Micro Shocks and Macro Fluctuations in the Information Network*

Zhen Huo † Minghao Li ‡ Jieran Wu § Huan Xiong ¶

July, 2025

Abstract

In this paper, we study how heterogeneity in firms' attention allocation contributes to the amplification of micro-level shocks into aggregate economic fluctuations. Empirically, we document that the attention firms receive follows a fat-tailed distribution based on their browsing activities related to each other's electronic filings on EDGAR. This finding suggests that a relatively small subset of firms attracts a disproportionately large share of attention, consistent with the fact that the browsing-weighted measure of sales growth exhibits substantial predictive power for macroeconomic forecasts based on survey data. Theoretically, we extend the noisy business cycle framework (Angeletos and La'O, 2010) to incorporate asymmetric attention at the firm level through a directed graph. We derive conditions under which granular effects emerge even when firms are homogeneous in size and analyze the interplay between attention allocation and firm-size heterogeneity in magnifying micro-level shocks. Quantitatively, our analysis reveals that attention heterogeneity alone can drive significant aggregate fluctuations resulting from micro shocks. Nevertheless, the degree to which attention asymmetry complements firm-size heterogeneity depends critically on the correlation between these two factors.

Keywords: Information Network, Business Cycle Fluctuations, Granularity, Graph Theory

*We are grateful for comments from Manuel Amador, Marios Angeletos, Marco Bassetto, V.V Chari, Joel Flynn, Xavier Gabaix, Simon Mongey, and Christopher Tonetti. The paper benefits from the discussion with seminar participants at Federal Reserve Bank of Minneapolis, Inter-American Development Bank, BSE Summer Forum: Macro Fluctuations with Micro Frictions, and Rice University.

†Yale University, Department of Economics, zhen.huo@yale.edu

‡Peking University, National School of Development, mhli@nsd.pku.edu.cn

§Zhejiang University, School of Economics, jw5ya@zju.edu.cn

¶Harbin Institute of Technology, Institute for Advanced Study in Mathematics, huan.xiong.math@gmail.com

1. Introduction

One of the most important questions in modern macroeconomics concerns the source and origins of aggregate fluctuations at the microeconomic level. According to conventional wisdom, firm-specific shocks' impacts on the aggregate economy wash out on average. This wisdom has been challenged by allowing for micro-level heterogeneity in firm size, as in Gabaix (2011), or by allowing for input-output linkages among sectors, as in Acemoglu et al. (2012). In a similar vein, firms also differ significantly from each other in terms of the amount of attention they receive. Some salient firms are frequently mentioned in the media and draw a lot of public attention, while most others are silent. A natural question is what makes a firm important? Is it because they are large in size, or because they are 'star' firms in the spotlight, closely watched by others? Can the heterogeneity along the dimension of firms' salience also give rise to granular effects? How will these different heterogeneities interact with each other?

This paper makes three contributions in addressing these questions: Empirically, we document that the attention firms receive follows a fat-tailed distribution based on their browsing activities related to each other's electronic filings on EDGAR (Electronic Data Gathering, Analysis, and Retrieval). This suggests that a small number of firms receive disproportionately large amounts of attention. Consistently, the browsing-weighted measure of sales growth demonstrates significantly stronger predictive power for macroeconomic forecasts in the survey data compared to the sales-weighted measure. Theoretically, we extend the noisy business cycle framework (Angeletos and La'O, 2010) to incorporate asymmetric attention at the firm level through a directed graph. We provide the condition on the in-degree and out-degree of this information network that allows for granular effects, even with homogeneous firm sizes. We also provide a characterization of the interactive effects between the attention channel and the traditional firm-size channel under the power law distributions. Quantitatively, we show that the attention channel alone can generate sizable aggregate fluctuations driven by micro shocks. However, whether the asymmetric attention channel complements the heterogeneous firm size channel can be ambiguous and hinges on the strength of the general equilibrium considerations.

Empirics. Our motivating facts utilize detailed firm-to-firm browsing information on documents filed at EDGAR. The bilateral browsing activities provide a direct measurement of firms' attention allocation. We use the total browsing requests received by a firm to measure its received attention, and the total browsing requests sent by a firm to measure its attention paid towards other firms.

We document four facts. First, the distributions of browsing received and sent both exhibit power laws. The browsings received follow almost an exact Zipf's law. The estimated tail parameter for the browsings sent is -0.92. That is, a small number of firms receive a disproportionately large amount of attention. Second, the browsing activities are weakly positively correlated with firm sizes. One may have the prior that the firms that receive most browsing requests are those biggest firms. However, the correlation between the browsings received and the sales share is only 0.29. Fewer than 20 firms that belong to the top 100 largest firms also belong to the top 100 firms that receive the most browsing. Third, the browsing received and sent by a firm are positively correlated only for large firms. Outside the top 400 largest firms, the browsing received and sent become uncorrelated. Fourth, browsing-weighted sales growth strongly predicts forecasts of macroeconomic variables in the Survey of Professional Forecasters, whereas the sales-weighted measure exhibits negligible predictive power.

Theory. Motivated by these facts, we propose a framework to study an incomplete-information economy with a finite (potentially very large) number of firms. Firms' output choices depend on their own productivity shock, which consists of a common component and a firm-specific component, as well as on the aggregate output via standard demand linkages. We build on Angeletos and La'O (2010), with an important difference: firms' information structure is heterogeneous, and an information network describes the bilateral information flow between firms.

This information network is captured by a directed graph. Operationally, there is an information matrix where its (i, j)-element can be 0 or 1. Firm j's productivity shock is perfectly observed by firm i if the element in the information network equals 1, and is not observed by firm i otherwise. A firm's information set, therefore, consists of their firm-specific collection of observed productivities. The in-degree of this information network summarizes the amount of attention a firm receives, while the out-degree summarizes the total amount of attention a firm directs outward.

An individual firm's micro shock can affect the aggregate GDP in three ways: (1) the efficiency channel, that micro shocks directly affect firms' production; (2) the information channel, that when observed by other firms, micro shocks are informative about the aggregate productivity; (3) the coordination channel, that a salient firm's shock is useful in forecasting other firms' decisions. We provide an explicit solution to the GDP elasticities with respect to micro shocks in equilibrium that account for these forces.

Under what conditions can a firm's firm-specific component induce aggregate fluctuations when the number of firms goes to infinity? We answer this question in two steps. First, we define a key statistics in gauging firms' importance in the information network, a diluted in-degree index. For an firm i, the index is given by the sum of reciprocal of other firms' out-degrees which observes firm i. This index summarizes the effective attention received. Second, we provide an estimate of the bounds of the GDP elasticity with respect to a firm's micro shock, both of which hinge on the diluted in-degree index. The granular effects can emerge if only if such index is growing with the number of firms at the same rate. Intuitively, the granular effects take place if a firm i is salient in the economy in the sense that it receives attention from most of the other firms in the economy. However, this is not sufficient. It has to be that such attention is concentrated on firm i and few other firms can share this attention.

Firm size and attention under power law. To gain further insight on the interaction between the conventional firm-size channel and the attention channel, we extend the framework that allows for both types of heterogeneities. We follow Gabaix (2011) and assume that both the firm size distribution and the attention recevied measured by in-degrees follow power law distributions.[1] In this setting, the nature of the aggregate volatility driven by micro shocks hinges on the exponents and the correlation between the two power law distributions.

We demonstrate qualitatively that, regardless of the correlation between firm size and the attention each firm receives, the rate at which expected GDP volatility due to micro-level shocks decays is governed by the minimum of the two tail exponents. Equivalently, the distribution exhibiting the heavier tail exerts a dominant influence on the economy's aggregate volatility. Quantitatively, however, the correlation between these two distributions becomes crucial: GDP volatility reaches its maximum under perfect positive assortative matching between firm size and received attention, and diminishes as this matching weakens.

Quantification and Implications. To quantify the information channel in amplifying the effects of micro shocks, we adopt a random graph approach when calibrating the information matrix. The leverage on the patterns of the browsing activities in EDGAR to assign probabilities for an element in the information matrix to be 1. When a firm i engages in more browsing activities and when a firm j receives more browsing requests, the aforementioned probability becomes higher. We further extend our theoretical framework to allow for heterogeneity in firm sizes and the volatility of the firm-specific components.

Quantitatively, the heterogeneous information channel alone can generate sizable fluctuations due to micro shocks. The top 100 firms that receive most attention can account for more than 20% of aggregate volatility. In a model with perfet information or with cannonical dispersed information, such number is much more dampened. When allowing for firm size heterogeneity, the perfect information model can also generate sizable granular effects. Importantly, most of the fluctuations due to micro shocks are from those firms that are big in sizes. In contrast, adding information heterogeneity, roughly half of the fluctuations due to micro shocks are from those firms that are salient but not large. Moreover, the volatility of the top 100 largest firms are dampened. The information channel still plays an important role when the traditional firm-size channel is present, and they do not necessarily reinforce each other.

The key mechanism is that the salient firms act as semi-public signals. They are important in coordinating the decision with other firms, and many firms respond to their productivity changes when forming higher-order expectations. However, if a firm is big but is not observed by other firms, the GDP elasticity with respect to this firm will be lower, which is the standard dampening effect of informational frictions. Whether these information and coordination channels reinforce the standard firm-size channel or not depends on the correlation between the firms' attention received and their sizes. Due to the relatively low correlation between these two dimensions, the information network actually dampens the firm-size channel in generating granular effects.

[1] The out-degrees are uniformally distributed.

Related literature. The granular effect in macroeconomics, introduced by Gabaix (2011), posits that idiosyncratic shocks to large firms may have significant macroeconomic implications, challenging the conventional assumption that firm-level shocks average out in the aggregate. There are two strands of literature that provide micro-foundation for this granular effect. The first strand focuses on the heterogeneity in the distribution of firm size and market power. Gabaix's seminar work demonstrates that in economies with fat-tailed firm-size distributions, firm-specific shocks propagate to the macroeconomy and induce aggregate fluctuations. Carvalho and Gabaix (2013) extend the "granular hypothesis" by showing that the evolving patterns of sectoral concentrations amplify the impact of granular shocks and account for the dynamics of aggregate volatility. More recently, Eeckhout et al. (2020) highlights how the growing dominance of large firms affects productivity and business cycle dynamics. The second strand of literature explores the role of economy's input-output linkages, or production network, in translating idiosyncratic shocks into sizable macroeconomic fluctuations. Acemoglu et al. (2012), Acemoglu et al. (2016), and Baqaee and Farhi (2020) adopt this approach by emphasizing the role of intersectoral dependencies in propagating firm-level shocks through supply chains, thereby amplifying their macroeconomic effects or creating cascading failures (Baqaee, 2018). Relative to the literature, we propose a novel channel of granularity effect that relies on the information network among firms in the economy. Supported by the empirical evidence, we show that firms' idiosyncratic shocks are able to generate sizable aggregate fluctuations that are independent of size-heterogeneity and production-network structure.

Our paper is also related to an expanding literature that incorporates information frictions in the models of production networks (Chahrour et al., 2021; Auclert et al., 2020; La'O and Tahbaz-Salehi, 2022; Angeletos and Huo, 2021; Bui et al., 2024; Nikolakoudis, 2024). The subjects of these research span from aggregate fluctuations driven by sectoral public information, as in Chahrour et al. (2021), to the interactions between higher-order beliefs and production networks, as in Bui et al. (2024). Our earlier work Fang et al. (2024) employs the same EDGAR browsing data set to characterize firms' endogenous information acquisition behaviors on production network. To account for the empirical findings, Fang et al. (2024) provides a production network model featured with rational inattention and endogenous nominal rigidity, and studies the optimal monetary policy in this environment. Jamilov et al. (2024) uses a real production network model to study the impact of endogenous information acquisition (limited attention) on sentiment-driven aggregate fluctuations, arguing that attentions are centered on downstream firms due to their role as "information agglomerators". Relative to this literature, in this paper we propose a novel theory of information network that builds on bilateral attentions between firms. The existence of this network does not rely on inter-sectoral linkages.

2. Motivating Facts

This section presents four novel empirical facts about firms' attention sent and received, using internet traffic data that track firms' access to other firms' SEC filings on EDGAR. Specifically, we show that firms' browsing activity follows a power-law distribution and is only weakly positively correlated with firm size. Moreover, browsing sent and received is positively correlated only for large firms. Finally, browsing-weighted sales growth strongly predicts forecasts of macroeconomic variables in the Survey of Professional Forecasters, whereas the sales-weighted measure exhibits negligible predictive power.

2.1 Browsing-Based Attention Measure with EDGAR Data

Our analysis leverages detailed firm-to-firm browsing information from the EDGAR system—the Electonic Data Gathering, Analysis, and Retrieval platform maintained by the U.S. Securities and Exchange Commission (SEC).[2] Since 1996, all U.S. public companies have been required to file disclosures electronically through EDGAR. These filings, which are freely accessible online, serve as a primary source of firm-specific information for managers, investors, and analysts.

In 2003, the SEC's Division of Economic and Risk Analysis (DERA) began publishing internet traffic data for SEC filings, known as the "EDGAR Logfile." A distinctive feature of this dataset is its granularity: it records each instance of a document being viewed online, including partially anonymized IP addresses of the visitors, precise access timestamps, and unique document identifiers (Accession Numbers). This high-resolution data enables us to track firm-to-firm browsing behavior, which we use as a proxy for firms' attention allocation.

Attention measure. We define a *browser* as the viewer of a document and a *browsee* as the owner of that document. To identify firm-level browsing relationships, we follow the methodology in Chen et al. (2020) and use data from ip-info.io—a leading provider of IP information—to de-anonymize partially masked IP addresses and map them to firm identities. This process produces a firm-level, daily browser–browsee dataset.[3]

Using this dataset, we construct a firm-to-firm browsing (information) network for any period t, denoted by $\mathcal{B}_t$. Each element b_{ijt} of the network represents the number of times firm i (the browser) accessed SEC filings belonging to firm j (the browsee) during period t:

$$b_{ijt} = \sum_{s \in t} a_{ijs}, \tag{2.1}$$

where $a_{ijs} = 1$ if firm i accessed a document filed by firm j at time s. We define the total browsing sent by firm i during period t as $B^S_{it} = \sum_j b_{ijt}$, and the total browsing received by firm j as $B^R_{jt} = \sum_i b_{ijt}$.

We argue that our browsing-based measure offers a meaningful proxy for firm-level attention. As the SEC's official platform for corporate disclosures, EDGAR provides the most timely and comprehensive access to firm-specific information, including unscheduled material events (Form 8-K) and annual financial statements (Form 10-K).[4] Unlike paid services such as Bloomberg or Yahoo Finance,

[2] More information on EDGAR is available at https://www.sec.gov/edgar/about.
[3] For additional details on the data construction, see Fang et al. (2024).
[4] The most frequently browsed document types are listed in Table C.1 in the Appendix.

EDGAR is freely accessible and preserves critical disclosures that may be delayed or omitted on third-party platforms. These features make EDGAR a primary source for investors, analysts, and managers seeking authoritative and up-to-date firm information.

Our measure complements and extends text-based attention proxies widely used in recent literature. First, while many existing measures rely on the content of public filings—which may be shaped by firms' strategic disclosure incentives—ours captures actual daily information acquisition behavior, free from such distortions. Second, unlike media-based proxies that typically reflect aggregate attention, our approach enables a direct and granular mapping of firm-to-firm attention linkages.

Despite these conceptual differences, our measure is positively correlated with existing text-based proxies, offering external validation for its use as a credible attention metric. For instance, when comparing total attention received by a firm, our measure correlates with a media-based text proxy at 0.39 (standard error: 0.035).

Table 1: Baseline Sample: Summary Statistics

	Mean	Median	S.D.
No. Browsers per Year	3,626	3,608	251
No. Browsees per Year	6,254	6,280	321
No. Browsees viewd by a Browser per Year	180	176	25
No. Browsers viewing a Browsee per Year	105	74	115
Browsing Volume sent by a Browser per Year	11,356	9,881	7,764
Browsing Volume received by a Browsee per Year	5,897	5,389	4,869

Baseline Sample. Our baseline sample focuses on public firms, as they are the primary filers on EDGAR and can be linked to data sources like Compustat. We restrict the sample period to 2009-2016.[5] The resulting sample includes 713,157,510 unique IP addresses associated with 7,622 public companies, accounting for 52.4% of all EDGAR browsing activity during the sample period.[6]

Table 1 provides a summary of our baseline sample. Each year, the data include an average of 3,626 unique browsers and 6,254 unique browsees. On average, a browser views filings from 180 different firms, while each browsee is accessed by 105 distinct browsers. In total, a typical browser generates 11,356 views per year, while a typical browsee receives 5,897 views.

2.2 Facts on Firms' Browsing Activities

We now present a set of novel empirical facts that motivate our theoretical analysis. These facts focus on the distribution of browsing activity—both received and sent—and its relationship with firm size. For clarity, we illustrate these patterns using data from the year 2016.

[5] We start in 2009 due to limited pre-2009 data and end in 2017, when the SEC discontinued the release of traffic data.

[6] We exclude automated web crawlers following the method in Cao et al. (2023). All results are robust to their inclusion.

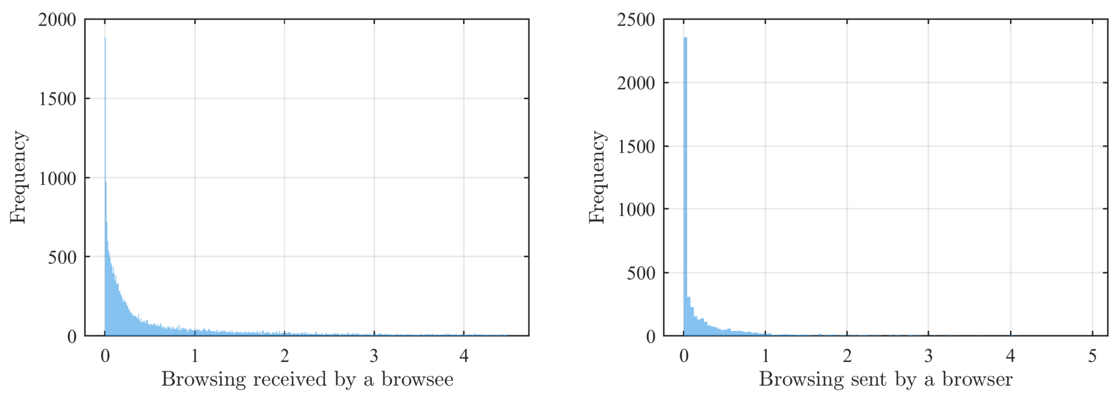
Fact 1. Distributions of browsing received and sent both exhibit power-law behavior.

We begin by examining the distributions of browsing received and sent. Figure 1 displays both distributions, which are highly right-skewed: while most firms engage in a relatively small number of browsing activities, a small subset of firms accounts for a disproportionately large share. The pronounced right tails indicate that both distributions are fat-tailed and potentially follow a power-law pattern.

Figure 1: Distribution of Firms' Browsings

(a) Browsing received by a browser

(b) Browsing sent by a browser



Notes: Browsings are computed using 2016 data. The sample includes 24,286 browsees and 4,542 browsers. The mean (median) browsing received per browsee is 72,895 (25,039), and the mean (median) browsing sent per browser is 2,888 (344).

To further examine the nature of the browsing distributions, we test whether they follow a power-law using standard methods from the literature. Following the approach in Gabaix and Ibragimov (2011), we rank firms by the total number of browsings received or sent and estimate the power-law exponent using the top 500 firms. Specifically, we apply the Rank-1/2 regression:

$$\ln(\text{Rank}_i - 1/2) = c + \kappa \ln(B_i),$$

where Rank_i is the rank of firm i, and B_i denotes its total browsing activity (received or sent). The coefficient α captures slope of the regression in log-log space and serves as the power-law exponent.[7]

For browsing received, the estimated exponent is -0.999 with a standard error of 0.004 and an R^2 of 0.993. This implies that the distribution closely follows Zipf's Law, which states that the probability of a firm receiving more than K browsings is inversely proportional to K:

$$\text{Prob}(B_i > K) \propto 1/K^{\kappa}, \text{ with } \kappa \approx 1,$$

[7]As shown in Gabaix and Ibragimov (2011), the Rank-1/2 estimator mitigates bias in the estimation of power-law exponents.

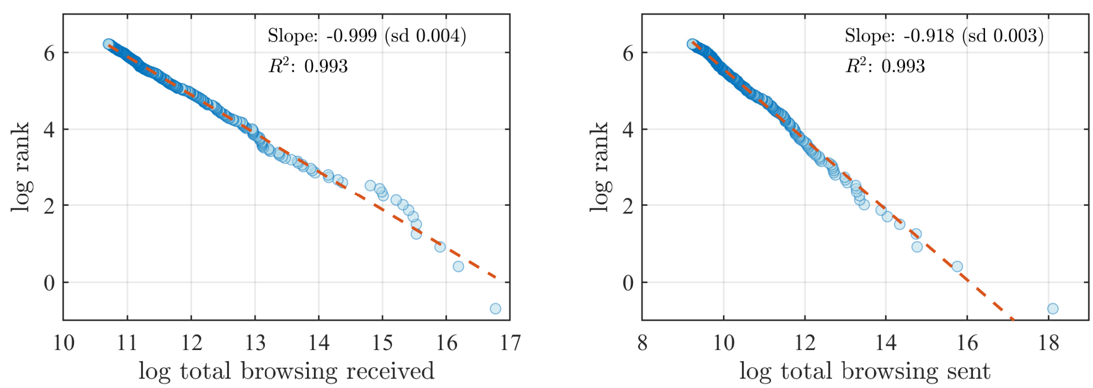
Figure 2a visually confirms this relationship.

For browsing sent, the estimated exponent is −0.918 with a standard error of 0.003 and an R^2 of 0.993, indicating a heavier (fatter) upper tail than in the case of browsing received. This suggests that a small number of firms engage in a disproportionately large amount of browsing activity. Figure 2b illustrates this power-law behavior.

These findings are robust to the choice of cutoff for the ranked firms. In Appendix Figure C.2, we show that our results continue to hold using the top 100 or top 1,000 firms. Our results are also robust to excluding financial firms, as shown in Appendix Figure C.1.

Figure 2: The Browsing Power-Law

(a) Power-Law (browsing received) (b) Power-Law (browsing sent)



Notes: This figure illustrates the log-log plot: the relationship between the rank of a firm's total browsing received (sent) and its total browsing received (sent). It visualizes the Rank-1/2 regression: $\ln(\text{Rank}_i - 1/2) = c + \alpha \ln(B_i)$.

Discussion. Several observations are worth noting. First, as this paper focuses on the macroeconomic implications of the observed power-law distribution in firm attention, we do not attempt to model the underlying determinants of attention heterogeneity across firms. Nonetheless, we do explore the sources of attention received. As shown in Appendix B.2, on average, firms tend to direct 25% of their attention to their suppliers and customers and 10% of their attention to their competitors.

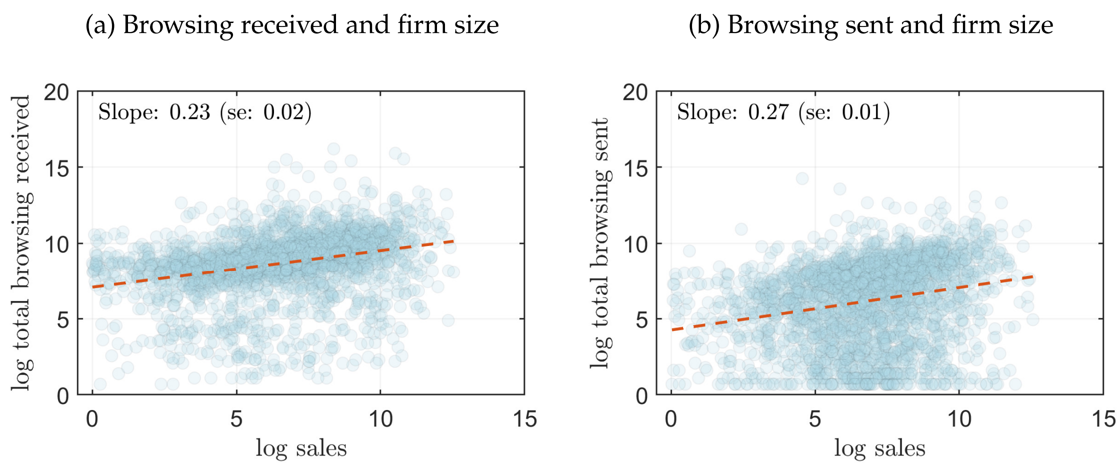
Lastly, we find moderate temporal persistence in browsing behavior, with an autocorrelation of 0.53 for attention sent and 0.40 for attention received.

Fact 2. Browsing activities are weakly positively correlated with firm size.

We next explore how firms' browsing activity relates to firm size, using sales as our size measure. We find that both attention received and attention sent are only weakly positively correlated with firm size. Specifically, the correlation between log browsing received and log sales is 0.29 (standard error: 0.02), while the correlation between log browsing sent and log sales is 0.22 (standard error: 0.02). Figure 3 displays scatter plots of these relationships, along with fitted linear trends.

This weak correlation suggests that the power-law patterns documented in Fact 1 are not simply a reflection of firm size. Indeed, Appendix Figure C.3 shows that the estimated power-law exponents remain largely unchanged even after excluding the top 100, 500, or 1,000 firms by sales.

