Granular Sentiments[†]

Rustam Jamilov **Alexandre Kohlhas** **Oleksandr Talavera** **Mao Zhang**

February 2025

Abstract

We propose an empirically-consistent theory of business cycles driven by fluctuations in the sentiment of a small number of firms. Using computational linguistics and analyst forecasts, we find that 50 firms account for over 60% of the variation in U.S. sentiment and macroeconomic outcomes. Our "Granular Sentiment Index", which measures the sentiment of these firms, is dominated by downstream firms that are close to the final consumer. Incorporating endogenous attention choice into a general equilibrium model with heterogeneous firms, we show that this heterogeneity arises because downstream firms act as natural "information agglomerators". A calibrated version of the model shows that sentiment shocks to the 20% most downstream firms explain 70% of sentiment-driven (and 20% of aggregate) fluctuations.

JEL Classification Codes: E32, D22, D83.

Keywords: business cycles, beliefs, incomplete information, granular hypothesis.

[†]For comments and useful suggestions, we thank Vladimir Asriyan, Florin Bilbiie, Saleem Bahaj, Vasco Carvalho, Ryan Chahrour, Wei Cui, Matthew Elliott, Cécile Fabre, Andrea Ferrero, Joel Flynn, Xavier Gabaix, Jordi Galí, Mishel Ghassibe, Yuriy Gorodnichenko, Albert Marcet, Michael McMahon, Kurt Mitman, Kristoffer Nimark, Franck Portier, Lukasz Rachel, Morten Ravn, Federica Romei, Victoria Vanasco, Jaume Ventura as well as seminar and conference participants at various venues. We further thank Varun Sharma for sharing their newspaper-based sentiment data. This research received financial support from *Handelsbanken Stiftelsen*. First draft: February, 2024.

Jamilov: University of Oxford. Email: rustam.jamilov@all-souls.ox.ac.uk.
Kohlhas: University of Oxford. Email: alexandre.kohlhas@economics.ox.ac.uk.
Talavera: University of Birmingham. Email: o.talavera@bham.ac.uk.
Zhang: University of St Andrews. Email: mz71@st-andrews.ac.uk.

"The varying expectations of businessmen, and not anything else, constitute the immediate [...] antecedents of industrial fluctuations." (Pigou, 1927)

1 Introduction

Effective macroeconomic policy hinges on the accurate identification and measurement of the origins of economic fluctuations. There is reason to believe that these origins extend beyond shocks to economic fundamentals, such as productivity. Indeed, since Pigou (1927)'s treatise on the sources of "industrial fluctuations", a frequently-proposed alternative driver of business cycles has been the coordinated waves of optimism and pessimism that often characterize people's views about the economy.

Over the past two decades, many advances in business-cycle theory have attempted to quantify the macroeconomic importance of such fluctuations in sentiment (e.g., Beaudry and Portier 2004; Blanchard et al. 2013; Chahrour et al. 2021). Much of this work has identified mechanisms by which erroneous movements in optimism (and pessimism) about the future cause demand-induced fluctuations, similar to those triggered by other types of demand shocks. However, despite the proliferation of this work—and the quantitative importance of sentiment shocks for driving business cycles that it finds—little, if any, evidence exists on *which* economic agents drive changes in overall sentiment.

In this paper, we focus on firms and investigate which firms drive business sentiment and whether shocks to the sentiment of these firms drive business cycles. Our aim is threefold. First, we wish to establish whether there exists a set of firms whose average sentiment about the future serves as a sufficient statistic for the economy-wide effects of sentiment shocks—that is, whether the canonical "Granular Hypothesis" (Gabaix, 2011) applies to sentiment-driven fluctuations. Second, we want to understand the characteristics of the firms that belong to this set: which features lead firms to be central for economy-wide sentiment and why? And finally, third, we wish to trace out the macroeconomic consequences of sentiment shocks to these firms.

Our main contribution is to provide a first-pass answer to these questions. We argue that the average sentiment towards future performance of fewer than 50 firms can account for 60% of the unconditional variation in U.S. sentiment and output over the past two decades. The "Granular Sentiment" measure, which captures the average sentiment of these 50 firms, is dominated by firms that are close to the final consumer—i.e., that are *downstream* (but not unusually large). We rationalize our results within a standard multi-sector economy to which we add endogenous attention choice. We show that attention gravitates toward downstream firms, as they emerge as natural "information agglomerators", whose sentiment matters disproportionally for aggregate fluctuations. When calibrated to match moments of U.S. data, the model shows that the 20 percent most downstream firms account for close to 70 percent of sentiment-driven output fluctuations.

Empirics: We begin our analysis by computing measures of firm-level sentiment towards future performance. To do so, we apply standard tools from computational linguistics to quarterly earnings calls of firms publicly listed in the U.S.. A significant fraction of each earnings call is a Q&A session between firm managers and industry analysts, which triggers unscripted conversations discussing financial results and beliefs about future performance. This contrasts with the often heavily formalized and scripted 10-K reports that listed firms also release. Using a well-established corpus of sentiment keywords from Loughran and McDonald (2011), we work through the transcript of every earnings call of every firm from every quarter to construct an index of firm-level sentiment about the future. We complement this baseline index with two auxiliary measures of firm sentiment. The first uses the FinBERT language model (Devlin et al., 2018) to create a finer text-based measure of sentiment at the firm level; the second, instead, uses financial analysts' forecasts to proxy changes in firm sentiment following each call.

Our first main result shows that the average sentiment of around 50 firms can account for over half of the variation in U.S. sentiment and output. Crucially, we document a non-monotone relationship between "firm informativeness"—as proxied by either the coefficient of determination (R^2) or the beta-coefficient from firm-level regressions of sentiment on macroeconomic outcomes—and the share of output fluctuations accountable. Initially, as we add firms into the average, the marginal informational benefit outweighs the additional firm-specific noise. However, beyond the roughly 50th most informative firm, the trade-off reverses. The "Granular Sentiment" index, which measures the average sentiment of these 50 firms, captures around 60% of the variation in economy-wide sentiment and output. Indeed, we find that the sentiment of these 50 firms is more informative than the naive average of the sentiment of all firms.

We show that our results extend to cases in which we partial out macroeconomic and firm-specific controls, and to the two auxiliary measures of firm-level sentiment. Importantly, we also construct the intersection of the resulting sets of firms from all the robustness specifications that we consider and find that, on average, 60% of the firms in each set are the same as in the baseline. The sentiment of a small number of *core* firms accounts for a consistently and disproportionally large share of aggregate fluctuations.

Our second key result characterizes this set of most informative firms. We take an agnostic approach and construct an array of potential explanatory variables, which includes firm size, market beta, idiosyncratic return volatilty, leverage, and sector, among several others. We document that sectoral *downstreamness*, i.e. proximity to the end consumer, is the strongest predictor of being a member of our set. The remaining variables, including firm size, matter little for the probability of being a member. All else equal, firms whose sentiment is more closely correlated with business-cycle fluctuations are more downstream. We conduct a battery of robustness tests which show that this finding is resilient to issues related to identification, alternative measures and specifications, as well as the inclusion of firm-level and aggregate controls.

The third and final step of our empirical analysis studies the economy-wide consequences of *innovations* in downstream-sector firms' sentiment. We construct a weighted average of sentiment of firms in our granular set, using previous year's downstreamness as weights. Using local projections, we show that shocks to this index have strong dynamic effects on economy-wide sentiment, output, unemployment, uncertainty, and inflation, conditional on important controls such as aggregate productivity and monetary policy. The effects are, on balance, large, statistically significant, and akin to those from the aggregate "noise" shocks commonly studied in the literature (e.g., Beaudry and Portier 2004; Lorenzoni 2009). To address potential concerns related to identification, we document that our results extend to circumstances in which we instrument for sentiment following the approach in Lagerborg et al. (2022), and when we structurally identify "true" sentiment shocks as in Chahrour and Jurado (2021).

Model: We proceed to provide a theory as to *why* downstream-sector firms' sentiment towards the future drives fluctuations in sentiment and output. We depart from a standard flex-price, multi-sector economy in the spirit of Acemoglu et al. (2012) and Chahrour et al. (2021) to which we add endogenous attention choice. In the model, firms in both upstream and downstream sectors choose inputs under imperfect information about sector-specific productivity. Sentiment shocks are modelled as mistaken views about other firms' productivity. We show that, in equilibrium, firms across all sectors tend to pay closer attention to (i.e., acquire less noisy information about) downstream-sector conditions for two reasons. First, downstream-sector firms are, all else equal, more central in the economy's production network. This, in turn, causes fluctuations in downstream productivity to be more important for overall demand, and hence for the optimal input choice of an individual firm (e.g., Baqaee and Farhi 2019). Second, because downstream-sector firms combine inputs from several upstream providers, their output also better aggregates the dispersed information that exists about shocks in the economy. In this sense, downstream-sector firms act as natural "information agglomerators", similar to the role played by market-clearing prices in Grossman and Stiglitz (1980)'s classical analysis.

We use our macroeconomic framework to explore the business-cycle implications of sentiment shocks to downstream-sector firms. To do so, we calibrate the parameters that govern the sentiment process to those from our Granular Sentiment Index. We show that, for standard parameter values, sentiment shocks towards the most downstream sectors drive a sizeable share of overall output fluctuations. Our main quantitative exercise involves calibrating the model economy to the 27 U.S. industries in Atalay (2017), derived from the BEA production tables. In this extended version of our model, we document that orthogonal innovations to sentiment about the productivity of firms in 20 percent of industries—those that are the most downstream—account for around 70 percent of sentiment-driven fluctuations in output. Overall, we find that sentiment shocks to the most downstream sectors explain close to 20 percent of output fluctuations. Tantalizingly, our results are therefore close to the famous "Pareto principle", stipulating that 20 percent of agents (here, firm sectors) drive more than 80 percent of actions (here, sentiment-driven output fluctuations) (Pareto, 1896).

Literature: Our research relates to two long-standing ideas in macroeconomics: (i) "sentiment-driven business cycles" and (ii) "the granular hypothesis".

First, we build on the literature that quantifies sentiment-driven macroeconomic fluctuations, following Pigou (1927)'s initial contribution. Prominent studies, among many others, are Beaudry and Portier (2004, 2006), Lorenzoni (2009), Blanchard et al. (2013), Angeletos and La'O (2013), Angeletos et al. (2018), Chahrour and Jurado (2018), Angeletos et al. (2020), Kohlhas and Walther (2021), Enders et al. (2021), Lagerborg et al. (2022), Chahrour and Ulbricht (2023), Maxted (2023), and Flynn and Sastry (2022, 2024). In this paper, we emphasize the role of a subset of firms, those which are close to the final consumer, in driving aggregate fluctuations in sentiment and output. Our work is, as such, closely related to Chahrour et al. (2021), who study how newspaper reporting about specific sectors helps account for aggregate fluctuations. The contribution of our paper, in this context, is to highlight which firm-level characteristics drive sentiment-driven fluctuations across sectors, and to propose a theory consistent with this evidence.

Second, our work builds on the seminal contribution of Gabaix (2011), who introduced the notion of "the granular origins of aggregate fluctuations". In a burgeoning literature, the granular hypothesis has been applied to the case of productivity shocks (Carvalho and Gabaix, 2013), international trade and finance (di Giovanni et al., 2014; Gaubert and Itskhoki, 2021), banking and insurance (Chodorow-Reich et al., 2020; Galaasen et al., 2023), and nominal rigidities (Pastén et al., 2024). To the best of our knowledge, we are the first to explore the implications of the granular hypothesis for beliefs and sentiment-driven business cycles. Relatedly, our paper also contributes to the literature on the macroeconomics of networks, which emphasizes the role of the input-output linkages for aggregate dynamics (e.g., Acemoglu et al. 2012, 2015; Liu 2019; Bigio and La'O 2020).

Finally, on the empirical front, our paper measures sentiment using textual analysis. This follows an expanding literature that uses text as data (e.g., Hansen and McMahon 2016; Hansen et al. 2017; Gentzkow et al. 2019). We build on the literature that measures sentiment with the dictionary-based approach of Loughran and McDonald (2011, 2016), which has been recently popularized by Hassan et al. (2019).[1] We augment the dictionary-based approach with a more sophisticated, natural language processing technique that detects and measures sentiment using BERT embeddings (Devlin et al., 2018). We particularly focus on FinBERT, a BERT-based sentiment classification model that has been pre-trained on financial information (Araci, 2019). The BERT class of models has been applied to multiple areas of economic research, ranging from central banking (Gorodnichenko et al., 2023) and environmental economics (Chava et al., 2021) to technology and innovation (Chava et al., 2020). Lastly, we also measure firm-level sentiment using financial analysts' forecasts of firm performance, an approach that has been used extensively by, for example, Bordalo et al. (2021) and Asriyan and Kohlhas (2024).

2 Data and Empirical Strategy

We begin by describing the data used in our analysis and by providing a broad overview of our empirical strategy. The next section then characterizes the heterogeneity that exists in firm-level sentiment as well as the factors responsible for it. In doing so, we present new evidence on the drivers of firm-level sentiment and its macroeconomic consequences.

2.1 Data and Measurement

We construct measures of firm-level sentiment using a variety of sources. Our baseline measure exploits transcripts of quarterly *earnings calls* from S&P Capital IQ. To comply with regulatory requirements, and to promote transparent communication with the investment community, publicly listed firms in the U.S. are mandated to hold conference

[1] In the context of earnings calls, this approach has recently been successfully applied to questions such as ”climate change risk” (Sautner et al., 2023), ”Brexit” (Hassan et al., 2023b), ”country risk” (Hassan et al., 2023a), and ”cyber risk” (Jamilov et al., 2023).

calls with financial analysts in conjunction with their quarterly earnings release.[2] Our dataset includes transcripts of all such earnings calls dating back to 2006 for firms in the S&P 500 index.[3] Our baseline sample, as a result, includes over 23,000 observations from 619 unique firms spanning the period 2006q4-2021q4.[4]

We leverage tools from textual analysis and use a *dictionary-based term-counting method*—used also in, for example, Baker et al. (2016)—as our baseline measure of sentiment. As a dictionary for positive and negative sentiment keywords, we use the corpus in Loughran and McDonald (2011). We define net-sentiment for firm $i = \{1, 2, ...N\}$ at time t as the difference between the number of positive- and negative-sentiment terms in the firm's earnings call transcript, scaled by the overall length of the transcript:

$$\xi_{it}^{1} \equiv \frac{\sum_{b}^{B_{it}} \left(1[b \in \mathcal{C}_{P}]\right) - \sum_{b}^{B_{it}} \left(1[b \in \mathcal{C}_{N}]\right)}{B_{it}}, \tag{1}$$

where $1[\cdot]$ is an indicator function, $\mathcal{C}_P$ and $\mathcal{C}_N$ denote the set of terms that belong to the positive and negative sentiment dictionary, respectively, and B_{it} is the total number of words in the transcript.

We later contrast and compare our baseline indicator with two alternative measures of firm-level sentiment: the first, ξ_{it}^{2} , uses the so-called FinBERT *natural language algorithm* to refine the basic dictionary-based approach, while the latter, ξ_{it}^{3}, relies on *analysts' firm-level forecasts* following the earnings call from the I/B/E/S Guidance database. Notice that all three measures proxy firms' sentiment about future performance. As such, the measures are closely related to one another—as in Pigou (1927)'s original idea—and, as we will show, contain information about future performance that is independent of fundamental firm-level and economy-wide characteristics, such as productivity.

We merge our measures of firm-level sentiment with basic income statement and balance sheet data from the CRSP-Compustat database. Combined, this provides us with a comprehensive dataset of firm-level sentiment and associated firm-specific characteristics.

An important variable for the analysis that follows is a firm's degree of *sectoral upstreamness*, measuring the average distance of firms' output in a sector to the final consumer. We define sectoral upstreamness as in Antràs et al. (2012) and extend the measure to 2021. In

[2]An earnings (conference) call usually begins with a management presentation delivered by key executives, typically the CEO, CFO, and occasionally other senior managers, detailing the company's financial results and future outlook. This is followed by a question and answer (Q&A) session between the management (e.g., CEO, CFO, investor relation officer, COO) and invited financial analysts (or investors). Each call is usually 45 minutes long and contains around 7,000-8,000 spoken words.

[3]See also, for example, Bordalo et al. (2023) and the discussions therein.

[4]From 2010 onwards, we are also able to obtain most of the underlying audio files, from which the text files are transcribed. We manually cross-verify the consistency of each transcript text.

compact matrix notation, upstreamness is defined as:

$$\mathbf{U}_t = \left[\mathbf{I} - \Delta_t\right]^{-1}\mathbf{1}, \tag{2}$$

where $\mathbf{1}$ is a column vector of ones, $\mathbf{I}$ is the identity matrix, and Δ_t is a square matrix with the numerator of the (s,k)-th entry, Δ_{skt}, being equal to the dollar value of commodity s used in k's production in period t. The denominator of Δ_{skt}, in contrast, equals the total sum of values less what is recorded under net exports (and net changes in inventories). The s-th entry of $\mathbf{U_t}$, U_{st}, contains the upstreamness measure for sector s in quarter t. We note that U_{st} is bounded below by unity by construction. A key advantage of the upstreamness measure in (2) is that its economic interpretation is simple: values of U_{st} correspond to the number of transaction rounds necessary for the average product to reach the final consumer. For example, the apparel sector is one of the most downstream in our sample with $U_{s,2021} = 1.05$. This suggests that firms in this industry sell their goods and services almost exclusively to the final consumer. On the other hand, the petrochemicals sector averages an upstreamness value of greater than 3.00, suggesting that it takes this sector's goods more than three rounds of sales, on average, to reach the end consumer.

We obtain all the necessary data to construct the upstreamness measure from the detailed U.S. I-O Tables, provided by the Bureau of Economic Analysis (BEA). Because upstreamness is defined at the level of a sector, we assume that U_{st} is common for all firms $i \in s$. Our baseline analysis uses 3-digit BEA industry codes, although we later show that all our main results are robust to alternative sectoral definitions.

Appendix A.1 provides further details on variable definitions and construction. Table A.1 in the Appendix presents basic summary statistics for our three main sentiment measures as well as other variables used in the empirical analysis.

2.2 Outline of the Empirical Approach

Our empirical strategy consists of three broad steps. First, we ask whether there exists a small number of firms whose average sentiment about the future is closely associated with the overall state of the economy. We wish to understand whether there exists sizable heterogeneity in the relationship between firm-level sentiment and the state of the economy. To this end, we estimate firm-level relationships between measures of sentiment and various indicators of business-cycle fluctuations, and average across firms.

Second, having determined the existence of a small number of firms whose average sentiment is closely correlated with the state of the economy, we wish to understand the characteristics of firms in that set. The "granular hypothesis", as articulated in Gabaix

(2011) in the context of productivity shocks, assumes that *firm size* is the relevant characteristic. Firm-level productivity shocks do not wash out in the aggregate if they hit *large* firms. However, in the context of sentiment-driven fluctuations, it is not obvious that firm size is still the correct characteristic to focus on. We take an agnostic approach and test an array of potential explanatory variables, including but not limited to firm size, leverage, cyclicality, and sectoral downstreamness. Our aim is to find the defining characteristic that best summarizes our set of "granular sentiment" firms.

Third, we move beyond the cross-section and analyze the business-cycle consequences of *innovations* in the sentiment of firms summarized by our defining characteristics. Our objective is to explore whether sentiment towards the future of a small number of firms that we identify not only reflects the state of the economy but also helps drive it forward. In order to address potential concerns related to endogeneity and omitted variable bias, in this step, we also provide a battery of extensions and robustness tests. The results from this step will also later provide crucial guidance for the calibration of our quantitative model in Section 7.

3 The Cross-Section of Firm-Level Sentiment

In this section, we carry out the first two steps of our empirical strategy outlined above. We show that the average sentiment of a small number of firms can account for most of output fluctuations at the business-cycle frequency, and that these firms all share a common characteristic—sectoral downstreamness.

3.1 Heterogeneity in Firm-Level Sentiment

We start by establishing a ranking of the average sentiment of subsets of firms, measuring how closely correlated average firm-level sentiment is to the business cycle. To start, for every firm in our sample, we run the following linear regression:

$$Y_t = \alpha_i + \beta_i \xi^1_{it} + \varepsilon_{it}, \quad \forall i, \tag{3}$$

where Y_t is HP-filtered (Ravn and Uhlig, 2002) aggregate output, ξ^1_{it} is the baseline dictionary-based measure of firm-level sentiment, and ε_{it} is an error term. We collect the $N \times 1$ vector of coefficients of determination (R^2_i) and rank firms in descending order of R^2_i, denoting the rank integer by K_i. As such, the firm whose sentiment is most closely correlated with the state of the economy has a rank integer of $K_i = 1$, the firm whose sentiment is the second-most closely correlated has a rank integer of $K_i = 2$, and so on.[5] There are two reasons why R_i^2 is our preferred measure of the "*informativeness*" of the i'th firm. First, by design, it directly captures the share of output variation that can be accounted for by a particular firm in our sample. Second, the same metric can easily be computed in any model, which facilitates a simple and exact model-to-data comparison. As we discuss below and show in Appendix A.3, our results do not change if we use (the absolute value of) regression betas, $|\beta_i|$, instead of the R_i^2 as a metric that summarizes informativeness. Notice that any measure needs to allow for both pro-cyclical and counter-cyclical sentiment at the firm level.

