# Micro Shocks and Macro Fluctuations in the Information Network*

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

July 2026

**Abstract**

In this paper, we study how heterogeneity in firms' attention allocation amplifies micro-level shocks into aggregate economic fluctuations. Empirically, using data on firms' browsing of one another's electronic filings on EDGAR, we document that the attention received by firms follows a fat-tailed distribution. A small subset of firms attracts a disproportionately large share of attention, consistent with our finding that the browsing-weighted measure of firm sales growth exhibits substantial predictive power for survey-based macroeconomic forecasts. Theoretically, we extend the classical dispersed-information framework to incorporate asymmetric firm-level attention through a directed graph. We derive conditions under which granular effects emerge even when firms are homogeneous in size, and we analyze how attention heterogeneity interacts with firm-size heterogeneity in magnifying micro-level shocks. Quantitatively, we find that attention heterogeneity alone can generate sizable aggregate fluctuations from micro-level shocks, and the extent to which attention asymmetry complements firm-size heterogeneity depends critically on the cross-sectional alignment between attention and firm size.

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

JEL codes: D83, E32

*We are grateful to Manuel Amador, George-Marios Angeletos, Marco Bassetto, Ian Dew-Becker, V.V. Chari, Joel Flynn, Xavier Gabaix, Christian Hellwig, Chen Lian, Wenlan Luo, Simon Mongey, Stephen Morris, Richard Rogerson, Thomas Sargent, Karthik Sastry, Alireza Tahbaz-Salehi, Christopher Tonetti, and Venky Venkateswaran for discussing our paper at the 2026 Cowles Summer Conference. The paper benefited from the discussion with seminar participants at Federal Reserve Bank of Minneapolis, Rice University, Inter-American Development Bank, 2025 BSE Summer Forum: Macro Fluctuations with Micro Frictions, 2025 NBER SI: Impulse and Propagation Mechanisms, ESWC 2025, WESEAMS 2025, Jinan University, Tsinghua University, Renmin University of China, Shandong University, NTU Singapore, National University of Singapore, Singapore Management University, CKGSB, INSEAD, Virtual Asian Seminar Series on Expectations in Macroeconomics, Princeton University, Tsinghua PBCSF, Peking University HSBC Business School, Federal Reserve Bank of Chicago, Federal Reserve Bank of Cleveland, Yale Cowles Summer Conference, IMF, and 9th HKUST and Jinan Joint Workshop on Macroeconomics. We thank Yushu Zhang and Allison Yu for their superb research assistance.

†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 central questions in modern macroeconomics concerns the microeconomic origins of aggregate fluctuations. According to conventional wisdom, the effects of firm-specific shocks should wash out in the aggregate. This view has been challenged by theories that allow for micro-level heterogeneity in firm size, as in Gabaix (2011), or allow for input-output linkages across sectors, as in Acemoglu et al. (2012). In a similar vein, firms also differ substantially in the amount of attention they receive. Some salient firms are frequently mentioned in the media and are closely watched by other firms, while most others receive little attention. A natural question is what makes a firm important. Is it because firms are large, or because they are 'star' firms in the spotlight, closely watched by others? Can the heterogeneity in firm salience also give rise to granular effects? How do 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, using data on firms' browsing of one another's electronic filings on EDGAR (Electronic Data Gathering, Analysis, and Retrieval). This suggests that a small number of firms receive a disproportionately large share of aggregate attention. Importantly, the browsing-weighted sales growth of the most-attended firms has substantially greater predictive power for macroeconomic forecasts in the survey data than the sales-weighted growth of the largest firms. Theoretically, we extend the noisy business cycle framework (Angeletos and La'O, 2010) to incorporate asymmetric firm-level attention through a directed graph. We provide the condition on the in-degree and out-degree of this information network under which granular effects arise, even when firms are homogeneous in size. We also characterize the interaction between the attention channel and the traditional firm-size channel under power-law distributions. Across these two dimensions, the heavier-tailed distribution qualitatively determines the economy's granular properties, while granular effects are larger when firm size and attention are more positively sorted. Quantitatively, we show that the attention channel alone can generate sizable aggregate fluctuations driven by micro shocks. When the firm-size channel is also present, the attention channel still accounts for about one-third of the granular effects.

**Empirics.** Our motivating facts use detailed firm-to-firm browsing information on documents filed on EDGAR following the procedure by Chen et al. (2020). These bilateral browsing activities provide a direct measure of firms' attention allocation. We use the total number of browsing requests received by a firm to measure its received attention, and the total number of browsing requests sent by a firm to measure its attention directed towards other firms.

We highlight three facts. First, the distributions of attention received and attention sent both exhibit power laws. In particular, attention received is close to Zipf's law, while attention sent is even more fat-tailed. These patterns imply that a small number of firms receive a disproportionately large share of attention.

Second, attention-based salience is empirically distinct from firm size. Although larger firms tend to receive and send more browsing requests, the relationship is quantitatively weak. The elasticity of attention received with respect to sales is only about 0.23, and the corresponding elasticity for attention sent is about 0.27. Moreover, the set of firms that receive the most attention overlaps only partially with the set of the largest firms. This suggests that the concentration of attention is not merely a relabeling of the concentration of economic activity, but instead captures a separate dimension of firm-level heterogeneity.

Third, the information network is related to both aggregate expectations and realized macroeconomic outcomes. We construct an attention-weighted measure of sales growth among firms receiving the most attention and compare it with the conventional sales-weighted measure based on the largest firms. The attention-weighted measure has substantially stronger predictive power for professional forecasters' expectations about future macroeconomic outcomes. Importantly, this pattern is not limited to forecasts: attention-weighted sales growth also helps explain realized fluctuations in these macroeconomic variables themselves. These findings point to a theory in which firm salience constitutes an independent source of granularity.

**Theory.** Motivated by these facts, we develop a framework for studying an incomplete-information economy populated by a finite, potentially large, number of firms. Firms' output choices depend on their own productivity shocks and on aggregate output through standard demand linkages. We build on Angeletos and La'O (2010), but depart from their framework in one central respect: firms differ in their information sets, and bilateral information flows across firms are represented by an information network. In this network, a firm's in-degree measures the amount of attention it receives, while its out-degree measures the amount of attention it directs toward other firms. These objects provide the model counterparts to attention received and attention sent in the empirical analysis.

An individual firm's micro shock can affect GDP in three ways: (1) the efficiency channel, where the shock directly affects the firm's own production; (2) the information channel, where a firm's shock is informative about its output and induces comovements when observed by other firms; (3) the coordination channel, where a salient firm's shock is useful in forecasting other firms' decisions. We provide an explicit solution for the elasticity of GDP with respect to each firm's micro shock in equilibrium, accounting for all three forces.

Under what conditions can a firm's idiosyncratic component induce aggregate fluctuations when the number of firms goes to infinity? We answer this question in two steps. First, we define a key statistic that summarizes a firm's importance in the information network: the diluted in-degree index. For a given firm, this index is defined as the sum of the reciprocal of the out-degrees of all firms that observe it. It therefore measures the effective attention received by a firm. Second, we derive bounds of the elasticity of GDP with respect to a firm's micro shock, both of which hinge on the diluted in-degree index. Granular effects emerge if and only if this index grows at the same rate as the number of firms. Intuitively, a firm can generate granular effects only if it is observed by a sizable fraction

of other firms and if the attention of those firms is sufficiently concentrated on it. This property is consistent with the empirical evidence on firms' browsing behavior.

**Firm size and attention under power law.** To gain further insight into the interaction between the conventional firm-size channel and the attention channel, we extend the framework to allow for both types of heterogeneity. We assume that both the firm size distribution and the attention received, measured by in-degrees, follow power-law distributions. In this setting, the nature of aggregate volatility driven by micro shocks depends on the tail 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 heavier-tailed distribution exerts a dominant influence on the economy's granular properties. 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 leverage the patterns of browsing activities in EDGAR to discipline the information matrix. The model matches the salient tail properties of the attention distribution, as well as key properties of forecast errors in the survey data.

Quantitatively, the heterogeneous information channel alone can generate sizable fluctuations from micro shocks. In our baseline calibration, the top 100 firms that receive the most attention account for about 16 percent of aggregate volatility. In a model with perfect information or with symmetrically dispersed information, this contribution is substantially muted, only about 2 percent. Once firm-size heterogeneity is introduced, the perfect-information model can generate sizable granular effects, and GDP volatility arising from micro-level shocks is driven entirely by large firms. By contrast, when both the attention and size channels are present, firms that are salient but not necessarily large remain important contributors to aggregate fluctuations.

The key mechanism is that the salient firms act as semi-public signals. They are important in coordinating decisions across firms, and many firms respond to salient firms' productivity changes when forming higher-order expectations. However, if a firm is large but is not observed by other firms, the GDP elasticity with respect to this firm's shock will be lower, reflecting the standard dampening effect of informational frictions. Whether these information and coordination channels reinforce the standard firm-size channel depends on the correlation between the firms' received attention and their sizes. Because these two dimensions are only weakly correlated in the data, the information network generates a distinct source of granularity rather than simply reinforcing the traditional firm-size channel.

**Related literature.** This paper complements the literature on the granular effect in macroeconomics. The first strand of this literature emphasizes the role of firm-size heterogeneity, large-firm dynamics, and market power (Gabaix, 2011; Carvalho and Gabaix, 2013; Di Giovanni et al., 2014; Carvalho and Grassi, 2019; Eeckhout et al., 2020). The second strand highlights production-network linkages as a channel through which microeconomic shocks propagate across sectors and firms (Horvath, 1998; Foerster et al., 2011; Acemoglu et al., 2012; Baqaee, 2018; Baqaee and Farhi, 2020; Dew-Becker, 2023). Di Giovanni and Levchenko (2012, 2013) and Di Giovanni et al. (2014) further extend these insights in a global trade setting. Relative to these contributions, we identify a distinct source of granular effects arising from firms' information networks and show empirically that idiosyncratic firm-level shocks can generate sizable aggregate fluctuations independently of firm-size heterogeneity and the structure of production networks.

This paper also contributes to the large literature on macroeconomic models with incomplete information. The classical framework is typically formulated as a beauty contest game where all agents are symmetric in terms of payoff and information structure (Morris and Shin, 2002; Woodford, 2003; Lorenzoni, 2009; Angeletos and La'O, 2010). Our model extends this framework by allowing for an asymmetric information structure via a directed graph. This extension accommodates the possibility that firms' idiosyncratic fundamentals become semi-public signals that capture both the salience channel and the traditional informational-friction channel when affecting shock propagation. In addition, the technical results in analyzing the interaction between general equilibrium forces with random graphs may be applicable to other related problems.

Our paper belongs to the expanding literature that incorporates information frictions into models of networks (Bergemann et al., 2017; Golub and Morris, 2020; Leister et al., 2022; La'O and Tahbaz-Salehi, 2022; Nikolakoudis, 2024; Afrouzi et al., 2025; Hajdini et al., 2025; Cuciniello and Paciello, 2026). On the empirical side, our earlier work Fang et al. (2024) employs the same EDGAR browsing data set to characterize firms' endogenous information acquisition behaviors along the production network and studies optimal monetary policy in this environment. This approach complements the measurement of firms' attention allocation based on text analysis (Flynn and Sastry, 2024; Song and Stern, 2025) and the measurement of received attention based on media coverage (Chahrour et al., 2021; Bui et al., 2024). On the theory side, Jamilov et al. (2024) use a real production network model to study the impact of endogenous information acquisition (limited attention) on sentiment-driven aggregate fluctuations, arguing that attention is centered on downstream firms because of their roles as "information agglomerators". Relative to this literature, our empirical findings provide direct evidence of attention concentration and highlight the salience channel in affecting market participants' beliefs. Our theory builds on the bilateral attention between firms, and the granular effects do not rely on inter-sectoral linkages.

# 2. Motivating Facts

This section establishes three stylized facts about the cross-section of attention flows among firms, using novel internet traffic data that track firms' access to other firms' SEC filings on EDGAR. These facts collectively characterize the distribution of firm-level attention, its relationship to firm size, and its macroeconomic relevance.

First, we show that the distributions of both attention received and attention sent are fat-tailed and follow power laws: a small fraction of firms accounts for a disproportionate share of total attention flows. Second, despite this extreme concentration, attention scales only weakly with firm size. Third, an attention-weighted measure of firm-level sales growth strongly predicts professional forecasters' expectations about future GDP, industrial production, and other key macroeconomic aggregates. By contrast, sales growth weighted by firms' sales share has little predictive power.

## 2.1 Measuring attention with EDGAR Browsing Data

Our analysis leverages detailed firm-to-firm browsing information from EDGAR, the Electronic Data Gathering, Analysis, and Retrieval platform maintained by the U.S. Securities and Exchange Commission (SEC). Since 1996, all U.S. public companies have been required to file disclosures electronically through EDGAR. These filings, including annual reports (Form 10-K), quarterly reports (Form 10-Q), and material event disclosures (Form 8-K), are freely accessible online and serve as a primary source of firm-specific information for managers, investors, and analysts.[1]

**Data source.** In 2003, the SEC's Division of Economic and Risk Analysis began publishing the "EDGAR Logfile," a high-resolution dataset that records each instance when a document is viewed online. For each access, the log file records a partially anonymized IP address, a precise timestamp, and a unique document identifier (the Accession Number). This detailed information allows us to track browsing behavior at the firm-to-firm level.

**Constructing firm-to-firm attention flows.** Following Chen et al. (2020), we map partially anonymized IP addresses to firm identities using data from *ip-info.io*, a leading provider of IP information. This procedure yields a daily panel of browser-browsee pairs. For any period $t$, we construct a firm-to-firm browsing network $\mathcal{B}_t$, where each element

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

represents the number of times firm $i$ (the browser) accessed SEC filings of firm $j$ (the browsee) during period $t$, with $a_{ijs} = 1$ if such an access occurred at time $s$. From this network, we define two firm-level measures: total attention sent by firm $i$, $B^S_{it} = \sum_j b_{ijt}$, and total attention received by firm $j$, $B^R_{jt} = \sum_i b_{ijt}$.

[1] Appendix Table A.1 lists the most frequently browsed document types.

We argue that this browsing-based measure provides a meaningful proxy for firm-level attention. First, as the SEC's official disclosure platform, EDGAR provides the most timely and comprehensive access to firm-specific information. Material events must be filed on EDGAR before they appear on company websites or third-party platforms, making EDGAR a first-order information source for market participants.[2] Second, unlike proprietary platforms such as Bloomberg or Capital IQ, EDGAR is freely accessible, which mitigates selection concerns related to subscription costs. Third, because EDGAR usage reflects active information acquisition rather than passive exposure to news mentions, it provides a direct measure of firms' information acquisition behavior.

**Relation to existing attention measures.** Our browsing-based measure complements and extends text-based proxies widely used in the literature. Existing measures often rely on the content of public filings, which may be shaped by firms' strategic disclosure incentives. By contrast, our measure captures actual daily information acquisition behavior, free from such distortions. Moreover, unlike media-based proxies that typically reflect aggregate attention to a firm, our approach allows us to construct direct firm-to-firm attention linkages.

Despite these conceptual differences, our measure is positively correlated with existing text-based proxies, providing external validation. For example, the correlation between total attention received and a media-based text proxy is 0.39, with a standard error of 0.035.[3]

**Limitations.** Our measure captures the frequency of browsing events, but not their duration or intensity. If each visit conveys the same amount of information, total visits would be proportional to information acquisition. To the extent that visit duration varies systematically, for example because users spend more time on complex filings, our frequency-based measure will contain measurement error. Directly measuring the information acquired per visit is almost infeasible with available data, making our frequency-based measure a practical compromise.

**Baseline Sample.** Our analysis focuses on public firms, which are the primary filers on EDGAR and can be linked to Compustat. The sample period spans 2006 to 2016, beginning when EDGAR traffic data became consistently available and ending when the SEC discontinued the data series.[4] The resulting sample comprises 7,622 public firms and more than 713 million IP addresses, accounting for 52.4 percent of all EDGAR browsing activity during this period.

Table 1 reports summary statistics. Two features of the attention network stand out. First, the network is dense on average: the typical browser views filings from 180 distinct firms each year, while the typical browsee is accessed by 105 distinct browsers. Second, browsing volume is substantial and highly skewed. The mean browser generates 11,356 views annually, compared with a median of 9,881 views. Similarly, the mean browsee receives 5,897 views, compared with a median of 5,389 views.

[2] For example, according to a 2015 *Wall Street Journal* article (https://www.wsj.com/articles/the-109-894-word-annual-report-1433203762), General Electric's 2003 10-K received 21,987 downloads through EDGAR, compared with just 800 from the company's website.

[3] Appendix A.1 describes the construction of the text-based measure.

