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

Wenting Song
Bank of Canada

Samuel Stern
University of Michigan

October 14, 2022

Abstract

This paper provides direct evidence of the importance of firm attention to macroeconomic dynamics. We construct a text-based measure of firm attention to macroeconomic 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.

JEL: D83, E44, E52
Keywords: Rational inattention, monetary policy, natural language processing

*Song (wsong@bank-banque-canada.ca): Bank of Canada; Stern (sternsa@umich.edu): University of Michigan. This paper is based on Song's job market paper. 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 Herreño, Yulei Luo, Alistair Macauley, Albert Marcet, Michael McMahon, Laura Veldkamp, and the participants at numerous seminars and conferences for helpful comments and discussions. The views expressed herein are those of the authors and not necessarily those of the Bank of Canada.

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) 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, yet 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 provides some of the first 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 with time-varying attention, the number of firms that referenced macroeconomic news rose 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, Sack and Swanson, 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 examine how attention affects firm performance under varying degrees of aggregate uncertainty. We construct an uncertainty index using forecast dispersion from the Survey of Professional Forecasters and measure performance along three dimensions: profitability, financial performance, and survival. The resulting estimates show that attentive firms outperform their inattentive competitors under increased uncertainty. Interestingly, attention appears to weakly reduce performance in low-uncertainty environments.

Finally, we present a quantitative rational inattention model calibrated using our new measure to study the aggregate implications of incomplete firm attention. In the model, 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. Consistent with our empirical findings, attentive firms have higher output semi-elasticities to expansionary monetary shocks and lower semi-elasticities to contractionary shocks. We incorporate observed countercyclicality of firm attention to show that the efficacy of monetary policy declines as the share of attentive firms rises and more firms set prices closer to the optimum. This new interpretation of attention-dependent monetary policy implies that central banks should expect policy interventions 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[1] and consequently the need for empirical measures[2]. Existing research has successfully measured attention in lab experiments (Reutskaja, Nagel, Camerer and Rangel, 2011), field experiments (Bartoš, Bauer, Chytilová and Matějka, 2016; Fuster, Perez-Truglia, Wiederholt and Zafar, 2018), and for individual consumers (McCaulay, 2020). Our methodology complements those measures as well as survey-based evidence on firm expectations by Tanaka, Bloom, David and Koga (2019), Coibion, Gorodnichenko and Kumar (2018), and Candia, Coibion and Gorodnichenko (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, 2020; Yang, 2022), asset prices (Van Nieuwerburgh and Veld-

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

[2] See Gabaix (2019) and Mackowiak, Matejka and Wiederholt (2021) for comprehensive surveys of existing measure of attention.

kamp, 2009), discrete choices (Matějka and McKay, 2015; Caplin, Dean and Leahy, 2019), aggregate dynamics (Maćkowiak and Wiederholt, 2015; Afrouzi and Yang, 2021), and reconciling micro and macro evidence (Auclert, Rognlie and Straub, 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 relates to existing literature on state dependencies of monetary policy. Tenreyro and Thwaites (2016) estimate non-linear responses in monetary policy which are weaker in recessions than in expansions. Vavra (2014), McKay and Wieland (2019) and Ottonello and Winberry (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.

Fourth, 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 has used textual analysis to study financial constraints (Buehlmaier and Whited, 2018), central bank communication (Hansen, McMahon and Prat, 2018), firm-level political risk (Hassan, Hollander, Van Lent and Tahoun, 2019), inflation expectation formation (Larsen, Thorsrud and Zhulanova, 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.

In a related paper, Flynn and Sastry (2021) independently and contemporaneously develop a text-based measure of firm attention to macroeconomic topics. They show, like we do, that their measure is countercyclical. Whereas we show that the stock prices of more attentive firms rise relative to less attentive firms in response to both positive and negative monetary shocks, they compare firms' labor market choices to those of a neoclassical model with full information, and show the gap between model and firm behavior is negatively correlated with firm attention both over time and across firms. Together these two papers present compelling evidence that our text-based measures contain information that is useful in predicting economic outcomes, and that these predictions are consistent with interpreting these measures as measures of attention.

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 present the mitigating effects of attention for uncertainty; in Section 6 we construct a quantitative model of rational inattention and conduct policy counterfactuals; Section 7 concludes.

2. Textual Measure of Attention

This section presents our measure of firm attention, conducts preliminary validation exercises, and documents stylized facts about attention. We show that cross-industry patterns of attention and their correlation with price adjustment are consistent with predictions about attention behavior. We then highlight two key stylized facts: aggregate attention moves countercyclically over the sample period, and the majority of firms remain polarized between never and always paying attention. The section concludes with some reflections on the limitations and promise of textual analysis as a tool for measuring attention.

2.1. Data and methodology

10-K filings Our analysis uses all electronically available 10-K filings by publicly listed US companies between 1994 and 2019[3]. Under Regulation S-K, the U.S. Securities and Exchange Commission (SEC) requires all public companies to disclose audited financial statements and a description of business conditions in these filings each year. Companies were phased

[3] Our methodology is also well-suited for quarterly 10-Q filings. However, we exclude these filings because they are typically less descriptive and do not require audited financial statements.

into mandatory electronic filing between 1993 and 1996, implying that our sample covers the universe of filers since 1996[4]. The final sample contains 201,751 documents submitted by 35,655 unique firms. Table 1 summarizes the length of these documents and unique vocabulary used by filers.

Table 1: 10-K filing length and vocabulary size

	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: (i) recent or future firm performance and (ii) the risk factors that shareholders face by investing in the company. The former context usually appears in Item 7, which requires managers to discuss the firm's financial conditions and results of operations. This section is written as a narrative and its length varies widely across firms (for instance, Item 7 of Alphabet's 2020 10-K filing is 17 pages long). Economic conditions as a source of risk appears in Items 1A and 7A, which detail general firm risks and near-term market risks, respectively.

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

[4] See SEC Release No. 33-7427 for more information about the phase-in process.

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

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

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

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

Total word count is generated by following the parsing strategy in Loughran and McDonald (2011): each text is stripped of all numbers and "stop words" such as articles and then mapped onto a dictionary of all words that appear in our sample of 10-K filings.

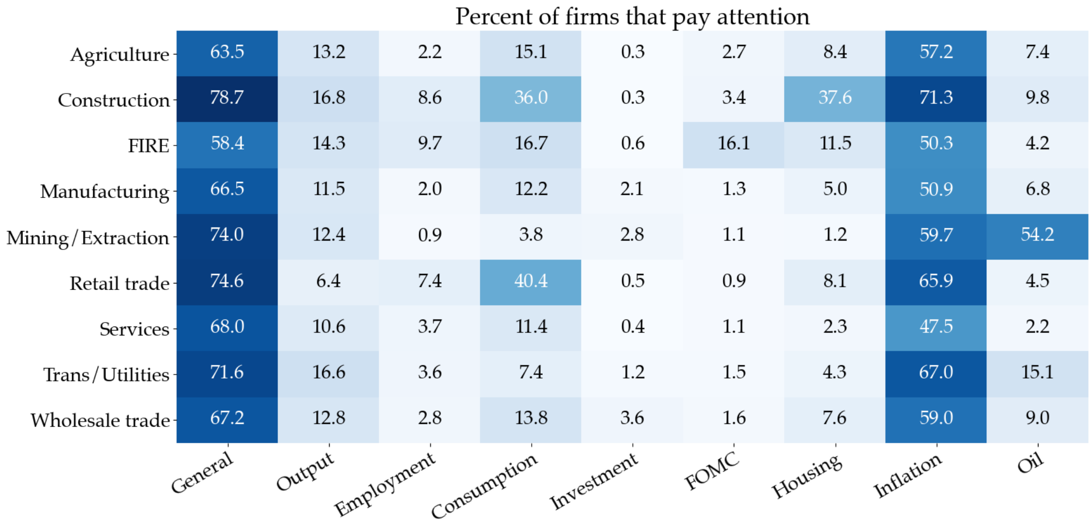
We treat *prevalence* as our baseline measure of firm attention in the majority of the paper. Since both measures are closely related, this avoids presenting duplicate results unnecessarily. The prevalence measure is also less susceptible to contamination by changes unrelated to firm attention. For instance, *intensity* will decrease mechanically if a firm expands its annual filing without mentioning any of the topics listed above. Nonetheless, *intensity* is essential for understanding the intensive margin of attention and features in documenting countercyclical attention below.

2.2. Preliminary validation exercises

This section uses cross-industry variation to test whether our attention measure behaves as expected for each economic topic and is consistent with a fundamental prediction of incomplete information models. We interpret the results as preliminary evidence of our measure's accuracy before presenting firm-level evidence in Section 4.

Cross-industry patterns of attention We first check whether attention to the nine topics listed above is concentrated among commonly associated industries. Figure 1 reports the share of firms in each industry that pays attention to each topic, where industry is defined

Figure 1: Average industry attention by topic



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.

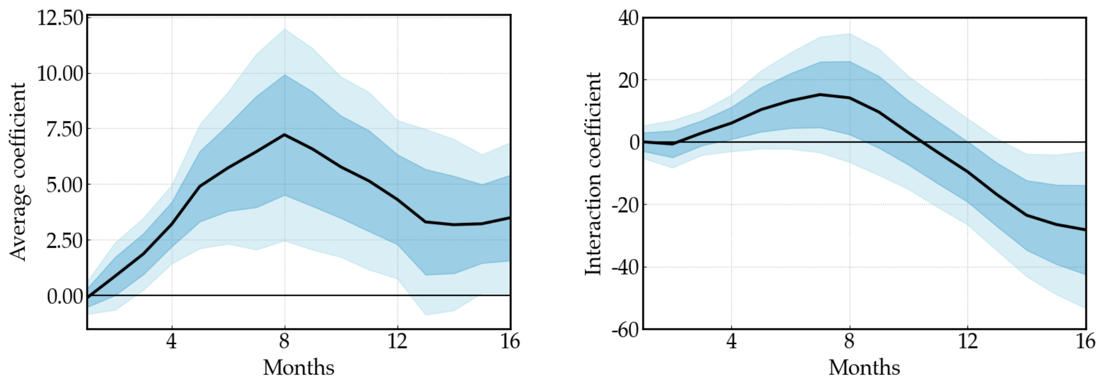
using 2-digit NAICS codes from `Compustat`. The quality of our attention measure varies by topic so results should be interpreted across industries rather than across topics.

By and large, attention appears highest among the industries related to a given topic: mining/extraction has the highest share of firms that pay attention to oil prices; retail trade pays the greatest attention to consumption; and the financial sector (FIRE) pays the most attention to monetary policy (FOMC). This cross-industry pattern serves as a common sense check for our prevalence measure and suggests that textual analysis methods are capable of identifying firm attention.

Price adjustment following monetary shocks We next test whether more attentive industries adjust prices faster following monetary policy shocks as predicted by models of incomplete information (Mankiw and Reis, 2002; Mackowiak and Wiederholt, 2009; Woodford, 2009). The effect of attention on an industry's price response is estimated using the interaction between high-frequency monetary shocks – constructed as in Gorodnichenko and Weber (2016)[5] – and average industry attention in a local projection model (Jordà, 2005).

[5]See Section 4.1 for a detailed description of their methodology.

Figure 2: Attention and price adjustment



(a) Average responses to monetary policy **(b)** Marginal effects of attention

Notes: Panel (a) and (b) reports the average and marginal coefficients, β_ν^h and $\beta_{d\nu}^h$, respectively, from estimating Equation (1) over months $h = 1, \cdots, 16$. We exclude finance and utility industries. Standard errors are double clustered by sector and year. Confidence intervals of 65% and 90% are reported. We have normalized the sign of monetary shocks so that positive shocks are expansionary.