We next construct N *portfolios* of sentiment, measuring the average sentiment of subsets of firms. We start from the 1st-ranked firm and proceed iteratively, such that the 2nd portfolio is the average sentiment of firms ranked first and second, the third portfolio is the average sentiment of firms ranked first, second, and third, and so on. Finally, we run regressions of output on portfolio-level sentiment:

$$Y_t = \alpha_p + \beta_p \xi_{pt} + \varepsilon_{pt}, \quad \forall p, \tag{4}$$

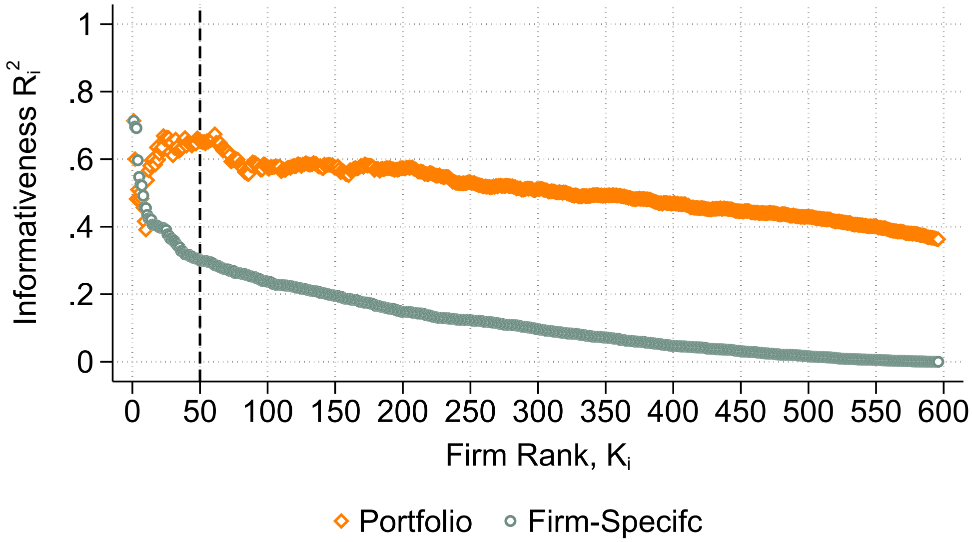
where ξ_{pt} denotes average sentiment of the p-th portfolio and ε_{pt} is an error term. Our object of interest is the portfolio-level R_p^2, which measures how closely correlated sentiment in the portfolio ranked K_i is with the business cycle—or put differently, how *informative* firms' average sentiment in the pth-portfolio is about the state of the economy.

Figure 1 presents the outcome of this exercise. The fact that firm-level informativeness, R_i^2, is monotonically decreasing in K_i is by construction. Notice that the relationship between portfolio-level informativeness, R_p^2, and firm rank, K_i, is increasing at first, but then inflects at around $K_i = 50$ and declines monotonically thereafter. The systematic relationship between portfolio-level informativeness and firm rank is close to being "inverse-U shaped", showcasing substantial heterogeneity between firms.

Starting from the first portfolio that includes just the firm ranked K_1, adding more firms into the average initially delivers portfolio-wide informational benefit that exceeds any firm-specific noise. However, this stops being the case at the inflection point, around $K_i = 50$, after which the informational costs from adding noisier firms outweigh their benefit. At the peak, the 50th portfolio accounts for more than 60 percent of business-cycle variation in output, exceeding meaningfully the equivalent share that arises from including *all* firms in the sample. We conclude that the average sentiment of 50 firms is

[5] Because regressions with a low number of observations could skew our results, we only retain firms which appear in our sample for at least 10 quarters. Our conclusions do not change if we increase (decrease) this restriction to 20 (5) quarters instead.

Figure 1: The Cross-Section of Firm-Level Sentiment



Notes: Firm- (R_i^2) and portfolio-level (R_p^2) coefficients of determination with respect to HP-filtered aggregate output.

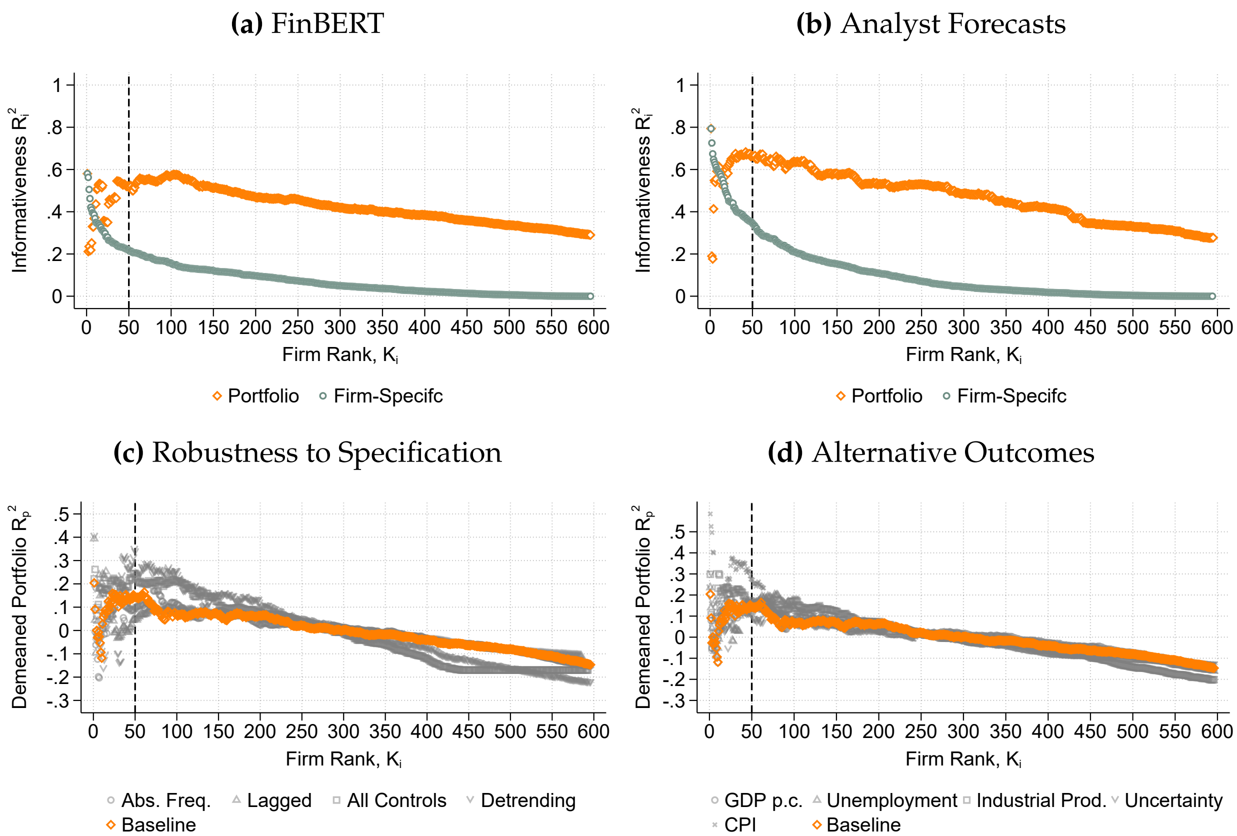
sufficient to account for the majority of business-cycle fluctuations in output.

Following the terminology in Gabaix (2011), we refer to firms with rank $K_i \leq K_{50}$—that is, those that are in the 50th portfolio—as belonging to a set Γ of *Granular Sentiment Firms*. These firms, whose average sentiment can account for most of business-cycle fluctuations, comprise only around 8 percent of firms in our sample. In this sense, sentiment towards these firms is close to being a *sufficient statistic* for understanding the macroeconomy.

Clearly, the evidence in Figure 1 is not an expression of causal evidence. Firm-level sentiment could both be driven by and itself drive output fluctuations. Notwithstanding such possible confounding, the inverse-U shape between portfolio-level informativeness, R_p^2, and firm rank, K_i, is suggestive that the marginal effect of innovations to firm-level sentiment—-that is, firm-specific sentiment shocks—on business-cycle fluctuations is unlikely to be uniformly distributed. A flat curve in Figure 1 would instead have been consistent with this alternative interpretation. Taken at face value, the evidence instead suggests substantial *heterogeneity*: the sentiment towards a select number of firms, at first pass, seems to matter more compared to others.

Robustness: Recall that we consider two alternative measures of firm-level sentiment. The first one, ξ_{it}^2, employs a sophisticated language processing algorithm—specifically, a pre-trained financial BERT (FinBERT) model—to potentially refine the basic dictionary-

Figure 2: The Cross-Section of Firm-Level Sentiment—Robustness



Notes: Firm- (R_i^2) and portfolio-level (R_p^2) coefficients of determination with respect to a macroeconomic aggregate with firm rank K_i on the horizontal-axis under alternative specifications. Panels (a) and (b) report the results using FinBERT sentiment ξ_{it}^2 and analyst forecast errors ξ_{it}^3, respectively. Panels (c) and (d) report the results under alternative specifications and alternative outcomes.

based approach. The second, ξ_{it}^3, computes firm-level sentiment directly using analysts' forecasts. We obtain analysts' forecasts from the I/B/E/S Guidance database and focus on one-quarter-ahead earnings forecasts issued within 90 days after a quarterly earnings call. We quantify sentiment as the difference between the median predicted earnings and its realized value, both scaled by the stock price.[6] Greater values indicate more optimistic earnings forecasts relative to what materializes.[7] Results in Panels (a) and (b) of Figure 2 lend credence to our earlier finding. In each case, the average sentiment of around 50 firms once more accounts for the bulk of output fluctuations. Indeed, the estimated relationships between portfolio-level informativeness and firm rank are remarkably similar to those in Figure 1: inverse-U shaped with an inflection point around $K_i = 50$.

A key concern about our analysis so far is that the relationship between firm-level

[6]We check that our results are robust to alternative windows for which we could have computed analysts' forecasts (e.g., 5 days and 30 days) and find no substantial difference.

[7]See Appendix A.1 for more details on data construction and Appendix A.2 for more details on FinBERT.

sentiment and business-cycle fluctuations might be due to ”reverse causality” (e.g., due to the cyclicality of firm productivity). That is, once one accounts for how business-cycle factors themselves affect firm-level sentiment, the relationship between portfolio-level informativeness, R^2_p, and firm rank, K_i, becomes flat and uniform.[8] To address this issue, we residualize our baseline sentiment measure, ξ^1_{it}, from potentially confounding factors before estimating (3) and (4). Specifically, we residualize ξ^1_{it} from firm fixed effects, aggregate productivity (Fernald, 2014), and (revenue-based) firm-level productivity.[9] We refer to the resulting robust measure as $\hat{\xi}^1_{it}$. Specifications that leverage ξ^1_{it} and $\hat{\xi}^1_{it}$ will be henceforth labeled as the “baseline” and “all controls”, respectively. Although the *level* of R^2_p differs across the different specifications, Panel (c) in Figure 2 reaffirms the basic pattern in the data: there is an inverse-U shaped relationship between portfolio-level informativeness, R^2_p, and firm rank, ameliorating concerns about reverse causation.

We now describe three additional sets of robustness exercises to supplement our cross-sectional finding. First, Panel (c) in Figure 2 shows that our result is also robust to issues related to the contemporaneous nature of the relationship in (3), concerns about the scaling of our sentiment measure, and alternative de-trending methods.[10] Second, consistent with the average sentiment towards a small number of firms being close to a ”sufficient statistic” for general aggregate fluctuations, Panel (d) in Figure 2, furthermore, demonstrates that a small number of firms can also account for the bulk of the variation in other macroeconomic outcomes—specifically, GDP per capita, industrial production, unemployment, inflation, and macroeconomic uncertainty (Baker et al., 2016). The pattern remains unchanged: informativeness, R^2_p, first increases and then decreases with the number of firms added to the portfolio, with the inflection point somewhere between $40 \leq K_i \leq 60$. Third and finally, Panel (a) of Figure A.1 in the Appendix documents a very similar pattern for the case of using (the absolute values of) betas, β_i, as opposed to the R^2_i, as a metric of informativeness in both (3) and (4).

Lastly, to investigate whether the set of informative firms remains the *same* across all the different specifications that we study in Figure 1 and Panels (a)-(d) in Figure 2, we construct the intersection of all the resulting sets of firms. Our finding is that, on average,

[8] This is particularly important for our quantitative analysis, since our theoretical model studies true, aggregate productivity shocks separately from sentiment shocks.

[9] See Appendix A.1 for details on how we estimate firm-level productivity. In the literature our measure is commonly referred to as revenue-based productivity or TFPR. Crucially, this measure also captures the impact of demand-side factors on the firm.

[10] In Panel (c) of Figure 2, we run a separate specification where firm-level sentiment, $\xi^1_{i,t-1}$, in equation (3) is one-quarter lagged (”Lagged”). We also plot the results from a specification in which we do not scale our sentiment measure, ξ^1_{it}, by transcript length, B_{it} (”Abs. Freq.”), as well as a specification that instead of the HP-filter detrends output by residualizing from time-fixed effects (”Detrending”).

60% of the firms in each set are the same as in the baseline. This suggests a considerable degree of overlap across the different specifications in the identified firms. All else equal, changes in the sentiment of a small number of core firms systematically account for a large share of business-cycle fluctuations. We return to this important issue later in Section 4.

This concludes the first step of our empirical analysis. In sum, we have demonstrated the presence of systematic *heterogeneity* in firm-level sentiment. Remarkably, the average sentiment of only around 50 firms accounts for the bulk of business-cycle fluctuations. The data refute the assumption that sentiment of each individual firm contributes to the characterization of business cycles equally. Indeed, including more firms into our average measure only reduces the proportion of fluctuations attributable to sentiment. Building on these findings, the next section delves into the specific characteristics of these 50 firms.

3.2 Sentiment and Downstreamness

We proceed to analyze the *characteristics* of the firms whose average sentiment is closely associated with the state of the business cycle. We pursue an intentionally agnostic approach to this question. A significant body of work, studying sectoral and firm-specific productivity shocks, suggests that firm size is the relevant characteristic determining whether individual productivity shocks lead to aggregate fluctuations (e.g., Carvalho and Grassi, 2019). However, it is not immediately clear that *size* is also the relevant factor for sentiment-driven fluctuations. In fact, recent studies have highlighted other primary characteristics (e.g., price stickiness) for explaining the aggregate consequences of non-productivity-related disturbances (e.g., Afrouzi and Bhattarai, 2023).

We construct an array, $\mathbf{X}_i$, of potential explanatory variables for each firm. The variables we collect and/or compute include: (i) firm size (measured by the logarithm of total assets and the logarithm of overall sales); (ii) cyclicality (measured by the market beta); (iii) book-to-market ratio; (iv) investment intensity; (v) market valuation; (vi) leverage; (vii) liquidity; (viii) Tobin's Q; and (ix) firm idiosyncratic return volatility. Additionally, we also include (x) our measure of sectoral upstreamness, as discussed in Section 2.1. Each of these variables could potentially be the reason for why some firms are more informative than others. To simplify the analysis, we collapse each variable into the $N \times 1$ dimension by averaging firm-specific values across sample years. We denote the resulting array by $\mathbf{X}_i$. Appendix A.1 provides further details on variable definitions and construction.

We next regress our firm-specific measure of informativness, R_i^2, onto the array of controls, $\mathbf{X}_i$, using three different specifications: (1) a linear regression; (2) a probit regression, where the dependent variable, G_i, takes the value of 1 if the firm is in our granular informativeness set Γ and 0 otherwise; and (3) an ordinal probit regression, where the

Table 1: Sentiment and Downstreamness

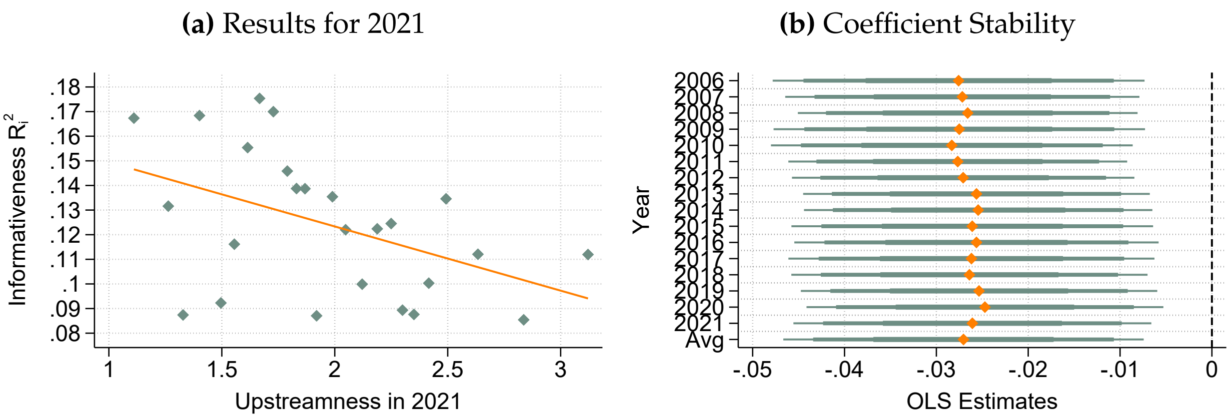
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	Probit	Probit	Ordinal Probit	Ordinal Probit
Upstreamness in 2021	-0.024**	-0.026***	-0.397***	-0.418***	-0.181*	-0.219**
	(0.009)	(0.010)	(0.150)	(0.161)	(0.097)	(0.099)
Log (Assets)	0.015	0.018	0.076	-0.282	0.144	0.247*
	(0.011)	(0.017)	(0.142)	(0.358)	(0.120)	(0.147)
Log (Sales)	-0.011	-0.014	-0.118	-0.218	-0.070	-0.083
	(0.009)	(0.010)	(0.147)	(0.156)	(0.096)	(0.096)
Market Beta		-0.009		-0.012		-0.128*
		(0.008)		(0.109)		(0.068)
Book-to-Market		-0.007		0.018		-0.070
		(0.009)		(0.149)		(0.079)
Investment		-0.013**		-0.039		-0.125***
		(0.005)		(0.093)		(0.047)
Valuation		0.003		0.487*		-0.083
		(0.017)		(0.256)		(0.157)
Leverage		-0.013**		-0.089		-0.138**
		(0.006)		(0.113)		(0.054)
Liquidity		-0.020**		-0.196*		-0.166**
		(0.009)		(0.111)		(0.073)
Tobin's Q		0.004		-0.116		0.045
		(0.011)		(0.184)		(0.102)
Return Volatility		0.021*		0.249**		0.153**
		(0.012)		(0.114)		(0.075)
Observations	531	469	531	469	531	469

Note: Columns (1) and (2) report results from LS cross-sectional regressions with R_i^2 as the dependent variable. Columns (3) and (4) report results from Probit cross-sectional regressions with a binary indicator, G_i, that takes the value of one for firms belonging to Γ as the dependent variable. Columns (5) and (6) report results from Ordinal Probit cross-sectional regressions with rank indicator K_i as the dependent variable. Standard errors (in parentheses) are clustered at the industry level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

dependent variable is the rank integer K_i itself. Table 1 presents the estimates.

In all cases, the only variable that is systematically significant and economically important is *sectoral upstreamness*. All other controls, including firm size (measured by either book assets or sales) and cyclicality, matter little for the informativeness of firms. All else equal, the more downstream a firm is—that is, the closer it is to the end consumer—the more fluctuations in its sentiment are associated with the business cycle. The magnitude of the estimated effect of upstreamness is, moreover, substantial. To ease interpretation, Figure A.2 in the Appendix reports predictive margins from the probit model. Increasing downstreameness by one unit (moving one step closer to the final consumer) increases the probability of being in our high-correlation set by 10% percent. Variations in upstreamness modify the informativeness of firms by considerable amounts.

Figure 3: Downstreamness and Sentiment: OLS Estimates

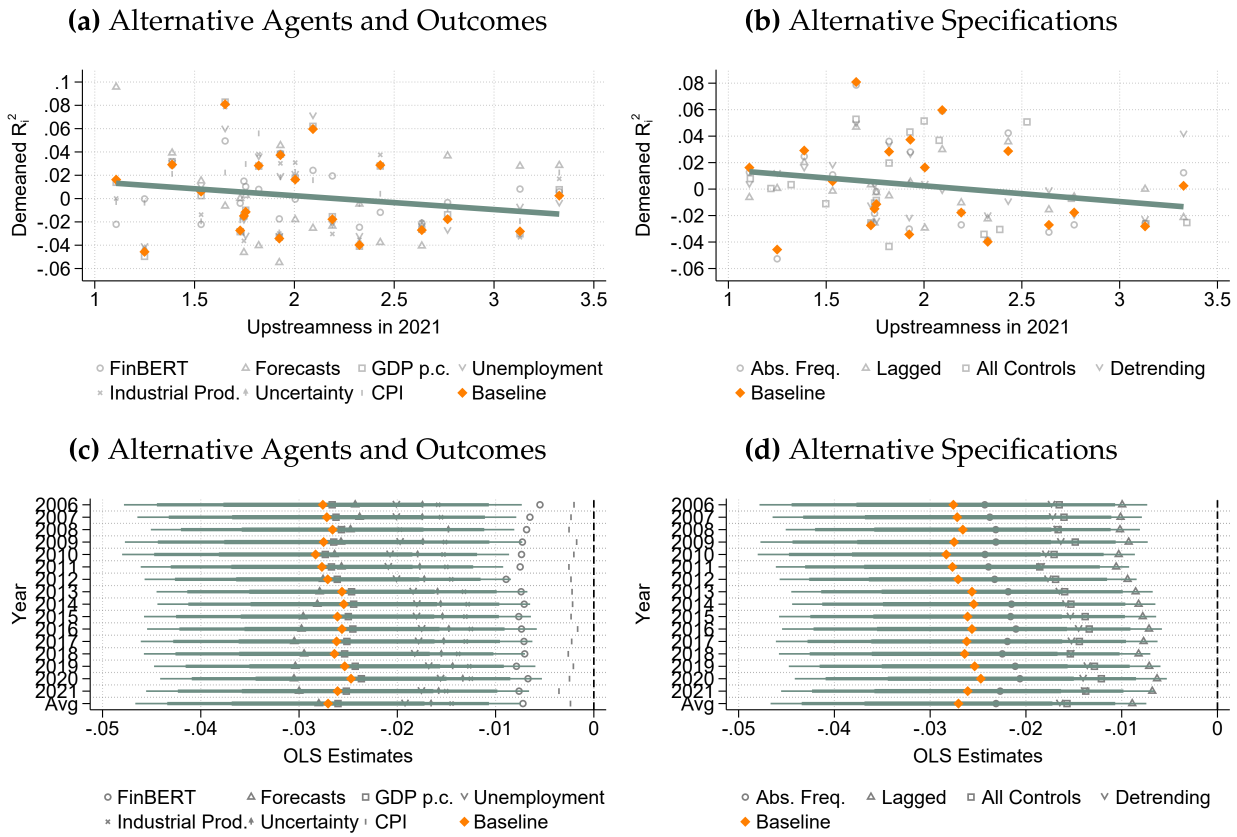


Notes: Results from cross-sectional LS regressions of upstreamness U_{st} on firm-level informativeness R_i^2. Panel (a) reports a binned-scatter relationship for 2021. Panel (b) shows point estimates and confidence intervals for each year. Horizontal bars are 68%, 90%, and 95% confidence intervals, clustered at the industry level.