[4] Following Cao et al. (2023), we exclude automated web crawlers; all results are robust to their inclusion.

These summary statistics foreshadow the strong concentration in firm-level attention documented below.

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 viewed 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 |

## 2.2 Facts on Firms' Browsing Activities

We now present a set of novel empirical facts that motivate our theoretical analysis. These facts characterize the cross-sectional distribution of attention, both received and sent, its relationship with firm size, and its macroeconomic relevance. For clarity, we illustrate the main patterns using data from 2016. Appendix B.4 shows that all results are robust when pooling observations across years.

**Fact 1. The distributions of attention received and sent are fat-tailed and follow power laws.**

We begin by characterizing the cross-sectional distributions of attention received and attention sent. Figure 1 plots the two distributions. Both are highly right-skewed: while most firms engage in relatively little browsing activity, a small subset accounts for a disproportionately large share of total attention flows. This pronounced skewness suggests that both distributions may be well described by power laws.

To quantify the degree of tail fatness, we test for power-law behavior using the Rank-1/2 estimator proposed by Gabaix and Ibragimov (2011). For each distribution, we rank firms by total browsing activity and estimate the power-law exponent $\kappa$ from the regression

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

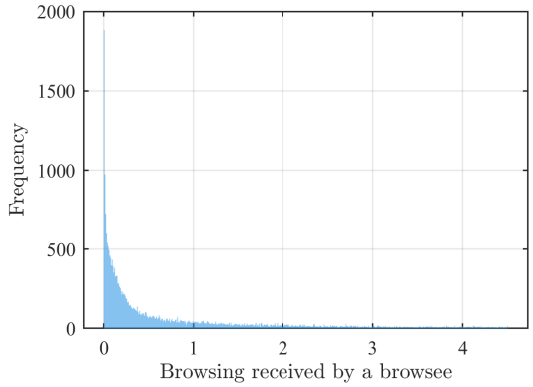
where $\text{Rank}_i$ is the rank of firm $i$, and $B_i$ denotes its total browsing activity, either received or sent. A power law implies $\text{Prob}(B_i > K) \propto K^{-\kappa}$, with $\kappa = 1$ corresponding to Zipf's Law.

For attention received, we estimate $\kappa = 0.999$ with a standard error of 0.063 and an $R^2$ of 0.993, using the top 500 firms.[5] This estimate is statistically indistinguishable from unity, indicating that

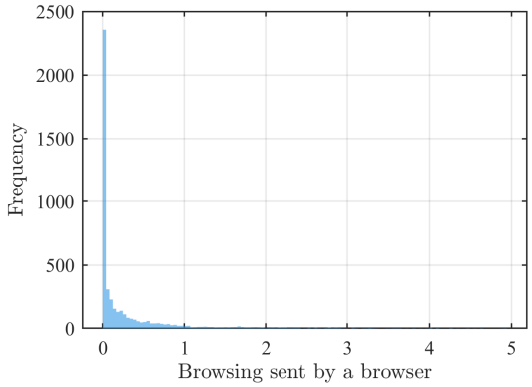
[5]The Rank-1/2 estimator mitigates the small-sample bias inherent in log-log rank-size regressions (Gabaix and Ibragimov, 2011).

Figure 1: Distribution of Firms' Browsing Activity

(a) Attention Received




(b) Attention Sent




*Notes*: Browsing activity is measured using 2016 data.

attention received closely follows Zipf's Law: the probability that a firm receives more than $K$ units of attention is inversely proportional to $K$. Figure 2a confirms this relationship visually.

For attention sent, the estimated exponent is $\kappa = 0.918$ with a standard error of 0.058 and an $R^2$ of 0.993. The point estimate is below one, suggesting an even fatter tail than Zipf's law. In other words, a small set of firms accounts for a particularly large share of total attention sent. Figure 2b illustrates this power-law pattern.

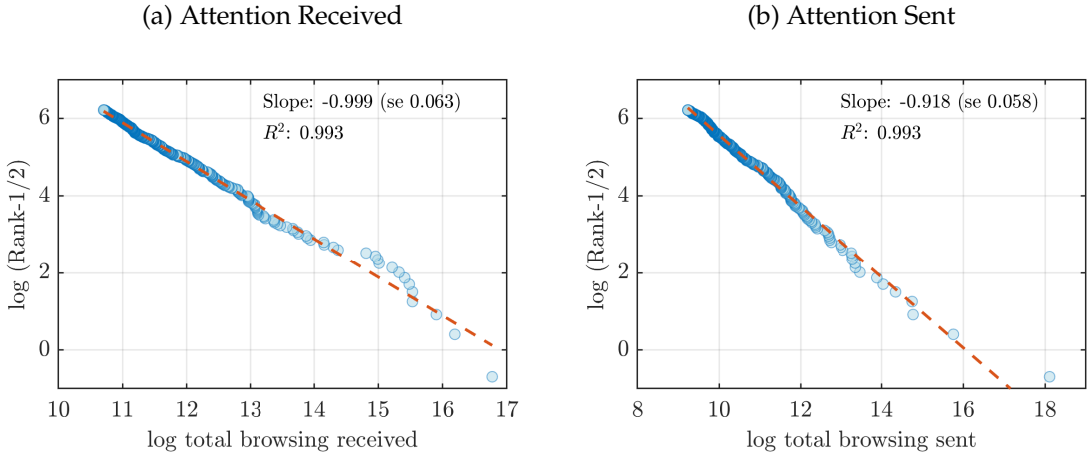
Appendix B.4 shows that these results are robust across alternative sample cuts and specifications. The estimates change little when we vary the rank cutoff, using either the top 100 or top 1,000 firms, exclude financial firms, or condition on industry. The power-law characterization also holds when we normalize browsing activity by the number of SEC filings, addressing the concern that our results may merely reflect heterogeneity in disclosure frequency.

**Discussion.** Two additional observations are useful for interpreting these patterns. First, the correlation between sending and receiving attention is positive only among the largest firms. For firms in the top 350 by sales, the correlation is approximately 0.1 and statistically significant; among smaller firms, attention received and attention sent are essentially uncorrelated. Appendix B.1 provides details, and we use this threshold relationship when calibrating the attention network in our quantitative model. Second, attention flows are moderately persistent over time: the annual autocorrelation is 0.53 for attention sent and 0.40 for attention received. Firms that are highly visible or highly attentive in one year tend to remain so in subsequent years.

**Fact 2. Attention scales with size at an elasticity far below one.**

A central prediction of granularity theories is that idiosyncratic shocks to large firms can affect aggregate fluctuations because these firms account for a disproportionate share of economic activity

Figure 2: The Browsing Power-Law




*Notes*: The figure plots the rank of each firm's total browsing activity against its total browsing activity on a log-log scale, using 2016 data. The fitted line corresponds to the Rank-1/2 regression: $\ln(\text{Rank}_i - 1/2) = c - \kappa \ln(B_i)$. Standard errors for $\widehat{\kappa}$ are computed as $(2/N)^{1/2}\widehat{\kappa}$ following Gabaix and Ibragimov (2011), where $N$ is the number of observations.

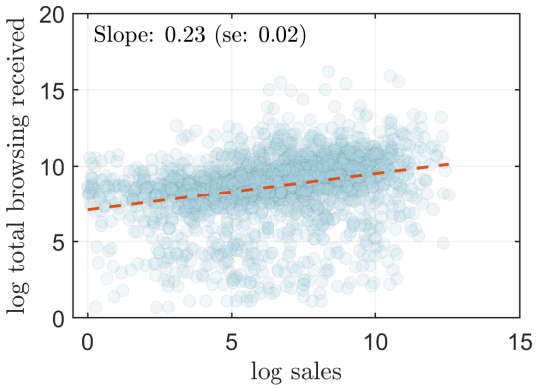
(Gabaix, 2011). Our attention-based notion of granularity operates through a different mechanism: firms matter if they are highly visible to other firms. To relate these two notions of firm importance, we examine the relationship between firms' browsing activity and firm size. We measure firm size using sales, which Appendix A.3 shows is the strongest predictor of both attention received and attention sent among a rich set of firm characteristics, including age, profitability, market valuation, tangibility, R&D intensity, and sales growth.

Despite the importance of large firms in aggregate production, attention only scales weakly with size. The elasticity of attention received with respect to sales is 0.23, with a standard error of 0.02. The corresponding elasticity for attention sent is 0.27, with a standard error of 0.01. Figure 3 presents scatter plots of these relationships together with fitted linear trends. These elasticities are substantially below one: a 10 percent increase in firm size is associated with less than a 3 percent increase in either attention received or attention sent. Attention-based granularity is therefore not simply a reflection of size-based granularity: large firms receive and send disproportionately less attention than their share of economic activity would predict.

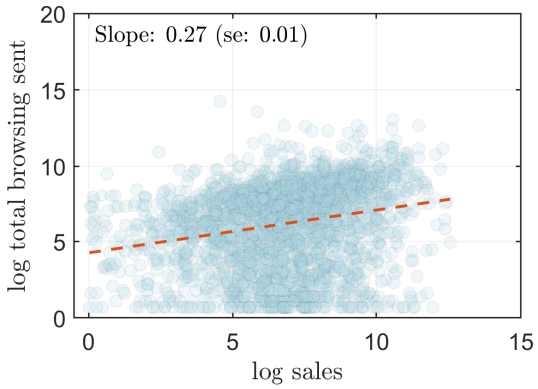
This weak scaling has direct implications for the power-law patterns documented in Fact 1. If the fat tails in attention were driven primarily by a few very large firms, then excluding the largest firms by sales should attenuate or eliminate the power-law behavior in attention received and attention sent. Appendix Figure B.10 shows that this is not the case: the estimated power-law exponents remain largely unchanged after dropping the top 100, 500, or 1,000 firms by sales. The concentration of attention is therefore distinct from the concentration of economic activity.

Figure 3: Attention and Firm Size

(a) Attention received




(b) Attention sent




*Notes*: The figure plots log total attention received (left panel) and log total attention sent (right panel) against log sales for the cross-section of firms in 2016. The estimated elasticities are substantially below one: 0.23 for attention received and 0.27 for attention sent.

**Fact 3. Attention-weighted firm sales growth predicts professional forecasters' macroeconomic expectations**

If a small set of firms receives a disproportionate share of attention, their sales performance may disproportionately influence how forecasters form expectations about aggregate outcomes such as GDP, industrial production, and unemployment. We test this hypothesis by constructing an attention-weighted measure of firm-level sales growth and examining whether it predicts mean forecasts from the Survey of Professional Forecasters (SPF).

For each quarter $t$, we rank firms by total browsing received and focus on the top $K = 100$ firms.[6] We construct the attention-weighted measure of sales growth for these firms:

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

where $\omega_{i,t}^{\text{attention}} = B_{i,t}^R / \sum_{j=1}^{K} B_{j,t}^R$ is firm $i$'s share of browsing received among the top $K$ firms in quarter $t$, and $B_{i,t}^R$ is total browsing received by firm $i$ in that quarter. Firm-level sales growth is measured as the year-over-year log change, $g_{i,t} = s_{i,t} - s_{i,t-4}$. We subtract the cross-sectional mean, $\overline{g}_t$, to isolate the firm-specific component of sales growth. To mitigate the influence of outliers, we winsorize sales growth at the 1st and 99th percentiles.[7]

[6] The results are robust to alternative choices of $K$, including 50, 200, and 500.

[7] The results are robust to alternative winsorization thresholds, including 5%/95% and the 20% winsorization of de-meaned growth rates used by Gabaix (2011).

For comparison, we construct an analogous sales-weighted measure:

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

where $\omega_{i,t}^{\text{sales}}$ is firm $i$'s sales share among the top $K$ firms with the largest sales. Appendix B.2.4 shows that the results are similar when we instead use the top $K$ firms with the largest absolute sales growth.

We then estimate the following regression specification:

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

where $f_t^{t+h}$ is the mean forecast, made in quarter $t$, for macroeconomic variable $f$ over a horizon of $h$ quarters.[8] Following Gabaix (2011), the control vector $X_t$ includes lagged macroeconomic variables and shocks: oil supply shocks, monetary policy shocks, the 3-month Treasury bill rate, and the term spread. The baseline specification uses a one-quarter horizon ($h = 1$) and four lags ($J = 4$). The results are robust to alternative horizon and lag lengths.

Table 2 presents the results for real GDP forecasts. Several findings stand out. First, attention-weighted sales growth strongly predicts one-quarter-ahead GDP forecasts. The coefficients on the first three lags are positive, and the first two are statistically significant at the 1 percent level. Second, the predictive power of the attention-weighted measure is economically large. Taken alone, it explains 55 percent of the variation in one-quarter-ahead GDP forecasts, as shown in column (1). By contrast, the sales-weighted measure explains only 10 percent of the variation, and none of its individual coefficients is statistically significant, as shown in column (2).

Columns (3)–(6) show that the predictive power of the attention-weighted measure is not subsumed by standard macroeconomic controls. Including both the attention-weighted measure and macro controls (column 5) yields an $R^2$ of 0.81—a 15 percentage point improvement over macro controls alone (column 3). Adding the sales-weighted measure provides almost no further explanatory power (column 6). These results indicate that the attention received by firms, rather than firm size alone, captures an important link between firm-level performance and macroeconomic expectations.

Figure 4 shows that these findings extend beyond GDP to other key macro aggregates: industrial production, the unemployment rate, real personal consumption expenditures, and real nonresidential fixed investment. For each series, the attention-weighted measure delivers a substantially higher $R^2$ than the sales-weighted measure, as shown in the left panel. The right panel plots the coefficient on the first lag for each measure. The attention-weighted coefficients have the expected signs and are statistically significant, while the sales-weighted coefficients are not. Appendix Table B.2 provides the complete regression results for each target variable.

[8] Results using median forecasts are quantitatively similar, as shown in Appendix Table B.1.

Table 2: Attention-Weighted Sales Growth Predicts GDP Forecasts

| | Mean Forecast of real GDP in $t+1$ | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $G^{\text{attention}}_{t-1}$ | 0.019*** | | | | 0.0081 | 0.0084 |
| | (0.0048) | | | | (0.0050) | (0.0050) |
| $G^{\text{attention}}_{t-2}$ | 0.016*** | | | | 0.0074 | 0.0066 |
| | (0.0049) | | | | (0.0048) | (0.0049) |
| $G^{\text{attention}}_{t-3}$ | 0.0090* | | | | 0.0094** | 0.012** |
| | (0.0049) | | | | (0.0043) | (0.0044) |
| $G^{\text{attention}}_{t-4}$ | 0.0013 | | | | 0.0063 | 0.0056 |
| | (0.0047) | | | | (0.0043) | (0.0043) |
| $G^{\text{sales}}_{t-1}$ | | -0.016 | | -0.0026 | | -0.0073 |
| | | (0.020) | | (0.018) | | (0.014) |
| $G^{\text{sales}}_{t-2}$ | | -0.018 | | -0.0075 | | -0.0071 |
| | | (0.023) | | (0.021) | | (0.016) |
| $G^{\text{sales}}_{t-3}$ | | 0.0030 | | -0.015 | | -0.020 |
| | | (0.023) | | (0.019) | | (0.015) |
| $G^{\text{sales}}_{t-4}$ | | 0.027 | | 0.00032 | | 0.014 |
| | | (0.019) | | (0.018) | | (0.014) |
| Macro controls (X) | | | ✓ | ✓ | ✓ | ✓ |
| $R^2$ | 0.55 | 0.10 | 0.66 | 0.70 | 0.81 | 0.86 |
| Adjusted $R^2$ | 0.50 | 0.00 | 0.51 | 0.50 | 0.68 | 0.70 |
| No. Observations | 40 | 40 | 40 | 40 | 40 | 40 |

*Notes*: The dependent variable is the mean SPF forecast of real GDP growth for the next quarter. Sales growth is winsorized at 1% and 99%. Macroeconomic controls include four lags each of oil supply shocks, monetary policy shocks, the 3-month Treasury bill rate, and the term spread. Sample: 2006:Q1–2016:Q4. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

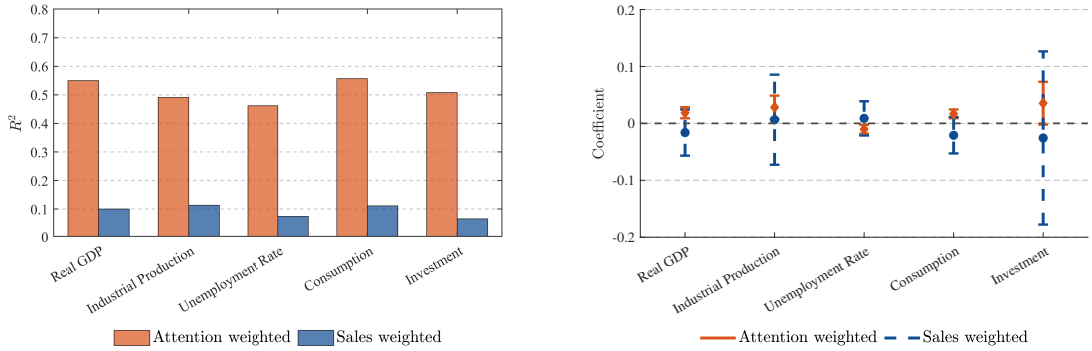
**Decomposing the source of predictive power.** The attention-weighted measure differs from the sales-weighted measure along two dimensions: the weights assigned to firms and the composition of the firm set. To separate these two channels, we conduct two auxiliary exercises.