Figure 3: Browsing Received, Browsing Sent, and Firm Size



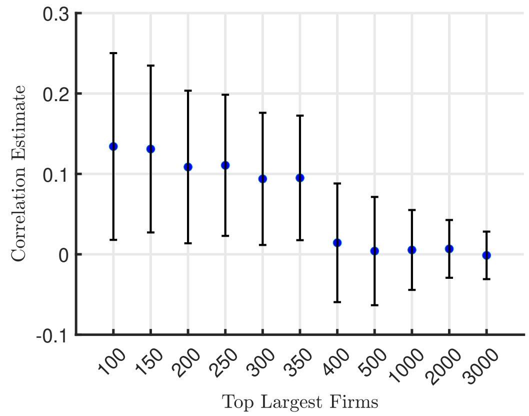
Notes: This figure visualizes the relationship between firms' log total browsing received (sent) and their log sales, along with the linear fit. Both display relatively weak correlation.

Fact 3. Browsing received and sent are positively correlated only for large firms

Do firms that browse others more intensively also receive more attention? The short answer is yes—but only among large firms. Let B_i^R and B_i^S denote the total browsing received and sent by firm i, respectively. We compute the correlation $\text{corr}(B_i^R, B_i^S)$ for the top N largest firms, varying N from 100 to 3,000.

Figure 4 plots the estimated correlations with their 90% confidence intervals. A clear pattern emerges: a significant positive correlation exists only among the largest firms, specifically the top 350, with a correlation of around 0.1. For smaller firms, the relationship between attention sent and received is effectively zero.

Figure 4: Browsing Received-Sent Correlation and Firm Sizes



Notes: This figure illustrates the computed correlation corr(B_i^R, B_i^S) for the top N largest firms, varying N from 100 to 3,000, with 90% confidence intervals

Fact 4. Browsing-weighted sales growth strongly predicts forecasts in the SPF

If a small fraction of firms receives disproportionally large attention, their sales performance may disproportionately influence forecasters' expectations about macroeconomic variables such as industrial production, GDP, or unemployment. We test this hypothesis and confirm the intuition.

Specifically, in each quarter t, we rank firms based on their total browsing received and focus on the top $K = 100$ firms, For these firms, we compute a browsing-weighted measure of sales growth:

$$G_t^{\text{browsing}} = \sum_{i=1}^{K} \omega_{i,t}^{\text{browsing}}(g_{i,t} - \overline{g}_t), \tag{2.2}$$

where $\omega_{i,t}^{\text{browsing}} = B_{i,t}^R / \sum_{j=1}^{100} B_{j,t}^R$ denotes the attention share of firm i among these top K firms, and $B_{i,t}^R$ is the total browsing received by firm i in quarter t. To adjust for seasonality, we define firm i's sales growth as the year-over-year change in log sales: $g_{i,t} = s_{i,t} - s_{i,t-4}$, where $s_{i,t}$ is the log sales of firm i in quarter t. To isolate the firm-specific component, we subtract the average sales growth across all firms, $\overline{g}_t = \sum_i^N g_{i,t}/N$, from $g_{i,t}$, where N denotes the total number of firms. To address the concern for outliers, we winsorize the sales growth at the percentile of 1% and 99%.[8]

For comparison, we also compute a conventional sales-weighted measure of sales growth as follows,

$$G_t^{\text{sales}} = \sum_{i=1}^{K} \omega_{i,t}^{\text{sales}}(g_{i,t} - \overline{g}_t), \tag{2.3}$$

where $\omega_{i,t}^{\text{sales}}$ is the sales share of firm i among the top K firms.

[8]Our results are robust if we use the whole sample, or we use 5% and 95% for winsorization.

We then estimate the following regression specification:

$$f_t^{t+h} = \alpha + \sum_{j=1}^{J} \beta_{1,j} G_{t-j}^{\text{browsing}} + \sum_{j=1}^{J} \beta_{2,j} G_{t-j}^{\text{sales}} + \gamma X_t + \varepsilon_t, \tag{2.4}$$

where f_t^{t+h} is the mean forecast, made in quarter t, for the variable f over a horizon of h quarters.[9] We obtain these forecasts from the Survey of Professional Forecasters (SPF). Following Gabaix (2011), the control vector X_t includes lagged macroeconomic variables and shocks with J lags—specifically, oil supply shocks, monetary policy shocks, the 3-month nominal T-bill rate, and the term spread.[10] Our baseline uses $h = 1$ horizon and $J = 4$ lags, though the results are robust to alternative horizon and lag lengths.

Table 2: Predictive Power of Browsing-weighted Sales Growth

	Mean Forecast of Industrial Production in $t+1$					
	(1)	(2)	(3)	(4)	(5)	(6)
$G_{t-1}^{\text{browsing}}$	0.028***				0.022*	0.024*
	(0.0100)				(0.011)	(0.012)
$G_{t-2}^{\text{browsing}}$	0.033***				0.022**	0.022*
	(0.010)				(0.010)	(0.012)
$G_{t-3}^{\text{browsing}}$	0.018*				0.017*	0.022*
	(0.010)				(0.0094)	(0.011)
$G_{t-4}^{\text{browsing}}$	-0.0024				0.0056	0.0035
	(0.0099)				(0.0095)	(0.011)
G_{t-1}^{sales}		0.0041		0.0047		-0.018
		(0.041)		(0.043)		(0.037)
G_{t-2}^{sales}		-0.038		-0.00053		-0.0021
		(0.047)		(0.049)		(0.042)
G_{t-3}^{sales}		-0.0079		-0.026		-0.024
		(0.047)		(0.044)		(0.039)
G_{t-4}^{sales}		0.075*		0.0033		0.035
		(0.041)		(0.042)		(0.036)
Macro controls (X)			✓	✓	✓	✓
R^2	0.49	0.11	0.57	0.61	0.76	0.78
Adjusted R^2	0.43	0.00	0.40	0.34	0.60	0.55
No. Observations	40	40	44	40	40	40

Notes: This Table presents the estimates for regression equation (2.4). The sales growth rates are winsorized at 1% and 99%. The macroeconomic control variables $\{X_{t-j}\}_{j=1}^4$ include oil supply shocks, monetary policy shocks, the 3-month nominal T-bill rate, and the term spread. The sample periods are from 2006 1st quarter to 2016 4th quarter. *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

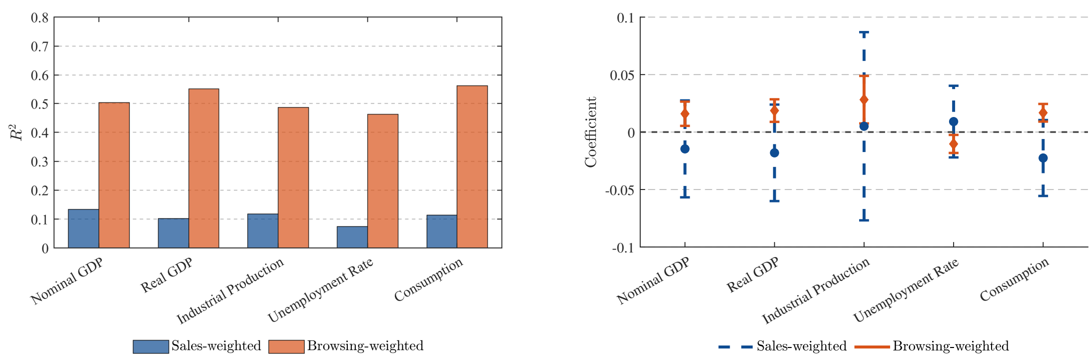
Table 2 presents regression results using industrial production as the forecast target, providing

[9] Our results are quantitatively unchanged if we use the median forecast, as shown in Appendix Table C.2.

[10] The term spread is defined as the 5-year bond rate minus the 3-month bond rate.

support for our hypothesis.[11] An increase in browsing-weighted sales growth in prior quarters is associated with higher mean forecasts of industrial production in the subsequent quarter. Moreover, this measure exhibits strong explanatory power: the adjusted R^2 in column (1) is 43%, whereas column (2) shows that the sales-weighted measure has negligible predictive power. Columns (4)–(6) further demonstrate that the browsing-weighted measure improves model fit by approximately 20 percentage points beyond what is achieved using standard macroeconomic controls, while the sales-weighted measure contributes little.

Figure 5: The Predictive Power of Browsing-weighted Sales Growth



Notes: The left and right panel in this figure plot the R^2 and regression coefficient for the regression (2.4), respectively. Blue color represents the case where only $\{G^{\text{sales}}_{t-j}\}$ are included as regressors, and red color represents the case where only $\{G^{\text{browsing}}_{t-j}\}$ are included as regressors.

Figure 5 summarizes the forecasting results for key macroeconomic aggregates, including GDP, the unemployment rate, and consumption. The left panel illustrates that the predictive power, as measured by the R^2, is consistently and significantly higher when using browsing-weighted sales growth compared to sales-weighted sales growth. The right panel shows that the regression coefficients on lagged sales growth in period $t-1$ are statistically significant in the case of browsing-weighted measures, while those based on sales-weighted measures are generally insignificant.

Summary. The first three facts we document provide a detailed view of firms' attention received and sent, captured by their browsing activity, and their relationship with firm size. Both firms' attention received and sent are dominated by power-law patterns, and these patterns are not merely a reflection of firm size. The fourth fact provides strong empirical support that the performance of a small fraction of firms, which receive disproportaionlly large attention, significantly influences forecasters' expectations about macroeconomic variables. Together, these empirical patterns motivate and discipline our subsequent theoretical and quantitative analysis .

[11] As shown in Appendix Table C.3, the results are robust to alternative forecast variables, including real GDP, the unemployment rate, real personal consumption expenditures, and real nonresidential fixed investment.

3. Framework

In this section, we present the baseline model, which features incomplete information, strategic uncertainty, and potential granular effects. Unlike the majority of the literature, the economy in our model is populated by a finite but potentially large number of firms that are heterogeneous in their attention allocation. We specify attention allocation in terms of directed graphs, and we show how the aggregate impacts of micro shocks interact with the underlying information network structure.

3.1 Environment

Time is discrete and indexed by t. The economy consists of a number of N information islands, each of which is inhabited by a firm indexed by $i \in \{1, 2, ...N\}$. These firms are responsible for producing their respective varieties. A representative household enjoys consumption of final goods and also allocates workers to different information islands for production, a setup akin to Angeletos and La'O (2010).

Each period consists of two stages. In the first stage, workers and firms meet in their local labor markets, where labor demand, labor supply, and firms' outputs are determined under incomplete information about the underlying states of the economy. In the second stage, the centralized goods market opens, and all transactions are conducted under perfect information.

Household. The preference of the representative household is given by

$$U(C_t) - \sum_{i=1}^{N} V(\ell_{it}) = \frac{C_t^{1-\gamma}}{1-\gamma} - \sum_{i=1}^{N} \frac{\ell_{it}^{1+\nu^{-1}}}{1+\nu^{-1}}$$

where the parameter γ controls the income effects and ν parameterizes the Frisch elasticity of labor supply. Here, ℓ_{it} is the labor supply of the workers who are sent to island i. The final goods consumption C_t is a composite of all the varieties, which admits a constant elasticity of substitution (CES) structure,

$$C_t = \left[\sum_{i=1}^{N} \left(\frac{1}{N} \right)^{\frac{1}{\rho}} C_{it}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}},$$

where ρ determines the elasticity of substitution across goods produced by different firms.[12] It follows that the aggregate price index and the corresponding demand schedule for a good i can be specified as

$$P_t = \left[\sum_{i=1}^{N} \frac{1}{N} P_{it}^{1-\rho} \right]^{\frac{1}{1-\rho}}, \qquad C_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\rho} C_t.$$

[12] In our quantitative analysis, we allow this parameter to vary across different firms so that it can help capture the firm-size heterogeneity.

Finally, the budget constraint of the household is

$$\sum_{i=1}^{N} P_{it} C_{it} = \sum_{i=1}^{N} W_{it} \ell_{it} + \sum_{i=1}^{N} \pi_{it},$$

where W_{it} and π_{it} are the local wage and profit in island i, respectively. In the second stage, the household chooses their optimal consumption plan given their income.

Firms. The production technology of firm is given by

$$Q_{it} = \exp(z_{it})\, \ell_{it}^{\theta} \tag{3.1}$$

where z_{it} is firm i's productivity shock at period t and θ represents the output elasticity with respect to labor. We assume that the productivity shock obeys the following process

$$z_{it} = \xi_t + \varepsilon_{it}, \tag{3.2}$$

where $\xi_t \sim \mathbb{N}(0, \tau_{\xi}^{-1})$ is an economy-wide common component and $\varepsilon_{it} \sim \mathbb{N}(0, \tau_{\varepsilon}^{-1})$ is the firm-specific component, respectively.

In the first stage of a period, firms need to determine the amount of labor input and therefore their output scale subject to informational frictions. We assume that firm i and workers sent to island i share the same information. They can observe the productivity z_{it} in their own island, but may not observe productivities in other islands. Consequently, their decisions have to depend on their imperfect local information. We use $\mathbb{E}_{it}[\cdot]$ to denote the local expectation operator, and will specify the details of the information set in Section 3.3.

In the local labor market of the information island i, firm i post wages to attract workers. Given a posted wage W_{it}, workers decide their optimal labor supply according to their intra-temporal Euler equation. Firm i's problem then can be written as

$$\max_{\ell_{it}} \mathbb{E}_{it} \left[\frac{U'(C_t)}{P_t} \left(P_{it} Q_{it} - W_{it} \ell_{it} \right) \right] \tag{3.3}$$

subject to

$$Q_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\rho} \frac{1}{N} C_t, \quad \text{and} \quad \ell_{it}^{\nu^{-1}} = W_{it} \mathbb{E}_{it} \left[\frac{U'(C_t)}{P_t} \right],$$

where the first constraint is the goods demand schedule and the second constraint corresponds to the workers' labor supply schedule. Facing uncertainty about other islands in the economy, firms use households' marginal utility when evaluating their profits across different states. The first-order condition gives rise to the optimal labor demand condition

$$\ell_{it}^{\nu^{-1}} = \theta \frac{\rho - 1}{\rho} \mathbb{E}_{it} \left[U'(C_t) C_t^{\frac{1}{\rho}} N^{-\frac{1}{\rho}} \exp(z_{it})^{1-\frac{1}{\rho}} \ell_{it}^{\theta(1-\frac{1}{\rho})-1} \right], \tag{3.4}$$

which equates the marginal cost and marginal benefit of hiring one more unit of labor. Form condition (3.4), it can be seen that the amount of labor employed is increasing in the firm's own productivity z_{it}, as well as the expected aggregate demand C_t due to demand linkages (provided that the income effect is not too strong). This observation will become more transparent in firms' best response functions.

3.2 Best Response

The goods market clearing condition implies that the $C_{it} = Q_{it}$. Denote GDP in this economy as Q_t which equals the aggregation of individual firms' outputs.

$$Q_t = \left[\sum_{i=1}^{N} \left(\frac{1}{N} \right)^{\frac{1}{\rho}} Q_{it}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}.$$

We define q_{it} and q_t as the log deviation of Q_{it} and Q_t from their steady-state values. Combining the production technology (3.1), the labor demand condition (3.4), and the market-clearing condition above produces the familiar best-response function.

Lemma 3.1. *Firms' optimal responses satisfy the following beauty-contest game*

$$q_{it} = \delta z_{it} + \alpha \mathbb{E}_{it}[q_t], \qquad q_t = \frac{1}{N} \sum_{i=1}^{N} q_{it}, \tag{3.5}$$

where the coefficients are

$$\delta = \frac{1 + \nu}{1 + \nu + \theta\nu \left(\frac{1}{\rho} - 1 \right)}, \quad \textit{and} \quad \alpha = \frac{\theta\nu \left(\frac{1}{\rho} - \gamma \right)}{1 + \nu + \theta\nu \left(\frac{1}{\rho} - 1 \right)}$$

Condition (3.5) summarizes different forces that influence firms' output choices. The coefficient δ parameterizes the partial equilibrium effects, in the sense that holding others' actions constant, how a firm would respond to its own productivity z_{it}. The coefficient α on the other hand parameterizes the general equilibrium effects. As the aggregate demand results from the outputs across all firms, the dependence on it gives rise to the strategic complementarity in firms' output decisions.

To see exactly how an individual firm forms expectations about other firms' outputs, it is useful to iterate on condition (3.5) and reach the following representation

$$q_{it} = \delta z_{it} + \delta\alpha\mathbb{E}_{it}\left[\frac{1}{N}\sum_{j=1}^{N} z_{jt}\right] + \delta\alpha^2\mathbb{E}_{it}\left[\frac{1}{N}\sum_{j=1}^{N}\mathbb{E}_{jt}\left[\frac{1}{N}\sum_{k=1}^{N} z_{kt}\right]\right] + \dots \tag{3.6}$$

That is, firms need to form expectations on others' productivities, as well as expectations on other firms' expectations on productivities, and so on. Because firms receive potentially asymmetric information, the standard law of iterated expectations no longer holds. As a result, aggregate output depends not only on first-order but also on higher-order expectations about other firms' productivities. Different from the envrionment with a continnum of agents (Morris and Shin, 2002) where agents only need to care about the economy average objects, firms in our model economy need to explicitly consider each firm's expectation due to the potential heterogeneity in information structure.

These expectations are crucial in determining how an individual firm's productivity shock transmits to the rest of the economy and affects the aggregate GDP. For example, if no firms observe z_{jt}, then firm j's shock will have no impact on the rest of the economy. In contrast, if a large number firms observes z_{jt}, then these firms can respond to it, with the understanding that other firms are also likely to respond to it. The extent to which an individual firm's shock affects the aggregate economy therefore hinges on the underlying information structure, which we outline next.

3.3 Information Network

Now we specify the information set for firms located on different information islands. The key object is the $N \times N$ matrix $\mathbf{A}$, which we refer to as the information or attention matrix. The elements of $\mathbf{A}$ are either 1 or 0, where

$$A_{ij} = \begin{cases} 1, & \text{if } i \text{ perfectly observe } j\text{'s fundamental } z_{jt}; \\ 0, & \text{otherwise.} \end{cases}$$

As firms can perfectly observe their own productivity, it follows that $A_{ii} = 1$ for all i. It follows that the local information set $\mathcal{I}_{it}$ is the collection of observed productivities: $\mathcal{I}_{it} = \{z_{jt} : A_{ij} = 1\}$.

The information network matrix allows for rich patterns of informational interactions among firms. In contrast to canonical symmetric dispersed-information models, such as Lucas (1973), where agents' information differs only due to their distinct shock realizations, the signal structure represented by $\mathbf{A}$ can be highly asymmetric. Varying the entries between 0 and 1 results in differing degrees of publicity regarding a firm's productivity. Specifically, some firms may receive substantial attention from others, whereas others may become effectively invisible within the economy.

Meanwhile, this formulation nests several commonly used information structures as special cases. For example, by setting $A_{ij} = 0$ for all $i \neq j$, we recover the symmetric dispersed-information model. Alternatively, by setting $A_{ij} = 1$ for all i and j, productivity shocks across all islands become common knowledge, corresponding to the frictionless case.

In-degree and out-degree. To better understand how the information matrix $\mathbf{A}$ governs the allocation of firms' attention, it will be useful to view it as a direct graph and consider its in-degrees and out-degrees. The in-degrees and out-degrees are related to the counterparts of the browsing received and browsing sent in the EDGAR browsing data.

Given an information matrix $\mathbf{A}$, for firm i, it collects a set $\mathcal{D}_i$ of signals, where $\mathcal{D}_i = \{j \in \mathbf{Z} : A_{ij} = 1\}$. We define firm i's total number of browsees as its network out-degree,

$$d_i \equiv \sum_{j=1}^{N} A_{ij} = \sum_{j \in \mathcal{D}_i} A_{ij}. \tag{3.7}$$

Intuitively, d_i measures the amount of attention firm i pays to other firms.

Similarly, for firm i, it is observed by a set $\mathcal{M}_i$ of firms, where $\mathcal{M}_i = \{j \in \mathbf{Z} : A_{ji} = 1\}$. We define firm i's total number of browsers as its network in-degree

$$m_i \equiv \sum_{j=1}^{N} A_{ji} = \sum_{j \in \mathcal{M}_i} A_{ji}. \tag{3.8}$$

Here, m_i measures the amount of attention that firm i receives.

As an example, consider the following information matrix with $N = 4$:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

It is easy to see that $m_4 = 4$, which implies that firm 4 receives most attention from other firms. Meanwhile, $d_1 = 4$, which implies that firm 1 pays most attention to other firms.

Inference. How do individual firms form expectations about other firms' productivities in the information network? The construction of the information matrix $\mathbf{A}$ and the shock process (3.2) imply that

$$\mathbb{E}_{it}[z_{jt}] = \begin{cases} z_{jt}, & \text{if } j \in \mathcal{D}_i, \\ \mathbb{E}_{it}[\xi_t], & \text{if } j \notin \mathcal{D}_i. \end{cases} \tag{3.9}$$

When z_{jt} is not directly observed by firm i, the optimal inference is the expectation of the common component ξ_t. The collection of productivities observed by firm i, denoted as $\mathcal{I}_{it} = \{z_{jt} : j \in \mathcal{D}_i\}$, is informative about ξ_t and effectively provides d_i signals regarding ξ_t. By applying the standard

Bayesian updating formula, we obtain

$$\mathbb{E}_{it}[\xi_t] = \lambda_i \sum_{j=1}^{N} A_{ij} z_{jt} = \lambda_i \sum_{j \in \mathcal{D}_i} z_{jt}, \quad \text{where} \quad \lambda_i = \frac{1}{d_i + \tau_\xi / \tau_\varepsilon}. \tag{3.10}$$

Here, λ_i corresponds to the signal-to-noise ratio and depends on the relative volatility between the common and firm-specific shocks. Importantly, a higher out-degree d_i lowers λ_i, as attention to each observed signal becomes diluted.

In the end, both $\{m_i\}$ and $\{d_i\}$ play a critical role in shaping firms' equilibrium responses to shocks. In what follows, we show that a suitably constructed diluted in-degree index serves as an approximate sufficient statistic for characterizing the influence of the information network on aggregate outcomes.

4. Macro Elasticity to Micro Shocks

In this section, we study how the micro shocks affect aggregate GDP. By construction, GDP is ultimately a function of all firms' productivities. We can express GDP as

$$q_t = \frac{1}{N} \sum_{i=1}^{N} q_{it} = \sum_{i=1}^{N} g_i z_{it},$$

where g_i is the GDP elasticity with respect to firm i's idiosyncratic shock. These elasticities are results of firms' partial and general equilibrium considerations, as well as the expectations shaped by the information network. In this section, we characterize the equilibrium expression for g_i and derive the conditions under which granularity arises—that is, when idiosyncratic shocks to individual firms can influence aggregate outcomes even as the number of firms approaches infinity.

4.1 GE and Higher-Order Expectations

Intuitively, there are three channels at work in determining the magnitude of g_i: (1) the efficiency channel: variations in z_{it} directly changes the production efficiency; (2) the information channel: z_{it} is informative about the aggregate component ξ_t and other firms' productivities; (3) the coordination channel: z_{it} of a salient firm i is more useful in forecasting other firms' decisions.

To understand these channels, we revisit firms' best response in terms of higher-order expectations as in condition (3.6). Aggregating across firms, the GDP can then be expressed as

$$q_t = \delta \sum_i \frac{1}{N} z_{it} + \delta\alpha \sum_i \frac{1}{N} \mathbb{E}_{it}\left[\sum_j \frac{1}{N} z_{jt}\right] + \delta\alpha^2 \sum_i \frac{1}{N} \mathbb{E}_{it}\left[\sum_j \frac{1}{N} \mathbb{E}_{jt}\left[\sum_k \frac{1}{N} z_{kt}\right]\right] + \dots \tag{4.1}$$

The first term in condition (4.1) corresponds to firms' response to their own fundamentals, which represents the direct effects of productivities or the efficiency channel.

The second term in condition (4.1) involves the first-order expectations about others' productivities. As illustrated in Section 3.3, such expectations are either perfect which coincide with the productivities or equal to the expectations about the aggregate component. Based on the definition of $\mathbf{A}$, it follows that

$$\delta\alpha \sum_{i} \frac{1}{N} \mathbb{E}_{it} \left[\sum_{j} \frac{1}{N} z_{jt} \right] = \delta\alpha \sum_{i=1}^{N} \frac{1}{N} \left(\sum_{j \in \mathcal{D}_i} \frac{1}{N} z_{jt} + \sum_{j \notin \mathcal{D}_i} \frac{1}{N} \lambda_i \sum_{k \in \mathcal{D}_i} z_{kt} \right).$$

Notice that the out-degrees and in-degrees play distinctive roles in the inference problem. The out-degree d_i determines the loading on observed signals through λ_i for firm i's own inference, while the in-degreee m_i determines the frequency of firm i's productivity is used for others' inference. In shaping the behavior of the average first-order expectations, the more salient firms with a larger in-degree in the information network plays a more important role: first, when other firms forming expectations about the salient firms, their productivities will be directly utilized; second, when inferring unobserved productivities, salient firms' productivities will be indirectly utilized as informative signals.

