\beta_{d\nu}$	Shock × Attn		1.07		
			(0.64)		
β_{ν_+}	Shock × $\mathbb{1}_{\nu_t>0}$			4.93*	6.54**
				(2.74)	(2.75)
β_{ν_-}	Shock × $\mathbb{1}_{\nu_t<0}$			-3.57	-0.95
				(3.72)	(3.69)
$\beta_{d\nu_+}$	Shock × Attn ×$\mathbb{1}_{\nu_t>0}$			2.02***	1.55**
				(0.72)	(0.72)
$\beta_{d\nu_-}$	Shock × Attn ×$\mathbb{1}_{\nu_t<0}$			-5.87*	-5.77*
				(3.18)	(3.30)
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				0.026	0.050

Notes: We have normalized the sign of the monetary shock ν_t so that a positive shock is expansionary (corresponding to a decrease in interest rates). Column (1) reports the average effect of monetary shocks from estimating $r_{it} = \delta_j + \beta_\nu \nu_t + \Gamma' X_t + \varepsilon_{it}$. Column (2) estimates the exposure model $r_{it} = \delta_j + \delta_j' v_t + \beta_\nu \nu_t + \beta_d d_{it} + \beta_{dv}(d_{it}\nu_{it}) + \Gamma_1' X_t + \Gamma_2' X_t v_t + \varepsilon_{it}$. Column (3) estimates the baseline attention model (6):

$$r_{it} = \delta_j + \delta_j' \nu_t + \beta_d d_{it} + \beta_1 \mathbb{1}_{\nu_t>0} + \beta_{\nu_+} \nu_t \mathbb{1}_{\nu_t>0} + \beta_{\nu_-} \nu_t \mathbb{1}_{\nu_t<0} \\ + \beta_{d\nu_+}(d_{it} v_t \mathbb{1}_{\nu_t>0}) + \beta_{d\nu_-}(d_{it}\nu_t \mathbb{1}_{\nu_t<0}) + \Gamma_1' X_t + \Gamma_2' X_t \nu_t + \varepsilon_{it},$$

where ν_t is the monetary shock, d_{it} is the prevalence attention measure, δ_j is an industry fixed effect and $\delta_j' \nu_t$ is its interaction with the shock, X_t contains firm-level controls of size, age, and leverage. The vector $X_t \nu_t$ contains the interactions between firm controls and the shock. Column (4) estimates Equation (6) for the sample up to 2007. Standard errors are clustered at the shock level and reported in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

column, which is also at odds with our empirical findings. Another concern is that investor attention is more important to stock price than firm attention. Inattentive investors would then systematically under-react to both positive and negative shocks, which fails to explain the observed asymmetry. A final concern is that firms may differ in price stickiness even within a narrow sector beyond controlled characteristics. For price stickiness to explain our empirical results, it must be correlated with macro keyword counts in SEC filings, which seems unlikely.

For additional robustness, the next subsection shows that our results are robust when controlling for firm management quality, past exposure to monetary policy, information effects of monetary policy, and macro variables.

The suboptimal responses to monetary policy by inattentive firms reported in Table 3, together with the large fraction of inattentive firms documented in Figure 3, provide some of the first direct evidence of the empirical consequences of firm inattention in the US. We estimate that inattentive firm returns rise by 2% less following positive shocks and drop by 6% more following negative shocks compared to those of their attentive peers. These differences are substantial given the average stock return response of 5%.

4.4 Additional empirical results

The Appendix contains two sets of additional empirical results.

Robustness checks Appendix B checks whether our baseline results in Section 4 are sensitive to potentially confounding factors. Tables A.4 and A.6 control for management quality[8] and exposure to monetary shocks, respectively, and show that our baseline asymmetric semi-elasticities are robust in each case.

Two concerns that have been raised about high frequency monetary shocks are that i) an "information effect" confounds the direct effects of a change to interest rates (Nakamura and Steinsson, 2018), and ii) monetary shocks may be correlated with business cycle fluctuations. Following Miranda-Agrippino and Ricco (2021), we control for each FOMC announcement's information effect using Greenbook forecast revisions between FOMC announcements and show that our main results are little changed in Table A.7. We then incorporate macro controls in Table A.8, including lagged unemployment, real output growth, and inflation. Again, our main results are robust for these controls.

[8]Data on management quality substantially restricts our sample, which is why we keep it as a robustness check rather than including it in our main results.

Limitations and promise of textual measures Recycled or *boilerplate* language is a key concern when using regulatory filings to measure firm attention. 10-K filings are often written collaboratively among managers and legal departments, and evidence suggests that firms include certain statements within 10-K filings to appease investors or lower liability (Cao et al., 2020). Moreover, firms likely save time and resources by revising their filing from the prior year rather than starting from scratch. Boilerplate language is a concerning source of measurement error when it includes keywords that identify firm attention. In Appendix C.3, we test whether boilerplate language contaminates our main results by measuring the diversity in filing language with a Jaccard score of lexical similarity and restricting our analysis to the most linguistically diverse 10-K sections. Results in Table A.9 are qualitatively and quantitatively similar to the baseline results.

Even greater measurement error may come from misidentifying attentive firms as inattentive (Type II error), which raises concerns about *underestimating* overall firm attention. False negatives may occur if our text analysis fails to identify discussion of economic topics due to our method's limited sophistication, or if attention is not uniformly publicized in 10-K filings across firms. For the purposes of this paper, underestimated attention would attenuate our results and imply that our current estimate for the cost of information frictions serves as a lower bound.

Text analysis methods also hold tremendous promise for uncovering a more refined depiction of firm attention and expectations formation. We illustrate these capabilities with two approaches for identifying the context in which firms discuss economic conditions. The first approach (Appendix C.2) uses a Latent Dirichlet Allocation (LDA) unsupervised model to categorize words that neighbor a given keyword. The second approach (Appendix C.1) uses the itemized structure of 10-K filings to identify which sections of the filing contain the most keywords.

5 Quantitative Model

Motivated by the empirical heterogeneity in firm attention, we now construct a general-equilibrium model with rationally-inattentive firms to understand the aggregate implications of heterogeneous firm attention. Key parameters of the model are calibrated using the attention measure and empirical moments from the sections above. Using the model, we quantify the state dependency of monetary policy as a result of attention and show the aggregate importance of information frictions.

5.1 Model environment

The model mechanism is an extension of 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 (2021). Attention is modeled with the Shannon mutual information following Sims (2003) and is an endogenous choice by the firm (Luo et al., 2017). 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,

$$\max_{C_{it},N_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t (\log C_t - \psi N_t), \tag{7}$$

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

$$C_t = \left(\int_0^1 C_{it}^{\frac{\varepsilon_p - 1}{\varepsilon_p}} dj \right)^{\frac{\varepsilon_p}{\varepsilon_p - 1}}, \tag{8}$$

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 receive a lump-sum transfer D_t from the government. The household budget constraint is given by:

$$\int_0^1 P_{it} C_{it} di + B_t \leq W_t N_t + (1 + \iota_t) B_{t-1} + D_t. \tag{9}$$

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:

$$\Delta \log Q_t = \rho \Delta \log Q_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2). \tag{10}$$

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:

$$\mathcal{I}(Q_{i,t|t-1}, Q_{i,t|t}) = \frac{1}{2} \log \frac{\sigma^2_{i,t|t-1}}{\sigma^2_{i,t|t}}, \tag{11}$$

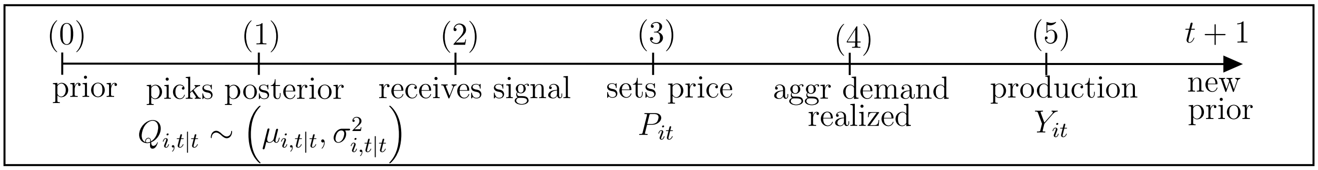
which captures the expected reduction in entropy from prior $Q_{i,t|t-1}$ to posterior $Q_{i,t|t}$. The Shannon mutual information is decreasing in the posterior variance so that more precise posteriors are more expensive. The marginal cost of information per nat, $2\omega_i$, is heterogeneous across firms and can be either high or low:

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

This heterogeneity is motivated by our empirical finding of polarized firm attention.