First, we hold the firm set fixed, using the top 100 firms by attention received, but replace the attention weights with sales weights. The resulting measure has little predictive power for GDP forecasts, as shown in Appendix Figure B.2 panel (a). Second, we retain the attention weights but shift the firm set to firms ranked 101–200 by attention received.[9] This measure retains some predictive power, but its explanatory power is substantially smaller than that of the top-100 attention-weighted measure, as shown in Appendix Figure B.2 panel (b).

Together, these exercises indicate that both the choice of weights and the selection of firms are important. Attention-weighted growth predicts macroeconomic forecasts not simply because it assigns attention-based weights, nor simply because it focuses on firms that receive high attention. Rather, its predictive power comes from combining the two: it identifies the firms that are most closely watched

[9]The results are similar when using any 100-firm set outside the top 100.

Figure 4: Attention-Weighted Growth Predicts Forecasts Across Multiple Macro Variables




*Notes*: The left panel compares $R^2$ from regressions using only attention-weighted growth and only sales-weighted growth for five macroeconomic aggregates. The right panel plots the coefficient on the first lag, $t-1$, for each measure, with 95 percent confidence intervals. Attention-weighted growth consistently outperforms sales-weighted growth.

and weights them by the amount of attention they receive. Appendix B.2.3 provides complete details.

**Explaining current economic activity.** The analysis above shows that attention-weighted firm growth predicts professional forecasters' expectations about future macroeconomic conditions. A natural question is whether it also helps explain contemporaneous fluctuations in economic activity. Following the granularity literature (Gabaix, 2011), we regress current-quarter growth rates of real GDP and other macroeconomic aggregates on current and lagged values of attention-weighted and sales-weighted sales growth. Table B.3 in Appendix B.3 reports the results.

Consistent with Gabaix (2011), sales-weighted growth alone explains a meaningful share of GDP fluctuations, with an $R^2$ of 0.26. However, attention-weighted growth delivers substantially greater explanatory power, with an $R^2$ of 0.45 when used alone. It also has significant incremental explanatory power when added to specifications that already include the sales-weighted measure. Appendix Figure B.4 shows that this pattern extends to industrial production, unemployment, consumption, and investment. Thus, attention-weighted aggregate sales growth contains information not only about macroeconomic forecasts, but also about current economic activity. This information goes beyond what is captured by size-based granularity alone.

### Summary

The first two facts provide a detailed characterization of firms' attention received and sent, as measured by the browsing activity, and of the relationship between attention and firm size. Both attention received and sent exhibit power-law patterns, and these patterns are not merely a reflection of firm size. The third fact shows that the performance of a small set of highly attended firms has substantial predictive power for professional forecasters' expectations about macroeconomic variables. Together, these empirical patterns motivate and discipline the theoretical and quantitative analysis that follows.

# 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 $N$ information islands, each of which is inhabited by a firm indexed by $i \in \{1, 2, \ldots, N\}$. 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

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

where $\nu$ parameterizes the Frisch elasticity of labor supply. Here, $\ell_{it}$ is the labor supply of the workers who are sent to island $i$. Final consumption $C_t$ is a CES aggregate of differentiated varieties,

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

where $\rho$ determines the elasticity of substitution across goods produced by different firms. The associated aggregate price index and demand schedule for a good $i$ are given by

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

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 $\pi_{it}$ are the local wage and profit in island $i$, respectively. In the second stage, the household chooses its optimal consumption plan given its income.

**Firms.** The production technology of a firm is given by

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

where $z_{it}$ is firm $i$'s productivity shock at period $t$. 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 $i$ posts 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 can then 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

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

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

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

which equates the marginal cost and marginal benefit of hiring one more unit of labor. From 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. 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*

$$\delta = \frac{\rho(1+\nu)}{\rho+\nu}, \qquad \text{and} \qquad \alpha = \frac{\nu}{\rho+\nu}.$$

Condition (3.5) summarizes different forces that influence firms' output choices. The coefficient $\delta$ parameterizes the partial equilibrium effects, that is, how a firm would respond to its own productivity $z_{it}$ when others' actions are held constant. The coefficient $\alpha$, on the other hand, parameterizes the general equilibrium effects. Because aggregate demand depends on output across all firms, it gives rise to 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

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

That is, firms need to form expectations about others' productivities, as well as expectations about other firms' expectations about 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. Unlike the environment with a continuum of agents (Morris and Shin, 2002) where agents only need to care about the economy's 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 firm observes $z_{jt}$, then firm $j$'s shock will have no impact on the rest of the economy. In contrast, if a large number of firms observe $z_{jt}$, 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.

### 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

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

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.

A brief discussion of the endogeneity of the information network is warranted. Our analysis is primarily positive in nature, focusing on shock transmission and propagation. For this purpose, whether the information network is taken as exogenously given or endogenously chosen is of secondary importance. However, for policy analysis or for studying the effects of changes in market structure, it is necessary to internalize the endogenous formation of the information network.

**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 directed graph and consider its in-degrees and out-degrees. The in-degrees and out-degrees correspond to 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 $\mathcal{D}_i = \left\{j \in \mathbf{Z} : A_{ij} = 1\right\}$.

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 the most attention from other firms. Meanwhile, $d_1 = 4$, which implies that firm 1 pays the 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 $\xi_t$. The collection of productivities observed by firm $i$, denoted as $\mathcal{I}_{it} = \{z_{jt} : j \in \mathcal{D}_i\}$, is informative about $\xi_t$ and effectively provides $d_i$ signals regarding $\xi_t$. 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, $\lambda_i$ 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 $\lambda_i$, 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 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. 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 change the production efficiency; (2) the information channel: $z_{it}$ is informative about the aggregate component $\xi_t$ 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

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

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 $\lambda_i$ for firm $i$'s own inference, while the in-degree $m_i$ determines the frequency with which firm $i$'s productivity is used in 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 play a more important role: first, when other firms form expectations about the salient firms, their productivities will be directly utilized; second, when inferring unobserved productivities, the salient firms' productivities will be indirectly utilized as informative signals.

The remaining terms in condition (4.1) are related to higher-order expectation 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 the 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 $\mathbf{T}_1$ captures 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 processes. The expansion of condition (4.2) uncovers these effects

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

In this expansion, the first term $\frac{\delta}{N}\mathbf{I}$ corresponds 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 ($\delta$) and the GE consideration ($\alpha$) is irrelevant.

In the canonical 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, $\delta$ and $\alpha$ start to play different roles 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 $j \in \mathcal{M}_i$, 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 $\lambda_i$. 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 on GDP elasticity with respect 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 $\mu_i$ succinctly summarizes these forces and is the key statistic to estimate an individual firm's quantitative importance. To see how to obtain the bounds, one needs to leverage the expansion (4.3). The lower bound includes only the two terms—the shock itself and the first-order expectations—-and bounds it from below. In contrast, the upper bound includes all the terms in the expansion and bounds it from above. In addition, the ratio between the upper bound and the lower bound of $g_i$ is a constant independent of $N$.[10]

A notable 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 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.

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

### 4.3 Granularity

Equipped with the bound estimates, we now 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* $\max_i \mu_i = \Theta(N)$.[11]

Proposition 4.3 establishes that a necessary and sufficient condition for the economy to exhibit granular effects is that 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.** *The diluted in-degree index $\mu_i$ satisfies the following properties:*

1. *The sum of diluted indices equals $N$,*
$$\sum_{i=1}^{N} \mu_i = N,$$
*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 $\max_i m_i = \Theta(N)$.*

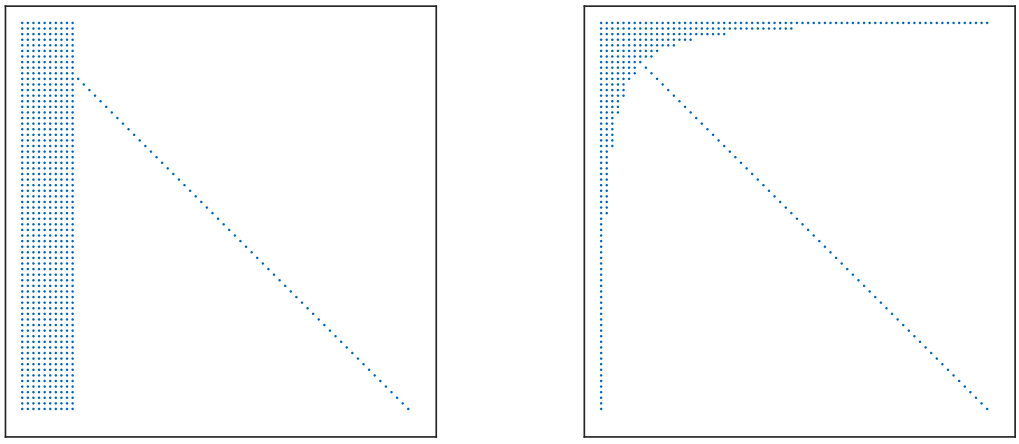
Part 1 of Corollary 4.1 implies that $\mu_i$ is bounded from above by $N$, and there can only be a small number of firms (relative to the size of the economy) that are able to generate granular effects.

Part 2 underscores the importance of a large in-degree for the granular effect, which has to grow at the same rate as $N$. 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 observe a fixed fraction $\omega$ of firms' productivities in the economy, in addition to their own. The left panel of Figure 5 displays the corresponding information matrix. Note that for the first $\omega N$ firms, they receive attention from all the firms, which implies their in-degree satisfies $m_i = N$. On the other hand, the out-degrees in this economy are given by $d_i \approx \omega N$.

[11] 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 $c_1 q(N) \leq f(N) \leq c_2 q(N)$ for all $N$ large enough.

Figure 5: 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 Zipf's law.

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

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

Clearly, the diluted in-degree index is not growing with $N$. Though the first $\omega N$ 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

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

which is visualized by the right panel of Figure 5. 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

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

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 $\zeta_i$ 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 a 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 $\omega_i$ is given by*

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

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 demand and supply forces is less important.[12]

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 elasticities 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*

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

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

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

which capture the general equilibrium forces via direct and indirect inference problems, respectively. The product 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.

[12] Our baseline analysis abstracts from the possibility that large firms internalize their potential market power. Allowing for this channel would substantially complicate the theoretical analysis in this section, and it will not likely alter the theoretical result as the relative size of any individual firm tends to converge to zero; nevertheless, its quantitative implications can be explored numerically.

**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 the index's 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 $\omega_i$ tends to converge to zero, causing the influence index $\mu_i^{\omega}$ 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,b), which is a generalization of the classical Erdos-Renyi model of a random graph. Instead of assuming a deterministic $\mathbf{A}$, information linkages among different firms

are now stochastic. Specifically, firm $i$ observes 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 $\{p_{ij}\}$ such that edges are independently assigned to each pair of vertices (firms) $(i, j)$. We maintain the assumption that each firm observes its own productivity with certainty such that (5.3) holds for any realization of network structure $\mathbf{A}$.[13]

The random-graph approach accommodates stochastic realizations of the information network structure and is naturally connected to our quantitative strategy of identifying browsing volumes on EDGAR with probabilities of information revelation in Section 6.2. At the same time, it preserves the central feature of our theoretical framework, namely that the expected in-degrees follow a power-law distribution. This property is stated formally in the following lemma. [14]

**Lemma 5.2.** *The expected in-degrees of the random graph are given by*

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

The details of the expected out-degree are spelled out in Appendix S.4. In a nutshell, when $N$ is sufficiently large, the expected out-degree can be viewed as uniformly distributed.

**Firm size.** Without loss of generality, we index firms so that their expected in-degrees are ordered by firm index. Importantly, this indexing need not coincide with the ranking of firms by size. We define the following permutation that connects the two orderings.

**Definition 3.** *A permutation $\pi$ is a bijective map that assigns a size ordering $j$ for each firm $i$ with corresponding expected in-degree $\mathbb{E}[m_i]$,*

$$\pi: i\in[1,2,\ldots,N]\longmapsto j\in[1,2,\ldots,N]. \tag{5.7}$$

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

$$\omega_{\pi(i)}=\frac{\pi(i)^{-\frac{1}{\beta_\omega}}}{\sum_{k=1}^{N}k^{-\frac{1}{\beta_\omega}}} \qquad \text{such that} \qquad \sum_{j=1}^{N}\omega_j=1. \tag{5.8}$$

[13] When $N$ is sufficiently large, the off-diagonal probabilities are given by $p_{ij}\approx\frac{1}{(j)^{\frac{1}{\beta_m}}}$. More details on this formulation are provided in Online Supplementary Appendix S.4.

[14] Our random-graph approach also shares a similar spirit with the rational inattention literature, such as Matějka and McKay (2015) and Maćkowiak et al. (2023), where actions and signals are modeled as stochastic choices. In the quantitative model, we adopt a logistic specification for the generation of probabilities, in a manner closely related to formulations used in the rational inattention literature.

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

A sequence of variables $\{x_k\}_{k\in\{1,2,\dots,N\}}$ satisfy power-law relations if $x_k = ck^{-\frac{1}{\kappa}}$ for some $\kappa > 0$. By construction, both the expected in-degree $\mathbb{E}[m_i]$ and the firm size $\omega_i$ are power-law sequences with exponents $\beta_m$ and $\beta_\omega$, respectively. Note that our setup is closely related to Gabaix (2011). In Gabaix (2011), firm sizes are assumed to be random variables drawn from a Pareto distribution. The following lemma then establishes the relationship between 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.

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

$$\mathbb{P}(X > x) = \frac{c^\kappa}{N} \cdot x^{-\kappa}, \tag{5.9}$$

*defined on support $x \in \left[\frac{c}{N^{\frac{1}{\kappa}}}, \infty\right]$ with $\kappa \geq 0$.*

1. *When $k$ is sufficiently large, the $k$th-order statistic admits an asymptotic expansion*

$$x_{(k)} \approx \mathbb{E}\left[x_{(k)}\right] = ck^{-1/\kappa} + o\left(k^{-1/\kappa}\right), \quad a.s. \quad k \to \infty. \tag{5.10}$$

2. *When $N \to \infty$, if $\kappa > 1$, $\mathbb{E}\left[x_{(k)}\right] = \Theta\left(ck^{-\frac{1}{\kappa}}\right)$, $\forall k \geq 1$.*

**GDP volatility.** For our purposes, the key primitives of the economy are given by the distributions of $m_i$ and $\omega_{\pi(i)}$, along with the ordering function $\pi(i)$, 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.

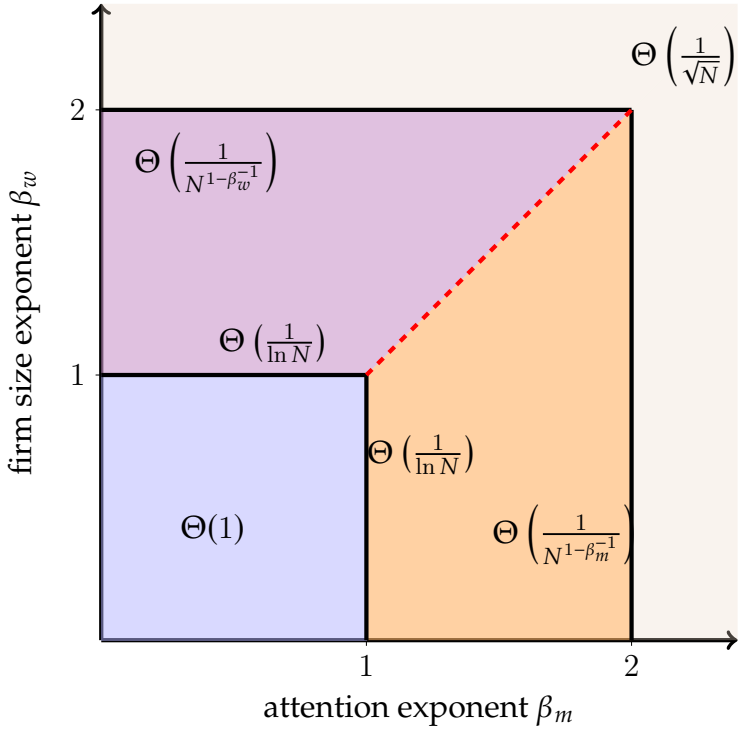
**Proposition 5.3.** *Denote $\beta \equiv \min\{\beta_m, \beta_\omega\}$. Independent of the ordering of $\{\pi(i)\}$, as $N \to \infty$, the decay rate of GDP volatility conditional on micro shocks follows*

1. $\sigma_{GDP} = \Theta\left(\frac{1}{\sqrt{N}}\right)$ *when* $\beta > 2$, *a.s.*,
2. $\sigma_{GDP} = \Theta\left(\frac{\sqrt{\ln N}}{\sqrt{N}}\right)$ *when* $\beta = 2$, *a.s.*,
3. $\sigma_{GDP} = \Theta\left(\frac{1}{N^{1-\beta^{-1}}}\right)$ *when* $1 < \beta < 2$, *w.h.p.*,
4. $\sigma_{GDP} \geq \Theta\left(\frac{1}{\ln N}\right)$ *when* $\beta = 1$, *w.h.p.*,
5. $\sigma_{GDP} = \Theta_p(1)$ *when* $\beta < 1$, *bounded in probability.*[15]

[15] See Appendix D.1 for the definitions of "w.h.p." (with high probability) and "bounded in probability".

