CENTRE FOR RESPONSIBLE BANKING & FINANCE

CRBF

University of St Andrews Business School

WORKING PAPERS IN RESPONSIBLE BANKING & FINANCE

Granular Sentiments

By *Rustam Jamilov, Alexandre Kohlhas, Oleksandr Talavera, and Mao Zhang*

Abstract: We propose an empirically-motivated theory of business cycles, driven by fluctuations in sentiment towards a small number of firms. We measure firm-level sentiment with computational linguistics and analyst forecast errors. We find that 50 firms can account for over 70% of the unconditional variation in U.S. sentiment and output over the period 2006-2021. The "Granular Sentiment Index", measuring sentiment towards the 50 firms, is dominated by firms that are closer to the final consumer, i.e. are downstream. To rationalize our findings, we embed endogenous information choice into a general equilibrium model with heterogeneous upstream and downstream firms. We show that attention centers on downstream firms because they act as natural "information agglomerators". When calibrated to match select moments of U.S. data, the model shows that orthogonal shocks to sentiment of the 20% most downstream firms explain more than 90% of sentiment-driven (and 20% of total) aggregate fluctuations.

WP Nº 25-007

2nd Quarter 2025

University of St Andrews
Scotland's first university

600 YEARS
1413 – 2013

Granular Sentiments[†]

Rustam Jamilov Alexandre Kohlhas Oleksandr Talavera Mao Zhang

February, 2024

Abstract

We propose an empirically-motivated theory of business cycles, driven by fluctuations in sentiment towards a small number of firms. We measure firm-level sentiment with computational linguistics and analyst forecast errors. We find that 50 firms can account for over 70% of the unconditional variation in U.S. sentiment and output over the period 2006-2021. The "Granular Sentiment Index", measuring sentiment towards the 50 firms, is dominated by firms that are closer to the final consumer, i.e. are *downstream*. To rationalize our findings, we embed endogenous information choice into a general equilibrium model with heterogeneous upstream and downstream firms. We show that attention centers on downstream firms because they act as natural "information agglomerators". When calibrated to match select moments of U.S. data, the model shows that orthogonal shocks to sentiment of the 20% most downstream firms explain more than 90% of sentiment-driven (and 20% of total) aggregate fluctuations.

[†]We thank Vladimir Asriyan, Florin Bilbiie, Saleem Bahaj, Vasco Carvalho, Wei Cui, Matthew Elliott, Cécile Fabre, Andrea Ferrero, Xavier Gabaix, Jordi Galí, Mishel Ghassibe, Albert Marcet, Michael McMahon, Marta Morazzoni, Franck Portier, Lukasz Rachel, Morten Ravn, Federica Romei, Victoria Vanasco, Jaume Ventura and seminar participants at various venues for valuable comments and suggestions. This research received financial support from *Handelsbanken Stiftelsen.* Jamilov: University of Oxford (rustam.jamilov@all-souls.ox.ac.uk). Kohlhas: University of Oxford (alexandre.kohlhas@economics.ox.ac.uk). Talavera: University of Birmingham (o.talavera@bham.ac.uk). Zhang: University of St Andrews (mz71@st-andrews.ac.uk).

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

Many economic fluctuations are attributable to the incompressible grains of economic activity [...] the granular hypothesis. (Gabaix, 2011)

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 or non-systematic monetary policy. Indeed, since the treatise of Pigou (1927) on the drivers of "industrial fluctuations", a frequently-proposed alternative driver of business cycles has been the coordinated waves of optimism and pessimism that often characterizes people's views about the state of 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); Lagerborg et al. (2022)). 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 ask which firms drive business sentiment, and whether shocks to them trigger business cycles. Our aim is threefold. First, we wish to establish whether the "Granular Hypothesis" (Gabaix, 2011) applies to sentiment-driven fluctuations—i.e., whether there exists a subset of firms which constitute a sufficient statistic for the economy-wide effects of sentiment shocks. Second, we want to understand the characteristic of the firms that belong to this subset — which features cause firms to be central for overall sentiment and why? And finally, third, we wish to understand 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 less than 50 firms can account for over 2/3 of the unconditional variation in U.S. sentiment and output over the past two decades. The "Granular Sentiment" measure, which captures sentiment towards these 50 firms, is dominated by firms that are close to the final consumer, i.e., that are *downstream*. We rationalize our results within a standard multi-sector model to which we add endogenous information choice. We show that attention gravitates to downstream firms, as they emerge as natural "information agglomerators", whose sentiment matters disproportionally for aggregate dynamics. When calibrated to match moments of U.S. data, the model shows that the 20 percent most downstream firms account for close to 90 percent of sentiment-driven output fluctuations.

Empirics: We start by computing a firm-level measure of sentiment. To do so, we apply standard tools from computational linguistics to publicly listed U.S. firms' quarterly earnings calls. A significant fraction of each earnings call is a Q&A session between firm managers and industry analysts, which triggers unscripted conversations discussing beliefs about futureperformance. This contrasts with the scripted and often heavily formalized 10-K reports that listed firms also release. Using a well-established corpus of positive and negative sentiment words from Loughran and McDonald (2011), we work through every transcript of every firm from every quarter to construct an index of firm-level sentiment.

We complemented our sentiment index with two auxiliary measures. The first employs the FinBERT machine learning model (Devlin et al., 2018; Araci, 2019) to the earnings calls to create a finer text-based measure of firm-level sentiment. The second measure instead uses analysts' one-quarter-ahead earnings per share forecasts following each earnings call. The revision to forecasts, and the associated realized errors, allow us to proxy for analysts' sentiment following the call.

Next, we use our firm-level measures of sentiment to analyze which firms are most informative about macroeconomic and sentiment fluctuations. We first construct a ranking of "firm informativeness", which closely follows the approach that Gabaix (2011) takes to firm size. Specifically, we run firm-level regressions with our measure of sentiment as the main independent variable and a macroeconomic outcome (e.g., output) as the dependent variable. We collect the coefficients of determination from these regressions and rank firms in descending order. We next construct portfolios of sentiment, starting with the sentiment of the most informative firm, then considering the average sentiment of the first two ranks of informativeness, and so on. Finally, we re-estimate our regressions using portfolio-level sentiment as the main independent variable and collect the vector of portfolio-level coefficient of determination.

Our first main result shows that sentiment towards 50 firms can account for over 60 percent of the variation in aggregate U.S. sentiment and output. In particular, we document that there is a non-monotone relationship between firm informativeness and the share of output fluctuations accountable. Initially, as we add additional firms into our portfolio the marginal informational benefit outweighs the additional firm-specific noise. However, after the around 50th ranked firm, the trade-off reverses. The "Granular

Sentiment" index, measuring sentiment towards these 50 firms and captures around 60 percent of the variation in U.S. sentiment and output, is, through this lens, a "sufficient statistic" for sentiment-induced fluctuations. Indeed, we find that sentiment towards these 50 firms is substantially more informative than the naive average of the sentiment of all firms. We show that our results extends to cases in which we partial out for the effects of other macroeconomic and firm-specific variables as well as alternative salient economic shocks, and extend to our auxiliary measures of firm sentiment.

Our second key result characterizes the most informative firms. We take an agnostic approach and construct an array of potential explanatory variables, which includes a firm's size, market beta, sector, idiosyncrati revenue volatilty, and leverage, among many others. We document that sectoral *downstreamness*, i.e. the proximity to the end consumer, is the strongest predictor of being a member of our set. All else equal, firms whose sentiment is more informative about aggregate fluctuations are more downstream. We conduct a battery of robustness tests, which show that this result is resilient to issues related to identification, alternative measures and specifications, and robust to the inclusion of firm-level as well as aggregate controls.

The third and final step of our empirical analysis studies the time-series implications of innovations to granular sentiments. 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 aggregate sentiment, output, unemployment, and inflation, conditional on important controls such as productivity or uncertainty shocks. The effects are, furthermore, large, and statistically significant but not overly persistent. Comfortingly, our results thus show that granular sentiment shocks have similar macroeconomic effects to the aggregate "noise" shocks studied in the literature (e.g., Lorenzoni (2009)). To address potential concerns with identification, we show that our results extend to circumstances in which we instrument sentiment with fatalities from U.S. mass shootings, following the approach in Lagerborg et al. (2022).

Model: We proceed to rationalize our empirical results within the context of 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 information choice. This allows us to explore the quantitative potential of sentiment shocks towards a small set of firms in driving economy-wide fluctuations. In the model, firms in both upstream and downstream sectors choose inputs under imperfect information about sector-specific 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 (productivity and output) 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 different upstream sectors, their output also better agglomerates the dispersed information that exists about productivity in the economy. In this sense, downstream-sector firms act as natural "information agglomerators", similar to the role of market-clearing prices in Grossman and Stiglitz (1980) classical analysis.

We use our macroeconomic model to explore the business cycle implications of firms' asymmetric attention choices. We show that, for standard parameter values, calibrated to match the BEA production tables, sentiment shocks towards the most downstream sectors drive a sizeable share of overall output fluctuations. Specifically, our main quantitative exercise involves calibrating the model economy to the 27 U.S industries in Atalay (2017) , derived from the BEA productions structure. We show that orthogonal shocks to sentiment about the productivity of firms in 20 percent of industries—those that are the most downstream—can account for over 90 percent of sentiment-driven fluctuations in output. Overall, we find that sentiment shocks to the most downstream sector explain around 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, output fluctuations).

Literature: Our research relates to two long-standing ideas in macroeconomics: (i) "animal spirits-driven business cycle fluctuations" and (ii) "the granular hypothesis". We review how our work contributes to each strand of research below, in addition to that on natural language processing in economics.

First, we build on the literature that quantifies animal spirits-driven business cycles, following Pigou (1927)'s initial idea. 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. (2021b), Lagerborg et al. (2022), Chahrour and Ulbricht (2023), Maxted (2023), Angeletos and Sastry (2019), Flynn and Sastry (2022, 2024). We emphasize the role of a subset of firms, those which are close to the final consumer, in driving aggregate fluctuations in sentiment. Our work is, as such, closely related to Chahrour et al. (2021), who study how newspaper reporting about specific sectors help account for aggregate fluctuations. The contribution of our paper, in this context, is to highlight which firm-level characteristics drive sentiment-driven fluctuations, across different 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 business cycles (Carvalho and Gabaix, 2013), international trade and finance (di Giovanni et al., 2014; Gaubert and Itskhoki, 2021), banking and insurance (Chodorow-Reich et al., 2021; Galaasen et al., 2023), and nominal rigidities (Pasten et al., 2023). To the best our knowledge, we are the first to explore the implications of the granular hypothesis for beliefs and sentiment-driven business cycles. Relatedly, our paper contributes to the literature on the network origins of macroeconomic fluctuations that emphasizes the role of the input-output structure for aggregate dynamics (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 (Hansen and McMahon, 2016; Hansen et al., 2018; 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, machine learning technique that detects and measures sentiment using the BERT algorithm (Devlin et al., 2018). We particularly focus on the version of the BERT language model that has been pre-trained on financial (natural language) 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), environmental economics (Chava et al., 2021), and 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, Coibion and Gorodnichenko (2012), Bordalo et al. (2021), Enders et al. (2021a), and Asriyan and Kohlhas (2024).

2 Empirical Analysis

In this section, we describe the data used for our analysis, our empirical strategy, and our main empirical results. We close the section with a discussion of robustness test and additional results. In doing so, we present new evidence on the drivers of firm-level

[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).

sentiment and its macroeconomic implications.

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, listed U.S. firms are mandated to hold conference calls with financial analysts in conjunction with their quarterly earnings release.[2] Our dataset includes the transcripts of the repository of such earnings call dating back to 2006 for U.S. firms that are part of the S&P 500 index. Our baseline sample, as a result, includes over 23,000 observations with 61 unique quarters and 619 unique firms from 2006q4-2021q4.[3]

Our baseline approach to measuring firm-level sentiment is a *dictionary-based term-counting method*, used also in, for example, Baker et al. (2016). As a dictionary for positive and negative sentiment words, 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 its earnings call, scaled by the overall length of the transcript. We confirm that our choice of scaling does not influence our results (Section 2.5).[4] We denote this measure for firm i at time t by ξ_{it}.

We contrast and compare our baseline measure with two auxiliary measures of firm-level sentiment: the first of which relies on *natural language processing*; the latter uses *financial analysts' firm-level forecasts*. In particular, our second measure, ξ^1_{it}, employs a sophisticated machine-learning algorithm to potentially refine the basic dictionary-based measure. We manually process and split each earnings call transcript into individual text segments, where each segment is either an analyst's question, a manager's answer, or a

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

[4]More formally, let $\mathcal{C}_P$ and $\mathcal{C}_N$ denote the set of all terms that belong to the positive and negative sentiment dictionary, respectively. Our sentiment measure counts the number of times terms from $\mathcal{C}_P$ and $\mathcal{C}_N$ appear in each earnings-call transcript i and quarter t. Denote with B_{it} the total number of words in each transcript. The baseline net-sentiment measure is defined as positive minus negative sentiment:

$$\xi_{it} \equiv \frac{\sum_b^{B_{it}} \left(1[b \in \mathcal{C}_P]\right) - \sum_b^{B_{it}} \left(1[b \in \mathcal{C}_N]\right)}{B_{it}} \tag{1}$$

where $1[\cdot]$ is an indicator function. Following Hassan et al. (2019), we multiply ξ_{it} by 100,000.

separate paragraph in the management presentation. We feed each text segment into a pre-trained financial BERT (FinBERT) model (see Appendix A.2 for more details). Each text segment obtains a predicted probability score, corresponding to each sentiment classification (positive, neutral, or negative), ranging from 0 to 1. We then assign each segment to the sentiment category with the highest predicted probability. We aggregate the scores to the firm-quarter level as the difference between positive and negative segments, scaled by the number of total words in each transcript.[5]

Our third and final measure, ξ^2_{it}, computes firm-level sentiment using financial analysts' forecasts. We obtain analysts' forecasts from the I/B/E/S database. We focus on one-quarter-ahead earnings forecasts issued within 90 days after a quarterly earnings call in order to gauge analysts' beliefs about future firm earnings. We quantify analysts' sentiment as the difference between the median predicted earnings and its realized value, both scaled by the stock price.[6] Greater values of ξ^2_{it} indicate more optimistic earnings forecasts relative to what materializes. We check that our results are robust to alternative windows for which we could have collected analysts forecasts (e.g., 5 days and 30 days).