Figure 5 shows a firm's timeline. It enters each period with a prior belief about aggregate nominal demand. Then it chooses its posterior distribution. Since the Shannon mutual information in (11) does not depend on the posterior mean, it is optimal for a firm to center the posterior distribution around the true mean. Therefore, the firm's only information choice is the posterior variance $\sigma^2_{i,t|t}$. Based on the chosen posterior distribution, the firm receives a signal of aggregate demand and sets its price, P_{it}, based on its posterior belief. Finally, aggregate demand is realized, and the firm produces and enters the next period with a new

Figure 5: Firm's timeline



prior.

A firm's value function is given by

$$V(\sigma^2_{i,t|t-1}) = \max_{P_{it}, \sigma^2_{i,t|t}} \mathbb{E}_t \left[\underbrace{\frac{Y_{it}}{P_t} (P_{it} - MC_t)}_{\text{flow op. profits}} - \underbrace{2\omega_i \mathcal{I}(Q_{i,t|t-1}, Q_{i,t|t})}_{\text{info costs}} + \beta \underbrace{V(\sigma^2_{i,t+1|t})}_{\text{cont. value}} \middle| \, \sigma^2_{i,t|t} \right], \quad (12)$$

which consists of flow operational profits that are maximized when firms successfully track 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:

$$Y_{it} = (P_{it}/P_t)^{-\varepsilon_p} C_t \quad \text{(demand)}$$

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

$$0 \leq \sigma^2_{i,t|t} \leq \sigma^2_{i,t|t-1} \quad \text{(no forgetting)}$$

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 (7);

(ii) Given an initial prior $\sigma^2_{i,0|-1}$, prices, and the households' choices, firms optimize (12);

(iii) Monetary policy follows (10);

(iv) All markets clear.

Model solution Following Mackowiak and Wiederholt (2009) and Afrouzi and Yang (2021), we approximate a firm's flow profits with second order log approximations around the full-information steady state.[9] 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 (`DRIPs`) developed in Afrouzi and Yang (2021).

5.2 Calibration

Calibration features two sets of parameters: standard parameters unrelated to information frictions 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

[9] Details of the approximation can be found in Appendix E.1. 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.

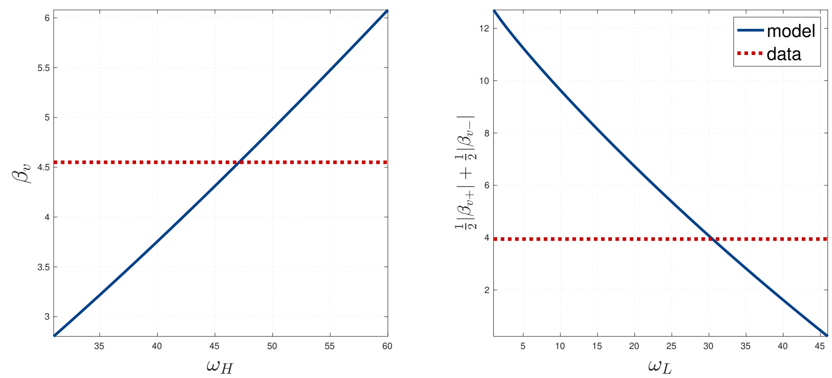
Information-friction parameters The bottom panel of Table 4 contains calibrations for parameters $(\theta, \omega_L, \omega_H)$. We use our text-based measure of attention to calibrate these important parameters governing the degree of information frictions in the model.

The fraction of attentive firms, θ, is set to equal 65% to match the average fraction of firms that have paid attention to the "general" topic over the sample period. Attention to the general economic conditions conveys firm attention to aggregate demand, which is a direct counterpart of the model state variable that firms track.

To calibrate the costs of attention, ω_L and ω_H, we target regression coefficients in Table 3. 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 (12), $r_{it} = \log V_{it} - \log \mathbb{E}_{t-1}(V_{it})$. 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.[10] Finally, we use ν_t as the monetary shocks.

[10] The empirical prevalence measure is binary: a value of 0 (below the cross-sectional mean) is interpreted as inattention, while a value of 1 (above the cross-sectional mean) is interpreted as attention. To map the empirical measure to the model counterpart, we make the assumption that the frequency of macro keywords in 10Ks is increasing in firm attention. This allows us to match the cross-sectional distribution of firm attention without explicitly modeling the writing process of 10Ks.

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 1,000 periods with 100 periods burn-ins. Simulated moments are generated with regressions discussed in the text.

We simulate the model for a panel of 100 firms and for 1,000 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) 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 measures 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 $\frac{1}{2}|\hat{\beta}_{dv+}| + \frac{1}{2}|\hat{\beta}_{dv-}|$, 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.5 in the appendix shows how simulated estimates of β_{dv+} and β_{dv-} change as a function of ω_L. β_{dv+} is positive and β_{dv-} is negative, suggesting that the stock returns of attentive firms outperform those of their inattentive peers for both positive and negative monetary shocks. As ω_L increases and the gap between the information costs for attentive and inattentive firms narrows, β_{dv+} decreases and β_{dv-} increases, implying a smaller difference in attention between attentive and inattentive firms.

The information cost parameters are calibrated to $\omega_L = 30$ and $\omega_H = 47$.[11] It may appear surprising that it is costly for firms to collect macro information considering macro data is freely available. However, as plant-level evidence in 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) estimates the data production function, which takes labor and capital inputs to process unstructured data into structured data and analyze data to produce knowledge. It requires significant effort 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 represent the costs of both acquiring and processing information.

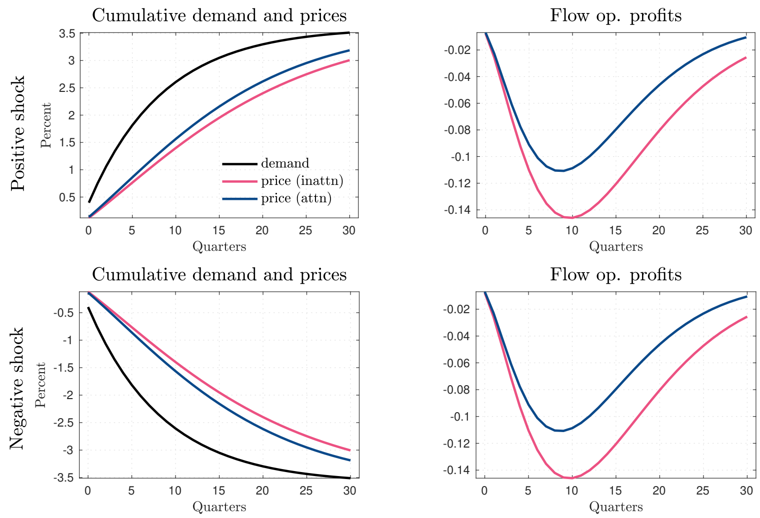
5.3 Model dynamics

We now study how firm inattention results in monetary non-neutrality. Figure 7 shows the impulse responses of individual firms to expansionary and contractionary monetary shocks of one standard deviation. Inattentive firms are shown in red, and attentive firms are shown