Over an h-month horizon, our model takes the form,

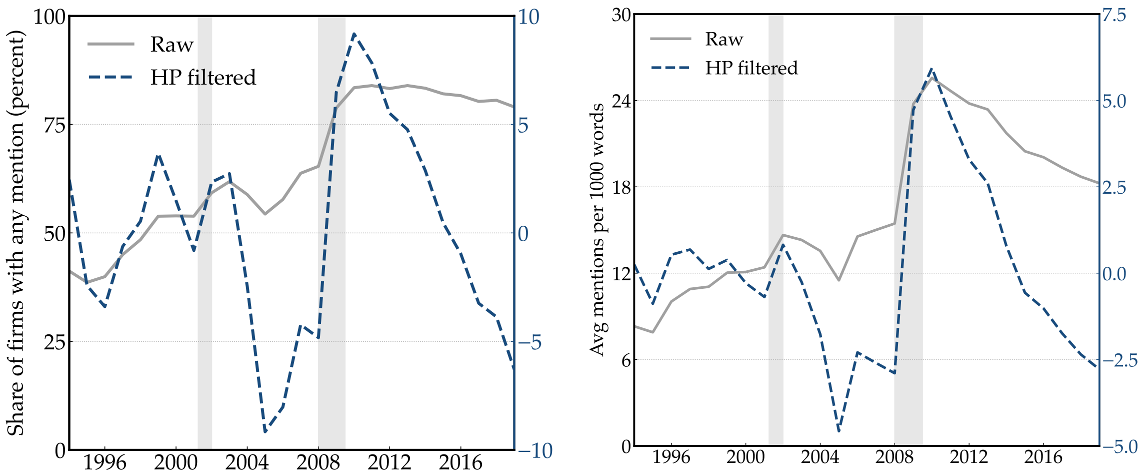
$$\log P_{s,t+h} - \log P_{s,t} = \alpha_s + \alpha_t + \beta_\nu^h \nu_t^{\mathrm{M}} + \beta_d^h d_{st} + \beta_{d\nu}^h d_{st} \nu_t^{\mathrm{M}} + \Gamma' Z_{st} + \varepsilon_{st} \tag{1}$$

where $P_{s,t}$ is the BLS Producer Price Index (PPI) for industry s (4-digit NAICS) in month t, ν_t^{M} denotes the monetary shock in month t, d_{st} denotes average industry attention, and Z_{st} is a vector of controls including industrial production, a recession indicator, and industry size. We include industry and time fixed effects, $\{\alpha_s, \alpha_t\}$, and cluster standard errors by both industry and year. For ease of interpretation, monetary shocks are normalized so that positive values correspond to expansionary shocks.

Figure 2 Panel (a) plots the estimated average price response, β_ν^h, and shows that prices rise in a hump-shaped manner following an expansionary monetary shock. At its peak, an unanticipated 25 basis point rate cut causes prices to rise by 1.8%. Panel (b) plots the marginal effect of attention on an industry's price response, $\beta_{d\nu}^h$. More attentive industries raise prices faster in the first 10 months after a monetary shock, though the effect begins to decline after about seven months as less attentive industries appear to catch up. This result is consistent with imperfect information models that predict faster price adjustment

Figure 3: Time series of attention to "economic conditions"

(a) Prevalence (d_{it}) **(b)** Intensity (s_{it})



Notes: Time series of firm attention to the keyword "economic conditions". Left panel plots the prevalence measure and reports the share of firms that mention the keyword. The right panel plots the intensity measure and reports the average mentions of the keyword per 1,000 words. "Raw" refers to the unfiltered series and "HP filtered" refers to the cyclical components of the HP-filtered series with smoothing factor $\lambda = 400$. Shares are reported in percent.

by attentive firms (e.g., Mackowiak and Wiederholt, 2009).

2.3. Stylized facts about firm attention

This section explores how attention varies over time and between firms. We document two key features – countercyclicality and polarization – and then summarize the relationship between attention and other firm characteristics.

Countercyclical attention to economic conditions Both the share of firms that mention economic keywords and the intensity with which they are discussed vary countercyclically over our sample period. This is illustrated in Figure 3, which plots the annual average *prevalence* and *intensity* measures for the phrase "economic conditions," as well as detrended versions using an HP-filter ($\lambda = 400$).

Panel (a) shows that the share of firms mentioning "economic conditions" has steadily increased since 1994, with particularly rapid growth during the Great Recession. Between 2008 and 2010, aggregate attention jumped by about 15 percentage points and remained elevated for the rest of the sample period. Average intensity in Panel (b) similarly spiked during the Great Recession but declined faster in subsequent years. We point to these sharp dynamics around the Great Recession as evidence of countercyclical attention, but we also acknowledge that our sample is limited and a longer time series is needed to confirm this result.

Some models with endogenous firm attention predict the countercyclicality displayed in Figure 3. Mackowiak and Wiederholt (2009) considers imperfect information firms that allocate attention between aggregate and idiosyncratic state variables. Increased aggregate uncertainty (itself countercyclical) induces these firms to shift attention toward aggregate conditions. Chiang (2021) decomposes the impact of lower expected productivity on attention into income and substitution effects. Countercyclical attention emerges among goods-producing agents when their marginal utility rises faster than the returns to attention falls under lower productivity.

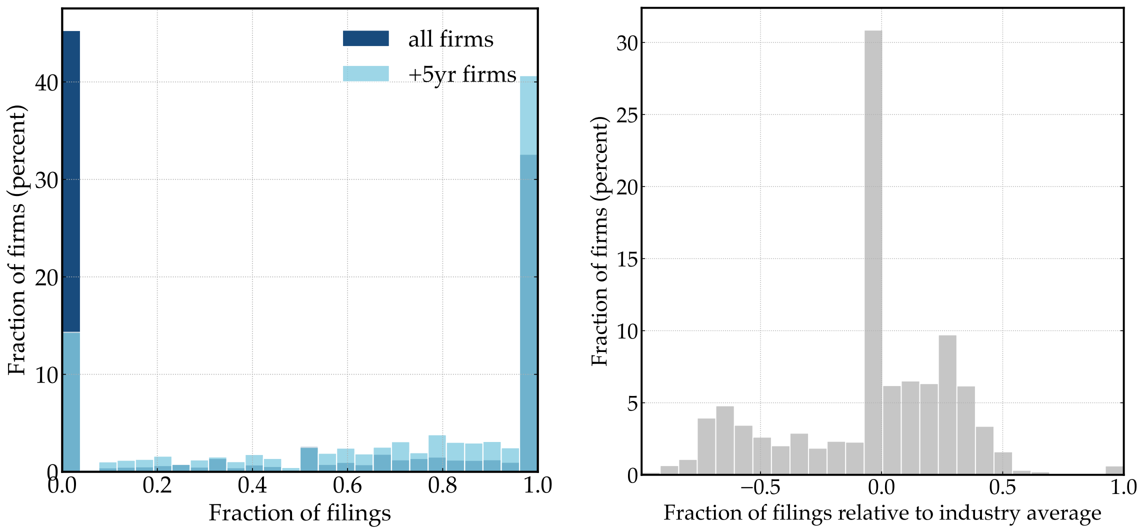
Polarization in firm attention Despite the countercyclical dynamics documented above, most firms in our sample are polarized between either always or never discussing economic conditions in their 10-K filings. Figure 4 illustrates this by plotting each firm's share of filings that mention the same key phrase, "economic conditions." The resulting distribution in Panel (a) is heavily concentrated at each extreme, with about three quarters of firms taking values of either 0 or 1. This suggests that most variation in attention occurs across firms rather than within, and countercyclical variation is caused by a relatively small subset of filers.

To test whether polarized attention is driven by firms with few filings, we overlay a second histogram in Panel (a) that restricts to firms with at least five years of data. This adjustment greatly reduces the share of firms that never pay attention, yet over half remain polarization between always- and never-attentive firms.

We also test whether cross-industry differences in attention are responsible for polarization. Panel (b) in Figure 4 demeans by industry, which explains approximately one quarter of attention variation. The distribution now contains a large mass of firms around their industry average (i.e. industries with little attention dispersion), while the remaining firms

Figure 4: Share of filings that mention "economic conditions"

(a) All firms **(b)** Within industry



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.

form a bimodal pattern consistent with polarization.

The presence of any inattentive firms may be surprising given that most US macroeconomic data are readily available. However, this result is consistent with a broader interpretation of attention that includes information processing, communication, and optimal response in addition to information acquisition. As highlighted in Reis (2006), firms likely require significant resources and expertise to process, summarize, and forecast macroeconomic series into sufficient statistics that informs firm decision-making. This is consistent with plant-level evidence from zbaracki2004managerial. To this end, Abis and Veldkamp (2020) estimates a data production function that takes labor and capital inputs to to produce knowledge from unstructured data.

Firm characteristics and attention Finally, we explore how attention relates to firm size, age, and leverage. The first set of results focuses on cross-sectional differences between attentive and inattentive firms, while the second set uses within-firm variation to study the

Table 2: Firm characteristics and attention

(a) Cross-firm 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) Within-firm 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 $d_{it} = \delta_t + \delta_j + \beta \cdot x_{it} + \varepsilon_{it}$, 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.

comovement of attention and other characteristics. Results are estimated using annual firm data and the following pair of regressions,

$$\text{Cross-firm variation:} \quad d_{it} = \delta_t + \delta_j + \beta \cdot x_{it} + \varepsilon_{it} \tag{2}$$

$$\text{Within-firm variation:} \quad d_{it} = \delta_i + \beta \cdot x_{it} + \varepsilon_{it} \tag{3}$$

where x_{it} is either firm size, age, or leverage[6], and d_{it} is our prevalence measure of attention. Equation (2) includes time and industry (4-digit NAICS) fixed effects, $\{\delta_t, \delta_j\}$, to highlight cross-sectional variation, while Equation (3) includes firm fixed effects, δ_i, to isolate within-firm variation.

Data on firm characteristics are from `Compustat`. Size is measured as the log of total assets, age as the number of years since a firm first appeared in Compustat, and leverage as the debt-to-asset ratio[7]. All firm covariates are standardized so their units are in standard deviations.

Results for this analysis are displayed in Table 2. Panel (a) shows that attention to general economic conditions is higher among larger, older, and less leveraged firms. Panel (b) shows

[6] Existing literature has found each of these characteristics to be relevant for the transmission of macroeconomic policy (Gertler and Gilchrist, 1994; Cloyne, Ferreira, Froemel and Surico, 2018; Ottonello and Winberry, 2020)

[7] See Appendix Table A.2 for summary statistics of firm characteristics by attention.

that a firm's attention increases with age and size, but does not have a strong relationship with leverage. Notably, these estimates are larger within-firm than across, suggesting a strong lifecycle pattern. One possible explanation for why attention increases with age is that firms are "scarred" by previous episodes when increased attention was necessary. This would also explain why our *prevalence measure* remains elevated following the Great Recession in Figure 3a.

2.4. Limitations and promise of textual measures

Boilerplate language is a key concern of using regulatory filings to measure firm attention. Filings are often written collaboratively between managers and legal departments, and evidence suggests that firms include certain statements within 10-K filings to appease investors or lower liability (Cao, Jiang, Yang and Zhang, 2020). Moreover, firms likely save time and resources by revising previous filings rather than starting from scratch each year.

The methods used above cannot distinguish between authentic attention to macroeconomic conditions and "cheap talk" references or recycled language that does not reflect current management practices. We attempt to address this shortcoming in Appendix C.1 by measuring the diversity in filing language with a Jaccard score of lexical similarity and testing whether our main findings are robust within the most linguistically diverse 10-K sections. Table A.7 confirms that our key findings are not driven by the most repetitive and standardized sections.