The remaining terms in condition (4.1) are related to higher-order expectations terms. Consider for example the element $\mathbb{E}_{it}[\mathbb{E}_{jt}[z_{kt}]]$ and suppose that $A_{jk} = 0$. In this case,

$$\mathbb{E}_{it} \left[\mathbb{E}_{jt}[z_{kt}] \right] = \mathbb{E}_{it} \left[\lambda_j \sum_{q \in \mathcal{D}_j} z_{qt} \right].$$

The productivities of more salient firms will be more heavily used for another reason: knowing that other firms will be more likely to rely on salient firms' productivities when making inferences, an individual firm itself will also rely more on salient firms' productivities as they are more useful in forecasting others' forecasts. Intuitively, the fundamentals of salient firms provide an additional coordination role as semi-public signals.

Though these channels can be complex due to rich heterogeneity of the attention allocation, the following proposition provides a compact solution to the equilibrium elasticities.

Proposition 4.1. *The GDP elasticities of micro shocks, $\boldsymbol{g} = \begin{bmatrix} g_1 & \dots & g_N \end{bmatrix}'$, are given by*

$$\boldsymbol{g} = \frac{\delta}{N} \left(\mathbf{I} - \alpha \frac{1}{N} \operatorname{diag}(\boldsymbol{m}) - \alpha \frac{1}{N} \mathbf{A}' \operatorname{diag}(\boldsymbol{\lambda}) (\mathbf{1} - \mathbf{A}) \right)^{-1} \mathbf{1}_{N \times 1}, \tag{4.2}$$

where $\boldsymbol{m} = [m_1, \dots, m_N]$ are the vector of in-degrees and $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_N]$ are the vector of signal-to-noise ratios.

To unpack Proposition 4.1, it is useful to define two auxiliary matrices contained in condition (4.2):

$$\mathbf{T}_1 = \alpha \frac{1}{N} \operatorname{diag}(\boldsymbol{m}) \qquad \text{and} \qquad \mathbf{T}_2 = \alpha \frac{1}{N} \mathbf{A}' \operatorname{diag}(\boldsymbol{\lambda}) (\mathbf{1} - \mathbf{A}).$$

The matrix $\mathbf{T}_1$ capture the effects of directly observed productivities when forming expectations, and in-degrees $\boldsymbol{m}$ measure firms' salience. The matrix $\mathbf{T}_2$ captures the effects of indirect inference about unobserved productivities when forming expectations, and the signal-to-noise ratios $\boldsymbol{\lambda}$ control firms' responsiveness to their signals.

The general equilibrium consideration further requires firms to take into account all other firms' decision making process. The expansion of condition (4.2) uncovers these effects

$$\boldsymbol{g} = \frac{\delta}{N}\Big(\mathbf{I} + (\mathbf{T}_1 + \mathbf{T}_2) + (\mathbf{T}_1 + \mathbf{T}_2)^2 + (\mathbf{T}_1 + \mathbf{T}_2)^3 + \dots\Big)\mathbf{1}_{N\times 1}. \tag{4.3}$$

In this expansion, the first term $\frac{\delta}{N}\mathbf{I}$ correspond to the partial equilibrium effects that firms respond to their own productivity. The $k+1$-th term $(\mathbf{T}_1 + \mathbf{T}_2)^k$ then corresponds to the k-th order higher-order expectations in condition (4.1). As semi-public signals, salient firms play a disproportionately important role in higher-order network effects due to their coordination function, thereby amplifying their influence on aggregate economic outcomes. In Section 6.2, we provide a quantitative assessment of the role of salient firms in driving each of the k-th order network effects.

Special cases. It is useful to consider two special cases with symmetric information structures: a perfect information economy and a canonical dispersed-information economy where firms only observe their own productivities.

In the perfect-information economy, $A_{ij} = 1$ for all i and j and the in-degrees $m_i = N$ for all i. It follows that the GDP is proportional to the average productivities,

$$q_t = \sum_{i=1}^{N} g_i z_{it}, \qquad \text{where} \qquad g_i = \frac{\delta}{N}(1-\alpha)^{-1}.$$

The multiplier $\delta(1-\alpha)^{-1}$ determines the overall responsiveness, but the distinction between the PE consideration (δ) and the GE consideration (α) is irrelevant.

In the cannoical dispersed-information economy, A_{ij} equals the identity matrix, $m_i = d_i = 1$ for all i, and $\lambda_i = \lambda = \frac{1}{1+\tau_\xi/\tau_\varepsilon}$. We have

$$q_t = \sum_{i=1}^{N} g_i z_{it}, \qquad \text{where} \qquad g_i = \frac{\delta}{N}\left(1 - \frac{\alpha}{N} - \frac{\alpha}{N}(N-1)\lambda\right)^{-1}.$$

In this case, g_i becomes much smaller, which is due to the standard dampening effects of informational frictions with $\lambda < 1$. Meanwhile, δ and α starts to play a different role as higher-order expectations differ from first-order ones.

Although the magnitudes of g_i differ across the two cases, they share the feature that $g_i \to 0$ as $N \to \infty$, implying that firm-level idiosyncratic shocks vanish in the aggregate in the limit of a large economy. In both cases, the information matrix $\mathbf{A}$ is symmetric. This raises the question: can an asymmetric information structure give rise to granular effects? In what follows, we address this question directly.

4.2 Bounds on the Elasticity g_i

As suggested in the preceding discussion, within an information network, firms with high in-degree tend to exert greater influence. However, the extent to which other firms respond to a salient firm also depends on their own out-degrees, as attention is diluted when they monitor many others. To capture a firm's effective importance in this environment, we define a diluted in-degree index that incorporates both of these opposing forces.

Definition 1. *An attention diluted in-degree index for firm i is*

$$\mu_i = \sum_{j \in \mathcal{M}_i} \frac{1}{d_j}. \tag{4.4}$$

To illustrate the intuition behind this index, consider the case where all firms observing firm i do not observe any other firms. In this case, $d_j = 1$ for all $j \in \mathcal{D}_i$, and the diluted in-degree index reduces to the original in-degree, $\mu_i = \sum_{j \in \mathcal{M}_i} 1 = m_i$. In contrast, as d_j increases—that is, as observers of firm i pay attention to more firms—the weight placed on firm i diminishes, with each observer allocating only $\frac{1}{d_j}$ of its attention to i. This dilution is reminiscent of the role out-degree plays in shaping the signal-to-noise ratio λ_i. Importantly, this diluted in-degree index turns out to be exactly what is needed to characterize the bounds on the elasticity g_i.

Proposition 4.2. *The bounds of GDP elasticity to z_{it} are given by*

$$\frac{\delta}{N}\left(1 + \frac{\alpha}{1 + \tau_\xi/\tau_\varepsilon} \cdot \mu_i\right) \leq g_i \leq \frac{\delta}{N}\left(1 + \frac{\alpha}{(1-\alpha)^2} \cdot \mu_i\right) \tag{4.5}$$

In principle, g_i relies on the entire information matrix, including information on both in-degrees and out-degrees. Proposition 4.2 shows that the diluted in-degree index μ_i succinctly summarizes these forces and is the key statistics to estimate an individual firm's quantitative importance. To see how to obtain the bounds, one needs to leverage on the expansion (4.3). The lower bound include only the two terms—the shock itself and the first-order expectations—-and bounded it from below. In contrast, the upper bound include all the terms in the expansion and bound it from above. In addition, the ratio between the upper bound and the lower bound of g_i is a constant independent of N.[13]

A noteble feature of Proposition 4.2 is that it provides estimates for g_i without the need to solve the equilibrium. Computing these bounds only requires known structural parameters. Additionally, these bounds allow one to identify the role of these structural parameters, particularly in quantitative

[13]Since $0 < \alpha < 1$, it is easy to deduce that $1 \leq \frac{1+\frac{\alpha}{(1-\alpha)^2}\cdot\mu_i}{1+\frac{\alpha}{1+\tau_\xi/\tau_\varepsilon}\cdot\mu_i} \leq \frac{1+\tau_\xi/\tau_\varepsilon}{(1-\alpha)^2}$.

analysis. For instance, reducing the relative volatility, as measured by $\tau_\xi/\tau_\varepsilon$, tends to increase the lower bound, as both the information and coordination channels are strengthened.

4.3 Granularity

Equipped with the bound estimates, now we discuss the limiting case where the number of firms, N, approaches infinity. In standard macroeconomic models, the impact of an individual firm on aggregate GDP tends to diminish due to the law of large numbers. This effect can be seen in Proposition 4.2, as both the lower and upper bounds are multiplied by $\frac{1}{N}$. This reasoning also applies to the two special cases involving perfect information or symmetric dispersed information. To allow for granular effects, it is necessary that, for some firms, the diluted in-degree index increases with N at a fast enough rate.

Proposition 4.3. *The micro shocks can generate aggregate fluctuations if and only if* $\max_i \mu_i = \Theta(N)$.[14]

Proposition 4.3 establishes that a necessary and sufficient condition for the economy to exhibit granular effects is that the firm with the largest diluted in-degree index increases with N at the same rate. Intuitively, for a firm i to experience granular effects, it must be observed by many other firms, while at the same time, the attention of these firms must be concentrated on firm i. We further illustrate these implications with the following corollary.

Corollary 4.1. *Given the diluted in-degree index μ_i defined in*

1. *The diluted in-degree index satisfies the following property*

$$\sum_{i=1}^{N} \mu_i = N,$$

and there are at most $K = o(N^a)$ firms that can have granular effects where $a \in (0, 1]$.

2. *A necessary condition to admit granular effects is that* $\max_i m_i = \Theta(N)$.

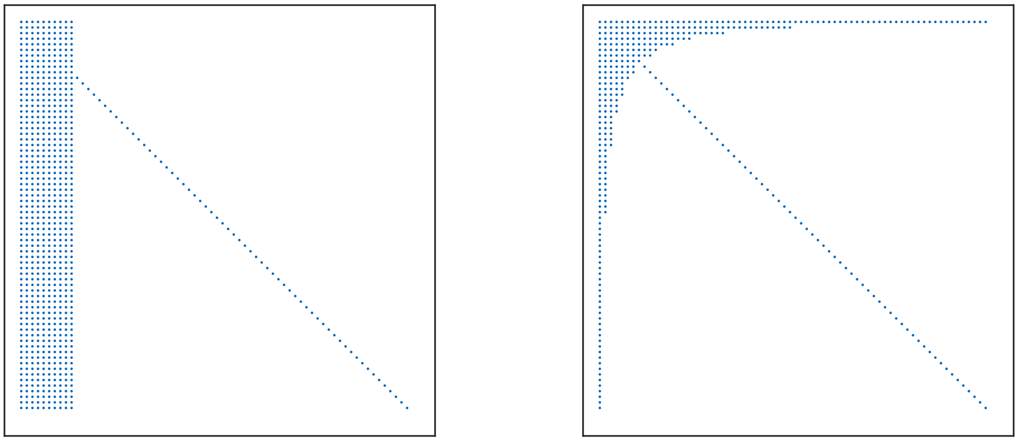
Part 1 of Corollary 4.1 implies that μ_i is bounded from above by N, and there can only be a small number of firms (relative to the size of the economy) that can generate granular effects.

Part 2 underscores the importance of a large in-degree for the granular effect, which has to grow at the same rate of N. However, this is not sufficient to guarantee the granular effect, as illustrated in the following examples.

Example 1. In the first example, we assume that all firms to observe a fixed fraction ω of firms' productivities in the economy, in addition to their own. The left panel of Figure 6 displays the corresponding information matrix. Note that for the first ωN firms, they receive attention from all

[14]For two functions $f(N)$ and $q(N)$, we write $f(N) = \Theta(q(N))$ if there exists some positive constants c_1, c_2 such that $c_1 q(N) \le f(N) \le c_2 q(N)$ for all N large enough.

Figure 6: Examples of Information Matrix



Notes: The left panel displays the information matrix where a fixed fraction of firms are observed by all other firms. The right panel displays the information matrix where the in-degree and out-degree follow the Zipf's law.

the firms, which implies their in-degree satisfies $m_i = N$. On the other hand, the out-degrees in this economy are given by

$$d_i \approx \omega N.$$

Can the first ωN firms exhibit granular effects? The answer is no. Notice that in this economy $\max_i \mu_i = \mu_1$, where

$$\mu_1 = \sum_{j \in \mathcal{M}_1} \frac{1}{d_j} \approx N \frac{1}{\omega N} = \omega.$$

Clearly, the diluted in-degree index is not growing with N. Though the first ωN firms receive a lot of attention, the attention is diluted due to the increasing out-degrees.

Example 2. In the second example, we assume that firms' in-degrees and out-degrees are perfectly correlated, and the in-degrees follow a discrete Zipf's law distribution. Specifically, the information matrix is constructed as follows

$$A_{ij} = 1 \quad \text{if} \;\; i \cdot j \leq N,$$

which is visualized by the right panel of Figure 6. The corresponding in-degrees and out-degrees are given by

$$d_i = m_i = \frac{N}{i}.$$

Can micro shocks generate aggregate fluctuations in this economy? The answer is yes. Focus on the diluted in-degree index for firm 1:

$$\mu_1 = \sum_{j \in \mathcal{M}_1} \frac{1}{d_j} = \sum_{j=1}^{N} \frac{j}{N} = \frac{1+N}{2} = \Theta(N),$$

which is proportional to N. In this economy, firm 1 receives a lot of attention from other firms, and most of these firms only pay attention to a small number of firms, leaving their attention towards firm 1 undiluted. This example is particularly informative as the attention allocation in the EDGAR browsing data displays a similar pattern.

5. Attention and Firm Sizes

In this section, we extend the baseline model to allow for firms of different sizes. We examine how aggregate GDP volatility is jointly influenced by heterogeneity in the distributions of firm size and attention. Additionally, we present an example where both follow power law distributions, enabling direct comparison with the results in Gabaix (2011).

5.1 Firm Size Heterogeneity

In our setting, differences in firm size can arise from either the supply side—through production efficiency—or the demand side—due to consumer preferences. Assume first that the production technology for firm i is

$$Q_{it} = \exp\left(\bar{z}_i + z_{it}\right) n_{it}^{\theta}.$$

Here, $\bar{z}_i$ represents firm i's steady-state level of productivity.

Meanwhile, we modify the final goods consumption bundle as

$$C_t = \left[\sum_{i=1}^{N} (\zeta_i)^{\frac{1}{\rho}} C_{it}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}},$$

where ζ_i is the taste parameter for products from firm i.

With these changes, the steady-state levels of firms' market shares differ, creating room for size effects to amplify micro shocks. In this environment, the best-response condition and the corresponding GDP accounting equation are stated in the following lemma.

Lemma 5.1. *The equilibrium output for firm i satisfies linear best response function,*

$$q_{it} = \delta z_{it} + \alpha \mathbb{E}_{it}[q_t], \qquad q_t = \sum_{i=1}^{N} \omega_i q_{it}, \tag{5.1}$$

where the steady-state sales share ω_i is given by

$$\omega_i = \frac{\zeta_i^{\varkappa}\left(\exp(\bar{z}_i)\right)^{(\rho-1)\varkappa}}{\sum_{j=1}^{N}\zeta_j^{\varkappa}\left(\exp(\bar{z}_j)\right)^{(\rho-1)\varkappa}}, \qquad \textit{with} \qquad \varkappa = \frac{\nu}{\rho\nu + (1-\rho)\theta}.$$

Note that the best-response condition remains unchanged, while firm size is now reflected in the GDP aggregation. The combination of demand and supply heterogeneities shapes the steady-state firm size $\{\omega_i\}$. For our purpose of understanding the effects of micro shocks on aggregate fluctuations, what matters is the distribution of $\{\omega_i\}$, while the exact split between the demand and the supply force is less important.

Consider momentarily the perfect information benchmark. In this case, the inference about others' decisions and therefore the aggregate output is perfect, $\mathbb{E}_{it}[q_t] = q_t$. It follows that the GDP can be expressed as

$$q_t = \sum_{i=1}^{N}\omega_i(\delta z_{it} + \alpha q_t) = \frac{\delta}{1-\alpha}\sum_{i=1}^{N}\omega_i z_{it} \tag{5.2}$$

That is, the GDP elasitcities with respect to micro shocks, $\{g_i\}$, are proportional to their size. Whether granular effects take place or not hinges on the property of the distribution of $\{\omega_i\}$, as shown in Gabaix (2011). The scalar $\frac{\delta}{1-\alpha}$ captures the general equilibrium effects, which are uniform across firms.

When information is incomplete, the aggregate fluctuations are jointly shaped by the information matrix and the firm-size distribution. Given an information network $\mathbf{A}$, the solution to the fixed-point problem (5.1) remains tractable.

Proposition 5.1. *Given information network* $\mathbf{A}$ *and the vector of firm size* $\boldsymbol{\omega} = \begin{bmatrix}\omega_1 & \dots & \omega_N\end{bmatrix}'$, *the GDP elasticities of micro shocks,* $\boldsymbol{g} = \begin{bmatrix}g_1 & \dots & g_N\end{bmatrix}'$, *are given by*

$$\boldsymbol{g} = \delta\left[\mathbf{I} - \alpha\,\mathrm{diag}(\mathbf{A}'\boldsymbol{\omega}) - \alpha\mathbf{A}'\,\mathrm{diag}(\boldsymbol{\omega})\,\mathrm{diag}(\boldsymbol{\lambda})\,(\mathbf{1}-\mathbf{A})\right]^{-1}\boldsymbol{\omega}. \tag{5.3}$$

Solution (5.3) highlights the interaction between firm size, represented by $\boldsymbol{\omega}$, and the information network, $\mathbf{A}$. Parallel to the analysis in Section 4.1, we define

$$\mathbf{T}_1 = \alpha\,\mathrm{diag}(\mathbf{A}'\boldsymbol{\omega}), \quad \text{and} \quad \mathbf{T}_2 = \alpha\mathbf{A}'\,\mathrm{diag}(\boldsymbol{\omega})\,\mathrm{diag}(\boldsymbol{\lambda})\,(\mathbf{1}-\mathbf{A}),$$

which capture the general equilibrium forces via direct and indirect inference problems, respectively. The products of the information matrix and the firm size represents their interactive effects, which are absent in our baseline analysis.

Compared to the benchmark scenario with perfect information (5.2), the general equilibrium effects are no longer $\frac{\delta}{1-\alpha}$, but become heterogeneous across firms. A large firm's influence on GDP is amplified when it attracts considerable attention within the network. Conversely, this influence is attenuated if the firm receives relatively little attention.

Influence index and granularity. Parallel to previous analysis in Section 4, we introduce a firm influence index that summarizes the effects of both firm size and the attention diluted in-degree.

Definition 2. *The influence index for firm i is defined by*

$$\mu_i^\omega = \omega_i + \sum_{j \in \mathcal{M}_i} \frac{\omega_j}{d_j}. \tag{5.4}$$

In this influence index, the first component mechanically captures the firm size channel. The second component mirrors the in-degree index in Definition 1 and reflects the effective attention received by firm i, measured as a weighted in-degree that accounts for both the out-degree and the size of the firms observing i. This index is informative because the asymptotic behavior of individual firms' aggregate impact hinges on its limiting properties.

Proposition 5.2. *With both heterogeneous firm sizes and the information network, micro shocks can generate aggregate fluctuations if and only if*

$$\max_i \mu_i^\omega = \max_i \left(\omega_i + \sum_{j \in \mathcal{M}_i} \frac{\omega_j}{d_j} \right) = \Theta(1).$$

As the number of firms N grows, the relative size of any individual firm ω_i tends to converges to zero, causing the influence index μ_i^ω to diminish accordingly. For a firm to retain aggregate relevance in the limit, it must either remain sufficiently large relative to the overall economy or receive a disproportionately large share of attention from other firms.

As a special case, when $\omega_i = \frac{1}{N}$ for all i, the index simplifies to

$$\mu_i^\omega = \frac{1}{N} + \frac{1}{N} \sum_{j \in \mathcal{M}_i} \frac{1}{d_j}.$$

It is straightforward to verify that the condition $\max_i \mu_i^\omega = \Theta(1)$ in Proposition 5.2 holds if and only if $\max_i \sum_{j \in \mathcal{M}_i} \frac{1}{d_j} = \Theta(N)$, thereby recovering the result in Proposition 4.3.

5.2 The Case with Power Law Distributions

To further explore the interplay between the attention and firm-size channels, we analyze an example where firm-level attention measures $\{m_i\}$ and firm sizes $\{\omega_i\}$ follow power-law distributions, while network degrees $\{d_i\}$ are uniformly distributed. In this context, the key economic properties hinge upon the tail exponents of the distributions and their correlation structure. We then provide qualitative and quantitative characterizations of how the GDP volatility depends on these parameters.

Random graph formulation. We follow Gabaix (2011) and consider a series of economies indexed by N. For the information matrix, we adopt a probabilistic formulation of the information network based on models of Chung et al. (2003a) and Chung et al. (2003b), which is a generalization of the classical Erdos-Renyi model of random graph. Instead of assuming a deterministic $\mathbf{A}$, information linkages among different firms are now stochastic. Specifically, firm i observe firm j's productivity with probability

$$\mathbb{P}\left(A_{ij}=1\right)=p_{ij}=\begin{cases}\frac{N\,j^{-\beta_m^{-1}}-1}{N-1}\in[0,1], & \forall j\neq i,\\ 1, & j=i,\end{cases} \tag{5.5}$$

where $\beta_m \geq 0$. We define the random graph as a collection of Bernoulli random variables $\mathbf{A}$ with heterogeneous probabilities $\{p_{ij}\}_{i,j\in\{1,2,\ldots,N\}}$ such that edges are independently assigned to each pair of vertices (firms) (i,j).[15] We maintain the assumption that each firm observes its own productivity with certainty such that (5.3) holds for any realization of network structure $\mathbf{A}$.[16]

With this formulation, the expected in-degrees is spelled out in the following lemma.

Lemma 5.2. *The expected in-degrees of the random graph is given by a sequence* $\{m_1, m_2, \ldots m_N\}$ *such that*

$$m_j=\sum_{i=1}^{N}\mathbb{E}\left[A_{ij}\right]=Nj^{-\frac{1}{\beta_m}}, \qquad \text{for } j=1,2\ldots N \tag{5.6}$$

The details of the expected out-degree are spelled out in Appendix A.7. In a nutshell, when N is sufficiently large, the expected out-degree can be viewed as uniformally distributed.

Firm size. We assume that the size of the firms $\{\omega_{\iota(1)},\ldots,\omega_{\iota(N)}\}$ in each economy are defined as a sequence

$$\omega_j=\frac{\iota(j)^{-\frac{1}{\beta_\omega}}}{\sum_{k=1}^{N}k^{-\frac{1}{\beta_\omega}}}=\frac{1}{\zeta\left(\frac{1}{\beta_\omega},N\right)}\iota(j)^{-\frac{1}{\beta_\omega}}\in[0,1],^{17} \qquad \text{such that} \qquad \sum_{i=1}^{N}\omega_i=1. \tag{5.7}$$

In this context, firm j in the information network is ranked $\iota(j)$ in terms of firm size. Crucially, $\iota(j)$ may differ from j, indicating that firms receiving the most attention are not necessarily the largest in size.

A sequence of variables $\{x_k\}_{k\in\{1,2,\ldots N\}}$ satisfy power law relations if $x_k = ck^{-\frac{1}{\kappa}}$ for some $\kappa > 0$. By construction, both the expected in-degree $\{m_i\}$ and the firm size $\{\omega_i\}$ are power-law sequences with exponents β_m and β_ω, respectively. The following lemma then establishes the relationship between

[15]If the edge probability is homogeneous, Prob $\left(A_{ij}=1\right)=p$, the prescribed expected degree sequences follows

$$\boldsymbol{m}=\boldsymbol{d}=Np\mathbf{1}$$

where $p\in[0,1]$ is the probability parameter, the model reduces to the standard Erdos-Renyi random graph.

[16]When N is sufficiently large, the off-diagonal probabilities are given by $p_{ij}\approx\frac{1}{(j)^{\frac{1}{\beta_m}}}$.

[17]Here, $\zeta(\beta_\omega^{-1},N)$ is the $N-$th order partial sum of the Riemann-Zeta function.

the power-law sequence and the Pareto distribution. In short, for a sequence of random variables drawn from a Pareto distribution, the power-law sequence is approximately the sample average of repeated draws from the Pareto distribution.

Lemma 5.3. *Suppose a sequence of random variables $x_{(1)} > x_{(2)} > .... > x_{(N)}$ are drawn from the Pareto distributions.*

$$\mathbb{P}(X > x) = \frac{c^{\kappa}}{N} \cdot x^{-\kappa}, \tag{5.8}$$

defined on support $x \in \left[\frac{c}{N^{\frac{1}{\kappa}}}, \infty\right]$ with $\kappa \geq 0$. Then up to a first-order approximation, when $N \to \infty$

$$\mathbb{E}\left[x_{(k)}\right] = x_k = ck^{-\frac{1}{\kappa}}; \qquad \forall k = 1, 2, ...N \tag{5.9}$$

where the expectation is defined with respect to the Pareto distribution.

GDP volatility. For our purposes, the key primitives of the economy are given by the distributions of m_i and $\omega_{\iota(i)}$, along with the ordering function $\iota(i)$, which governs the matching pattern between these two characteristics. When both sources of heterogeneity are present, a natural question arises: how does GDP volatility, conditional on micro shocks, behave as $N \to \infty$? The following proposition addresses this question.