Robustness: These results do not depend on outlier observations, the year used to measure upstreameness, or the precise econometric specification. Panel (a) in Figure 3 presents a binned scatter plot of the relationship between upstreamness in 2021, $U_{s,2021}$, and our measure of informativeness, R_i^2. Importantly, the effects of all other firm controls have been partialled out. All else equal, we find that firms that belong to a less upstream sector have a higher R_i^2, i.e., are more closely associated with business-cycle fluctuations. Across the range of the upstreamness measure, more *downstream* firms are more informative. Panel (b) in Figure 3 plots point estimates along with 68%, 90%, and 95% confidence intervals for *every* year. We observe remarkable coefficient stability across time. Figure A.3 in the Appendix reports similar results for the probit regression and its ordinal counterpart.

Furthermore, we expand on the cross-sectional relationship between measured informativeness and sectoral upstreamness and perform three sets of exercises that are in line with those conducted in Section 3.1. First, we consider the same alternative specifications that we explained in detail in Section 3.1, which, among others, residualize our sentiment measure from potentially confounding factors before estimating (3) and (4). Second, we consider the two alternative measures of firm-level sentiment, based on FinBERT and analyst forecasts, respectively. And third, we consider a broader range of macroeconomic outcomes than output. Figure 4 presents the results. Panels (a) and (b) show binned scatter plots for the year 2021, while Panels (c) and (d) present point estimates and confidence bounds for all years in the sample. On balance, we find that our main result is highly robust. With the exception of the specification that uses inflation as the dependent variable, all point estimates lie within the confidence interval of our baseline model. The negative

Figure 4: Downstreamness and Sentiment—Robustness



Notes: Results from cross-sectional LS regressions of upstreamness U_{st} on firm-level informativeness R_i^2 under alternative specifications. Panels (a) and (b) report binned-scatter relationships for 2021. Panels (c) and (d) show point estimates and confidence intervals for each year. Horizontal bars are 68%, 90%, and 95% confidence intervals. Standard errors are clustered at the industry level.

relationship between upstreamness and informativeness is a general feature of the data, even after the inclusion of additional controls and the use of alternative measures.[11] All in all, we conclude that the relationship between informativeness and downstreamness is robust to outlier observations, reverse-causality concerns, different detrending techniques, various regression specifications, and extends to alternative macro outcomes. Lastly, Appendix A.3 demonstrates that our results are also robust to alternative sectoral definitions and to the use of betas, $|\beta_i|$, as a measure of informativeness instead of the R_i^2.

In summary, in this section, we have demonstrated that, in the panel of listed firms, business-cycle fluctuations are strongly associated with the average sentiment of a select group of roughly 50 firms, and characterized the main driver of this set of firms—sectoral

[11] Appendix A.5 presents further robustness exercises. In particular, we run placebo regressions for our baseline cross-sectional exercise and find that, in each year of the sample, it is highly unlikely that the relationship between informativeness and downstreamness was obtained by chance.

downstreamness. Our results, all else equal, suggest that sectoral downstreamness is a dominant characteristic for understanding sentiment-induced fluctuations by firms. We now move on to the time-series level, in order to show that *innovations* to a simple index of the sentiment of firms in our granular set indeed help drive business-cycle fluctuations.

4 Granular Sentiment and the Business Cycle

Thus far, we have established that the sentiment towards future performance of a small subset of firms can account for much of the variation in macroeconomic outcomes. We have also identified the key characteristic of this subset—sectoral downstreamness. In this section, we proceed to the third and final step of our empirical strategy. We examine the dynamic consequences of innovations in the sentiment of firms in our granular set and demonstrate that these shifts drive business-cycle fluctuations.

4.1 An Index of Granular Sentiment

Similar to Gabaix (2011), we start by computing a weighted average of firm-level sentiment, ξ^1_{it}, using previous-year downstreamness as weights, $\omega_{i,t-4}$.[12] Crucially, we restrict the set of firms to those in our granular set Γ. We denote the resulting index by $\mathcal{S}_t$:

$$\mathcal{S}_t \equiv \sum_{i \in \Gamma} \omega_{i,t-4} \xi^1_{it}, \tag{5}$$

where the weights, ω_{it}, sum to unity in each period, and refer to the index as the *Granular Sentiment Index (GSI)*. To further safeguard against the selection on factors other than downstreamness driving our results, we also compute a weighted average of the sentiment of firms that belong to the highest downstreamness decile:

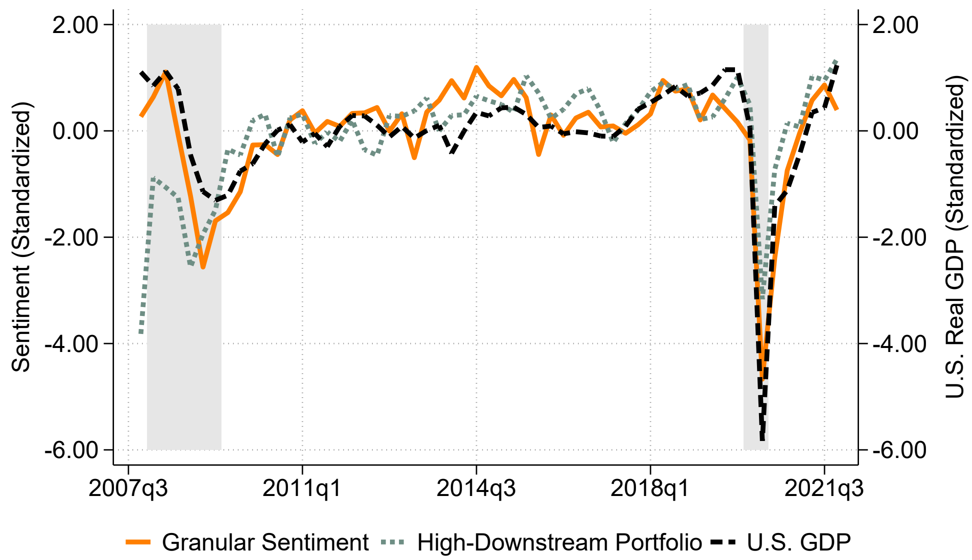
$$\mathcal{P}_t = \sum_{i \in P10_t} \omega_{i,t-4} \xi^1_{it}, \tag{6}$$

where $i \in P10_t$ signifies firms that belong to the top decile of firms sorted by downstreamness in quarter t. We refer to $\mathcal{P}_t$ as the *High-Downstreamness Sentiment Index (DSI)*.[13] We

[12] Compared to an equally-weighted average, using donwstreamness as weights attaches greater importance to more downstream firms' sentiment, although this does not affect our results. We normalize the inverse of the upstreamness measure, U^{-1}_{it}, and use this quantity as the weight, ω_{it}.

[13] Notice that any empirical test that involves $\mathcal{P}_t$ is substantially harder to pass than its equivalent $\mathcal{S}_t$-counterpart, as it involves sorting based on a *pre-defined* characteristic.

Figure 5: Granular Sentiment over Time



Notes: Time-series plots of the granular sentiment index (GSI) $\mathcal{S}_t$, the high-downstreamness index (DSI) $\mathcal{P}_t$, and detrended output.

also aggregate up the robust firm-level sentiment measure, $\hat{\xi}^1_{it}$, which has been orthogonalized from confounding factors, and refer to the resulting indices as $\hat{\mathcal{S}}_t$ and $\hat{\mathcal{P}}_t$.

Figure 5 plots the time-series behavior of $\mathcal{S}_t$, $\mathcal{P}_t$, and aggregate output, where all series have been standardized for comparability purposes. Panels (a) and (b) of Figure A.6 in the Appendix showcase the two alternative measures of sentiment that are derived from the FinBERT algorithm and analyst forecasts, respectively. In Panels (c) and (d), the Figure also plots the dynamics of the robust measures. The close relationship over time between sentiment and output, visible in Figure 5, carries over to all of these alternative measures.

We study the relationship between our two indices, $\mathcal{S}_t$ and $\mathcal{P}_t$, and various indicators of macroeconomic conditions, as well as several economic- and market-based indicators of aggregate sentiment. Importantly, the alternative measures of economy-wide sentiment that we consider are constructed from underlying datasets that differ from ours. Instead of picking a singular index in an ad-hoc manner, we instead focus on the first principal component (PC) from an array of five well-known indices: the OECD Business Confidence Index (BCI), the University of Michigan Index of Consumer Sentiment, the ISM Purchasing Managers' Index (PMI), the Sentix sentiment index, and the news-based economic sentiment index by van Binsbergen et al. (2024).[14] These five measures capture sentiment towards the future of distinct economic agents: businesses, consumers,

[14] See Appendix A.1 for further details.

market participants, and newspaper outlets. The first PC of these, which we denote by $\mathcal{K}_t$ or "economy-wide sentiment", should thus represent a balanced measure of overall sentiment.

Table A.3 in the Online Appendix shows that our two indices, $\mathcal{S}_t$ and $\mathcal{P}_t$, across the board, are highly contemporaneously correlated with different measures of macroeconomic conditions and sentiment. In particular, the GSI, by itself, can account for around 40% and 70% of the unconditional variation in economy-wide sentiment and output, respectively.[15] The relationship between the DSI and these variables is less strong, as should be expected, but still commensurable. Notably, both indices are *positively* correlated with output and inflation, albeit the relationship with inflation is noisy. This suggests, at first pass, that autonomous changes in sentiment could trigger "demand-type" fluctuations, consistent with a Piguvian view of business cycles. We turn to this possibility next.

4.2 Dynamic Effects on the Macroeconomy

We estimate the dynamic effects of changes in our granular sentiment indices on the macroeconomy. To this end, we run Jordà (2005)-style linear local projections:

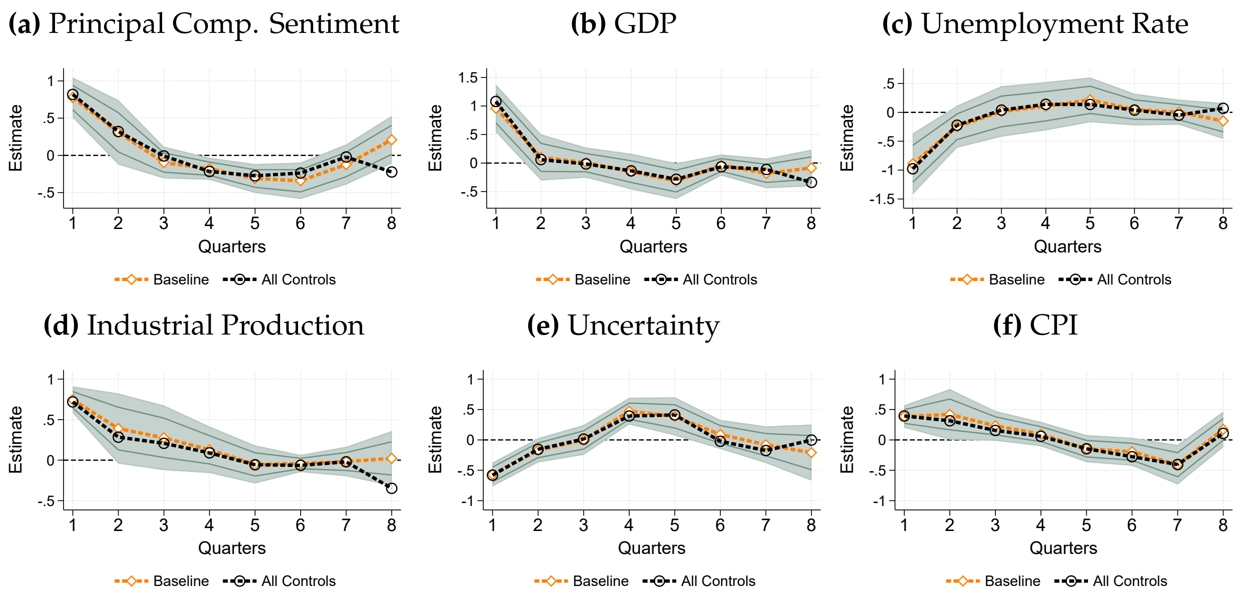
$$Y_{t+h} = \delta_h + \beta_h \times Z_t + \sum_{\ell=1}^{L} \gamma_{h\ell} \mathbf{X}_{t-\ell} + u_{ht} \tag{7}$$

where δ_h are horizon-specific fixed effects. The main coefficient of interest, β_h, is that on the granular sentiment measure $Z_t \in \{\mathcal{S}_t, \mathcal{P}_t, \hat{\mathcal{S}}_t, \hat{\mathcal{P}}_t\}$. As outcome variables, Y_{t+h}, we consider the aforementioned business-cycle variables (output, industrial production, unemployment, uncertainty, and inflation) and the economy-wide sentiment measure, $\mathcal{K}_t$. Crucially, in every specification that we consider, we saturate the linear projection with the same vector of controls, $\mathbf{X}_t$, which includes economy-wide TFP (Fernald, 2014), the Federal Funds Rate, and the (real) Nasdaq market return. All of these variables are standard additions to macro Vector Autoregressions (VARs) and help address concerns about omitted variable bias. We include $L = 2$ lags of all variables in the projections. Figures 6 and 7 depict the resulting impulse response functions for horizons up to $h = 8$ quarters ahead from a 1-standard deviation increase in the baseline and robust measures.

In all cases, a positive innovation in granular sentiment induces "business-cycle-like" synchronized movements: economic sentiment spikes, output and industrial production

[15]Table A.2 in the Appendix presents a matrix of correlations across all the sentiment indices that we consider: $\mathcal{S}_t$, $\mathcal{P}_t$, $\mathcal{K}_t$, and the five underlying indices of $\mathcal{K}_t$. We see that across the board the correlations are very high. Interestingly, $\mathcal{S}_t$ is most closely associated with the OECD BCI which is intuitive since $\mathcal{S}_t$ captures the sentiment towards listed firms.

Figure 6: Dynamic Effects of Granular Sentiment on the Macroeconomy

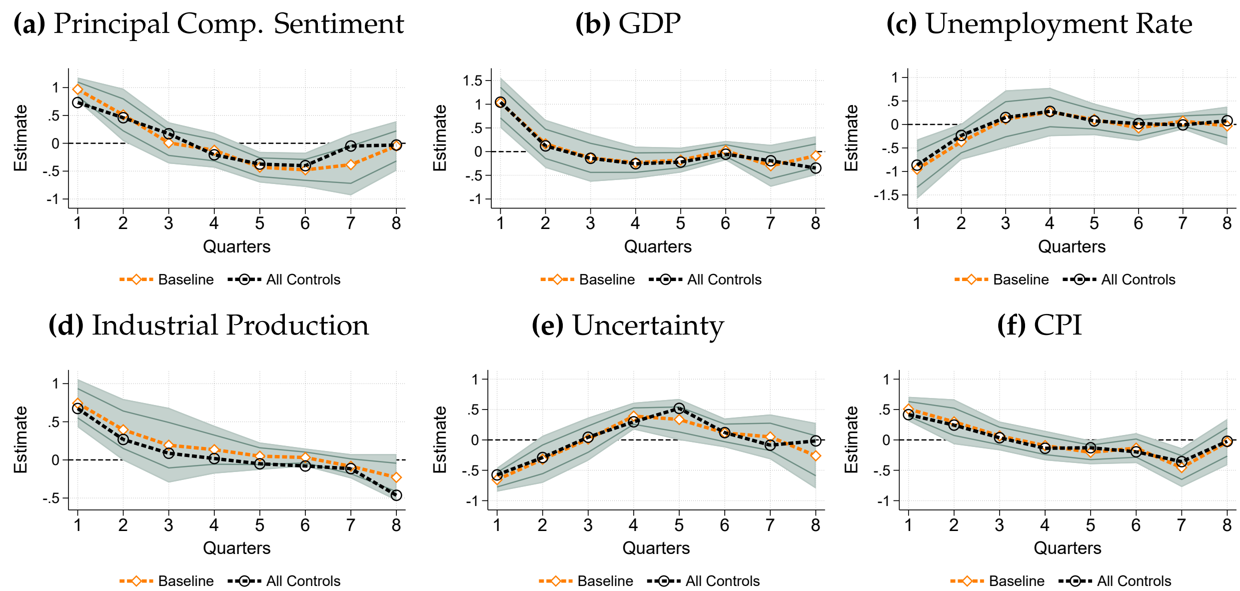


Notes: Local-projection estimates of the Granular Sentiment Indices on macro and sentiment aggregates. The "baseline" and "all controls" specifications correspond to the use of measures $\mathcal{S}_t$ and $\hat{\mathcal{S}}_t$, respectively. Lines correspond to 68%, and shaded areas to 90%, robust (Newey-West) confidence bands, respectively.

rise, unemployment falls, uncertainty goes down, and inflation rises. These effects are, on balance, economically large and statistically significant for around two-three quarters. For example, over the course of the first year, a 1-standard deviation increase in the GSI leads to a cumulative increase in output equivalent to roughly a third of the variable's standard deviation. Such magnitudes are remarkable given that the impulse responses already control for two critical sources of business-cycle variation: aggregate productivity and monetary policy. Moreover, in the case of the robust measures, $\hat{\mathcal{S}}_t$ and $\hat{\mathcal{P}}_t$, these effects are additionally robust to the presence of potential confounders such as firm fixed effects and time-varying firm-level productivity. Taken together, fluctuations in sentiment towards only around 50 downstream firms lead to coordinated spikes in activity and inflation closely reminiscent of those triggered by standard demand shocks.

The results in Figure 6 and 7 are consistent with a large body of work studying the macroeconomic effects of innovations in *economy-wide* sentiment (e.g., Beaudry and Portier, 2004; Lorenzoni, 2009; Barsky and Sims, 2012; Chahrour and Jurado, 2018). Our estimated impulse responses are, like those found in this literature, comparatively short-lived, up to one year in length, and resemble those caused by standard demand shocks. Our results, thus, confirm that innovations in granular, firm-driven sentiment have similar macroeconomic propagations to those of economy-wide sentiment shocks.

Figure 7: Dynamic Effects of High-Downstreamness Sentiment on the Macroeconomy



Notes: Local-projection estimates of the high-downstreamness indices on macro and sentiment aggregates. The "baseline" and "all controls" specifications correspond to the use of measures $\mathcal{P}_t$ and $\hat{\mathcal{P}}_t$, respectively. Lines correspond to 68%, and shaded areas to 90%, robust (Newey-West) confidence bands, respectively.

Notwithstanding this similarity, one shortcoming of the approach which studies innovations in overall sentiment is that the associated business-cycle responses are often attributed to "black box" drivers, containining statistical summaries that are not clearly interpretable at the micro level. By contrast, the measures of granular sentiment proposed above offer a source of business-cycle variation that is directly measured at the micro-level of individual firms, and is clearly interpretable in terms of the optimism or pessimism related to specific firms. In this sense, the two indices above offer a useful tool to evaluate the nature of estimated sentiment shocks.

Robustness: We now consider several extensions of our baseline specification, the themes of which we first detailed in Section 3.1. We consider, among other extensions, alternative measures of sentiment, macro outcomes, and econometric specifications. Figures A.7 and A.8 in the Appendix present the results in two parts for ease of interpretation. The estimated impulse responses are remarkably in consonance with our earlier results. We find that both the contemporaneous and dynamic effects of innovations in granular sentiment on the macroeconomy practically everywhere lie within the confidence band of our baseline specification. The estimated impact, furthermore, usually remains statistically significant for several quarters. In addition, Figure A.9 in the Appendix presents a full set of local projections using an alternative de-trending approach (in which we residualize all macro variables from time fixed effects instead of HP-filtering) and once more shows that the estimates of dynamic effects are robust. We conclude that the dynamic consequences of innovations to our granular sentiment indices on the macroeconomy are robust to alternative measures, outcome variables, and specific treatments of the raw data.

4.3 Causality

To further support the interpretation that innovations in sentiment drive business cycles rather than merely reflect them, we address the issue of causality with two different and complementary approaches. First, we employ a novel identification strategy to examine the effects of structurally identified sentiment shocks within a VAR framework. Second, we directly instrument changes in sentiment using a recently proposed instrument designed to proxy for sentiment-driven fluctuations. Across both approaches, the results are consistent with those obtained from the baseline linear projections.

4.3.1 Structural Shocks to Sentiment

We begin with a structural identification approach that closely follows the recently proposed methodology in Chahrour and Jurado (2021). We estimate a VAR with the same variables and lag order as our baseline local projections.

The key identifying assumption of this exercise is as follows. Fluctuations in sentiment are driven by either structural innovations to sentiment itself or by other innovations which drive people's noisy expectations about *future* sentiment that may or may not realize. We refer to the former as "true sentiment" shocks and to the latter as "idiosyncratic noise" shocks. Granular sentiment and the economy can thus respond to true sentiment shocks on impact and to idiosyncratic shocks with some pre-specified lag. In slight anticipation of our structural model, the true sentiment shock closely corresponds to the object ξ_{it}, which in the model represents the common component of firm beliefs about the productivity of a particular sector. As such, it is precisely ξ_{it} that is the cause of sentiment-driven fluctuations, and our goal here is to isolate it from idiosyncratic noise.[16]

Figure 8 demonstrates these identifying assumptions in the form of impulse response functions with 10 periods before and after the shock. The left panel plots the response of granular sentiment, $\mathcal{S}_t$, to the true sentiment shock. The response is large on impact

[16] We also need to specify a "target horizon" for the noise shocks. In line with the literature, we set the target horizon to 20, which is sufficiently long for agents to be able to forecast future changes in sentiment (Chahrour et al., 2024). Our results do not change materially if we vary this parameter.