We merge measures of firm-level sentiment with basic income statement and balance sheet data from the CRSP-Compustat merger (WRDS) for the years 2006-2021. An essential variable for our analysis is the degree of sectoral downstreamness, which we define as in Antràs et al. (2012) and extend to 2021. In compact matrix notation, *upstreamness* is defined as follows:

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

where $\mathbf{1}$ is a column vector of ones, $\mathbf{I}$ is the identity matrix, and Δ is a square matrix with the numerator of the (s,k)-th entry of Δ, $d_{sk}Y_k$, the dollar value of commodity s used in k's production. The denominator $Y_s - X_s + M_s$ is computed as the total row-sum of values, less what is recorded under net exports and net changes in inventories. The s-th entry U_{st} of $\mathbf{U_t}$ contains the upstreamness measure for industry s in quarter t. We note that U_{st} is bounded below by unity. We obtain all the necessary data to construct Δ from the detailed Use Tables of the U.S. I-O Tables, provided by the Bureau of Economic Analysis (BEA). Because upstreamness is defined at the level of an industry, we assume that U_{st} is common for all firms $i \in s$. Our baseline analysis uses 3-digit BEA industry codes, although we

[5]Formally, let N_{ipt} and N_{int} denote the number of positive and negative segments in the earnings-call related to firm i at time t, respectively. Then, $\xi^1_{it} \equiv \frac{N_{ipt}-N_{nt}}{B_{it}}$. We multiply ξ^1_{it} by 100,000, as with the dictionary-based measure. We have also experimented, following Gorodnichenko et al. (2023), with an alternative definition: $\zeta_{it} = \frac{N_{ipt}-N_{int}}{N_{ipt}+N_{int}}$. Results (not shown) do not change.

[6]Let EPS_{it} denote firm i's realized earnings per share (EPS) in quarter t, and let $f^h[EPS_{it}]_{iht}$ denote the associated median analyst forecast that was issued within h days of the call date. Our third and final firm-level sentiment measure is computed as: $\xi^2_{iht} = f^h[EPS_{it}]_{iht} - EPS_{it}$ for a given h.

show that our main results are robust to alternative sectoral definitions (Section 2.5).[7]

A main advantage of the measure in (2) is that the economic interpretation is simple: values of U_{st} correspond to the number of transaction rounds necessary for a product to reach the final consumer. For example, the apparel sector is one of the most downstream in the sample with $U_{s,20021} = 1.05$ as of 2021. This suggest 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 level of greater than 3.00 suggesting that it takes this sector's goods more than three rounds of sales to reach the consumer.

Finally, Table A.1 in the Appendix presents basic summary statistics for our three sentiment measures, in addition to firm-level characteristics. Appendix A.1 provides further details on variables' construction and definitions.

2.2 The Cross Section of Firm-Level Sentiment

The first step of our empirical analysis involves constructing a ranking of the *informativeness* of firm-level sentiment for the macroeconomy. We then proceed to studying the determinants of firm-level informativeness in the next subsection. To start, for every firm in our sample, we run the following linear regression:

$$Y_t = \alpha_i + \beta_i \xi_{it} + \varepsilon_{it}, \quad \forall i \tag{3}$$

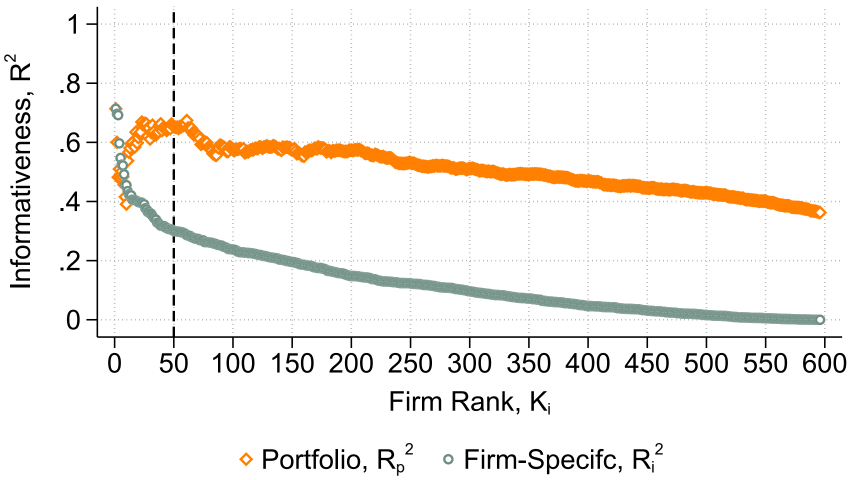
where Y_t is Hodrick-Prescott filtered (Ravn and Uhlig, 2002) real GDP, ξ_{it} is the baseline dictionary-based measure of firm-level sentiment, and ε_{it} is an error term.[8] We collect the $N \times 1$ vector of coefficients of determination R_i^2 from equation (3). There are at least three reasons why this is our preferred measure of informativeness. First, by design, it captures the share of output variation that can be attributed to a particular firm in our sample. Second, the measure is well defined on a non-negative interval, which will be useful for the second step of our analysis. Finally, the same metric can easily be computed in our calibrated model, which facilitates a simple and exact model-to-data comparisons.

We rank firms in the descending order of informativeness R_i^2 and denote the rank integer by K_i. We then construct N portfolios of sentiment, starting from the 1st ranked firm and proceeding iteratively, such that the 2nd portfolio is the average sentiment for firms ranked first and second, the third portfolio is the average of sentiment for firms ranked first, second, and third, and so on. We last run *portfolio-level* regressions of output

[7] Appendix A.3 provides additional distributional information and lists company names of the most downstream firms in the sample.

[8] In Section 2.5 we discuss extensively how our results from this exercise do not change if we use real GDP per capita, unemployment rate, CPI, or industrial production in place for Y_t.

Figure 1: The Cross Section of Firm-Level Sentiment



Notes: The figure demonstrates firm- (R_i^2) and portfolio-level (R_p^2) coefficients of determination with respect to HP-filtered U.S. real GDP with firm rank K_i on the horizontal axis.

on sentiment:

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

where ε_{pt} denotes average sentiment of the pth portfolio. Our object of interest is portfolio-level informativeness R_p^2.[9] We present the outcome of this exercise in Figure 1.

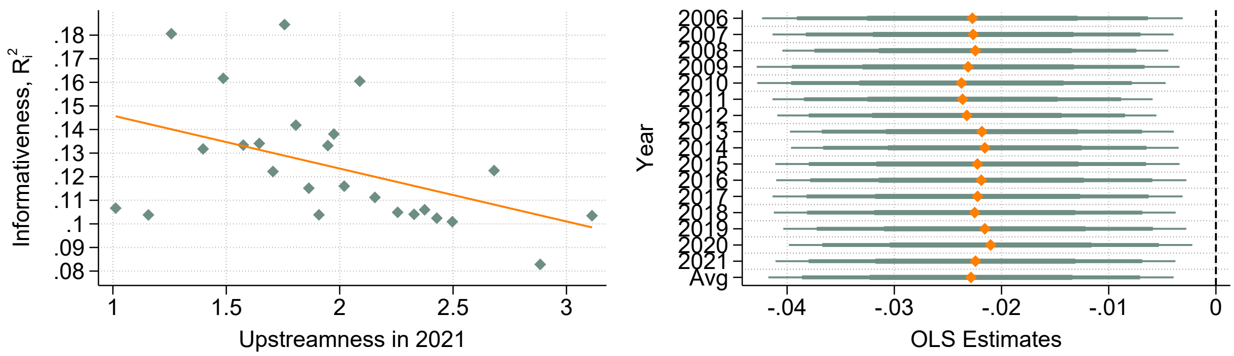
The curve demonstrating the relationship between portfolio-level informativeness and firm rank is increasing at first, but then inflects at around $K_i = 50$ and declines monotonically thereafter. The relationship between the two measures is almost "inverse-U shaped". Starting from the first portfolio that includes just the firm ranked K_1, adding more firms into the portfolio comes, at first, with a portfolio-wide informational benefit that exceeds idiosyncratic firm-level noise. This stops being the case at the inflection point after which the marginal cost from adding noisier firms outweighs any benefit. On balance, this pattern strongly suggests that the marginal effect of firm-level sentiment on business cycle fluctuations is not uniformly distributed. A flat curve in Figure 1 would instead have been consistent with this alternative interpretation. The sentiment towards some firms systematically seems to matter more than others.

Following the terminology in Gabaix (2011), we refer to firms with rank $K_i \leq K_{50}$ — that is, those that are in the $K_{p=50}$-portfolio — as belonging to a set Γ of *granular*

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

Figure 2: Downstreamness and Sentiment: OLS Estimates

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



Notes: Results from cross-sectional OLS 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. Standard errors are clustered at the sector level.

informativeness. These firms, whose average sentiment accounts for more than 60 percent of business cycle fluctuations, comprise around 8 percent of all firms in our sample. In this sense, we view the set of informationally granular firms as a *sufficient statistic* for sentiment-driven business cycles.[10]

Notwithstanding the aforementioned benefits of our informativeness measure, our approach potentially raises concerns in relation to identification and robustness. The above finding in Figure 1 that a small number of firms' sentiment is sufficient to capture business cycle dynamics is, nevertheless, highly robust. In Section 2.5, we show that this result survives a battery of extensions and robustness tests including, but not limited to, adding fixed effects—capturing e.g. firm-level differences in cyclicality—as well additional firm and aggregate controls to our regression specifications (e.g., firm-level and aggregate total factor productivity). This lends credence to the idea that a small subset of firms can capture the bulk of sentiment-driven business cycle fluctuations.

2.3 Downstreamness and Sentiment

In the previous subsection, we showed that the average sentiment of around 50 firms is highly correlated with the state of the U.S. business cycle. In this subsection, we analyze what characteristics these granular firms possess. We pursue an agnostic approach to answer this question. In the seminal work by Gabaix (2011), the absorbing characteristic, determining whether granular shocks produce aggregate fluctuations, is firm size. How-

[10]Notice that $R^2_{p=50}$ is substantially larger than $R^2_{p=N}$, where the latter includes all firms in the economy.

ever, it is not *prima facie* clear whether size is the dominant dimension for sentiments. We construct an array $\mathbf{X}_i$ of potential explanatory variables for each firm. We collect and/or compute the following variables: firm size (measured by log of total assets), market beta, the book-to-market ratio, investment intensity, market valuation, leverage, liquidity, Tobin's Q, as well as firms' idiosyncratic return volatility. Lastly, we include our measure of downstreamnes, discussed in Subsection 2.1. All of these variables could potentially account for our earlier findings. We collapse each variable into the $N \times 1$ dimension by taking the firm-specific average across sample years. We denote the resulting array with $\mathbf{X}_i$. Appendix A.1 provides further details on variable definitions and construction.

We then regress our firm-specific measure of informativness, R_i^2, onto the array of controls, $\mathbf{X}_i$, using three different specifications: (i) a linear OLS regression; (ii) a probit regression, where the dependent indicator variable takes a value of unity if the firm is in our granular informativeness set Γ and zero otherwise; and (iii) an ordinal probit regression, where the dependent variable is the rank integer K_i itself.

Panel (a) in Figure 2 presents the result graphically in two steps for the linear regression estimates. In particular, Panel (a) in Figure 2 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 from $\mathbf{X}$ are partialled out in the figure. All else equal, we find that firms that belong to a less upstream sector in 2021 have a higher R_i^2, i.e., are more closely correlated with business fluctuations. More *downstream* firms are more informative.

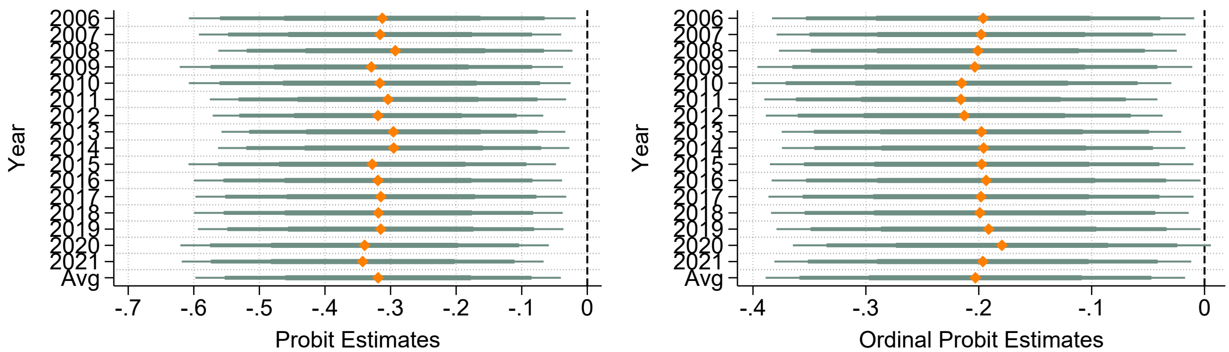
Table A.3 in the Appendix, instead, presents the results in a table format. In all cases, the only control that is systematically significant and economically important is sectoral upstreamness. All other variables, including size, matter little. All else equal, the more downstream a firm is the more likely it is to be included in the granular set Γ. The magnitude of the estimated effect is, furthermore, substantial: all else equal, increasing downstreamenss by one (that is, moving one step closer to the final consumer) increases the probability of being in the granular set by more than 30 percent (Column 4 in Table A.3).

The above result does not depend on the precise year used to measure upstreamness, or on the precise econometric specification. Panel (b) of Figure 2 plots coefficient estimates for *every* year in the sample. That is, we run 16 separate regressions with the relevant U_{it} for the corresponding year t. We observe remarkable coefficient stability across time.

Figure 3 shows analogous results for the probit regression and its ordinal probit counterpart. Furthermore, the results are, in each case, virtually unchanged from those reported in Table A.3. The relationship between informativeness and downstreamness is

Figure 3: Downstreamness and Sentiment: Probit and Ordinal Probit Estimates

(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.

robust along the time-series dimension and across different regression specifications.

In summary, the results in this subsection show that changes to the sentiment of firms that are downstream, all else equal, are closely correlated with aggregate fluctuations. The correct absorbing characteristic, which gives traction to the granular hypothesis in the context of sentiments, is donwstreamness. Section 2.5 offers a plethora of additional extensions and robustness tests for this cross-sectional relationship without affecting any of our conclusions.