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. If $\beta_m > 2$, 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) where the firm-size channel dominates. If the opposite applies, then the stochastic properties of the economy will be determined by the salient firms instead. Figure 6 then visualizes Proposition 5.3.

Figure 6: Decay Rates of GDP Volatility with Power Law Distributions




Importantly, the result in Proposition 5.3 does not hinge on the correlation between the two distributions. Whether attention is concentrated on large firms or on small firms is irrelevant for the decay rate of GDP volatility. This does not imply, however, that the correlation between the two distributions is quantitatively unimportant. To the contrary, it is a key determinant of the magnitude of granular effects, especially in finite samples and when the two tail exponents are similar, as suggested by the data.

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

**Proposition 5.4.** *Assume $\beta = \min\{\beta_\omega, \beta_m\} > 1$. Up to a leading-order approximation, the expected volatility of GDP conditional on micro shocks satisfies the following properties when $N \to \infty$:*

1. $\mathbb{E}[\sigma_{GDP}]$ *is maximized when the ordering of* $\{m_i\}$ *coincides with that of* $\{\omega_{\pi(i)}\}$.

2. $\mathbb{E}[\sigma_{GDP}]$ *is minimized when the ordering of* $\{m_i\}$ *is the exact opposite of* $\{\omega_{\pi(i)}\}$.

3. $\mathbb{E}[\sigma_{GDP}]$ *decreases as the number of inversions between* $\{m_i\}$ *and* $\{\omega_{\pi(i)}\}$ *increases.*[16]

*The expectation is defined with respect to the random graph* $\mathbf{A}$.[17]

Proposition 5.4 implies that the contribution of micro shocks to aggregate fluctuations depends on the joint distribution of firm size and attention. On the one hand, asymmetric attention weakens the traditional firm-size channel, since fewer firms observe productivity changes at large firms. On the other hand, for large firms that receive substantial attention, the attention channel amplifies other firms' responses to those productivity changes. The quantitative importance of micro shocks therefore depends critically on how these two attributes are matched, a question we examine in the next section. We also show there that the same insight extends to the case $\beta \leq 1$ and to the exact general equilibrium effects summarized by $(\mathbf{I} - \mathbf{T}_1 - \mathbf{T}_2)^{-1}$, rather than its first-order approximation $\mathbf{I} + \mathbf{T}_1 + \mathbf{T}_2$.

**Remark on proofs.** In Online Supplementary Appendix S.6 - S.9, we present detailed analyses and proofs for Proposition 5.3 and 5.4. The key strategy of our analysis is to propose a set of regularity conditions on the structure of the information network $\mathbf{A}$ and the distribution of firm size $\boldsymbol{\omega}$. These conditions are consistent with our previous insights on the diluted in-degree index, and are satisfied probabilistically in our stochastic power-law model. They are then used to estimate the GDP volatility as well as the impact of the interactions between size and salience on aggregate volatility.

From a methodological perspective, we also introduce a set of mathematical tools for the proofs of Proposition 5.3 and 5.4, such as the inhomogeneous Erdos-Renyi random graph theory, rearrangement inequalities in combinatorics, and asymptotic approximations of the Riemann zeta function. We believe these tools may be useful more broadly for analyzing models with complex information structures.

# 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. Finally, we conduct additional robustness tests 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

[16] The number of inversions is equal to the number of pairs $i < j$ such that $\left(\omega_{\pi(i)} - \omega_{\pi(j)}\right) \cdot (m_i - m_j) < 0$.

[17] In Online Supplementary Appendix S.9, we present two extensions that sharpen Proposition 5.4 with almost-sure statements of volatility comparisons. The first extension imposes additional restrictions on $(\beta_m, \beta_\omega)$. The second extension considers aggregate volatility that is driven by a fraction of firms receiving the largest attention. When $N \to \infty$, the number of these high-salience firms also grows to infinity.

the bilateral browsing matrix constructed from EDGAR, we denote $\widetilde{m}_i$ 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. We define the following linear predictor of browsing incidence between firm $i$ and firm $j$ as

$$f_{ij} = \Gamma_0 + \Gamma_1 \log \widetilde{m}_j + \Gamma_2 \log\left(\varrho_0 + \varrho_1 \widetilde{d}_i\right). \tag{6.1}$$

In the model economy, $A_{ij}$ can either be 0 or 1, and we specify the probability that $A_{ij} = 1$ through a standard logistic probability transformation

$$\Pr\left(A_{ij} = 1\right) = \frac{1}{1 + \exp\left(-f_{ij}\right)}.$$

This specification captures the idea that (1) if firm $j$ receives more browsing requests in EDGAR, it is more likely that firm $i$ observes $j$; (2) if firm $i$ sends more browsing requests in EDGAR, it is more likely that firm $i$ will also observe firm $j$. The parameter $\varrho_1$ further influences the correlations between in-degrees and out-degrees of the random graph.

We calibrate the parameters in equation (6.1) to jointly match key moments of the random graph that represents the information matrix: the estimated power-law exponents of the in-degree and out-degree distributions, as well as the correlation between in-degree and out-degree among the top 350 largest firms. We also target the forecast dispersion and the mean absolute forecast errors for firms' GDP forecasts in the Livingston data, to ensure that the overall degree of informational frictions in the model is comparable to that observed in the data. Table 4 shows that the calibrated model reproduces these moments closely.

In Section 6.3, we consider an alternative calibration strategy in which bilateral browsing intensities are mapped into firms' signal precisions about one another's fundamentals. The results are similar to those under our baseline calibration.

**Other structural parameters.** The structural parameters are calibrated based on standard values in the literature. Specifically, we set the Frisch elasticity of labor supply to $\nu = 2$, which lies in the middle of available micro and macro estimates. We set the elasticity of substitution across goods to $\rho = 3.5$, implying a steady-state markup of 1.4, within the range of estimates in Edmond et al. (2023) and Eeckhout et al. (2020). Taken together, these parameters imply a degree of strategic complementarity of $\alpha = 0.36$ in the best-response function and a partial-equilibrium response to firm-level productivity shocks of $\delta = 1.9$.

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

Table 3: Parameterization

| Param. | Value | Data Source | Empirical Target |
|---|---|---|---|
| Exogenously determined parameters | | | |
| $\nu$ | 2 | — | Frisch elasticity |
| $\rho$ | 3.5 | — | markup |
| $\{\omega_i\}_{i=1}^N$ | — | Compustat | sales' share |
| $\overline{\sigma_i}$ | 0.176 | Compustat | weighted average of idiosyncratic shock volatility |
| $\{\sigma_i\}_{i=1}^N$ | — | Compustat | idiosyncratic productivity shock volatility |
| $\sigma_\xi$ | 0.02 | San Francisco Fed | aggregate TFP shock volatility |
| Endogenously determined parameters | | | |
| $\Gamma_0$ | -29.20 | Livingston | abs forecast error, mean |
| $\Gamma_1$ | 2.01 | EDGAR | in-degree power-law coefficient |
| $\Gamma_2$ | 3.6 | EDGAR | out-degree power-law coefficient |
| $\varrho_1$ | 0.00011 | EDGAR | corr between out-degree and in-degree (top 350 firms) |
| $\varrho_0$ | 2.40 | Livingston | forecast, dispersion |

Table 4: Targeted Moments

| Empirical Target | Data | Model |
|---|---|---|
| in-degree power-law coefficient | -1.00 | -1.00 |
| out-degree power-law coefficient | -0.92 | -0.92 |
| corr between out-degree and in-degree top 350 firms | 0.10 | 0.12 |
| abs forecast error, mean | 0.025 | 0.032 |
| forecast, dispersion | 0.018 | 0.019 |

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.[18] 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.

## 6.2 Results

Table 5 summarizes the comparison across various model economies. For the information network, we consider three versions: the heterogeneous-information model corresponds to our baseline calibration that targets the EDGAR browsing pattern; the perfect-information economy corresponds to

[18] 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.

the one with $A_{ij} = 1$ for all $i$ and $j$; the Lucas-island economy corresponds to the canonical 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 5: Aggregate and Idiosyncratic Volatility

| | Size and Volatility | | |
|---|---|---|---|
| | equal size + equal volatility | hetero. size + equal volatility | hetero. size + hetero. volatility |
| **Heterogeneous Info.** | | | |
| total volatility | 4.63 | 5.13 | 5.57 |
| top-100 firms/total | 15.98% | 40.96% | 44.34% |
| top-100 firms | 0.74 | 2.10 | 2.47 |
| top-100-sales firms | - | 2.01 | 2.23 |
| top-100-attention firms | 0.74 | 1.22 | 1.31 |
| **Perfect Info.** | | | |
| total volatility | 6.36 | 7.03 | 7.29 |
| top-100 firms/total | 2.20% | 41.96% | 48.42% |
| top-100 firms | 0.14 | 2.95 | 3.55 |
| top-100-sales firms | - | 2.95 | 3.32 |
| top-100-attention firms | - | 1.19 | 1.35 |
| **Lucas Island** | | | |
| total volatility | 4.07 | 4.51 | 4.71 |
| top-100 firms/total | 2.21% | 41.91% | 48.41% |
| top-100 firms | 0.09 | 1.89 | 2.28 |
| top-100-sales firms | - | 1.89 | 2.13 |
| top-100-attention firms | - | 0.76 | 0.87 |

*Notes*: "—" indicates missing or unavailable data. Values of volatilities are multiplied by 100.

In the economy without heterogeneity in firm sizes and shock volatilities (the first column of Table 5), we isolate the role of the information network. 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 0.74, which is approximately five times higher than in the perfect-information economy and eight times higher than in the Lucas-island economy.[19] In addition, the GDP volatility attributable to the top-100 firms is roughly $16\%$, substantially larger than $2\%$ in the other two benchmark settings. Unsurprisingly, these influential firms correspond to the "top-100-attention firms" that receive the most attention from others in the network.

When firm size heterogeneity is introduced, the second column of Table 5 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

[19] 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.

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 substantially 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. Later in this section, we show that this pattern is largely attributable to the weak correlation between firm size and the attention firms receive.

**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

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

The $k-$th term on the right-hand side of this equation can be interpreted as the $k-$th order network effects, which are related to the $(k-1)$-th order of higher-order expectations. One mechanical force that applies to all 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, which involves only the expectations.

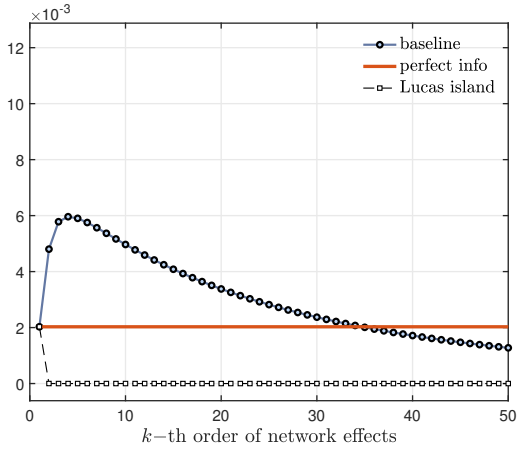
The introduction of a heterogeneous information network generates two countervailing forces relative to the perfect information benchmark. On one hand, limited information tends to attenuate agents' responsiveness to underlying fundamentals, as uncertainty dampens their reactions (Angeletos and Lian, 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 7 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 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 $k = 1$.

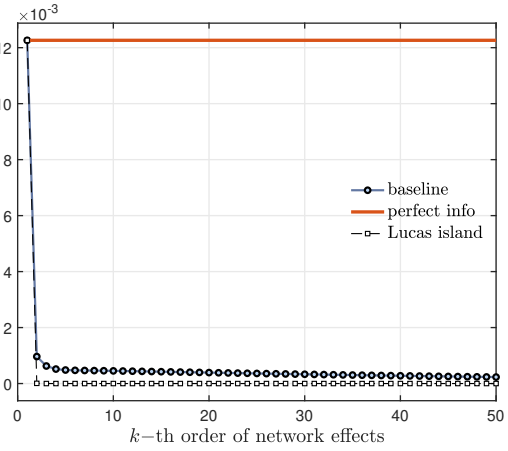
In our baseline economy, the "top-100-attention firms" attract significant attention and function

Figure 7: Role of Top Firms in Higher-Order Network Effects

(a) variance of top-100 attention firms




(b) variance of top-100 sales firms




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 5.

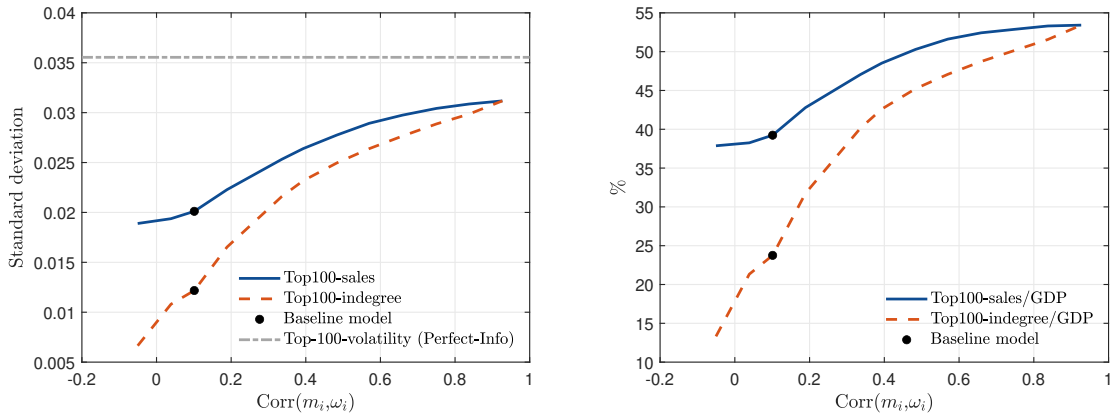
The right panel of Figure 7 illustrates the variance of $k$-th order network effects attributed to the top 100 firms ranked by their sales shares. Mechanically, since these firms are larger than the "top-100-attention firms", the network effect at $k = 1$ is correspondingly higher. In contrast to the 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. We further investigate the role of this correlation in the next exercise.

**Interaction: size and attention.** As suggested theoretically in Section 5.2, the sorting pattern between the firm size distribution and the attention distribution matters quantitatively for granular effects. The granular effects become stronger as firms' received attention and size become more positively assortative. To investigate this channel in our quantitative model, we fix the information matrix—and thus the attention distribution—while reshuffling firm sizes to generate different correlations between firm size $\omega_i$ and received attention $m_i$.

Figure 8 shows how the standard deviation of GDP attributable to idiosyncratic shocks to the top 100 firms (both in terms of sales and in terms of attention) varies with the correlation between the two dimensions, $\text{Corr}(m_i, \omega_i)$. In the baseline calibration, this correlation is approximately 0.15. As

the correlation increases, the contribution of top firms' shocks to GDP volatility rises substantially. If the two distributions were perfectly correlated, the volatility contribution would increase threefold. In this scenario, salience effects due to attention concentration almost offset the dampening effects arising from informational frictions, leaving the contribution of the top 100 firms by sales comparable to that under perfect information.