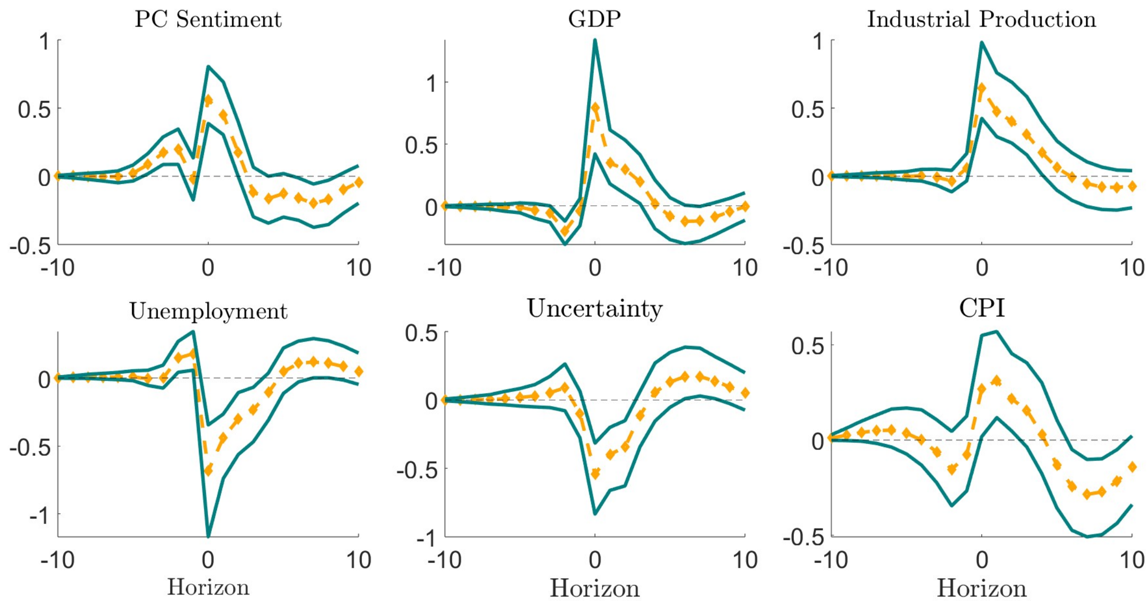
Figure 8: Structural VAR Identification

Notes: This figure plots impulse response functions of granular sentiment, $\mathcal{S}_t$, to a one standard deviation true granular sentiment shock (left panel) and noise shock (right panel) at time t = 0. The dashed lines are point estimates and the solid lines are 90% bias-corrected bootstrap confidence bands. All variables are standardized.

with a gradual dissipation within roughly 5 quarters. The right panel shows the response of granular sentiment to the noise shock, which is a noisy zero. As such, our approach clearly isolates true sentiment from idiosyncratic, noisy disturbances.

Our aim is to quantify the response of the endogenous variables to the true granular sentiment shock, which has been orthogonalized from the influence of other shocks. Figure 9 illustrates our results. We identify three key takeaways. First, both qualitatively and quantitatively, our results mirror closely those obtained using local projections. A positive, true granular sentiment shock leads to an increase in overall sentiment, $\mathcal{K}_t$, output and industrial production, along with a decline in unemployment and uncertainty, and a rise in inflation. Second, the effects of these responses persist for approximately 4-5 quarters, which is consistent with but slightly longer than the durations observed in the local projection estimates. Third, we document virtually no advance movement from any variable prior to the shock. This pattern suggests that shocks to granular sentiment are not easily anticipated. This finding further supports the argument that innovations to sentiment act as a driver of the business cycle rather than merely reflecting it.

Figure 9: Structural VAR Responses to the True Granular Sentiment Shock



Notes: This figure plots impulse response functions to a one standard deviation shock in true granular sentiment at time t = 0. The dashed lines are point estimates and the solid lines are 90% bias-corrected bootstrap confidence bands. All variables are standardized.

4.3.2 Instrumental Variable Approach

Our second approach for addressing causality involves instrumenting changes in our granular sentiment indices using mass shooting fatalities in the U.S. Lagerborg et al. (2019) extensively argue that the exclusion restriction for this instrument is verified, as fatal shootings in the U.S. are largely randomly assigned across space and time. The instrument has, furthermore, been argued to also be highly relevant: Lagerborg et al. (2022) document it to be highly correlated with measures of overall consumer and firm sentiment. Because our micro-based measures, $\mathcal{S}_t$ and $\mathcal{P}_t$, themselves are good proxies for economy-wide sentiment, as shown earlier, it follows that mass shooting fatalities are also relevant for our context.

We run 2SLS regressions with mass shooting fatalities as an instrument for granular sentiment $\mathcal{S}_t$. Table 2 reports the results from the second stage as well as first-stage F-statistics. The top and bottom panels report the results without and with additional time-series controls, respectively. We find that our granular sentiment index, when instrumented with mass shootings fatalities, has significant, contemporaneous effects on economy-wide sentiment, output, and unemployment. The impacts on industrial production and uncertainty are economically similar albeit with larger standard errors. The impact on inflation, however, cannot be statistically distinguished from a noisy zero. The

Table 2: IV Regression with U.S. Mass Shootings

	Independent Variable: Granular Sentiment instrumented by Mass Shootings					
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	PC Sentiment	GDP	Unemployment	Industrial Production	Uncertainty	CPI
Sentiment	0.545***	0.487***	-0.659***	0.607**	-0.431	-0.083
	(0.192)	(0.114)	(0.155)	(0.295)	(0.327)	(0.291)
Observations	39	39	39	39	39	39
All Controls	✗	✗	✗	✗	✗	✗
First-Stage F-stat	15.325	15.325	15.325	15.325	15.325	15.325
	Independent Variable: Granular Sentiment instrumented by Mass Shootings					
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	PC Sentiment	GDP	Unemployment	Industrial Production	Uncertainty	CPI
Sentiment	0.576***	0.361***	-0.558***	0.305	-0.492	-0.326
	(0.210)	(0.089)	(0.139)	(0.245)	(0.353)	(0.367)
Observations	39	39	39	39	39	39
All Controls	✓	✓	✓	✓	✓	✓
First-Stage F-stat	15.848	15.848	15.848	15.848	15.848	15.848

Notes: Results from IV regressions of granular sentiment $\mathcal{S}_t$, instrumented by U.S. mass shooting fatalities, on macroeconomic aggregates. Top and bottom panels report results without and with controls, respectively. Controls include aggregate TFP, the Fed Funds Rate, and real Nasdaq returns. Newey-West standard errors are in parentheses. *p < 0.1; **p < 0.05; ***p < 0.01.

relatively muted effect on inflation was already evident in Figures 6 and 7 above.[17] Table A.4 in the Appendix reports very similar results for the DSI, $\mathcal{P}_t$.

Overall, we view the findings from the above two exercises as providing crucial independent support for a plausibly causal effect of innovations in granular sentiment on the macroeconomy.

4.4 Discussion

We have documented that changes in the sentiment of approximately 50 downstream firms can influence business-cycle fluctuations. An increase in our measures of granular sentiment triggers a "Pigouvian-style expansion," characterized by a surge in output and industrial production, a reduction in unemployment and uncertainty, and a possible rise in inflation. Identifying the purely exogenous sources of business cycles remains a significant challenge (e.g., Angeletos et al., 2020). Nevertheless, the detailed microeconomic approach

[17] An important qualifier to this discussion is the relatively low values of first-stage F-statistics, even though they are above the rule-of-thumb threshold of 10. Lagerborg et al. (2022) also note that instrument power can fluctuate depending on the exact specification and definition of the mass shootings instrument.

to measuring sentiment that we follow, combined with comprehensive cross-sectional and time-series evidence that includes structural and instrumental-variable identification methods, paint a robust picture which supports the idea that sentiment-driven fluctuations of comparatively few firms affect the macroeconomy.

That said, our empirical results have so far been mute on two pertinent questions: (i) *why* is downstream firms' sentiment important for aggregate fluctuations instead of, for example, the sentiment that is accorded to more upstream firms?; and (ii) what is the overall contribution of orthogonal innovations in downstream firms' sentiment to business-cycle fluctuations? To provide a first-pass answer to these questions, we proceed to lay down an empirically-motivated multi-sector economy in which sentiment shocks to downstream firms drive business-cycle dynamics. The outcome of the model will also allow us to quantify the separate channels which cause downstream-sector firms to be central for sentiment-driven fluctuations.

5 Theoretical Framework

In this section, we study sentiment-driven fluctuations in a multi-sector model with imperfect information. A representative household decides how much labor to supply and how much to consume. Firms decide how much labor and intermediate inputs to demand and to use in production. To start, there are 2 sectors in the economy: an upstream (u) and a downstream (d) sector, where each sector consists of a continuum of firms that sell their goods in competitive markets. Sector $i \in \{u, d\}$ is defined by how good i enters in the production function of other sectors and into household consumption. The modeling structure is similar to what has been used in Acemoglu et al. (2012) and Chahrour et al. (2021), but with the crucial difference that firms make *endogenous* attention choices.

Sectors and Firms: A firm in sector i at time $t = 1, 2, ...$ uses the production function

$$Q_{it} = Z_{it}\left(\Pi_j X_{ijt}^{\alpha_{ij}}\right) L_{it}^{\delta_i} \tag{8}$$

to produce output, Q_{it}, where the variable Z_{it} is a sector-specific productivity shock, X_{ijt} is the intermediate input used by sector i (which is produced by sector j), and L_{it} is sector i's labor input. Coefficients $\alpha_{ij} \geq 0$ denote the factor share of good j used in the production of good i. The production function exhibits *decreasing returns to scale,* so that $\sum_j \alpha_{ij} + \delta_i = 1 - \gamma$, where $\delta_i \geq 0$ and $\gamma \in (0, 1)$. Importantly, the upstream sector ($i = u$) does not use downstream goods ($j = d$) in production ($\alpha_{ud} = 0$); only the converse is true.

Firms in sector i choose labor L_{it} and intermediate inputs $\left\{X_{ijt}\right\}$ to maximize profits,

$$\Pi_{it} = P_{it}Q_{it} - W_t L_{it} - \sum_j P_{jt}X_{ijt}, \tag{9}$$

where P_{it} denotes the price of goods produced by sector i, and W_t is the wage rate.

Households: The representative household decides how much to work and how much to consume in each period. Its preferences are described by the utility function

$$\mathcal{U}_t = C_t - \frac{\left(\sum_i L_{it}\right)^{1+1/\nu}}{1+1/\nu}, \tag{10}$$

where C_t denotes the consumption of the downstream good, and $\nu > 0$. We normalize the price of consumption to one. The household's budget constraint is:

$$C_t = W_t \sum_i L_{it} + \sum_i \Pi_{it}. \tag{11}$$

The representative household's objective is to maximize its utility (10) subject to (11).

Timing and Information Structure: Each period is comprised of two stages.

In the *first stage*, to capture that production decisions are taken under imperfect information about demand, firms choose labor inputs before production takes place and before equilibrium prices are observed. In place, firms commit to their labor choices based on noisy signals about their own and others' productivity. In particular, a firm in sector i at time t is assumed to observe:

$$s^z_{ijt} = z_{jt} + \xi_{it} + m^z_{ij}\varepsilon_{ijt}, \tag{12}$$

where $z_{jt} \equiv \log Z_{jt}$, $m^z_{ij} \geq 0$, and $j = \{u, d\}$. The shocks $\xi_{it} \sim \text{AR1}\left(\rho_\xi, \sigma^2_\xi\right)$ and $\varepsilon_{ijt} \sim \mathcal{N}(0,1)$ are independent of each other, sectoral productivities, and across time. Nature draws

$$z_{jt} = \vartheta_t + u_{jt}, \tag{13}$$

where $\vartheta_t \sim \text{AR1}\left(\rho_\vartheta, \sigma^2_\vartheta\right)$ and $u_{jt} \sim \mathcal{N}\left(0, \sigma^2_u\right)$ are independent of each other and all other disturbances. Crucially, the shocks $\{\xi_{it}\}$ in equation (12) capture sector-specific "sentiment shocks" to firms' information about their own and others' productivity—the common modeling-device that we use to capture sentiment-related shocks to firms' information about others in the economy (e.g., Angeletos and La'O, 2013; Acharya et al., 2021). Importantly, the shocks $\{\xi_{it}\}$ are orthonogal to realized firm productivity z_{jt} in (12), and hence correspond closely to the residualized empirical sentiment measure, $\hat{\xi}_i^1$, analyzed in Section 3 and 4.

In addition to the information observed about sectoral productivities, firms in sector i at time t are also assumed to observe a noisy signal of the previous period's output:

$$s_{ijt}^{q} = q_{jt-1} + m_{ij}^{q} e_{ijt}, \tag{14}$$

where $m_{ij}^{q} \geq 0$ and $e_{ijt} \sim \mathcal{N}\left(0, \sigma_e^2\right)$ is independent of all other disturbances.

We summarize the information firms in sector i base their labor choices on in the information set $\Omega_{it} \equiv \left\{s_{i0} \cup (s_{is})_{s=1}^{s=t}\right\}$, where $s_{it} \equiv \left(s_{iit}^{z}, s_{ijt}^{z}, s_{iit}^{q}, s_{ijt}^{q}\right), j \neq i$.[18]

In a *second stage*, after labor choices are sunk, firms choose intermediate inputs and pay a wage that induces the household to supply the amount of labor inputs chosen in the first stage. Production takes places and the household consumes. From firms' perspective, labor inputs may be *ex post* suboptimal, because of information frictions, while from the household's perspective labor supply is always optimal.

Finally, we assume that, in the initial period $t = 0$, each firm i first chooses its attention vector $m_i \equiv \begin{pmatrix} m_{id}^{z} & m_{iu}^{z} & m_{id}^{q} & m_{iu}^{q} \end{pmatrix} \in \mathbb{R}_{+}^{4}$ to maximize expected profits, subject to an attention cost function, $K(m)$, where $K(\cdot)$ is positive, decreasing in all elements of m, and convex. In line with much of the related literature, the firm makes this choice *ex ante*, behind the veil of ignorance (e.g. Veldkamp, 2023). The firm then receives an (infinitely) long sequence of signals, denoted by s_{i0}. The latter assumption ensures that the firm's signal extraction problem is initialized in the steady state (Maćkowiak et al., 2023).

Discussion: Before characterizing the equilibrium, it is useful to discuss a few key concepts that will play a central role in our subsequent analysis.

First, note that we can summarize the input-output linkages between sectors with a matrix $\mathbf{A} = \left[\alpha_{ij}\right]$, which with some abuse of terminology we will refer to as the economy's input-output matrix. Notice also that $\mathbf{A}$ is a triangular matrix, because of the upstream-downstream structure of the economy. We define Λ_i as the ith element of the column vector $\Lambda \equiv (\mathbf{I} - \mathbf{A}')^{-1}\beta$, where $\beta \equiv \begin{pmatrix} 1 & 0 \end{pmatrix}'$.[19] The coefficient Λ_i is a measure of the *Bonacich centrality* of a sector, weighted by the sector's share of final consumption (i.e., the *Domar*

[18] We assume that in the initial period $t = 0$, all firms receive an (infinitely) long sequences of signals from equations (12) and (14), which we denote by s_{i0}. This assumption follows the convention in the literature (see, e.g., Maćkowiak et al., 2023). By allowing firms to observe an infinite history of signals initially, we ensure that their signal extraction problem is initialized in steady state.

[19] Due to issues related to the invertibility of matrices, it will be useful to have $\beta_1 = \varepsilon$, where ε is a small rather than actual zero. We abstract from this issue here, as it does not affect any of the results that follow.

Weight). It is the dot-product of the transpose of the ith column of the Leontief inverse $(\mathbf{I}-\mathbf{A})^{-1} = \mathbf{I}+\mathbf{A}+\mathbf{A}^2+...$, in which element (i,j) captures the direct and indirect importance of sector j as a supplier for sector i, and β, which measures final sales shares. We later discuss the relationship between Λ_i and our measure of *upstreamness* from Section 2.

Second, although our information structure is simple, it allows us to highlight two distinct channels that cause attention to center on the downstream sector. As we will show, the first channel is due only to the downstream sector being closer to the final consumer, and hence closer to final demand. This channel operates even without the presence of the *endogenous signals* in equation (14). The second channel, by contrast, is due to the downstream sector also being a better "information agglomorator" than the upstream sector, and rests on the presence of endogenous information. We discuss these channels in detail further below.

Lastly, notice that *aggregate value-added* in our economy is equal to consumption C_t of the downstream good. This is the measure of output that we focus on.[20]

6 Equilibrium and Solution

We proceed to characterize the equilibrium of the model. To do so, we first derive household and firm optimality conditions. We then use these to highlight the two different channels by which attention choices $\{m_i\}$ center on the downstream sector.

6.1 Optimality Conditions

We start with optimality conditions related to the labor and goods markets. We solve and discuss firms' optimal attention choices in the next subsection.

We solve the firm's two-stage problem by backward induction. Conditional on its labor and attention choices, in the *second stage*, the firm maximizes profits in equation (9), conditional on the production technology in equation (8), by equating the marginal product of the intermediate good with its marginal cost:

$$\alpha_{ij}\frac{P_{it}Q_{it}}{X_{ijt}} = P_{jt}. \tag{15}$$

[20] The market clearing relationships for goods are, respectively,

$$\begin{aligned} C_t + X_{ddt} &= Q_{dt} \quad \text{(downstream)} \\ X_{uut} + X_{udt} &= Q_{ut} \quad \text{(upstream).} \end{aligned}$$

In the *first stage*, firms internalize these choices when determining their optimal labor demand. As a result, the labor input is set so as to maximize expected profits $\mathbb{E}_{it}[\Pi_{it}]$ after the realization of information, where $\mathbb{E}_{it}[\cdot] = \mathbb{E}[\cdot \mid \Omega_{it}]$ denotes firm i's conditional expectation. This is done by equating the expected marginal product of labor with its marginal cost, the real wage, which results in:

$$L_{it} = \delta_i \cdot \frac{\mathbb{E}_{it}[P_{it}Q_{it}]}{\mathbb{E}_{it}[W_t]}. \tag{16}$$

In contrast to firms, the household only makes a decision in the second stage. It sets its labor supply until the marginal utility of consuming the real wage equals the marginal disutility of working,

$$L_t^{\frac{1}{\nu}} = W_t, \tag{17}$$

where $L_t = \sum_i L_{it}$ is total labor supplied to the upstream and downstream sectors.

6.2 Equilibrium Conditions

We can use the above optimality conditions to derive the equilibrium of the economy. As in Chahrour et al. (2021), the market-clearing conditions allow us to substitute firm revenue in (16) with economy-wide output, C_t. In order to solve for firms' labor choices, what remains is to solve for C_t. Since labor inputs are chosen in the first stage, labor can be treated as a fixed factor in the second stage. Appendix B.1 shows that, conditional on the first-stage labor choices, the (log of) the economy-wide output can be written as:

$$c_t = \Lambda' (\mathbf{z}_t + \varkappa_0 + \varkappa_1 \ell_t), \tag{18}$$

where $\mathbf{z}_t \equiv (z_{ut}, z_{dt})'$, $\ell_t \equiv \log(L_{ut}, L_{dt})'$, and the coefficients $\varkappa_0 \equiv \left[\sum_j \alpha_{ij} \log \alpha_{ij}\right]_i - \operatorname{diag}(\delta_i + \gamma) \log \Lambda$ and $\varkappa_1 \equiv \operatorname{diag}(\delta_i) > 0$. Inserting equation (18) back into the first-stage labor choice in condition (16), along with the expression for the equilibrium wage, then results in sector i's optimal labor choice:

Proposition 1. *The optimal labor choice for sector* $i = \{u, d\}$ *satisfies:*

$$L_{it} = (\delta_i \lambda_i) \cdot \frac{\mathbb{E}_{it}[\exp(\Lambda' [\mathbf{z}_t + \varkappa_0 + \varkappa_1 \ell_t])]}{\mathbb{E}_{it}\left[\left(\sum_j \exp(\ell_{jt})\right)^{1/\nu}\right]}, \tag{19}$$

where $\Lambda = [\lambda_i]_i = (\mathbf{I} - \mathbf{A}')^{-1} \beta$, $\ell_t = \log\left(L_{ut}, L_{dt}\right)'$, *and* $\varkappa_0$ *and* $\varkappa_1$ *are defined above.*

Proposition 1 characterizes firms' optimal labor choices. Given attention choices, $\{\mathbf{m}_i\}$, equilibrium labor choices $\{L_i\}$ are given by the fixed-point of equation (19). Importantly, however, Proposition 1 also helps characterize the drivers of firms' attention choices, $\{\mathbf{m}_i\}$, and thus the consequences of sentiment-driven fluctuations.

6.3 Optimal Attention Choices

In this subsection, we examine two forces that drive firms to focus their attention on the downstream sector. This increased focus, in turn, causes variations in downstream-sector sentiment, ξ_{dt}, to play a larger role for business-cycle fluctuations. Before discussing the decomposition of attention choices into these two forces, we, however, briefly turn to the question of which firms have a large incentive to pay more attention altogether. For simplicity, we focus on the case in which $\nu \to \infty$ and the real wage stays constant.

Proposition 2. *Let* $\tilde{x}_t \equiv \log X_t - \log X$ *denote the log-deviation of a variable from its value at the non-stochastic solution, and let* $\Pi(\cdot)$ *denote the profit function of a firm in terms of log-deviations. Suppose, further, that* $\nu \to \infty$*. Then, a second-order approximation shows that:*

$$\mathbb{E}\left[\Pi\left(\tilde{l}_{it}, \tilde{c}_t\right) - \Pi\left(\tilde{l}_{it}^{\star}, \tilde{c}_t\right)\right] \propto -\lambda_i^2 \delta_i \cdot \mathbb{E}\left[(\tilde{c}_t - \mathbb{E}_{it}[\tilde{c}_t])^2\right], \tag{20}$$

where $\tilde{l}_{it}^{\star}$ *denotes a firm in sector* i*'s ideal labor choice under full information.*

Proposition 2 characterizes the determinant of a firm's attention choice—the difference between the ex-ante profits that the firm achieves and that which it could achieve if it made a fully-informed input choice—and shows that this differential depends on the mean-squared error of the firm's aggregate demand expectation. Notice that this mean-squared error is multiplied by $\delta_i \lambda_i^2$. As a result, any loss of accuracy, due to a lack of attention, matters more for firms who (i) use labor more intensively in production (i.e., have a high δ_i); and (ii) are more central in the production network (i.e., have a large value of λ_i). This is because these firms, in effect, use labor more intensively in production—either directly or indirectly—and, as a result, have a larger benefit from paying closer attention and more accurately choosing their labor input.

1. Attention and Downstreamness: The first reason firms, all else equal, prefer to pay attention to the downstream sector can be seen from Proposition 2 and equation (18). Notice that, in (18), the expectation of sectoral productivity, $\mathbf{z}_t$, is multiplied by the vector of centrality weights, Λ. The log of consumption demand in the economy, and hence the demand for a firm's output, depends on the dot-product between the vector of centrality weights, Λ, and the vector of sector-specific productivity shocks, $\mathbf{z_t}$. This result carries over to substantially more general frameworks than the model we consider here, due to Hulten's celebrated theorem (Hulten, 1978). Because of this dot-product, firms in our economy have an incentive to pay more attention to more central sectors (i.e., those that have a large λ_i). Productivity fluctuations in these sectors, all else equal, move aggregate demand around by more, and thus have the potential to create larger errors in a firm's input choice. By paying closer attention (i.e., choosing a smaller value of m_i) to more central sectors, an individual firm is, thus, better able to align its input choice with the overall demand for its product and achieve higher profits.