2.4 Granular Sentiment over Time

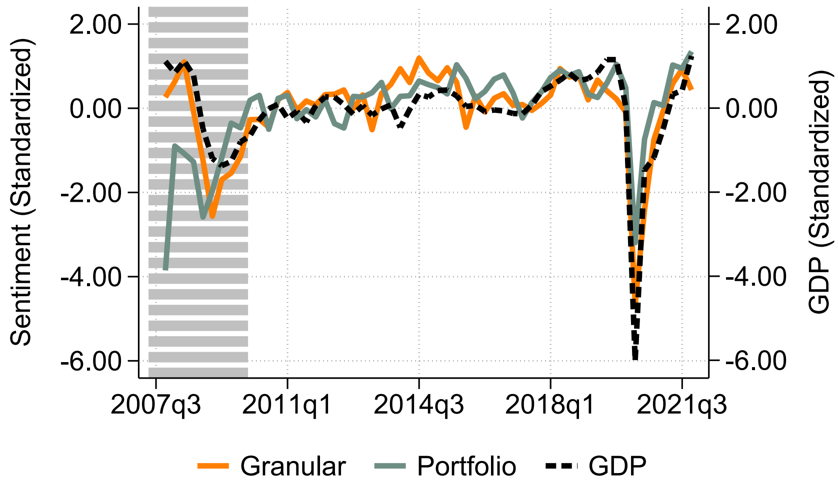
We next turn to the time-series implications of changes in granular sentiment. Following Gabaix (2011), we begin by computing a weighted average of firm-level sentiment ξ_{it} using previous-year donwstreamness as weights, $\omega_{i,t-4}$. Importantly, we restrict the set of firms to those in our granular informativeness set Γ. We denote the resulting time-series measure by $\mathcal{S}_t$:

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

where weights $\omega_{it} \equiv 1/U_{it}$ sum up to unity in each period.[11] To further safeguard against the selection on factors other than downstreamness driving our results, we also compute

[11]Compared to an equally-weighted average, using donwstreamness as weights attaches greater importance to movements in more downstream firms' sentiment, although this does not materially affect our results.

Figure 4: Granular Sentiment over Time



Notes: Time-series plots of granular sentiment $\mathcal{S}_t$, sentiment of the high-downstreamness portfolio $\mathcal{P}_t$, and output.

a weighted average, $\mathcal{P}_t$, of the sentiment of the 10 percent most downstream firms alone:

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

where $i \in P10_t$ signifies firms that belong to the top decile of all firms sorted by downstreamness in quarter t. Any empirical test that involves $\mathcal{P}_t$ is harder to pass, as it involves sorting based on a pre-defined characteristic. Figure 4 plots the time-series behavior of $\mathcal{S}_t$, $\mathcal{P}_t$, and detrended US output. All series have been standardized. On balance, our measures of granular sentiment track output fluctuations closely. Appendix A.4 plots additional time-series plots for indices that use our alternative measures of sentiment, derived using the FinBERT algorithm and analyst EPS forecasts, respectively. The tight relationship over time between sentiment and output carries over to these alternative measures.

Contemporaneous Relationship: We first explore the contemporaneous relationship between our measures of granular sentiment, $\mathcal{S}_t$ and $\mathcal{P}_t$, and various aggregate indicators. Importantly, in every exercise, we control for aggregate TFP (Fernald, 2014), overall macroeconomic uncertainty (Baker et al., 2016), the Federal Funds Rate, and real Nasdaq market returns. Table 1 presents the results.

The upper panel of Table 1 shows that $\mathcal{S}_t$ is strongly positively associated with economic- and market-based indices of aggregate sentiment — specifically, the Confer-

Table 1: Granular Sentiment and the Macroeconomy

	Independent Variable: Granular Sentiment, $\mathcal{S}_t$					
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Conference Board LEI	sentix Sentiment	GDP	Unempl.	Ind. Prod.	CPI
Sentiment	0.659***	0.522***	0.761***	-0.725***	0.886***	0.347***
	(0.151)	(0.087)	(0.130)	(0.162)	(0.088)	(0.124)
Observations	57	57	57	57	57	57
All Controls	✓	✓	✓	✓	✓	✓
aR^2 without Controls	0.619	0.654	0.731	0.658	0.720	0.270
aR^2 with Controls	0.701	0.713	0.829	0.743	0.813	0.277
	Independent Variable: High-Downstreamness Sentiment, $\mathcal{P}_t$					
Sentiment	0.507***	0.467***	0.388***	-0.386**	0.383***	0.054
	(0.112)	(0.101)	(0.127)	(0.153)	(0.124)	(0.135)
Observations	57	57	57	57	57	57
All Controls	✓	✓	✓	✓	✓	✓
aR2 without Controls	0.180	0.396	0.197	0.164	0.089	0.008
aR2 with Controls	0.654	0.727	0.652	0.590	0.525	0.221

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, uncertainty, the Fed Funds Rate, and real Nasdaq returns. Robust standard errors are in parentheses. *p < 0.1; **p < 0.05; ***p < 0.01.

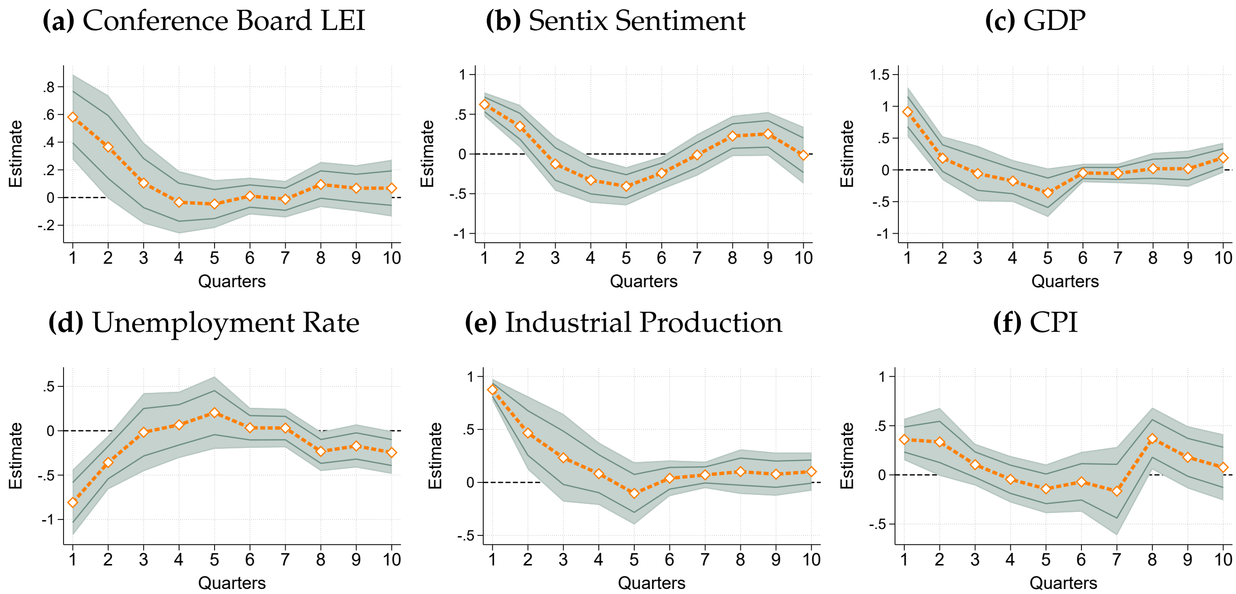
ence Board's Leading Economic Indicator (LEI) and the sentix Sentiment Index.[12] This suggests that $\mathcal{S}_t$, which is constructed from micro data alone, is consistent with prominent, aggregate measures of sentiment.[13] Next, we observe that $\mathcal{S}_t$ is also highly correlated with different measures of macroeconomic conditions: de-trended output, the unemployment rate, industrial production, and inflation. In each column of Table 1, we also present (adjusted) R^2 with and without the controls. By itself, $\mathcal{S}_t$ accounts for 62% and roughly 70% of the unconditional variation in aggregate sentiment and macroeconomic variables, respectively. Notably, $\mathcal{S}_t$ is *positively* associated with inflation. In conjunction with the positive correlation with aggregate quantities, this suggests that autonomous changes in sentiment of a small number of firms could trigger "demand-type" noise shocks (Lorenzoni, 2009). We return to this idea later in this section.

The bottom panel of Table 1 presents the results for the downstreamness index, $\mathcal{P}_t$. We find that, with the exception of inflation, our results do not change much: sentiment

[12] The LEI index provides an early indication of significant turning points in the business cycle and is available at https://www.conference-board.org/us/. The sentix Sentiment Index represents investors' market expectations over future months and is available at https://www.sentix.de/index.php/en/.

[13] There is also a high, positive association with other indicators of aggregate sentiment such as the Index of Consumer Sentiment from the University of Michigan (not shown).

Figure 5: Dynamic Effects of Granular Sentiment on the Macroeconomy



Notes: Local-projection estimates of granular sentiment $\mathcal{S}_t$ on macroeconomics aggregates. Lines correspond to 68% and shaded areas are 90% confidence intervals. Standard errors are clustered at the quarterly level.

of the most downstream firms is highly correlated with measures of aggregate sentiment and macroeconomic indicators. While the (adjusted) R^2 understandably drop, as this is a more stringent empirical test, we still find that $\mathcal{P}_t$ accounts for around 28% and 15% of fluctuations in aggregate sentiment and macroeconomic dynamics, respectively.

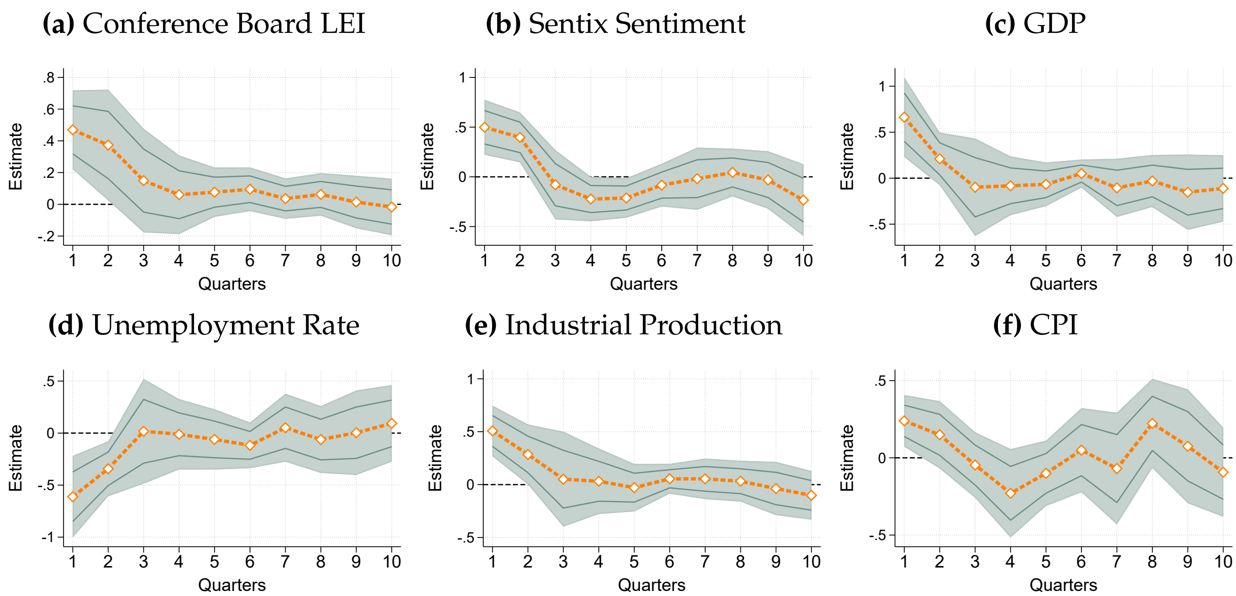
Dynamic Relationship: We proceed to estimate the dynamic effects of granular sentiment on the macroeconomy. To this end, we run Jordà (2005)-style local projections:

$$Y_{t+h} = \delta_h + \beta_h \times \mathcal{S}_t + \sum_{\ell=1}^{2} \gamma_{h\ell} \mathbf{X}_{t-\ell} + u_{ht}$$

where δ_h are horizon-specific fixed effects. The main coefficient of interest, β_h is that on granular sentiment $\mathcal{S}_t$. Outcome variables, Y_{t+h}, are once again aggregate sentiment and macroeconomic indicators. We use the same vector of controls $\mathbf{X}_t$ as before, and include 2 lags of every control variable in the regressions.

Figure 5 depicts the resulting impulse response functions for horizons of up to $h = 10$ quarters ahead. Positive changes in granular sentiment, all else equal, induce "business-cycle-like" synchronized movements: economic and market sentiment spike, output and industrial production rise, unemployment falls, and inflation rise. These effects are

Figure 6: Dynamic Effects of High-Downstreamness Portfolio on Macroeconomy



Notes: Local-projection estimates of the high-downstreamness portfolio sentiment $\mathcal{P}_t$ on macro aggregates. Lines are 68% and shaded areas are 90% confidence intervals. Standard errors are clustered at the quarterly level.

generally economically large and statistically significant for up to around three quarters. Indeed, fluctuations in sentiment towards just 50 downstream firms appear to lead to coordinated spikes in economic activity and inflatation closely reminiscent of those triggered by ”economy-wide noise shocks” (Lorenzoni, 2009) — both in terms of cross-variable correlation and dynamics and size. Figure 6 reports local projections for the high-downstreamness index, $\mathcal{P}_t$, and shows similar results and interpretations to those we arrive at in Figure 5.