Figure 8: Assortative Matching between $m_i$ and $\omega_i$




**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:

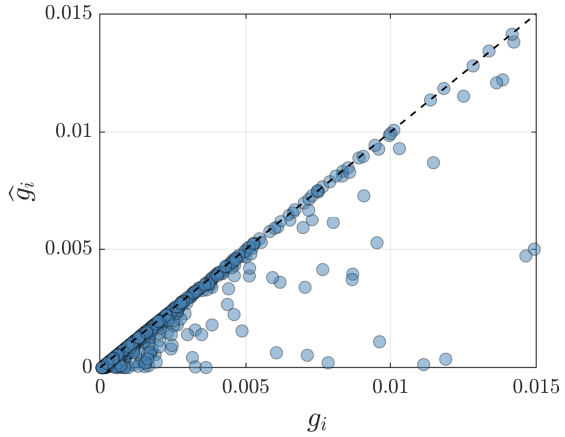
$$g_i = \beta_0 + \beta_1 \omega_i + \beta_2 m_i + \varepsilon_i, \tag{6.2}$$

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

Panel A in Table 6 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 that eliminates informational frictions while preserving the other sources of heterogeneity. In our baseline model, the regression coefficients on both $m_i$ and $\omega_i$ are positive and significant. The $R^2$ is 0.84. 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— a 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 $\omega_i$ as in condition (5.2). As a result, there is no variation in $m_i$, and $\omega_i$ accounts for all the variations in $g_i$.

Figure 9: Decomposition of $g_i$ between Size and Attention

(a) predicted $g_i$ with only size




(b) predicted $g_i$ with size and attention

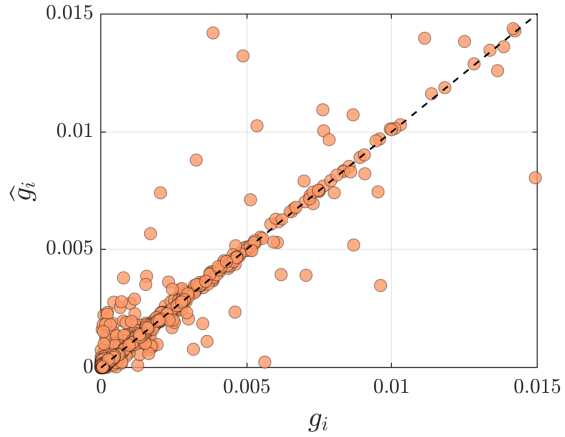



Table 6: Decomposition of $g_i$ between Size and Attention

| | Baseline | | | Perfect Info | | |
|---|---|---|---|---|---|---|
| | $m_i$ | $\omega_i$ | | $m_i$ | $\omega_i$ | |
| **Panel A** | | | Total | | | Total |
| $R^2$ | | | 0.84 | | | 1.00 |
| $\beta$ | 0.52*** | 2.01*** | | — | 3.00*** | |
| **Panel B: Explanatory Power** | | | | | | |
| $R^2$(incremental inclusion) | 28.9% | 71.1% | 100% | — | 100% | 100% |
| $R^2$(independent inclusion) | 26.4% | 73.6% | 100% | — | 100% | 100% |

Panel B quantifies the relative contributions of attention and firm size in explaining the variation in $g_i$. We adopt two complementary approaches. First, we evaluate the incremental explanatory power of each component by sequentially adding $m_i$ and $\omega_i$ to a baseline regression that includes only a constant, as in equation (6.2), and compute the corresponding increase in $R^2$.[20] Second, we estimate separate regressions including each component individually and report the resulting $R^2$ relative to the baseline specification. Across both approaches, we find that attention accounts for approximately 30% and firm size for 70% 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 visualizes their relative importance. In the left panel, when using the estimated 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, including both the size and the attention channel brings the predicted GDP elasticities much closer to the actual ones.

[20] 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.

**Explaining macroeconomic forecasts.** We revisit Fact 3 in Section 2, which shows that the attention-weighted measure of firm sales growth has substantial predictive power for forecasts of aggregate variables and outperforms the traditional sales-weighted measure. We now ask whether the same pattern arises in the model.

Using model-generated data, we estimate the following specification:

$$y_t = \beta_0 + \beta_1 G_t^{\text{attention}} + \beta_2 G_t^{\text{sales}} + \varepsilon_t, \tag{6.3}$$

where $y_t = N^{-1} \sum_{i=1}^{N} \mathbb{B}_{it}[q_t]$ is the average forecast across firms about current aggregate output. The variable $G_t^{\text{attention}}$ is the attention-weighted measure of sales growth among the top 100 firms by attention received, while $G_t^{\text{sales}}$ is the size-weighted measure of sales growth among the top 100 firms by sales. To facilitate a direct comparison, we also estimate the analogous regression using forecasts from the Survey of Professional Forecasters.

Table 7: Explaining GDP Forecasts

| | Baseline Model | | | Data | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Browsing-weighted | 0.036*** | | 0.037*** | 0.022*** | | 0.022*** |
| | (0.002) | | (0.006) | (0.006) | | (0.006) |
| Sales-weighted | | 0.003 | -0.004 | | 0.012 | 0.012 |
| | | (0.003) | (0.003) | | (0.015) | (0.013) |
| $R^2$ | 0.200 | 0.000 | 0.200 | 0.241 | 0.014 | 0.255 |

*Notes*: In the baseline-model columns, the dependent variable is $y_t = N^{-1} \sum_{i=1}^{N} \mathbf{E}_{it}[q_t]$, the average forecast across firms about current aggregate output. In the data columns, the dependent variable is the mean SPF forecast of GDP growth. The weighted sales-growth measures are constructed as $G_t = \sum_{i=1}^{100} \omega_{i,t-1}(g_{i,t} - \overline{g}_t)$, where $g_{i,t}$ is firm $i$'s log sales growth and $\omega_{i,t-1}$ is either its attention weight among the top 100 firms by attention received or its sales weight among the top 100 firms by sales. Since the baseline model contains substantially fewer firms than the empirical economy, GDP growth is more volatile in the model. We therefore rescale $y_t$ so that its standard deviation matches that of the corresponding empirical series. Standard errors are reported in parentheses. *** $p < 0.001$.

Table 7 reports the regression results, together with the corresponding empirical estimates. The model reproduces the central empirical pattern. The browsing-weighted measure of sales growth is strongly associated with GDP forecasts. By contrast, the sales-weighted measure has little explanatory power. Quantitatively, the model also generates an $R^2$ close to that observed in the data.

## 6.3 Robustness

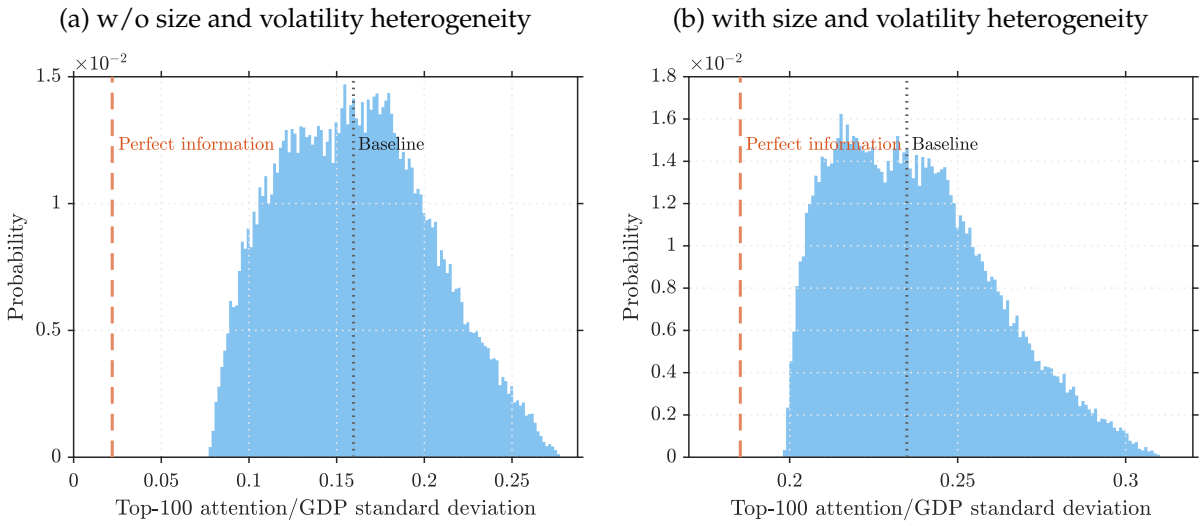
In this subsection, we explore the robustness of the quantitative results.

**Varying parameter values.** The Frisch elasticity $\nu$ and the elasticity of substitution $\rho$ influence equilibrium outcomes only through their effect on strategic complementarity, $\alpha$. To evaluate the robustness of our quantitative findings, we draw 100,000 independent parameter combinations from uni-

form distributions over $\nu \in [1, 3]$ and $\rho \in [2.5, 4.5]$.

Figure 10 reports the distribution of model-implied moments across these parameter draws. The left panel reports results for the specification without heterogeneous firm sizes and volatilities (Column 1 of Table 5), and the right panel reports results for the specification with heterogeneous firm sizes and volatilities (Column 3 of Table 5). Note that under perfect information, the relative contribution of the top 100 firms by attention to GDP volatility is independent of the value of $\alpha$. The vertical dotted lines mark the corresponding baseline values, while the red dashed line represents the values from the perfect-information model. Across most parameter draws, the top 100 firms by attention remain important for aggregate fluctuations, and contribute much more than in the perfect-information benchmark. The distance between the baseline and the perfect-information benchmark is smaller in the right panel due to the presence of heterogeneous firm sizes: some of these firms are also large, and their granular effects need not operate through the attention channel alone.

Figure 10: Robustness on Parameterization




**Adding public signals.** In the baseline model, firms learn about the aggregate component, $\xi_t$, solely through their observations of other firms' productivities. To allow for the possibility that firms also learn about $\xi_t$ from aggregate statistics, we extend the model to allow firms to observe an additional public signal

$$s_t = \xi_t + v_t, \qquad v_t \sim \mathbb{N}(0, \tau_v^{-1}).$$

Introducing a public signal reduces the effective prior variance of the common shock $\xi_t$ to $\widehat{\sigma}_\xi^2 = (\tau_\xi + \tau_v)^{-1}$. As expected, a lower prior variance of $\xi_t$ diminishes the informational and coordination roles played by the productivities of salient firms. Appendix Figure C.1 shows that the granular effects arising from heterogeneous attention remain quantitatively important over a relatively broad range of prior variances.

**Alternative information structure.** In the baseline calibration, the browsing intensity in the data is mapped to the probability that an element of the information matrix equals one. We now consider an alternative specification in which asymmetric attention is captured by heterogeneity in the precision of firms' private signals about other firms' fundamentals, rather than through the information matrix itself. Under this setup, bilateral browsing intensity maps directly into signal precision.

In particular, we consider an alternative specification in which, in each period, firm $i$ observes noisy signals about other firms' productivities:

$$x_{ijt} = z_{jt} + u_{ijt}, \qquad u_{ijt} \sim \mathbb{N}(0, \tau_{ij}^{-1}).$$

Here, asymmetric attention is captured by heterogeneity in signal precision rather than by the information matrix. A larger $\tau_{ij}$ means that firm $i$ observes firm $j$'s fundamental more precisely, consistent with more intensive browsing of firm $j$'s documents.

We parameterize $\tau_{ij}$ as

$$\tau_{ij} = (\widetilde{d}_i + \Gamma) \left( \frac{\widetilde{m}_j}{\max_j \widetilde{m}_j} \right)^{\chi}, \tag{6.4}$$

where $\widetilde{d}_i$ and $\widetilde{m}_j$ are, respectively, the numbers of browsing requests sent and received by a firm in EDGAR. Thus, signal precision rises with the browsing intensity of firm $i$ and the salience of firm $j$ in the browsing network. We choose the auxiliary parameters $\Gamma$ and $\chi$ to match the properties of Livingston forecasts as in the baseline calibration.

Table C.3 and Table C.4 in Appendix C.2 present the calibrated parameters and corresponding simulation results under the alternative calibration strategy. The main quantitative findings are largely unchanged: the attention channel alone continues to generate significant granular effects and remains quantitatively important once firm heterogeneity is introduced.

### 6.4 News-Based Measures of Attention

In this section, we provide an alternative calibration of the information network using firms' visibility in news media. While EDGAR browsing data provide a direct measure of firm-to-firm information acquisition, such data are not readily available for many researchers or policymakers who may want to implement the framework in other settings or in real time. By contrast, media-based measures of firm salience are widely available and can be constructed at high frequency across firms, countries, and time periods (Chahrour et al., 2021; Bui et al., 2024).

We measure firm $j$'s attention received, $\widetilde{m}_j$, by the number of Yahoo Finance news articles that mention firm $j$ during the sample period.[21] Since news data do not provide a natural measure of attention sent by each firm, we assume that attention sent is uniformly distributed. Thus, heterogeneity in the information network is driven entirely by heterogeneity in firms' media visibility. Following

[21] Please see Appendix A.1 for description of this data.

the random-graph approach, we specify

$$\Pr\big(A_{ij}=1\big)=\frac{1}{1+\exp\big(-(\Gamma_0+\Gamma_1\log\widetilde{m}_j)\big)}.$$

This specification preserves the key feature of the baseline information network: firms that receive more attention are more likely to be observed by others.

The details of the calibration are provided in Appendix C.1. Table 8 reports the quantitative results. The news-based calibration delivers results close to, and slightly stronger than, the baseline EDGAR calibration. When firm sizes and idiosyncratic shock volatilities are equalized, the top 100 firms account for 17.01 percent of aggregate volatility in the heterogeneous-information economy. With firm-size heterogeneity, the attention channel remains quantitatively important. In the heterogeneous information economy, the top 100 attention firms generate volatility of 1.33 when firm sizes are heterogeneous and idiosyncratic shock volatilities are equal, and 1.37 when both firm sizes and shock volatilities are heterogeneous.

Table 8: Aggregate and Idiosyncratic Volatility

| | Size and Volatility | | |
|---|---|---|---|
| | equal size + equal volatility | hetero. size + equal volatility | hetero. size + hetero. volatility |
| **Heterogeneous Info.** | | | |
| total volatility | 4.82 | 5.29 | 5.78 |
| top-100 firms/total | 17.01% | 40.83% | 43.08% |
| top-100 firms | 0.82 | 2.16 | 2.49 |
| top-100-sales firms | — | 2.06 | 2.27 |
| top-100-attention firms | 0.82 | 1.33 | 1.37 |
| **Perfect Info.** | | | |
| total volatility | 6.36 | 7.03 | 7.29 |
| top-100 firms/total | 2.20% | 41.96% | 48.42% |
| top-100 firms | 0.14 | 2.95 | 3.55 |
| top-100-sales firms | — | 2.95 | 3.32 |
| top-100-attention firms | — | 1.19 | 1.35 |

*Notes*: "—" indicates missing or unavailable data. Values of volatilities are multiplied by 100.

The somewhat larger effects relative to the EDGAR baseline reflect two forces highlighted by the theory. First, news-based attention received is more fat-tailed: its power-law coefficient is $-0.82$, compared with approximately $-1.00$ in the EDGAR calibration. A fatter tail concentrates visibility among a smaller set of firms and strengthens the role of salient firms as information nodes. Second, media visibility is more strongly correlated with firm size. The correlation between firm size and news-based attention received is $0.34$ (s.e. $0.04$), compared with $0.23$ (s.e. $0.02$) for EDGAR attention. This stronger alignment increases the interaction between the size channel and the attention channel, consistent with Proposition 5.4.

Overall, the exercise shows that the quantitative importance of the information channel is not specific to EDGAR browsing data. Media attention is less direct as a measure of firm-to-firm informa-

tion acquisition, but it is widely available and captures an economically relevant component of firm salience. The news-based calibration therefore provides a practical way to implement the framework when direct browsing data are unavailable.

## 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.

Afrouzi, H., B. Hèbert, J. La'O, and J. Leahy (2025). Sectoral salience. Technical report. mimeo.

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.

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.

Bergemann, D., T. Heumann, and S. Morris (2017). Information and interaction.

Billingsley, P. (1995). *Probability and Measure* (3 ed.). Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons.

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.

Carvalho, V. M. and B. Grassi (2019). Large firm dynamics and the business cycle. *American Economic Review 109*(4), 1375–1425.

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.

Cuciniello, V. and L. Paciello (2026). Who talks? disclosure and price inference in production networks. Working paper.

Dew-Becker, I. (2023). Tail risk in production networks. *Econometrica 91*(6), 2089–2123.

Di Giovanni, J. and A. A. Levchenko (2012). Country size, international trade, and aggregate fluctuations in granular economies. *Journal of Political Economy 120*(6), 1083–1132.

Di Giovanni, J. and A. A. Levchenko (2013). Firm entry, trade, and welfare in zipf's world. *Journal of International Economics 89*(2), 283–296.

Di Giovanni, J., A. A. Levchenko, and I. Mejean (2014). Firms, destinations, and aggregate fluctuations. *Econometrica 82*(4), 1303–1340.