Proposition 5.3. *Denote $\beta \equiv \min\{\beta_m, \beta_\omega\}$. Independent of the ordering of $\{\iota(i)\}$, as $N \to \infty$, the GDP volatility conditional on micro shocks follow*

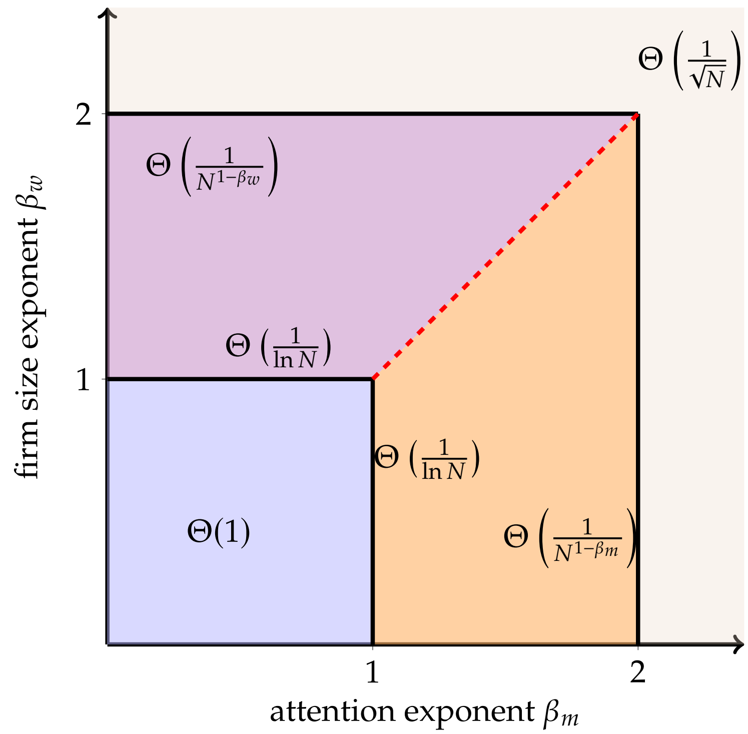
1. $\mathbb{E}[\sigma_{GDP}] = \Theta\left(\frac{1}{\sqrt{N}}\right)$ *when* $\beta > 2$.
2. $\mathbb{E}[\sigma_{GDP}] = \Theta\left(\frac{\sqrt{\ln N}}{\sqrt{N}}\right)$ *when* $\beta = 2$.
3. $\mathbb{E}[\sigma_{GDP}] = \Theta\left(\frac{1}{N^{1-\beta}}\right)$ *when* $1 < \beta < 2$.
4. $\mathbb{E}[\sigma_{GDP}] = \Theta\left(\frac{1}{\ln N}\right)$ *when* $\beta = 1$.
5. $\mathbb{E}[\sigma_{GDP}] = \Theta(1)$ *when* $\beta < 1$.

The expectation is defined with respect to the joint probability measure on random graph $\mathbb{P}(\mathbf{A})$.

This proposition makes it clear that qualitatively, to understand the decay rate of GDP volatility driven by micro shocks, it is sufficient to pay attention to the distribution with a smaller exponent. Supposing that $\beta_m > 2$, or the attention distribution is close enough to a uniform distribution, then the behavior of the economy is exactly the same as that in Gabaix (2011). If the opposite applies, then the stochastic property of the economy will be dominated by the salient firms instead. Figure 7 then visualizes Proposition 5.3.

Importantly, the result in Proposition 5.3 does not depend on the correlation between the two distributions. Whether big firms receive the most attention or small firms receive the most attention is

Figure 7: Decay Rates of Volatility with Power Law Distributions



irrelevant in determining the decay rate of the GDP volatility. However, this should not be interpreted as meaning that the correlation between the two distributions is unimportant. In contrast, correlation is an important factor in determining the quantitative relevance of granular effects. This is particularly important when the sample size is finite and when the exponents of the two distributions are similar to each other, which is the case empirically.

The following proposition further emphasizes the role of the positive assortative matching between firm size and received attention.

Proposition 5.4. *Let $N \to \infty$ and fix two sequences for $\{m_i\}$ and $\{\omega_i\}$. Suppose (i) $\sum_{i \in \mathcal{M}_k} \omega_i \approx \sum_{i \in \mathcal{M}_k} \frac{1}{N} = \frac{m_k}{N}$ for each $1 \le k \le N$, and (ii) $d_i \approx d$ for some $d = o(N)$ and each $1 \le i \le N$. Then up to first-order approximation of $(\mathbf{I} - \mathbf{T}_1 - \mathbf{T}_2)^{-1}$, the GDP volatility conditional on micro shocks satisfies*

1. *σ_{GDP} is maximized when the ordering of $\{m_i\}$ coincides with that of $\{\omega_{\iota(i)}\}$.*
2. *σ_{GDP} is minimized when the ordering of $\{m_i\}$ is the exact opposite of $\{\omega_{\iota(i)}\}$.*
3. *σ_{GDP} decreases as the number of inversions between $\{m_i\}$ and $\{\omega_{\iota(i)}\}$ increases.*[18]

Under the random graph formulation given by (5.5), when $\beta_m > 1$, the two assumptions $\sum_{i \in \mathcal{M}_k} \omega_i \approx \sum_{i \in \mathcal{M}_k} \frac{1}{N} = \frac{m_k}{N}$ and $d_i \approx d$ hold almost surely with probability one.

[18]Here the number of inversions comes from the inversion number of a permutation, which in our setting is equal to the number of pairs $i < j$ such that $(\omega_{\iota(i)} - \omega_{\iota(j)}) \cdot (m_i - m_j) < 0$.

6. Quantification

In this section, we aim to quantify the extent to which micro-level shocks propagate into aggregate fluctuations through the information network. We first outline the calibration strategy, followed by an analysis of the role of information heterogeneity alone and its interactive effects with the traditional firm-size channel. In the end, we provide additional testable implications of the model.

6.1 Calibration

Information Matrix. To connect the browsing activities from EDGAR and the information matrix in our framework, we continue to adopt the random graph approach as in Section 5.2. Starting with the bilateral browsing matrix constructed from EDGAR, we define $\widetilde{m}_i$ as the number of browsing requests received by firm i and $\widetilde{d}_i$ as the number of browsing requests sent by firm i in the sample. In the model economy, A_{ij} can either be 0 or 1, and we specify the probability that $A_{ij} = 1$ as

$$\mathbb{P}(A_{ij} = 1) = \min\left(c\left(\frac{\widetilde{m}_j}{\max_k \widetilde{m}_k}\right)^{\kappa}(1 + \varrho_j \widetilde{d}_i),\ \bar{P}\right). \tag{6.1}$$

In formula (6.1), the term $\left(\frac{\widetilde{m}_j}{\max_k \widetilde{m}_k}\right)^{\kappa}$ captures the idea that if firm j receives more browsing requests in EDGAR, it is more likely that firm i observes j. The term $(1 + \varrho_j \widetilde{d}_i)$ captures the idea that if firm i sends more browsing requests in EDGAR, it is more likely that firm i will also observe firm j.

The parameter κ controls for the tail distributions of the degrees in the random graph, we set $\kappa = 1.8$ to target the power-law coefficient estimated for the attention received in the browsing data. In particular, for $j = 1, 2, ...N$,

$$\mathbb{E}\left[m_j\right] = \sum_{i=1}^{N} \mathbb{P}(A_{ij} = 1) \propto \left(\frac{\widetilde{m}_j}{\max_k \widetilde{m}_k}\right)^{\kappa}.$$

The parameters $\{\varrho_j\}$ determines the correlations between in-degrees and out-degrees of the random graph. Motivated by Fact 3 in Section 2, we employ the following specification,

$$\varrho_j = \begin{cases} \varrho & \text{if } j \in \{\text{top-400-sales firm}\}, \\ 0 & \text{otherwise,} \end{cases}$$

where ϱ is set to 0.003 and it allows us to match the degree correlations of the top 400-sales firms in the data.

Finally, we calibrate the constant c in our model to match the observed fraction of active information linkages in the data—that is, the proportion of browser–browsee firm pairs with nonzero interactions. We set $\bar{P} = 0.95$ to discipline the upper bound on in-degrees in the random graph, aligning it with the maximum number of unique visitors observed for any single firm. Together, the parameters

Table 3: Calibrated Parameters

Param.	Value	Data Source	Empirical Target
Exogenously determined parameters			
γ	0.2	—	income elasticity
ν	2	—	Frisch elasticity
ρ	0.4	—	elasticity of substitution
θ	0.66	—	labor share
$\{\omega_i\}_{i=1}^N$	—	Compustat	sales' share
$\overline{\sigma_i}$	0.176	Compustat	weighted average of idiosyncratic productivity shock volatility
$\{\sigma_i\}_{i=1}^N$	—	Compustat	idiosyncratic productivity shock volatility
σ_ξ	0.02	San Francisco Fed	aggregate TFP shock volatility
Endogenously determined parameters			
c	0.2631	EDGAR	fraction of browser-browsee linkages
κ	1.8	EDGAR	power law exponent in the degree sequences of browsing matrix
ϱ	0.003	EDGAR	correlation between in-degree and out-degree in the browsing matrix
$\bar{P}$	0.95	EDGAR	maximal number of visitors for one single firm

$(\kappa, \varrho, c, \bar{P})$ fully characterize the random graph structure in our model.

Structural Parameters. The structural parameters are calibrated based on standard values in literature. Specifically, we set the income elasticity $\gamma = 0.2$ consistent with Angeletos and La'O (2010), the Frisch elasticity of labor supply to $\nu = 1$ which is in the middle of various micro and macro estimates, the elasticity of substitution among different goods to $\rho = 0.4$ which is within the range of the estimates obtained in Boehm et al. (2019), and the labor share $\theta = 0.66$. These structural parameters jointly imply that in the best response function, the degree of strategic complementarity is $\alpha = 0.50$ and the PE response to firms' own productivity shocks is $\delta = 0.67$.

In the quantitative exercise, we further allow for heterogeneous shock volatilities. Firm i's idioysncratic shock ε_{it} now follows $\varepsilon_{it} \sim \mathbb{N}(0, \tau_i^{-1})$ where τ_i can be different across firms. As a result, firms' shocks will mechanically have heterogeneous contribution to the GDP volatility. In addition, in the inference problem, firms' productivity shocks will not be equally informative about the common shock.

We calibrate the firm size distribution and shock volatilities using firm-level data from Compustat and aggregate TFP statistics from the Federal Reserve Bank of San Francisco.[19] Firm sizes are measured as average annual sales. To estimate firm-level shock volatilities, we follow the standard approach in the literature: we first compute annual firm-level TFP, then calculate the standard deviation of its growth rate for each firm. These firm-level volatilities are subsequently averaged using sales weights to obtain the aggregate measure of firm-level shock volatility.

[19] We use the real-time, quarterly series on total factor productivity (TFP) for the U.S. business sector, adjusted for variations in factor utilization—specifically labor effort and capital workweek—published by the San Francisco Fed.

6.2 Results

Table 4 summarizes the comparison across various model economies. For the information network, we consider three versions: the heterogeneous-information model corresponds our baseline calibration that targeted to the EDGAR browsing pattern; the perfect-information economy corresponds to the one with $A_{ij} = 1$ for all i and j; the Lucas-island economy corresponds to the cannonical dispersed-information economy with $A_{ij} = 1$ only when $i = j$. For all these economies, we start with the case where we mute the heterogeneity in firm sizes and shock volatilities, and we then allow these heterogeneities to play a role.

Table 4: Aggregate and Idiosyncratic Volatility

	Size and Volatility		
	equal size + equal volatility	hetero. size + equal volatility	hetero. size + hetero. volatility
Heterogeneous Info.			
total volatility	1.71	2.03	2.27
idiosyncratic/total	28.37%	45.02%	44.10%
top-100 firms	0.44	0.86	0.97
top-100-sales firms	–	0.79	0.85
top-100-attention firms	0.44	0.62	0.62
Perfect Info.			
total volatility	2.75	3.06	3.19
idiosyncratic/total	14.13%	46.20%	52.16%
top-100 firms	0.06	1.34	1.61
top-100-sales firms	–	1.33	1.50
top-100-attention firms	–	0.54	0.61
Lucas Island			
total volatility	1.36	1.52	1.60
idiosyncratic/total	14.13%	46.31%	51.65%
top-100 firms	0.03	0.66	0.80
top-100-sales firms	–	0.66	0.75
top-100-attention firms	–	0.27	0.30

Notes: "—" indicates missing or unavailable data.

In the economy without heterogeneity in firm sizes and shock volatilities (the first column of Table 4), we isolate the role of the information network. Due to the finite sample size—approximately 4,000 firms—the law of large numbers does not hold exactly. Nevertheless, the share of GDP variance attributable to idiosyncratic shocks in the heterogeneous-information economy is substantially larger than in the other two benchmark settings. The standard deviation of the contribution to GDP volatility by "the top-100 firms"—the most influential firms in terms of GDP volatility contribution—is 0.44, which is approximately seven times higher than in the perfect-information economy and fifteen times higher than in the Lucas island economy.[20] Unsurprisingly, these influential firms correspond to the "top-100-attention firms" that receive the most attention from others in the network.

[20] When firm sizes and idiosyncratic shock volatilities are equalized, firms in both the perfect-information and Lucas-island economies are symmetric. As a result, any subset of 100 firms contributes identically to aggregate volatility.

When firm size heterogeneity is introduced, the second column of Table 4 reveals a markedly different pattern. Under perfect information or in the Lucas-island case, micro-level shocks now account for a substantial portion of aggregate GDP volatility, consistent with the firm-size channel emphasized by Gabaix (2011). In this case, the fluctuations due to micro-shocks are mostly driven by those largest firms, that is, "top-100-sales firms". Note that the firms receiving the most attention contribute substantially less to GDP volatility, with their limited impact primarily driven by overlap with the top 100 firms by sales.

In contrast, within our baseline heterogeneous-information economy, both the "top-100-sales firms" and the "top-100-attention firms" contribute similarly to aggregate fluctuations, indicating that the information channel remains quantitatively important. Furthermore, the magnitude of the GDP volatility driven by the "top-100-sales firms" is smaller than in the perfect-information economy, suggesting that the information channel partially dampens the amplification effect of firm-size heterogeneity.

Mechanism: information network effects. To further understand these results, it is useful to revisit the discussion on higher-order expectations in Section 4.1. In the extended model with firm-size heterogeneity, GDP can be expressed as

$$q_t = \delta \underbrace{\sum_i \omega_i z_{it}}_{\text{1st-order}} + \delta\alpha \underbrace{\sum_i \omega_i \mathbb{E}_{it}\left[\sum_j \omega_j z_{jt}\right]}_{\text{2nd-order}} + \delta\alpha^2 \underbrace{\sum_i \mathbb{E}_{it}\left[\sum_j \omega_j \mathbb{E}_{jt}\left[\sum_k \omega_k z_{kt}\right]\right]}_{\text{3rd-order}} + \dots$$

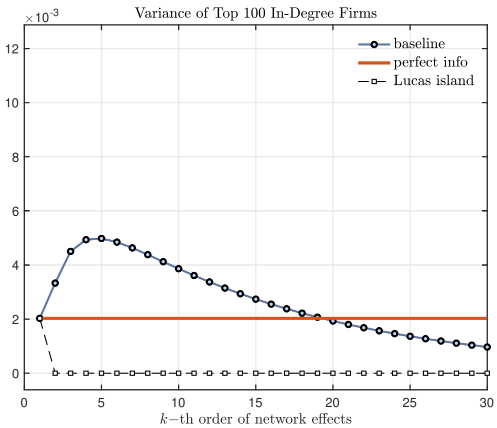
The $k-$th term on the right-hand side of this equation can be interpreted as the $k-$th order network effects, which is related to the $(k-1)$-th order of higher-order expectations. One mechnical force that applies to all the three economies is that as the order increases, its importance declines, captured by $\delta\alpha^k$. In what follows, we focus on the part that is highlighted in blue color, which involves only the expectations.

The introduction of a heterogeneous information network generates two countervailing forces relative to the perfect information benchmark. On one hand, limited information tends to attenuate agents' responsiveness to underlying fundamentals, as uncertainty dampens their reactions (Angeletos and Lian, 2023). On the other hand, higher-order expectations amplify the influence of firms that receive greater attention or visibility, effectively assigning them disproportionate weight in aggregate expectations (Morris and Shin, 2002; Bui et al., 2024). The interplay of these forces governs the extent to which micro-level shocks propagate to the macroeconomy.

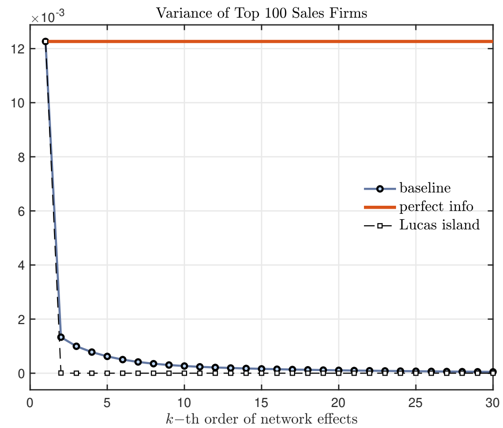
The left panel of Figure 8 displays the variance of the k-th order network effects attributable to the "top-100-attention firms". For the first-order term, the variance is identical across all economies because it is independent of the information network—hence all three lines begin at the same point. In contrast, as the order increases, information heterogeneity becomes consequential. Under perfect

Figure 8: Role of Top Firms in Higher-Order Network Effects

(a) top-100 attention firms



(b) top-100 sales firms



information, higher-order expectations coincide with first-order expectations and equal the average productivity, resulting in a flat red line across different orders. In the Lucas-island economy, higher-order expectations respond much less to fundamentals than first-order expectations—a key insight from Morris and Shin (2002)—leading to a steep decline in the dashed line after $k = 1$.

In our baseline economy, the "top-100-attention firms" attract significant attention and function similarly to *semi-public* signals. If their productivities were fully public, the blue circular line would continue to rise with k, since public signals increasingly shape higher-order beliefs. As semi-public signals, however, their influence initially amplifies with k before declining at higher orders. This initial rise constitutes a strong amplification channel, which helps explain the substantial volatility attributed to top in-degree firms in Table 4.

The right panel of Figure 8 illustrates the variance of k-th order network effects attributed to the top 100 firms ranked by their sales shares. Mechanically, since these firms are larger than the "top-100-attention firms", the network effect at $k = 1$ is correspondingly higher. In contrast to the left panel, the blue circular line declines monotonically with k, although at a slower rate than in the Lucas-island economy. This difference arises because not all high-attention firms are large; in fact, only about 20% of the "top-100-attention firms" also rank among the "top-100-sales firms". The relatively low correlation between firm size and attention implies that the information channel partially offsets, rather than reinforces, the firm-size channel. These two channels amplify each other only when in-degree and firm size are sufficiently positively correlated, as discussed in Section 5.2.

Decomposition: size versus attention. We now present a concise assessment of the relative contributions of firm size and received attention in determining the elasticity of GDP to micro-level shocks. Particularly, we consider the following regression:

$$g_i = \beta_0 + \beta_1 \omega_i + \beta_2 m_i + \varepsilon_i, \tag{6.2}$$

which can be viewed as a linear approximation of the GDP elasticity g_i with firm size ω_i and attention m_i.

Panel A in Table 5 provides the results on the regression coefficients and the associated R^2. The "Baseline" specification refers to the model economy with heterogeneous information, firm sizes, and shock volatilities. In contrast, the "Perfect Info" model corresponds to the one eliminates informational frictions while preserving the other sources of heterogeneity. In our baseline model, the regression coefficients on both m_i and ω_i are positive and significant. The R^2 is 0.73. On the one hand, the results indicate that these two factors explain a substantial portion of the variation in g_i. On the other hand, they also suggest that, through general equilibrium effects, g_i depends on the joint distribution of firm sizes and the information network—an channel emphasized in Proposition 5.1 but omitted from specification (6.2). This is in contrast with "Perfect Info" economy, where the GE effect is homogeneous across firms and g_i is proportional to ω_i as in condition (5.2). As a result, there is no variation in m_i and ω_i accounts for all the variations in g_i.

Table 5: Decomposition of g_i between Size and Attention

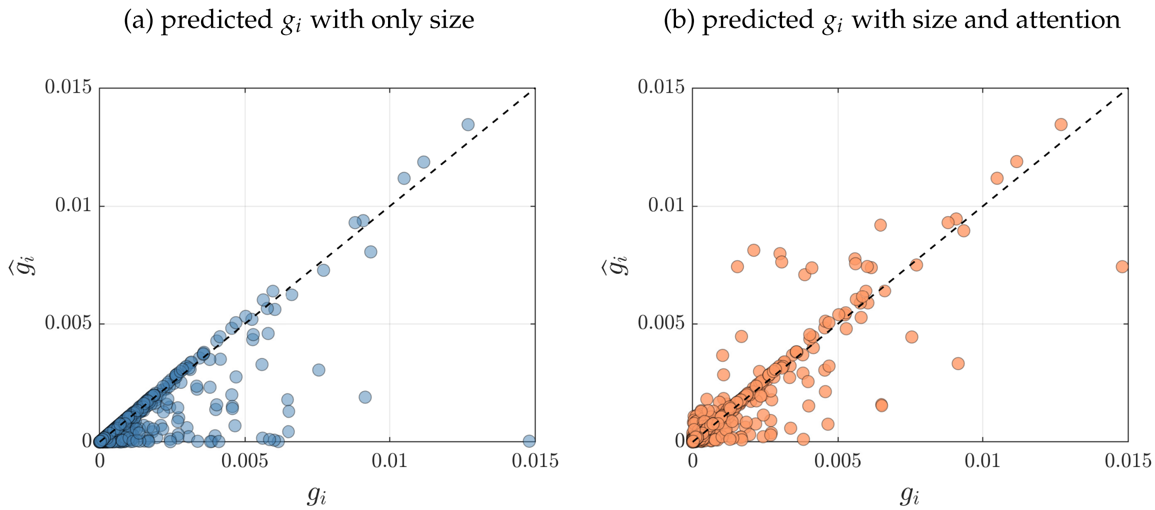
	Baseline			Perfect Info		
	m_i	ω_i		m_i	ω_i	
Panel A			Total			Total
R^2			0.73			1.00
β	0.20***	0.72***		—	1.36***	
Panel B: Explanatory Power						
$\beta \times$ *standard deviation of component*	41.7%	58.3%	100%	—	100%	100%
R^2(incremental inclusion)	41.1%	58.9%	100%	—	100%	100%
R^2(independent inclusion)	37.4%	62.6%	100%	—	100%	100%

Panel B quantifies the relative contributions of attention and firm size in explaining the variation in g_i. We adopt three complementary approaches. First, for each component, we compute a variation-adjusted coefficient defined as the product of its estimated regression coefficient and standard deviation, $\beta \times$ std(component). We then report each component's share in the sum of these variation-adjusted coefficients. Second, we evaluate the incremental explanatory power of each component by sequentially adding m_i and ω_i to a baseline regression that includes only a constant, as in equation (6.2), and compute the corresponding increase in R^2.[21] Third, we estimate separate regressions including each component individually and report the resulting R^2 relative to the baseline specification. Across all approaches, we find that attention accounts for approximately 40% and firm size for 60% of the variation in g_i, highlighting the quantitative importance of the attention channel in shaping the sensitivity of GDP to micro-level shocks.

Figure 9 further visualize their relative importance. In the left panel, when using the estimated

[21] A potential concern is that the incremental R^2 may depend on the order of inclusion. As a robustness check, we verify that our results are highly robust to alternative ordering.

Figure 9: Decomposition of g_i between Size and Attention



coefficients and setting $\widehat{g}_i = \beta_0 + \beta_1 \omega_i$, the predicted GDP elasticity $\widehat{g}_i$ is on average lower than the actual g_i, indicating that the size channel alone is not sufficient. In the right panel, when including both the size and the attention channel, it brings the predicted GDP elasticities much closer to the actual ones.

6.3 Testable Implication

We revisit Fact 4 in Section 2, which demonstrates that the browsing-weighted measure of sales growth exhibits substantial predictive power for forecasts of aggregate variables, outperforming the traditional sales-weighted measure.

In the same spirit, we run the following specification using our model-generated data

$$\frac{1}{N}\sum_{i=1}^{N} \mathbf{E}_{it}[q_t] = \beta_0 + \beta_1 G_t^{\text{attention}} + \beta_2 G_t^{\text{sales}} + \varepsilon_t, \tag{6.3}$$

where $\frac{1}{N}\sum_{i=1}^{N} \mathbf{E}_{it}[q_t]$ is the average forecasts across firms about current GDP, $G_t^{\text{attention}}$ is the attention-weighted measure of sales growth of top-100 attention firms, and G_t^{sales} is the size-weighted measure of sales growth of top-100 sales firms. To facilitate a direct comparison, we also estimate the same regression using data from the Survey of Professional Forecasters.

Table 6: Predictive Power of Browsing-weighted Sales Growth

	Baseline			Perfect Info			Data		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Browsing-weighted	0.350		0.258	0.674		0.137	0.022		0.022
	(0.008)		(0.009)	(0.016)		(0.009)	(0.006)		(0.006)
Sales-weighted		0.367	0.174		0.845	0.743		0.012	0.012
		(0.012)	(0.011)		(0.007)	(0.010)		(0.015)	(0.014)
R^2	0.645	0.503	0.718	0.634	0.935	0.946	0.218	0.013	0.237

Notes: This table presents regression results for the specification: $f_t^{\text{GDP}} = \beta_1 G_{t-1}^{\text{browsing}} + \beta_2 G_{t-1}^{\text{Sales}} + \varepsilon_t$, where f_t^{GDP} is the mean forecast for GDP in period t, and G_{t-1} represents the weighted sales of the top-100 sales or browsing-received firms in period t, $G_{t-1} = \sum_{i=1}^{100} \omega_{i,t-1}(g_{i,t} - \overline{g}_t)$. We define $g_{i,t}$ as firm i's log sales growth in period t and $\omega_{i,t-1}$ is firm i's weight constructed using variables in period $t-1$.