Corollary 1 connects a firm's network centrality measure, λ_i, to our Antràs et al. (2012)-based measure of sectoral upstreamness, U_i

Corollary 1. *Let* $U \equiv [U_u, U_d]' = \left[(\mathbf{I} - \mathbf{A}')^{-1} \cdot \Lambda\right] \otimes \Lambda^{-1}$ *denote sectoral upstreamness, where* $\Lambda = [\lambda_u, \lambda_d]' = (\mathbf{I} - \mathbf{A}')^{-1} \beta$ *is network centrality. Then,* $U_d < U_u$ *and* $\lambda_d > \lambda_u$ *iff.* $\alpha_{uu} + \alpha_{du} < 1$.

Corollary 1 shows that the downstream sector is both more "downstream" and more "central" when the sum of the upstream sector's factor shares used in production is less than one. This condition is always satisfied in any plausible calibration of our model. Consequently, firms in both the upstream and downstream sectors naturally pay greater attention to the downstream sector, as fluctuations in downstream-sector productivity have a larger impact on aggregate demand, all else equal.

Strategic Interactions: The influence of downstreamness on firms' sectoral attention choices is modulated by the degree of strategic interactions between them. Proposition 1 shows that there are two channels by which other firms' input choices affect a given firm's labor decision, and hence its attention choice. The numerator in equation (19) shows that labor choices are, all else equal, *strategic complements* as $\varkappa_1 > 0$: when other firms increase labor demand, this increases household income, and thus the overall demand for goods in the economy (equation 18). By contrast, the denominator in equation (19) shows that labor choices can also be *strategic substitutes*: when other firms increase labor demand, this increases the real wage when $\nu \in \mathbb{R}_+$ and dampens the incentive to employ labor. Similarly to Chahrour et al. (2021), within our framework, the increase in demand dominates that from the real wage for all $\nu > 1$; that is for all standard calibrations of the Frisch elasticity. The presence of strategic complementarities, in this case, further amplifies firms' preference to pay attention to the downstream sector through the demand-side channel (Proposition 2). As in Hellwig and Veldkamp (2009), strategic complementarities in actions drive strategic complementarities in information choice.

2. Attention and Information Agglomeration: The second reason firms can prefer to pay attention to the downstream sector follows from the information structure. Each period, firms in sector i receive a signal about previous period's output in sector j (equation 14). These signals are informative because they reveal information about the previous period's (and hence current period's) productivity. But now notice that the downstream sector utilizes the upstream sector's input , while the converse does not occur:

$$q_{dt} = \log z_{dt} + \alpha_{dd} x_{dd,t} + \alpha_{du} x_{du,t} + \delta_d l_{dt} \tag{21}$$

$$q_{ut} = \log z_{ut} + \quad 0 \quad + \alpha_{ud} x_{uu,t} + \delta_u l_{ut}. \tag{22}$$

Thus, the downstream sector's output, q_{dt}, embeds information about the upstream sector's productivity, z_{ut}, through the downstream sector's acquisition of upstream sector goods, $X_{du,t}$. This implants additional information into the downstream sector's output, and makes its output—as we will see shortly below—a more "information-cost efficient" signal to pay attention to. In this sense, downstream-sector output plays a similar role to that of market-clearing prices in Grossman and Stiglitz's famous analysis (Grossman and Stiglitz, 1980): it agglomerates the dispersed bits of imperfect information that exist about fundamentals in the economy, which causes firms' attention to, all else equal, center on it.

7 Quantitative Analysis

In this section, we quantitatively evaluate the business-cycle effects of sentiment shocks on firms. Our analysis demonstrates that attention gravitates towards firms that are more downstream, as they are more central to the economy and serve as better agglomerators of dispersed information. Indeed, we find that sentiment shocks to downstream firms drive close to 70 percent of sentiment-driven fluctuations, and around 20 percent of the overall business cycle. We close the section with an extension of our baseline framework that allows for the 27 sectors used in the Atalay (2017)'s sectoral definitions for the U.S. economy. Our results, on balance, confirm those from the baseline model.

7.1 Numerical Solution and Parametrization

Numerical Solution: The limited attention framework that we have laid down does not permit an analytical solution. Instead, we solve the model numerically, using a parameterized expectations algorithm akin to that proposed in Chahrour et al. (2021). Solving the model requires finding values for the loadings of the expectations in Proposition 1 onto current and past signals, as well as firms' attention choices $\{\mathbf{m}_i\}$, which are consistent with firm optimality, Bayesian updating of expectations, and market clearing. We do so by first truncating the set of past signals we consider. We truncate the signal vector $\mathbf{s}_{it}$ at $\mathbf{s}_{i,t-H}$, where $H = 20$. That said, our numerical results are already stable from around $H = 10$. We then iterate on the following two steps until convergence.

First, we keep firms' attention choices $\{\mathbf{m}_i\}$ fixed and derive equilibrium labor choices $\{L_{it}\}$ and expectations in (19) by solving them with the help of the parameterized expectations algorithm. We iterate on firms' labor choices until convergence. Second, we keep firms' labor choices $\{L_{it}\}$ fixed and derive new values for firms' optimal attention choices $\{\mathbf{m}_i\}$. We do so by deriving expressions for firms' *ex-ante* profits as a function of their attention choices. We then maximize these expressions with respect to firms' attention vectors $\{\mathbf{m}_i\}$. We halt the iteration between the two steps when the discrepancy between the set of attention choices in two consecutive iterations is small in terms of the absolute difference. Appendix B.3 contains further details on the implementation of the algorithm.

Parametrization: We set $\gamma = 0.10$, consistent with modern estimates for the share of decreasing returns at the firm level, and set $\nu = 5$ (Rogerson and Wallenius, 2009).[21] We rank the 63 sectors in the BEA sectoral definition (excl. the public sector) according to their measure of upstreamness. We let $\mathbf{A}$ be determined by the average input shares among sectors that are above and below median upstreamness, respectively. The upstream sector's factor shares, as such, equal those for sectors that are above the median measure of upstreamness. The vector $\beta \equiv \begin{pmatrix} 1 & 0 \end{pmatrix}'$ by assumption. We take standard values for the productivity process: we set $\sigma_\vartheta = 0.005$, $\sigma_u = 1.00$ and $\rho_\vartheta = 0.80$. These values are all within the range used in standard dynamic stochastic general equilibrium models.

Our baseline parametrization, in addition, targets moments of our Granular Sentiment Index, $\mathcal{S}_t$—in particular, we set the autocorrelation and standard of innovations in line with the corresponding components of the data ($\rho_\xi = 0.54$ and $\sigma_\xi = 0.20$). For the attention cost function, we use $K(\mathbf{m}_i) = \mu \sum_n \mathbf{m}_i[n]^{-2}$, $n = \{1, 2, 3\}$; that is, a marginal cost of information, μ, multiplied by the sum of signal precisions, $\mathbf{m}_i[n]^{-2}$ (Veldkamp, 2023). The free parameter μ in this expression determines the extent of limited attention. For example, if μ is equal to zero, our economy collapses back to its full-information counterpart, as firms can obtain infinitely precise signals at zero cost. We set μ such that

[21] We choose a comparatively large value of the Frisch elasticity, as labor is the only factor of production in our economy subject to information frictions. Clearly, the overall factor supply subject to information frictions is larger than that attributable to labor only. In a similar framework to that studied here, Asriyan and Kohlhas (2024) show that a Frisch elasticity equal to 5 corresponds roughly to the overall elasticity of factor supply that one would derive from an economy in which both capital as well as relatively inelastic labor are used in production and both are subject to information frictions.

Table 3: Attention Choices in the Baseline Model

Sector	Std. TFP	Λ	Attention Choices
Downstream	1.30	1.17	(1.80, 3.10, 2.20, 10.05)
Upstream	1.30	0.17	(2.80, 10.30, 3.00, 3.15)

Notes: This table summarizes general-equilibrium attention allocation across the two sectors. Columns 2-4 represent, in order, the standard deviation of sectoral productivity shocks, the measure of production network centrality, and the attention choice vector $\{\mathbf{m}_i\}$, listed in order $(m_d^z, m_u^z, m_d^q, m_u^q)$.

we match the average accuracy (measured by the root-mean-squared error) of firms' GDP forecasts in the Duke-CFO Survey (Graham et al., 2020). This results in $\mu = 2.5e^{-4}$.

7.2 Implied Attention Choices

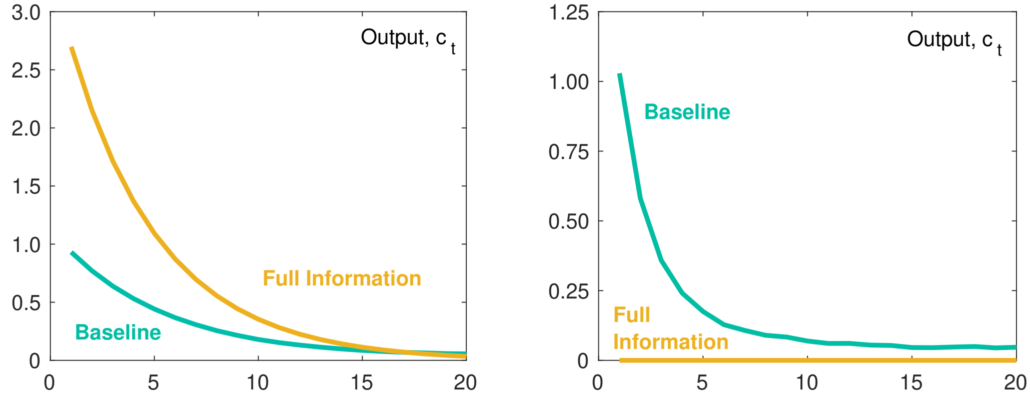
Recall that attention choices, $\{\mathbf{m_i}\}$, center on the downstream sector (i) because it is more central in the production network; and (ii) because its output better agglomerates sectoral information. Table 3 demonstrates these mechanisms in general equilibrium.

The table shows that, despite the equal volatility of productivity shocks, across both sectors, firms' attention gravitates towards the downstream sector. Indeed, in our baseline calibration, firms in the upstream sector pay closer attention to downstream productivity than their *own* sector's productivity ($m_d^z < m_u^z$ in column one).[22] This is, in part, because the downstream sector is around six times more central in the production network, as measured by the discrepancy between λ_d and λ_u. Finally, Table 3 shows that the upstream sector also pays a comparable amount of attention to the downstream sector's output as it does to its own productivity ($m_d^q \approx m_u^z$). This is in line with our earlier discussion demonstrating that the downstream sector's output better agglomerates sectoral information about productivity.

On balance, the results in Table 3 are consistent with our empirical findings in Section 2: because of the closer attention paid to the downstream sector, innovations to sectoral information about the downstream sector, all else equal, have a greater impact on firms' expectations about aggregate productivity and output. Any erroneous information about downstream sector productivity—i.e., any sentiment shock—, as a result, also has larger effects. We explore the implications of this asymmetry in attention below.

[22] Specifically, in our baseline calibration, firms in the upstream sector pay around 3 times more attention to downstream firms' productivity ($m_d^z = 2.80$ vs $m_u^z = 10.30$).

Figure 10: Output Responses to Productivity and Sentiment Shocks



Panel (a): Productivity   Panel (b): Downstream Sentiment

Notes: The left panel showcases the impulse response function of output, c_t, to a one-standard deviation shock to aggregate productivity, ϑ_t. The figure depicts this both for the benchmark economy with limited information and for its perfect-information counterpart in which $\mu = 0$. The right figure instead showcases the response of c_t to a one-standard deviation shock to the sentiment of the downstream sector, ξ_{dt}. All responses are in percentage terms.

7.3 Shock Propagation and Sentiment

We start with innovations to economy-wide productivity, ϑ_t. Because of limited attention ($\mu > 0$), the response of the economy to fundamental productivity shocks is dampened. The left panel of Figure 10 demonstrates this result by comparing the output response in our economy to a one-standard deviation increase in economy-wide productivity, ϑ_t, with that under full information ($\mu = 0$). As in Sims (2003), a positive productivity shock raises output upon impact. However, due to the presence of information frictions, the output response under limited attention is dampened and substantially more persistent.

The flip-side of the decreased responsiveness to fundamental productivity shocks is an increased responsiveness to sentiment shocks. The right panel of Figure 10 showcases the output response to a one-standard-deviation increase in sentiment towards the downstream sector, ξ_{dt}, and compares it to the full-information counterfactual, in which there is no response. The effects of a sentiment shock on the upstream sector are similar, albeit more muted. In response to a positive sentiment shock, firms across the economy believe the downstream sector to be more productive (see equation 12). This, in turn, increases overall demand. Thus, firms in *both* sectors increase the demand for inputs and increase production, which in turn raises the equilibrium level output in the economy—a short-lived boom occurs similar to that which we estimated in Section 4. However, as firms in both sectors start to learn from the observation of additional signals that downstream productivity has, in fact, not increased, they reverse their earlier decisions and the boom

Table 4: Business Cycle Attribution

	Share of Output Variance $\mathbb{V}[y_t]$		
	Benchmark	Full Information	27 Atalay Sectors
Productivity	0.81	1.00	0.81
Sentiment	0.19	0.00	0.19
– downstream sector	0.18	0.00	0.13
– upstream sector	0.01	0.00	0.06
Total N of sectors	2	2	27
N of downstream sectors	1	1	5
N of upstream sectors	1	1	22

Notes: The table shows the decomposition of economy-wide output fluctuations into productivity shocks and sectoral sentiment shocks. The table includes estimates from an extension of our baseline model that includes 27 different sectors, in accordance with the Atalay (2017)-sectoral definitions.

slowly subsides back down.

In sum, Figure 10 shows that both fundamental and sentiment-driven fluctuations in output can arise in the equilibrium of our model. We next turn to a decomposition of the relative importance of the two types of disturbances for business-cycle fluctuations. We further decompose sentiment-driven fluctuations into those that originate from the downstream and the upstream sectors, respectively.

7.4 Granular Sentiment and Business Cycles

We leverage our calibrated model to quantitatively assess the share of business-cycle fluctuations driven by sector-specific sentiment shocks, ξ_{it}. We compare the results under our baseline calibration to those under full information, in which $\mu = 0$, in Table 4. The middle column shows that, under full information, all fluctuations in output are caused by fundamental disturbances—in this case, productivity shocks. This contrasts with the results under our baseline calibration in which around 1/5 of business-cycle fluctuations are driven by changes in sentiment. This estimate for the share of business-cycle fluctuations attributable to innovations in sentiment towards *firms* is at the lower-end of the range of existing estimates in the literature that studies the consequences of *economy-wide* sentiment shocks (e.g., Blanchard et al., 2013 and Chahrour and Ulbricht, 2023). These works, on average, find that around 1/6-1/3 of business-cycle fluctuations can be attributed to sentiment shocks.

However, crucially, Table 4 also provides the decomposition of sentiment-driven fluctuations into those that originate from the downstream versus the upstream sector, respec-

Table 5: Business Cycle Attribution—Decomposition

	Share of Output Variance $\mathbb{V}[y_t]$			
	Two Sector Economy		Multi Sector Economy	
	Benchmark	Symmetric I/O	Benchmark	Symmetric I/O
Productivity	0.81	0.97	0.81	0.95
Sentiment	0.19	0.03	0.19	0.05
– downstream sector	0.18	0.02	0.13	0.03
– upstream sector	0.01	0.01	0.06	0.02
Total N of sectors	2	2	27	27
N of downstream sectors	1	1	5	5
N of upstream sectors	1	1	22	22

Notes: The table shows the decomposition of economy-wide output fluctuations into productivity shocks and sectoral sentiment shocks. The table includes estimates from the symmetric I/O model.

tively. Column one of the Table shows that the lion's share of the overall effect of sentiment shocks is due to sentiment shocks to the downstream sector. Indeed, close to 95 percent of overall sentiment-induced fluctuations are due to fluctuations in sentiment towards the downstream sector. This showcases one important consequence of firms' asymmetric attention choices. Despite the equal volatility of productivity, firms across both sectors pay closer attention to the downstream sector. Yet, the flip-side of this increased attention towards the downstream sector is that downstream-specific sentiment shocks are substantially more potent, driving almost all sentiment-induced output dynamics.

We can further decompose the consequences of the asymmetric attention choices in Table 4 into the two aforementioned forces: (i) the differences in network centrality; and (ii) the differences in the agglomeration of information. To do so, we assume a symmetric input-output structure, equal to the average of the upstream and downstream sector, but still feed in the differential informativeness of sectoral output from the earlier calibration. We then re-run our analysis. Consistent with our earlier discussion, we find that attention choices are now less asymmetric. The consequences can be seen in Table 5. Overall, sentiment shocks now drive substantially less of output fluctuations. As can be seen from columns one and two, the relative share attributable to the downstream sector has, furthermore, now declined from around 95 percent to 66 percent. The remaining share of the asymmetry (16 pp. from full equality) is due to the differing informativeness of sectoral output.

Clearly, results in the first column of Table 4 provided only a first pass at a quantitative assessment of the effects of sector-specific sentiment shocks. Our framework allows for an intensive margin in the attention allocated to more and less downstream sectors. But our results are limited in that we have only so-far considered a two-sector economy. Consequently, sector-specific shocks have effects similar to those of economy-wide shocks, and there are natural limits to the asymmetry in attention attached to different sectors. To address these concerns, we now turn to an extension of our baseline framework which allows for a more detailed production network with *many* sectors.

7.5 Sentiment Shocks in a Multi Sector Model

We consider an extension of our baseline framework to the 27 different sectors studied in Atalay (2017). We cross-walk from the BEA industries used in our empirical analysis to the 27 sectors in Atalay (2017) and calibrate the input-shares matrix $\mathbf{A}$ to the resulting input shares from the production network that arises. We also take the final sales vector β from this exercise. Finally, we recalibrate the cost of information to once more target the average accuracy from the Duke-CFO Survey. All other details of the calibration are identical to those described earlier. The third column of Table 4 presents the results.

The decrease in the size of sectors (as a result of the increase in numbers) has two opposing effects on our quantitative results. On the one hand, sector-specific disturbances, all else equal, become less important.[23] This, in turn, makes any sector-specific disturbance—including sentiment shocks—less important for output fluctuations. On the other hand, however, as we increase the number of sectors, we also increase differences in measures of downstreamness and network centrality. This, in turn, elevates the extent of asymmetric attention, as firms economize on attention costs by focusing only on a handful of sectors, which increases the effects of sentiment shocks towards these sectors.

The third column of Table 4 shows that, on average, these two forces are close to canceling each other out in equilibrium. Sentiment shocks still account for around 19 percent of the variation in output, 13 percentage points of which are due to sentiment shocks to the 5 most downstream sectors. Sentiment shocks towards the remaining 22 sectors contribute only around 6 percentage points to output fluctuations. The sentiment-induced business cycle is, in this sense, *granular*: roughly 20 percent of sectors are responsible for 70 percent of the sentiment-driven output variation. The ubiquitous "Pareto Principle" (the 80-20 rule), in this sense, also provides a decent proxy for sentiment-driven fluctuations.[24] Fi-

[23] Clearly, in the limit in which the number of sectors tends towards infinity, sector-specific shocks do not affect economy-wide outcomes.

[24] The Pareto Principle would, all else equal, stipulate that 80 percent of fluctuations would be caused by 20 percent of sectors.

Table 6: Sensitivity Analysis: Output Fluctuations

	Sentiment/Downstream	*Frisch Elasticity* 3.00	5.00	10.00
	33 % less	0.04/0.02	0.09/0.04	0.17/0.13
λ	Baseline model	0.10/0.05	0.19/0.13	0.28/0.19
	33 % more	0.08/0.07	0.20/0.15	0.25/0.21

Notes: The table shows the share of output fluctuations, $\mathbb{V}[y_t]$, attributable to (downstream sector) sentiment shocks. The table computes these quantities for various model parameters detailed in the main text

nally, the fourth column of Table 5 shows that the main reason for this asymmetry is once more the heterogeneity that exists in the production network.

All in all, the results in Table 4 are also comfortingly consistent with our earlier empirical findings, showing that sentiment towards a small set of firms can cause sentiment-induced aggregate fluctuations. Our results in this section have further shed light on the precise mechanisms through which these aggregate fluctuations arise.

7.6 Discussion and Extensions

The literature studying the effects of microeconomic shocks on macroeconomic outcomes commonly views data through the lenses of power-law densities (e.g., Gabaix, 2009). We can use the above attribution estimates—the share of sentiment fluctuations that can be attributed to downstream firms—to infer the approximate degree of "sentiment granularity" in the data even *without* a complete measure of the characteristic itself. In a standard model, one would fit a Pareto (α) distribution on the characteristic, e.g. firm size as measured by sales or total assets, and recover an estimate $\hat{\alpha}$. Generally, an $\hat{\alpha}$ of below 2 corresponds to a thick-tailed density in which the variance is not defined and standard laws of large numbers break down (Gabaix, 2011). In addition, an estimate of 1.16 corresponds exactly to the 80-20 Pareto rule. In our baseline multi-sector model, we uncover a 70-20 pattern, which is, all else equal, consistent with an $\hat{\alpha}$ of somewhere in the $(1.2, 2)$ region. Sentiment-driven fluctuations are, therefore, granular but less concentrated than what a standard Pareto principle would indicate, albeit only moderately.

That said, the precise share attributable to the downstream sector clearly depends on details of the network structure—in particular, the distribution of network centrality weights, $\{\lambda\}_i$—and hence the level of aggregation. Table 6 shows the sensitivity of our attribution exercise with respect to changes in input-output structure. To do so, we stretch (or dampen) the distances in λ_i between pairs of sectors by 1/3[25]. We focus on the extended 27-sector model, where we once more classify the 5 most downstream sectors as the "downstream". The Table further shows the consequences of changes to the supply elasticity of labor, governing how easily the economy adjusts to new information. Across the different calibrations, sentiment-driven fluctuations attributable to the downstream sectors account for around 50-88 percent of overall sentiment-driven fluctuations. The more asymmetric the network structure is, the larger, all else equal, the share attributable to the downstream sector. In all cases, fewer than 20 percent of sectors account for the dominant share of sentiment-driven fluctuations. As with the propagation of fundamental disturbances, such as productivity shocks, some sectors matter substantially more than others for sentiment-driven dynamics.