Even greater measurement error may arise from misidentifying firms as inattentive, which raises concerns about underestimating aggregate attention. False negatives can result from methodological limitations or variation in the amount of information that firms choose to disclose. 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. Furthermore, firm managers are obligated under Regulation S-K to disclose any material risk factors.

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. Appendix C.2 uses a Latent Dirichlet Allocation (LDA) unsupervised model to categorize words neighboring each keyword and produces nine unlabeled "topics" in which keywords appear. Appendix C.3 uses the itemized structure of 10-K filings to identify sections that contain the most keywords. This analysis shows that keywords typically appear in sections that discuss firm risk factors (Item 1A) and business operations (Item 7A).

3. Illustrative Framework

This section derives a testable implication of firm attention to address a key identification challenge for our text-based measure: whether it captures *exposure* rather than attention to macroeconomic conditions. We present a stylized model in which firms are heterogeneous in both attention and exposure to an aggregate state variable, and then consider how firm outcomes vary with each source of heterogeneity. The model predicts contradictory responses to aggregate shocks under varying attention and exposure, which guides our empirical design in Section 4. The model environment is kept minimal to highlight key mechanisms before Section 6 incorporates 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[8]:

$$\hat{\pi}(\hat{s},\hat{a}) = \pi_s(\bar{s},\bar{a})\bar{s}\hat{s} + \frac{1}{2}\left(\pi_{ss}(\bar{s},\bar{a})\bar{s}^2 - \pi_{aa}(\bar{s},\bar{a})\bar{a}^2\right)\hat{s}^2 + \frac{1}{2}\pi_{aa}(\bar{s},\bar{a})\bar{a}^2(\hat{a}-\hat{s})^2 \tag{4}$$

where $\bar{s}$ and $\bar{a}$ denote the steady-state values, $\hat{\pi}$, $\hat{s}$ and $\hat{a}$ denote the log deviations from the

[8]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.

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.

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

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:

$$r = \hat{v} - \hat{v}_{-1} \tag{5}$$

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 (expansionary) 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:*

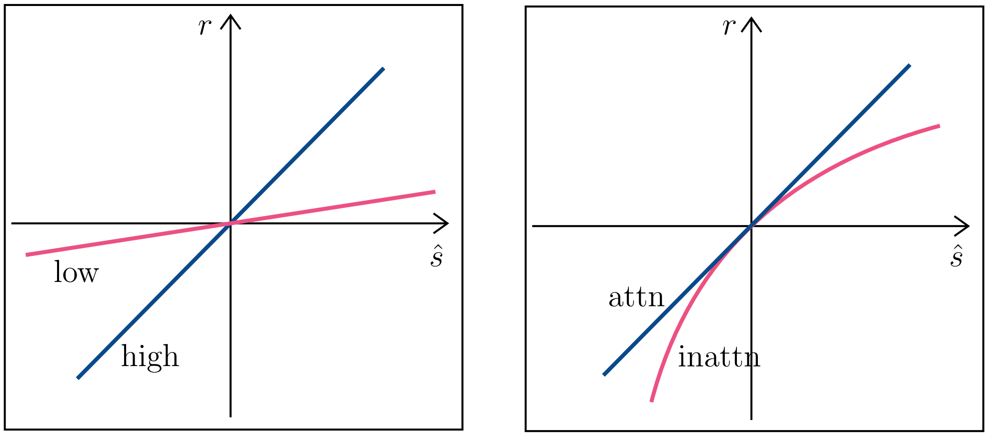
$$\begin{cases} \frac{\partial r_i}{\partial \hat{s}} > \frac{\partial r_j}{\partial \hat{s}} & \textit{if } \hat{s} > 0 \\ \frac{\partial r_i}{\partial \hat{s}} = \frac{\partial r_j}{\partial \hat{s}} & \textit{if } \hat{s} = 0 \\ \frac{\partial r_i}{\partial \hat{s}} < \frac{\partial r_j}{\partial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

Proof. See Appendix D.2 ■

Figure 5 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,

Figure 5: Model predictions for exposure vs attention

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



Notes: Illustration of model predictions of return elasticity with respect to aggregate shocks. Verticle axes represent conditional realized return, and horizontal axes represent the magnitude of shocks. Left panel shows return elasticity for firms that are 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.

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. Asymmetric Response to Monetary Shocks

We now test the hypothesis that attentive firms respond better to aggregate shocks using a high-frequency identification strategy. Shocks are constructed as plausibly exogenous monetary policy surprises following FOMC announcements, and resulting changes in firm value are measured using stock prices. We use our *prevalence* measure to estimate the relative performance of attentive firms, and then test whether they fare better following both positive and negative shocks[9]. Results in this section serve the dual purpose of validating our text-based attention measure and quantifying the expected benefits of attention to economic conditions.

Stock prices are a particularly informative outcome variable because they are forward-looking and similarly high frequency as our monetary shocks. The cumulative effect of a rate surprise on expected future profits will be reflected quickly in a firm's stock price. By restricting to a narrow window around the shock, we isolate this price effect while avoiding other confounding factors. In comparison, a firm's investment and hiring decisions will be smoothed over a longer horizon and any low-frequency response is confounded by other factors that influence these choices. These limitations are exacerbated by the low statistical power of high-frequency monetary shocks, preventing precise estimates of investment and hiring responses[10].

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 the exposure to monetary policy.

4.1. Data

Monetary policy shocks are constructed using the high-frequency identification strategy developed by 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

[9]This testable implication from Section 3 works for any aggregate shock with a related attention measure. We use high-frequency monetary shocks as "proof of concept" because they are familiar and well-identified. See Ramey (2016) for a comprehensive survey of alternative aggregate shocks.

[10]See Nakamura and Steinsson (2018) for further discussion of this "power problem."

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,

$$\nu_t = \frac{D}{D-\tau}(ff^0_{t+\Delta t^+} - ff^0_{t-\Delta t^-}), \tag{6}$$

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 by 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,

$$\begin{aligned} r_{it} = {} & \beta_d d_{it} + \beta_1 \mathbb{1}_{\nu_t>0} + \beta_{\nu_+}\nu_t \mathbb{1}_{\nu_t>0} + \beta_{\nu_-} v_t \mathbb{1}_{\nu_t<0} \\ & + \beta_{d\nu_+}(d_{it}\nu_t \mathbb{1}_{\nu_t>0}) + \beta_{d\nu_-}(d_{it}\nu_t \mathbb{1}_{\nu_t<0}) + \delta_j + \delta_j \nu_t + \Gamma_1' X_t + \Gamma_2' X_t \nu_t + \varepsilon_{it}, \end{aligned} \tag{7}$$

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 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 one percent 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 (7) are presented in the third column. We test whether

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 (FOMC)		-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 (7):

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

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 (7) for the sample up to 2007. Standard errors are clustered at the shock level and reported in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

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. Column 4 shows that this result is not driven by outsized monetary surprises during the Great Recession nor unconventional monetary policy at the zero lower bound by ending the sample in 2007.

The *asymmetric* response to positive and negative shocks rules out alternative interpretations of the textual measure that predict a symmetric effect. The foremost alternative discussed in Section 3 is that the textual measure identifies exposure to monetary shocks rather than attention. Any such symmetric effect would also appear in the interaction coefficient, $\beta_{d\nu}$, in Column 2, which is only weakly positive. Appendix B.1 further shows that directly estimating and controlling for exposure to monetary shocks leaves our main findings unchanged.

Suboptimal responses to monetary policy by inattentive firms reported in Table 3, together with the large fraction of inattentive firms documented in Figure 4, provide some of the first direct evidence on 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%.

Ruling out alternative sources of asymmetry We now consider alternative explanations for the asymmetric price response documented above. Each explanation is tested by augmenting our baseline model to include interaction terms for a confounding variable, c_{it}, that mirror the interaction terms for firm attention, d_{it}. The resulting "horse race" model takes the form,

$$\begin{aligned} r_{it} = &\delta_j + \delta_j \nu_t + \beta_1 \mathbb{1}_{\nu_t>0} + \beta_{\nu_+} \nu_t \mathbb{1}_{\nu_t>0} + \beta_{\nu_-} v_t \mathbb{1}_{\nu_t<0} \\ &+ \beta_d d_{it} + \beta_{d\nu_+}(d_{it}\nu_t \mathbb{1}_{\nu_t>0}) + \beta_{d\nu_-}(d_{it}\nu_t \mathbb{1}_{\nu_t<0}) \\ &+ \beta_c c_{it} + \beta_{c\nu_+}(c_{it}\nu_t \mathbb{1}_{\nu_t>0}) + \beta_{c\nu_-}(c_{it}\nu_t \mathbb{1}_{\nu_t<0}) + \Gamma_1' X_t + \Gamma_2' X_t \nu_t + \varepsilon_{it}, \end{aligned} \tag{8}$$

where, as in the baseline specification, we include industry fixed effects, industry-by-shock fixed effects, a vector of firm controls, and its interaction with monetary shocks. If the main result, $\beta_{d\nu_-} < 0 < \beta_{d\nu_+}$, holds true then we rule out c_{it} as a confounding source of asymmetry.

The first factor considered is productivity. Van Nieuwerburgh and Veldkamp (2006) presents a model in which higher productivity increases learning as well as production. If productivity determines both information acquisition and the response to aggregate shocks, it could explain the asymmetric result found above. Productivity is constructed following the control function approach from Olley and Pakes (1996) and estimated using GMM as in Wooldridge (2009)[11].

Management quality is another potential confounder that could explain both attention and firm performance. Effective managers who capitalize on opportunities during expansionary shocks and mitigate losses from contractionary shocks will generate the same asymmetric performance pattern documented in our main results. We approximate a firm's management quality using the share of board members who hold a graduate degree since existing research documents a strong relationship between education and management quality (Bloom and Van Reenen, 2010)[12]. Data on the educational attainment is from `BoardEx`, which covers publicly traded US firms.

The third variable considered is a firm's financial performance measured using return on assets (ROA). Managers may feel compelled to cite macroeconomic conditions when explaining recent performance, and a tendency for well-performing firms to cite such conditions could generate the asymmetry observed.

Finally, we control for the length of a firm's SEC filing as a measure of its preference for information provision. Longer filings – measured using log word count – offer more opportunities for managers to mention macro keywords and signal commitment to due diligence. If thorough due diligence engenders investor confidence, then stocks should outperform following positive and negative monetary shocks.

Table 4 reports the estimates for Equation 8 using each factor described above: productivity, management quality, profitability, and filing length. As in our baseline results,

[11] Firm output is measured as total sales (SALE) deflated by the BEA's implicit price deflator and labor is defined as total number of employees. Firm capital is constructed using the perpetual inventory method. Capital stock for each firm is initialized as Gross Property, Plants, and Equipment (PPEGT), and annual net investment in all subsequent years is defined as the change in Net Property, Plants, and Equipment (PPENT). Capital each period is defined as the sum of capital from the previous period and net investment. Finally, nominal capital is deflated using the BEA's investment price deflator (FRED: A008RD3A086NBEA).

[12] Graduate degrees include: MBA, MS, MSC, MA, JD, MD, MPA, MSE, PHD, and any degree names which include "master" or "doctor."