Edmond, C., V. Midrigan, and D. Y. Xu (2023). How costly are markups? *Journal of Political Economy 131*(7), 1619–1675.

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*.

Flynn, J. P. and K. Sastry (2024). Attention cycles. Technical report, National Bureau of Economic Research.

Foerster, A. T., P.-D. G. Sarte, and M. W. Watson (2011). Sectoral versus aggregate shocks: A structural factor analysis of industrial production. *Journal of Political Economy 119*(1), 1–38.

Frieze, A. and M. Karoński (2026). *Introduction to Random Graphs*. Cambridge University Press.

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.

Golub, B. and S. Morris (2020). Expectations, networks, and conventions. *arXiv preprint arXiv:2009.13802*.

Hajdini, I., S. Kumar, S. Malik, J. J. Norris, and M. Pedemonte (2025). Supply chain networks and the macroeconomic expectations of firms. Technical report, IDB Working Paper Series.

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

Horvath, M. (1998). Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks. *Review of Economic Dynamics 1*(4), 781–808.

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.

Leister, C. M., Y. Zenou, and J. Zhou (2022). Social connectedness and local contagion. *The Review of Economic Studies 89*(1), 372–410.

Lorenzoni, G. (2009). A theory of demand shocks. *American economic review 99*(5), 2050–2084.

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

Maćkowiak, B., F. Matějka, and M. Wiederholt (2023). Rational inattention: A review. *Journal of Economic Literature 61*(1), 226–273.

Matějka, F. and A. McKay (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review 105*(1), 272–298.

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.

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

Song, W. and S. Stern (2025). Firm inattention and the efficacy of monetary policy: A text-based approach. *Review of Economic Studies 92*(5), 3438–3469.

Woodford, M. (2003). Imperfect common knowledge and the effects of monetary policy. *Knowledge, information, and expectations in modern macroeconomics: In honor of Edmund S. Phelps 25*(1), 4.

# Online Appendix

Micro Shocks and Macro Fluctuations in the Information Network

Zhen Huo Minghao Li Jieran Wu Huan Xiong

**A Data Appendix** **A-2**
A.1 Text-Based Attention Measure . . . A-2
A.2 Most Viewed Filings . . . A-2
A.3 Attention and Firm Characteristics . . . A-3

**B Additional Empirical Results** **A-5**
B.1 Attention Received and Sent Are Positively Correlated Only for Large Firms . . . A-5
B.2 Predicting Macroeconomic Forecasts: Additional Results . . . A-5
B.3 Explaining Current Macroeconomic Fluctuations . . . A-9
B.4 Robustness of the Power-Law Results . . . A-10

**C Additional Quantitative Results** **A-15**
C.1 Quantitative Results Using News Media . . . A-15
C.2 Alternative Information Structure . . . A-15
C.3 Adding Public Signals . . . A-16

**D Proofs of Results in Section 3 and Section 4** **A-17**
D.1 Asymptotic and Probability Notations . . . A-17
D.2 Proof of Lemma 3.1 . . . A-18
D.3 Proof of Proposition 4.1 . . . A-19
D.4 Proof of Proposition 4.2 . . . A-21
D.5 Proof of Proposition 4.3 and Corollary 4.1 . . . A-24

# A. Data Appendix

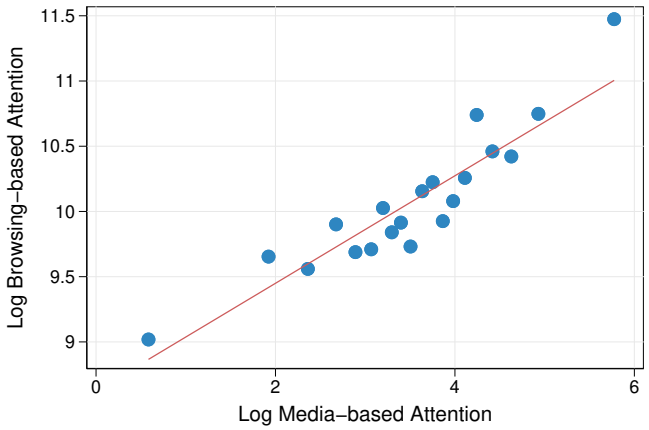
## A.1 Text-Based Attention Measure

This appendix constructs a news-based attention measure used to validate the EDGAR browsing measure in Section 2. We collect newswire-style articles from Yahoo Finance in 2016 and identify firm mentions using Named Entity Recognition (NER) from the Natural Language Toolkit (NLTK). We extract organization entities and match them to Compustat firms using standardized firm names, common aliases, and fuzzy string matching to account for naming variants such as "IBM" and "International Business Machines" or "AT&T" and "American Telephone and Telegraph." When an article mentions multiple firms, we retain all firm–article pairs.

For each firm $i$, text-based attention is defined as the total number of Yahoo Finance articles mentioning the firm in 2016. This measure captures the intensity of public media coverage and provides an external benchmark for our browsing-based measure of attention.

Figure A.1 compares text-based attention with browsing-based attention, measured by total EDGAR views received. Both variables are in logs. The cross-sectional correlation is 0.39, with a bootstrap standard error of 0.035. Thus, firms that receive more attention on EDGAR also tend to receive more attention in news media. The correlation is nevertheless far from one, consistent with the distinction between active information acquisition on EDGAR and broader public exposure through news coverage.

Figure A.1: Correlation Between Browsing-Based and Text-Based Attention




*Notes*: The figure presents a binscatter of log total EDGAR views received against log total Yahoo Finance news mentions for the cross-section of firms in 2016. The correlation is 0.39, with a bootstrap standard error of 0.035.

## A.2 Most Viewed Filings

Table A.1 reports the ten most accessed EDGAR form types in 2016. The three most viewed forms are 8-K, 10-Q, and 10-K, which together account for more than 2.9 billion requests, or about 70% of browsing volume among the top ten forms. This pattern indicates that EDGAR attention is concentrated on both timely event disclosures and periodic financial reports. Form descriptions follow the SEC's EDGAR documentation; see https://www.sec.gov/edgar/about.

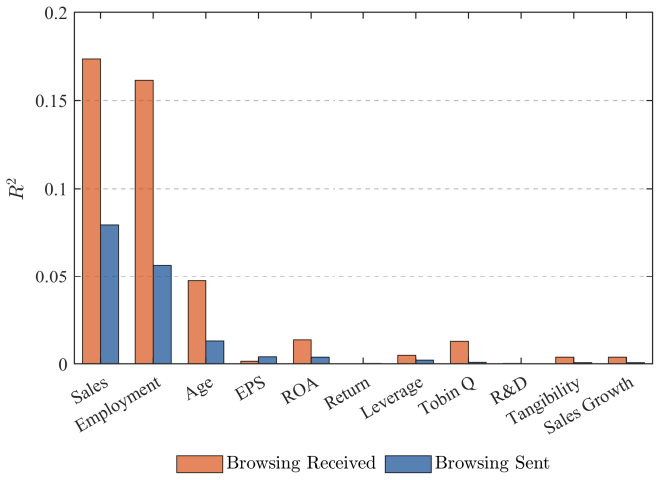
Table A.1: Most Viewed Filings

| Form | Total Browsing | Description |
|---|---|---|
| 8-K | 1,102,683,079 | Current report for material events or corporate changes |
| 10-Q | 1,034,609,909 | Quarterly financial report |
| 10-K | 778,512,251 | Annual report and comprehensive overview of the registrant |
| 424B2 | 272,580,533 | Prospectus supplement for securities offerings |
| 497K | 208,204,363 | Summary prospectus for investment companies |
| 6-K | 207,299,363 | Current report for certain foreign private issuers |
| 485 | 190,956,241 | Registration statement amendment for investment companies |
| 497 | 138,954,321 | Prospectus filing for investment companies |
| DEF14A | 93,240,818 | Definitive proxy statement |
| 485APOS | 66,608,633 | Post-effective amendment to a registration statement |

*Notes*: This table lists the ten EDGAR form types with the largest number of browsing requests in 2016.

## A.3 Attention and Firm Characteristics

Figure A.2: Explanatory Power of Firm Characteristics for Attention




This appendix relates firm-level attention to observable firm characteristics and motivates our use of sales as the baseline size measure in Fact 2. For each characteristic $x_{i,t}$, we first estimate

$$B_{i,t} = \alpha + \beta x_{i,t} + \varepsilon_{i,t},$$

separately for attention received, $B^R_{i,t}$, and attention sent, $B^S_{i,t}$. We consider sales, employment, age, earnings per share, return on assets, excess returns, leverage, Tobin's Q, R&D intensity, tangibility, and sales growth. All variables are winsorized at the 1st and 99th percentiles.

Figure A.2 reports the univariate $R^2$ values. Sales is the strongest predictor of both attention received and attention sent, with $R^2$ values of 0.17 and 0.08, respectively. Employment is the next strongest predictor, while all other characteristics explain substantially less variation.

Table A.2: Attention Received/Sent and Firm Characteristics

| | Attention Received | | Attention Sent | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Age | 0.036*** | 0.035*** | -0.054*** | -0.043*** |
| | (0.0043) | (0.0038) | (0.014) | (0.013) |
| Sales | 0.029*** | 0.029*** | 0.060*** | 0.058*** |
| | (0.0018) | (0.0019) | (0.0053) | (0.0054) |
| Earnings per Share | -0.0027*** | -0.0023*** | -0.0013 | -0.00092 |
| | (0.00080) | (0.00087) | (0.0018) | (0.0019) |
| ROA | -0.0013 | -0.00086 | 0.087*** | 0.095*** |
| | (0.0033) | (0.0030) | (0.031) | (0.030) |
| Excess Return | 0.064*** | 0.050*** | 0.029** | -0.0055 |
| | (0.0043) | (0.0045) | (0.012) | (0.012) |
| Leverage | 0.0020 | 0.0019 | 0.0026 | 0.0027 |
| | (0.0012) | (0.0012) | (0.0036) | (0.0038) |
| Tobin's Q | 0.00011 | 0.00020 | 0.0090 | 0.013** |
| | (0.00036) | (0.00031) | (0.0062) | (0.0066) |
| R&D Intensity | 0.16*** | 0.16*** | 0.44*** | 0.44*** |
| | (0.027) | (0.022) | (0.12) | (0.11) |
| Tangibility | -0.051*** | -0.058*** | -0.099** | -0.093** |
| | (0.013) | (0.012) | (0.041) | (0.040) |
| Sales Growth | 0.015** | 0.028*** | -0.055* | -0.018 |
| | (0.0070) | (0.0071) | (0.029) | (0.028) |
| L.Sales Growth | 0.028*** | 0.030*** | 0.018 | 0.0050 |
| | (0.0065) | (0.0066) | (0.026) | (0.026) |
| L.Attention Received | 0.90*** | 0.89*** | | |
| | (0.0033) | (0.0041) | | |
| L.Attention Sent | | | 0.85*** | 0.85*** |
| | | | (0.0045) | (0.0049) |
| Firm FE | ✓ | ✓ | ✓ | ✓ |
| Year FE | | ✓ | | ✓ |
| Adjusted $R^2$ | 0.84 | 0.85 | 0.75 | 0.76 |
| No. Observations | 50,379 | 50,379 | 19,425 | 19,425 |

*Notes*: The table reports panel regressions of attention received and attention sent on firm characteristics. Variables in logs are attention received, attention sent, sales, and age. All variables are winsorized at the 1st and 99th percentiles. Standard errors, clustered at the firm level, are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

We then estimate panel regressions of the form

$$B_{i,t} = \alpha_i + \alpha_t + \boldsymbol{\beta}' \boldsymbol{X}_{i,t} + \varepsilon_{i,t},$$

where $\alpha_i$ and $\alpha_t$ denote firm and year fixed effects, and $\boldsymbol{X}_{i,t}$ includes the full set of firm characteristics. Table A.2 reports the results. Sales remains positive and statistically significant for both attention received and attention sent. Attention is also highly persistent, with coefficients on lagged attention between 0.85 and 0.90. R&D intensity is positively associated with both measures, while age is positively associated with attention received but negatively associated with attention sent. Overall, sales is the most important observable correlate of attention, but the concentration of attention documented in Fact 1 is not mechanically explained by size alone.

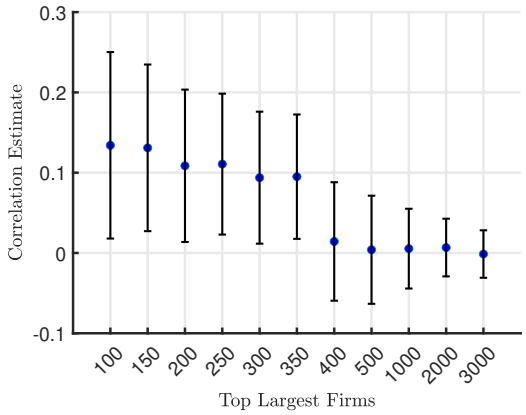
# B. Additional Empirical Results

## B.1 Attention Received and Sent Are Positively Correlated Only for Large Firms

This section examines whether firms that browse others more also receive more attention. Let $B_i^R$ and $B_i^S$ denote total browsing received and sent by firm $i$, respectively. Figure B.1 plots the correlation $\text{corr}(B_i^R, B_i^S)$ for the top $N$ firms by sales, with $N$ varying from 100 to 3,000, together with 90% confidence intervals.

The correlation is positive and statistically significant only among the largest firms. For samples up to roughly the top 350 firms by sales, the correlation reaches about 0.1. Once smaller firms are included, the correlation becomes economically small and statistically indistinguishable from zero. Thus, among the largest firms, attention is more bidirectional: firms that attract more attention also tend to acquire more information about others. Among smaller firms, by contrast, attention received and attention sent are largely unrelated.

Figure B.1: Correlation Between Attention Sent and Received, by Firm Size




*Notes*: The figure plots $\text{corr}(B_i^R, B_i^S)$ for the top $N$ firms by sales, where $N$ ranges from 100 to 3,000. Shaded areas denote 90% confidence intervals. The correlation is positive and statistically significant only for the largest firms, approximately the top 350 by sales.

## B.2 Predicting Macroeconomic Forecasts: Additional Results

This appendix reports additional forecasting results for Fact 3. The results show that the predictive power of attention-weighted firm growth is robust to using median rather than mean SPF forecasts, extends beyond GDP to other macroeconomic aggregates, and is driven by the combination of attention-based weights and the extreme visibility of the top attention-receiving firms.

### B.2.1 Median Forecasts

Table B.1 repeats the baseline forecasting exercise using the median SPF forecast of real GDP growth. The results are close to those in Table 2. Lagged attention-weighted growth remains positive and statistically significant in most specifications, while the sales-weighted measure is small and insignificant. The attention-weighted measure alone explains 56% of the variation in one-quarter-ahead median GDP forecasts, and the specification with macro controls reaches an $R^2$ of 0.85.

Table B.1: Predictive Power for Median GDP Forecasts

| | Median Forecast of Real GDP in $t+1$ | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $G_{t-1}^{\text{attention}}$ | 0.020*** | | | | 0.0096* | 0.0100* |
| | (0.0049) | | | | (0.0052) | (0.0052) |
| $G_{t-2}^{\text{attention}}$ | 0.016*** | | | | 0.0075 | 0.0071 |
| | (0.0051) | | | | (0.0051) | (0.0052) |
| $G_{t-3}^{\text{attention}}$ | 0.0092* | | | | 0.0099** | 0.013** |
| | (0.0050) | | | | (0.0045) | (0.0047) |
| $G_{t-4}^{\text{attention}}$ | 0.00062 | | | | 0.0056 | 0.0047 |
| | (0.0049) | | | | (0.0045) | (0.0045) |
| $G_{t-1}^{\text{sales}}$ | | -0.015 | | -0.0018 | | -0.0077 |
| | | (0.021) | | (0.019) | | (0.015) |
| $G_{t-2}^{\text{sales}}$ | | -0.022 | | -0.011 | | -0.010 |
| | | (0.024) | | (0.022) | | (0.017) |
| $G_{t-3}^{\text{sales}}$ | | 0.0057 | | -0.012 | | -0.017 |
| | | (0.023) | | (0.020) | | (0.016) |
| $G_{t-4}^{\text{sales}}$ | | 0.028 | | 0.00078 | | 0.015 |
| | | (0.020) | | (0.019) | | (0.015) |
| Macro controls | | | ✓ | ✓ | ✓ | ✓ |
| $R^2$ | 0.56 | 0.11 | 0.66 | 0.70 | 0.81 | 0.85 |
| Adjusted $R^2$ | 0.51 | 0.00 | 0.51 | 0.49 | 0.67 | 0.69 |
| No. Observations | 40 | 40 | 41 | 40 | 40 | 40 |

*Notes*: The dependent variable is the median SPF forecast of real GDP growth for the next quarter. Attention-weighted and sales-weighted growth are constructed as in Section 2. Macro controls include four lags each of oil supply shocks, monetary policy shocks, the 3-month T-bill rate, and the term spread. Sales growth is winsorized at the 1st and 99th percentiles. Sample: 2006:Q1–2016:Q4. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

### B.2.2 Other Macroeconomic Variables

Table B.2 shows that the predictive power of attention-weighted growth extends to industrial production, unemployment, consumption, and investment forecasts. The coefficients have the expected signs: positive for industrial production, consumption, and investment, and negative for unemployment. Across variables, attention-weighted growth explains 46–56% of the variation in one-quarter-ahead SPF forecasts, while the sales-weighted measure explains little.