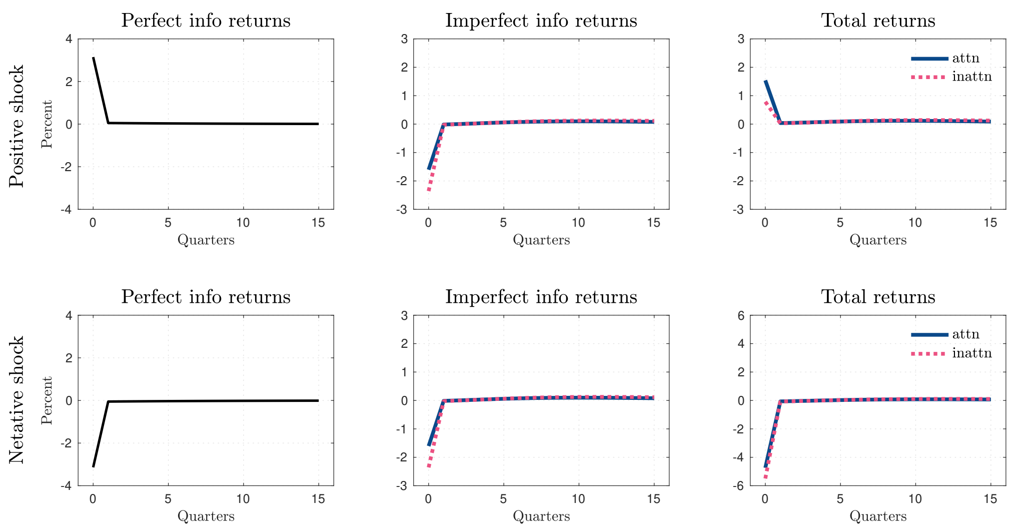
[11] The only preceding calibration for firm cost of attention is Afrouzi (2020), which studies the rational inattention problem of New Zealand firms under strategic complementarity and calibrates $\omega = 0.3$ using firm beliefs reported in New Zealand surveys. Our calibration differs by the sample of US firms and the abstraction from strategic complementarity. Flynn and Sastry (2021) builds upon our calibration strategy of matching conditional regression moments to study countercyclical attention.

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, and "inattn" refers to the impulse responses of inattentive firms.

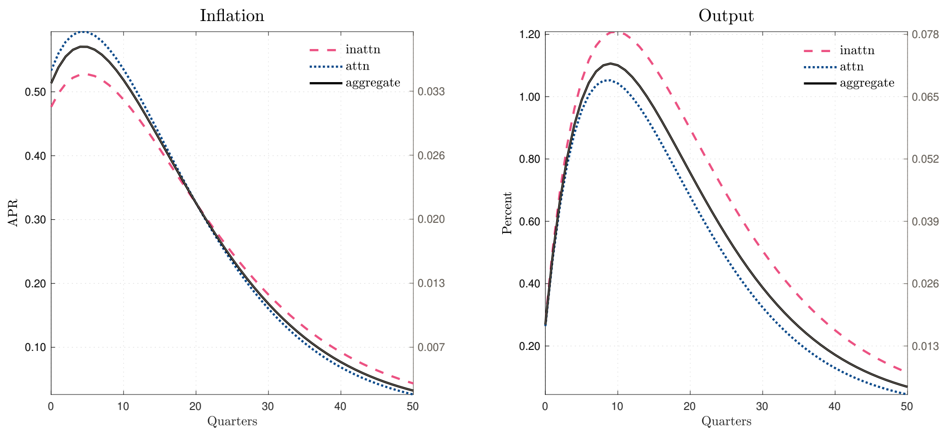
in blue. Panel (a) shows the responses of firms' prices and flow operating profits. As nominal aggregate demand rises, firms' prices respond sluggishly, reflecting partial incorporation of noisy signals about demand. Attentive firms track aggregate demand better than inattentive firms and exhibit more responsive prices. Since we approximate firm profits around the full-information steady state, any deviation from the full-information benchmark results in a loss. Inattentive firms experience greater operational losses 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. With a constant marginal cost of information, firms' equilibrium choice of attention is not time-varying and therefore does not result in a change in returns.

Panel (b) shows the responses of stock returns. Following an expansionary monetary shock, full-information equity returns of both attentive and inattentive firms increase, since firms are monopolistically competitive. Returns of attentive firms increase by more than those of inattentive firms because attentive firms track the optimal price more closely. Returns of both imperfect-information firms are lower than those of a full-information firm that sets the optimal price. Following a contractionary shock, returns of attentive firms drop by less than those of inattentive firms.

In Figure 8, we study the aggregate responses of output and inflation by aggregating attentive and inattentive firms. In response to a one standard deviation expansionary monetary shock to the nominal aggregate demand (Q), annualized inflation rises by 0.55% and output rises by 1.1% at their respective peak. Since inattentive firms mischaracterize the nominal monetary shock as a real shock, the aggregate output response to monetary policy is stronger in an economy with a larger fraction of inattentive firms.

To compare the magnitude of aggregate impulse responses to standard benchmarks such as Christiano et al. (2005), we convert the nominal aggregate demand shock in our inattention model to a nominal interest rate shock used in Christiano et al. (2005). We do so by estimating the passthrough of interest rate on the nominal aggregate demand in Appendix E.3. The right scales in Figure 8 show the impulse responses to a monetary shock equivalent to a 25 basis point interest rate cut. Annualized inflation and output increase by 0.04% and 0.07% at their peak, respectively. As a benchmark, Christiano et al. (2005) estimates the annualized peak

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.

effect of monetary policy shocks as 0.2% for inflation and 0.5% for output. With information as the only source of friction, our model generates about one seventh of their output response.

5.4 Inattention and the efficacy of monetary policy

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

The countercyclicality of aggregate attention suggests an important insight about the efficacy of monetary policy: When the Federal Reserve cuts rates during an ongoing recession, monetary policy is less powerful because firms are likely paying more attention to central bank decision-making. With a higher fraction of attentive firms, information frictions are less severe, monetary policy is closer to neutral, and monetary stimulus has a smaller effect on output. In contrast, preemptive monetary policy measures aimed at averting a potential recession are more powerful because a smaller fraction of firms are likely to respond

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.

attentively.

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

$$d_t = \tau_t + \zeta_t + \xi_t,$$

where τ_t, ζ_t, and ξ_t denote the trend, cyclical, and error components of the attention measure d_t, respectively. The series frequency is annual and the 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 maximum of the HP-filtered prevalence measure in the left panel of Figure 2. 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. The average output response to monetary policy is 5% weaker in the most attentive calibration compared to the least attentive calibration. This suggests that 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. This result is consistent with studies on the state dependency of monetary policy that find US monetary policy to be weaker in recessions than in expansions (Tenreyro and Thwaites, 2016).

6 Conclusion

The empirical evidence of information frictions that we document in this paper, along with growing evidence in the literature (Candia et al., 2021), highlights firms' deviations from full-information rational expectations (FIRE). To discipline models without FIRE, researchers require an understanding of firms' information sets and expectations formation processes.

In that direction, this paper presents a new text-based measure of firm attention to macroeconomic news, which will be made available publicly and updated on an ongoing basis. We validate that the measure indeed captures firm attention by testing for an asymmetric prediction of rational inattention on monetary policy transmission. We show that firms that pay attention to the FOMC have larger increases in stock returns after positive monetary shocks and smaller decreases in stock returns after negative monetary shocks, providing direct empirical evidence for the consequences of firm inattention.