Table 4: Controlling for alternative explanations of asymmetry

Control Variable:	**Productivity**	**Mgmt Quality**	**Reported Profit**	**Filing Length**
Shock $\times \mathbb{1}_{v_t>0}$	6.88**	0.41	4.55*	-2.59
	(2.87)	(2.63)	(2.66)	(6.40)
Shock $\times \mathbb{1}_{v_t<0}$	-1.39	-10.53**	-4.04	21.86
	(4.04)	(4.37)	(3.92)	(14.24)
Attention	-0.13	-0.10*	-0.07	-0.07
	(0.10)	(0.06)	(0.06)	(0.05)
Attn $\times$ Shock $\times \mathbb{1}_{v_t>0}$	3.08**	2.59***	2.00***	1.81***
	(1.43)	(0.91)	(0.71)	(0.60)
Attn $\times$ Shock $\times \mathbb{1}_{v_t<0}$	-5.45***	-8.18**	-5.78*	-5.32*
	(1.78)	(3.34)	(3.16)	(2.95)
Control Var	0.04***	-0.07	0.04*	-0.01
	(0.01)	(0.06)	(0.02)	(0.04)
Control $\times$ Shock $\times \mathbb{1}_{v_t>0}$	0.10	1.53**	-0.97	0.80
	(0.14)	(0.73)	(1.67)	(0.49)
Control $\times$ Shock $\times \mathbb{1}_{v_t<0}$	-0.04	-2.54	-8.36**	-2.68*
	(0.22)	(2.58)	(3.93)	(1.59)
Observations	376644	324154	574804	575667
R^2	0.027	0.041	0.026	0.026
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	no
Wald test p-value: Attention	**0.001**	**0.003**	**0.027**	**0.031**
Wald test p-value: Control	0.387	0.143	0.126	0.041

Notes: This table augments Column 3 of Table 3 to control for four potential confounding sources of asymmetry. The estimated regression has the form,

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

where c_{it} represents the alternative "Control" variable. As with attention, the control variable is interacted with both positive and negative monetary shocks. All other features of the model specification remain unchanged from Table 3. The four control variables considered are: (1) firm productivity estimated as in Olley and Pakes (1996); (2) management quality approximated with board member educational attainment; (3) profit measured as earnings before extraordinary items over total assets; and (4) filing length measured as the log word count of the 10-K filing. The final two rows report p-values of Wald tests for $H_0 : \beta_{d\nu_+} = \beta_{d\nu_-}$ and $H_0 : \beta_{c\nu_+} = \beta_{c\nu_-}$, respectively.

attentive firms experience a larger increase in market value following an expansionary shock and a smaller contraction following a contractionary shock. The estimates are statistically significant, with similar magnitudes as those in Table 3. While some of the control variables (e.g. filing length) also display an asymmetric effect on firms' responses to monetary shocks, the explanatory power of firm attention remains under all specifications. All four Wald tests for $H_0 : \beta_{d\nu+} = \beta_{d\nu-}$ are rejected at 5% significance. In Appendix Table A.3, we show that these results are also robust to excluding zero-lower-bound periods.

Additional robustness checks Further robustness analysis pertaining to the identification of high frequency monetary shocks can be found in Appendix B. Appendix B.2 controls for the information effect of FOMC announcements using Greenbook forecast revisions, and Appendix B.3 tests whether aggregate conditions confound the estimated effect of high frequency shocks. In each case, our main results remain robust.

5. Attention, Performance, and Aggregate Uncertainty

This section explores how attention affects firm performance under varying levels of aggregate uncertainty. One implication of our illustrative model is that the performance gap between attentive and inattentive firms widens with the magnitude of nominal demand shocks. Returns to attention should therefore increase in periods of greater uncertainty and larger shocks[13].

We test this prediction by estimating the interaction effect between attention and uncertainty on firm performance. Aggregate uncertainty is measured using the interquartile range of quarterly forecasts for real GDP, inflation, and unemployment from the Survey of Professional Forecasters. Each series is standardized over our sample period (1994-2019) and then averaged into a composite uncertainty index.

Firm performance is measured along three dimensions: profitability, financial performance, and survival. Profitability is measured as a firm's return on assets (ROA), which we construct using earnings before extraordinary items over total assets. Financial perfor-

[13]See Appendix D.3 for an extended illustrative framework that incorporates time-varying uncertainty.

mance is measured as return on equity (ROE) using earnings before extraordinary items over market capitalization. Finally, survival is defined as whether a firm remains in operation in the next year. Each variable is constructed using annual Compustat data, and ROA and ROE are winsorized at 1%.

Our regression model takes the following form,

$$y_{it} = \alpha_j + \beta d_{it} + \delta\sigma_t + \gamma d_{it} \cdot \sigma_t + \Gamma' Z_{it} + \varepsilon_{it},$$

where y_{it} represents one of the three performance variables defined above, d_{it} is our binary attention measure, σ_t is aggregate uncertainty, α_j captures industry fixed effects by 4-digit NAICS, and Z_{it} is a vector of firm controls including size, age, and 10-K filing length (as previously defined). Standard errors are clustered by both year and industry. We extend the model to future outcomes, $y_{i,t+h}$ to capture any lagged effects of attention on performance.

Results from this analysis are reported in Table 5. On average, aggregate uncertainty reduces profitability, financial performance, and the probability of survival, which is consistent with existing models of uncertainty (e.g., Bloom, Bond and Van Reenen, 2007).[14] Attention to macroeconomic conditions, however, mitigates the negative effects of uncertainty: in periods of high uncertainty, attentive firms have higher profitability, better financial performance, and higher probability of survival. Interestingly, the first row in Table 5 suggests that attention reduces firm performance under low uncertainty, consistent with models of imperfect information in which firms face a cost of attention and reap the benefit in states with large realized shocks (such as Reis, 2006).

Section 4 showed that attentive firms respond more optimally to monetary shocks. This section finds that these same firms outperform less attentive competitors under elevated aggregate uncertainty. Together, they paint a picture of attentive firms as more responsive to evolving macroeconomic conditions, and highlight the benefits gained for their diligence.

[14] See Leahy and Whited (1996) for a general discussion of firm decisions under uncertainty.

Table 5: Effects of attention on firm responses to uncertainty

	ROE		ROA		Survival	
	Impact	Peak	Impact	Peak	Impact	Peak
Attention (General)	-0.01	-0.02*	-0.02	-0.02**	-0.00	-0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Uncertainty (SPF IQR)	-0.03**	-0.02	-0.04**	-0.03*	-0.01	-0.02*
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Attention × Uncertainty	0.02*	0.03**	0.05**	0.05***	0.01	0.03**
	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Observations	104507	92023	110267	97180	111637	66813
R^2	0.163	0.156	0.246	0.236	0.034	0.028
Clustered SE	yes	yes	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes	yes	yes
4-digit NAICS FE	yes	yes	yes	yes	yes	yes

Notes: The table reports results from estimating $y_{it}^h = \alpha_j + \beta d_{it} + \delta\sigma_t + \gamma d_{it} \cdot \sigma_t + \Gamma' Z_{it} + \varepsilon_{it}$ for horizons $h = 1, \cdots, 5$. The dependent variables y_t include (i) profitability measured with return on assets (ROA, i.e. net income over total assets), (ii) financial performance measured with return on equity (ROE, i.e. net income over equity), and (iii) an indicator variable for firm survival. Independent variables include: the prevalence attention to general economic conditions, d_{it}; macroeconomic uncertainty, σ_t^2, measured as the interquartile range of quarterly growth rate forecasts for real GDP, inflation, and unemployment from the Survey of Professional Forecasters; the interaction between attention and uncertainty; industry fixed effects δ_j; and firm controls Z_{it}. We standardize the interquartile range of each series over our observed sample period, take the absolute average deviation each quarter, and then average these quarterly values each year. The on-impact effect corresponds to the estimates for $h = 1$. The peak effect corresponds to the estimated marginal effect over the 5-year horizon. Standard errors are clustered at the shock level and reported in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

6. Quantitative Model

Our attention measure can inform model-based analysis as well as empirical studies. This section presents a quantitative model in which macro inattention drives monetary non-neutrality. Both the rate of attentive firms and the cost of inattention are calibrated using the prevalence measure presented in Section 2. We use the model to quantify the importance of attention. Our main finding is that the efficacy of monetary policy depends on aggregate attention when firms face information frictions.

6.1. Environment

We start with a canonical dynamic general-equilibrium model with rationally inattentive firms as in Mackowiak and Wiederholt (2009) and Afrouzi and Yang (2021). Time is discrete and infinite. The economy consists of a representative household, firms, and a central bank. Households and the central bank have full information about the economy, while firms pay a cost proportional to information obtained (measured using Shannon mutual information as in Sims, 2003). We allow firms to differ ex-ante in their marginal costs of information, which is motivated by the heterogeneity documented in Section 2.3.

Household A representative household maximizes its life-time utility,

$$\max_{C_{it},N_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t (\log C_t - \psi N_t), \tag{9}$$

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, $C_t = \left(\int_0^1 C_{it}^{\frac{\varepsilon-1}{\varepsilon}} dj\right)^{\frac{\varepsilon}{\varepsilon-1}}$, where ε 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 firms' profits Π_t. The household budget constraint is given by:

$$\int_0^1 P_{it} C_{it} di + B_t \leq W_t N_t + (1 + \iota_t) B_{t-1} + \Pi_t. \tag{10}$$

Central bank The central bank targets aggregate money supply similar to Caplin and Spulber (1987) and Gertler and Leahy (2008). Nominal aggregate demand follows the process

$$\Delta \log Q_t = \rho \Delta \log Q_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2). \tag{11}$$

Firms There is a unit mass of monopolistically-competitive firms indexed by $i \in [0, 1]$. Firms are owned by the household and operate a decreasing-returns-to-scale production technology with labor as its only input,

$$Y_{it} = N_{it}^{\gamma}. \tag{12}$$

Firms can flexibly set prices but do not observe aggregate demand. Each period, they choose the precision of a noisy signal about demand, s_{it}, whose cost is proportional its information content,

$$2\omega_i \cdot \mathcal{I}(Q_t; s_{it}). \tag{13}$$

The Shannon mutual information, $\mathcal{I}(Q_t; s_{it})$, measures the information gained about aggregate demand from observing s_{it}. Specifically, it measures the expected reduction in entropy between a firm's prior and posterior beliefs about Q_t. A more precise signal results in higher mutual information and is therefore costlier to firms. The information represented by $\mathcal{I}$ can be thought of as a firm's *attention* to the unknown aggregate state.

We assume that firms are heterogeneous in their information-processing technology and face either high or low costs of attention,

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

In the economy, a fraction $\theta \in (0, 1)$ of firms have low information-processing costs, while the other $1 - \theta$ face high costs.

Each period, firms jointly set prices and choose a demand signal to maximize their value,

$$V(S_i^{t-1}) = \max_{P_{it}, s_{it}} \mathbb{E}_t \left[(1/P_t)((\underbrace{P_{it}Y_{it} - \mathcal{C}_t(Y_{it})}_{\text{operational profits}}) - \underbrace{2\omega_i \mathcal{I}(Q_t; s_{it}))}_{\text{info costs}} + \underbrace{\beta \Lambda_{t,t+1} V(S_i^{t+1})}_{\text{cont. value}} \,\middle|\, S_i^t \right], \tag{14}$$

$$\text{s.t.} \qquad Y_{it} = (P_{it}/P_t)^{-\varepsilon} C_t$$

$$S_i^t = S_i^{t-1} \cup s_{it}$$

which consists of operational profits, information costs, and a continuation value. The state variable, S_i^t, represents the history of signals that a firm has received up to time t and reflects its current information set. The function, $\mathcal{C}_t(\cdot)$, is the cost of production and $\Lambda_{t,t+1}$ denotes the stochastic discount factor.