8 Conclusion

Because macroeconomic outcomes, such as output, are combinations of more disaggregated variables, it is reasonable to conjecture that macroeconomic fluctuations have their origins in idiosyncratic, microeconomic shocks to individual sectors or agents (Acemoglu et al., 2017). Sentiment-induced fluctuations are, in this respect, no different. While the "micro origins" hypothesis has been applied to fundamental shocks like productivity, an application to sentiment-driven fluctuations has so far been missing. In this paper, we have attempted to fill this gap.

Our focus—both empirically and theoretically—has been on firms and how the sentiment of a comparatively small number of these about the future drives aggregate expectations and fluctuations. We have demonstrated how the sentiment of a small subset of downstream firms correlates closely with the overall state of the economy, and how innovations to the sentiment of these firms themselves help drive business-cycle fluctuations. Indeed, in a quantitative-theoretical multi-sector economy, calibrated to match our firm-level dataset on sentiment, we found that around 1/5 of output fluctuations can be attributable to innovations in the sentiment of the 20% most downstream firms. Our model suggests that around 70% of overall sentiment-induced fluctuations is driven by sentiment towards only the 20% most downstream firms.

One natural direction for future research is an extension of our framework, and the types of "Pareto principle" computations, to other agents in the economy: e.g., households or investors. Beliefs of a small subset of these could, in principle, likewise matter for aggregate fluctuations. Another is to expand the list of model outcomes considered. As

[25] We do so by changing the final goods share vector, β.

we have shown empirically, sentiment-driven fluctuations also help drive unemployment dynamics, among others, although the mechanisms that operate on these variables may differ from those we have highlighted for output. We leave this for future research.

Finally, as mentioned at the start of this paper, the design and conduct of optimal policy rely crucially on the accurate measurement of the true origin of economic fluctuations. Our research concludes that it is proximity to the final consumer that makes a firms' sentiment important for the business cycle. Regulation of information disclosure by listed companies should, as a consequence, pay particularly close attention to central, downstream firms, as disclosures by such informationally granular firms, all else equal, have larger impacts on the broader economy.

References

Acemoglu, Daron, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "Systemic Risk and Stability in Financial Networks," *The American Economic Review*, 2015, *105.*

_ , _ , and _ , "Microeconomic Origins of Macroeconomic Tail Risks," *The American Economic Review*, 2017, *107.*

_ , Vasco Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "The Network Origins of Aggregate Fluctuations," *Econometrica*, 2012, *80* (5).

Acharya, Sushant, Jess Benhabib, and Zhen Huo, "The Anatomy of Sentiment-driven Fluctuations," *Journal of Economic Theory*, 2021, *195*, 105280.

Afrouzi, Hassan and Saroj Bhattarai, "Inflation and GDP Dynamics in Production Networks: A Sufficient Statistics Approach," Working Paper 31218, National Bureau of Economic Research May 2023.

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

_ , Fabrice Collard, and Harris Dellas, "Quantifying Confidence," *Econometrica*, 2018, *86* (5), 1689–1726.

_ , _ , and _ , "Business-cycle Anatomy," *American Economic Review*, October 2020, *110* (10), 3030–70.

Antràs, Pol, Davin Chor, Thibault Fally, and Russell Hillberry, "Measuring the Upstreamness of Production and Trade Flows," *American Economic Review*, May 2012, *102* (3), 412–16.

Araci, Dogu T, "FinBERT: Financial Sentiment Analysis with Pre-trained Language Models," *Manuscript*, 2019.

Asriyan, Vladimir and Alexandre Kohlhas, "The Macroeconomics of Firm Forecasts," *Oxford University Working Paper*, 2024.

Atalay, Enghin, "How Important Are Sectoral Shocks?," *American Economic Journal: Macroeconomics*, October 2017, *9* (4), 254–80.

Baker, Scott R, Nicholas Bloom, and Steven J Davis, "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 07 2016, *131* (4), 1593–1636.

Baqaee, David R and Emmanuel Farhi, "The Macroeconomic Impact of Microeconomic Shocks: beyond Hulten's Theorem," *Econometrica*, 2019, *87* (4), 1155–1203.

Barsky, Robert B and Eric R Sims, "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence," *American Economic Review*, 2012, *102* (4), 1343–1377.

Beaudry, Paul and Franck Portier, "An Exploration into Pigou's Theory of Cycles," *Journal of Monetary Economics*, 2004, *51* (6), 1183–1216.

_ and _ , "Stock Prices, News, and Economic Fluctuations," *American Economic Review*, September 2006, *96* (4), 1293–1307.

Bigio, Saki and Jennifer La'O, "Distortions in Production Networks," *The Quarterly Journal of Economics*, 05 2020, *135* (4), 2187–2253.

Blanchard, Olivier J, Jean-Paul L'Huillier, and Guido Lorenzoni, "News, Noise, and Fluctuations: an Empirical Exploration," *American Economic Review*, December 2013, *103* (7), 3045–70.

Bordalo, Pedro, Nicola Gennaioli, Andrei Shleifer, and Stephen J Terry, "Real Credit

Cycles," *NBER Working Paper*, 2021.

__ , __ , **Rafael La Porta, Matthew O'Brien, and Andrei Shleifer**, "Long Term Expectations and Aggregate Fluctuations," Working Paper 31578, National Bureau of Economic Research August 2023.

Carvalho, Vasco and Basile Grassi, "Large Firm Dynamics and the Business Cycle," *American Economic Review*, April 2019, *109* (4), 1375–1425.

__ **and Xavier Gabaix**, "The Great Diversification and its Undoing," *American Economic Review*, 2013, *103(5).*

Chahrour, Ryan and Kyle Jurado, "News or Noise? the Missing Link," *American Economic Review*, July 2018, *108* (7), 1702–36.

__ **and** __ , "Recoverability and Expectations-Driven Fluctuations," *The Review of Economic Studies*, 03 2021, *89* (1), 214–239.

__ **and Robert Ulbricht**, "Robust Predictions for Dsge Models with Incomplete Information," *American Economic Journal: Macroeconomics*, 2023, *15* (1), 173–208.

__ **, Kristoffer Nimark, and Stefan Pitschner**, "Sectoral Media Focus and Aggregate Fluctuations," *American Economic Review*, 2021, *111* (12), 3872–3922.

__ **, Vito Cormun, Pierre De Leo, Pablo A Guerrón-Quintana, and Rosen Valchev**, "Exchange Rate Disconnect Revisited," *NBER Working Paper 32596*, 2024.

Chava, Sudheer, Wendi Du, and Nikhil Paradkar, "Do Managers Walk the Talk on Environmental and Social Issues?," *Manuscript*, 2020.

__ **,** __ **, and** __ , "More than Buzzwords? Firms' Discussions of Emerging Technologies in Earnings Conference Calls," *Manuscript*, 2021.

Chodorow-Reich, Gabriel, Andra Ghent, and Valentin Haddad, "Asset Insulators," *The Review of Financial Studies*, 05 2020, *34* (3), 1509–1539.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Manuscript*, 2018.

di Giovanni, Julian, Andrei A. Levchenko, and Isabelle Mejean, "Firms, Destinations, and Aggregate Fluctuations," *Econometrica*, 2014, *82* (4), 1303–1340.

Enders, Zeno, Michael Kleemann, and Gernot J Müller, "Growth Expectations, Undue Optimism, and Short-run Fluctuations," *Review of Economics and Statistics*, 2021, *103* (5), 905–921.

Fernald, John, "A Quarterly, Utilization-adjusted Series on Total Factor Productivity," *Federal Reserve Bank of San Francisco Working Paper*, 2014, *19.*

Flynn, Joel and Karthik Sastry, "The Macroeconomics of Narratives," *Working Paper*, 2022.

__ **and** __ , "Attention Cycles," *Working Paper*, 2024.

Gabaix, Xavier, "Power Laws in Economics and Finance," *Annual Review of Economics*, 2009, *1* (Volume 1, 2009), 255–294.

__ , "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 2011, *79* (3), 733–772.

Galaasen, Sigurd, Ragnar Juelsrud, Rustam Jamilov, and Hélène Rey, "Granular Credit Risk," *NBER Working Paper 27994*, 2023.

Gaubert, Cecile and Oleg Itskhoki, "Granular Comparative Advantage," *Journal of Political Economy*, 2021, *129* (3), 871–939.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, "Text as Data," *Journal of Economic*

Literature, September 2019, *57* (3), 535–74.

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera, "The Voice of Monetary Policy," *American Economic Review*, February 2023, *113* (2), 548–84.

Graham, John, Brent Meyer, Nicholas B Parker, and Sonya Ravindranath Waddell, "Introducing the Cfo Survey," *Federal Reserve Bank of Atlanta*, 2020.

Grossman, Sanford J and Joseph E Stiglitz, "On the Impossibility of Informationally Efficient Markets," *The American economic review*, 1980, *70* (3), 393–408.

Hansen, Stephen and Michael McMahon, "Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication," *Journal of International Economics*, 2016, *99*, S114–S133.

_ , _ , and Andrea Prat, "Transparency and Deliberation within the Fomc: a Computational Linguistics Approach," *The Quarterly Journal of Economics*, 10 2017, *133* (2), 801–870.

Hassan, Tarek A, Jesse Schreger, Markus Schwedeler, and Ahmed Tahoun, "Sources and Transmission of Country Risk," *Review of Economic Studies*, 2023, *91(4).*

_ , Stephan Hollander, Laurence van Lent, and Ahmed Tahoun, "Firm-level Political Risk: Measurement and Effects," *The Quarterly Journal of Economics*, 08 2019, *134* (4), 2135–2202.

_ , _ , _ , and _ , "The Global Impact of Brexit Uncertainty," *Journal of Finance*, 2023, *79(1).*

Hellwig, Christian and Laura Veldkamp, "Knowing What Others Know: Coordination Motives in Information Acquisition," *The Review of Economic Studies*, 2009, *76* (1), 223–251.

Hulten, Charles R, "Growth Accounting with Intermediate Inputs," *The Review of Economic Studies*, 1978, *45* (3), 511–518.

Jamilov, Rustam, Hélène Rey, and Ahmed Tahoun, "The Anatomy of Cyber Risk," *NBER Working Paper 28906*, 2023.

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

Kohlhas, Alexandre N and Ansgar Walther, "Asymmetric Attention," *American Economic Review*, September 2021, *111* (9), 2879–2925.

Lagerborg, Andresa, Evi Pappa, and Morten O Ravn, "Does Economic Insecurity Really Impact on Gun Violence at Us Schools?," *Nature: Human Behavior*, 2019, *3.*

_ , _ , and _ , "Sentimental Business Cycles," *The Review of Economic Studies*, 08 2022, *90* (3), 1358–1393.

Liu, Ernest, "Industrial Policies in Production Networks," *The Quarterly Journal of Economics*, 08 2019, *134* (4), 1883–1948.

Lorenzoni, Guido, "A Theory of Demand Shocks," *American Economic Review*, December 2009, *99* (5), 2050–84.

Loughran, Tim and Bill McDonald, "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-ks," *Journal of Finance*, 2011, *66 (1).*

_ and _ , "Textual Analysis in Accounting and Finance: a Survey," *Journal of Accounting Research*, 2016, *54.*

Maćkowiak, Bartosz, Filip Matějka, and Mirko Wiederholt, "Rational Inattention: a Review," *Journal of Economic Literature*, 2023, *61* (1), 226–273.

Maxted, Peter, "A Macro-finance Model with Sentiment," *The Review of Economic Studies*, 03 2023, *91* (1), 438–475.

Pareto, Vilfredo, "Cours D'économie Politique. Tome Premier.," *Geneva: Librairie Droz*, 1896.

Pastén, Ernesto, Raphael Schoenle, and Michael Weber, "Sectoral Heterogeneity in Nominal Price Rigidity and the Origin of Aggregate Fluctuations," *American Economic Journal: Macroeconomics*, April 2024, *16* (2), 318–52.

Pigou, Arthur C, "Industrial Fluctuations," *New York: Macmillan*, 1927.

Ravn, Morten O and Harald Uhlig, "On Adjusting the Hodrick-prescott Filter for the Frequency of Observations," *The Review of Economics and Statistics*, 2002, *84* (2), 371–376.

Rogerson, Richard and Johanna Wallenius, "Micro and Macro Elasticities in a Life Cycle Model with Taxes," *Journal of Economic theory*, 2009, *144* (6), 2277–2292.

Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang, "Firm-level Climate Change Exposure," *The Journal of Finance*, 2023, *78* (3), 1449–1498.

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

van Binsbergen, Jules H, Svetlana Bryzgalova, Mayukh Mukhopadhyay, and Varun Sharma, "(Almost) 200 Years of News-Based Economic Sentiment," *NBER Working Paper 32026*, 2024.

Veldkamp, Laura L, "Information Choice in Macroeconomics and Finance," *Princeton University Press*, 2023.

Online Appendix for "Granular Sentiments"

Rustam Jamilov   Alexandre Kohlhas   Oleksandr Talavera   Mao Zhang

February 2025

Contents

A Empirical Appendix

A.1 Data Construction and Definitions

We obtain information on stock prices from the Center for Research in Security Prices (CRSP) and basic income statement and balance sheet data from Standard and Poors' Compustat. We begin with the description of firm-level variables. Firm size is defined as the log of total assets (log(atq)) and the log of sales (log(saleq)). Book-to-market ratio is computed as the ratio of total book common equity to the market value of common shares outstanding (ceqq/(prccq · cshoq)).

We compute market betas by running daily firm-level linear regressions of excess returns on the market factor and a constant. Excess returns are defined as the difference between firm (log) returns and the risk-free rate. The risk-free rate and the market factor are obtained from Kenneth French's website. Investment intensity is computed as the quarterly difference in the total (net) property plant and equipment value scaled by total assets ($(\mathrm{ppentq}_t - \mathrm{ppentq}_{t-1})/\mathrm{atq}_{t-1}$). Market value is defined in logs (log(mkvaltq)). The leverage ratio is defined as the ratio of the sum of long-term and short-term liabilities to total assets ((dlttq + dlcq)/atq). The liquidity ratio is defined as the ratio of cash and short-term investments to total assets (cheq/atq). Tobin's Q, which proxies future growth opportunities, is approximated with the following formula: (total assets - total equity book value + total equity market value) / total assets ($(\mathrm{atq} - \mathrm{ceq} + \mathrm{prccq} * \mathrm{cshoq})/\mathrm{atq}$). Realized stock volatility is obtained by computing 60 calendar-day rolling standard deviations of firm (log) returns.

We estimate firm-level productivity by first running linear regressions of (log) sales on a constant, (log) total assets, and (log) selling, general and administrative expenses (xsgaq), which we use to proxy the labor input. Firm productivity is then defined as the residual from this regression. To limit the influence of outliers, every variable has been winsorized at the 1% level in each quarter.

Both dictionary- and FinBERT-based firm-level sentiment measures $\xi^1_{i,t}$ and $\xi^2_{i,t}$ are scaled by the number of total words in each transcript and multiplied by 100,000. To limit the influence of outliers, every sentiment measure—including $\xi^1_{i,t}$, $\xi^2_{i,t}$, $\xi^3_{i,t}$ and every refined/robust measure—has been winsorized at the 2.5% level in each quarter.

We obtain total factor productivity data from Fernald (2014). Aggregate uncertainty is proxied by the Economic Policy Uncertainty (EPU) index of Baker et al. (2016). Real Gross Domestic Product (GDP), GDP per capita, industrial production, the Consumer Price Index (CPI), and unemployment rate were obtained from the St. Louis Federal Reserve online database.

The five existing sentiment indices are: the OECD Business Confidence Index (BCI)[1], the University of Michigan Index of Consumer Sentiment that is constructed in the Survey of Consumers[2], the Purchasing Managers Index by the Institute for Supply Management[3], the Sentix market sentiment index[4], and the news sentiment index that is based on nearly 200 years of newspaper data and is developed by van Binsbergen et al. (2024). The first principal component of the five aforementioned sentiment measures is the economy-wide sentiment index $\mathcal{K}_t$ in the main text.

All aggregate variables have been de-trended with the Hodrick-Prescott filter following Ravn and Uhlig (2002). An alternative de-trending approach involves residualizing every aggregate variable from the time fixed effect and the results using this robustness test are described in this Appendix.

We obtain the data on mass shootings in the U.S. from Lagerborg et al. (2022). The data is available over 2006q4-2018q4. We use the number of victims from mass shootings as our instrument. The variable has been HP-filtered. In order to limit the influence of outliers, such as the tragic 2017 Las Vegas shooting, we trim the instrument at the 5% level.

Table A.1 provides summary statistics for every key variable used in the paper.

[1] The index is publically available here.

[2] The index is publically available here.

[3] We have obtained the ISM PMI Services index from the Bloomberg Terminal. The underlying data can be accessed here.

[4] We have obtained the sentix index from the Bloomberg Terminal. The underlying data can be accessed here.

Table A.1: Summary Statistics

	Observations	Mean	Std. Dev.	Min	Max
		Firm Data			
Baseline Dictionary Sentiment, ξ_{it}^1	23387	1.858	1	-1.647	4.255
FinBERT Sentiment, ξ_{it}^2	23387	1.457	1	-1.931	5.631
Analyst Sentiment, ξ_{it}^3	23357	-0.299	1	-7.510	5.295
Sentiment, Absolute Frequency	23387	1.710	1	-1.679	4.268
Sentiment, Lagged	21832	1.865	1	-1.815	4.286
Sentiment, Firm FE	23386	0.000	1	-5.997	4.403
Sentiment, Firm Controls	19464	0.000	1	-5.476	4.446
Sentiment, Aggregate Controls	23386	0.000	1	-5.958	4.293
Sentiment, All Controls	19464	0.000	1	-5.435	4.338
Log Assets	23387	6.530	1	2.559	10.388
Log Sales	23384	5.76	1	-1.454	9.008
Market Beta	21595	3.189	1	-1.057	7.834
Book Market Ratio	23386	1.136	1	0.000	24.653
Investment Intensity	22897	0.136	1	-19.186	33.742
Market Value	23384	7.690	1	2.097	11.774
Leverage Ratio	22528	1.633	1	0.000	5.481
Liquidity Ratio	23387	0.966	1	0.000	6.676
Tobin's Q	23386	1.281	1	0.296	20.029
Return Volatility	21440	1.631	1	0.388	18.327
	Sectoral Data				
Upstreamness	802	2.360	0.670	1.000	3.466
	Aggregate Data				
Principal Component of Sentiments	61	0.037	1	-2.819	2.345
OECD Sentiment	61	0.004	1	-3.731	1.275
Michigan Sentiment	61	0.072	1	-2.381	2.134
ISM PMI Index	61	0.056	1	-3.510	1.751
Sentix Sentiment	61	-0.006	1	-3.728	1.371
News Sentiment	53	0.000	1	-2.578	2.006
Real GDP	61	-0.026	1	-6.039	1.214
Real GDP per capita	61	-0.03	1	-5.963	1.262
Unemployment Rate	61	3E-09	1	-1.363	5.555
Industrial Production	61	-0.04	1	-3.8	1.593
Uncertainty	61	0.035	1	-1.345	3.96
CPI	61	-0	1	-2.098	3.723
Mass Shootings	43	-0.453	1	-2.561	2.042

Notes: This table provides basic summary statistics for main variables used in the empirical analysis. See Appendix A.1 for details on data definitions and construction.

A.2 Details on FinBERT

To analyze sentiment that is contained in the texts of quarterly earnings conference calls, we also employ FinBERT: a pre-trained language model based on bidirectional encoder representations from a transformers model (BERT) for financial natural language processing tasks. Specifically, we extract word embeddings from the earnings call texts, utilizing BERT's contextual embeddings.

In traditional word embedding models like GloVe, each word is mapped to a single vector representation, which does not account for different meanings a word might have based on the context. In contrast, BERT generates contextualized embeddings for words, considering the meaning of word in the specific context. Furthermore, when compared to other contextual embeddings models like ELMo, which considers context sequentially by generating embeddings based on left-to-right and right-to-left contexts, BERT captures a broader context by considering all words in the sentence simultaneously. These features make BERT a highly effective and accurate tool for interpreting text (Devlin et al., 2018). The BERT class of models has been applied to multiple areas of economic research, ranging from central banking (Gorodnichenko et al., 2023) and environmental economics (Chava et al., 2021) to technology and innovation (Chava et al., 2020).

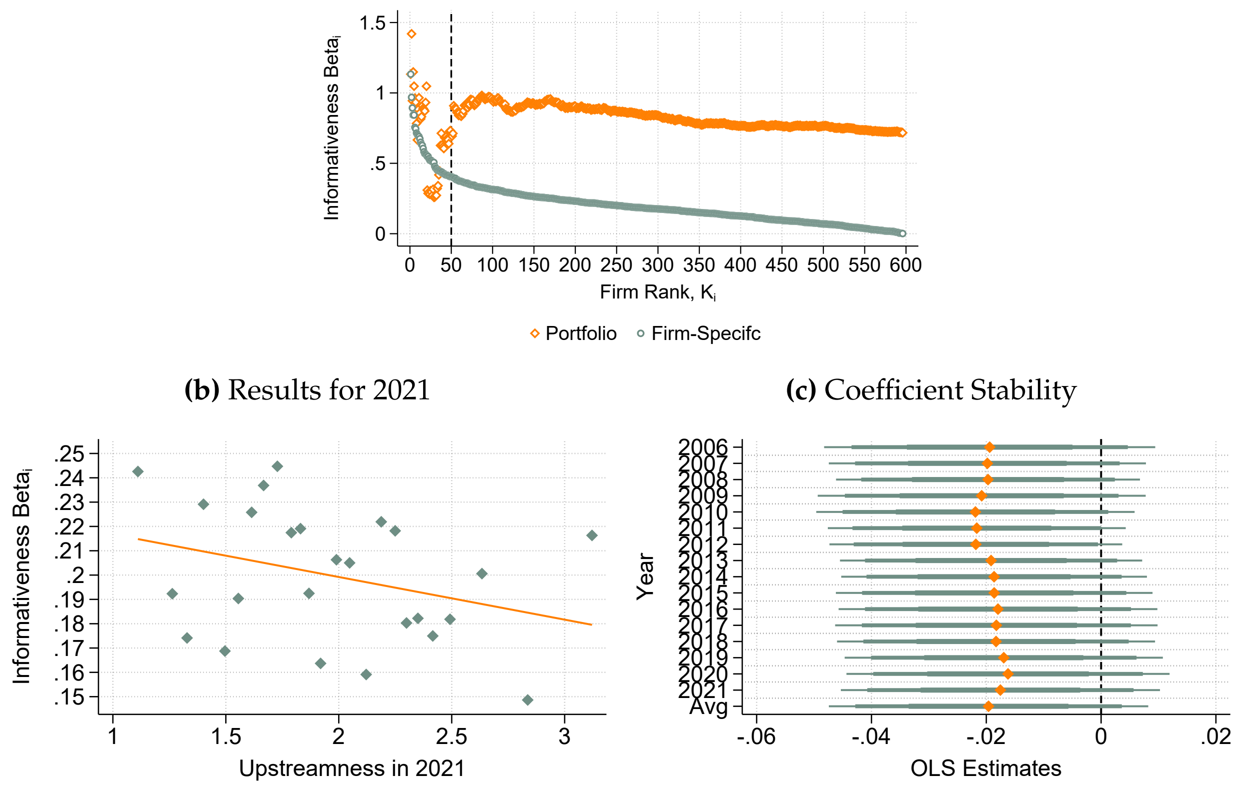
FinBERT, specifically fine-tuned for financial language and introduced by Araci (2019), utilizes BERT to classify the text sentiment. The training process involves using a financial corpus which consists of 1.8M news articles from Reuters, for further pre-training of BERT. Regarding the sentiment analysis, the Financial Phrasebank dataset is employed, comprising 4,845 English sentences randomly selected from financial news on LexisNexis. These sentences are manually labelled by 16 researchers with finance and business backgrounds, classifying them into positive, neutral, and negative sentiment.