Table 6 presents regression results from two model variants differing in their information structures, along with the corresponding empirical estimates. The empirical findings indicate that the browsing-weighted measure of sales growth is more informative for predicting GDP forecasts. Consistent with this pattern, in our baseline model with a heterogeneous information matrix, the attention-weighted measure of sales growth proves more influential than the traditional sales-weighted measure in accounting for variations in GDP forecasts. By contrast, in the perfect-information model, the sales-weighted measure is more important. The attention-weighted measure retains some relevance only due to partial overlap between the sets of top firms identified by attention and by size.

7. Conclusion

This paper documents substantial heterogeneity in both the attention firms receive and the attention they allocate to others. Using data on firms' browsing behavior related to one another's electronic filings, we show that attention follows a fat-tailed distribution. We develop a theoretical framework that embeds an information network into a noisy business cycle model. Within this framework, we derive conditions on the in-degree and out-degree distributions of the information network under which granular effects emerge. Quantitatively, we find that the information channel can generate sizable aggregate fluctuations from micro-level shocks. However, it also dampens the traditional firm-size channel, as the correlation between firm size and received attention is not sufficiently strong.

Two directions for future research appear particularly promising. First, the information network in our model is taken as exogenous. While this serves as a useful benchmark for positive analysis, it limits the scope for policy evaluation. Incorporating endogenous information acquisition, for instance through the lens of rational inattention, would provide a richer foundation for both normative and positive analysis. Second, our framework abstracts from input-output linkages across firms. Exploring the interaction between the information network and the production network—especially in settings where firm-level transaction data are available—could yield further insights into the propagation of micro shocks.

References

Acemoglu, D., V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi (2012). The network origins of aggregate fluctuations. *Econometrica 80*(5), 1977–2016.

Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2016). Networks, shocks, and systemic risk. In Y. Bramoullé, A. Galeotti, and B. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 569–592. Oxford University Press.

Angeletos, G.-M. and Z. Huo (2021). Myopia and anchoring. *American Economic Review 111*(4), 1166–1200.

Angeletos, G.-M. and J. La'O (2010). Noisy business cycles. *NBER Macroeconomics Annual 24*, 319–378.

Angeletos, G.-M. and C. Lian (2023). Dampening general equilibrium: incomplete information and bounded rationality. In *Handbook of economic expectations*, pp. 613–645. Elsevier.

Auclert, A., M. Rognlie, and L. Straub (2020). Micro jumps, macro humps: Monetary policy and business cycles in an estimated hank model. *Working Paper*.

Baqaee, D. R. (2018). Cascading failures in production networks. *Econometrica 86*(5), 1819–1838.

Baqaee, D. R. and E. Farhi (2020). Productivity and misallocation in general equilibrium. *The Quarterly Journal of Economics 135*(1), 105–163.

Boehm, C. E., A. Flaaen, and N. Pandalai-Nayar (2019). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 tōhoku earthquake. *Review of Economics and Statistics 101*(1), 60–75.

Bui, H., Z. Huo, A. Levchenko, and N. Pandalai-Nayar (2024). Noisy global value chains. *Working Paper*.

Cao, S., W. Jiang, B. Yang, and A. L. Zhang (2023). How to talk when a machine is listening: Corporate disclosure in the age of ai. *The Review of Financial Studies 36*(9), 3603–3642.

Carvalho, V. M. and X. Gabaix (2013). The great diversification and its undoing. *American Economic Review 103*(5), 1697–1727.

Chahrour, R., K. Nimark, and S. Pitschner (2021). Sectoral media focus and aggregate fluctuations. *American Economic Review 111*(12), 3872–3922.

Chen, H., L. Cohen, U. Gurun, D. Lou, and C. Malloy (2020). Iq from ip: Simplifying search in portfolio choice. *Journal of Financial Economics 138*(1), 118–137.

Chung, F., L. Lu, and V. Vu (2003a). Eigenvalues of random power law graphs. *Annals of Combinatorics 7*(1), 21–33.

Chung, F., L. Lu, and V. Vu (2003b). Spectra of random graphs with given expected degrees. *Proceedings of the National Academy of Sciences 100*(11), 6313–6318.

Eeckhout, J., J. D. Loecker, and G. Unger (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics 135*(2), 561–644.

Fang, Y., Z. Huo, M. Li, and J. Wu (2024). Inattentive network: Evidence and theory. *Working Paper*.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*(3), 733–772.

Gabaix, X. and R. Ibragimov (2011). Rank- 1/2: a simple way to improve the ols estimation of tail exponents. *Journal of Business & Economic Statistics 29*(1), 24–39.

Graham, R. L., D. E. Knuth, and O. Patashnik (1994). *Concrete Mathematics: A Foundation for Computer Science* (2nd ed.). Addison-Wesley.

Horn, R. A. and C. R. Johnson (2012). *Matrix analysis*. Cambridge university press.

Jamilov, R., A. Kohlhas, O. Talavera, and M. Zhang (2024, February). Granular Sentiments. Discussion Papers 2414, Centre for Macroeconomics (CFM).

La'O, J. and A. Tahbaz-Salehi (2022). Optimal monetary policy in production networks. *Econometrica 90*(3), 1295–1336.

Lucas, R. E. (1973). Some international evidence on output-inflation tradeoffs. *The American economic review 63*(3), 326–334.

Morris, S. and H. S. Shin (2002). Social value of public information. *american economic review 92*(5), 1521–1534.

Nikolakoudis, G. (2024). Incomplete information in production networks. Working Paper.

Pitman, J. (2012). *Probability*. Springer Science & Business Media.

Plemmons, R. J. (1977). M-matrix characterizations. i—nonsingular m-matrices. *Linear Algebra and its applications 18*(2), 175–188.

Titchmarsh, E. C. and D. R. Heath-Brown (1986). *The Theory of the Riemann Zeta-Function* (2nd ed.). Oxford University Press.

Online Appendix

Micro Shocks and Macro Fluctuations in the Information Network

Zhen Huo   Minghao Li   Jieran Wu   Huan Xiong

A Proofs of Results **A-2**

A.1 Proof of Lemma 3.1 . . . . A-2

A.2 Proof of Proposition 4.1 . . . . A-3

A.3 Proof of Proposition 4.2 . . . . A-5

A.4 Proof of Proposition 4.3 and Corollary 4.1 . . . . A-10

A.5 Proof of Proposition 5.2 . . . . A-12

A.6 Proof of Lemma 5.3 . . . . A-14

A.7 Proof of Lemma 5.2 . . . . A-15

A.8 Proof of Proposition 5.3 . . . . A-16

A.9 Proof of Proposition 5.4 . . . . A-24

B Data Appendix **A-28**

B.1 Text-Based Attention Measure A-28

B.2 Composition of Browsing Activity A-29

C Additional Tables and Figures **A-30**

A. Proofs of Results

A.1 Proof of Lemma 3.1

Proof. The household optimization problem leads to the following first-order condition for labor supply:

$$\left(\frac{1}{N}\right)^{\frac{1}{\rho}} \ell_{it}^{\frac{1}{\nu}} = W_{it}\mathbb{E}_{it}\left[\frac{U'(C_t)}{P_t}\right] = W_{it}\mathbb{E}_{it}\left[\frac{U'(Q_t)}{P_t}\right]$$

where we apply the market-clearing condition for aggregate output. Substitute the optimal labor supply to the firm i's optimization problem,

$$\max_{n_{it}} \mathbb{E}_{it}\left[\frac{U'(Q_t)}{P_t}\left(P_{it}Q_{it} - \frac{\ell_{it}^{\frac{1}{\nu}}}{\mathbb{E}_{it}\left[\frac{U'(Q_t)}{P_t}\right]}\ell_{it}\right)\right]$$

subject to

$$Q_{it} = \left(\frac{P_{it}}{P_t}\right)^{-\rho}\frac{1}{N}Q_t$$
$$Q_{it} = \exp(z_{it})\ell_{it}^{\theta}$$

Since we normalize the aggregate price-index $P_t = 1$, we obtain the first-order condition,

$$\left(\frac{1}{N}\right)^{\frac{1}{\rho}}\left(1+\frac{1}{\nu}\right)\ell_{it}^{\frac{1}{\nu}} = \theta\frac{\rho-1}{\rho}\mathbb{E}_{it}\left[U'(Q_t)\left(\frac{Q_t}{N}\right)^{\frac{1}{\rho}}\exp(z_{it})^{1-\frac{1}{\rho}}\ell_{it}^{\theta(1-\frac{1}{\rho})-1}\right]$$

Taking log to both side of the equation and ignore constant terms and higher-order terms, we obtain log-linearized condition,

$$\frac{1}{\nu}\ell_{it} = \left(\frac{1}{\rho}-\gamma\right)\mathbb{E}_{it}[q_t] + \frac{\rho-1}{\rho}z_{it} + \left(\theta\frac{\rho-1}{\rho}-1\right)\ell_{it}$$

Note that in log-linear form,

$$q_{it} = z_{it} + \theta\ell_{it}$$

Replacing ℓ_{it} by q_{it}, we have

$$q_{it} = \delta z_{it} + \alpha\mathbb{E}_{it}[q_t]$$

where the two composite parameters are defined as

$$\delta = \frac{1+\nu}{1+\nu+\theta\nu\left(\frac{1}{\rho}-1\right)}, \quad \text{and} \quad \alpha = \frac{\theta\nu\left(\frac{1}{\rho}-\gamma\right)}{1+\nu+\theta\nu\left(\frac{1}{\rho}-1\right)}$$

as in the main-text. Finally, by construction, the aggregate output is given by

$$Q_t = \left[\sum_{i=1}^{N}\left(\frac{1}{N}\right)^{\frac{1}{\rho}} Q_{it}^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}.$$

we log-linearize this equation around the symmetric steady state, leading to

$$q_t = \frac{1}{N}\sum_{i=1}^{N} q_{it}, \tag{A.1}$$

and the proof is now complete. □

A.2 Proof of Proposition 4.1

Proof. Now suppose we are given an attention graph $\mathbf{A}$, conjecture the following solution for individual firm decision,

$$q_{it} = \sum_{j=1}^{N} A_{ij} M_{ij} z_{jt} \tag{A.2}$$

where we define the model solution as a $N \times N$ response matrix $\mathbf{M} = [M_{ij}]$. In vector form,

$$q_t = \mathbf{M}\boldsymbol{z}_t \tag{A.3}$$

Inspecting the beauty-contest representation (3.5), the optimal decision q_{it} must be measurable with respect to firm i's signal set $S_{\mathcal{D}_i} = \{z_{jt} : j \in \mathcal{D}_i\}$. Therefore, measurability requires

$$M_{ij} = 0; \text{ if } A_{ij} = 0 \tag{A.4}$$

As such, the equilibrium conjecture follows

$$q_{it} = \sum_{j \in \mathcal{D}_i}^{N} M_{ij} z_{jt} \tag{A.5}$$

To derive the solution, it is necessary to use general representation (A.2). Using the in-out degree construction in the main text, we obtain a compact representation for the expected aggregate output,

$$\frac{1}{N}\mathbb{E}_{it}\left[\sum_{j=1}^{N} q_{jt}\right] = \frac{1}{N}\left[\sum_{j\in\mathcal{D}_i}\left(\sum_{k=1}^{N} A_{kj}M_{kj}\right)z_{jt} + \left\{\sum_{h\notin\mathcal{D}_i}\sum_{k=1}^{N} A_{kh}M_{kh}\right\}\mathbb{E}_{it}\left[\xi_t|S_{\mathcal{D}_i}\right]\right] \tag{A.6}$$

Next, we apply the inference formula (3.10),

$$\frac{1}{N}\mathbb{E}_{it}\left[\sum_{j=1}^{N} q_{jt}\right] = \frac{1}{N}\sum_{j\in\mathcal{D}_i}\left[\sum_{k=1}^{N} A_{kj}M_{kj} + \lambda_i\sum_{h\notin\mathcal{D}_i}\left(\sum_{k=1}^{N} A_{kh}M_{kh}\right)\right]z_{jt}$$

where $\lambda_i = \frac{1}{d_i+\tau_\xi/\tau_\varepsilon}$ as in the main-text. Substitute the conditional expectation into the best-response function, we write the optimal decision rule as

$$q_{it} = \delta z_{it} + \alpha\frac{1}{N}\sum_{j\in\mathcal{D}_i}\left[\sum_{k=1}^{N} A_{kj}M_{kj} + \lambda_i\sum_{h\notin\mathcal{D}_i}\left(\sum_{k=1}^{N} A_{kh}M_{kh}\right)\right]z_{jt} \tag{A.7}$$

given that $i \in \mathcal{D}_i$, $\forall i \in \mathcal{I}$. Comparing (A.7) to the refined conjecture (A.4) and (A.5), we obtain a set of fixed-point systems for the individual-level equilibrium using the method of undetermined coefficients. For each individual firm $i \in \mathcal{I}$, its response vector $\mathbf{M}_i$ is given by

$$M_{ij} = \begin{cases} 0; & j \notin \mathcal{D}_i \\ \alpha\frac{1}{N}\left[\sum_{k\in\mathcal{M}_j} A_{kj}M_{kj} + \lambda_i\sum_{h\notin\mathcal{D}_i}\left(\sum_{k\in\mathcal{M}_h} A_{kh}M_{kh}\right)\right]; & j \in \mathcal{D}_i,\ j \neq i \\ \delta + \alpha\frac{1}{N}\left[\sum_{k\in\mathcal{M}_j} A_{kj}M_{kj} + \lambda_i\sum_{h\notin\mathcal{D}_i}\left(\sum_{k\in\mathcal{M}_h} A_{kh}M_{kh}\right)\right]; & j = i \in \mathcal{D}_i, \end{cases} \tag{A.8}$$

Note here we apply the in-degree property,

$$\sum_{k=1}^{N} A_{kj}M_{kj} = \sum_{k\in\mathcal{M}_j} A_{kj}M_{kj}$$

The fixed-point system (A.8) for matrix $\mathbf{M}$ is a linear equation system; however, characterizing for individual-level solution in closed-form is hard, if impossible. It would require keeping track of the exact attention distribution of $\mathbf{A}$, but the system can be solved numerically without any difficulties.

Now we derive a closed-form solution for the aggregate output. Recall from (A.4), we define the aggregate output as

$$q_t = \frac{1}{N}\sum_{i=1}^{N} q_{it} = \frac{1}{N}\sum_{j=1}^{N}\left(\sum_{k\in\mathcal{M}_j} M_{kj}\right)z_{jt}$$

Therefore, it is natural to define the aggregate-equilibrium solution as a $N \times 1$ vector $\mathbf{g}$ such that

$$g_j \equiv \frac{1}{N} \sum_{k \in \mathcal{M}_j} M_{kj}$$

Now fix a column in matrix $\mathbf{M}$ that corresponds to shock z_{jt}, we sum (A.8) over $k \in \mathcal{M}_j$, we get

$$g_j = \frac{1}{N} \sum_{k \in \mathcal{M}_j} M_{kj} = \frac{1}{N} \left(\delta + \alpha \frac{1}{N} m_j \sum_{k \in \mathcal{M}_j} M_{kj} + \alpha \frac{1}{N} \sum_{k \in \mathcal{M}_j} \left[\lambda_k \sum_{h \notin \mathcal{D}_k} \left(\sum_{l \in \mathcal{M}_h} M_{lh} \right) \right] \right) \tag{A.9}$$

We apply algebra tricks to simplify the last term,

$$\frac{1}{N} \sum_{h \notin \mathcal{D}_k} \left(\sum_{l \in \mathcal{M}_h} M_{lh} \right) = \frac{1}{N} \sum_{h=1}^{N} (1 - A_{kh}) \left(\sum_{l \in \mathcal{M}_h} M_{lh} \right) = \begin{bmatrix} 1 - A_{k1} & 1 - A_{k2} & \dots & 1 - A_{kN} \end{bmatrix} \begin{bmatrix} \frac{1}{N} \sum_{l \in \mathcal{M}_1} M_{l1} \\ \frac{1}{N} \sum_{l \in \mathcal{M}_2} M_{l2} \\ \vdots \\ \frac{1}{N} \sum_{l \in \mathcal{M}_N} M_{lN} \end{bmatrix} = \left(\mathbf{1}_{1 \times N} - \mathbf{A}_{(k,:)} \right) \cdot \mathbf{g}$$

where $\cdot$ denotes inner-product. Therefore, the last term is simplified to

$$\frac{1}{N} \sum_{k \in \mathcal{M}_j} \left[\lambda_k \sum_{h \notin \mathcal{D}_k} \left(\sum_{l \in \mathcal{M}_h} M_{lh} \right) \right] = \sum_{k \in \mathcal{M}_j} \lambda_k \left[\left(\mathbf{1}_{1 \times N} - \mathbf{A}_{(k,:)} \right) \cdot \mathbf{g} \right] = \sum_{k=1}^{N} \lambda_k A_{kj} \left[\left(\mathbf{1}_{1 \times N} - \mathbf{A}_{(k,:)} \right) \cdot \mathbf{g} \right] = \mathbf{e}_j \mathbf{A}^T \operatorname{diag}(\boldsymbol{\lambda}) (\mathbf{1} - \mathbf{A}) \mathbf{g}$$

where $\operatorname{diag}(\boldsymbol{\lambda})$ is the diagonal matrix of signal loading vector $\{\lambda_1, \lambda_2, \dots . \lambda_N\}$, which depends on the attention network's out-degree $\{d_1, d_2, \dots . d_N\}$. $\mathbf{e}_j$ is jth standard basis row vector. $\mathbf{1}$ is $N \times N$ matrix of 1s. Now equation (A.9 becomes

$$g_j = \frac{1}{N} \delta + \alpha \frac{1}{N} m_j g_j + \alpha \frac{1}{N} \mathbf{e}_j \mathbf{A}^T \operatorname{diag}(\boldsymbol{\lambda}) (\mathbf{1} - \mathbf{A}) \mathbf{g} \tag{A.10}$$

where we use the definition $g_j \equiv \frac{1}{N} \sum_{k \in \mathcal{M}_j} M_{kj}$ to the second term. Stacking over $j = 1, 2, \dots N$, we obtain the unique solution for $\mathbf{g}$,

$$\mathbf{g} = \frac{\delta}{N} \left(\mathbf{I} - \frac{\alpha}{N} \operatorname{diag}(\boldsymbol{m}) - \alpha \frac{1}{N} \mathbf{A}^T \operatorname{diag}(\boldsymbol{\lambda}) (\mathbf{1} - \mathbf{A}) \right)^{-1} \mathbf{1}_{N \times 1} \tag{A.11}$$

where $\boldsymbol{m}$ is the vector of attention network's in-degrees. The proof is now complete. □

A.3 Proof of Proposition 4.2

Let us first fix some notations and provide some preliminary results which will be needed later. For two functions $f(n)$ and $g(n)$, we write $f(n) = O(g(n))$ if there exists some positive constant $c > 0$ such that $f(n) \leq c g(n)$ for all n larger than some constant; $f(n) = \Omega(g(n))$ if there exists some positive constant c such that $f(n) \geq c g(n)$ for all n large enough; and $f(n) = \Theta(g(n))$ if there exists some positive constants c_1, c_2 such that $c_1 g(n) \leq f(n) \leq c_2 g(n)$ for all n large enough. Also, we write $f(n) = o(g(n))$ if $\lim_{n \to +\infty} \frac{f(n)}{g(n)} = 0$.

We will need the following results from the matrix analysis. The first one is the Gershgorin Circle

Theorem (see Horn and Johnson (2012)), which can be used to characterize the range of eigenvalues of a given matrix.

Theorem 1 (Gershgorin Circle Theorem). *Let* $\mathbf{A} = (A_{ij})_{n\times n}$ *be an* $n \times n$ *matrix and* λ *be an eigenvalue of* $\mathbf{A}$*. Then there exists some* $1 \le i \le n$ *such that*

$$|\lambda - A_{ii}| \le \sum_{j\neq i} |A_{ij}|.$$

Similarly, there exists some $1 \le i \le n$ *such that*

$$|\lambda - A_{ii}| \le \sum_{j\neq i} |A_{ji}|.$$

We also need the following definition of M-matrix.

Definition 3 (Plemmons (1977)). *An* $n\times n$ *matrix* $\mathbf{A}$ *is called an M-matrix if it can be written as* $A = k\mathbf{I}_n - \mathbf{B}$ *where* $\mathbf{I}_n$ *is the identity matrix,* B *is a non-negative matrix (every element of* B *is non-negative) whose spectral radius (the largest absolute value or complex modulus of its eigenvalues) is strictly less than* k*.*

The following are some properties of M-matrix.

Theorem 2 (Plemmons (1977)). *Let* $\mathbf{A}$ *be an M-matrix. Then* $\mathbf{A}$ *is non-singular and the inverse of* $\mathbf{A}$ *is non-negative (every element of* $\mathbf{A}^{-1}$ *is non-negative).*

Let $\mathbf{A}_i'$ (the transpose of the column vector $\mathbf{A}_i$) be the i-th row of the information matrix $\mathbf{A}$. Then $\mathbf{A}_i$ is the i-th column of $\mathbf{A}'$ (the transpose of $\mathbf{A}$). Thus we obtain

$$\mathbf{A}' \operatorname{diag}(\boldsymbol{\lambda})(\mathbf{1} - \mathbf{A}) = \sum_{i=1}^{N} \lambda_i \mathbf{A}_i(\mathbf{1} - \mathbf{A}_i'). \tag{A.12}$$

By the definition of $\mathbf{A}$, it is easy to check that the following lemma holds.

Lemma A.1. $\mathbf{A}_i(\mathbf{1} - \mathbf{A}_i') = \mathbf{U}_i$ *where* $\mathbf{U}_i = (u_{kj})_{N\times N}$ *such that*

$$u_{kj} = \begin{cases} 1, & \text{if } A_{ik} = 1 \text{ and } A_{ij} = 0 \\ 0 & \text{otherwise} \end{cases}. \tag{A.13}$$

In particular, the diagonal elements of $\mathbf{U}_i$ *are all equal to* 0*.*

Next we study the properties of $\mathbf{V} = \mathbf{T}_1 + \mathbf{T}_2$. By Lemma A.1 we derive the following result.

Lemma A.2. *Let*

$$\mathbf{V} = (\mathbf{v}_{ij})_{N\times N} = \mathbf{T}_1 + \mathbf{T}_2.$$

Then we have:

(1) Every column sum of $\mathbf{V}$ *is at most* $\alpha < 1$*, and thus* $\mathbf{I} - \mathbf{V}$ *is an* M*-matrix.*

(2) Let $k \geq 1$*. Then the k-th row sum* v_k *of* $\mathbf{V}$ *satisfies*

$$\frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \cdot \mu_k = \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \cdot \sum_{i \in \mathcal{M}_k} \frac{1}{d_i} \leq v_k \leq \alpha \cdot \sum_{i \in \mathcal{M}_k} \frac{1}{d_i} = \alpha \cdot \mu_k.$$

(3) For the k-th row of $\mathbf{V}$*, every non-diagonal element is upper bounded by* $\frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} \frac{1}{d_i}$*, thus is smaller than or equal to the diagonal element* $\frac{\alpha}{N} \cdot m_k$*.*

(4) Let $j \geq 1$*. Then every column sum of* $\mathbf{V}^j$ *(the j-th power of* $\mathbf{V}$*) is at most* α^j*. Then* $\lim_{j \to +\infty} \mathbf{V}^j = \mathbf{0}$ *and thus* $(\mathbf{I} - \mathbf{V})^{-1} = \sum_{j=0}^{+\infty} \mathbf{V}^j$*.*

(5) Let $j \geq 1$*, and* $k \geq 1$*. Then the k-th row sum* $v_k^{(j)}$ *of* $\mathbf{V}^j$ *is upper bounded by*

$$v_k^{(j)} \leq j\alpha^j \cdot \mu_k.$$

(6) Let $k \geq 1$*. Then the k-th row sum of* $(\mathbf{I} - \mathbf{V})^{-1} = \sum_{j=0}^{+\infty} \mathbf{V}^j$ *is upper bounded by*

$$1 + \frac{\alpha}{(1-\alpha)^2} \cdot \mu_k,$$

and lower bounded by

$$1 + \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \cdot \mu_k.$$

Proof. (1) Let's consider the j-th column in $\mathbf{V}$ and show that its j-th column sum is at most $\alpha < 1$. First, by Lemma A.1, the j-th diagonal element of $\mathbf{V}$ is the j-th diagonal element of $\frac{\alpha}{N} \operatorname{diag}(\boldsymbol{m})$, thus is equal to

$$\frac{\alpha}{N} \cdot m_j.$$

For the non-diagonal element j-th column sum of $\mathbf{V}$, by Lemma A.1, it only involves

$$\lambda_i \mathbf{A}_i (\mathbf{1} - \mathbf{A}_i')$$

where $A_{ij} = 0$, thus $j \notin \mathcal{D}_i$ and $i \notin \mathcal{M}_j$. Then the non-diagonal element j-th column sum of $\mathbf{V}$ is equal to

$$\frac{\alpha}{N} \cdot \sum_{i \notin \mathcal{M}_j} \sum_{k \in \mathcal{D}_i} \lambda_i. \tag{A.14}$$

Add the above diagonal and non-diagonal elements together, we obtain that the j-th column sum of $\mathbf{V}$ is equal to

$$\frac{\alpha}{N} \cdot m_j + \frac{\alpha}{N} \cdot \sum_{i \notin \mathcal{M}_j} \sum_{k \in \mathcal{D}_i} \lambda_i$$

$$
\begin{aligned}
&= \frac{\alpha}{N} \cdot m_j + \frac{\alpha}{N} \cdot \sum_{i \notin \mathcal{M}_j} d_i \lambda_i \\
&\leq \frac{\alpha}{N} \cdot m_j + \frac{\alpha}{N} \cdot \sum_{i \notin \mathcal{M}_j} 1 \\
&= \frac{\alpha}{N} \cdot m_j + \frac{\alpha}{N} \cdot (N - m_j) \\
&= \alpha.
\end{aligned}
$$

Then by Theorem 1, we know that the spectral radius of $\mathbf{V}$ is at most α, thus by Definition 3 we obtain that $\mathbf{I} - \mathbf{V}$ is an M-matrix.