Equilibrium Given nominal aggregate demand $\{Q_t\}_t$ and an initial set of signals, the equilibrium is a set of households' allocations $\{C_t, N_t, B_t\}_t$; firms' allocations $\{s_{it}, P_{it}, N_{it}\}$; and prices $\{\iota_t, P_t, W_{it}\}_t$ such that

i. Given prices, households' allocations solve (9); given prices and an initial set of signals, firms' allocations solve (14);

ii. All markets clear—i.e., for $t \geq 0$ and $i \in [0,1]$, $B_t = 0$, $Y_{it} = C_{it}$, $Y_t = C_t$, and $N_t = \sum_i N_{it}$.

Solution We approximate firm's flow profits with second-order log approximations around the full-information steady state.[15] 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 based on the algorithm for dynamic rational inattention problems (`DRIPs`) developed in Afrouzi and Yang (2021).[16]

6.2. Calibration

We calibrate the model in two steps. First, we endogenously assign a subset parameters. Then, we calibrate the remaining parameters that govern information frictions to match empirical moments.

The top panel of Table 6 shows the calibration for assigned parameters. We calibrate the model quarterly and set the discount rate to be $\beta = 0.96^{1/4}$. The stochastic process

[15] Details of the approximation can be found in Appendix E.2. Log-quadratic approximation is a common simplifying assumption in rational inattention models (see, for example, Mackowiak and Wiederholt, 2009; Afrouzi and Yang, 2021) 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.

[16] Appendix E.3 contains details of the implementation.

for aggregate demand, $\{\rho, \sigma_\nu\}$, is calibrated using quarterly US nominal output between 1994 and 2019. We restrict to manufacturing to match our empirical specification, which compares firms within their sector. The elasticity of substitution is set to $\varepsilon = 10$, implying a steady-state markup of 11%, and the disutility of labor is set to $\psi = 0.90$ to offset the steady-state distortions from monopolistic competition. Finally, We set returns to scale $\gamma = 0.83$ according to the estimate in Basu and Fernald (1997) for US private economy.

Table 6: Calibration

Parameter	Description	Value
Assigned parameters		
β	discount rate	$0.96^{1/4}$
ρ	shock persistence	0.89
σ_ν	shock std. dev.	0.034
ε	elasticity of substitution	10
ψ	disutility of labor	0.9
γ	returns to scale	0.83
Information parameters		
θ	fraction of attentive firms	65%
ω_L	cost of information	1×10^{-6}
ω_H	cost of information	0.62

The bottom panel of Table 6 presents parameters that govern information frictions in the model and are calibrated using our text-based measure of attention. The share of firms with low information costs, θ, matches the share of attentive firms in Figure 3. Since the marginal cost of information is constant and the marginal benefit of information is increasing in prior uncertainty, firms pay a constant level of attention—decreasing in ω_i—to maintain a reservation level of uncertainty. The share of firms with ω_L, therefore, corresponds to the share of attentive firms.

We set ω_L close to zero to reflect that attentive firms face minimal information frictions. We set ω_H, the remaining parameter for cost of information, to match the heterogeneous responses to monetary shocks estimated in Table 3. Stock returns in the model are defined as the log change in a firm's value function in Equation (14), $r_{it} = \log V_{it} - \log \mathbb{E}_{t-1}(V_{it})$. Since our main empirical specification uses the prevalence attention measure, we define a corresponding indicator variable, $d_{it} = \mathbb{1}(\mathcal{I}_{it} > \bar{\mathcal{I}}_t)$, for firm attention above the cross-sectional mean in a given period.[17] Finally, we use ν_t as the monetary shocks.

We simulate the model for a panel of 100 firms and for 1000 quarters, discarding the first 100 quarters as burn-in. With the simulated data, we estimate

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

We set ω_h to target the elasticity $\frac{1}{2}|\hat{\beta}_{dv+}| + \frac{1}{2}|\hat{\beta}_{dv-}|$ from Column (3) in Table 3, which measures the relative stock return losses of firms that do not pay attention. Appendix Figure A.6a shows how ω_H is identified. As ω_H increases and the gap between ω_L and ω_H widens, the simulated elasticity monotonically increases, implying greater heterogeneity between attentive and inattentive firms. The resulting estimate for ω_H is 0.62 per nat, representing 8% of steady-state firm values under perfect information.

The average response to monetary shocks is an untargeted moment. Our model produces an average return semi-elasticity of 6.44, as estimated with the coefficient β_ν from $r_{it} = c + \beta_\nu \nu_t + \varepsilon_{it}$, which is consistent with our empirical estimates from Column (1) in Table 3 as well as from Gürkaynak et al. (2005).

6.3. Attention and the efficacy of monetary policy

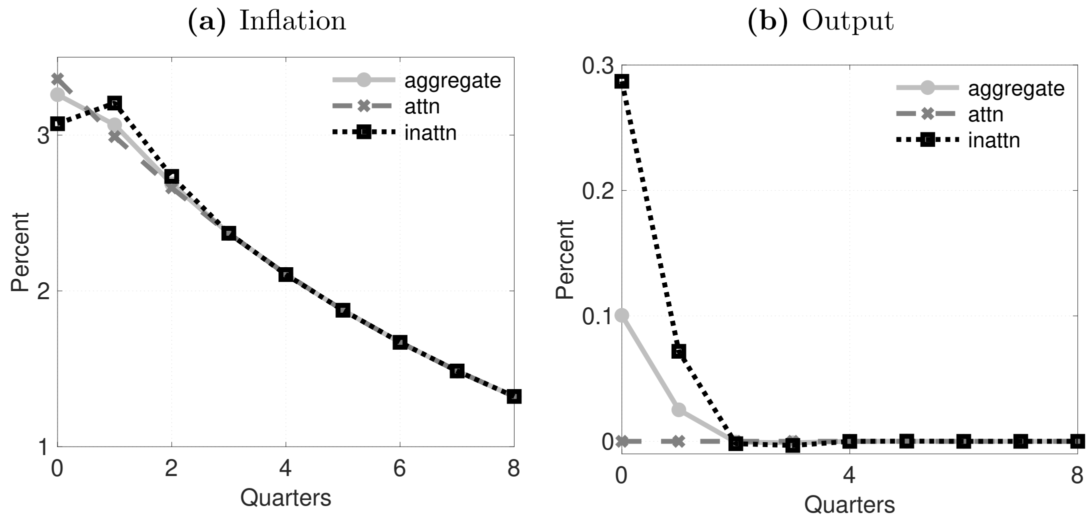
Figure 6 plots the aggregate responses to a one standard deviation expansionary shock to nominal aggregate demand growth. Panel (a) shows that inattentive firms under-adjust prices, reflecting partial incorporation of noisy signals about demand. Attentive firms track aggregate demand better than inattentive firms and exhibit more responsive prices.[18]

Panel (b) highlights that inattentive firms are responsible for increased output following an expansionary shock. These firms mischaracterize the nominal demand shock as a real shock and respond by raising output, while attentive firms correctly identify the nominal

[17] To connect Shannon with the text-based attention measure, we assume that the frequency of macro keywords in 10Ks is strictly increasing in firm attention, which allows us to match the cross-sectional distribution of firm attention without explicitly modelling the writing process of 10Ks.

[18] Appendix Figure A.5 shows individual firms' impulse responses for prices, profits, attention, and stock returns (including full-information returns, imperfect-information returns, and total returns) in response to both expansionary and contractionary monetary shocks.

Figure 6: Aggregate responses to expansionary monetary shock



Notes: Impulse responses of inflation and output. 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.

nature of the shock and respond by raising prices.[19]

A key implication of Panel (b) is that the aggregate output response to monetary policy increases with the share of inattentive firms. To illustrate the quantitative scope of the effect, we exogenously vary the fraction of attentive firms and compare output responses in our baseline calibration against two alternatives, $\check{\theta} = 56\%$ and $\hat{\theta} = 73\%$, which correspond to the minimum and maximum fraction of attentive firms over the sample period.

Table 7 reports the aggregate responses to monetary policies change as the fraction of attentive firms in the economy changes. The response of output growth to monetary policy is 7 basis points weaker in the most attentive calibration compared to the least attentive calibration. This suggests that expansionary policy in the depth of a recession when more firms are paying attention will be weaker than a preemptive interest rate (i.e. leaning against the wind) when aggregate attention is lower. This pattern is consistent with existing studies on the state dependency of monetary policy (e.g., Tenreyro and Thwaites, 2016). Similarly, monetary tightening imposes a smaller contractionary effect on output when more firms are

[19] To compare the magnitude of aggregate impulse responses to standard benchmarks, we convert the nominal aggregate demand shock in our inattention model to a nominal interest rate shock used in Christiano, Eichenbaum and Evans (2005). The passthrough of interest rate on the nominal aggregate demand is estimated in Appendix E.5. In response to a 25 basis point interest rate cut, output increases by around 0.1% on impact, roughly in line with the estimate in Christiano et al. (2005).

attentive to monetary news, which highlights the importance of clear central-bank communication (stressed, for example, in Haldane, Macaulay and McMahon, 2021).

Table 7: Attention and monetary non-neutrality

	Least attentive	Baseline	Most attentive
Fraction of attentive firms (θ)	56%	65%	73%
Output response	0.12%	0.09%	0.8%

Notes: Dependence of output responses on the fraction of attentive firms in the economy. Output responses are calculated as percent deviation from the steady state in response to a 25 basis point rate cut. Calibration for the least and most attentive economy is described in the main text.

6.4. Robustness

For robustness, Appendix E.6 implements an alternative calibration strategy that uses industry-level price adjustment estimates from Figure 2. It finds that attention remains quantitatively important and that the efficacy of monetary policy weakens under higher rates of attention.

7. 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.

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.

Basu, Susanto and John G Fernald, "Returns to cale in US production: Estimates and implications," *Journal of Political Economy*, 1997, *105* (2), 249–283.

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

Bloom, Nicholas and John Van Reenen, "Why do Management Practices Differ Across Firms and Countries?," *Journal of Economic Perspectives*, 2010, *24* (1), 203–24.

Bloom, Nick, Stephen Bond, and John Van Reenen, "Uncertainty and investment dynamics," *Review of Economic Studies*, 2007, *74* (2), 391–415.

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.

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 Swanson, "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*, 2005.

Haldane, Andrew, Alistair Macaulay, and Michael McMahon, "The Three E's of Central-Bank Communication with the Public," 2021.

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.

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.

Leahy, John V. and Toni M. Whited, "The Effect of Uncertainty on Investment: Some Stylized Facts," *Journal of Money, Credit and Banking*, 1996, *28* (1), 64–83.

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

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

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, "Learning asymmetries in real business cycles," *Journal of monetary Economics*, 2006, *53* (4), 753–772.

_ **and** _ , "Information Immobility and the Home Bias Puzzle," *The Journal of Finance*, 2009, *64* (3), 1187–1215.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

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.

Wooldridge, Jeffrey M, "On estimating firm-level production functions using proxy variables to control for unobservables," *Economics letters*, 2009, *104* (3), 112–114.

Yang, Choongryul, "Rational inattention, menu costs, and multi-product firms: Micro evidence and aggregate implications," *Journal of Monetary Economics*, 2022, *128*, 105–123.

Appendices

A. Additional Tables and Figures

Table A.1 contains the list of keywords used in frequency search under each topic. The keywords are based on `Econoday`, which provides notifications for major economic news and is the service behind `Bloomberg` economic calendar.

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 contains the summary statistics of firm characteristics by attention. In this table, a firm is attentive if its prevalence attention to the general topic is nonzero in any year in the sample period.

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: In this table, 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.