A.3 Additional Cross-Sectional Results

In this section we present additional cross-sectional evidence that complements the main text. Figure A.1 presents results from a series of robustness exercises where instead of the informativeness R_i^2, as in the baseline, we use β_i (which are in absolute values and scaled by 100) from regressions of firm-level sentiment on HP-filtered aggregate output. Figure A.2 depicts predictive margins from the baseline panel probit model. Figure A.3 reports the results from the cross-sectional analysis of sentiment and downstreamness using probit and ordinal probit models and for all years in the sample. Finally, Figures A.4 and A.5 present the results from the cross-sectional analysis of sentiment and downstreamness under the Atalay (2017) and Chahrour et al. (2021) definitions of industries, respectively. As discussed also in the main text, the baseline finding of a negative relationship between upstreamness and informativeness does not change.

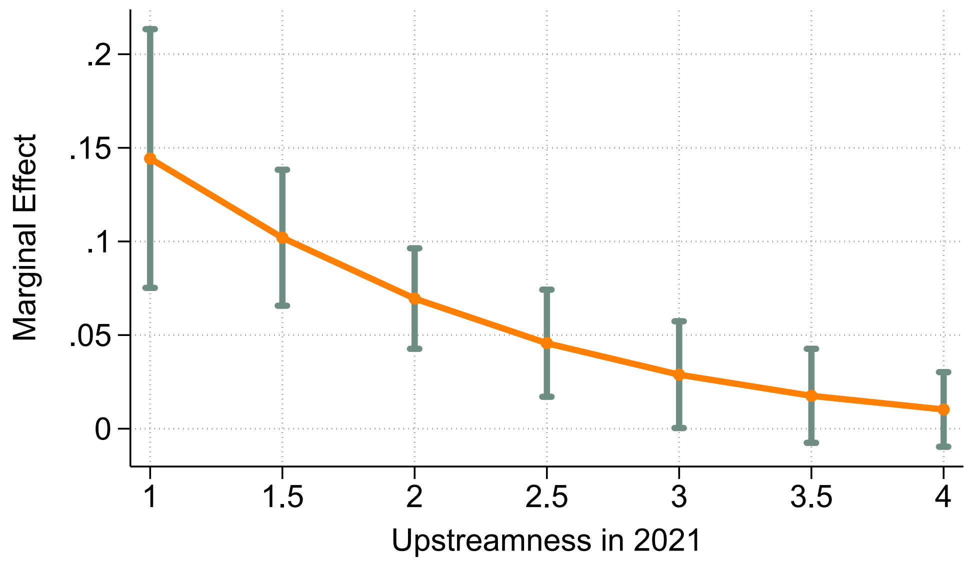
Figure A.1: Robustness Results with the Informativeness Beta

(a) The Cross-Section of Firm-Level Sentiment



Notes: Panel (a) shows firm- and portfolio-level betas with respect to HP-filtered output. Betas are in absolute values and scaled by 100. Panels (b) and (c) show results from cross-sectional regressions of upstreamness U_{us} on firm-level informativeness $|\beta_i|$. Horizontal bars in Panel (c) are 68%, 90%, and 95% confidence intervals. Standard errors are clustered at the industry level.

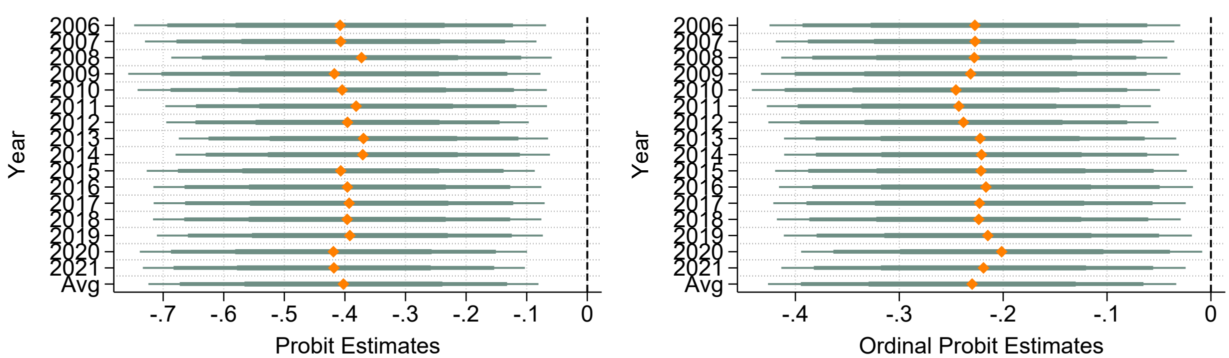
Figure A.2: Downstreamness and Sentiment: Predictive Margins of Probit Model



Notes: Marginal effects from a probit cross-sectional regression of upstreamness, $U_{s,t=2021}$, on the binary indicator of informativeness that takes the value of unity if firms are in the granular informativeness set, Γ, and 0 otherwise. The specification includes all the usual firm controls.

Figure A.3: Downstreamness and Sentiment: Alternative Econometric Approaches

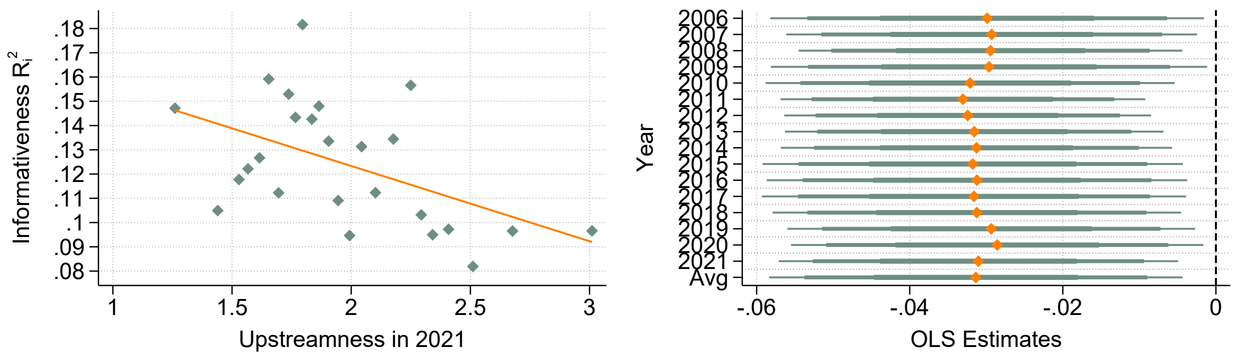
(a) Probit **(b)** Ordinal Probit



Notes: Results from cross-sectional regressions of upstreamness U_{st} on the binary indicator of informativeness (Probit model, Panel (a)) and the rank indicator K_i (Ordinal Probit model, Panel (b)) for each year. Horizontal bars are 68%, 90%, and 95% confidence intervals. Standard errors are clustered at the sector level.

Figure A.4: Downstreamness and Sentiment: Atalay (2017) Industry Definition

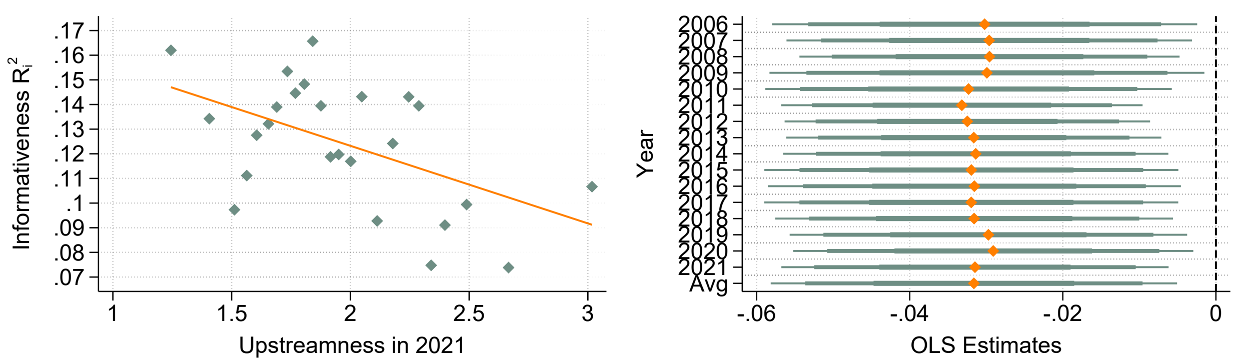
(a) Results for 2021 **(b)** Coefficient Stability



Notes: Results from cross-sectional regressions of upstreamness U_{us} on firm-level informativeness R_i^2. U_{us} is constructed based on Atalay (2017) sectors. Horizontal bars in Panel (b), are 68%, 90%, and 95% confidence intervals. Standard errors are clustered at the industry level.

Figure A.5: Downstreamness and Sentiment: Chahrour et al. (2021) Industry Definition

(a) Results for 2021 **(b)** Coefficient Stability



Notes: Results from cross-sectional regressions of upstreamness U_{us} on firm-level informativeness R_i^2. U_{us} is constructed based on Chahrour et al. (2021) sectors. Horizontal bars in Panel (b), are 68%, 90%, and 95% confidence intervals. Standard errors are clustered at the industry level.

A.4 Additional Time-Series Results

This section presents additional time-series results that complement the main text. In Figure A.6, we plot time-series dynamics of alternative variations of the granular sentiment index $\mathcal{S}_t$ and the high-downstreamness portfolio sentiment $\mathcal{P}_t$. Panels (a) and (b) plot the indices that are constructed under the FinBERT (ξ_t^2) and analyst forecast (ξ_t^3) definitions of sentiment, respectively. Panels (c) and (d) plot the indices under alternative robust specifications. All measures, including the real GDP series, are standardized.

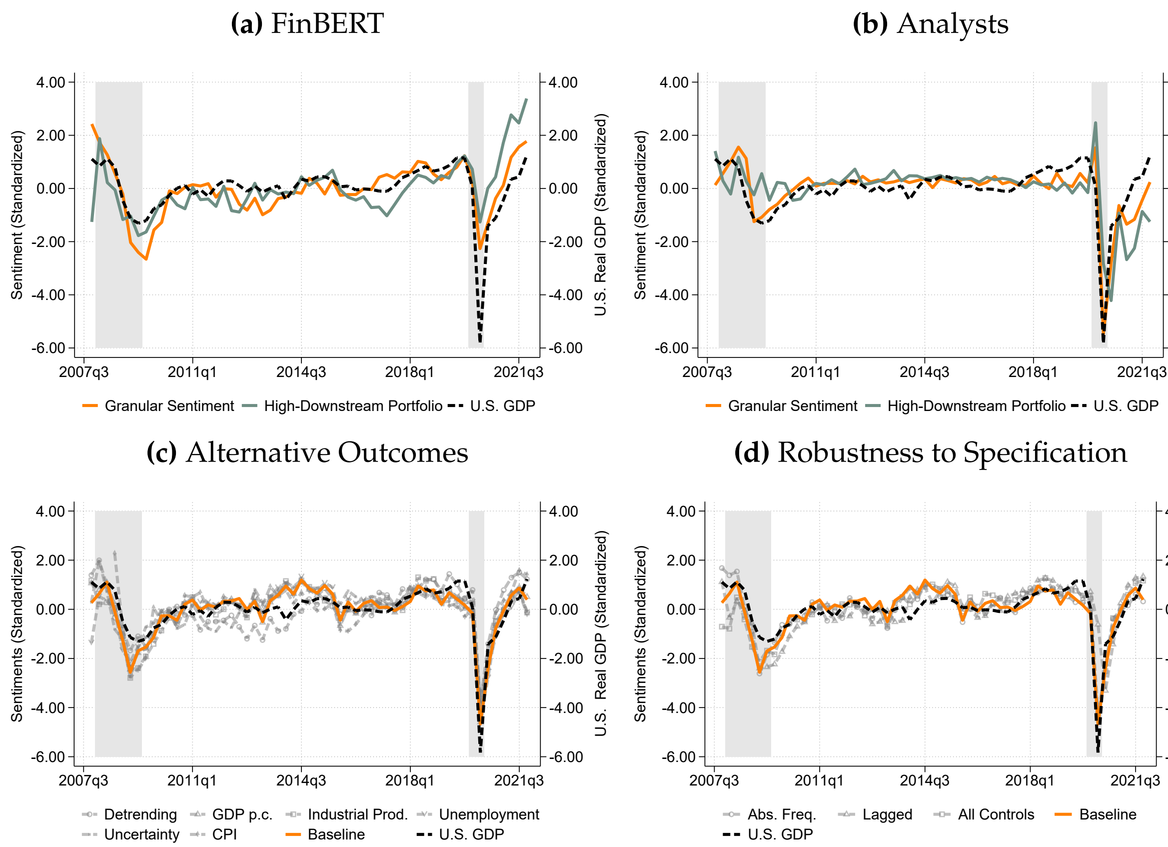
Figures A.7 and A.8 present local projection estimates under alternative specifications. We consider various extensions and variations of our baseline specification, the theme of which is in line with what was detailed first in Section 3.1. We consider alternative measures of firm-level sentiment, other macro outcomes variables than output, residualized sentiment measures which control for potentially confounding firm-level factors, among others. Figure A.9 presents the full set of local projection results under the alternative de-trending scheme: we do not employ the HP-filter but instead remove the time fixed effect from every aggregate series.

Table A.2 presents a matrix of correlations between all the sentiment indices that are used in the paper, including the baseline measures of granular sentiment $\mathcal{S}_t$, the high-downstreamness portfolio sentiment $\mathcal{P}_t$, the first principal component of five existing sentiment indices $\mathcal{K}_t$, and the five underlying sentiment measures that constitute $\mathcal{K}_t$.

In Table A.3 we report results from time-series regressions of $\mathcal{S}_t$ and $\mathcal{P}_t$ on macroeconomic outcomes and economy-wide sentiment $\mathcal{K}_t$. All specifications include the usual set of controls: the Fed Funds Rate, real Nasdaq returns, and aggregate TFP. Standard errors are robust to arbitrary heteroskedasticity and autocorrelation.

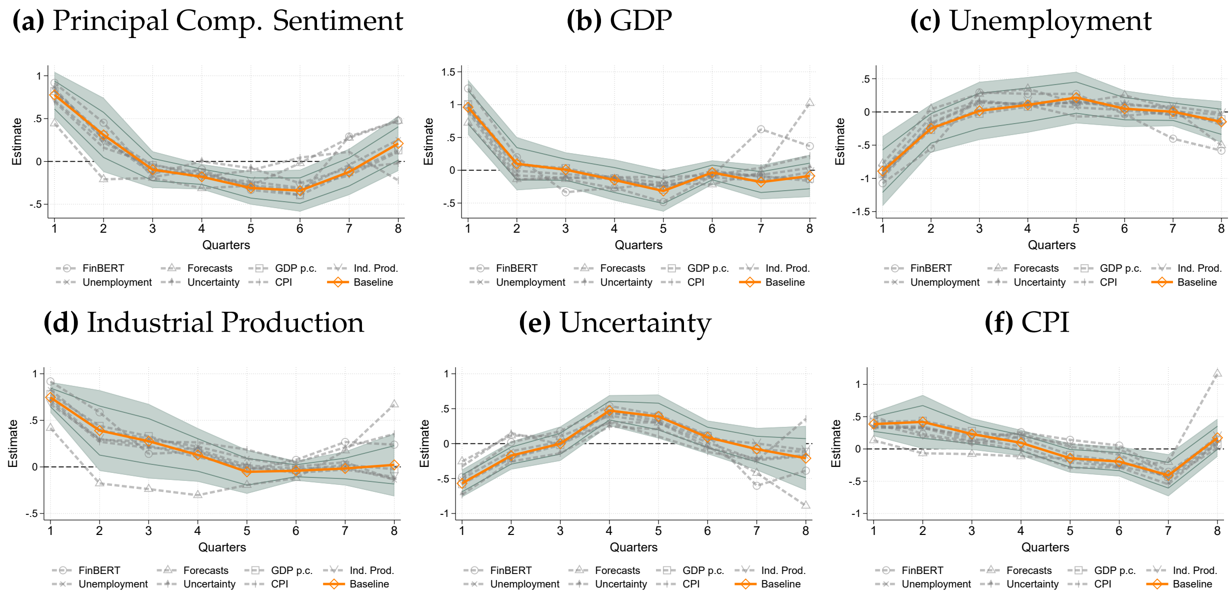
Table A.4 reports results from 2SLS regressions with mass shooting fatalities as an instrument for the high-downstreamness sentiment measure $\mathcal{P}_t$. The top and bottom panels report results without and with additional time-series controls, respectively. The usual set of controls includes the Fed Funds Rate, real Nasdaq returns, and aggregate TFP. Standard errors are robust to arbitrary heteroskedasticity and autocorrelation.

Figure A.6: Granular Sentiment Time Series: Alternative Specifications



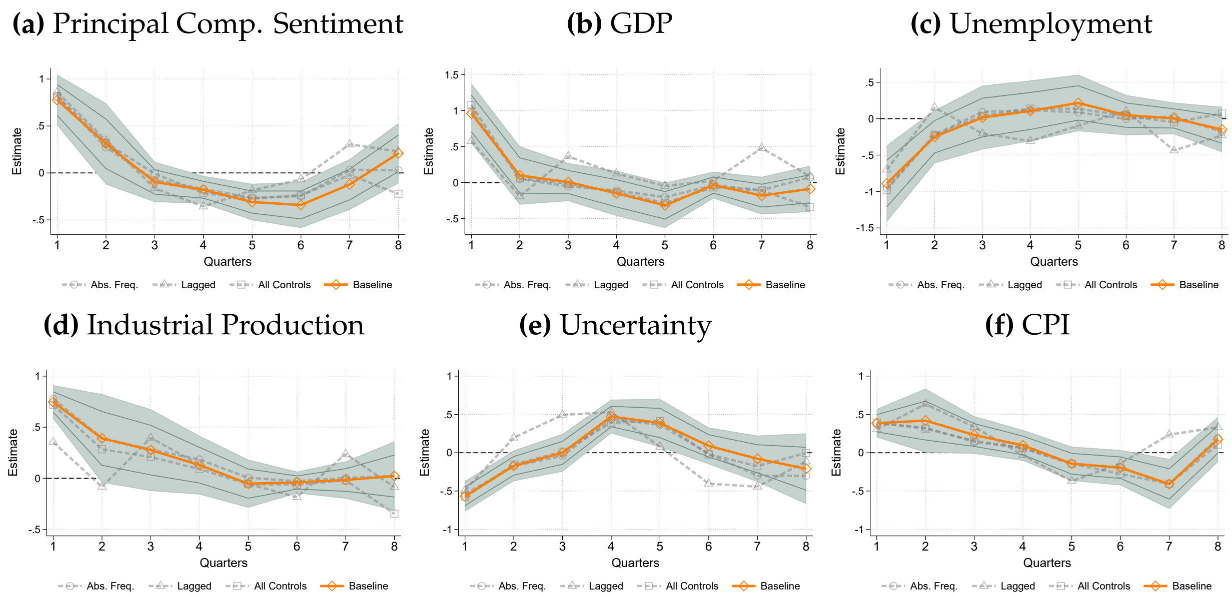
Notes: Time-series plots of granular sentiment, $\mathcal{S}_t$, and sentiment of the high-downstreamness portfolio, $\mathcal{P}_t$, for alternative measures of firm-level sentiment (Panels (a) and (b)) and under alternative specifications (Panels (c) and (d)).

Figure A.7: Dynamic Effects Robustness—Alternative Sentiments and Macro Outcomes



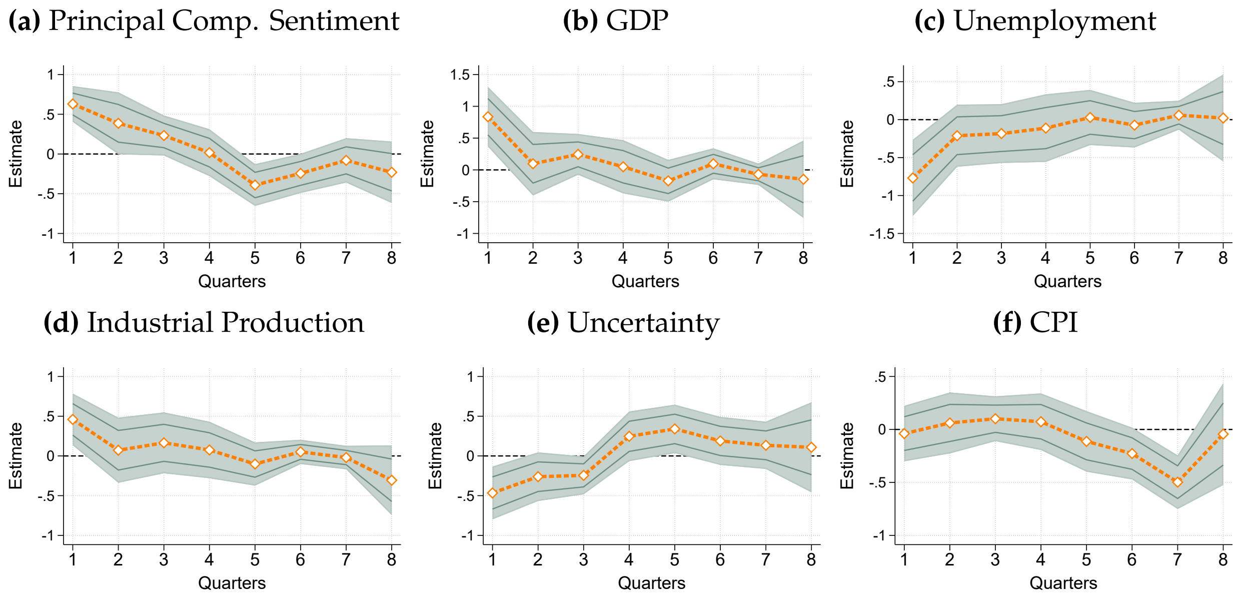
Notes: Local projection estimates of granular sentiment, $\mathcal{S}_t$, on macro aggregates under alternative firm-level sentiment measures and macro outcomes, as described in main text. Lines correspond to 68% and shaded areas to 90% Newey-West confidence bands, respectively.