The empirical measure can be used in combination with imperfect-information models to ground those theories with data. We demonstrate the value of this measure in a quantitative rational inattention model by showing that time variation in firm attention has important implications for the state dependency of monetary policy. In the model, average inattention drives the degree of monetary non-neutrality. The countercyclical nature of firm attention to macroeconomic news implies that the efficacy of monetary policy is weaker during recessions and should be considered in policy design.

References

Abis, Simona and Laura Veldkamp, "The Changing Economics of Knowledge Production," *Unpublished Manuscript*, 2020.

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

_ **and Choongryul Yang**, "Dynamic Rational Inattention and the Phillips Curve," *Unpublished Manuscript*, 2021.

Angeletos, George-Marios and Jennifer La'O, "Sentiments," *Econometrica*, 2013, *81* (2), 739–779.

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.

Bartoš, Vojtěch, Michal Bauer, Julie Chytilová, and Filip Matějka, "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition," *American Economic Review*, 2016, *106* (6), 1437–75.

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.

Candia, Bernardo, Olivier Coibion, and Yuriy Gorodnichenko, "The Inflation Expectations of US Firms: Evidence from a new survey," Technical Report, National Bureau of Economic Research 2021.

Cao, Sean, Wei Jiang, Baozhong Yang, and Alan L Zhang, "How to Talk When a Machine is Listening: Corporate Disclosure in the Age of AI," Technical Report, National Bureau of Economic Research 2020.

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.

Chiang, Yu-Ting, "Strategic Uncertainty over Business Cycle," *Unpublished Manuscript*, 2021.

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.

Cloyne, James, Clodomiro Ferreira, Maren Froemel, and Paolo Surico, "Monetary Policy, Corporate Finance and Investment," Technical Report, National Bureau of Economic Research 2018.

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.

Farhi, Emmanuel and Iván Werning, "Monetary policy, bounded rationality, and incomplete markets," *American Economic Review*, 2019, *109* (11), 3887–3928.

Flynn, Joel P and Karthik Sastry, "Attention cycles," *Unpublished Manuscript*, 2021.

Fuster, Andreas, Ricardo Perez-Truglia, Mirko Wiederholt, and Basit Zafar, "Expectations with endogenous information acquisition: An experimental investigation," *The Review of Economics and Statistics*, 2018, pp. 1–54.

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.

_ **and Simon Gilchrist**, "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms," *The Quarterly Journal of Economics*, 1994, *109* (2), 309–340.

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 A, Stephan Hollander, Laurence Van Lent, and Ahmed Tahoun, "Firm-level Political Risk: Measurement and Effects," *The Quarterly Journal of Economics*, 2019, *134* (4), 2135–2202.

Jaimovich, Nir and Sergio Rebelo, "Can News about the Future Drive the Business Cycle?," *American Economic Review*, 2009, *99* (4), 1097–1118.

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

Larsen, Vegard H, Leif Anders Thorsrud, and Julia Zhulanova, "News-Driven Inflation Expectations and Information Rigidities," *Journal of Monetary Economics*, 2021, *117*, 507–520.

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.

Lucas, Robert, "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 1972, *4* (2), 103–124.

Luo, Yulei, Jun Nie, Gaowang Wang, and Eric R Young, "Rational Inattention and the Dynamics of Consumption and Wealth in General Equilibrium," *Journal of Economic Theory*, 2017, *172*, 55–87.

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

Maćkowiak, Bartosz and Mirko Wiederholt, "Business Cycle Dynamics under Rational Inattention," *The Review of Economic Studies*, 2015, *82* (4), 1502–1532.

Mackowiak, Bartosz, Filip Matejka, and Mirko Wiederholt, "Rational inattention: A review," 2021.

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.

Miranda-Agrippino, Silvia and Giovanni Ricco, "The Transmission of Monetary Policy Shocks," *American Economic Journal: Macroeconomics*, 2021, *13* (3), 74–107.

Nakamura, Emi and Jón Steinsson, "High Frequency Identification of Monetary Non-neutrality: 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," *Econometrica*, 2020, *88* (6), 2473–2502.

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.

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.

Online Appendix

A 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 `Econoday`, complemented with macroeconomic words and phrases from popular press.

Table A.2: Summary statistics of firm characteristics by 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: This table reports 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.

B Additional Robustness

This appendix investigates additional firm characteristics that may drive attention to monetary policy. We then show that our baseline results in Table 3 are robust when controlling for these additional factors. We also show that our results are robust when controlling for potentially confounding factors of monetary shocks by controlling for business-cycle variables and the information effect of monetary shocks.

B.1 Results robust when controlling for management quality

Management quality is a part of a firm's infrastructure that determines its information-processing capacity. We obtain data on publicly-traded firms' board members and their education levels from `BoardEx`. Management quality, m_{it}, is measured as the fraction of firm i's board members in year t who have a master's degree or above.[12]

Table A.3 first shows that indeed, firms whose board members have attained higher levels of education are more attentive to monetary policy than their peers. Column (2) shows that over the lifecycle of a firm, it is more likely to be attentive when it has a highly-educated board.

Table A.3: Attention and firm management

	(1)	(2)
Management	0.0110***	0.0585***
	(0.0039)	(0.0054)
Observations	65392	65393
R^2	0.422	0.756
Time-industry FE	yes	no
Firm FE	no	yes

Notes: Column (1) reports the estimated coefficient β from $d_{it} = \delta_t + \delta_j + \beta \cdot m_{it} + \varepsilon_{it}$, and Column (2) reports the estimated coefficient β from $d_{it} = \delta_i + \beta \cdot m_{it} + \varepsilon_{it}$, described as in the main text. d_{it} is the prevalence attention to FOMC news, m_{it} is the fraction of board members who have a master's degree or above, δ_i is a firm fixed effect, δ_j is an industry fixed effect (4-digit NAICS), and δ_t is a time fixed effect.

[12] Degrees counted as master-level or above include: MBA, MS, MSC, MA, JD, MD, MPA, MSE, PHD, and degree names that include "master" or "doctor."

Table A.4 then shows that our baseline results are robust when controlling for management quality. Since good management can capitalize on expansionary shocks and mitigate contractionary shocks, we allow the controls for management to interact with the monetary shocks asymmetrically. The specification estimated in Table A.4 is:

$$
\begin{aligned}
r_{it} = \delta_j + \delta_j' v_t + \beta_{v+} v_t \mathbb{1}_{v_t>0} + \beta_{v-} v_t \mathbb{1}_{v_t<0} + \beta_d d_{it} + \beta_m m_{it} \\
+ \beta_{dv+} d_{it} v_t \mathbb{1}_{v_t>0} + \beta_{dv-} d_{it} v_t \mathbb{1}_{v_t<0} + \beta_{mv+} m_{it} v_t \mathbb{1}_{v_t>0} + \beta_{mv-} m_{it} v_t \mathbb{1}_{v_t<0} \\
+ \Gamma_1' X_t + \Gamma_2' X_t v_t + \varepsilon_{it},
\end{aligned} \tag{13}
$$

where m_{it} denotes management quality and d_{it} denotes our baseline prevalence attention measure. Column (1) of Table A.4 reports the results from our baseline specification using only the sample that overlaps with `BoardEx` data. Column (2) reports no significant effects of management quality on responses to monetary policy. Column (3) reports the effects of attention controlling for management quality. The Wald test for the null hypothesis that β_{dv+} and β_{dv-} are equal is rejected at 1%, suggesting that the finding of asymmetric responses to monetary policy by attention is robust when controlling for management quality.