Table A.3: Controlling for alternative explanations of asymmetry (excl. ZLB)

Control Variable:	**Productivity**	**Mgmt Quality**	**Reported Profit**	**Filing Length**
Shock $\times \mathbb{1}_{v_t>0}$	8.54***	2.13	6.01**	3.01
	(2.79)	(2.82)	(2.65)	(5.33)
Shock $\times \mathbb{1}_{v_t<0}$	1.33	-8.05*	-1.52	27.23*
	(3.89)	(4.52)	(3.60)	(13.93)
Attention	-0.09	-0.05	-0.03	-0.04
	(0.10)	(0.07)	(0.06)	(0.06)
Attn $\times$ Shock $\times \mathbb{1}_{v_t>0}$	2.38	2.35***	1.55**	1.51**
	(1.52)	(0.85)	(0.72)	(0.64)
Attn $\times$ Shock $\times \mathbb{1}_{v_t<0}$	-5.85***	-8.21**	-5.75*	-5.09*
	(1.69)	(3.42)	(3.26)	(3.06)
Control Var	0.04***	-0.05	0.89***	-0.00
	(0.01)	(0.07)	(0.16)	(0.03)
Control $\times$ Shock $\times \mathbb{1}_{v_t>0}$	0.02	1.06	-3.70*	0.38
	(0.13)	(0.67)	(1.91)	(0.34)
Control $\times$ Shock $\times \mathbb{1}_{v_t<0}$	-0.14	-3.01	0.54	-2.97*
	(0.20)	(2.57)	(4.59)	(1.54)
Observations	280381	192900	431724	432458
R^2	0.038	0.032	0.027	0.027
Clustered SE	yes	yes	yes	yes
Firm controls	yes	yes	yes	yes
4-digit NAICS FE	yes	yes	yes	yes
excl. ZLB	yes	yes	yes	yes
Wald Test p-value: Attention	**0.001**	**0.005**	**0.048**	**0.060**
Wald Test p-value: Control	0.294	0.173	0.460	0.035

Notes: This table augments Column 3 of Table 3 to control for four potential confounding sources of asymmetry. Estimates are for the sample up to 2007 to exclude zero-lower–bound periods. The estimated regression has the form,

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

where c_{it} represents the alternative "Control" variable. As with attention, the control variable is interacted with both positive and negative monetary shocks. All other features of the model specification remain unchanged from Table 3. The four control variables considered are: (1) firm productivity estimated as in Olley and Pakes (1996); (2) management quality approximated with board member educational attainment; (3) profit measured as earnings before extraordinary items over total assets; and (4) filing length measured as the log word count of the 10-K filing. The final two rows report p-values of Wald tests for $H_0 : \beta_{d\nu_+} = \beta_{d\nu_-}$ and $H_0 : \beta_{c\nu_+} = \beta_{c\nu_-}$, respectively.

B. Additional Robustness

This appendix includes additional robustness tests of our main results in Table 3. We show that these results are robust to controlling for firm-specific exposure to monetary shocks, the information effect of high frequency shocks, and potentially confounding business cycle fluctuations.

B.1. Results robust to controlling for monetary exposure

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

To measure a firm's exposure to the monetary policy at date τ, we estimate the sensitivity of its stock prices to prior FOMC announcements over a 5-year rolling window using $t \in [\tau - 1826, \tau)$:

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

where v_t is the high-frequency monetary shock, and r_{it} is the close-to-close returns of firm i at date t. We also estimate sensitivity while controlling for the market factor (r^M) and Fama-French 3 factors (r^M, SML, and HML) using daily data on factors from Kenneth French's website. Exposure is defined as the absolute value of estimated sensitivity,

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

Table A.4 presents our two interaction coefficients of interest after controlling for exposure, θ_{it}^{λ}. The Wald tests for our null hypothesis, $\beta_{dv+} = \beta_{dv-}$, remains rejected at 5% for all three exposure measures. This confirms that our results are not driven by firms' exposure to monetary policy.

Table A.4: 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 (7) with additional controls for monetary exposure, θ_{it}^{λ}, $\lambda \in \{\text{baseline, CAPM, FF3}\}$. Standard errors are clustered at the shock level and reported in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

B.2. 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.5 show that our main results are robust to controlling for Greenbook forecast revisions.

Table A.5: 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 (7) with additional controls for Greenbook forecast revisions. Column (1) displays the baseline results from Table 3. Columns (2) - (4) adds 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.3. Results robust to controlling for macro fluctuations

While high-frequency monetary shocks v_t are considered exogenous, we conduct additional 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.6 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.6: 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 (7) with an additional vector of macro control Z_{t-1}, where Z_{t-1} include 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.

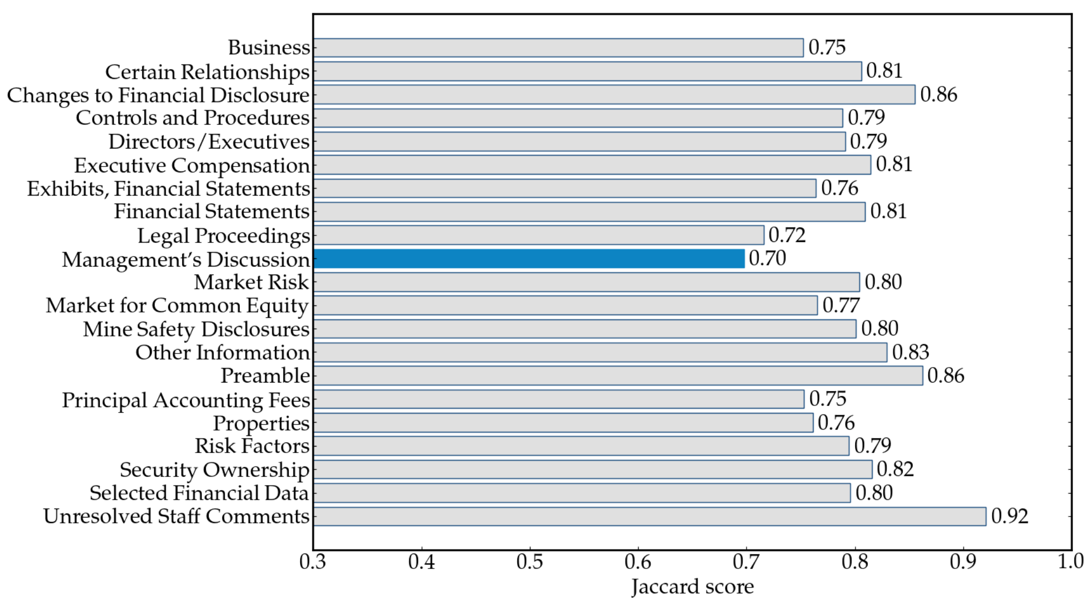
C.1. Lexical similarity

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

Figure A.1: 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.

Table A.7: 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 (7), restricting to 10-K items that discuss firm operations (Items 1 and 7). Standard errors are in parentheses. * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$).

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.7.

C.2. LDA: context of macro discussions

To enable automated context detection, we use the Latent Dirichlet Allocation (LDA) model to uncover topics firms tend to discuss in conjunction with macro news. LDA (Blei, Ng and

Jordan, 2003) is an unsupervised learning algorithm aimed at grouping words in documents into meaningful topics. We apply LDA to texts in earning filings within 20 words surrounding a macroeconomic keyword and set the number of topics to be 10.

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

We model topics surrounding each of the nine macro categories for the attention measure, as well as an aggregate category containing keywords from all categories. Figures A.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 represent a topic clustered by LDA, and the darkness of the cell within a topic represent the likelihood of a word to appear in the topic. Figure A.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 into their daily operations; Topic 2 relates to demand, as firms track and gauge the aggregate demand; Topic 6 relate to financing costs, as firms pay attention to how monetary policy affect their financial costs, investment decisions, and portfolio holdings; Topic 10 relates to labor costs, as firms assess the tightness of the labor market. Rest of the topics relate to housing, currency, and risk factors.

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

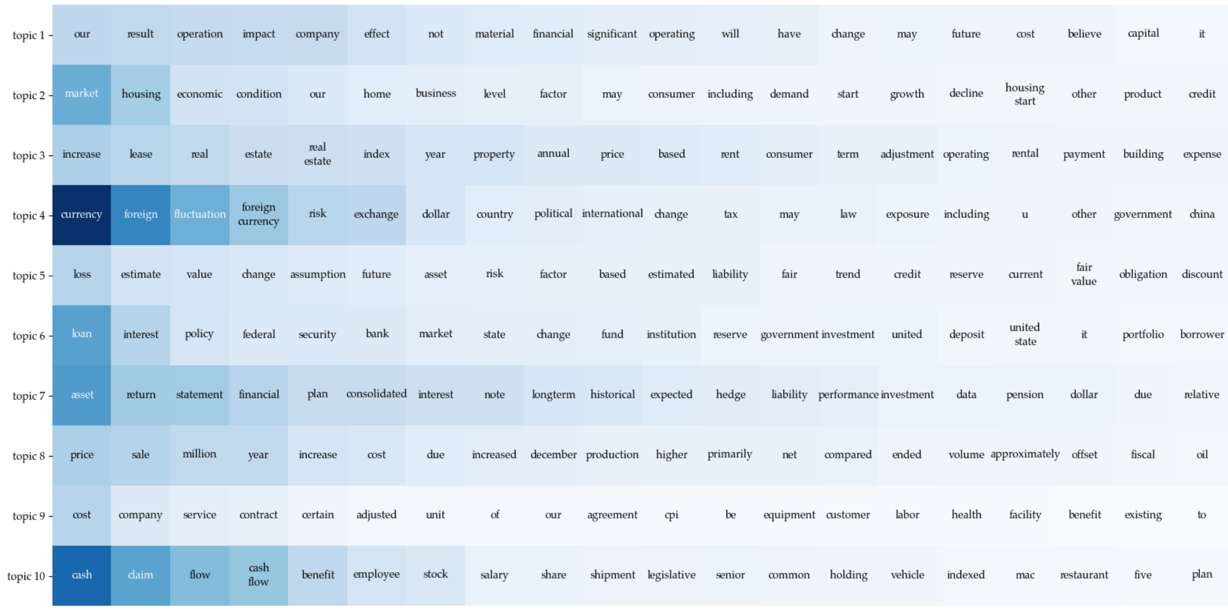
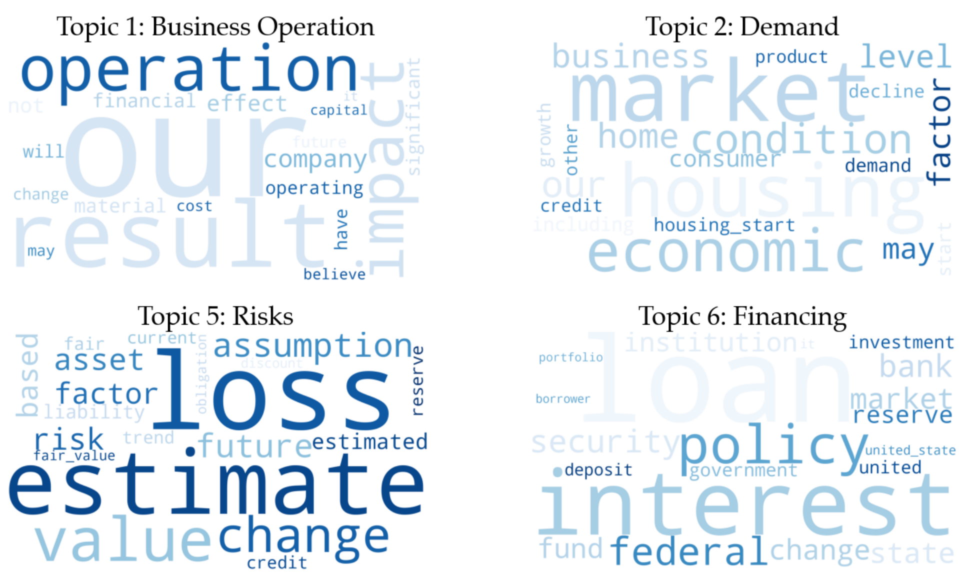