Figure A.8: Dynamic Effects Robustness—Alternative Specifications



Notes: Local projection estimates of granular sentiment, $\mathcal{S}_t$, on macro aggregates under alternative econometric specifications. Lines correspond to 68% and shaded areas to 90% Newey-West confidence bands, respectively.

Figure A.9: Dynamic Effects Robustness: Alternative De-trending



Notes: Local projection estimates of granular sentiment, $\mathcal{S}_t$, on macro aggregates. Macro aggregates have been residualized from the time fixed effect instead of being HP-filtered. Lines correspond to 68% and shaded areas to 90% Newey-West confidence bands, respectively.

Table A.2: Matrix of Correlation Between Sentiment Indices

	Granular Sentiment	High-Downstream	Principal Component	OECD	Michi-gan	ISM PMI	Sentix	News
Granular Sentiment	100%							
High-Downstream	47%	100%						
Principal Component	61%	50%	100%					
OECD	80%	35%	81%	100%				
Michigan	34%	53%	65%	42%	100%			
ISM PMI	55%	42%	88%	78%	54%	100%		
Sentix	68%	53%	94%	82%	65%	78%	100%	
News	30%	28%	57%	37%	49%	45%	47%	100%

Notes: This table presents the matrix of correlations between various sentiment indices used in the paper. Variables, in order, correspond to baseline granular sentiment $\mathcal{S}_t$, baseline sentiment of the high-downstreamness portfolio $\mathcal{P}_t$, the first principal component of five existing sentiment indices $\mathcal{K}_t$, and the five sentiment indices that underlie $\mathcal{K}_t$—the OECD Business Confidence Index, the University of Michigan Index of Consumer Sentiment, the ISM Purchasing Managers' Index (PMI), the Sentix sentiment index, and news-based economic sentiment by van Binsbergen et al. (2024).

Table A.3: Granular Sentiment and the Macroeconomy—Contemporaneous Relationship

	Independent Variable: Granular Sentiment					
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	PC Sentiment	GDP	Unemployment	Industrial Production	Uncertainty	CPI
Granular Sentiment	0.514***	0.814***	-0.803***	0.785***	-0.627***	0.378***
	(0.090)	(0.140)	(0.187)	(0.085)	(0.167)	(0.106)
Observations	57	57	57	57	57	57
All Controls	✓	✓	✓	✓	✓	✓
aR^2 without Controls	0.438	0.729	0.653	0.721	0.460	0.266
aR^2 with Controls	0.565	0.823	0.729	0.804	0.440	0.286
	Independent Variable: High-Downstreamness Portfolio Sentiment					
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	PC Sentiment	GDP	Unemployment	Industrial Production	Uncertainty	CPI
Granular Sentiment	0.548***	0.539***	-0.547**	0.448***	-0.497***	0.167
	(0.121)	(0.192)	(0.229)	(0.118)	(0.149)	(0.140)
Observations	57	57	57	57	57	57
All Controls	✓	✓	✓	✓	✓	✓
aR^2 without Controls	0.375	0.199	0.165	0.096	0.216	0.01
aR^2 with Controls	0.647	0.599	0.526	0.532	0.378	0.205

Notes: Results from OLS regressions of granular sentiment $\mathcal{S}_t$ (Panel (a)) and high-downstreamness portfolio sentiment $\mathcal{P}_t$ (Panel (b)) on macroeconomic aggregates. Controls include aggregate TFP, the Fed Funds Rate, and real Nasdaq returns. Newey-West standard errors are in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table A.4: IV Regression with U.S. Mass Shootings—High-Downstreamness Portfolio

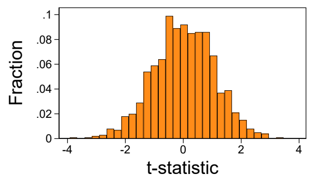
	Independent Variable: High-Downstream Portfolio Instrumented by Mass Shootings					
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	PC Sentiment	GDP	Unemployment	Industrial Production	Uncertainty	CPI
Sentiment	0.483**	0.431**	-0.584**	0.538	-0.382	-0.074
	(0.221)	(0.215)	(0.255)	(0.410)	(0.274)	(0.245)
Observations	39	39	39	39	39	39
All Controls	✗	✗	✗	✗	✗	✗
First-Stage F-stat	8.673	8.674	8.675	8.676	8.677	8.678
	Independent Variable: High-Downstream Portfolio Instrumented by Mass Shootings					
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	PC Sentiment	GDP	Unemployment	Industrial Production	Uncertainty	CPI
Sentiment	0.382***	0.239***	-0.370***	0.202	-0.326*	-0.216
	(0.127)	(0.083)	(0.110)	(0.202)	(0.193)	(0.222)
Observations	39	39	39	39	39	39
All Controls	✓	✓	✓	✓	✓	✓
First-Stage F-stat	13.889	13.890	13.891	13.892	13.893	13.894

Notes: Results from IV regressions of the DSI ($\mathcal{P}_t$), instrumented by U.S. mass shooting fatalities, on macroeconomic aggregates. Top and bottom panels report results without and with controls, respectively. Controls include aggregate TFP, the Fed Funds Rate, and real Nasdaq returns. Newey-West standard errors are in parentheses. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

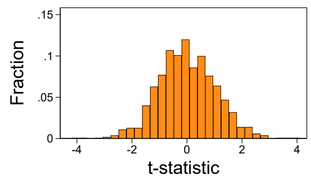
A.5 Further Robustness Tests

This section presents additional robustness tests that complement the main text. Figure A.10 presents the outcome of a placebo test for the documented cross-sectional relation between the informativeness R_i^2 and sectoral downstreamness. We run specifications where downstreamness is randomly re-assigned across sectors within a time period, with replacement. T-statistics from 1,000 of these Monte-Carlo simulations are presented for every year of the analysis, and each regression includes the usual set of controls. Each distribution of t-statistics is centered around zero with a small minority (less than 2%) of cases falling above the rule-of-thumb threshold of |2|. Thus, it is highly unlikely that our cross-sectional result was obtained by pure chance.

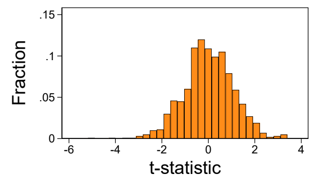
Figure A.10: Placebo Cross-sectional Regressions of Informativeness on Upstreamness



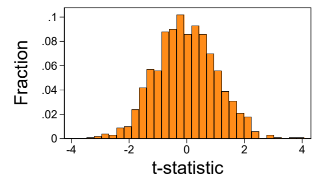
(a) 2006



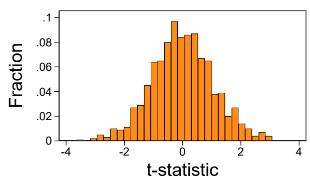
(b) 2007



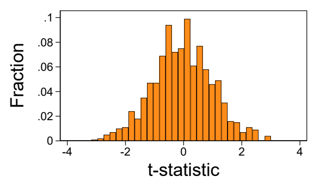
(c) 2008



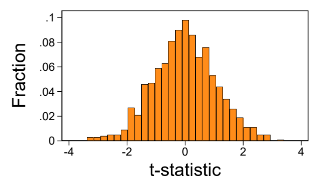
(d) 2009



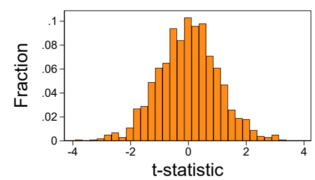
(e) 2010



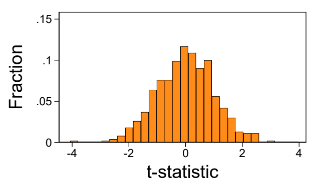
(f) 2011



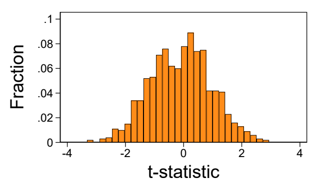
(g) 2012



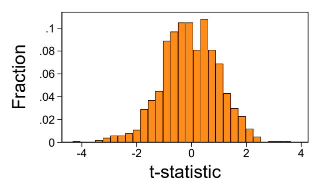
(h) 2013



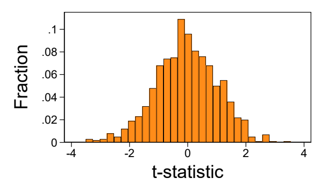
(i) 2014



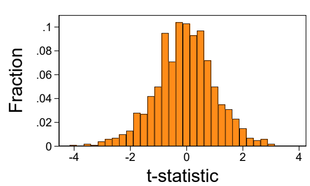
(j) 2015



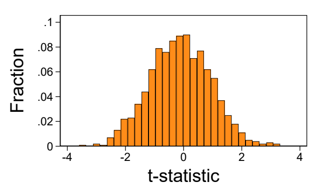
(k) 2016



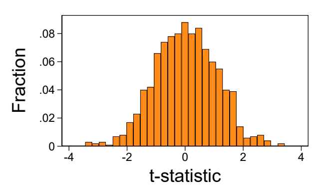
(l) 2017



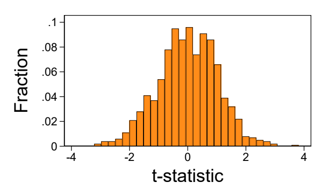
(m) 2018



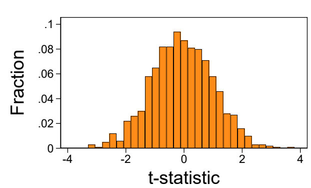
(n) 2019



(o) 2020



(p) 2021



(q) Average (2006-2021)

Notes: Each panel reports histograms of t-statistics from non-parameteric Monte-Carlo permutations with 1,000 simulations for cross-sectional regressions of firm-level R_i^2 on upstreamness U_{st} for different years. In each regression upstreamness is randomly re-assigned with replacement. Each specification includes the usual controls. Standard errors are clustered at the industry level.

B Model Appendix

B.1 Equilibrium Characterization

This appendix derives the equilibrium condition in (19).[5]

To start, notice that the market-clearing conditions for the upstream and downstream sectors imply the following expressions for their revenue functions, respectively:

- Downstream ($\beta_d = 1$):

$$\begin{aligned} Q_d &= C + X_{dd} \\ P_d Q_d &= P_d C + P_d X_{dd} = \beta_d C + \alpha_{dd} P_d Q_d \end{aligned}$$

- Upstream ($\beta_u = 0$):

$$\begin{aligned} Q_u &= X_{uu} + X_{ud} \\ P_u Q_u &= P_u \left(\alpha_{uu} \frac{P_u Q_u}{P_u} + \alpha_{du} \frac{P_d Q_d}{P_u} \right) = \beta_u C + \left(\alpha_{uu} P_u Q_u + \alpha_{du} P_d Q_d \right), \end{aligned}$$

where we have also used the demand conditions for intermediate goods in (15).

We conclude that

$$V = (\mathbf{I} - \mathbf{A'})^{-1} \mathbf{b} \cdot C \equiv \Lambda \cdot C, \tag{B.1}$$

where $V = [V_i]_i$ with sector revenue $V_i \equiv P_i Q_i = \beta_i C + \sum_j \alpha_{ji} V_j$.

It remains to find an expression for economy-wide output C. We proceed as follows: the production function implies that the revenue for sector can also alternatively be written as

$$V_i = P_i Z_i \left[\Pi_j \left(\alpha_{ij} \frac{V_i}{P_j} \right)^{\alpha_{ij}} \right] L_i^{\delta_i}.$$

Taking logs, we arrive at

$$\begin{aligned} v_i &= p_i + z_i + \sum_j \alpha_{ij} \left(\log \alpha_{ij} + v_i - p_j \right) + \delta_i l_i \\ (\gamma + \delta_i) v_i &= z_i + \tau_i + \delta_i l_i + p_i - \sum \alpha_{ij} p_j, \end{aligned}$$

where lower-case letters denote the log of their upper-case counterparts and τ_i is defined

[5] For ease of notation, we abstract from time subscripts in this appendix

in the main text. Thus, stacking results shows that

$$\operatorname{diag}(\gamma+\delta_i)\,v = z + \tau + \operatorname{diag}(\delta_i)\ell + (\mathbf{I}-\mathbf{A})p. \tag{B.2}$$

Now, combining the two equations for revenue shows the market-clearing price vector is

$$p = (\mathbf{I}-\mathbf{A})^{-1}\left[\operatorname{diag}(\gamma+\delta_i)\log\mathbf{L} - \operatorname{diag}(\delta_i)\,\ell + \operatorname{diag}(\gamma+\delta_i)\,\mathbf{1}_n c - z - \tau\right]. \tag{B.3}$$

The vector of log-consumption allocations $\mathbf{c}$ is therefore

$$\begin{aligned}\mathbf{c} &= \log\mathbf{b} + c\mathbf{1}_n - \mathbf{p}\\ &= \log\mathbf{b} + c\mathbf{1}_n - (\mathbf{I}-\mathbf{A})^{-1}\left[\operatorname{diag}(\gamma+\delta_i)\log\mathbf{L} - \operatorname{diag}(\delta_i)\,\ell + \operatorname{diag}(\gamma+\delta_i)\,\mathbf{1}_n c - z - \tau\right].\end{aligned}$$

The consumption index is

$$c = \mathbf{b}'(\mathbf{c} - \log\mathbf{b}) \tag{B.4}$$

Inserting the expression for the log-consumption allocations into the consumption index in (B.4), and rearranging terms shows that economy-wide consumption equals

$$c = \beta'(\mathbf{I}-\mathbf{A})^{-1}\left(z + \varkappa_0 - \mathbf{k_1}\ell\right), \tag{B.5}$$

where the coefficients $\varkappa_0 \equiv \mathbf{t} - \operatorname{diag}(\delta_i+\gamma)\log\Lambda$ with $\tau \equiv \left[\sum_j \alpha_{ij}\log\alpha_{ij}\right]_i$ and $\mathbf{k_1} \equiv \operatorname{diag}(\delta_i)$.

B.2 Proofs of Results

Proof of Proposition 1: The result in the proposition follows immediately from inserting (18) and (17) into (16), using also that $P_{it}Q_{it} = \lambda_i C_t$. □

Proof of Corollary 1: First, notice that $\mathbf{A} = [\alpha_{uu}\ 0; \alpha_{du}\ \alpha_{dd}]$ and $\beta = [0\ 1]'$. Thus,

$$\Lambda = \begin{pmatrix}\frac{\alpha_{du}}{(1-\alpha_{uu})(1-\alpha_{dd})}\\ \frac{1}{1-\alpha_{dd}}\end{pmatrix},\quad U = \begin{pmatrix}\frac{2-\alpha_{dd}-\alpha_{uu}}{(1-\alpha_{uu})(1-\alpha_{dd})}\\ \frac{1}{1-\alpha_{dd}}\end{pmatrix}. \tag{B.6}$$

Notice that $u_u > u_d$ so that the upstream sector is always more upstream than the downstream sector in the production chain. We conclude that $\lambda_d > \lambda_u$ iff. $1-\alpha_{uu} > \alpha_{du}$. □

Proof of Proposition 2: The result follows from similar steps to those in Maćkowiak and Wiederholt (2009). The profit function of a firm is:

$$\Pi(L_i, C) = P_i Q_i - wL_i = \lambda_i C - L_i, \tag{B.7}$$

which expressed in terms of log-deviations can be written as:

$$\pi(l_i, c) = \Pi\left(\bar{L}_i \exp^{l_i}, \bar{C} \exp^{c}\right),$$

where $\bar{X}$ denotes the steady-state value of the variable X. A second-order approximation around the steady-state then shows that

$$\pi\left(l_i, c\right) \approx \pi\left(0,0\right) + \pi_1 l_i + \pi_2 c + \frac{1}{2}\pi_{11} l_i^2 + \frac{1}{2}\pi_{22} c^2 + \pi_{12} l_i c \tag{B.8}$$

$$= \pi\left(0,0\right) + \pi_2 c + \frac{1}{2}\pi_{11} l_i^2 + \frac{1}{2}\pi_{22} c^2 + \pi_{12} l_i c. \tag{B.9}$$

The full-information solution is thus given by $l_i^\star = \frac{\pi_{12}}{|\pi_{11}|} c$, while the imperfect-information solution can be characterized by $l_i = \mathbb{E}_i\left[l_i^\star\right]$. It now follows that:

$$\pi(l_i, c) - \pi(l_i^\star, c) = \frac{|\pi_{11}|}{2}(l_i - l_i^\star)^2 = \frac{|\pi_{11}|}{2}\left[\left(\frac{\pi_{12}}{\pi_{11}}\right) \cdot \left(c - \mathbb{E}_i\left[c\right]\right)\right]^2.$$

Using equation (B.7) then provides us with the result. □

B.3 Numerical Solution

We solve the model numerically. This entails solving two fixed point problems: (i) an inner fixed point problem, which solves for firms' labor choices given their attention choices; and (ii) an outer fixed point problem, which solve for firms' attention choices given input choices. Below we detail how we solve each problem. We halt the iteration between the two steps that solve their respective fixed-point problems when the attention vectors $\{\mathbf{m}\}_i$ have converged in the sense of maximum absolute difference..

Inner Fixed Point: We solve for the vector of labor choices using a parameterized expectations algorithm akin to that used in Chahrour et al. (2021). Let the determinants of a firm's labor choice in (19) be decomposed into:

$$\log V_{num,t} = \Lambda'(\mathbf{z}_t + \mathbf{k}_0 + \mathbf{k}_1 \cdot \ell_\mathbf{t})$$
$$\log V_{den,t} = \frac{1}{\nu} \log \left(\sum_j \exp \ell_\mathbf{t} \right),$$

and let $\mathbf{L}_t \equiv [L_{1,t}, ..., L_{N,t}]$ for $t = \{1, 2, , ...T\}$. We then approximate $\mathbb{E}\left[V_{num,t} \mid \Omega_{i,t}\right]$ and $\mathbb{E}\left[V_{den,t} \mid \Omega_{i,t}\right]$ using a linear regression of $\log V_{num,t}$ and $\log V_{den,t}$, respectively, onto $\Omega_{i,t}$ for each $i = \{1, 2, ..., N\}$. We draw shocks and compute firms' signals conditional on their attention choices $\{\mathbf{m_i}\}$. We use the fitted values $\log V_{num,t}$ and $\log V_{den,t}$ in place of expectations in (19). We subsequently update the sequence of labor choices $\{\mathbf{L}_t\}$ and check if the history $\{\mathbf{L_1}, \mathbf{L_2}, ..., \mathbf{L_T}\}$ has converged. We initialize the algorithm using the full-information solution $\{\mathbf{L}^\star_\mathbf{1}, \mathbf{L}^\star_\mathbf{2}, ..., \mathbf{L}^\star_\mathbf{T}\}$. Throughout, we set $T = 100,000$. We do not find that larger values of T change our results.

Outer Fixed Point: We solve for firms' attention choices $\{\mathbf{m}_i\}$. We conjecture a vector of attention choices for each firm $\{\mathbf{m}_i\}$. Given these attention choices, we solve for firms' labor inputs—that is, we solve the inner fixed point problem—using the above algorithm. Given equilibrium labor choices, and hence equilibrium consumption C_t and wages W_t in the economy, we then, for each $i = \{1, 2, .., N\}$, solve a firm's attention choice problem given the cost function $K(\mathbf{m}_i)$. This provides us with updated values of $\{\mathbf{m}_i\}$. We halt the algorithm that solves for the outer fixed-point problem when the series of attention choices $\{\mathbf{m}_i\}$ has converged.

References

Araci, Dogu T, "FinBERT: Financial Sentiment Analysis with Pre-trained Language Models," *Manuscript*, 2019.

Atalay, Enghin, "How Important Are Sectoral Shocks?," *American Economic Journal: Macroeconomics*, October 2017, *9* (4), 254–80.

Baker, Scott R, Nicholas Bloom, and Steven J Davis, "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 07 2016, *131* (4), 1593–1636.

Chahrour, Ryan, Kristoffer Nimark, and Stefan Pitschner, "Sectoral Media Focus and Aggregate Fluctuations," *American Economic Review*, 2021, *111* (12), 3872–3922.

Chava, Sudheer, Wendi Du, and Nikhil Paradkar, "Do Managers Walk the Talk on Environmental and Social Issues?," *Manuscript*, 2020.

_ , _ , and _ , "More than Buzzwords? Firms' Discussions of Emerging Technologies in Earnings Conference Calls," *Manuscript*, 2021.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, "Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Manuscript*, 2018.

Fernald, John, "A Quarterly, Utilization-adjusted Series on Total Factor Productivity," *Federal Reserve Bank of San Francisco Working Paper*, 2014, *19.*

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera, "The Voice of Monetary Policy," *American Economic Review*, February 2023, *113* (2), 548–84.

Lagerborg, Andresa, Evi Pappa, and Morten O Ravn, "Sentimental Business Cycles," *The Review of Economic Studies*, 08 2022, *90* (3), 1358–1393.

Maćkowiak, Bartosz and Mirko Wiederholt, "Optimal sticky prices under rational inattention," *American Economic Review*, 2009, *99* (3), 769–803.

Ravn, Morten O and Harald Uhlig, "On Adjusting the Hodrick-prescott Filter for the Frequency of Observations," *The Review of Economics and Statistics*, 2002, *84* (2), 371–376.

van Binsbergen, Jules H, Svetlana Bryzgalova, Mayukh Mukhopadhyay, and Varun Sharma, "(Almost) 200 Years of News-Based Economic Sentiment," *NBER Working Paper 32026*, 2024.