Table A.4: Controlling for management quality

	(1)	(2)	(3)
Attention	-0.10		-0.10*
	(0.06)		(0.06)
Shock × Attn $\times \mathbb{1}_{v_t>0}$	2.48***		2.59***
	(0.86)		(0.91)
Shock × Attn $\times \mathbb{1}_{v_t<0}$	-7.53**		-8.18**
	(3.52)		(3.34)
Management		-0.04	-0.07
		(0.06)	(0.06)
Shock × Mgmt $\times \mathbb{1}_{v_t>0}$		1.24	1.53**
		(0.76)	(0.73)
Shock × Mgmt $\times \mathbb{1}_{v_t<0}$		-0.68	-2.54
		(2.88)	(2.58)
Observations	324154	324154	324154
R^2	0.038	0.038	0.038
Clustered SE	yes	yes	yes
Firm controls	yes	yes	yes
4-digit NAICS FE	yes	yes	yes
Wald Test p-value	0.008	0.528	0.003

Notes: Results from estimating the specification in (13), with variables as defined in the main text: Column (1) shows our baseline results, using only sample that overlaps with BoardEx data. Column (2) shows the effects of management quality. Column(3) shows the effects of attention controlling for management quality. Standard firm controls include age, size, and leverage. Standard errors are clustered at the shock level and reported in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

B.2 Results robust when controlling for monetary exposure

The theoretical prediction of asymmetry from Section 3 confirms the baseline effects in Table 3 are driven by firm attention rather than firm exposure to monetary policy. Nevertheless, we conduct additional robustness in this section to directly control for firms' exposure to monetary policy.

To measure a firm's exposure to the monetary policy at date τ, we first estimate the sensitivity of its stock prices to prior FOMC announcements over a 5-year rolling window using $t \in [\tau - 5\text{yr}, \tau)$:

$$
\begin{aligned}
&\text{Baseline model:} \quad & r_{it} &= \alpha_{i\tau} + \beta_{i\tau}^{\texttt{baseline}} v_t + \varepsilon_{it} \\
&\text{CAPM model:} \quad & r_{it} - r_t^f &= \alpha_{i\tau} + \beta_{i\tau}^{\texttt{capm}} v_t + \beta_{i\tau}^{M}(r_t^M - r_t^f) + \varepsilon_{it} \\
&\text{FF3 model:} \quad & r_{it} - r_t^f &= \alpha_{i\tau} + \beta_{i\tau}^{\texttt{ff3}} v_t + \beta_{i\tau}^{1}(r_t^M - r_t^f) + \beta_{i\tau}^{2} SMB_t + \beta_{i\tau}^{3} HML_t + \varepsilon_{it},
\end{aligned}
$$

where v_t is the high-frequency monetary shock, and r_{it} is the close-to-close returns of firm i at date t. In addition to the baseline model, we also estimate a stock's sensitivity when controlling for the market factor (r^M) and Fama-French 3 factors (r^M, SML, and HML), to isolate the sensitivity to monetary policy. We obtain the daily data on factors from Kenneth French's website.

Based on the estimated sensitivity, we then measure a firm's exposure to monetary policy as the absolute values of the beta's

$$
\theta_{i\tau}^{\lambda} = |\beta_{i\tau}^{\lambda}| \quad \text{for } \lambda \in \{\text{baseline, CAPM, FF3}\}
$$

Table A.5 shows that firm attention increases with the exposure to monetary policy, both in the cross section and over the time series. The relationship is robust for the measures of monetary exposure. In Appendix B, we incorporate additional controls for monetary exposure in the baseline specification to show that even though exposure drives a firm's attention, the baseline results of differential monetary transmission by attention are not driven by a firm's exposure to monetary policy.

Table A.5: Attention and exposure to monetary policy

Panel A: Time-industry level

	(1)	(2)	(3)
Exposure (baseline model)	0.0016***		
	(0.0006)		
Exposure (CAPM model)		0.0036***	
		(0.0006)	
Exposure (FF3 model)			0.0036***
			(0.0008)
Observations	74272	73649	72509
R^2	0.034	0.035	0.035
Time-industry FE	yes	yes	yes
Firm FE	no	no	no
Firm controls	yes	yes	yes

Panel B: Firm level

	(1)	(2)	(3)
Exposure(baseline model)	0.0016***		
	(0.0005)		
Exposure (CAPM model)		0.0012**	
		(0.0006)	
Exposure (FF3 model)			0.0019***
			(0.0006)
Observations	74280	73657	72520
R^2	0.567	0.567	0.560
Time-industry FE	no	no	no
Firm FE	yes	yes	yes
Firm controls	yes	yes	yes

Notes: Panel A reports the estimated coefficient β from $d_{it} = \delta_t + \delta_j + \beta \cdot \theta_{it}^{\lambda} + \varepsilon_{it}$, and Panel B reports the estimated coefficient β from $d_{it} = \delta_i + \beta \cdot \theta_{it}^{\lambda} + \varepsilon_{it}$, with θ_{it}^{λ} denoting the exposure to monetary policy with $\lambda \in \{\text{baseline, CAPM, FF3}\}$ and constructed as described in the main text. d_{it} is the prevalence attention to FOMC news, δ_i is a firm fixed effect, δ_j is a sector fixed effect (4-digit NAICS), and δ_t is a time fixed effect. Firm controls include age, size and leverage.

Table A.6 then controls for each measure of monetary exposure, θ_{it}^{λ}, for $\lambda \in$ {baseline, CAPM, FF3}. For all three measures, the Wald tests for the null hypothesis that $\beta_{dv+} = \beta_{dv-}$ are rejected at 5%, showing that our results are not driven by firms' exposure to monetary policy.

Table A.6: Controlling for exposure to monetary policy

	(1)	(2)	(3)
Shock × Attn $\times \mathbb{1}_{v_t>0}$	2.03***	2.03***	2.03***
	(0.73)	(0.72)	(0.72)
Shock × Attn $\times \mathbb{1}_{v_t<0}$	-5.99*	-5.99*	-5.94*
	(3.25)	(3.25)	(3.24)
Observations	572884	571708	568169
R^2	0.026	0.026	0.026
Clustered SE	yes	yes	yes
Firm controls	yes	yes	yes
4-digit NAICS FE	yes	yes	yes
Monetary sensitivity control	baseline model	CAPM model	FF3 model
Wald Test p-value	0.027	0.027	0.027

Notes: Results from estimating the baseline specification (6) with additional controls for monetary exposure, θ_{it}^{λ}, $\lambda \in$ {baseline, CAPM, FF3}, defined in Appendix B.2. Standard firm controls include age, size, and leverage. Standard errors are clustered at the shock level and reported in parentheses.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

B.3 Results not driven by information effect of monetary policy

Nakamura and Steinsson (2018) documents that FOMC announcements release information about the economic fundamentals, in addition to monetary policy. Following Miranda-Agrippino and Ricco (2021), we control for the information effects of monetary policy by including as controls the Greenbook forecast revisions between FOMC meetings. We obtain data on Greenbook forecasts from the Federal Reserve Bank of Philadelphia. Table A.7 show that our main results are robust when controlling for Greenbook forecast revisions.