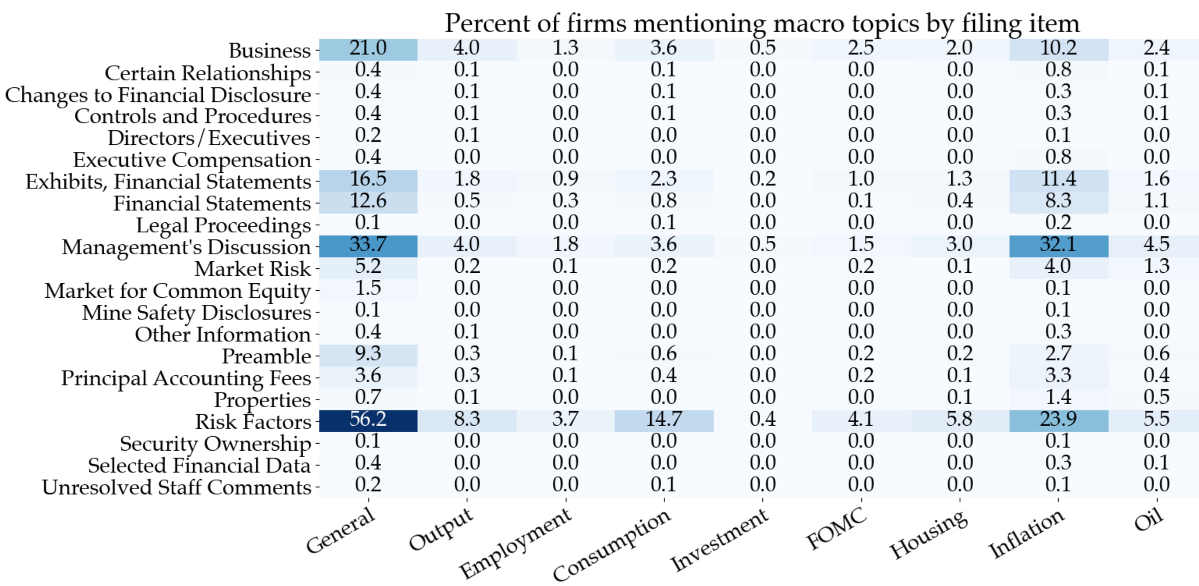
Figure A.3: LDA output for texts surrounding all macro keywords: Selected topics



C.3. Itemized frequency search

10-K filings have standard formats and are organized in sections. We perform refined frequency counts for each of the section, or "items", to see where attention is concentrated in. Results of frequency counts of macroeconomic keywords by filing item are shown in Figure A.4 in the Appendix. Discussions of the macroeconomy are concentrated in Description of Business (Item 1), Risk Factors (Item 1A) and Management Discussion and Analysis of Financial Condition and Results of Operations (Item 7A).

Figure A.4: 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.4 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.

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 added and subtracted $\frac{1}{2}\pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\hat{s}_t^2$ to complete squares and used the fact that $\bar{a} = \bar{s}$ in the steady state. The resulting expression is equation (4).

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 (4) we can derive the return elasticity with respect to the aggregate shock to be:

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

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

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

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

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

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

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

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

which is also independent of the magnitude of ε.

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

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

$$\frac{\partial r_j}{\partial \varepsilon} = \pi_s(\bar{s}, \bar{a})\bar{s} + \left(\pi_{ss}(\bar{s}, \bar{a})\bar{s}^2 - \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2\right)\varepsilon + \pi_{aa}(\bar{s}, \bar{a})\bar{a}^2(f_j(\varepsilon) - \varepsilon)(f_j'(\varepsilon) - 1) \tag{16}$$

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 $f(0) = 0$, (15) and (16) 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 $\varepsilon_t < f_j(\varepsilon_t) < 0$,

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

■

D.3. Model with time-varying uncertainty

We provide a framework to illustrate the effects of attention on firm profits when macroeconomic uncertainty is time-varying. This gives rise to the test we perform in Table 5.

Environment The aggregate state variable, y_t, follows an autoregressive process:

$$y_t = \rho y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \nu_t^2)$$

where ν_t^2 denotes the volatility of the aggregate state and is time-varying.

The firm's objective is to track the state variable as closely as possible and sets its prices, x_t, accordingly. The loss function is given by

$$\mathcal{L} = (x_t - y_t)^2$$

To track the macroeconomy, the firm chooses a noisy signal centered around the true state

$$s_t = y_t + u_t, \quad u_t \sim N(0, \tau_t^2)$$

Following Sims (2003), the level of noise contained in signal implies the flow of information of

$$\kappa_t = \frac{1}{2}\log_2 \frac{\tau_t^2 + \nu_t^2}{\tau_t^2}$$

Firms are constrained in its cognitive bandwidth to process information

$$\kappa_t \leq \kappa$$

Optimization Given this set up, the price a firm sets given the signal is

$$\begin{aligned} x_t = \mathbb{E}[y_t | y_{t-1}, s_t] &= \frac{\tau_t^2}{\tau_t^2 + \nu_t^2}\rho y_{t-1} + \frac{\nu_t^2}{\tau_t^2 + \nu_t^2} \\ &= 2^{-2\kappa_t}\rho y_{t-1} + (1 - 2^{-2\kappa_t})(y_t + u_t) \end{aligned}$$

A firm's chooses the level of information to obtain in order to minimize the expected loss

$$\begin{aligned} \min_{\kappa_t} \mathbb{E}\,\mathcal{L} &= \mathbb{E}[2^{-2\kappa_t}\varepsilon_t - (1 - 2^{-2\kappa_t})u_t]^2 \\ &= \text{Var}\,[2^{-2\kappa_t}\varepsilon_t - (1 - 2^{-2\kappa_t})u_t] \\ &= 2^{-2\kappa_t}\nu_t^2 \end{aligned}$$

subject to its bandwidth constraint

$$\kappa_t \leq \kappa$$

Therefore, the firm's realized loss is

$$\mathcal{L} = 2^{-2\kappa}\nu_t^2$$

As the economy becomes more uncertain, firms' loss increases

$$\frac{\partial \mathcal{L}}{\partial \nu_t^2} = 2^{-2\kappa} > 0$$

However, attentive firms suffer a smaller loss compared to inattentive firms as uncertainty rises

$$\frac{\partial \frac{\partial \mathcal{L}}{\partial \nu_t^2}}{\partial \kappa} = -(2\log 2)2^{-2\kappa} < 0$$

E. Additional Details for the Quantitative Model

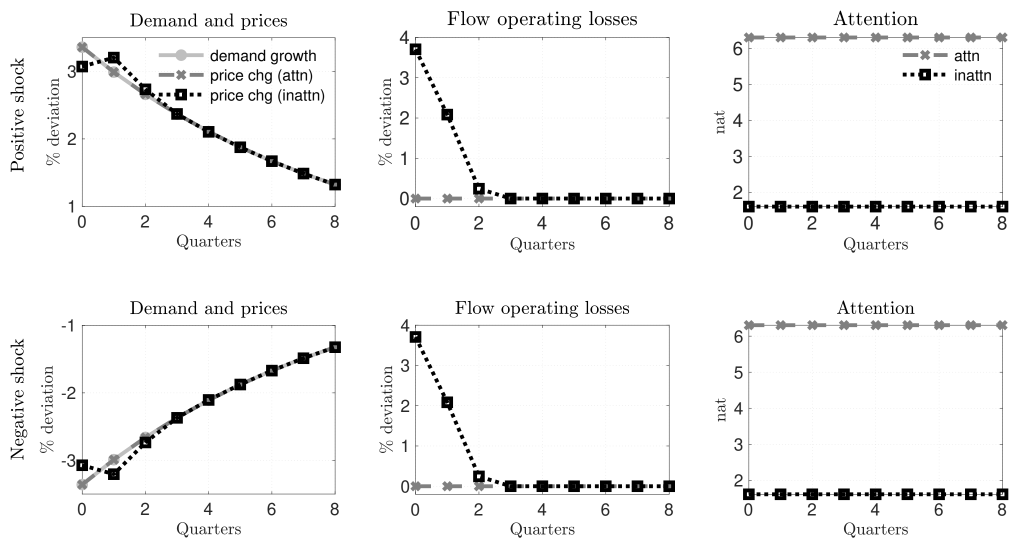
E.1. Firm impulse responses

Figure A.5 shows the impulse responses of individual firms to monetary shocks of one standard deviation. Panel (a) shows that as nominal aggregate demand rises, inattentive firms under-adjust prices, reflecting partial incorporation of noisy signals about demand. Attentive firms track aggregate demand better than inattentive firms and exhibit more responsive prices. Because of the imprecise information about the aggregate demand, inattentive firms experience greater losses in flow profits from the full-information benchmark, in response to both expansionary and contractionary shocks. With a constant marginal cost of information, firms' equilibrium choice of attention is not time-varying. Even though inattentive firms pay less attention, the higher marginal costs they face result in a higher total information costs.

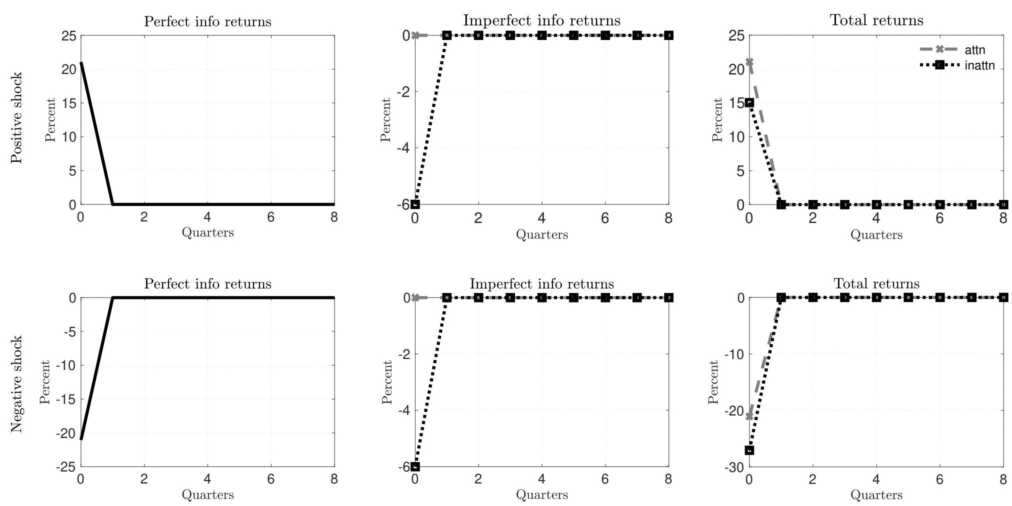
Panel (b) shows the responses of stock returns. Following an expansionary monetary shock, full-information equity returns of both attentive and inattentive firms increase, since firms are monopolistically competitive and have decreasing returns to scale. Returns of attentive firms increase by more than 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.

Figure A.5: 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.

E.2. 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.3. Numerical solution

We solve the rational inattention problem based on the `DRIPs` algorithm developed by Afrouzi and Yang (2021). Under log-quadratic approximation, the firm's problem is given by

$$\begin{aligned}
p_{it}^* &= \alpha p_t + (1-\alpha) q_t \\
\Delta q_t &= \rho \Delta q_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2) \\
p_t &= \int_0^1 p_{it} di.
\end{aligned}$$

Differencing out the unit root allows us to obtain the Wold representation of p_{it}^* as

$$p_{it}^* = (1-L)\Phi(L)\tilde{\nu}_t, \quad \tilde{\nu}_t = (1-L)^{-1}\nu_t = \sum_{j=0}^{\infty} \nu_{t-j},$$

where $\Phi(L)$ is the lag operator.