Instrumented Changes to Sentiment: Although our results in Table 1 and Figures 5 and 6 control for macroeconomic conditions (e.g., aggregate TFP and macroeconomic uncertainty), it could perceivably still be the case that fluctuations in sentiment are partially endogenous to the outcome variables we consider. To alleviate such concerns about identification, we instrument for changes in $\mathcal{S}_t$. In particular, following Lagerborg et al. (2022), we choose mass shooting fatalities in the U.S. as an instrument for sentiment fluctuations. The exclusion restriction for this instrument is verified since, as argued in Pappa et al. (2019) extensively, fatal shootings in the U.S. are largely randomly assigned across regions and time. Furthermore, the instrument is relevant; the literature has shown it to be highly correlated with measures of consumer and firm sentiment. It is therefore also relevant for $\mathcal{S}_t$, because our micro-based $\mathcal{S}_t$ is itself a good proxy of aggregate sentiment

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:	Conference Board LEI	sentix Sentiment	GDP	Unem-ployment	Ind. Prod.	CPI
Sentiment	0.769***	0.903***	0.521***	-0.744***	0.724**	-0.108
	(0.261)	(0.340)	(0.161)	(0.201)	(0.299)	(0.381)
Observations	45	45	45	45	45	45
All Controls	✗	✗	✗	✗	✗	✗
aR^2	0.725	0.431	0.587	0.258	0.638	-0.094
First stage F-stat	6.857	6.857	6.857	6.857	6.857	6.857
	Independent Variable: Granular Sentiment instrumented by Mass Shootings					
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Conference Board LEI	sentix Sentiment	GDP	Unem-ployment	Ind. Prod.	CPI
Sentiment	0.462**	0.668**	0.270*	-0.610***	0.428	-0.457
	(0.233)	(0.335)	(0.164)	(0.202)	(0.296)	(0.546)
Observations	45	45	45	45	45	45
All Controls	✓	✓	✓	✓	✓	✓
aR^2	0.808	0.581	0.791	0.676	0.749	-0.055
First stage F-stat	4.870	4.870	4.870	4.870	4.870	4.870

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

indicators, as shown earlier.

We run contemporaneous 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. In all cases, we find that granular sentiment, when instrumented with mass shootings, has significant effects on different measures of overall sentiment, output, and unemployment. The impact on industrial production is economically similar albeit with larger standard errors than before. The impact on inflation, however, cannot be statistically distinguished from a noisy zero. The relatively muted effect on inflation was already evident from the dynamic effects in Figure 5 and 6.[14]

Overall, we conclude that our previous results about the relationship between granular sentiment and the macroeconomy are not due to reverse causation, but can be interpreted

[14] An important qualifier to this discussion is the low values of first-stage F-statistics. Lagerborg et al. (2022) also note that depending on the exact specification and definition of the mass shootings instrument, instrument power can fluctuate.

causally. Orthogonal increases in granular sentiment appear to drive business cycle fluctuations in a manner reminiscent of standard demand shocks. A positive change in sentiment increases output and industrial production, and decreases unemployment and inflation. In the next subsection, we evaluate the general robustnesss of this finding, in addition to that of our previous cross-sectional result that a select number of downstream firms account for most of the fluctuations in aggregate sentiment.

2.5 Discussion, Extensions, and Robustness Tests

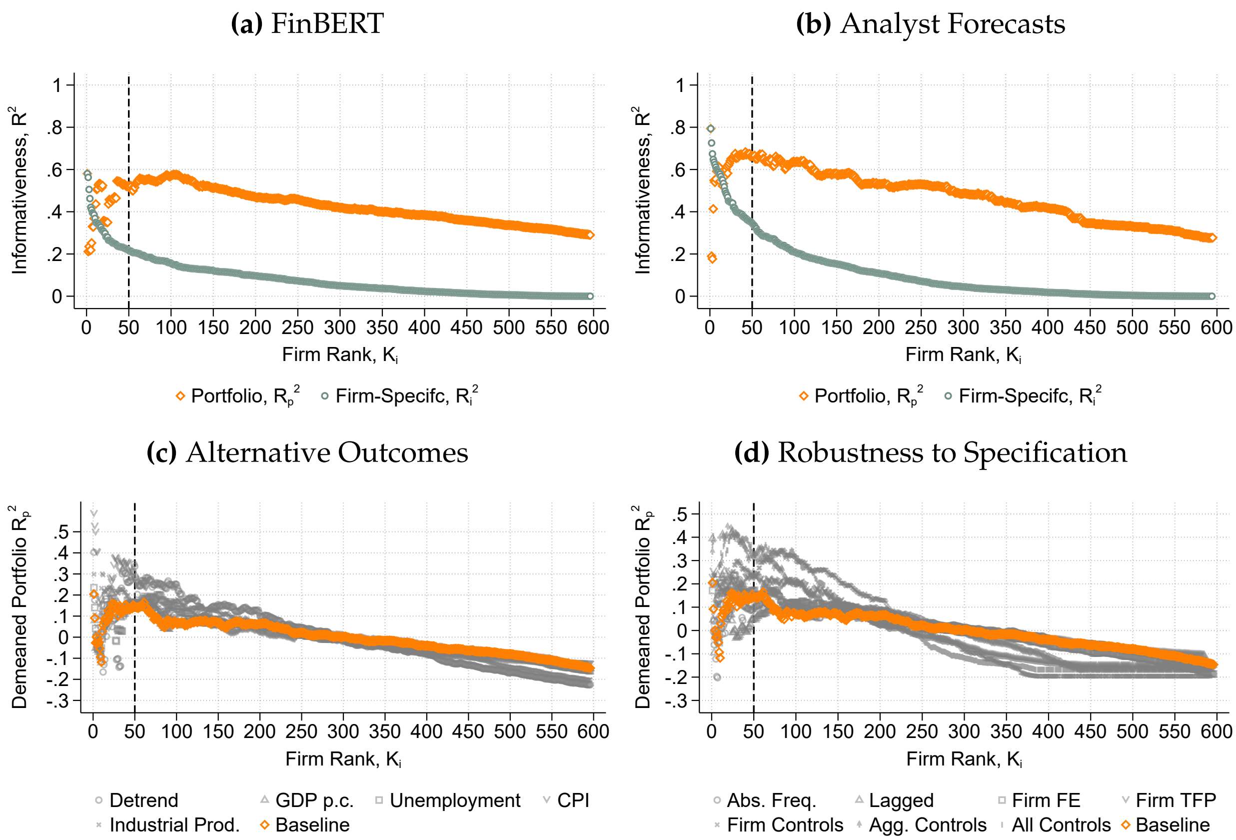
In this section, we discuss various extensions to our baseline empirical specifications. There are four general themes to our extensions, some of which we repeat at various layers of our analysis. First, we consider alternative measures of firm-level sentiment — specifically, the FinBERT measure, ξ^1_{it}, and the analyst forecast-based measure, ξ^2_{it}. Second, we consider alternative dependent variables in equation (3) — in particular, the unemployment rate, inflation, industrial production, as well as an alternative measure of detrended output. The alternative approach to detrending involves residualizing a variable from the time-fixed effect. We also here re-estimate equation (3) using lagged sentiment $\xi_{i,t-1}$, to account for any potential sluggishness in the relationship. Third, we adjust our measure of firm-level sentiment ξ_{it} by multiplying the measure with the number of words in each transcript, yielding the absolute frequency of net sentiment. Lastly, fourth, we residualize our sentiment measure ξ_{it} from potentially confounding factors to confirm that firm-level sentiment is truly "idiosyncratic". To this end, we extract the component of firm-level sentiment that cannot be explained by firm fixed effects or time-varying firm-specific and aggregate controls. In particular, we partial out the effects of firm productivity, market beta, (log) total assets, the book-to-market ratio, investment intensity, market value, leverage, liquidity, Tobin's Q, as well as the firm's idiosyncratic return volatility.

Notice that, for our purposes, it is important to control for various measures of productivity. This is because our theoretical framework below allows for changes in sentiment that are independent from productivity. The empirical specification used to extract refined sentiment is the following:

$$\xi_{it} = \mu_i + \alpha_1 A_{it} + \alpha_2' \mathbf{X}_{it} + \alpha_3' \mathbf{Z}_t + \hat{\xi}_{it} \tag{7}$$

where μ_i is the firm fixed effect, A_{it} is firm productivity, $\mathbf{X}_{it}$ is the vector of firm controls exlcuding productivity, and $\mathbf{Z}_t$ is the vector of aggregate controls. We extract five versions of the residual $\hat{\xi}_{it}$ from the above regression by adding controls step-by-step: first, we

Figure 7: The Cross Section of Firm-Level Sentiment — Robustness

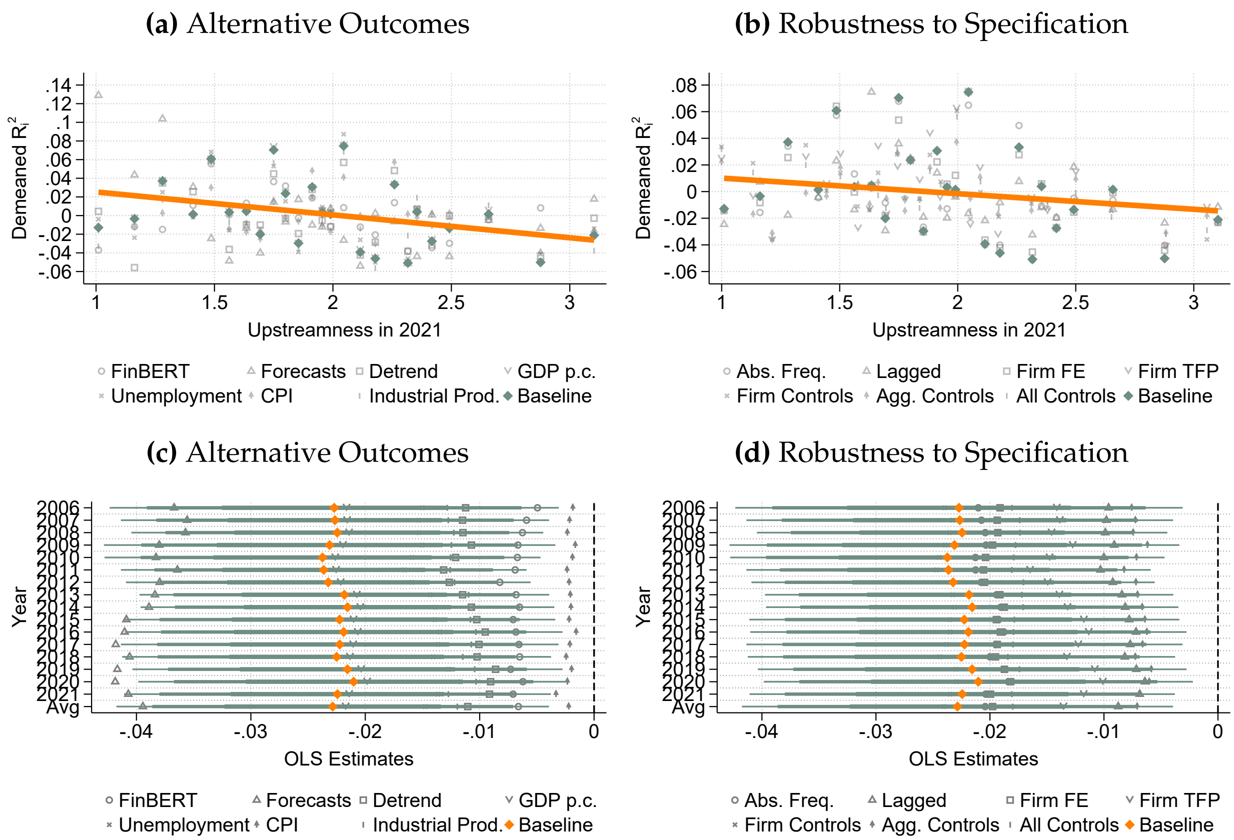


Notes: Firm- (R_i^2) and portfolio-level (R_p^2) coefficients of determination with respect to U.S. real GDP with firm rank K_i on the horizontal-axis under alternative specifications. Panels (a) and (b) report the results for FinBERT sentiment ξ^1_{it} and analyst forecast errors ξ^2_{it}, respectively. Panels (c) and (d) report the results under alternative outcomes and robustness to specification as described in the text.

add only μ_i; second, μ_i and A_{it}; third, μ_i, A_{it}, and $\mathbf{X}_{it}$; fourth, μ_i, A_{it}, and $\mathbf{Z}_t$; and finally, fifth, μ_i, A_{it}, X_{it}, and $\mathbf{Z}_t$. We denote the respective sentiment measures by $\hat{\xi}^j_{it}$, j=1,2,. . . ,5. We construct our measures step-by-step to understand which variables, if any, affect our results. We now proceed with our various extensions.

Granular Sentiment Firms: We begin by expanding on our first main result on the cross section of firm-level sentiment and the existence of a small number of firms whose sentiment is closely associated with the overall economy. We re-run equation (3) using alternative measures of firm-level sentiment and alternative dependent variables. We, in each case, compute the definition-and-outcome specific portfolio R-squared. Panels (a) and (b) in Figure 7 present the results for the two alternative measures of sentiment. Panels (c) and (d), by contrast, report the results using alternative macro outcomes, alternative means of de-trending, and various refinements of $\hat{\xi}_{it}$. As the *level* of R_p^2 differs across

Figure 8: Downstreamness and Sentiment — Robustness



Notes: Results from cross-sectional OLS 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.