Table B.2: Predictive Power for Other Macroeconomic Variables

| | Mean Forecasts in $t+1$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Industrial Production | | Unemployment | | Consumption | | Investment | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $G^{\text{attention}}_{t-1}$ | 0.028*** | | -0.010** | | 0.017*** | | 0.035* | |
| | (0.010) | | (0.0038) | | (0.0038) | | (0.019) | |
| $G^{\text{attention}}_{t-2}$ | 0.034*** | | -0.0096** | | 0.011*** | | 0.057*** | |
| | (0.010) | | (0.0040) | | (0.0039) | | (0.019) | |
| $G^{\text{attention}}_{t-3}$ | 0.019* | | -0.0073* | | 0.0052 | | 0.049** | |
| | (0.010) | | (0.0039) | | (0.0038) | | (0.019) | |
| $G^{\text{attention}}_{t-4}$ | -0.00088 | | -0.0064 | | 0.0039 | | 0.031* | |
| | (0.0099) | | (0.0038) | | (0.0037) | | (0.018) | |
| $G^{\text{sales}}_{t-1}$ | | 0.0065 | | 0.0088 | | -0.021 | | -0.026 |
| | | (0.039) | | (0.015) | | (0.016) | | (0.075) |
| $G^{\text{sales}}_{t-2}$ | | -0.032 | | 0.017 | | -0.011 | | -0.012 |
| | | (0.045) | | (0.017) | | (0.018) | | (0.086) |
| $G^{\text{sales}}_{t-3}$ | | -0.0060 | | -0.0074 | | 0.0044 | | -0.021 |
| | | (0.044) | | (0.017) | | (0.018) | | (0.085) |
| $G^{\text{sales}}_{t-4}$ | | 0.071* | | -0.0099 | | 0.013 | | 0.10 |
| | | (0.038) | | (0.014) | | (0.015) | | (0.073) |
| $R^2$ | 0.49 | 0.11 | 0.46 | 0.07 | 0.56 | 0.11 | 0.51 | 0.06 |
| Adjusted $R^2$ | 0.43 | 0.01 | 0.40 | -0.03 | 0.51 | 0.01 | 0.45 | -0.04 |
| No. Observations | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 |

*Notes*: The dependent variable is the mean SPF forecast for the indicated aggregate in quarter $t+1$. All specifications include four lags of either attention-weighted or sales-weighted growth, with no additional macro controls. Sales growth is winsorized at the 1st and 99th percentiles. Sample: 2006:Q1–2016:Q4. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

### B.2.3 Decomposing the Source of Predictive Power

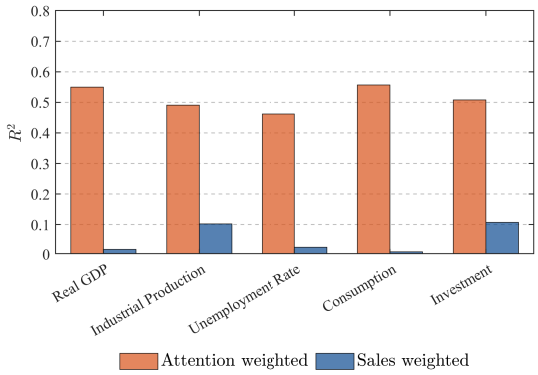
We next decompose whether the predictive power in Fact 3 comes from attention weights or from the composition of firms selected. Figure B.2 panel (a) compares two measures constructed from the same set of firms: the top 100 firms by attention received. The baseline measure weights these firms by attention, while the alternative weights them by sales. The attention-weighted measure consistently delivers much higher $R^2$ values, showing that the weights themselves are central to the forecasting result.

Figure B.2 panel (b) asks whether the result extends beyond the very top of the attention distribution. It constructs an attention-weighted measure using firms ranked 101–200 by attention received. Predictive power falls sharply relative to the baseline, indicating that the information content of attention-weighted growth is concentrated among the most visible firms.

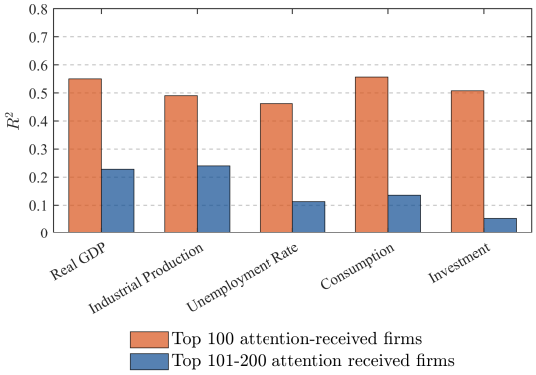
Together, these exercises show that the forecasting power of attention-weighted growth reflects both the attention weights and the selection of the most visible firms. Neither sales-weighting the top attention firms nor attention-weighting less visible firms generates comparable predictive power.

Figure B.2: Decomposition: Attention Weights and Firm Composition

(a) attention weights




(b) firm composition




*Notes*: Panel (a) compares attention and sales weights among the top 100 attention-receiving firms. Panel (b) compares attention-weighted growth for the top 100 firms and firms ranked 101–200 by attention received. Both panels report $R^2$ values from univariate regressions of one-quarter-ahead SPF forecasts.

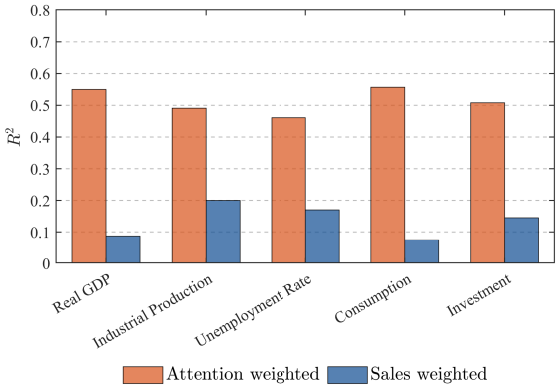
### B.2.4 Alternative Selection Criterion: Firms with Largest Absolute Sales Growth

Finally, we examine whether the result is driven by selecting firms with extreme sales growth rather than by selecting firms that receive extreme attention. For each quarter, we rank firms by absolute sales growth, select the top $K = 100$, and construct

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

where $\omega_{i,t}^{\text{sales}}$ is firm $i$'s sales share among the selected firms, $g_{i,t} = s_{i,t} - s_{i,t-4}$, and $\bar{g}_t$ is average sales growth across all firms. We then compare this measure with the baseline attention-weighted measure using the same specification as in Fact 3. Figure B.3 shows that selecting firms by extreme sales growth generates substantially weaker predictive power. Thus, the baseline result is not simply driven by extreme firm-level growth.

Figure B.3: Attention-Weighted versus Sales-Growth-Selected Measures




*Notes*: The figure reports $R^2$ values from univariate regressions of one-quarter-ahead SPF forecasts on attention-weighted growth and on an alternative measure that selects the top 100 firms by absolute sales growth and weights them by sales.

## B.3 Explaining Current Macroeconomic Fluctuations

This appendix examines whether attention-weighted firm growth also explains realized macroeconomic fluctuations. Following Gabaix (2011), we regress current-quarter growth rates of aggregate variables on contemporaneous and lagged attention-weighted and sales-weighted firm growth, with and without standard macroeconomic controls.

Table B.3 reports results for real GDP growth. The sales-weighted measure has meaningful explanatory power: used alone, it yields an $R^2$ of 0.26, consistent with the granular-residual evidence in Gabaix (2011). The attention-weighted measure performs better, with an $R^2$ of 0.45 when used alone. Its strongest coefficients appear at lags $t-2$ and $t-3$. When both growth measures and macro controls are included, the $R^2$ rises to 0.89, and attention-weighted growth remains significant. Thus, attention-weighted firm growth contains information about realized GDP growth beyond both standard macro controls and sales-weighted granular growth.

Table B.3: Explaining Growth Rates of Real GDP

| | GDP Growth Rate in Quarter $t$ | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $G_t^{\text{attention}}$ | 1.31 | | | | 0.74 | 0.57 |
| | (0.99) | | | | (1.61) | (1.40) |
| $G_{t-1}^{\text{attention}}$ | 1.80* | | | | 1.00 | 1.04 |
| | (1.03) | | | | (1.62) | (1.49) |
| $G_{t-2}^{\text{attention}}$ | 3.61*** | | | | 3.92** | 4.79*** |
| | (1.03) | | | | (1.45) | (1.24) |
| $G_{t-3}^{\text{attention}}$ | 4.33*** | | | | 6.05*** | 5.29*** |
| | (1.03) | | | | (1.32) | (1.39) |
| $G_t^{\text{sales}}$ | | 14.9*** | | 11.9*** | | 4.96 |
| | | (3.89) | | (3.28) | | (4.21) |
| $G_{t-1}^{\text{sales}}$ | | 14.8*** | | 12.9*** | | 12.2*** |
| | | (3.90) | | (3.35) | | (3.35) |
| $G_{t-2}^{\text{sales}}$ | | 8.57** | | 10.1*** | | 8.34** |
| | | (3.90) | | (3.69) | | (3.82) |
| $G_{t-3}^{\text{sales}}$ | | 4.76 | | 8.49** | | 2.23 |
| | | (3.82) | | (3.79) | | (4.03) |
| Macro controls | | | ✓ | ✓ | ✓ | ✓ |
| $R^2$ | 0.45 | 0.26 | 0.52 | 0.70 | 0.79 | 0.89 |
| Adjusted $R^2$ | 0.39 | 0.22 | 0.31 | 0.52 | 0.58 | 0.73 |
| No. Observations | 42 | 69 | 54 | 54 | 42 | 42 |

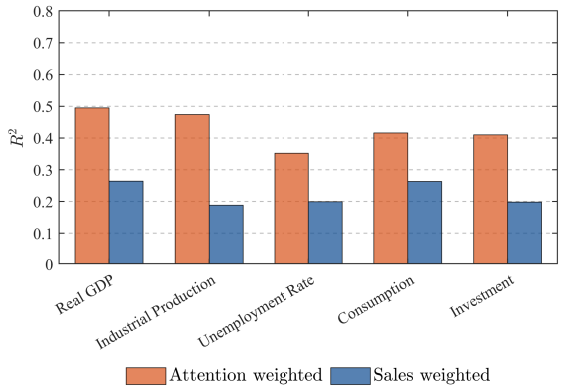
*Notes*: The dependent variable is quarterly real GDP growth. Macro controls include four lags each of oil supply shocks, monetary policy shocks, the 3-month T-bill rate, and the term spread. Sample periods vary with data availability for controls. Columns 1–2 use the full available sample, while columns 3–6 require complete control data. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure B.4 extends the analysis to industrial production, unemployment, consumption, and investment. For each variable, we report the $R^2$ from regressions using only attention-weighted growth or only sales-weighted growth, following the specification in column 1 of Table B.3. Sales-weighted growth explains a nontrivial share of aggregate fluctuations, with $R^2$ values between 0.19 and 0.27. Attention-weighted growth consistently de-

livers higher explanatory power, with $R^2$ values between 0.35 and 0.50.

Overall, attention-weighted firm growth helps explain realized macroeconomic fluctuations and does so more strongly than conventional sales-weighted granular growth. The sales-weighted measure performs in line with the granularity literature; the additional contribution is that attention weighting substantially increases explanatory power.

Figure B.4: Explaining Growth Rates of Macroeconomic Variables



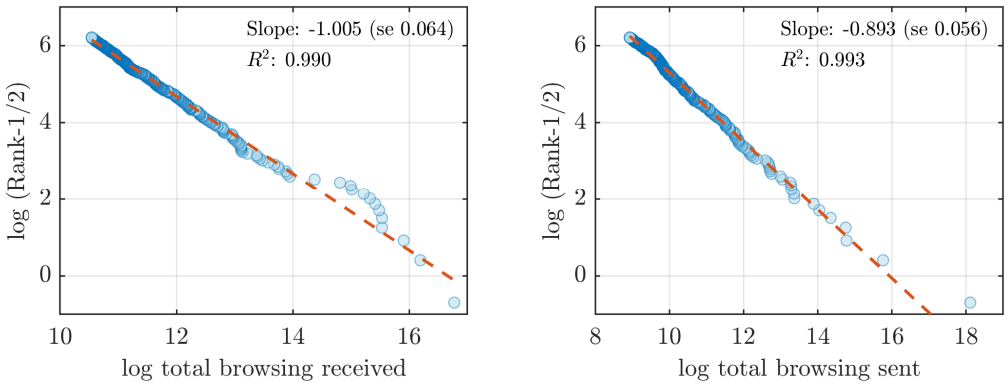

*Notes*: The figure reports $R^2$ values from regressions of current-quarter aggregate growth rates on contemporaneous and lagged attention-weighted or sales-weighted firm growth. The macroeconomic aggregates are real GDP, industrial production, unemployment, real personal consumption expenditures, and real nonresidential fixed investment.

## B.4 Robustness of the Power-Law Results

This appendix shows that the power-law patterns in Fact 1 are robust to alternative samples, normalizations, industries, and rank cutoffs. Across all specifications, the estimated exponent for attention received remains close to one, while the exponent for attention sent remains below one.

**Excluding financial firms.** Figure B.5 excludes financial firms, whose disclosure requirements and browsing patterns may differ from those of industrial firms. The estimated exponents are nearly unchanged.

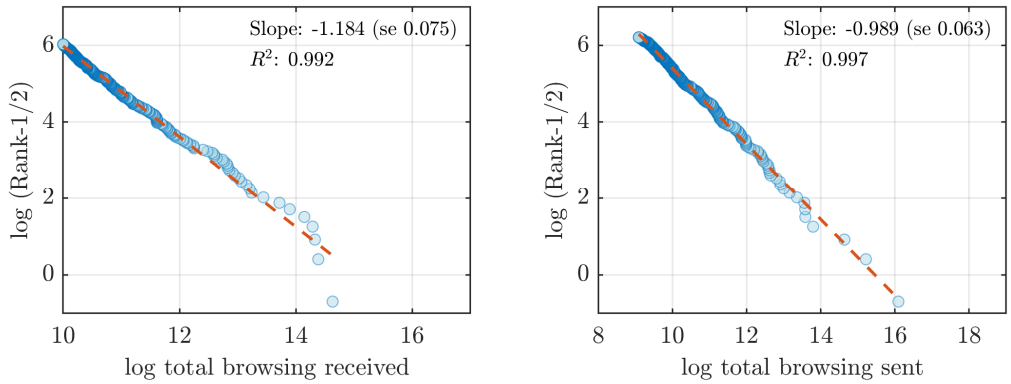
Figure B.5: Robustness: Excluding Financial Firms




*Notes*: The figure replicates the Rank-1/2 power-law estimation after excluding all financial firms (SIC codes 6000–6999). The estimates are nearly identical to those in the baseline sample, indicating that the results are not driven by the financial sector.

**Pooled sample across years.** Figure B.6 pools all available years from 2009 to 2016. The estimates remain close to the 2016 baseline, showing that the fat-tailed attention distributions are not specific to one year.

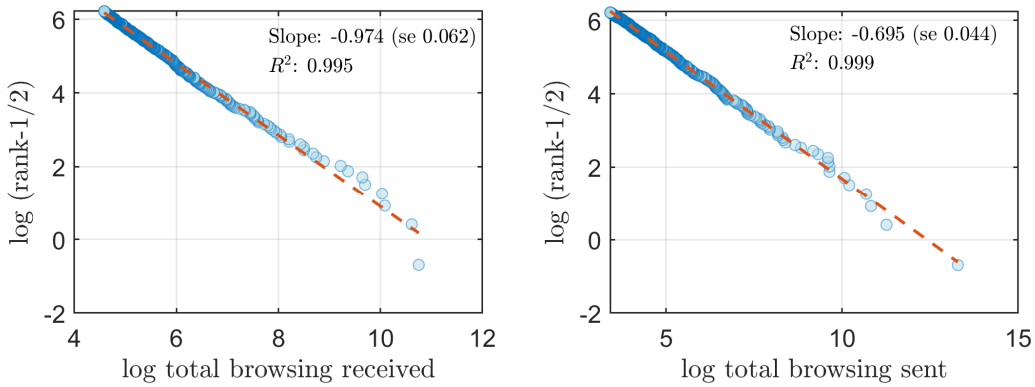
Figure B.6: Robustness: Pooled Sample (2009–2016)




*Notes*: The figure pools all years from 2009 to 2016 and estimates the power-law exponents on the combined sample. The estimates are nearly identical to the 2016 baseline, showing that the results are not year-specific.