Table A.7: Controlling for Greenbook forecast revisions

	(1)	(2)	(3)	(4)
Shock $\times$ Attn $\times \mathbb{1}_{v_t>0}$	2.02***	1.88**	1.94***	1.94***
	(0.72)	(0.75)	(0.72)	(0.72)
Shock $\times$ Attn $\times \mathbb{1}_{v_t<0}$	-5.87*	-5.47	-5.71	-5.71
	(3.18)	(3.58)	(3.68)	(3.68)
Observations	575667	575667	575667	575667
R^2	0.026	0.026	0.026	0.026
Clustered SE	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes
4-digit NAICS FE	yes	yes	yes	yes
Greenbook rev controls		rgdp	rgdp infl	rgdp infl unemp
Wald Test p-value	0.026	0.070	0.063	0.063

Notes: Results from estimating the baseline specification (6) with additional controls for Greenbook forecast revisions. Column (1) displays the baseline results from Table 3. Columns (2) - (4) add Greenbook forecast revisions for real GDP, inflation, and unemployment iteratively. Standard firm controls include age, size, and leverage. Standard errors are clustered at the shock level and reported in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

B.4 Results robust when controlling for macro fluctuations

While the high-frequency monetary shocks, v_t, are considered exogenous, we conduct additional tests for robustness, controlling for business-cycle fluctuations. Macro controls include: lagged real GDP growth, unemployment rate, and inflation, obtained from `FRED`. Column (1) of Table A.8 displays our baseline results without macro controls. Column (2) includes macro controls, controlling for aggregate fluctuations. Column (3) includes macro controls and their interactions with the monetary shock, controlling for differential firm sensitivity to aggregate fluctuations. Column (4) includes macro controls and their separate interactions with expansionary and contractionary monetary shocks, controlling for asymmetric firm sensitivity to aggregate fluctuations. Our main results are robust under all specifications.

Table A.8: Controlling for macroeconomic variables

	(1)	(2)	(3)	(4)
Shock × Attn $\times \mathbb{1}_{v_t>0}$	2.02***	2.06***	1.74**	1.74**
	(0.72)	(0.73)	(0.78)	(0.71)
Shock × Attn $\times \mathbb{1}_{v_t<0}$	-5.87*	-6.27*	-5.38	-7.31**
	(3.18)	(3.21)	(3.34)	(3.31)
Observations	575667	575667	575667	575667
R^2	0.026	0.028	0.028	0.028
Clustered SE	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes
4-digit NAICS FE	yes	yes	yes	yes
Macro controls	no	yes	yes	yes
+ interactions	no	no	yes	no
+ asym interactions	no	no	no	yes
Wald Test p-value	0.026	0.021	0.060	0.014

Notes: Results from estimating the baseline specification (6) with an additional vector of macro control Z_{t-1}, where Z_{t-1} includes lagged real GDP growth, unemployment rate, and inflation. Column (1) displays the baseline results from Table 3. Column (2) includes macro controls Z_{t-1},. Column (3) includes Z_{t-1} and $Z_{t-1}v_t$. Column (4) includes Z_{t-1} and $Z_{t-1}v_t\mathbb{1}_{v>0}$, and $Z_{t-1}v_t\mathbb{1}_{v<0}$. Standard firm controls include age, size, and leverage. Standard errors are clustered at the shock level and reported in parentheses.
* ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

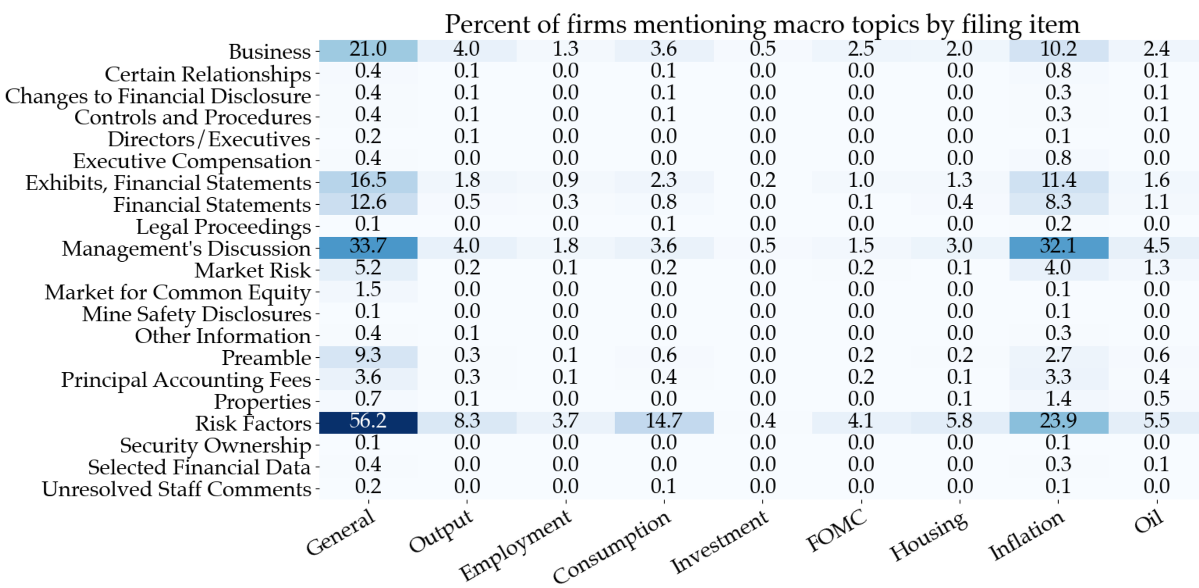
C Additional Results from Textual Analysis

This appendix contains a set of additional results using natural language processing to investigate the context in which firms discuss macro keywords in 10-K filings and to provide further validation of the text-based measures.

C.1 Itemized frequency search

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

Figure A.1: 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.

Results in Figure A.1 show that firms pay attention to macro news to assess the impact on their business operations and risks, consistent with assumptions that firms mentioning a macroeconomic topic do so in order to incorporate the news into their decision-making.

C.2 LDA: Context of macro discussions

To enable automated context detection, we use the Latent Dirichlet Allocation (LDA) model to uncover topics that 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 includes words and bigrams that appear at least 20 times. We filter out words that occur in fewer 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.2 and A.3 visualize the LDA output surrounding keywords in all categories. Figure A.2 shows the heat map of LDA outputs. Each row represents a topic clustered by LDA, and the darkness of the cell within a topic represents the likelihood of a word appearing in the topic. Figure A.3 highlights the word cloud of selected topics in A.2.

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 their daily operations; Topic 2 relates to demand, as firms track and gauge the aggregate demand; Topic 6 relates to financing costs, as firms pay attention to how monetary policy affects their financial costs, investment decisions, and portfolio holdings; Topic 10 relates to labor costs, as firms assess the tightness of the labor market. The remaining topics relate to housing, currency, and risk factors.