(2) First, by Lemma A.1, the k-th diagonal element of $\mathbf{V}$ is the k-th diagonal element of $\frac{\alpha}{N}\,\mathrm{diag}(\boldsymbol{m})$, thus is equal to

$$
\frac{\alpha}{N} \cdot m_k.
$$

For the non-diagonal element k-th row sum of $\mathbf{V}$, by Lemma A.1, it only involves

$$
\lambda_i \mathbf{A}_i (\mathbf{1} - \mathbf{A}_i')
$$

where $A_{ik} = 1$, thus $k \in \mathcal{D}_i$ and $i \in \mathcal{M}_k$. Thus the non-diagonal element k-th row sum of $\mathbf{V}$ is equal to

$$
\frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} \sum_{j \notin \mathcal{D}_i} \lambda_i.
$$

Add the above diagonal and non-diagonal elements together, we obtain that the k-th row sum of $\mathbf{V}$ is equal to

$$
\begin{aligned}
&\frac{\alpha}{N} \cdot m_k + \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} \sum_{j \notin \mathcal{D}_i} \lambda_i \\
&= \frac{\alpha}{N} \cdot m_k + \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} \lambda_i (N - d_i) \\
&= \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} (1 - \lambda_i d_i + N \lambda_i)) \\
&= \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} \frac{\tau_\xi + N \tau_\varepsilon}{\tau_\xi + d_i \tau_\varepsilon}.
\end{aligned}
$$

Then

$$
\frac{1}{\tau_\xi / \tau_\varepsilon + 1} \cdot \frac{N}{d_i} \leq \frac{\tau_\xi + N \tau_\varepsilon}{\tau_\xi + d_i \tau_\varepsilon} \leq \frac{N}{d_i}.
$$

Therefore,

$$
\frac{\alpha}{\tau_\xi / \tau_\varepsilon + 1} \cdot \sum_{i \in \mathcal{M}_k} \frac{1}{d_i} \leq v_k \leq \alpha \cdot \sum_{i \in \mathcal{M}_k} \frac{1}{d_i}.
$$

(3) Similarly to (2), the non-diagonal elements $\mathbf{v}_{kj}(1 \leq j \leq N,\ \ j \neq k)$ in the k-th row of $\mathbf{V}$ satisfies

$$\mathbf{v}_{kj} = \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k \text{ and } i \notin \mathcal{M}_j} \lambda_i \leq \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} \lambda_i \leq \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} \frac{1}{d_i} \leq \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} 1 \leq \frac{\alpha}{N} \cdot m_k. \tag{A.15}$$

(4) We prove it by induction. The case $j = 1$ is proved by (1). Now assume that it is true for j. Let $\mathbf{V}^j = (\mathbf{V}_1^j \ \ \mathbf{V}_2^j \ \ \dots \ \ \mathbf{V}_N^j)$ where $\mathbf{V}_i^j$ is the i-th column of $\mathbf{V}^j$. Then the first column of $\mathbf{V}^{j+1} = \mathbf{V}^j \cdot \mathbf{V}$ is

$$\sum_{i=1}^{N} \mathbf{v}_{i1} \cdot \mathbf{V}_i^j.$$

However, the sum of elements in the i-th column $\mathbf{V}_i^j$ in $\mathbf{V}^j$ is at most α^j by induction hypothesis, thus the sum of elements in $\sum_{i=1}^{N} \mathbf{v}_{i1} \cdot \mathbf{V}_i^j$ (i.e., the sum of elements in the first column $\mathbf{V}^{j+1}$) is at most

$$\alpha^j \cdot \sum_{i=1}^{N} \mathbf{v}_{i1} \leq \alpha^{j+1}.$$

For other columns, the proof is the same. Notice that $0 < \alpha < 1$, then $\lim_{j \to +\infty} \mathbf{V}^j = \mathbf{0}$ and thus $(\mathbf{I} - \mathbf{V})^{-1} = \sum_{j=0}^{+\infty} \mathbf{V}^j$.

(5) We prove it by induction. The $j = 1$ case is proved by (2). Now assume that it is true for j. Let $\mathbf{V}_{i,j}^*$ be the i-th row of $\mathbf{V}^j$. Then the first row of $\mathbf{V}^{j+1} = \mathbf{V} \cdot \mathbf{V}^j$ is

$$\sum_{i=1}^{N} \mathbf{v}_{1i} \cdot \mathbf{V}_{i,j}^*.$$

However, the sum $v_1^{(j)}$ of elements in the first row of $\mathbf{V}^j$ is at most $j\alpha^j \cdot \mu_1$ by induction hypothesis. Also, notice that $\mathbf{v}_{11} \leq \alpha$ and $\mathbf{v}_{1i} \leq \frac{\alpha}{N} \cdot \mu_1$ for $i \neq 1$, and by (4) we know the sum of all elements in $\mathbf{V}^j$ is at most $\alpha^j \cdot N$, thus the sum of elements in $\sum_{i=1}^{N} \mathbf{v}_{1i} \cdot \mathbf{V}_{i,j}^*$ (i.e., the sum $v_1^{(j+1)}$ of elements in the first row $\mathbf{V}^{j+1}$) satisfies

$$\begin{aligned}
v_1^{(j+1)} &\leq \alpha \cdot v_1^{(j)} + \alpha^j \cdot N \cdot \frac{\alpha}{N} \cdot \mu_1 \\
&= \alpha \cdot v_1^{(j)} + \alpha^{j+1} \cdot \mu_1 \\
&\leq \alpha \cdot j\alpha^j \cdot \mu_1 + \alpha^{j+1} \cdot \mu_1 \\
&= (j+1)\alpha^{j+1} \cdot \mu_1.
\end{aligned}$$

For other rows, the proof is the same.

(6) By (5) we obtain that the k-th row sum of $(\mathbf{I}-\mathbf{V})^{-1}=\sum_{j=0}^{+\infty}\mathbf{V}^j$ is upper bounded by

$$1+\sum_{j=1}^{+\infty} j\alpha^j\cdot\mu_k = 1+\frac{\alpha}{(1-\alpha)^2}\cdot\mu_k.$$

The lower bound can be directly implied by (2) since the k-th row sum of $(\mathbf{I}-\mathbf{V})^{-1}=\sum_{j=0}^{+\infty}\mathbf{V}^j$ is larger than or equal to the k-th row sum of $\mathbf{I}+\mathbf{V}$. □

Now we are ready to prove Proposition 4.2.

Proof of Proposition 4.2. Since g_i is equal to $\frac{\delta}{N}$ times the i-th row sum of $(\mathbf{I}-\mathbf{T}_1-\mathbf{T}_2)^{-1}=(\mathbf{I}-\mathbf{V})^{-1}=\sum_{j=0}^{+\infty}\mathbf{V}^j$, thus by (6) of Lemma A.2 we obtain that g_i has the following lower and upper bounds for $1\le i\le N$, i.e.,

$$\frac{\delta}{N}\left(1+\frac{\alpha}{\tau_\xi/\tau_\varepsilon+1}\cdot\mu_i\right)\le g_i\le\frac{\delta}{N}\left(1+\frac{\alpha}{(1-\alpha)^2}\cdot\mu_i\right).$$

□

A.4 Proof of Proposition 4.3 and Corollary 4.1

In this section, we first provide the following result on the lower and upper bounds for $\sum_{k=1}^N g_k^2$.

Lemma A.3. *Let $\mathbf{A}=(a_{ij})_{N\times N}$ be the attention graph where $a_{ii}=1$. Let $\mu_i=\sum_{j\in\mathcal{M}}\frac{1}{d_j}$ for $1\le i\le N$ and $c_{\xi,\varepsilon}=\frac{\tau_\xi}{\tau_\varepsilon}$. Then $\sum_{k=1}^N g_i^2$ is equal to*

$$\Theta\left(\frac{1}{N^2}\sum_{i=1}^N\mu_i^2\right).$$

More precisely, we have

$$\delta^2\cdot\left(\frac{1}{N}+\frac{\alpha^2}{(c_{\xi,\varepsilon}+1)^2}\cdot\frac{1}{N^2}\sum_{i=1}^N\mu_i^2\right)\;\le\;\sum_{i=1}^N g_i^2\;\le\;\delta^2\cdot\left(\frac{2}{N}+\frac{2\alpha^2}{(1-\alpha)^4}\cdot\frac{1}{N^2}\sum_{i=1}^N\mu_i^2\right)$$

and

$$\delta^2\cdot\left(\frac{\alpha^2}{(c_{\xi,\varepsilon}+1)^2}\cdot\frac{1}{N^2}\sum_{i=1}^N\mu_i^2\right)\;\le\;\sum_{i=1}^N g_i^2\;\le\;\delta^2\cdot\left(2+\frac{2\alpha^2}{(1-\alpha)^4}\right)\cdot\frac{1}{N^2}\sum_{i=1}^N\mu_i^2.$$

Proof. By the inequality $1+x^2\le(1+x)^2\le 2(1+x^2)$ (for $x\ge 0$) and Lemma A.2, we obtain

$$\delta^2\cdot\left(\frac{1}{N}+\frac{\alpha^2}{(c_{\xi,\varepsilon}+1)^2}\cdot\frac{1}{N^2}\sum_{i=1}^N\mu_i^2\right)\;\le\;\sum_{i=1}^N g_i^2\;\le\;\delta^2\cdot\left(\frac{2}{N}+\frac{2\alpha^2}{(1-\alpha)^4}\cdot\frac{1}{N^2}\sum_{i=1}^N\mu_i^2\right).$$

Since

$$\sum_{i=1}^N\mu_i=\sum_{i=1}^N\sum_{j\in\mathcal{M}_i}\frac{1}{d_j}=\sum_{i=1}^N\sum_{i\in\mathcal{D}_j}\frac{1}{d_j}=\sum_{i=1}^N 1=N,$$

we obtain

$$\sum_{i=1}^{N} \mu_i^2 \geq N.$$

Therefore, we have

$$\delta^2 \cdot \left(\frac{\alpha^2}{(c_{\xi,\varepsilon}+1)^2} \cdot \frac{1}{N^2} \sum_{i=1}^{N} \mu_i^2 \right) \leq \sum_{i=1}^{N} g_i^2 \leq \delta^2 \cdot \left(2 + \frac{2\alpha^2}{(1-\alpha)^4} \right) \cdot \frac{1}{N^2} \sum_{i=1}^{N} \mu_i^2$$

and thus

$$\sum_{i=1}^{N} g_i^2 = \Theta\left(\frac{1}{N^2} \sum_{i=1}^{N} \mu_i^2 \right).$$

□

Now we are ready to prove Proposition 4.3 and Corollary 4.1.

Proof of Proposition 4.3. Let $\mu = \max_i \mu_i$. Since

$$\sum_{i=1}^{N} \mu_i = \sum_{i=1}^{N} \sum_{j \in \mathcal{M}_i} \frac{1}{d_j} = \sum_{i=1}^{N} \sum_{i \in \mathcal{D}_j} \frac{1}{d_j} = \sum_{i=1}^{N} 1 = N,$$

we know $\sum_{i=1}^{N} \mu_i^2$ is maximal if there are $\frac{N}{\mu}$ number of μ_i equals μ and $N - \frac{N}{\mu}$ number of μ_i equals 0. Therefore,

$$\mu^2 = \max_i \mu_i^2 \leq \sum_{i=1}^{N} \mu_i^2 \leq \mu^2 \cdot \frac{N}{\mu} = \mu N.$$

Thus micro shocks can generate aggregate fluctuations if and only if $\sum_{i=1}^{N} g_i^2 = \Theta(1)$, by Lemma A.3 this is equivalent to $\sum_{i=1}^{N} \mu_i^2 = \Theta(N^2)$. Then by the above inequality it is equivalent to $\mu = \max_i \mu_i = \Theta(N)$. □

Proof of Corollary 4.1. First, we have

$$\sum_{i=1}^{N} \mu_i = \sum_{i=1}^{N} \sum_{j \in \mathcal{M}_i} \frac{1}{d_j} = \sum_{i=1}^{N} \sum_{i \in \mathcal{D}_j} \frac{1}{d_j} = \sum_{i=1}^{N} 1 = N.$$

Therefore, for any constant $c > 0$, we have the number of μ_i larger than cN is at most $\frac{1}{c} = \Theta(1)$. Also, the i-th firm can have granular effects if and only if $g_i = \Theta(1)$, by Proposition 4.2 this is equivalent to $\mu_i = \Theta(N)$, which means that $\mu_i > cN$ for some constant $c > 0$. Thus there are at most $K = \Theta(1)$ firms that can have granular effects.

By Proposition 4.3, we know the micro shocks can generate aggregate fluctuations if and only if

$\max_i \mu_i = \Theta(N)$. However,

$$\mu_i = \sum_{j \in \mathcal{M}_i} \frac{1}{d_j} \leq \sum_{j \in \mathcal{M}_i} 1 = m_i.$$

Thus $\max_i \mu_i = \Theta(N)$ implies $\max_i m_i = \Theta(N)$.

□

A.5 Proof of Proposition 5.2

Proof. In Lemma A.4(6), we will show that g_k has the following lower and upper bounds for $1 \leq k \leq N$, i.e.,

$$\delta\left(\omega_k + \alpha\omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j \notin \mathcal{D}_i} \omega_j\right) \leq g_k \leq \delta\left(\omega_k + \frac{\alpha}{(1-\alpha)^2} \cdot \left(\omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i}\right)\right). \tag{A.16}$$

(1) $\Longleftarrow$: Assume that $\mu_k^\omega = \omega_k + \sum_{j \in \mathcal{M}_k} \frac{\omega_j}{d_j} = \max_i \mu_i^\omega = \Theta(1)$. Let $w = \max_i \omega_i$ be the maximal firm size. Then either (i) $\omega_k = w = \Theta(1)$ for some $1 \leq k \leq N$; or (ii) $w = o(1)$ and $\sum_{j \in \mathcal{M}_k} \frac{\omega_j}{d_j} = \Theta(1)$ for some $1 \leq k \leq N$. Thus, we prove the claim for the two cases.

(i) If $\omega_k = \Theta(1)$, then by Eq. (A.16) we obtain

$$g_k \geq \delta\omega_k = \Theta(1).$$

Thus

$$\sum_{i=1}^N g_i^2 \geq g_k^2 = \Theta(1).$$

(ii) If $w = \max_i \omega_i = o(1)$ and $\sum_{j \in \mathcal{M}_k} \frac{\omega_j}{d_j} = \Theta(1)$, then by Eq. (A.16) we obtain

$$\begin{aligned}
g_k &\geq \delta\left(\omega_k + \alpha\omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j \notin \mathcal{D}_i} \omega_j\right) \\
&= \delta\left(\omega_k + \alpha\omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} - \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j \in \mathcal{D}_i} \omega_j\right) \\
&\geq \delta\left(\frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} - \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j \in \mathcal{D}_i} \omega_j\right) \\
&\geq \delta\left(\frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} - \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j \in \mathcal{D}_i} \omega\right) \\
&= \delta\left(\frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} - \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \cdot d_i\omega\right)
\end{aligned}$$

$$= \delta\left(\frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1}\sum_{i\in\mathcal{M}_k}\frac{\omega_i}{d_i} - \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1}\cdot\omega\sum_{i\in\mathcal{M}_k}\omega_i\right)$$

$$\geq \delta\left(\frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1}\sum_{i\in\mathcal{M}_k}\frac{\omega_i}{d_i} - \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1}\cdot\omega\right).$$

Since $w = o(1)$ and $\sum_{j\in\mathcal{M}_k}\frac{\omega_j}{d_j} = \Theta(1)$, by the above inequality we derive

$$g_k = \Theta(1).$$

Thus

$$\sum_{i=1}^N g_i^2 \geq g_k^2 = \Theta(1).$$

(2) $\Longrightarrow$: Assume $\sum_{i=1}^N g_i^2 = \Theta(1)$.
by Eq. (A.16) we obtain

$$\sum_{i=1}^N g_i^2 = O\left(\sum_{i=1}^N\left(\omega_i + \sum_{j\in\mathcal{M}_i}\frac{\omega_j}{d_j}\right)^2\right) = O\left(\sum_{i=1}^N(\mu_i)^2\right). \tag{A.17}$$

Also, we have

$$\begin{aligned}\sum_{i=1}^N \mu_i^\omega &= \sum_{i=1}^N \omega_i + \sum_{i=1}^N\sum_{j\in\mathcal{M}_i}\frac{\omega_j}{d_j}\\ &= 1 + \sum_{j=1}^N\frac{\omega_j}{d_j}\sum_{i\in\mathcal{D}_j}1\\ &= 1 + \sum_{j=1}^N\frac{\omega_j}{d_j}\cdot d_j\\ &= 1 + \sum_{j=1}^N\omega_j = 1 + 1 = 2.\end{aligned}$$

Let $\mu^\omega = \max_i \mu_i^\omega$. Therefore, we know $\sum_{i=1}^N(\mu_i^\omega)^2$ is maximal if there are $\frac{2}{\mu^\omega}$ number of μ_i^ω equals μ^ω and $N - \frac{2}{\mu^\omega}$ number of μ_i^ω equals 0. Therefore,

$$\sum_{i=1}^N(\mu_i^\omega)^2 \leq \frac{2}{\mu^\omega}\cdot(\mu^\omega)^2 = 2\mu^\omega.$$

Therefore, by Eq. (A.17) we have

$$\mu^{\omega} \geq \frac{1}{2}\sum_{i=1}^{N}(\mu_i^{\omega})^2 \geq \Theta\left(\sum_{i=1}^{N} g_i^2\right) = \Theta(1).$$

which means that

$$\max_i \mu_i^{\omega} = \mu^{\omega} = \Theta(1)$$

The proof is now complete. □

Remark. When $\omega_i = \frac{1}{N}$ for $1 \leq i \leq N$, we have $\max_i \mu_i^{\omega} = \Theta(1)$ if and only if $\max_i \left(\frac{1}{N} + \frac{1}{N}\sum_{j\in\mathcal{M}_i}\frac{1}{d_j}\right) = \Theta(1)$, which is equivalent to Proposition 4.3,

$$\max_i \mu_i = \max_i \sum_{j\in\mathcal{M}_i}\frac{1}{d_j} = \Theta(N).$$

A.6 Proof of Lemma 5.3

Proof. Since $x_{(1)} > x_{(2)} > \cdots > x_{(N)}$ are drawn from the Pareto distribution

$$\mathbb{P}(X > x) = \frac{c^{\kappa}}{N}\cdot x^{-\kappa},$$

it follows that

$$\mathbb{P}(\frac{c^{\kappa}}{N}\cdot X^{-\kappa} < y) = \mathbb{P}(X > \left(\frac{Ny}{c^{\kappa}}\right)^{-1/\kappa}) = \frac{c^{\kappa}}{N}\cdot\left(\left(\frac{Ny}{c^{\kappa}}\right)^{-1/\kappa}\right)^{-\kappa} = y.$$

Therefore, $\frac{c^{\kappa}}{N}\cdot X^{-\kappa}$ follows the uniform law. Given that $x_{(1)} > x_{(2)} > \cdots > x_{(N)}$, we use basic probability theory (Pitman (2012), Pages 327–328) to deduce that

$$\mathbb{E}\left(\frac{c^{\kappa}}{N}\cdot x_{(k)}^{-\kappa}\right) = \frac{k}{N+1}.$$

Now take logs on both sides of the equation,

$$\log\mathbb{E}\left(\frac{c^{\kappa}}{N}\cdot x_{(k)}^{-\kappa}\right) = \log\frac{k}{N+1}.$$

Up to a first-order approximation, we omit higher-order terms in Jessen's inequality,

$$\log\mathbb{E}\left[x_{(k)}\right] = \log\left[\left(\frac{N+1}{N}\right)\frac{c^{\kappa}}{k}\right]^{\frac{1}{\kappa}}$$

Now let $N \to \infty$, it follows that up to first-order approximation we have

$$\mathbb{E}\left[x_{(k)}\right] = ck^{-\frac{1}{\kappa}}$$

□

A.7 Proof of Lemma 5.2

Proof. First, for $1 \le j \le N$, by (5.5) we have

$$\begin{aligned} m_j &= \sum_{i=1}^{N} \mathbb{E}\left[A_{ij}\right] \\ &= p_{jj} + \sum_{i=1}^{j-1} p_{ij} + \sum_{i=j+1}^{N} p_{ij} \\ &= 1 + (N-1) \cdot \frac{N\, j^{-\beta_m^{-1}} - 1}{N-1} \\ &= 1 + N\, j^{-\beta_m^{-1}} - 1 \\ &= \frac{N}{j^{\frac{1}{\beta_m}}}. \end{aligned}$$

Similarly, for $1 \le i \le N$, we have

$$\begin{aligned} d_i &= \sum_{j=1}^{N} \mathbb{E}\left[A_{ij}\right] \\ &= p_{ii} + \sum_{j=1}^{i-1} p_{ij} + \sum_{j=i+1}^{N} p_{ij} \\ &= 1 - \frac{N\, i^{-\beta_m^{-1}} - 1}{N-1} + \sum_{j=1}^{N} \frac{N\, j^{-\beta_m^{-1}} - 1}{N-1} \\ &= \frac{N}{N-1}\left[\zeta(\beta_m^{-1}, N) - i^{-\beta_m^{-1}}\right]. \end{aligned}$$

where $\zeta\left(\beta_m^{-1}, N\right)$ is a finite-order approximation of the convergent Riemann-Zeta function $\zeta(\beta_m^{-1})$ for sufficiently large N, given by[22]

$$\zeta\left(\beta_m^{-1}, N\right) \approx \begin{cases} \frac{N^{1-\beta_m^{-1}}}{1-\frac{1}{\beta_m}}, & \beta_m > 1; \\ \ln N, & \beta_m = 1 \\ \zeta(\beta_m^{-1}), & \beta_m \in [0,1). \end{cases} \tag{A.18}$$

[22]We define $f(N) \approx g(N)$ if $\lim_{N\to+\infty} \frac{f(N)}{g(N)} = 1$.

Finally, (A.18) is a well-known result and can be found from Page 77 in Titchmarsh and Heath-Brown (1986) and Page 429 in Graham et al. (1994). □

A.8 Proof of Proposition 5.3

By Lemma A.1 we derive the following result.