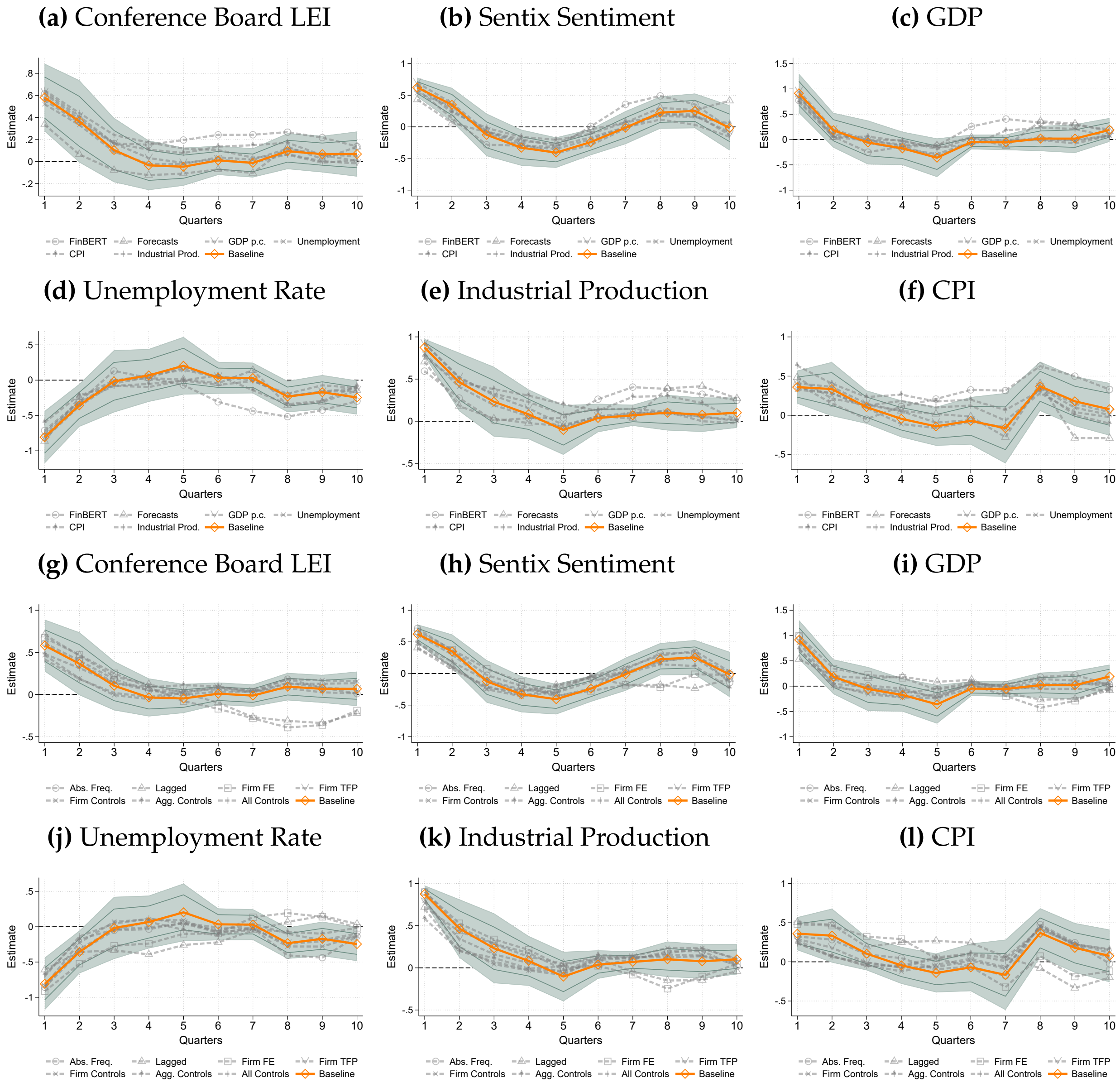
specifications, we de-mean each R_p^2 in Panels (c) and (d). Across all of the panels in Figure 7, the basic relationship between informativeness rank, K_i, and R_p^2 is the same, and identical to that in our baseline. There is, all else equal, an inverse-U-shaped relationship between informativeness and portfolio R^2 with inflection point at around 50 firms.

Cross-Sectional Analysis: Next, we present expanded results on the cross-sectional relationship between informativeness and sectoral upstreamness. Panels (a) and (b) of Figure 8 show binned scatter plots for the year 2021 using different specifications. Panels (c) and (d) report coefficients and confidence intervals for each year together with the overall sample average. On balance, we find that our main finding is highly robust. With the exception of the specification which uses inflation as the dependent variable in equation (3), all point estimates lie within the confidence interval of our baseline model. The negative association between upstreamness and informativeness, R_i^2, is a general feature of the data, even after the inclusion of additional controls and the use of alternative measures.

Time-Series Analysis: Finally, we analyze the robustness of our estimated dynamic effects of granular sentiment on the macroeconomy. For each alternative specification, we run the same local projection as in the baseline case. Figure 9 reports the results in two parts. First, Panels (a)-(f) show dynamic responses under the first and second extension theme (with the exception of alternative de-trending for which we present results separately in Appendix A.4). Second, Panels (g)-(l) show dynamic responses under the third and fourth extension themes. We find that the contemporaneous effects of granular sentiment on the macroeconomy are sizable, as practically every estimate lies within the confidence band of our baseline specification. The estimated impact, furthermore, usually remains statistically significant for up to 3 quarters. As a result, we conclude that the dynamic effects of granular sentiment on the macroeconomy are robust as well.

Further Robustness Tests: In the Appendix, we supplement the results discussed in the main text with various additional tests. First, Appendix A.3 reports cross-sectional estimates using an alternative definition of a "sector". We employ either the Atalay (2017) or the Chahrour et al. (2021) definition and find that our results do not change. This Appendix also provides supplementary tables and descriptive information on downstream industries and firms. Second, Appendix A.4 presents a full set of local projection estimates using the alternative de-trending approach. Finally, Appendix A.5 offers additional tests of robustness. In particular, we run placebo tests for our baseline cross-sectional exercise and find that, in each year of the sample, it is highly unlikely that the positive relationship between informativeness and downstreamness was obtained by chance.

Figure 9: Dynamic Effects of Granular Sentiment on the Macroeconomy—Robustness



(a) Conference Board LEI **(b)** Sentix Sentiment **(c)** GDP

(d) Unemployment Rate **(e)** Industrial Production **(f)** CPI

(g) Conference Board LEI **(h)** Sentix Sentiment **(i)** GDP

(j) Unemployment Rate **(k)** Industrial Production **(l)** CPI

Notes: Local-projection estimates of granular sentiment $\mathcal{S}_t$ on macro aggregates under alternative specifications. Panels (a)-(f) report the first set of results under alternative macro outcomes. Panels (g)-(l) report the second set of results under robustness to specification, as described in main text. Lines are 68% and shaded areas are 90% confidence intervals. Standard errors are clustered on time.

3 Theoretical Framework

We study the effects of sentiment shocks to firms in a simple multi-sector economy. 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. There are 2 sectors in the economy, an upstream (u) and a downstream (d) sector. Each sector consists of a continuum of firms that sell their goods in perfectly 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 modelling structure is similar to what has been used in Acemoglu et al. (2012) and Chahrour et al. (2021), but with a crucial exception that firms make *endogenous* attention choices.

Sectors and Firms: A firm in sector i at time $t = 1, 2, \ldots$ 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}. 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. The 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 observes

$$s^{z}_{ijt} = z_{jt} + \xi_{it} + m_{i,j}\varepsilon_{ijt}, \quad \xi_t \sim \mathcal{N}\left(0, \sigma^2_{\xi}\right), \quad \varepsilon_{ijt} \sim \mathcal{N}(0,1), \tag{12}$$

where $z_{jt} \equiv \log Z_{jt}$, $m_{i,j} \geq 0$, and $j = \{u, d\}$. The shocks $\left\{\varepsilon_{ijt}\right\}$ and $\{\xi_{it}\}$ are independent of each other, sectoral productivity, and across time. Nature draws

$$z_{jt} = \vartheta_t + u_{jt}, \quad u_{jt} \sim \mathcal{N}\left(0, \sigma^2_u\right) \tag{13}$$

where $\vartheta_t \sim \mathrm{AR1}\left(\rho_{\vartheta}, \sigma^2_{\vartheta}\right)$ is independent of all other shocks. Crucially, the disturbances ξ_{it} captures sector-specific "noise shocks" to firms' information about their own and others' productivity — i.e., sentiment shocks.

In addition to the information observed about sectoral productivity, each sector i also observes a noisy signal of previous period's output of the other sector $j \neq i$:

$$s^{q}_{ijt} = q_{jt-1} + m_{j,q}e_{it}, \quad e_{it} \sim \mathcal{N}\left(0, \sigma^2_e\right), \tag{14}$$

where $m_{j,q} \geq 0$. We summarize the information firms base labor choices on in the information set $\Omega_{it} \equiv \left\{s_{i0} \cup (s_{is})^{s=t}_{s=1}\right\}$, where $s_{it} \equiv \left(s^z_{iit}, s^z_{ijt}, s^q_{ijt}\right), j \neq i$.[15]

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, while for the household labor supply is optimal.

Finally, for each i, we assume that firm i's attention vector $m_i \equiv \begin{pmatrix} m_{iu} & m_{id} & m_{jq} \end{pmatrix} \in \mathbb{R}^3_+$ is chosen to maximize firms' *ex-ante* profits at the start of time, subject to an attention cost K(q), where $K(\cdot)$ is positive, decreasing in all elements of q, and convex.

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

[15] 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.

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 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 $\beta \equiv (\ 1 \quad 0\ )'$.[16] 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 Weight*). It is the dot product of the transpose of the ith column of 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 will later discuss the relationship between Λ_i and our measure of *downstreamness* from Section 2.

Second, although our information structure is simple, it allows us to highlight two distinct channels which cause attention to centre 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 channels operates even without the presence of the *endogenous information* in (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 the two different channels further below.

Finally, notice that *aggregate value added output* in our economy is equal to consumption C_t of the downstream good. This is the measure of output that we focus on.[17]

4 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.

[16]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.

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

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

4.1 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 backwards 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 condition (8), by equating the marginal product of the intermediate good with its marginal cost:

$$\alpha_{ij} \cdot \frac{P_i Q_i}{X_{ij}} = P_j. \tag{15}$$

In the *first stage*, firms internalize these choices when determining their optimal labor input. The labor input is set so as to maximize expected profits $\mathbb{E}_{it}[\Pi_{it}]$. This is done by equating the expected marginal product of labor with its marginal cost, the real wage, which results in:

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

where $\mathbb{E}_{it}[\cdot] = \mathbb{E}[\cdot \mid \Omega_{it}]$ denotes firm i's conditional expectation. 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.

4.2 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 and the vector of final sales shares $\beta = \begin{pmatrix} 1 & 0 \end{pmatrix}'$ can be used to show that $P_{it}Q_{it} = \lambda_i \cdot C_t$, which allows us to substitute for firm revenue in (16). In order to solve for firms' labor choices, what remains is to solve for economy-wide output 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 first stage labor choices, the (log of) the output of the downstream sector can be derived as

$$c_t = \beta'(\mathbf{I} - \mathbf{A})^{-1} \left(\mathbf{z}_t + \varkappa_0 + \varkappa_1 \ell_t\right), \tag{18}$$

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

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

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

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

Proposition 1 characterizes firms' input choices in our economy. 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 firms' attention choices, $\{\mathbf{m}_i\}$.

4.3 Optimal Attention Choices

In this subsection, we discuss the two separate forces that lead firms to, all else equal, pay closer attention to the downstream sector: (i) that fluctuations in downstream-sector productivity are more important for aggregate demand; and (ii) that the downstream sector's output provides a natural "information agglomerator". Before discussing the decomposition of attention choices into these two separate forces, we, however, briefly turn to the question of which firms have a large incentive to pay more attention altogether.

Proposition 1 shows that a firm's input choice, L_i depends on its expectation of sectoral productivities, $\exp(\Lambda' \mathbf{z}_t)$, as well as its expectation of other sector's input choices, $\sum_j \exp(\ell_{jt})$. All else equal, any errors in expectations, due to a lack of attention, matter more for firms who (i) use labor intensively in production (i.e., have a high δ_i), and (ii) are 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.

1. Attention and Downstreamness: The first reason firms, all else equal, prefer to pay attention to the downstream sector can be seen in Proposition 1. Notice that, in equation (19), the expectation of sectoral productivity z_t is multiplied by the vector of centrality weights Λ. To first order, the log of consumption demand in the economy, and hence, all else equal, the demand for a firm's product, depends on the dot-product between the vector of centrality weights, Λ, and the vector of sector-specific productivity shocks, $\mathbf{z}_t$ (see equation (18)). This result carries over to substantially more general frameworks than the model we consider here, due to Hulten's celebrated theorem (Hulten, 1978). As a consequence, firms in our economy, all else equal, have an incentive to pay closer attention to sectors that are central in the production network (i.e., have a large λ_i). Productivity fluctuations in these sectors, all else equal, move around aggregate demand by more, and hence create larger fluctuations in the demand for an individual firm's output. 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 demand for its product. Corollary 1 connects a firm's network centrality , 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 calibration of our model. In this case, both upstream and downstream firms naturally pay more attention to the downstream sector, as fluctuations in productivity of the downstream sector move around demand by more.

Strategic Interactions: The strength of the effect of downstreamness on sectoral attention is modulated by the degree of strategic interactions between firms. Proposition 1 shows there are two channels by which other firms' input choices affect a given firm's labor decision. 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. 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 and dampens the incentive 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 used in quantitative flex-price models (e.g., Rogerson and Wallenius (2009)). The presence of strategic complementarities in firms' labor choices further amplifies their preference to pay attention to the downstream sector: as in Hellwig and Veldkamp (2009), strategic complementarities in actions lead to strategic complementarities in information choice.

2. Attention and Information Agglomeration: The second reason firms 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 $j \neq i$ (see equation 14). These signals are informative for firms in the current period as they reveal information about previous period's (and hence current period's) productivity. But now notice that the downstream sector utilizes the upstream sector's input in production, while the converse does not occur (equation 8). As a result, the downstream sector's output, q_{dt}, embeds information about the upstream sector's productivity, z_{dt}. This makes the downstream sector's output, all else equal, a more "information-cost efficient" signal of overall demand and productivity. In this sense, the downstream-sector outcome in our economy plays a similar role to that of market-clearing prices in the Grossman and Stiglitz (1980) famous analysis: it agglomerates the dispersed bits of information that exist about productivity in the economy, which causes firms' attention to center on it.

5 Quantitative Analysis

In this section, we quantitatively evaluate the effects of sector-specific sentiment shocks. We show that attention centers on downstream firms, because they are more central and agglomerate dispersed information. As a result, we find that sentiment shocks to downstream firms drive close to 90 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 definition. Our results confirm those from the baseline model.

5.1 Numerical Solution and Parametrization

Numerical Solution: The limited attention model does not permit an analytical solution. Instead, we solve the model numerically, using a parameterized expectations algorithm akin to that used in Chahrour et al. (2021) (Appendix B.3). 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 $\{m_i\}$, 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 $\{L_i\}$ and expectations in (19) by solving it with the help of an parameterized expectations algorithm. We iterate on firms' labor choices until convergence in the sense of absolute differences of $\{L_i\}$.

Second, we keep firms' labor choices $\{L_i\}$ fixed and derive new values for firms' optimal attention choices $\{\mathbf{m}_i\}$. We do so by deriving an expression for a firm's *ex-ante* profits as a function of its attention choice. We then maximize this expression with respect to the firm's attention vector $\{\mathbf{m_i}\}$. We stop 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 $\gamma = 1/3$, consistent with the average share of capital income in the economy, and set $\nu = 3$. 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 α_{ij}, as such, equal those for sectors that are above the median measure of upstreamness. The vector $\beta \equiv \begin{pmatrix} 1 & 0 \end{pmatrix}'$ is by assumption. We take standard values for the productivity process: we set $\sigma_\vartheta = \sigma_u = 1$ and $\rho_\vartheta = 0.85$. These values are all within the range used in standard dynamic stochastic general equilibrium models.