Figure A.2: LDA output for texts surrounding all macro keywords

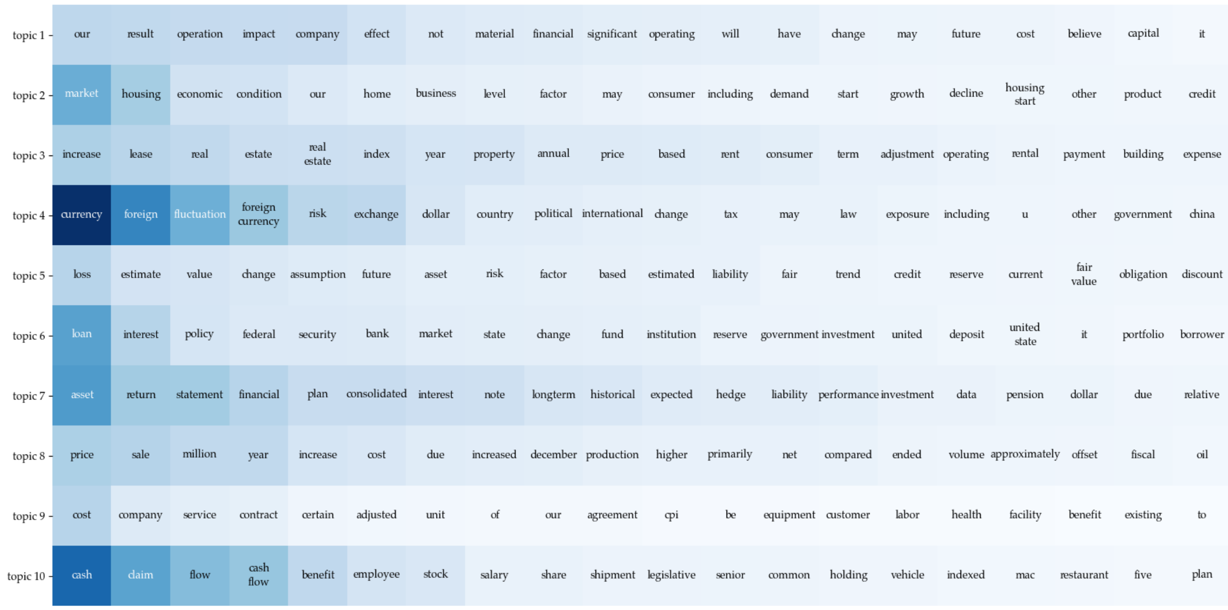
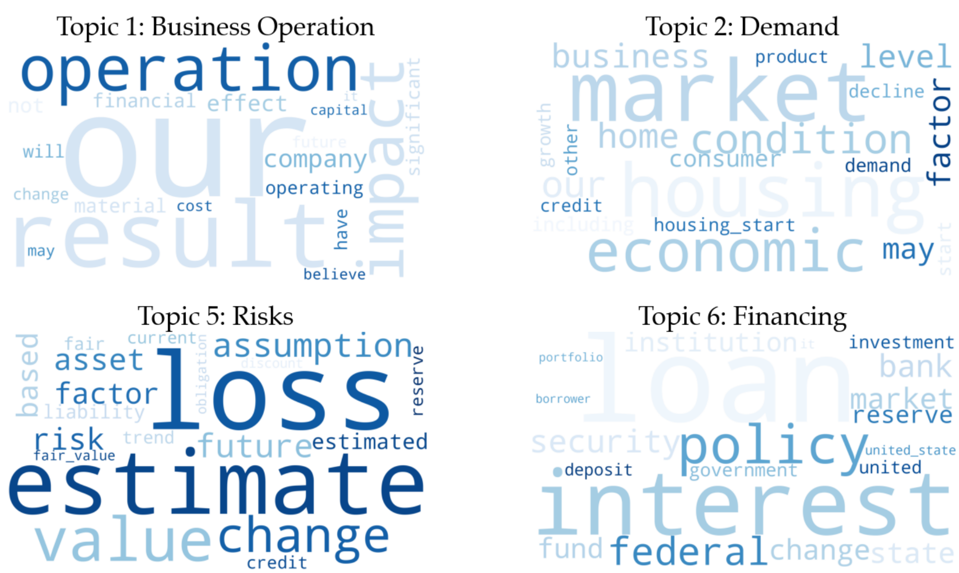


Figure A.3: LDA output for texts surrounding all macro keywords: Selected topics



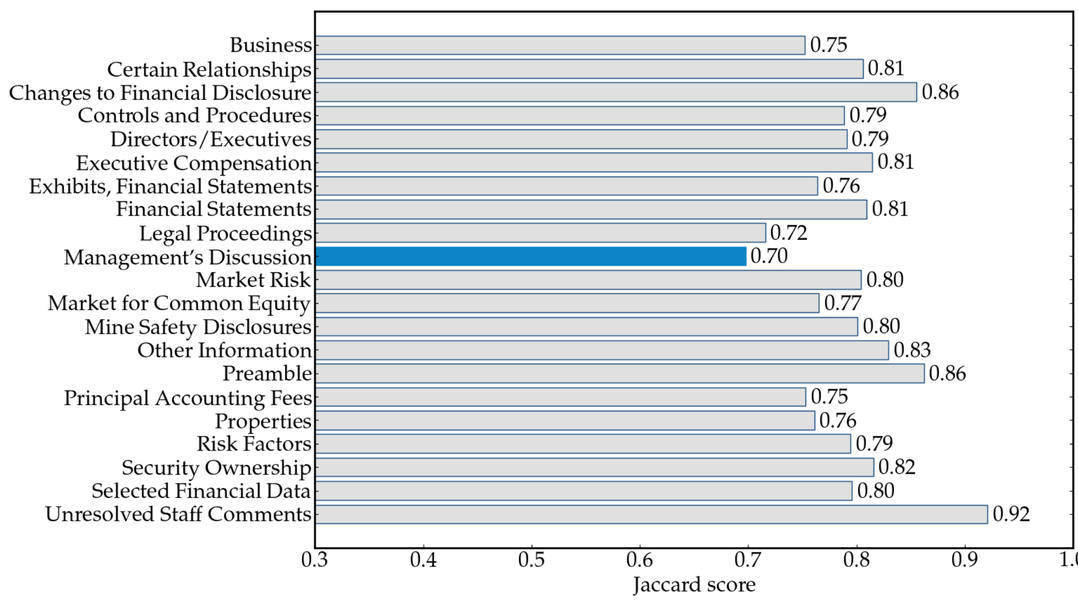
C.3 Lexical similarity

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.4 reports the average Jaccard score for each section of 10-K filings.

Figure A.4: 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.9.

Table A.9: Restricting attention to low lexical similarity 10-K sections

	(1) Average	(2) Exposure	(3) Attention	(4) excl ZLB
Shock	5.62***	4.13*		
	(1.22)	(2.42)		
Attention		-0.03	-0.08	-0.05
		(0.04)	(0.05)	(0.05)
Shock × Attn		0.02		
		(0.45)		
Shock × $\mathbb{1}_{v_t>0}$			4.55*	6.21**
			(2.65)	(2.66)
Shock × $\mathbb{1}_{v_t<0}$			-4.16	-1.45
			(3.72)	(3.69)
Shock × Attn × $\mathbb{1}_{v_t>0}$			0.79	0.50
			(0.56)	(0.54)
Shock × Attn × $\mathbb{1}_{v_t<0}$			-5.24**	-4.95*
			(2.48)	(2.53)
Observations	546596	546596	546596	409889
R^2	0.018	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			0.030	0.058

Notes: Results from variants of estimating the baseline specification in (6), restricting to 10-K items that discuss firm operations (Items 1 and 7). Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

D Additional Details for the Stylized Model

D.1 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 adds and subtracts $\frac{1}{2}\pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\hat{s}_t^2$ to complete squares and uses the fact that $\bar{a} = \bar{s}$ in the steady state. The resulting expression is equation (3).