Lemma A.4. *Let*

$$\mathbf{V} = (\mathbf{v}_{ij})_{N\times N} = \alpha\, \mathrm{diag}(\mathbf{A}'\boldsymbol{\omega}) + \alpha \mathbf{A}'\, \mathrm{diag}\,(\boldsymbol{\omega})\, \mathrm{diag}\,(\boldsymbol{\lambda})\,(\mathbf{1} - \mathbf{A})\,.$$

Then we have:

(1) Every column sum of $\mathbf{V}$ *is at most* $\alpha < 1$*, and thus* $\mathbf{I} - \mathbf{V}$ *is an M-matrix.*

(2) Let $k \geq 1$ *and* $\mu'_k = \sum_{i\in\mathcal{M}_k} \frac{N\omega_i}{d_i}$*. Then the k-th row sum* v'_k *of* $\mathbf{V}$ *satisfies*

$$\frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1}\cdot\mu'_k = \frac{\alpha N}{\tau_\xi/\tau_\varepsilon + 1}\cdot\sum_{i\in\mathcal{M}_k}\frac{\omega_i}{d_i} \leq v'_k \leq \alpha N\cdot\sum_{i\in\mathcal{M}_k}\frac{\omega_i}{d_i} = \alpha\cdot\mu'_k.$$

Let $\mu''_k = \omega_k\cdot\sum_{i\in\mathcal{M}_k}\omega_i + \sum_{i\in\mathcal{M}_k}\frac{\omega_i}{d_i}$*. Then the k-th weight row sum* $v_k = \sum_{i=1}^N \omega_i \mathbf{v}_{ki}$ *of* $\mathbf{V}$ *is equal to*

$$\Theta\left(\omega_k\cdot\sum_{i\in\mathcal{M}_k}\omega_i + \sum_{i\in\mathcal{M}_k}\frac{\omega_i}{d_i}\sum_{j\notin\mathcal{D}_i}\omega_j\right).$$

Moreover,

$$\alpha\omega_k\cdot\sum_{i\in\mathcal{M}_k}\omega_i + \frac{\alpha}{\tau_\xi/\tau_\varepsilon+1}\sum_{i\in\mathcal{M}_k}\frac{\omega_i}{d_i}\sum_{j\notin\mathcal{D}_i}\omega_j \leq v_k \leq \alpha\omega_k\cdot\sum_{i\in\mathcal{M}_k}\omega_i + \alpha\sum_{i\in\mathcal{M}_k}\frac{\omega_i}{d_i}\sum_{j\notin\mathcal{D}_i}\omega_j \leq \alpha\mu''_k.$$

(3) For the k-th row of $\mathbf{V}$*, every non-diagonal element is upper bounded by* $\frac{\alpha}{N}\cdot\mu'_k$*, thus is smaller than or equal to the diagonal element* $\alpha\cdot\mu''_k$*.*

(4) Let $j \geq 1$*. Then every column sum of* $\mathbf{V}^j$ *(the j-th power of* $\mathbf{V}$*) is at most* α^j*. Then* $\lim_{j\to+\infty}\mathbf{V}^j = \mathbf{0}$ *and thus* $(\mathbf{I}-\mathbf{V})^{-1} = \sum_{j=0}^{+\infty}\mathbf{V}^j$*.*

(5) Let $j \geq 1$ *and* $k \geq 1$*. Then the k-th weight row sum* $v_k^{(j)} = \sum_{i=1}^N \omega_i \mathbf{V}^j_{ki}$ *of* $\mathbf{V}^j$ *is upper bounded by*

$$v_k^{(j)} \leq j\alpha^j\cdot\mu''_k.$$

(6) Let $k \geq 1$*. Then the k-th element in* $(\mathbf{I}-\mathbf{V})^{-1}\boldsymbol{\omega} = \sum_{j=0}^{+\infty}\mathbf{V}^j\boldsymbol{\omega}$ *is upper bounded by*

$$\omega_k + \frac{\alpha}{(1-\alpha)^2}\cdot\mu''_k = \omega_k + \frac{\alpha}{(1-\alpha)^2}\cdot\left(\omega_k\cdot\sum_{i\in\mathcal{M}_k}\omega_i + \sum_{i\in\mathcal{M}_k}\frac{\omega_i}{d_i}\right),$$

and lower bounded by

$$\omega_k + \alpha\omega_k \cdot \sum_{i\in\mathcal{M}_k} \omega_i + \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i\in\mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j\notin\mathcal{D}_i} \omega_j.$$

Thus g_k has the following lower and upper bounds for $1 \le k \le N$, i.e.,

$$\delta\left(\omega_k + \alpha\omega_k \cdot \sum_{i\in\mathcal{M}_k} \omega_i + \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i\in\mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j\notin\mathcal{D}_i} \omega_j\right) \le g_k \le \delta\left(\omega_k + \frac{\alpha}{(1-\alpha)^2} \cdot \left(\omega_k \cdot \sum_{i\in\mathcal{M}_k} \omega_i + \sum_{i\in\mathcal{M}_k} \frac{\omega_i}{d_i}\right)\right).$$

Proof. (1) Let's consider the j-th column in $\mathbf{V}$ and show that its j-th column sum is at most $\alpha < 1$. First, by Lemma A.1, the j-th diagonal element of $\mathbf{V}$ is the j-th diagonal element of $\alpha\,\mathrm{diag}(\mathbf{A}'\boldsymbol{\omega})$, thus is equal to

$$\alpha \cdot \sum_{i\in\mathcal{M}_j} \omega_i.$$

For the non-diagonal element j-th column sum of $\mathbf{V}$, by Lemma A.1, it only involves

$$\alpha\omega_i\lambda_i\mathbf{A}_i(\mathbf{1} - \mathbf{A}_i')$$

where $A_{ij} = 0$, thus $j \notin \mathcal{D}_i$ and $i \notin \mathcal{M}_j$. Then the non-diagonal element j-th column sum of $\mathbf{V}$ is equal to

$$\alpha \cdot \sum_{i\notin\mathcal{M}_j} \sum_{k\in\mathcal{D}_i} \omega_i\lambda_i. \tag{A.19}$$

Add the above diagonal and non-diagonal elements together, we obtain that the j-th column sum of $\mathbf{V}$ is equal to

$$\begin{aligned}
&\alpha \cdot \sum_{i\in\mathcal{M}_j} \omega_i + \alpha \cdot \sum_{i\notin\mathcal{M}_j} \sum_{k\in\mathcal{D}_i} \omega_i\lambda_i \\
&= \alpha \cdot \sum_{i\in\mathcal{M}_j} \omega_i + \alpha \cdot \sum_{i\notin\mathcal{M}_j} d_i\omega_i\lambda_i \\
&\le \alpha \cdot \sum_{i\in\mathcal{M}_j} \omega_i + \alpha \cdot \sum_{i\notin\mathcal{M}_j} \omega_i \\
&= \alpha \cdot \sum_{i=1}^{N} \omega_i \\
&= \alpha.
\end{aligned}$$

Then by Theorem 1, we know that the spectral radius of $\mathbf{V}$ is at most α, thus by Definition 3 we obtain that $\mathbf{I} - \mathbf{V}$ is an M-matrix.

(2) First, by Lemma A.1, the k-th diagonal element of $\mathbf{V}$ is equal to

$$\alpha \cdot \sum_{i \in \mathcal{M}_k} \omega_i.$$

For the non-diagonal element k-th row sum of $\mathbf{V}$, by Lemma A.1, it only involves

$$\alpha \omega_i \lambda_i \mathbf{A}_i (\mathbf{1} - \mathbf{A}_i')$$

where $A_{ik} = 1$, thus $k \in \mathcal{D}_i$ and $i \in \mathcal{M}_k$. Thus the non-diagonal element k-th row sum of $\mathbf{V}$ is equal to

$$\alpha \cdot \sum_{i \in \mathcal{M}_k} \sum_{j \notin \mathcal{D}_i} \omega_i \lambda_i.$$

Add the above diagonal and non-diagonal elements together, we obtain that the k-th row sum v_k' of $\mathbf{V}$ is equal to

$$\begin{aligned}
&\alpha \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \alpha \cdot \sum_{i \in \mathcal{M}_k} \sum_{j \notin \mathcal{D}_i} \omega_i \lambda_i \\
&= \alpha \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \alpha \cdot \sum_{i \in \mathcal{M}_k} \omega_i \lambda_i (N - d_i) \\
&= \alpha \cdot \sum_{i \in \mathcal{M}_k} \omega_i \left(1 - \lambda_i d_i + N \lambda_i\right)) \\
&= \alpha \cdot \sum_{i \in \mathcal{M}_k} \omega_i \cdot \frac{\tau_\xi + N \tau_\varepsilon}{\tau_\xi + d_i \tau_\varepsilon}.
\end{aligned}$$

Also, we have

$$\frac{1}{\tau_\xi / \tau_\varepsilon + 1} \cdot \frac{N}{d_i} \leq \frac{\tau_\xi + N \tau_\varepsilon}{\tau_\xi + d_i \tau_\varepsilon} \leq \frac{N}{d_i}.$$

Therefore,

$$\frac{\alpha N}{\tau_\xi / \tau_\varepsilon + 1} \cdot \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \leq v_k' \leq \alpha N \cdot \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i}.$$

Furthermore, we have

$$v_k = \sum_{i=1}^{N} \omega_i \mathbf{v}_{ki} = \alpha \omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \alpha \cdot \sum_{i \in \mathcal{M}_k} \sum_{j \notin \mathcal{D}_i} \omega_i \lambda_i \omega_j. \tag{A.20}$$

Therefore,

$$\alpha \omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \frac{\alpha}{\tau_\xi / \tau_\varepsilon + 1} \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j \notin \mathcal{D}_i} \omega_j \leq v_k \leq \alpha \omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \alpha \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j \notin \mathcal{D}_i} \omega_j \leq \alpha \mu_k''.$$

(3) Similarly to (2), the non-diagonal elements $\mathbf{v}_{kj}(1 \le j \le N,\ \ j \ne k)$ in the k-th row of $\mathbf{V}$ satisfies

$$\mathbf{v}_{kj} = \alpha \cdot \sum_{i \in \mathcal{M}_k \text{ and } i \notin \mathcal{M}_j} \omega_i \lambda_i \le \alpha \cdot \sum_{i \in \mathcal{M}_k} \omega_i \lambda_i \le \alpha \cdot \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} = \frac{\alpha}{N} \cdot \mu'_k \le \alpha \mu''_k. \tag{A.21}$$

(4) We prove it by induction. The case $j = 1$ is proved by (1). Now assume that it is true for j. Let $\mathbf{V}^j = (\mathbf{V}^j_1\ \ \mathbf{V}^j_2\ \ \dots\ \ \mathbf{V}^j_N)$ where $\mathbf{V}^j_i$ is the i-th column of $\mathbf{V}^j$. Then the first column of $\mathbf{V}^{j+1} = \mathbf{V}^j \cdot \mathbf{V}$ is

$$\sum_{i=1}^{N} \mathbf{v}_{i1} \cdot \mathbf{V}^j_i.$$

However, the sum of elements in the i-th column $\mathbf{V}^j_i$ in $\mathbf{V}^j$ is at most α^j by induction hypothesis, thus the sum of elements in $\sum_{i=1}^{N} \mathbf{v}_{i1} \cdot \mathbf{V}^j_i$ (i.e., the sum of elements in the first column $\mathbf{V}^{j+1}$) is at most

$$\alpha^j \cdot \sum_{i=1}^{N} \mathbf{v}_{i1} \le \alpha^{j+1}.$$

For other columns, the proof is the same. Notice that $0 < \alpha < 1$, then $\lim_{j \to +\infty} \mathbf{V}^j = \mathbf{0}$ and thus $(\mathbf{I} - \mathbf{V})^{-1} = \sum_{j=0}^{+\infty} \mathbf{V}^j$.

(5) We prove it by induction. When $j = 1$, by (2) we have

$$v_k^{(1)} \le \alpha \mu''_k.$$

Thus (5) holds for $j = 1$. Now assume that it is true for j. Let $\mathbf{V}^*_{i,j}$ be the i-th row of $\mathbf{V}^j$. Then the k-th element $v_k^{(j+1)}$ of $\mathbf{V}^{j+1}\boldsymbol{\omega} = \mathbf{V} \cdot \mathbf{V}^j \boldsymbol{\omega}$ is equal to

$$v_k^{(j+1)} = \sum_{i=1}^{N} \mathbf{v}_{ki} \cdot \mathbf{V}^*_{i,j} \boldsymbol{\omega}.$$

However, the sum $v_k^{(j)}$ is at most $j\alpha^j \cdot \mu'_k$ by induction hypothesis. Also, notice that $\mathbf{v}_{kk} \le \alpha$ and $\mathbf{v}_{ki} \le \alpha \mu''_k$ for $i \ne k$, and by (4) we know the sum of all elements in $\mathbf{V}^j \boldsymbol{\omega}$ is at most $\alpha^j \cdot \sum_{i=1}^{N} \omega_i \le \alpha^j$, thus

$$\begin{aligned} v_k^{(j+1)} &\le \alpha \cdot v_k^{(j)} + \alpha^j \cdot \alpha \mu''_k \\ &= \alpha \cdot v_k^{(j)} + \alpha^{j+1} \cdot \mu''_k \\ &\le \alpha \cdot j\alpha^j \cdot \mu''_k + \alpha^{j+1} \cdot \mu''_k \\ &= (j+1)\alpha^{j+1} \cdot \mu''_k, \end{aligned}$$

which completes the proof.

(6) The upper bound is obtained directly by (5). The lower bound can be directly implied by (2)

since every element in $(\mathbf{I}-\mathbf{V})^{-1} = \sum_{j=0}^{+\infty} \mathbf{V}^j$ is larger than or equal to the corresponding element in $\mathbf{I}+\mathbf{V}$. □

Also, the following lemma is easy to prove.

Lemma A.5. *For $x \geq 0$ and $y \geq 0$, we always have*

$$x^2 + y^2 \leq (x+y)^2 \leq 2(x^2+y^2).$$

Now we are ready to prove Proposition 5.3.

Proof of Proposition 5.3. Under the random graph formulation

$$\mathbb{P}\left(A_{ij}=1\right) = p_{ij} = \begin{cases} \frac{\frac{N}{j^{1/\beta_m}}-1}{N-1} \in [0,1]; & \forall j \neq i \\ 1; \quad j = i \end{cases}$$

we obtain

$$\mathbb{E}[m_j] = \sum_{i=1}^{N} \mathbb{E}\left[A_{ij}\right] = \frac{N}{j^{1/\beta_m}}; \qquad j = 1, 2 \ldots N$$

and

$$\mathbb{E}[d_i] = \sum_{j=1}^{N} \mathbb{E}\left[A_{ij}\right] = \frac{N}{N-1}\left[\sum_{j=1}^{N} \frac{1}{j^{1/\beta_m}} - \frac{1}{i^{1/\beta_m}}\right] = \frac{N}{N-1}\left[\zeta(1/\beta_m, N) - \frac{1}{i^{1/\beta_m}}\right],$$

where the finite-order component of the Riemann-Zeta function can be approximated as

$$\zeta(1/\beta_m, N) \approx \begin{cases} \frac{N^{1-1/\beta_m}}{1-1/\beta_m}; & 0 \leq 1/\beta_m < 1; \\ \ln N; & 1/\beta_m = 1 \\ \zeta(1/\beta_m); & 1/\beta_m > 1 \end{cases}$$

for sufficiently large N, and ζ is the convergent Riemann-Zeta constant.

We assume that the size of the firms $\{\omega_{\iota(1)}, \ldots, \omega_{\iota(N)}\}$ in each economy are defined as a sequence

$$\omega_j = \frac{\iota(j)^{-\frac{1}{\beta_\omega}}}{\sum_{k=1}^{N} k^{-\frac{1}{\beta_\omega}}} = \frac{1}{\zeta\left(\frac{1}{\beta_\omega}, N\right)} \iota(j)^{-\frac{1}{\beta_\omega}} \in [0,1], \qquad \text{such that} \quad \sum_{i=1}^{N} \omega_i = 1, \tag{A.22}$$

where firm i in the information network is ranked $\iota(i)$ in terms of firm size.

Therefore, when $\beta_m > 0$, for each fixed $1 \leq i \leq N$, we always have

$$\mathbb{E}[\sum_{j \notin \mathcal{D}_i} \omega_j] = \frac{N-d_i}{N} \approx 1.$$

Then, by Lemma A.4 we obtain that

$$g_k = \Theta\left(\omega_k + \sum_{i\in\mathcal{M}_k} \omega_i \cdot \left(\omega_k + \frac{1}{d_i}\right)\right) = \Theta\left(\omega_k + \sum_{i\in\mathcal{M}_k} \frac{\omega_i}{d_i}\right).$$

Furthermore, by Lemma A.5 we derive

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta\left(\mathbb{E}[\sum_{k=1}^{N} \omega_k^2 + \sum_{k=1}^{N}\left(\sum_{i\in\mathcal{M}_k} \frac{\omega_i}{d_i}\right)^2]\right).$$

Then we divide the proof into the following cases.

(1) $\beta = \min\{\beta_m, \beta_\omega\} > 1$**.**

Assume $\beta = \min\{\beta_m, \beta_\omega\} > 1$. Under the random graph formulation, we have $\mathbb{E}[m_k] = \frac{N}{k^{1/\beta_m}}$ $(1 \leq k \leq N)$ follows the power law, while $\mathbb{E}[d_i] \approx \frac{N^{1-1/\beta_m}}{1-1/\beta_m}$ $(1 \leq i \leq N)$.

Then

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta\left(\sum_{k=1}^{N}\left(\frac{1-1/\beta_\omega}{k^{1/\beta_\omega} \cdot N^{1-1/\beta_\omega}}\right)^2 + \sum_{k=1}^{N}\left(\frac{1-1/\beta_m}{N^{1-1/\beta_m}} \cdot \frac{1}{k^{1/\beta_m}}\right)^2\right).$$

We divide the discussion into the following cases.

Case 1: $\beta_m \geq \beta_\omega > 2$.

Then

$$\begin{aligned}
\mathbb{E}[\sum_{k=1}^{N} g_k^2] &= \Theta\left(\sum_{k=1}^{N}\left(\frac{1-1/\beta_\omega}{k^{1/\beta_\omega} \cdot N^{1-1/\beta_\omega}}\right)^2 + \sum_{k=1}^{N}\left(\frac{1-1/\beta_m}{N^{1-1/\beta_m}} \cdot \frac{1}{k^{1/\beta_m}}\right)^2\right) \\
&= \Theta\left(\frac{1}{N^{2-2/\beta_\omega}} \sum_{k=1}^{N} \frac{1}{k^{2/\beta_\omega}}\right) \\
&= \Theta\left(\frac{N^{1-2/\beta_\omega}}{N^{2-2/\beta_\omega}}\right) \\
&= \Theta\left(\frac{1}{N}\right).
\end{aligned}$$

Case 2: $\beta_\omega \geq \beta_m > 2$.

Then

$$\begin{aligned}
\mathbb{E}[\sum_{k=1}^{N} g_k^2] &= \Theta\left(\sum_{k=1}^{N}\left(\frac{1-1/\beta_\omega}{k^{1/\beta_\omega} \cdot N^{1-1/\beta_\omega}}\right)^2 + \sum_{k=1}^{N}\left(\frac{1-1/\beta_m}{N^{1-1/\beta_m}} \cdot \frac{1}{k^{1/\beta_m}}\right)^2\right) \\
&= \Theta\left(\frac{1}{N^{2-2/\beta_m}} \sum_{k=1}^{N} \frac{1}{k^{2/\beta_m}}\right)
\end{aligned}$$

$$= \Theta\left(\frac{N^{1-2/\beta_m}}{N^{2-2/\beta_m}}\right)$$
$$= \Theta\left(\frac{1}{N}\right).$$

Case 3: $\beta_m \geq \beta_\omega = 2$.
Then

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta\left(\sum_{k=1}^{N}\left(\frac{1-1/\beta_\omega}{k^{1/\beta_\omega} \cdot N^{1-1/\beta_\omega}}\right)^2 + \sum_{k=1}^{N}\left(\frac{1-1/\beta_m}{N^{1-1/\beta_m}} \cdot \frac{1}{k^{1/\beta_m}}\right)^2\right)$$
$$= \Theta\left(\frac{1}{N}\sum_{k=1}^{N}\frac{1}{k}\right)$$
$$= \Theta\left(\frac{\ln N}{N}\right).$$

Case 4: $\beta_\omega \geq \beta_m = 2$.
Then

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta\left(\sum_{k=1}^{N}\left(\frac{1-1/\beta_\omega}{k^{1/\beta_\omega} \cdot N^{1-1/\beta_\omega}}\right)^2 + \sum_{k=1}^{N}\left(\frac{1-1/\beta_m}{N^{1-1/\beta_m}} \cdot \frac{1}{k^{1/\beta_m}}\right)^2\right)$$
$$= \Theta\left(\frac{1}{N}\sum_{k=1}^{N}\frac{1}{k}\right)$$
$$= \Theta\left(\frac{\ln N}{N}\right).$$

Case 5: $1 < \beta_\omega < 2$ and $\beta_m \geq \beta_\omega$.
Then

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta\left(\sum_{k=1}^{N}\left(\frac{1-1/\beta_\omega}{k^{1/\beta_\omega} \cdot N^{1-1/\beta_\omega}}\right)^2 + \sum_{k=1}^{N}\left(\frac{1-1/\beta_m}{N^{1-1/\beta_m}} \cdot \frac{1}{k^{1/\beta_m}}\right)^2\right)$$
$$= \Theta\left(\frac{1}{N^{2-2/\beta_\omega}}\sum_{k=1}^{N}\frac{1}{k^{2/\beta_\omega}}\right)$$
$$= \Theta\left(\frac{1}{N^{2-2/\beta_\omega}}\right).$$

Case 6: $1 < \beta_m < 2$ and $\beta_\omega > \beta_m$.
Then

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta\left(\sum_{k=1}^{N}\left(\frac{1-1/\beta_\omega}{k^{1/\beta_\omega} \cdot N^{1-1/\beta_\omega}}\right)^2 + \sum_{k=1}^{N}\left(\frac{1-1/\beta_m}{N^{1-1/\beta_m}} \cdot \frac{1}{k^{1/\beta_m}}\right)^2\right)$$

$$= \Theta\left(\frac{1}{N^{2-2/\beta_m}} \sum_{k=1}^{N} \frac{1}{k^{2/\beta_m}}\right)$$
$$= \Theta\left(\frac{1}{N^{2-2/\beta_m}}\right).$$

(2) $\beta_m \geq \beta_\omega = 1$**.**

Assume $\beta_m \geq \beta_\omega = 1$. Under the random graph formulation, we have $\mathbb{E}[m_k] = \frac{N}{k^{1/\beta_m}}$ $(1 \leq k \leq N)$ follows the power law, while $\sum_{k=1}^{N} \omega_k^2 = \sum_{k=1}^{N} \frac{1}{k^2 \cdot (\ln N)^2}$.

Case 7: $\beta_m \geq \beta_\omega = 1$**.** Then

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta\left(\frac{1}{(\ln N)^2} \sum_{k=1}^{N} \frac{1}{k^2}\right)$$
$$= \Theta\left(\frac{1}{(\ln N)^2}\right).$$

(3) $\beta_\omega \geq \beta_m = 1$**.**

Assume $\beta_\omega \geq \beta_m = 1$. Under the random graph formulation, we have $\mathbb{E}[m_k] = \frac{N}{k}$ $(1 \leq k \leq N)$ follows the power law, while $\mathbb{E}[d_i] \approx \ln N$ $(1 \leq i \leq N)$.

Case 8: $\beta_\omega \geq \beta_m = 1$**.** Then

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta\left(\frac{1}{(\ln N)^2} \sum_{k=1}^{N} \frac{1}{k^2}\right)$$
$$= \Theta\left(\frac{1}{(\ln N)^2}\right).$$

(4) $\beta_\omega < 1$**.**

Assume $\beta_\omega < 1$. Under the random graph formulation, we have $\mathbb{E}[m_k] = \frac{N}{k^{1/\beta_m}}$ $(1 \leq k \leq N)$ follows the power law, while $\sum_{k=1}^{N} \omega_k^2 = \sum_{k=1}^{N} \frac{1}{k^{2/\beta_\omega}}$.

Case 9: $\beta_\omega < 1$**.** Thus

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta\left(\sum_{k=1}^{N} \frac{1}{k^{2/\beta_\omega}}\right) = \Theta(1).$$

(5) $\beta_m < 1$**.**

Case 10: $\beta_m < 1$**.** Similarly, we have

$$\mathbb{E}[\sum_{k=1}^{N} g_k^2] = \Theta(1).$$

The above discussion completes the proof of Proposition 5.3.

□

A.9 Proof of Proposition 5.4

Proof. We break the proof of this proposition in two parts.

Step 1: Firstly, we have $m_1 > m_2 > \cdots > m_N$. We will show that when $\omega_k > \omega_j$ and $k < j$, then interchange ω_k and ω_j will decrease the GDP volatility.

By (A.20), the k-th weight row sum of $\mathbf{V}$ equals

$$v_k = \sum_{i=1}^{N} \omega_i \mathbf{v}_{ki} = \alpha\omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \alpha \cdot \sum_{i \in \mathcal{M}_k} \sum_{j \notin \mathcal{D}_i} \omega_i \lambda_i \omega_j.$$

Since $d_i \approx d = \Theta(N^a)$ for some $0 < a < 1$ and each $1 \le i \le N$, we have

$$\lambda_i = \frac{1}{d_i + \tau_\xi/\tau_\varepsilon} \approx \frac{1}{d}.$$

Therefore,

$$v_k \approx \alpha\omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \alpha \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d} \sum_{j \notin \mathcal{D}_i} \omega_j.$$

Now by Proposition 4.2 and Lemma A.4, $(\mathbf{I}-\mathbf{V})^{-1} = \mathbf{I}+\mathbf{V}+O\left(\mathbf{V}^2\right)$. Up to the first-order approximation, we drop higher-order terms. Then,

$$g_k = \delta(\omega_k + v_k) \approx \delta\omega_k + \alpha\delta\omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \alpha\delta \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d} \sum_{j \notin \mathcal{D}_i} \omega_j.$$

Also, since $d_i \approx d = o(N)$, we obtain that $\sum_{j \notin \mathcal{D}_i} \omega_j \approx 1$. Then we have

$$\begin{aligned} g_k &\approx \delta\omega_k + \alpha\delta\omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \alpha\delta \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} \\ &\approx \delta\omega_k + \alpha\delta\omega_k \cdot \frac{m_k}{N} + \frac{\alpha\delta}{d} \cdot \frac{m_k}{N}, \end{aligned}$$

and

$$g_j \approx \delta\omega_j + \alpha\delta\omega_j \cdot \frac{m_j}{N} + \frac{\alpha\delta}{d} \cdot \frac{m_j}{N}.$$

If we change the order of ω_k and ω_j, (i.e., let $\iota(j) = k$ and $\iota(k) = j$) and keep the order of m_k and m_j, we obtain

$$g'_{\iota(j)} = {g'}_k \approx \delta\omega_j + \alpha\delta\omega_j \cdot \frac{m_k}{N} + \frac{\alpha\delta}{d} \cdot \frac{m_k}{N},$$

and

$$g'_{\iota(k)} = {g'}_j \approx \delta\omega_k + \alpha\delta\omega_k \cdot \frac{m_j}{N} + \frac{\alpha\delta}{d} \cdot \frac{m_j}{N}.$$

Then we have

$$\begin{aligned}\left(g_k^2+g_j^2\right)-\left(g'^2_k+g'^2_j\right) &\approx \omega_k^2\cdot\left(\delta+\frac{\alpha\delta m_k}{N}\right)^2+\omega_j^2\cdot\left(\delta+\frac{\alpha\delta m_j}{N}\right)^2-\omega_j^2\cdot\left(\delta+\frac{\alpha\delta m_k}{N}\right)^2-\omega_k^2\cdot\left(\delta+\frac{\alpha\delta m_j}{N}\right)^2\\ &\quad+2\omega_k\cdot\left(\delta+\frac{\alpha\delta m_k}{N}\right)\cdot\frac{\alpha\delta}{d}\cdot\frac{m_k}{N}+2\omega_j\cdot\left(\delta+\frac{\alpha\delta m_j}{N}\right)\cdot\frac{\alpha\delta}{d}\cdot\frac{m_j}{N}\\ &\quad-2\omega_j\cdot\left(\delta+\frac{\alpha\delta m_k}{N}\right)\cdot\frac{\alpha\delta}{d}\cdot\frac{m_k}{N}-2\omega_k\cdot\left(\delta+\frac{\alpha\delta m_j}{N}\right)\cdot\frac{\alpha\delta}{d}\cdot\frac{m_j}{N}\\ &=(\omega_k^2-\omega_j^2)\cdot\left(\left(\delta+\frac{\alpha\delta m_k}{N}\right)^2-\left(\delta+\frac{\alpha\delta m_j}{N}\right)^2\right)\\ &\quad+2(\omega_k-\omega_j)\cdot\left(\left(\delta+\frac{\alpha\delta m_k}{N}\right)\cdot\frac{\alpha\delta}{d}\cdot\frac{m_k}{N}-\left(\delta+\frac{\alpha\delta m_j}{N}\right)\cdot\frac{\alpha\delta}{d}\cdot\frac{m_j}{N}\right)\\ &>0.\end{aligned}$$

Therefore, when $\omega_k > \omega_j$ and $m_k > m_j$ ($m_k > m_j$ implies $k < j$ in our setting), if we interchange the order of ω_k and ω_j and keep the order of m_k and m_j, then the summation $g_k^2 + g_j^2$ will decrease, while all other g_i^2 ($i \neq k, j$) remain the same. Thus it means that interchange ω_k and ω_j will decrease the GDP volatility.