The baseline parametrization also targets moments of our Granular Sentiment Index $\mathcal{S}_t$ — in particular, we set $\rho_\xi = 2/3$ and $\sigma_\xi = 1$. 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 the 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 $\mu = 2.1e^{-4}$.

5.2 Implied Attention Choices

Recall from our earlier discussion that attention $\{\mathbf{m_i}\}$ centers on the downstream sector (i) because it is more central in the production network; and (ii) because its output better agglomerates sectoral information about productivity. Table 4 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

Table 3: Attention Choices in the Baseline Model

Sector	Std. TFP	Λ	Attention (m_d, m_u, m_q)
Downstream	1.90	1.31	(1.00, 2.02, 10.00)
Upstream	1.90	0.34	(2.03, 3.05, 3.00)

Notes: This table summarizes general-equilibrium attention allocation across the two sectors. Columns two through four represent, in order, volatility of sectoral productivity shocks, production network centrality, and the attention choice vector $\{\mathbf{m}_i\}$.

than their *own* sector's productivity ($m_d < m_u$). Specifically, in our baseline calibration, firms in the upstream sector pay around 50 percent more attention to downstream firms ($m_d = 2.03$ vs $m_d = 3.05$). This is, in part, because the downstream sector is around four times more central in the production network, as measured by the discrepancy between $\lambda_d > \lambda_u$. Finally, Table 4 shows that the upstream sector pays a comparable amount of attention to the downstream sector's output as it does to its own productivity ($m_d \approx m_q$). This is in line with our earlier discussion demonstrating that the downstream sector's output better agglomerates sectoral information about productivity.

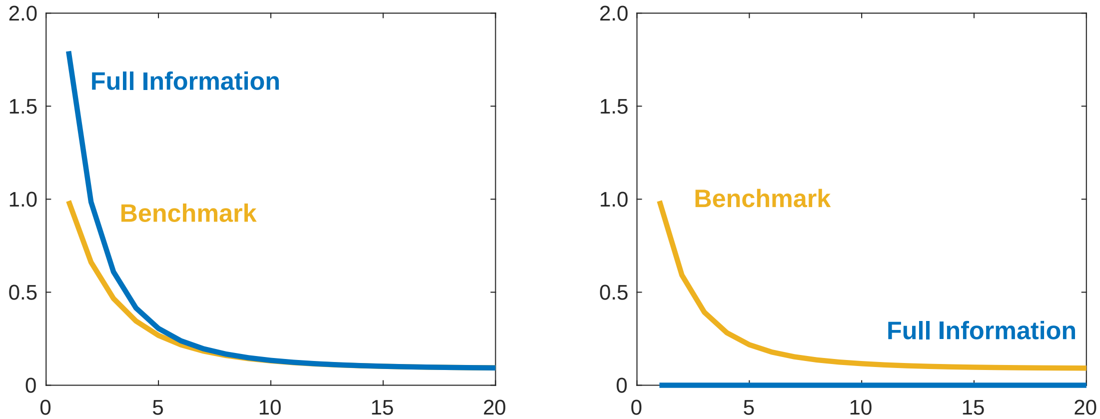
On balance, the results in Table 4 are consistent with our empirical findings in Section 2: because of the closer attention paid to the downstream sector, innovations to sectoral information about the downstream sector will, all else equal, have a greater impact on firms' expectation about aggregate productivity and output in the economy. As a result, any erroneous information about downstream sector productivity — i.e., any sentiment shocks — will also have larger effects. We now further explore the implications of this asymmetry in attention.

5.3 Shock Propagation and Sentiment

Because of the presence of limited attention ($\mu > 0$), the response of the economy to fundamental productivity shocks is dampened. The left panel of Figure 10 demonstrates this standard 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. But, due to the presence of information frictions, the output response under limited attention is dampened and more persistent.

The flip-side of the decreased responsiveness to fundamental disturbances is an increased responsiveness to sentiment shocks. The right panel of Figure 10 showcases the

Figure 10: Impulse Response to Productivity and Sentiment Shocks



Notes: The left panel showcases the impulse response function of (log) 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}.

output response to a one-standard-deviation increase in sentiment of 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 to the upstream sector are similar albeit more muted. In response to a positive sentiment shock, firms across the economy believe the downstream sector is more productive (see equation 12) and will produce more. This increases overall demand. As a result, firms in *both* sectors increase the demand for inputs and increase production, which in turn raises the equilibrium level output in the economy — a boom occurs. 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 subsides back down.

In sum, Figure 10 shows that both fundamental and sentiment-driven fluctuations in output can arise in the equilibrium in 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 sector, respectively.

5.4 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. We compare the results under our baseline calibration to those under full information in Table 4. The middle column shows that, under full information, all fluctuations in output are driven by changes in fundamental shocks — in this case, productivity. This contrasts with the results under our baseline calibration where more than 1/4 of business cycle fluctuations are driven by changes in sentiment. This estimate for the share of business cycles attributable to sentiment-driven shocks is consistent with the estimates of the effects of economy-wide sentiment shocks in e.g., Blanchard et al. (2013) and Chahrour and Ulbricht (2023), who find that around 1/6-1/3 of the business cycle can be attributed to sentiment-driven disturbances.

However, crucially, Table 4 also provides the decomposition of sentiment-driven fluctuations into those that originate from the downstream and the upstream sector, respectively. 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 90 percent of overall sentiment-induced fluctuations are due to the downstream sector. This showcases one important consequence of firms' asymmetric attention choices (Table 4). Despite the equal volatility of fundamental productivity, firms across both sectors pay substantially closer attention to the downstream sector. 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 fluctuations.

Clearly, the calibrated results in first column provide 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. As a result, 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.

Sentiment Shocks in an Extended 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. All other

Table 4: Business Cycle Attribution

	Share of Output Variance $\mathbb{V}[y_t]$		
	Benchmark	Full Information	27 Atalay Sectors
Productivity	0.72	1.00	0.81
Sentiment	0.28	0.00	0.19
– downstream sector	0.25	0.00	0.18
– upstream sector	0.03	0.00	0.01
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.

details of the calibration produce are identical to those described in Subsection 5.1. The third column of Table 4 presents the results.

The decrease in the size of sectors (as a result of the increase in their number) has two opposing effects on our quantitative results. First, sector-specific disturbances, all else equal, become less important.[18] This, in turn, makes sector-specific disturbances — including sentiment shocks — less important for output fluctuations in absolute terms. Second, as we increase the number of sectors, we also increase differences in measures of downstreamness and network centrality. In turn, this elevates the extent of asymmetric attention in our economy, as firms economize on their attention by only focusing on a handful of sectors. All else equal, this increases the effects of sentiment shocks towards these sectors.

The third column of Table 4 shows that, on balance, these two forces close to cancel each other out in equilibrium. The calibrated effects of sentiment shocks declines to around 20 percent of output fluctuations, 18 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 1 percentage point to output fluctuations. The sentiment-induced business cycle is, in this sense, *granular*: roughly 20 percent of sectors are responsible for 90 percent of the sentiment-driven aggregate fluctuation. We thus revisit the ubiquitous Pareto principle but now in the context of sentiments.

All in all, the results in Table 4 are comfortingly consistent with our earlier empirical

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

results, showing that a small share of firms account account for most of sentiment-induced fluctuations (Section 2). Clearly, the precise share of firms/sectors does not map one-to-one into our empirical analysis, because of differences in data coverage between firms in the Capital IQ sample and the overall economy. Notwithstanding such concerns, however, our results in this section have shed light on the mechanisms and quantitative potential for sentiment shocks to a subset of firms to drive aggregate fluctuations to a meaningful extent.

6 Conclusion

Because macroeconomic variables, 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-driven fluctuations and beliefs are, in this respect, no different. While the "micro origins" hypothesis has been applied to fundamental shocks like productivity, an application to sentiment and beliefs has so far been missing. In this paper, we have attempted to fill this gap.

Our focus—both empirically and theoretically—have been on firms and how business sentiment of a subset of these drives aggregate beliefs and fluctuations. One natural direction for future research is thus an extension of the Pareto principle and of our framework to other key agents in the economy: e.g., households or investors. Beliefs of subset of these could, in principle, matter for aggregate fluctuations. We leave this for future research.

Finally, as mentioned in the beginning, the design and conduct of optimal policy benefits from accurate measurements of the true origin of economic fluctuations. Our research concludes that it proximity to the final consumer— and not necessarily book or market size —that makes firm-level sentiment important. Regulation of information disclosure by listed companies should, as a consequence, pay particular close attention to central, downstream firms, as disclosures by such informationally granular firms, all else equal, have a larger impacts on the broader economy.

References

Acemoglu, D., A. Ozdaglar, and A. 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.*

— **, V. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi**, "The network origins of aggregate fluctuations," *Econometrica*, 2012, *80* (5).

Angeletos, G-M. and J. La'O, "Sentiments," *Econometrica*, 2013, *81.*

— **and K. Sastry**, "Inattentive Economies," *NBER Working Paper 26413*, 2019.

— **, F. Collard, and H. Dellas**, "Quantifying confidence," *Econometrica*, 2018, *86* (5), 1689–1726.

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

Antràs, P., D. Chor, T. Fally, and R. Hillberry', "Measuring the Upstreamness of Production and Trade Flows," *American Economic Review: Papers & Proceedings*, 2012, *102.*

Araci, D., "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, E., "How Important Are Sectoral Shocks?," *American Economic Journal: Macroeconomics*, 2017, *9.*

Baker, S., N. Bloom, S. Davis, and A. Notes, "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics*, 2016, *131(4).*

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

Beaudry, P. and F. Portier, "An exploration into Pigou's theory of cycles," *Journal of Monetary Economics*, 2004, *51.*

— **and** — , "Stock Prices, News, and Economic Fluctuations," *American Economic Review*, 2006, *96.*

Bigio, S. and J. La'O, "Distortions in Production Networks*," *Quarterly Journal of Economics*, 2020, *135.*

Blanchard, O., J-P. L'Huillier, and G. Lorenzoni, "News, Noise, and Fluctuations: An Empirical Exploration," *American Economic Review*, 2013, *103.*

Bordalo, Pedro, Nicola Gennaioli, Andrei Shleifer, and Stephen J Terry, "Real credit cycles," *NBER Working Paper*, 2021.

Carvalho, V. and X. 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 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.

Chava, S., W. Du, and N. 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, G., A. Ghent, and V. Haddad, "Asset Insulators," *Review of Financial Studies*, 2021, *34* (3).
Coibion, O. and Y. Gorodnichenko, "What Can Survey Forecasts Tell Us about Information Rigidities?," *Journal of Political Economy*, 2012, *120.*
Devlin, J., M-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Manuscript*, 2018.
di Giovanni, J., A. Levchenko, and A. Mejean, "Firms, destinations, and aggregate fluctuations," *Econometrica*, 2014, *82(4).*
Enders, Z., M. Kleemann, and G. J. Müller, "Growth Expectations, Undue Optimism, and Short-Run Fluctuations," *Review of Economics and Statistics*, 2021, *103.*
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, J., "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity," *Federal Reserve Bank of San Francisco Working Paper*, 2014, *19.*
Flynn, J. and K. Sastry, "The Macroeconomics of Narratives," *Working Paper*, 2022.
_ and _ , "Attention Cycles," *Working Paper*, 2024.
Gabaix, X., "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 2011, *79(3).*
Galaasen, S., R. Juelsrud, R. Jamilov, and H. Rey, "Granular Credit Risk," *NBER Working Paper 27994*, 2023.
Gaubert, Cecile and O. Itskhoki, "Granular Comparative Advantage," *Journal of Political Economy*, 2021, *129.*
Gentzkow, M., B. Kelly, and M. Taddy, "Text as Data," *Journal of Economic Literature*, 2019, *57.*
Gorodnichenko, Y., T. Pham, and O. Talavera, "The Voice of Monetary Policy," *American Economic Review*, 2023, *113.*
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, S. and M. McMahon, "Shocking language: Understanding the macroeconomic effects of central bank communication," *Journal of International Economics*, 2016, *99.*
_ , _ , and A. Prat, "Transparency and deliberation within the FOMC: a computational linguistics approach," *Quarterly Journal of Economics*, 2018, *133.*
Hassan, T., J. Schreger, M. Schwedeler, and A. Tahoun, "Sources and Transmission of Country Risk," *Review of Economic Studies*, 2023, *Forthcoming.*
_ , S. Hollander, L. v. Lent, and A. Tahoun, "Firm-Level Political Risk: Measurement and Effects," *Quarterly Journal of Economics*, 2019, *Forthcoming.*
_ , _ , _ , and _ , "The Global Impact of Brexit Uncertainty," *Journal of Finance*, 2023, *Forthcoming.*
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, R., H. Rey, and A. Tahoun, "The Anatomy of Cyber Risk," *NBER Working Paper 28906*, 2023.

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

Kohlhas, A. and A. Walther, "Asymmetric Attention," *American Economic Review*, 2021, *111.*

Lagerborg, A., E. Pappa, and M. Ravn, "Sentimental Business Cycles," *Review of Economic Studies*, 2022, *90.*

Liu, E., "Distortions in Production Networks*," *Quarterly Journal of Economics*, 2019, *134.*

Lorenzoni, G., "A Theory of Demand Shocks," *American Economic Review*, 2009, *99.*

Loughran, T. and B. 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).

Maxted, P., "A Macro-Finance Model with Sentiment," *Review of Economic Studies*, 2023, *90.*

Pappa, E., L. Lagerborb, and M. Ravn, "Does economic insecurity really impact on gun violence at US schools?," *Nature: Human Behavior*, 2019, *3.*

Pasten, E., R. Schoenle, and M. Weber, "Sectoral Heterogeneity in Nominal Price Rigidity and the Origin of Aggregate Fluctuations," *American Economic Journal: Macroeconomics*, 2023, *Forthcoming.*

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

Ravn, M. O. and H. Uhlig, "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *The Review of Economics and Statistics*, 2002, *84.*

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, Z., L. van Lent, G. Vilkov, and R. Zhang, "Firm-level climate change exposure," *Journal of Finance*, 2023, *78(3).*

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

Veldkamp, Laura L, "Information choice in macroeconomics and finance," *Princeton University Press*, 2023.