D.2 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 (3), 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 firm i is larger than 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:

$$\frac{\partial r_i}{\partial \varepsilon} = \pi_s(\bar{s}, \bar{a})\bar{s} + \left(\pi_{ss}(\bar{s}, \bar{a})\bar{s}^2 - \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\right)\varepsilon \tag{14}$$

$$\frac{\partial r_j}{\partial \varepsilon} = \pi_s(\bar{s}, \bar{a})\bar{s} + \left(\pi_{ss}(\bar{s}, \bar{a})\bar{s}^2 - \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\right)\varepsilon + \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2(f_j(\varepsilon) - \varepsilon)(f_j'(\varepsilon) - 1), \tag{15}$$

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 $f(0) = 0$, (14) and (15) 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 $\varepsilon_t < f_j(\varepsilon_t) < 0$,

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

■

E Additional Details for the Quantitative Model

E.1 Approximation of firms' value function

A firms' value function for its operating profits can be expressed as

$$\begin{aligned}
V^{op} &= \max_{\{P_{it},\sigma^2_{i,t|t},\sigma^2_{i,t+1|t}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t\, \mathbb{E}\left[\Pi(P_{it},P_t,Q_t)|\sigma^2_{i,0|-1}\right] \\
&= \max_{\{P_{it},\sigma^2_{i,t|t},\sigma^2_{i,t+1|t}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t\, \mathbb{E}\left[\frac{\Pi(P_{it},P_t,Q_t)}{\Pi^*(P^*_{it},P_t,Q_t)}\Pi^*(P^*_{it},P_t,Q_t)|\sigma^2_{i,0|-1}\right] \\
&= \max_{\{P_{it},\sigma^2_{i,t|t},\sigma^2_{i,t+1|t}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \Pi^*(P^*_{it},P_t,Q_t)\, \mathbb{E}\left[L(P_{it},P_t,Q_t)|\sigma^2_{i,0|-1},\right]
\end{aligned}$$

where $\Pi(P_{it},P_t,Q_t)$ denotes the firm's operating profits, and $L(P_{it},P_t,Q_t) \equiv \frac{\Pi(P_{it},P_t,Q_t)}{\Pi^*(P^*_{it},P_t,Q_t)}$ denotes the loss from imperfect information relative to full-information profits $\Pi^*(P^*_{it},P_t,Q_t)$. The last equality follows the fact that L is homogeneous of degree 1.

Under the second-order log approximation around the non-stochastic steady state, we can express the loss as:

$$\begin{aligned}
&\frac{\Pi(P_{it},P_t,Q_t)}{\Pi^*(P^*_{it},P_t,Q_t)} \approx \frac{\Pi(\bar{P},\bar{P},\bar{Q})}{\Pi(\bar{P},\bar{P},\bar{Q})} + (p_{it}-p^*_{it})\bar{P}\frac{\Pi_1(\bar{P},\bar{P},\bar{Q})}{\Pi(\bar{P},\bar{P},\bar{Q})} + (p^2_{it}-p^{*2}_{it})\frac{\bar{P}}{2}\frac{\Pi_{11}(\bar{P},\bar{P},\bar{Q})}{\Pi(\bar{P},\bar{P},\bar{Q})} \\
&\quad - p_{it}p^*_{it}\bar{P}^2\frac{\Pi_1(\bar{P},\bar{P},\bar{Q})^2}{\Pi(\bar{P},\bar{P},\bar{Q})^2} + (p_{it}-p^*_{it})p_t\bar{P}^2\frac{\Pi_{12}(\bar{P},\bar{P},\bar{Q})}{\Pi(\bar{P},\bar{P},\bar{Q})} + (p_{it}-p^*_{it})q_t\bar{P}\bar{Q}\frac{\Pi_{13}(\bar{P},\bar{P},\bar{Q})}{\Pi(\bar{P},\bar{P},\bar{Q})} \\
&= 1 + (p^2_{it}-p^{*2}_{it})\frac{\bar{P}}{2}\frac{\Pi_{11}(\bar{P},\bar{P},\bar{Q})}{\Pi(\bar{P},\bar{P},\bar{Q})} + (p_{it}-p^*_{it})p_t\bar{P}^2\frac{\Pi_{12}(\bar{P},\bar{P},\bar{Q})}{\Pi(\bar{P},\bar{P},\bar{Q})} + (p_{it}-p^*_{it})q_t\bar{P}\bar{Q}\frac{\Pi_{13}(\bar{P},\bar{P},\bar{Q})}{\Pi(\bar{P},\bar{P},\bar{Q})} \\
&= 1 + (p_{it}-p^*_{it})^2\frac{\bar{P}}{2}\frac{\Pi_{11}(\bar{P},\bar{P},\bar{Q})}{\Pi(\bar{P},\bar{P},\bar{Q}),}
\end{aligned}$$

where lowercase letters denote log deviations from the steady state. The second equality uses the fact that $\Pi_1 = 0$ from optimal choices. In addition, $\Pi_1(P^*_{it},P_t,Q_t) = 0$ implies $p^*_{it}\bar{P}\Pi_{11}(\bar{P},\bar{P},\bar{Q})+p_t\bar{P}\Pi_{12}(\bar{P},\bar{P},\bar{Q})+q_{it}\bar{Q}\Pi_{13}(\bar{P},\bar{P},\bar{Q}) = 0$, which leads to the third equality.

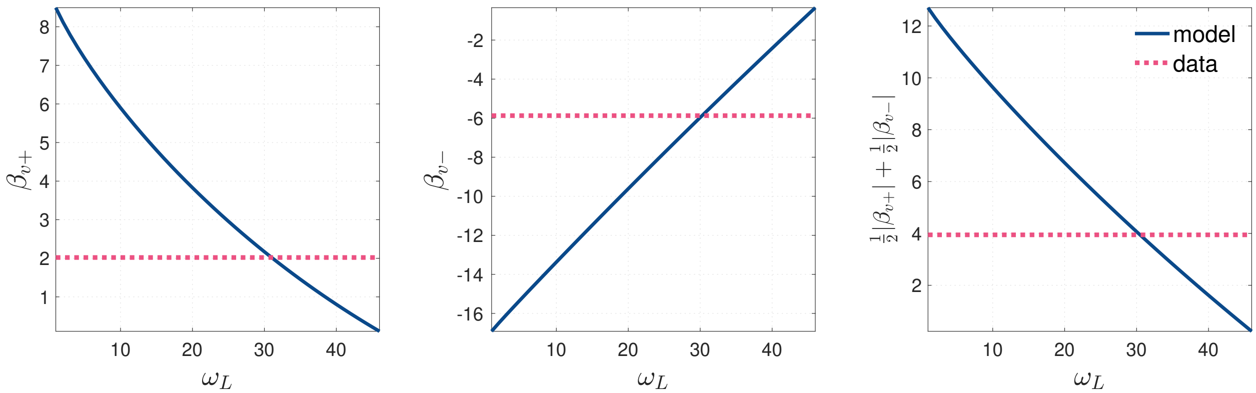
A firm's log operating value, v^{op}, can be decomposed into:

$$v^{op} = v^* + l,$$

consisting of v^*, the full-information value, and l, the loss in firm value from imperfect information approximated as above.

E.2 Details for model calibration

Figure A.5: Sensitivity of simulated moments to ω_L



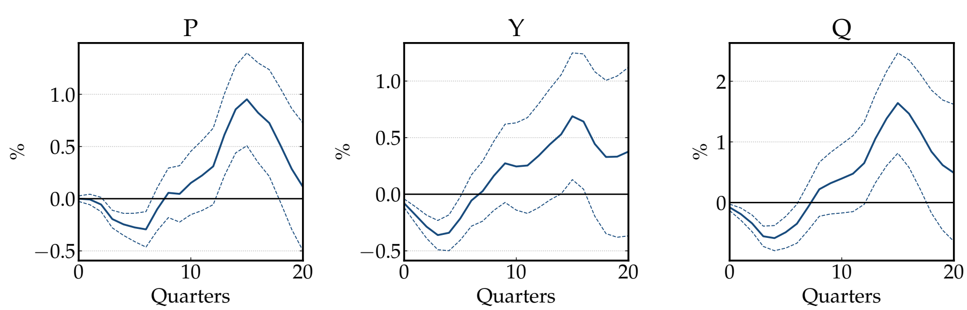
Notes: Calibration plots showing simulated moments for a range of costs of information parameters (ω_L). We simulate models for a panel of 100 firms and for 1,000 periods with 100 periods burn-ins. Simulated moments are generated with regressions discussed in the text in Section 5:

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

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.

E.3 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 US 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 a 1.6% peak increase in nominal demand.