We specify the length of truncation to be $L = 40$ and define $x_t = (\tilde{\nu}_t, \tilde{\nu}_{t-1}, \cdots, \tilde{\nu}_{t-(L+1)})$. Then, the state-space representation of the system is given by

$$\begin{aligned}
x_t &= \mathbf{A} x_{t-1} + \mathbf{Q}\nu_t \\
q_t &= \mathbf{H}_q' x_t \\
p_{it}^* &= \mathbf{H}' x_t
\end{aligned}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \sigma_\nu \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \text{ and } \quad \mathbf{H}_q = \begin{bmatrix} 1 \\ \rho \\ \rho^2 \\ \vdots \\ \rho^{L-1} \end{bmatrix}$$

To solve for $\mathbf{H}$, we proceed as follows. In the n-th iteration, we start with the guess from the previous iteration, $\mathbf{H}_{k,(n-1)}$, where the subscript $k \in \{l, h\}$ indexes firms with low and high marginal cost of information, respectively. Then, we solve the rational inattention problem and obtain an updated guess. The optimal price is given by

$$p_t^* = \alpha p_t + (1-\alpha) q_t = (1-\alpha)\sum_{j=0}^{\infty} \alpha^j q_t^{(j)} = (1-\alpha)\sum_{j=0}^{\infty} \alpha^j (\theta q_{lt}^{(j)} + (1-\theta) q_{ht}^{(j)}),$$

where $q_{kt}^{(j)}$ is the j-th order belief of type-k firms on average, and $k \in \{l, h\}$. Now we guess and verify the expression for $q_{kt}^{(j)}$. Suppose there exists a matrix $\mathbf{X_{kj}}$ such that $q_{kt}^{(j)} = \mathbf{H}_q' \mathbf{X}_{kj} x_t$. Then, we can solve $q_t^{(j+1)}$ forward as

$$\begin{aligned} q_t^{(j+1)} &= \int_{\theta} \mathbb{E}_{it,l}\, q_{lt}^{(j)} di + \int_{(1-\theta)} \mathbb{E}_{it,h}\, q_{ht}^{(j)} di \\ &= \mathbf{H}_q' \big(\theta \mathbf{X}_{lj} \mathbf{X}_{l,(n)} x_t + (1-\theta) \mathbf{X}_{hj} \mathbf{X}_{h,(n)}\big) x_t, \end{aligned}$$

where $\mathbf{X}_{k,(n)} = \sum_{j=0}^{\infty} ((\mathbf{I} - \mathbf{K}_{k,(n)} \mathbf{Y}_{k,(n)}') \mathbf{A})^j \mathbf{K}_{k,(n)} \mathbf{Y}_{k,(n)}' \mathbf{M}'^j$. Matrices $\mathbf{K}$ and $\mathbf{Y}$ are Kalman gains and loadings of optimal signals solved from the rational inattention problem and are as specified in Afrouzi and Yang (2021); and $\mathbf{M}$ is a shift matrix.

Setting $\mathbf{X}_{kj} = \mathbf{X}_{k,(n)}^j$ for all j implies that

$$q_{kt}^{(j)} = \mathbf{H}_q' \mathbf{X}_{k,(n)}^j x_t,$$

which verifies the guess for $q_{kt}^{(j)}$.

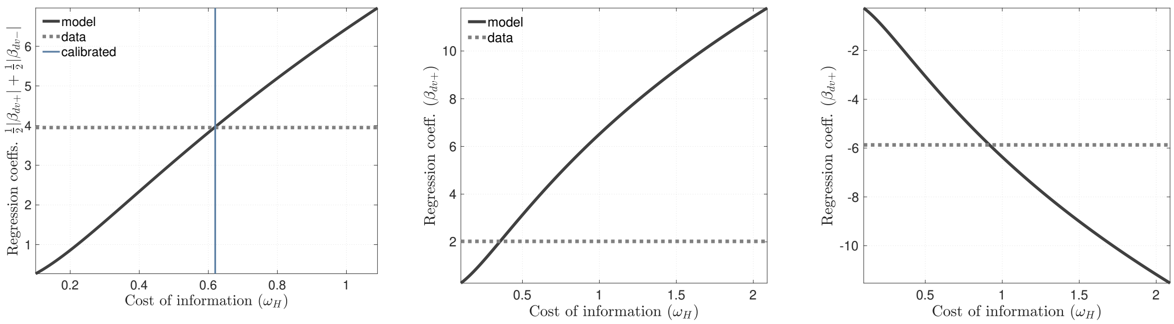
We therefore obtain the updated guess $\mathbf{H}_{k,(n)} = (1-\alpha) \mathbf{X}_{k,p,(n)}' \mathbf{H}_q$, where $\mathbf{X}_{k,p,(n)} = \sum_{j=0}^{\infty} \alpha^j \mathbf{X}_{(n)}^j$, and iterate until convergence.

E.4. Details for model calibration

Figure A.6 provides details on how calibrated parameters are identified. In Panel (a), we simulate the model for a range of ω_H. As ω_H increases and the gap between ω_L and ω_H widens, the simulated elasticity monotonically increases, implying greater heterogeneity between attentive and inattentive firms. In Panel (b) we report the sensitivity of ω_L and ω_H separately. As ω_H increases, the magnitude of both semi-elasticities increases monotonically.

Figure A.6: Sensitivity of simulated moments to calibrated parameters

(a) Heterogeneous responses **(b)** Semi-elasticities



Notes: Simulated moments for a range of parametrization for labor share γ and cost of information ω_H. We simulate models for a panel of 100 firms and for 1000 periods with 100 periods burn-ins. Simulated moments are generated with regressions discussed in the text. Panel (a) shows the sensitivity of average absolute values of $\beta_{d\nu+}$ and $\beta_{d\nu-}$ to changes in parameter values of ω_H; Panel (b) shows the sensitivity of $\beta_{d\nu+}$ and $\beta_{d\nu-}$ separately.

E.5. Passthrough regressions

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 $\Delta_h y_{t-1,t+h}$ is average percent changes in the variable of interest, and $\varepsilon_t^{\mathtt{i}}$ is the high-frequency monetary policy shock. The dependent variables are U.S. manufacturing output over the sample period of 1994 to 2019. We estimate the responses of manufacturing prices, real output and nominal output. Path of β_h informs the average cumulative changes in the dependent variable in response to the interest rate shock.

Figure A.7: Passthrough of rates to nominal demand

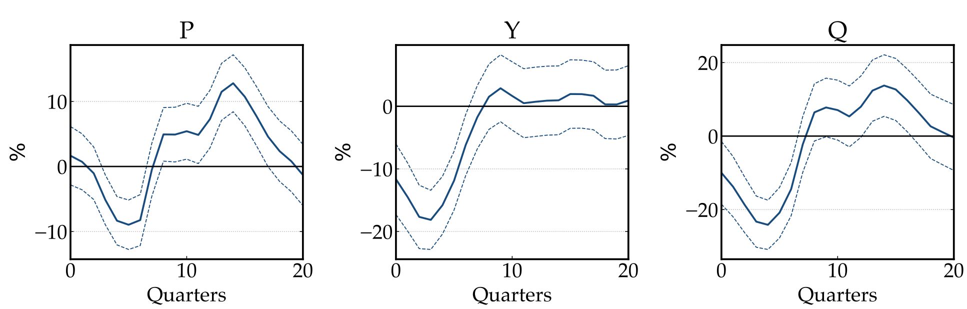


Figure A.7 shows the results of the local projection. A 25 basis point point expansionary shock to the interest rate leads to about 3.3 percent peak increase in nominal demand.

E.6. Alternative calibration with price adjustment

In this section, we provide an alternative calibration strategy for the quantitative model. We calibrate the model to match the speed of price adjustment rather than the results of Section 4 as in the baseline case.

Table A.8: Calibration with price adjustment

Parameter	Description	Value
Assigned parameters		
β	discount rate	$0.96^{1/4}$
ρ	shock persistence	0.89
σ_ν	shock std. dev.	4.23×10^{-2}
ε	elasticity of substitution	7
ψ	disutility of labor	0.86
γ	returns to scale	0.83
Information parameters		
θ	fraction of attentive firms	65%
ω_L	cost of information	58
ω_H	cost of information	53×10^3

We calibrate the model quarterly. The unit of analysis, i, represents a 4-digit NAICS sub-industry within manufacturing[20]. Table A.8 shows the parameter values. Assigned parameters that are unrelated to information frictions follow the baseline calibration in Section 6.2, with the exceptions of the elasticity of substitution. We set $\varepsilon = 7$, instead of $\varepsilon = 10$, to capture a lower elasticity of substitution across industries than across firms, following Gorodnichenko and Weber (2016).

We then calibrate costs of information, ω_L and ω_H, to target industry inflation responses to monetary shocks. To obtain empirical targets, we re-estimate Equation (1) for manufac-

[20]For example, NAICS 3331 represents "Agricultural, construction and mining machinery manufacturing", and NAICS 3332 represents "industrial machinery manufacturing".

Table A.9: Targeted moments

Targeted Moment	Data	Model
attentive price semi-elasticity to monetary shocks ($\hat{\beta}_\nu + \hat{\beta}_{d\nu}$)	8.79	8.79
inattentive price semi-elasticity to monetary shocks ($\hat{\beta}_\nu$)	3.41	3.41

turing sectors at quarterly frequency using

$$\Delta \log P_{s,t} = \alpha_s + \alpha_t + \beta_\nu \nu_t^{\mathrm{M}} + \beta_d d_{st} + \beta_{d\nu} d_{st} \nu_t^{\mathrm{M}} + \Gamma' Z_t + \varepsilon_{st}$$

where $P_{s,t}$ is the PPI of industry s (4-digit NAICS) in quarter t; ν_t is monetary shocks; d_{st} is sector average prevalence attention; Z_t and $\{\alpha_s, \alpha_t\}$ are our standard controls and fixed effects.

Table A.9 shows that in response to a 100 basis point expansionary monetary shock, attentive sectors raise prices by 8.8% in the first quarter, while inattentive sectors raise prices by 3.4%.

To match these empirical targets, we simulate the model for a range of ω values and obtain the impulse responses to a monetary shock equivalent to a 100 basis point rate cut. We set ω_L so that the simulated inflation of attentive industries in response to the monetary shock matches $\hat{\beta}_\nu + \hat{\beta}_{d\nu}$, the observed semi-elasticity of attentive industries. Similarly, we set ω_H so that the simulated inflation responses of inattentive industries matches $\hat{\beta}_\nu$, the observed semi-elasticity of inattentive industries. Table A.9 shows that the calibrated model matches the semi-elasticities of industry inflation in response to monetary shocks.

As in the baseline model, we quantify the importance of attention on the efficacy of monetary policy by varying the fraction of attentive firms. Table A.10 shows the response of output growth is 23 basis points weaker in the most attentive calibration compared to the least attentive calibration, which shows a quantitative importance consistent with the baseline calibration.

Table A.10: Attention and monetary non-neutrality

	Least attentive	Baseline	Most attentive
Fraction of attentive firms (θ)	56%	65%	73%
Output response	1.65%	1.53%	1.42%

Notes: Dependence of output responses on the fraction of attentive firms in the economy. Output responses are calculated as percent deviation from the steady state in response to a monetary shock equivalent to a 25 basis point rate cut. Calibration for the least and most attentive economy is described in Section 6.3 in the main text.