Online Appendix for "Granular Sentiments"

Rustam Jamilov Alexandre Kohlhas Sasha Talavera Mao Zhang

February, 2024

Contents

A Empirical Appendix

A.1 Data Construction

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)). 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 ($(\text{ppentq}_t - \text{ppentq}_{t-1})/\text{atq}_{t-1}$). 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 ((atq − ceq + prccq ∗ cshoq)/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 $\xi_{i,t}$ 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 $\xi_{i,t}$, $\xi^2_{i,t}$, $\xi^3_{i,t}$ 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 two proxies for aggregate sentiment are the Conference Board Leading Index (LEI), which provides early indications of changing business cycles, and the sentix Sentiment Index, which represents investors' market expectations over the next month. 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.

We use the number of victims from mass shootings as the instrument for sentiment. The variable has been HP-filtered. In order to limit the influence of outliers, such as the tragic 2017 Las Vegas shooting, we winsorize the instrument at the 5% level in every quarter. Data, obtained from Lagerborg et al. (2022), is available over 2006q4-2018q4.

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

Table A.1: Summary Statistics

	Observations	Mean	Std. Dev.	Min	Max
			Firm Data		
Sentiment, ξ_{it}	23387	1.858	1	-1.647	4.255
Sentiment, ξ_{it}^2	23387	1.457	1	-1.931	5.631
Sentiment, ξ_{it}^3	23357	-0.299	1	-7.510	5.295
Sentiment, Abs. Freq.	23387	1.710	1	-1.678	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 TFP	19464	0.000	1	-5.476	4.446
Sentiment, Firm Controls	17002	0.000	1	-5.732	5.224
Sentiment, Aggreg Controls	19464	0.000	1	-5.438	4.575
Sentiment, All Controls	17002	0.000	1	-5.647	4.818
Log Assets	23387	6.530	1	2.559	10.388
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		
Conf. Board LEI	61	0.012	1	-3.084	1.822
sentix Sentiment	61	-0.006	1	-3.728	1.371
Real GDP	61	-0.026	1	-6.039	1.214
Real GDP p.c.	61	-0.028	1	-5.963	1.262
Unemployment Rate	61	0.000	1	-1.363	5.555
CPI	61	0.000	1	-2.098	3.723
Industrial Production	61	-0.042	1	-3.800	1.593
Mass Shooting IV	49	-0.298	1	-2.062	1.989

Notes: This table provides basic summary statistics for main variables used in the empirical analysis.

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).

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

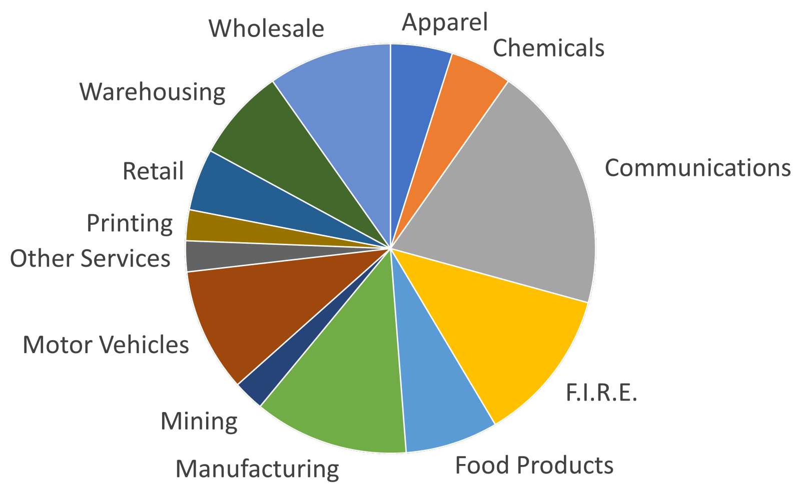
In this section we present additional cross-sectional results that complement the main text. Figure A.1 presents the sectoral distribution of firms that belong to the granular sentiment set Γ of the 50 most informative firms. The purpose of this Figure is to illustrate that most firms belong to sectors that are generally fairly downstream along the production chain: apparel, communications, food products, F.I.R.E., retail, etc.

Table A.2 lists the ten most informative firms based on the rank indicator K_i. The table showcases company names, the informativesness R_i^2, the level of upstreamness in 2021, and sector names. The purpose of this Table is to demonstrate that 80 percent of the 10 most informative firms are very downstream with U_i of less than 2, i.e. fewer than two transactional steps necessary to reach the final consumer.

Table A.3 reports results from the cross-sectional regressions of informativeness on upstreamness in 2021 along with the usual vector of controls. Columns (1) and (2) present results from the OLS, (3) and (4) from probit, and (5) and (6) from ordinal probit models, respectively. In columns (1) and (2) the dependent variable is the informativeness R_i^2. In columns (3) and (4) the dependent variable is an indicator that takes the value of unity for firms that belong to the granular informativeness set Γ. Finally, in columns (5) and (6) the dependent variable is the rank integer K_i.

Figure A.2 presents results from the cross-sectional analysis of sentiment and downstreamness under the Atalay (2017) definition of industries. Simiarly, Figure A.3 presents results from the cross-sectional analysis of sentiment and downstreamness under the Chahrour et al. (2021) definition of industries. As discussed also in the main text, the baseline finding of a negative relationship between upstreamness and informativeness does not change.

Figure A.1: Sectoral Distribution in Γ



Notes: This pie chart shows the distribution of sectoral representation in the baseline granular sentiment set Γ.

Table A.2: Select Downstream Firms

Rank K_i	R_i^2	Company	Upstreamness in 2021	Sector
1	0.714	IQVA Holdings	2.001	Services
2	0.696	Under Armour	1.045	Apparel
3	0.597	Honeywell Intl	1.670	Aerospace, technologies
4	0.547	CBRE Real Estate	1.821	Real Estate
5	0.525	Altari Engineering	1.227	Information Technology
6	0.522	Booking Holdings	1.821	Travel Technology
7	0.425	Canopy Growth Corp	1.670	Medical Cannabis
8	0.421	Citigroup	2.093	Financial Services
9	0.406	CDW Corp	1.930	Education Services
10	0.405	Tapestry Inc	1.045	Apparel

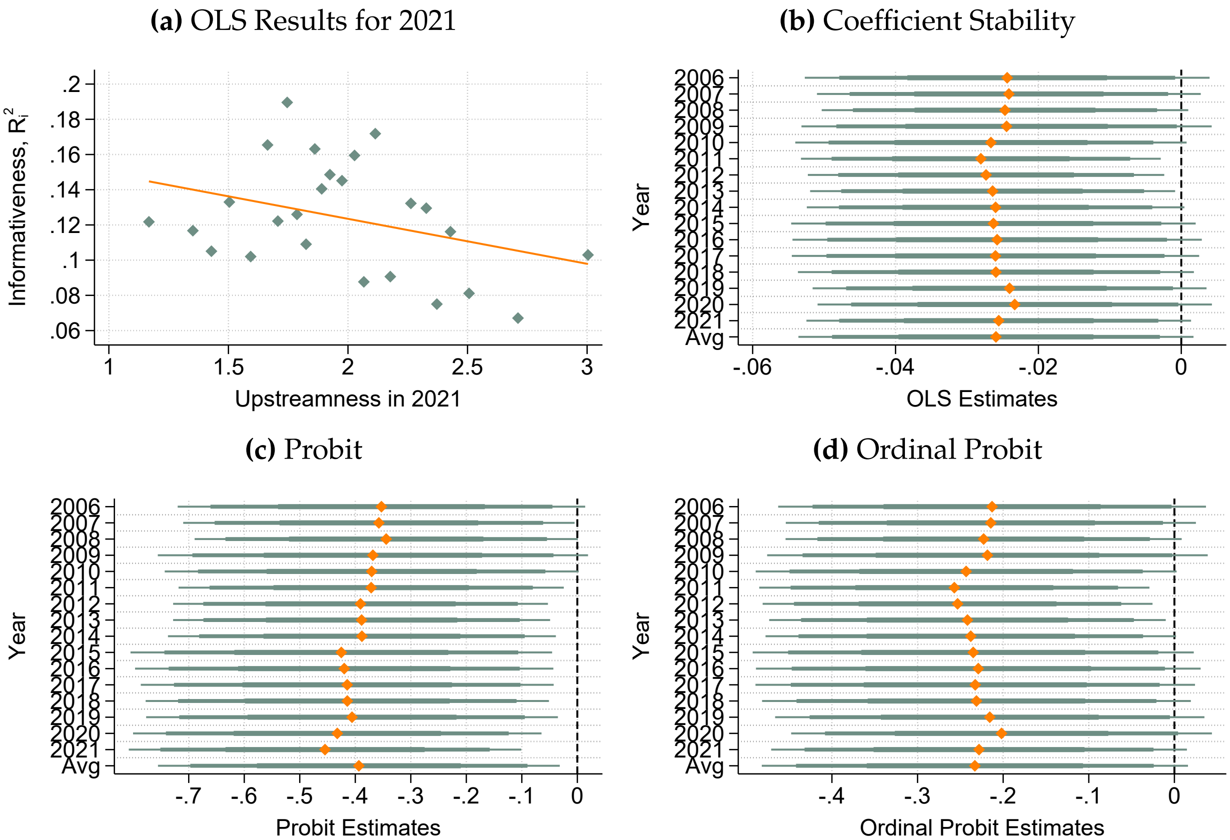
Notes: The ten most informative firms according to the rank indicator K_i. Columns 2-4 report firm-level R_i^2, company name, the level of upstreamness U_{st} in 2021, and sector name.

Table A.3: Downstreamness and Sentiment

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	Probit	Probit	Ordinal Probit	Ordinal Probit
Upstreamness in 2021	-0.020**	-0.022**	-0.342***	-0.343**	-0.157*	-0.196**
	(0.009)	(0.009)	(0.131)	(0.141)	(0.086)	(0.094)
Log (Assets)	0.006	0.013	-0.029	-0.391	0.085	0.217
	(0.006)	(0.016)	(0.085)	(0.332)	(0.057)	(0.138)
Market Beta		-0.009		-0.011		-0.125*
		(0.008)		(0.112)		(0.067)
Book-to-Market		-0.006		0.042		-0.062
		(0.009)		(0.152)		(0.080)
Investment		-0.013**		-0.036		-0.124***
		(0.005)		(0.094)		(0.046)
Valuation		-0.004		0.395		-0.127
		(0.013)		(0.251)		(0.126)
Leverage		-0.013*		-0.089		-0.136**
		(0.006)		(0.122)		(0.054)
Liquidity		-0.018**		-0.157		-0.153**
		(0.009)		(0.133)		(0.070)
Tobin's Q		0.007		-0.086		0.063
		(0.011)		(0.188)		(0.098)
Return Volatility		0.019		0.231**		0.141*
		(0.012)		(0.117)		(0.078)
Observations	531	469	531	469	531	469
R^2	0.011	0.047				

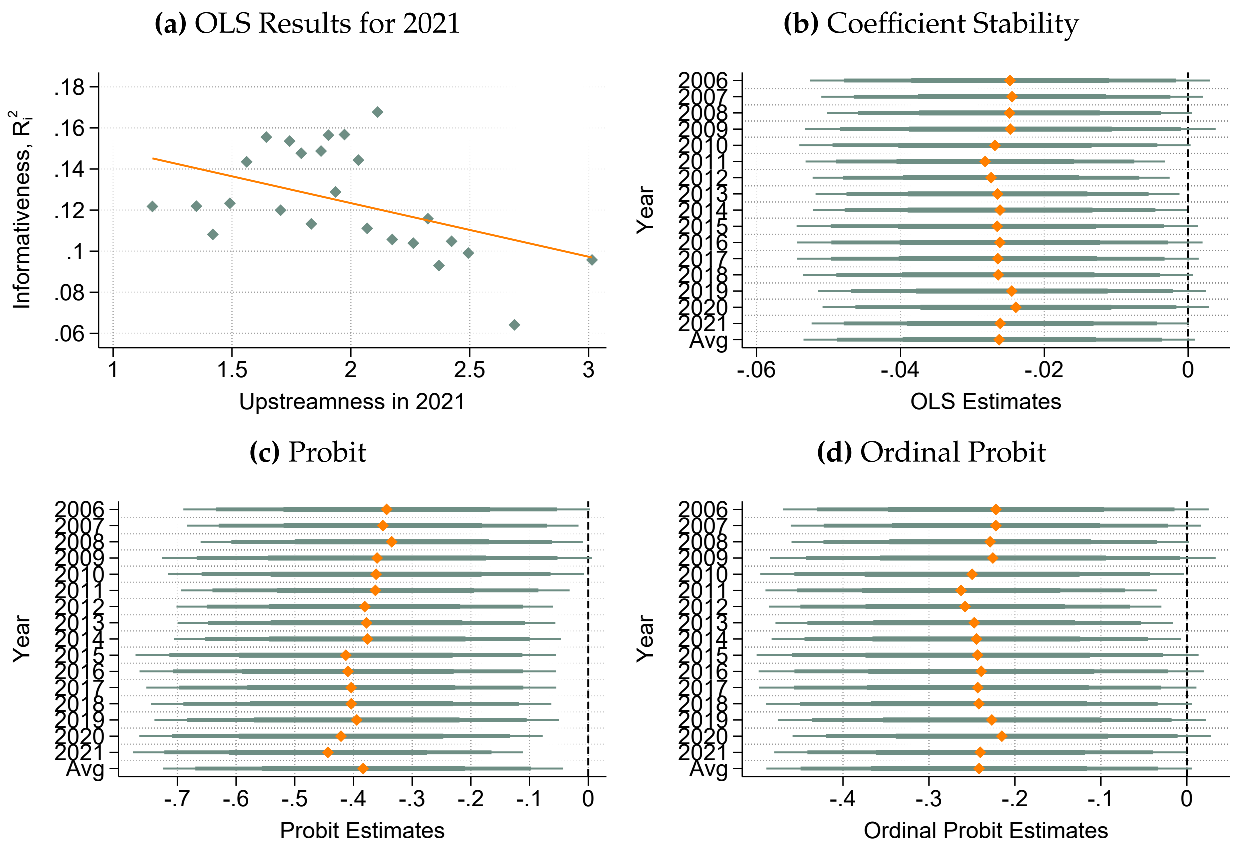
Note: Columns (1) and (2) report results from OLS cross-sectional regressions with R_i^2 as the dependent variable. Columns (3) and (4) report results from Probit cross-sectional regressions with a binary indicator that takes the value of unity for firms belonging to Γ as the dependent variable. Columns (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. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figure A.2: Downstreamness and Sentiment: Atalay (2017) Industry Definition