Also, notice that after the above interchange of ω_k and ω_j where $k < j$ by letting $\iota(j) = k$ and $\iota(k) = j$, the (k, j) pair becomes a new inversion pair (i.e., $(\omega_k - \omega_j)\cdot(m'_k - m'_j) < 0$). On the other hand, for all $i \neq j, k$, it is not difficult to see that the number of inversion pairs in $\{(k, s), (s, j)\}$ will not decrease. Therefore, after the above interchange of ω_k and ω_j where $k < j$, we will obtain that the number of inversion pairs strictly increases by at least one. Thus starting from any order of $\omega_1, \omega_2, \ldots, \omega_N$, if there exists some non-inversion pair (k, j), then we can repeat the above process by interchanging ω_k and ω_j, finally, we will always reach the case that $\omega_1 < \omega_2 < \ldots < \omega_N$, thus the σ_{GDP} is minimized when the ordering of $\{m_i\}$ is the exact opposite of $\{\omega_i\}$. Then we prove the second claim. The first and the third claim hold similarly.

We need the following concentration inequalities for the summation of independent Bernoulli random variables.

Lemma A.6 (Chung et al. (2003a)). *Assume that $X_1, X_2, \ldots, X_N$ are independent Bernoulli random variables with*

$$\mathbb{P}(X_i = 1) = p_i,\ \ \mathbb{P}(X_i = 0) = 1 - p_i.$$

For $X = \sum_i a_i X_i$ we have $E[X] = \sum_i a_i p_i$ and we define $\nu = \sum_i a_i^2 p_i$. Then we have

$$\mathbb{P}(X < \mathbb{E}[X] - t) \leq \exp\left(-\frac{t^2}{2\nu}\right), \tag{A.23}$$

$$\mathbb{P}\left(X > \mathbb{E}[X] + t\right) \leq \exp\left(-\frac{t^2}{2(Var(X) + at/3)}\right), \tag{A.24}$$

where $a = \max_k a_k$.

Step 2: Next, we show that when $\beta_m > 1$, the two technical conditions in the Proposition 5.4 holds almost surely.

Under the random graph formulation given by (5.5), when $\beta_m > 1$, we can verify that the two assumptions $\sum_{i\in\mathcal{M}_k} \omega_i \approx \sum_{i\in\mathcal{M}_k} \frac{1}{N} = \frac{m_k}{N}$ and $d_i \approx d$ for some $d = o(N)$ hold almost surely with probability one.

Actually, by 5.5 we obtain $\mathbb{E}[m_j] = \frac{N}{j^{1/\beta_m}}$. Thus by Lemma A.6 we have

$$\mathbb{P}\left(m_j < \frac{N}{j^{1/\beta_m}} - t\right) \leq \exp\left(-\frac{t^2}{2N/j^{1/\beta_m}}\right). \tag{A.25}$$

and

$$\mathbb{P}\left(m_j > \frac{N}{j^{1/\beta_m}} + t\right) \leq \exp\left(-\frac{t^2}{2N/j^{1/\beta_m} + 2t/3}\right). \tag{A.26}$$

Therefore,

$$\mathbb{P}\left(\frac{N}{j^{1/\beta_m}} - t \leq m_j \leq \frac{N}{j^{1/\beta_m}} + t\right) \geq 1 - \exp\left(-\frac{t^2}{2N/j^{1/\beta_m}}\right) - \exp\left(-\frac{t^2}{2N/j^{1/\beta_m} + 2t/3}\right). \tag{A.27}$$

Let $t = \left(\frac{N}{j^{1/\beta_m}}\right)^{2/3}$ in the above inequality, we obtain

$$\mathbb{P}\left(\frac{N}{j^{1/\beta_m}} - \left(\frac{N}{j^{1/\beta_m}}\right)^{2/3} \leq m_j \leq \frac{N}{j^{1/\beta_m}} + \left(\frac{N}{j^{1/\beta_m}}\right)^{2/3}\right) \tag{A.28}$$

$$\geq 1 - \exp\left(-\frac{1}{2}\left(\frac{N}{j^{1/\beta_m}}\right)^{1/3}\right) - \exp\left(-\frac{\left(N/j^{1/\beta_m}\right)^{4/3}}{2N/j^{1/\beta_m} + 2\left(N/j^{1/\beta_m}\right)^{2/3}/3}\right). \tag{A.29}$$

When N tends to infinity, the above probability tends to 1, thus

$$m_j = \frac{N}{j^{1/\beta_m}} + o\left(\frac{N}{j^{1/\beta_m}}\right)$$

holds almost surely with probability one when $N \to \infty$. Thus

$$m_j \approx \frac{N}{j^{1/\beta_m}}$$

holds almost surely with probability one.

Similarly, let $v(N) = \frac{N}{N-1}\left(\zeta(\frac{1}{\beta_m}, N) - \frac{1}{i^{\frac{1}{\beta_m}}}\right)$ and $d = \zeta(\frac{1}{\beta_m}, N)$. We obtain

$$\mathbb{P}\left(v(N) - t \leq d_j \leq v(N) + t\right) \geq 1 - \exp\left(-\frac{t^2}{2v(N)}\right) - \exp\left(-\frac{t^2}{2v(N) + 2t/3}\right). \tag{A.30}$$

Let $t = (v(N))^{2/3}$ in the above inequality, we obtain

$$\mathbb{P}\left(v(N) - (v(N))^{2/3} \leq d_j \leq v(N) + (v(N))^{2/3}\right) \tag{A.31}$$

$$\geq 1 - \exp\left(-\frac{1}{2}(v(N))^{1/3}\right) - \exp\left(-\frac{(v(N))^{4/3}}{2v(N) + 2(v(N))^{2/3}/3}\right). \tag{A.32}$$

When N tends to infinity, the above probability tends to 1, thus

$$d_j = v(N) + o(v(N)) = \zeta(\frac{1}{\beta_m}, N) + o\left(\zeta(\frac{1}{\beta_m}, N)\right)$$

holds almost surely with probability one when $N \to \infty$. Thus

$$d_j \approx \zeta(\frac{1}{\beta_m}, N) = d$$

holds almost surely with probability one.

The proof is now complete.

□

B. Data Appendix

B.1 Text-Based Attention Measure

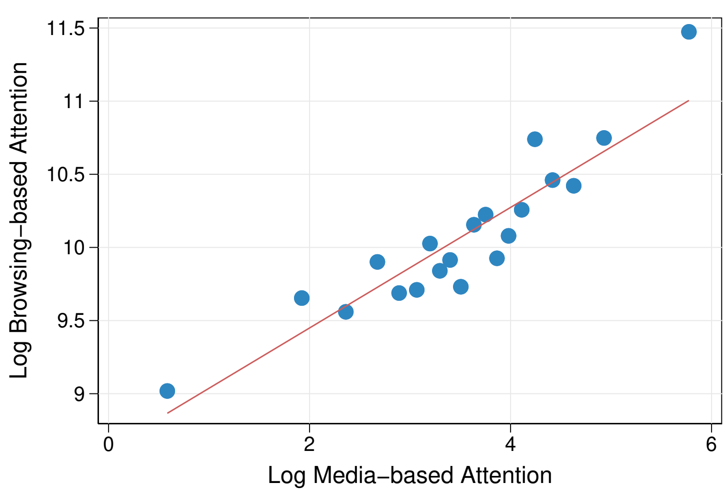
Construction from news media. To construct a text-based measure of attention from news media, we collect newswire-type articles published on Yahoo Finance in 2016. These articles are typically brief and distributed in real time, making them a timely source of firm-related information. We apply Named Entity Recognition (NER) using the Stanford NLTK (Natural Language Toolkit) package to extract company names from each article. NER is a natural language processing technique that identifies entities such as organizations, locations, and individuals within text.

Following entity extraction, we use a fuzzy matching algorithm to align the identified company names with those listed in the Compustat database, accounting for minor discrepancies in naming conventions. A news article is classified as relevant to a firm if it explicitly mentions the firm's name or any recognized variant.

This procedure yields a panel dataset that records the timestamp of each news article, its unique identifier, and the associated company or companies. Notably, a single news article may be linked to multiple firms when coverage spans more than one company. We aggregate the total number of reports on firm i, which measures firm i's attention received.

Correlation with Browsing-Based Attention. Figure B.1 binscatters the correlation between our browsing-based attention measure and the text-based attention measure from the news media at the firm level. The correlation is 0.39 with standard error of 0.35. This shows that our attention measure is significantly positively correlated with attention proxied by media news.

Figure B.1: Browsing Received-Sent Correlation and Firm Sizes

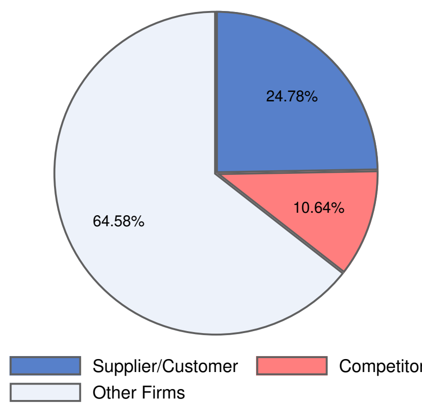


B.2 Composition of Browsing Activity

To examine how firms allocate their limited attention, we link our browsing data to the FactSet Revere Supply Chain database available via WRDS. This comprehensive resource provides detailed information on publicly disclosed supply chain relationships, including the identities of both suppliers and customers. Using these linkages, we identify each firm's direct supply chain partners. On average, a supplier is associated with 7.6 reported customer relationships, while a customer maintains 7.55 reported supplier relationships per year.

Figure B.2 illustrates how firms distribute their browsing activity across three categories: (i) supply chain partners (suppliers and customers), (ii) competitors, and (iii) other firms that are neither competitors nor supply chain partners. On average, firms direct 24.78% of their browsing toward suppliers and customers, and 10.6% toward competitors. The remaining share of browsing is directed toward other firms.

Figure B.2: Allocation of Browsing Activity



Note: This figure plots the allocation of a firm's browsing volume to its suppliers or customers, its competitors, and other firms in the Factset dataset in 2016.

C. Additional Tables and Figures

Table C.1: Filed Documents Being Viewed Most

Form	Avg Browsing	Description
S-4	273,887,040	Registration of securities issued in business combination transactions
POSASR	46,209,696	Post-effective Amendment to an automatic shelf registration statement on Form S-3ASR or Form F- 3ASR
S-3ASR	32,536,336	Automatic shelf registration statement of securities of well-known seasoned issuers
F-4	15,068,362	Registration statement for securities issued by foreign private issuers in certain business combination transactions
F-3	13,138,934	Registration statement for specified transactions by certain foreign private issuers
T-3	12,627,994	Initial application for qualification of trust indentures
F-3ASR	9,830,287	Automatic shelf registration statement of securities of well-known seasoned issuers
40APP	1,533,771	Applications under the Investment Company Act other than those reviewed by Office of Insurance Products
424B5	1,074,623	Prospectus filed pursuant to Rule 424(b)(5)
15-15D	1,003,571	Notice of suspension of duty to file reports pursuant to Section 13 and 15(d) of the Act

Note: This table lists the 10 forms filed on the EDGAR with the most browsings in 2016.

Table C.2: Predictive Power of Browsing-weighted Sales Growth (Median)

	Median Forecast of Industrial Production in $t+1$					
	(1)	(2)	(3)	(4)	(5)	(6)
$G_{t-1}^{browsing}$	0.027**				0.021*	0.022*
	(0.010)				(0.012)	(0.013)
$G_{t-2}^{browsing}$	0.036***				0.025**	0.023*
	(0.011)				(0.011)	(0.013)
$G_{t-3}^{browsing}$	0.020*				0.018*	0.022*
	(0.011)				(0.010)	(0.011)
$G_{t-4}^{browsing}$	-0.00039				0.0077	0.0069
	(0.010)				(0.0100)	(0.011)
G_{t-1}^{sales}		-0.0014		0.0026		-0.0084
		(0.042)		(0.043)		(0.037)
G_{t-2}^{sales}		-0.033		0.0032		-0.0011
		(0.048)		(0.050)		(0.043)
G_{t-3}^{sales}		-0.010		-0.037		-0.034
		(0.047)		(0.046)		(0.040)
G_{t-4}^{sales}		0.079*		0.0044		0.031
		(0.041)		(0.044)		(0.038)
Macro controls (X)			✓	✓	✓	✓
R^2	0.49	0.11	0.56	0.60	0.76	0.77
Adjusted R^2	0.43	0.01	0.39	0.33	0.58	0.53
No. Observations	40	40	44	40	40	40

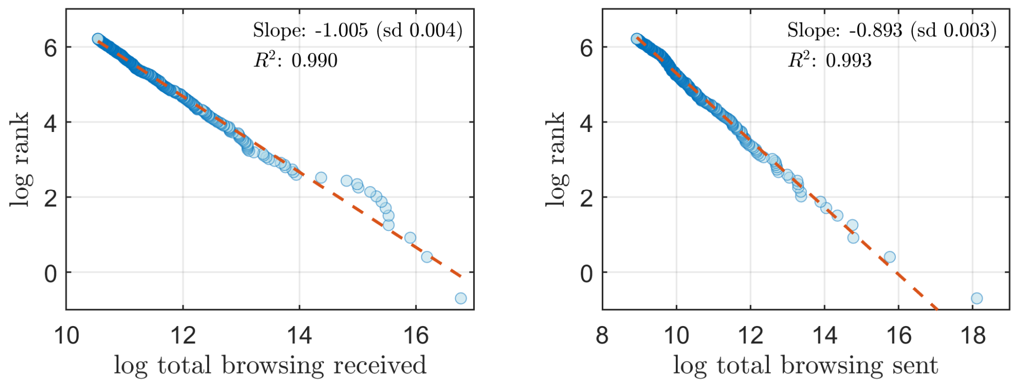
Note: This Table presents the estimates for regression equation (2.4) using median of the forecast. The sales growth rates are winsorized at 1% and 99%. The macroeconomic control variables $\{X_{t-j}\}_{j=1}^{4}$ include oil supply shocks, monetary policy shocks, the 3-month nominal T-bill rate, and the term spread. The sample periods are from 2006 1st quarter to 2016 4th quarter. *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.3: Predictive Power of Browsing-weighted Sales Growth (Alternative Targets)

	Mean Forecasts of macro variables in $t+1$							
	Real GDP		Unemployment		Consumption		Investment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$G^{\text{browsing}}_{t-1}$	0.019***		-0.17***		0.017***		0.036*	
	(0.0048)		(0.056)		(0.0038)		(0.019)	
$G^{\text{browsing}}_{t-2}$	0.016***		-0.13**		0.011***		0.057***	
	(0.0050)		(0.057)		(0.0039)		(0.019)	
$G^{\text{browsing}}_{t-3}$	0.0092*		-0.095		0.0053		0.050**	
	(0.0049)		(0.057)		(0.0038)		(0.019)	
$G^{\text{browsing}}_{t-4}$	0.0014		-0.060		0.0041		0.033*	
	(0.0048)		(0.055)		(0.0037)		(0.018)	
G^{sales}_{t-1}		-0.018		0.18		-0.023		-0.026
		(0.021)		(0.22)		(0.016)		(0.078)
G^{sales}_{t-2}		-0.019		0.24		-0.011		-0.015
		(0.024)		(0.25)		(0.019)		(0.089)
G^{sales}_{t-3}		0.0035		-0.10		0.0056		-0.021
		(0.024)		(0.25)		(0.019)		(0.089)
G^{sales}_{t-4}		0.028		-0.15		0.012		0.11
		(0.021)		(0.22)		(0.016)		(0.077)
R^2	0.55	0.10	0.45	0.09	0.56	0.11	0.51	0.06
Adjusted R^2	0.50	0.00	0.38	-0.02	0.51	0.01	0.45	-0.05
No. Observations	40	40	40	40	40	40	40	40

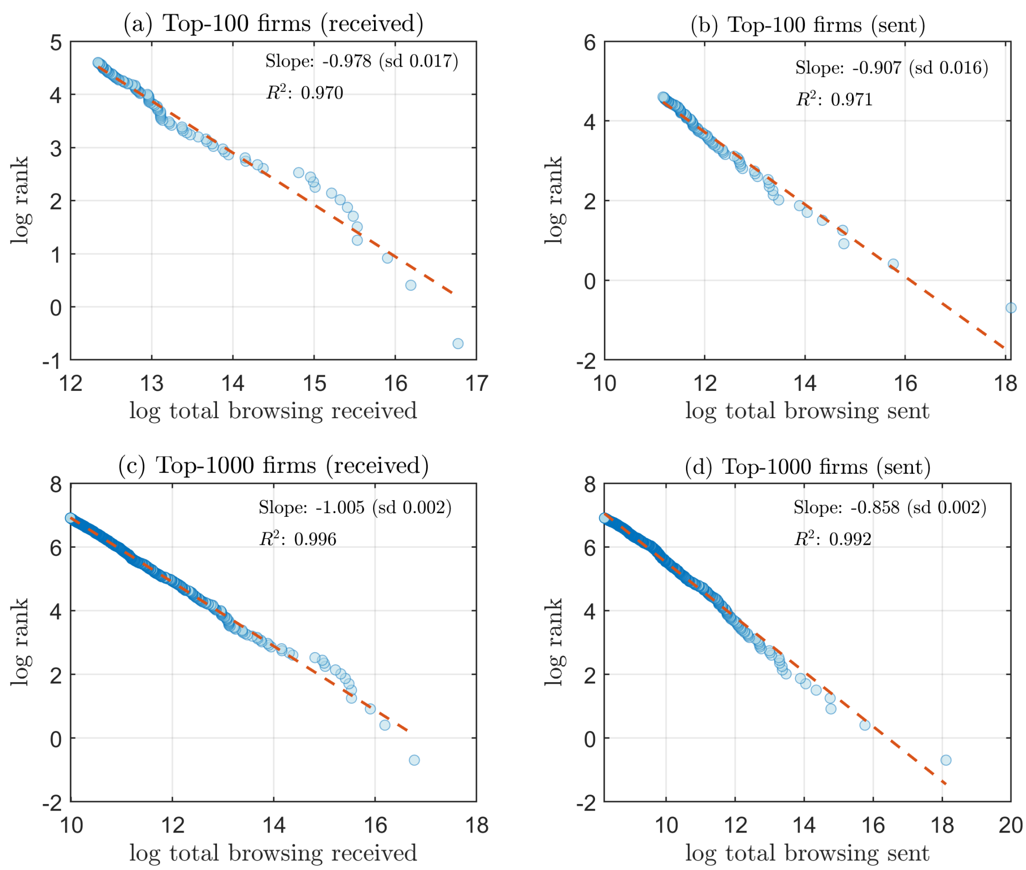
Note: This Table presents the estimates for regression equation (2.4) using other macroeconomic targets of the forecast. The sales growth rates are winsorized at 1% and 99%. The macroeconomic control variables $\{X_{t-j}\}_{j=1}^{4}$ include oil supply shocks, monetary policy shocks, the 3-month nominal T-bill rate, and the term spread. The sample periods are from 2006 1st quarter to 2016 4th quarter. *Significance*: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure C.1: Robustness: The Browsing Power-Law (excluding financial firms)



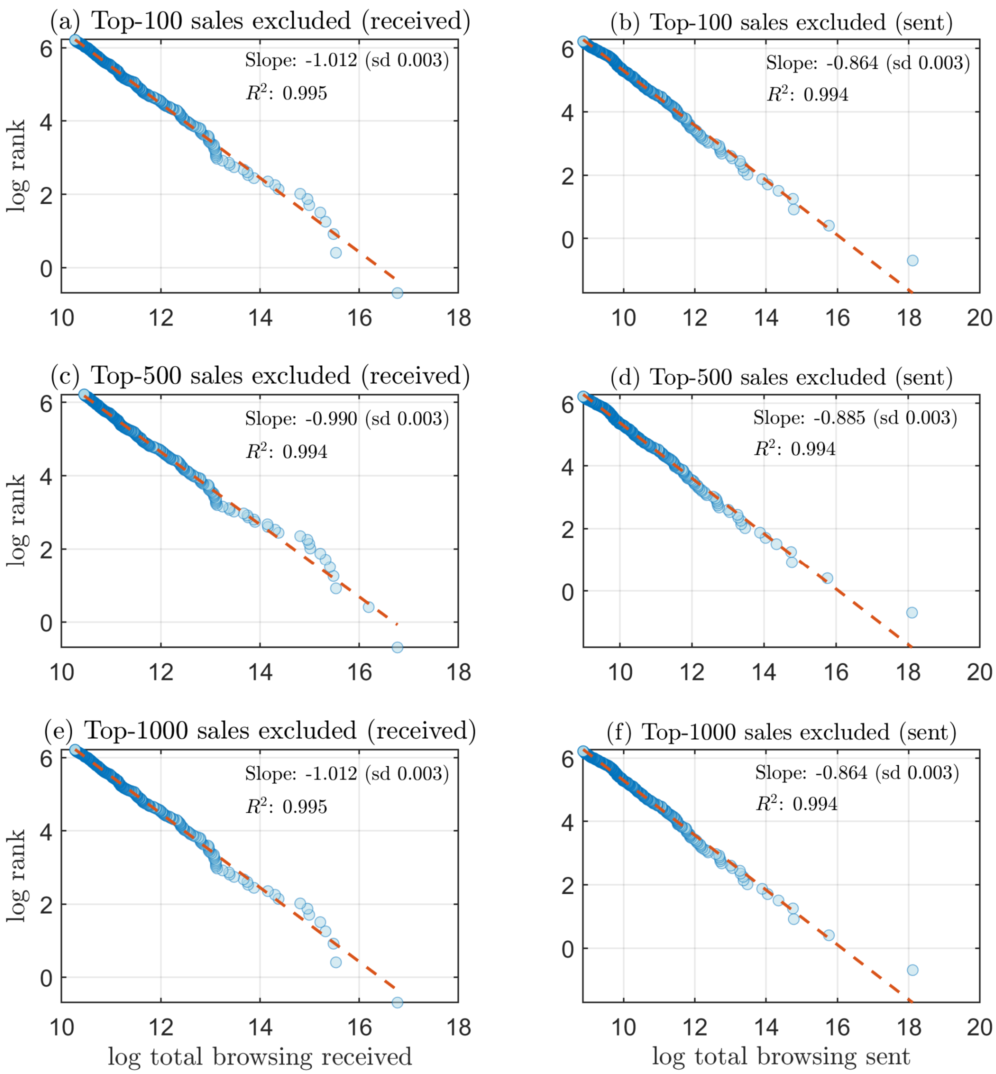
Note: This figure shows that the power-law patterns revealed in the main text continue to hold if we exclude financial firms from our sample.

Figure C.2: Robustness: The Browsing Power-Law (Top 100/1000)



Note: This figure shows that the power-law patterns revealed in the main text continue to hold if we consider top-100 or top-1000 firms of attention received or sent.

Figure C.3: Robustness: The Browsing Power-Law (Excluding Top-sales firms)



Note: This figure shows that the power-law patterns revealed in the main text continue to hold if we exclude the top-100, top-500, or top-1000 sales firms.