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

Figure A.3: Downstreamness and Sentiment: Chahrour et al. (2021) Industry Definition



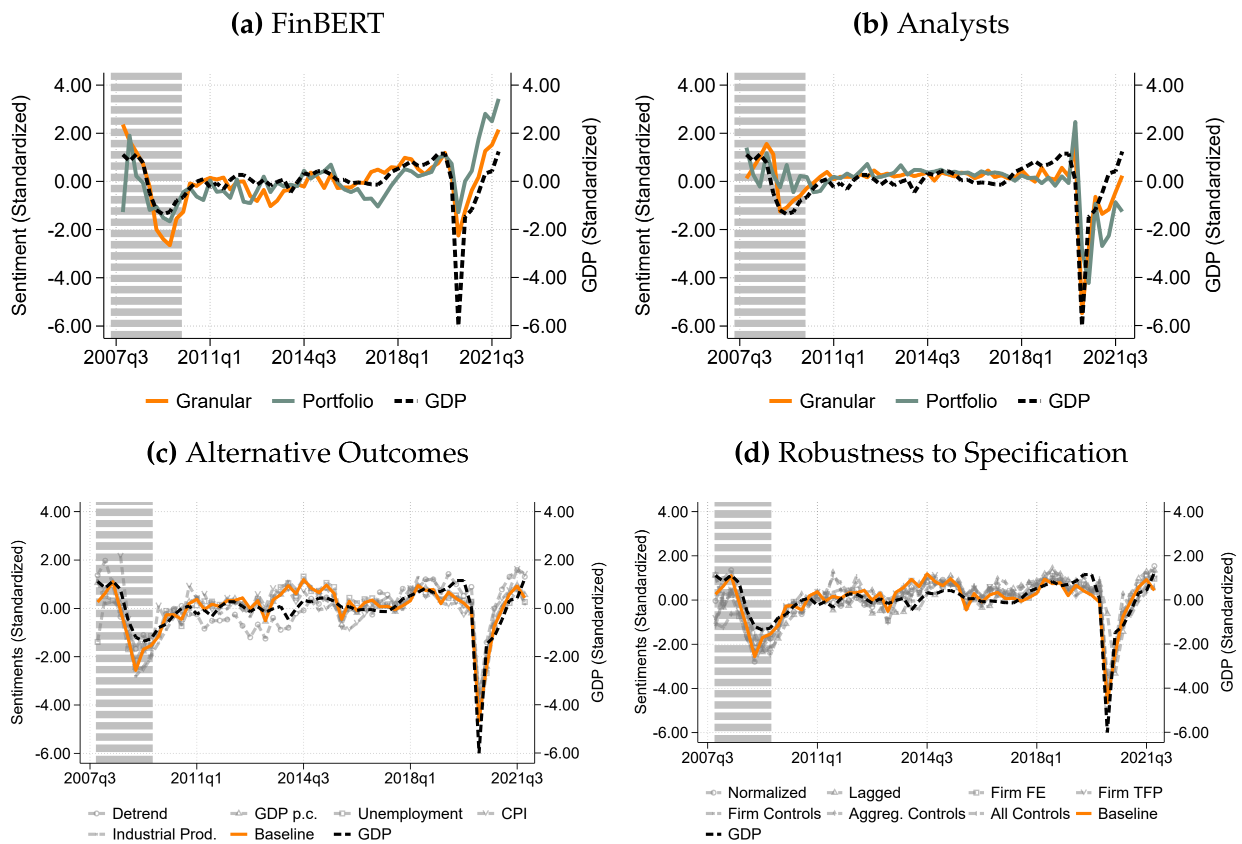
Notes: Results from cross-sectional regressions of upstreamness U_{us} on three metrics of informativeness: firm-level R_i^2 (Panels (a) and (b)), binary indicator G_i (Panel (c)), and rank indicator K_i (Panel (d)). U_{us} is constructed based on Chahrour et al. (2021) sectors. Horizontal bars in Panels (b), (c), and (d) 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.4 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^1) and analyst forecast (ξ_t^2) definitions of sentiment, respectively. Panels (c) and (d) plot the indices under alternative robust specifications. All measures, including the real GDP series, are standardized.

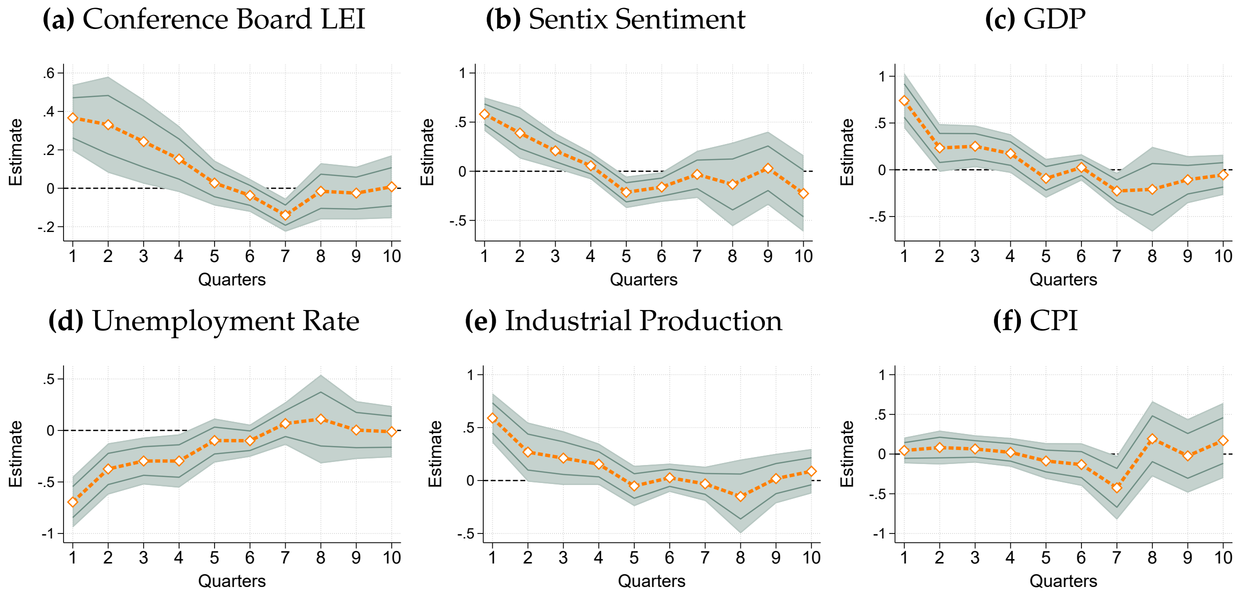
Figure A.5 presents the full set of local projection results under the alternative detrending scheme: we do not employ the HP-filter but instead remove the time fixed effect from every aggregate series. The purpose of this figure is to show that our result on a significant contemporaneous and dynamic relationship between granular sentiment $\mathcal{S}_t$ and the broader economy does not change.

Figure A.4: 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.5: Dynamic Effects of Granular Sentiment $\mathcal{S}_t$ on the Macroeconomy: Alternative De-trending

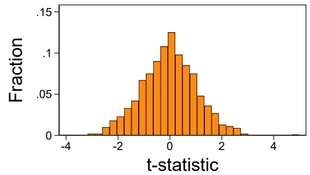


Notes: Local-projection estimates of granular sentiment $\mathcal{S}_t$ on macro aggregates. Macro aggregates have been residualized from the time fixed effect instead of being HP-filtered. Lines are 68% and shaded areas are 90% confidence intervals. Standard errors are clustered at the quarterly level.

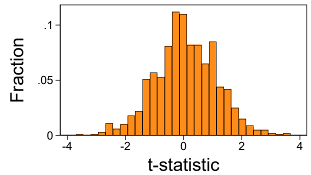
A.5 Further Robustness Tests

This section presents additional robustness tests that complement the main text. Figure A.6 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.

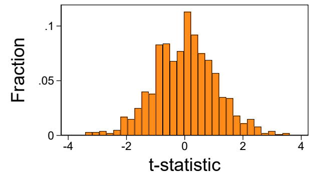
Figure A.6: Placebo Cross-sectional Regressions of Informativeness on Upstreamness



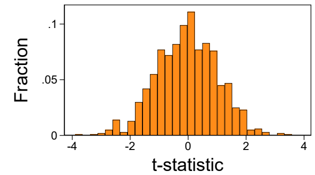
(a) 2006



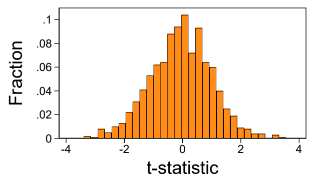
(b) 2007



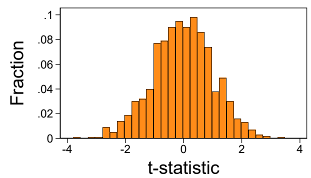
(c) 2008



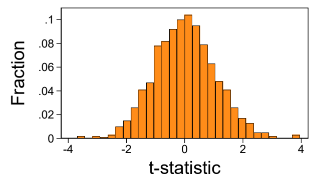
(d) 2009



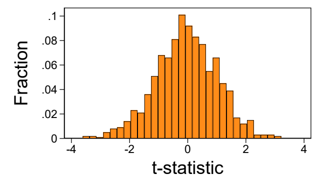
(e) 2010



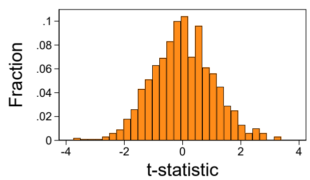
(f) 2011



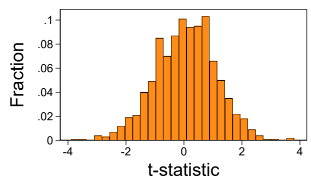
(g) 2012



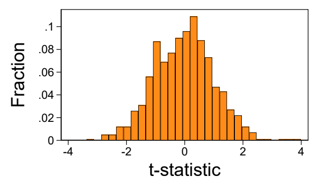
(h) 2013



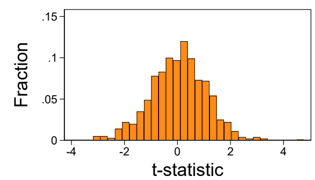
(i) 2014



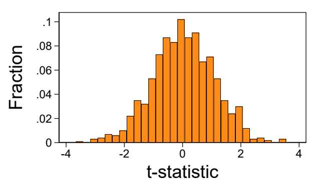
(j) 2015



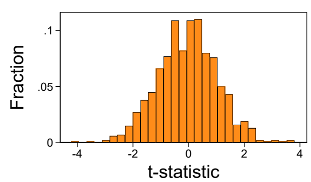
(k) 2016



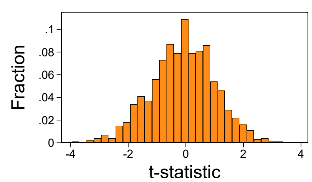
(l) 2017



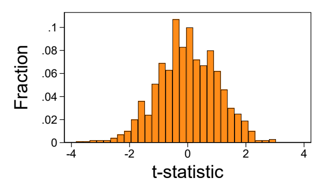
(m) 2018



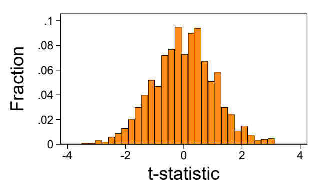
(n) 2019



(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).[1]

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

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

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

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

in the main text. Thus, stacking results shows that

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

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

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

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

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

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

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

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. $1 - \alpha_{uu} > \alpha_{du}$. □

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:

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

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

Atalay, E., "How Important Are Sectoral Shocks?," *American Economic Journal: Macroeconomics*, 2017, *9.*

Baker, S., N. Bloom, S. Davis, and A. Notes, "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics*, 2016, *131(4).*

Chahrour, Ryan, Kristoffer Nimark, and Stefan Pitschner, "Sectoral media focus and aggregate fluctuations," *American Economic Review*, 2021, *111* (12), 3872–3922.

Devlin, J., M-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Manuscript*, 2018.

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

Lagerborg, A., E. Pappa, and M. Ravn, "Sentimental Business Cycles," *Review of Economic Studies*, 2022, *90.*

Ravn, M. O. and H. Uhlig, "On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations," *The Review of Economics and Statistics*, 2002, *84.*

The Centre for Responsible Banking & Finance
CRBF Working Paper Series
Department of Finance, University of St Andrews
The Gateway, North Haugh,
St Andrews, Fife,
KY16 9RJ.
Scotland, United Kingdom
https://crbf.wp.st-andrews.ac.uk/

Recent CRBF Working papers published in this Series

Second Quarter | 2025

25-006 **Kirsten Schmidt and Lena Tonzer:** Banks' Foreign Homes.

First Quarter | 2025

25-005 **Edoardo Lanciano, Daniele Previati, and Ornella Ricci:** Crypto Ownership among Young People: The Effect of Financial Literacy and Risk Propensity.

25-004 **Pejman Abedifar, Soroush Kamyab, Steven Ongena, and Amine Tarazi:** Relationship Banking: The Borrower's Incentives Channel.

25-003 **Alfred Duncan, João Pedro De Camargo Mainente, and Charles Nolan:** Monetary and Financial Policy with Privately Optimal Risk Taking.

25-002 **Mehmet Selman Çolak, Yavuz Kılıç, Yusuf Yıldırım, and Muhammed Hasan Yılmaz:** The Effect of Personal Credit Card Limits on Card Use: Micro Data Evidence from an Emerging Market.

25-001 **Hans Degryse, Olivier De Jonghe, Leonardo Gambacorta, and Cédric Huylebroek:** Bank Specialization and Corporate Innovation.

Fourth Quarter | 2024

24-030 **Dimitris K. Chronopoulos, John O.S. Wilson, and Muhammed H. Yilmaz:** State Ownership and Bank Lending: Evidence from Turkish Banking.

24-029 **Anna L. Sobiech, Lechedzani Kgari, Linh Nguyen, and John O.S. Wilson:** Implicit Guarantees and Cross-Border Bank Stability: Evidence from Geographic Ringfencing.

University of St Andrews
Scotland's first university

600 YEARS
1413 – 2013