**Normalizing by filing frequency.** One potential concern is that firms with more SEC filings mechanically attract more browsing, potentially confounding the attention measure. To address this, we normalize each firm's total browsing received (or sent) by its number of filings, constructing a measure of attention per filing. Figure B.7 shows that the power-law characterization holds strongly for this normalized measure, suggesting that attention concentration reflects differences in visibility rather than filing frequency alone.

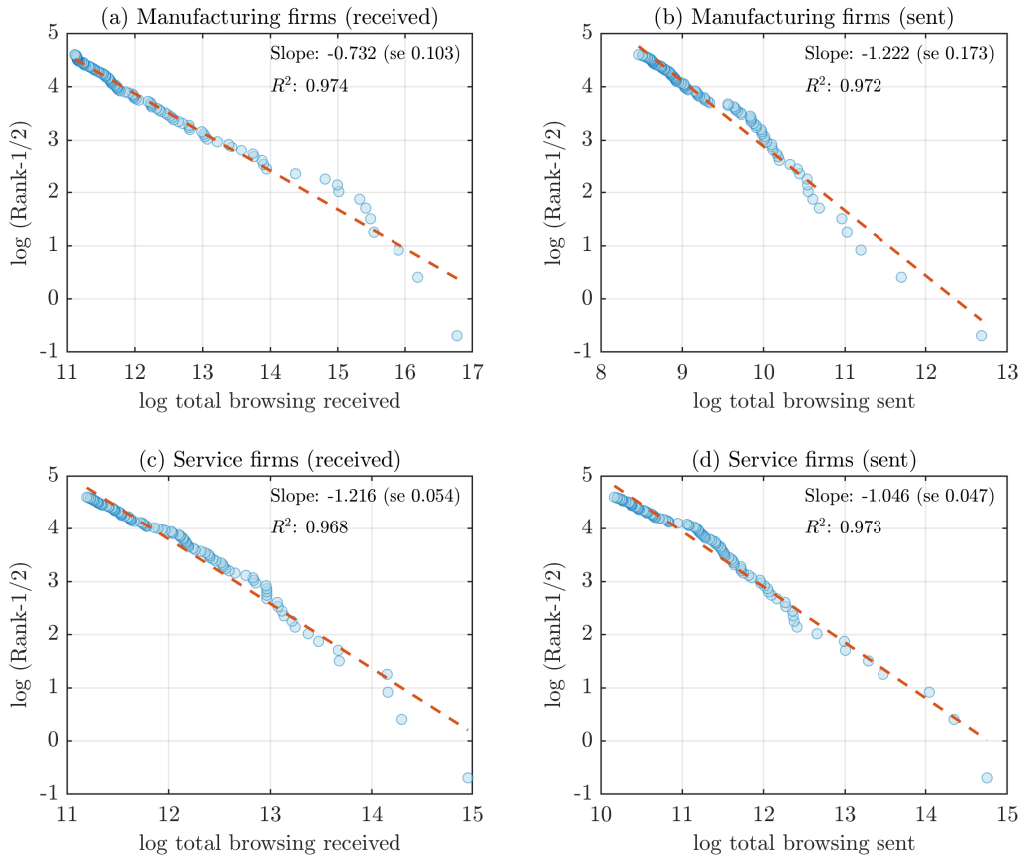
Figure B.7: Robustness: Attention per Filing




*Notes*: The figure normalizes each firm's total browsing by its number of SEC filings, measuring attention received (or sent) per filing. The power-law exponents remain close to the baseline, showing the results are not driven by heterogeneity in filing frequency.

**By industry.** Figure B.8 estimates the exponents separately for manufacturing and service firms, the two largest industry groups in the sample. Both sectors exhibit exponents similar to the full sample, indicating that the result is not driven by a single industry.

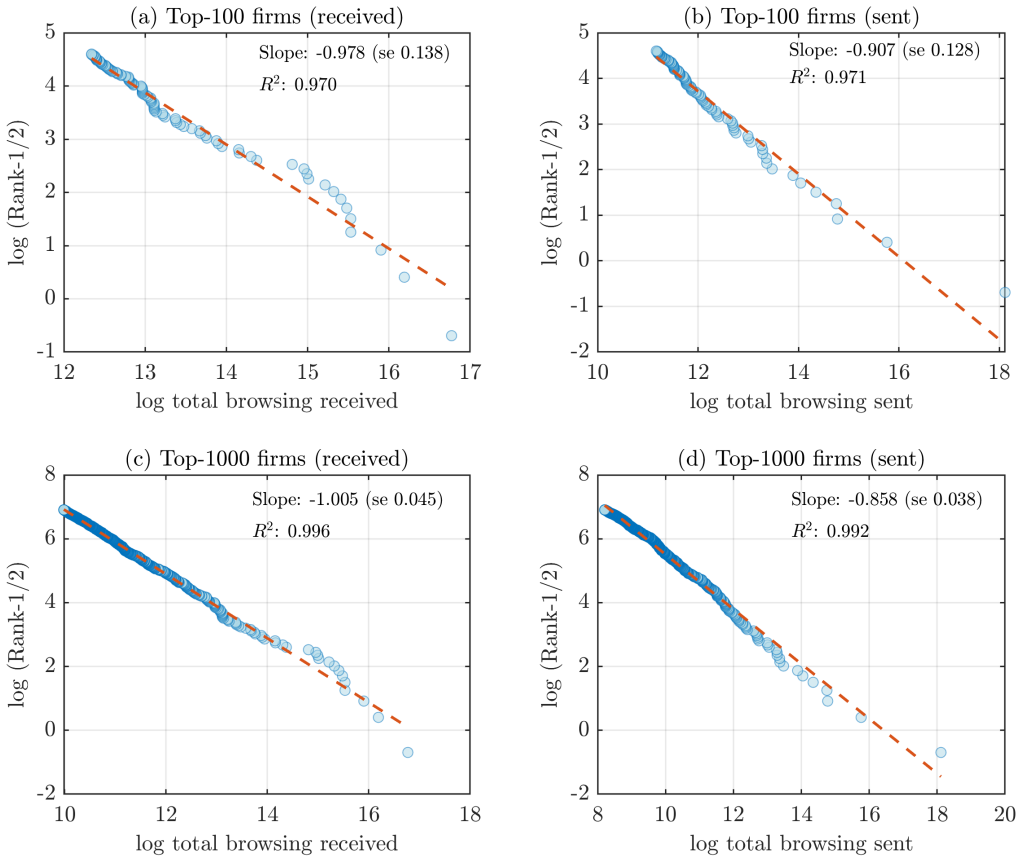
Figure B.8: Robustness: Manufacturing and Service Industries




*Notes*: The figure estimates power-law exponents separately for manufacturing firms (SIC 2000–3999) and service firms (SIC 7000–8999). Both industries display exponents close to the baseline estimates, confirming that the results are not industry-specific.

**Varying the rank cutoff.** Our baseline estimates use the top 500 firms by attention received or sent. Figure B.9 shows that the results are robust to alternative cutoffs: using the top 100 or top 1,000 firms yields nearly identical exponent estimates. The power-law fit is not sensitive to the precise number of firms included in the estimation.

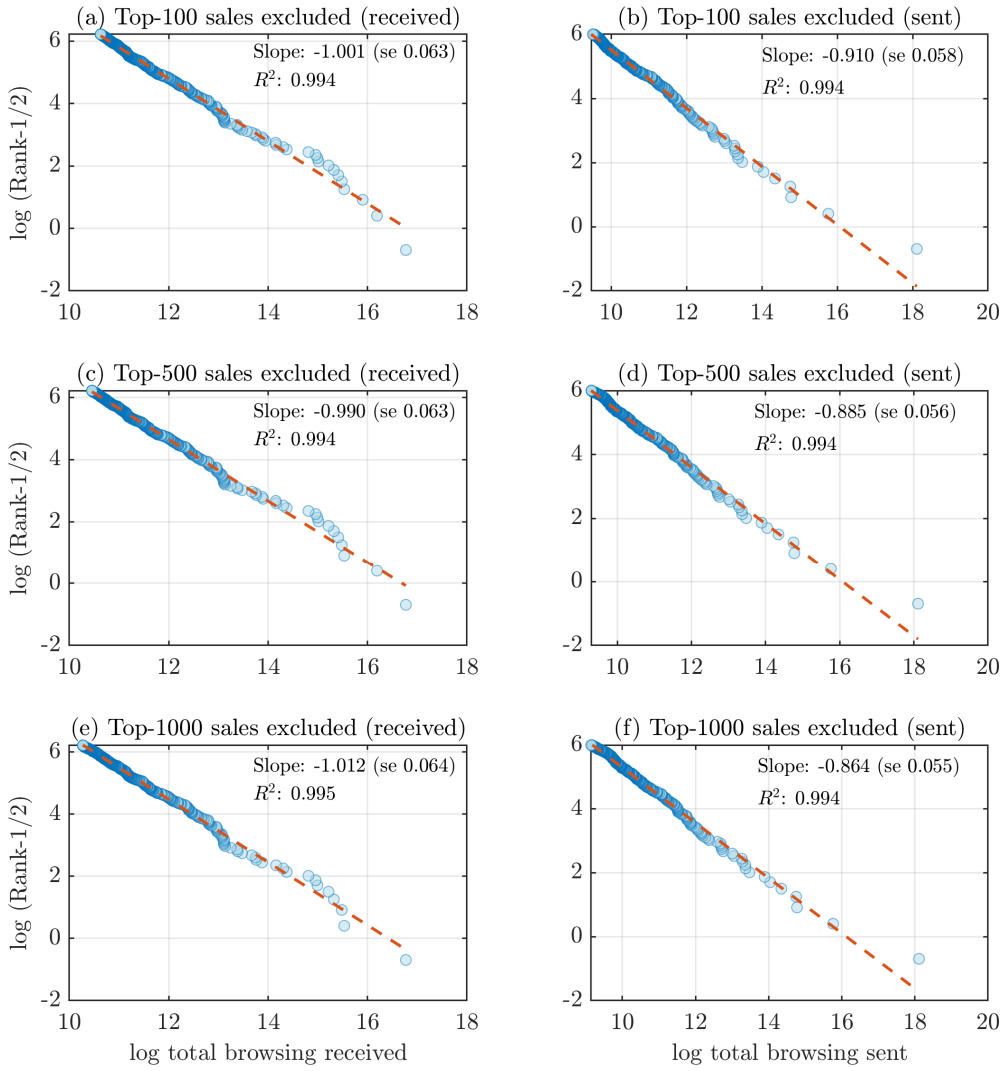
Figure B.9: Robustness: Top 100 and Top 1,000 Cutoffs




*Notes*: The figure estimates power-law exponents using the top 100 firms (left panels) and top 1,000 firms (right panels) by attention received or sent. The estimates are stable across cutoffs, indicating that the results are not sensitive to the choice of $N$ in the Rank-1/2 regression.

**Excluding the largest firms by sales.** Fact 2 shows that attention scales only weakly with firm size. A complementary test asks whether the power-law tails are driven by a few giant firms. Figure B.10 excludes the top 100, 500, and 1,000 firms by sales and re-estimates the power-law exponents on the remaining sample. The exponents remain close to their baseline values, confirming that the concentration of attention is not merely a reflection of the concentration of economic activity.

Figure B.10: Robustness: Excluding Top Firms by Sales




*Notes*: The figure excludes the top 100, 500, and 1,000 firms by sales and re-estimates the power-law exponents on the remaining sample. The estimates are nearly identical to the baseline, showing that the fat-tailed attention distributions are not driven by a few extremely large firms.

**Summary.** Across all these robustness checks—excluding financial firms, pooling years, normalizing by filings, splitting by industry, varying the rank cutoff, and excluding the largest sales firms—the power-law characterization remains stable. Attention received consistently exhibits an exponent near one (Zipf's Law), while attention sent displays an even fatter tail with an exponent significantly below one. These patterns are robust features of the data.

# C. Additional Quantitative Results

## C.1 Quantitative Results Using News Media

This subsection repeats the quantitative exercise using the news-based attention measure. Structural parameters are kept identical to the baseline calibration. The information-network parameters are chosen to match the power-law coefficient of news-based attention received and the mean absolute forecast error in the Livingston Survey. Tables C.1 and C.2 report the calibration and targeted moments.

Table C.1: Parameter Calibration Using News-Based Attention

| Param. | Value | Data Source | Empirical Target |
|---|---|---|---|
| Exogenously determined parameters | | | |
| $\nu$ | 2 | — | Frisch elasticity |
| $\rho$ | 3.5 | — | elasticity of substitution |
| $\{\omega_i\}_{i=1}^N$ | — | Compustat | sales share |
| $\overline{\sigma_i}$ | 0.176 | Compustat | weighted average idiosyncratic volatility |
| $\{\sigma_i\}_{i=1}^N$ | — | Compustat | idiosyncratic productivity volatility |
| $\sigma_\xi$ | 0.02 | San Francisco Fed | aggregate TFP volatility |
| Endogenously determined parameters | | | |
| $\Gamma_0$ | -25 | Livingston | mean absolute forecast error |
| $\Gamma_1$ | 2.01 | Yahoo Finance | in-degree power-law coefficient |

Table C.2: Targeted Moments Using News-Based Attention

| Empirical Target | Data | Model |
|---|---|---|
| In-degree power-law coefficient | -0.82 | -0.82 |
| Mean absolute forecast error | 0.025 | 0.030 |

## C.2 Alternative Information Structure

This subsection reports quantitative results under the alternative information structure described in the main text. Instead of a binary information matrix, firms observe noisy signals about other firms' fundamentals, with bilateral browsing intensity mapped into signal precision as in equation (6.4). All structural parameters governing preferences, technology, firm size, and shock processes are kept fixed at their baseline values. The information parameters $\Gamma$ and $\chi$ are chosen to match the mean absolute forecast error and forecast dispersion in the Livingston Survey. Table C.3 reports the calibration.

Table C.3: Parameter Calibration in Alternative Model

| Param. | Value | Data Source | Empirical Target |
|---|---|---|---|
| Exogenously determined parameters | | | |
| $\nu$ | 2 | — | Frisch elasticity |
| $\rho$ | 3.5 | — | elasticity of substitution |
| $\{\omega_i\}_{i=1}^N$ | — | Compustat | sales share |
| $\overline{\sigma_i}$ | 0.176 | Compustat | weighted average idiosyncratic volatility |
| $\{\sigma_i\}_{i=1}^N$ | — | Compustat | idiosyncratic productivity volatility |
| $\sigma_\xi$ | 0.02 | San Francisco Fed | aggregate TFP volatility |
| Endogenously determined parameters | | | |
| $\mathbf{B}$ | — | EDGAR | browsing matrix |
| $\Gamma$ | 8850 | Livingston | mean absolute forecast error: model 0.033, data 0.025 |
| $\chi$ | 1.9 | Livingston | forecast dispersion: model 0.019, data 0.018 |

Table C.4 shows that the main quantitative conclusions are unchanged. Heterogeneous information alone generates substantial aggregate volatility, and the top attention-receiving firms account for a disproportionate share of this volatility. With firm-size and volatility heterogeneity, both top-sales and top-attention firms contribute meaningfully to aggregate fluctuations. Relative to the perfect-information benchmark, dispersed information attenuates the dominance of the largest firms, consistent with the baseline model.

Table C.4: Aggregate and Idiosyncratic Volatility: Alternative Model

| | Size and Volatility | | |
|---|---|---|---|
| | Equal size, equal volatility | Hetero. size, equal volatility | Hetero. size, hetero. volatility |
| **Heterogeneous Information** | | | |
| Total volatility | 4.79 | 5.28 | 5.58 |
| Top-100 firms/total | 16.28% | 40.53% | 46.59% |
| Top-100 firms | 0.78 | 2.14 | 2.60 |
| Top-100-sales firms | — | 2.05 | 2.31 |
| Top-100-attention firms | 0.78 | 1.30 | 1.46 |
| **Perfect Information** | | | |
| Total volatility | 6.36 | 7.03 | 7.29 |
| Top-100 firms/total | 2.20% | 41.96% | 48.42% |
| Top-100 firms | 0.14 | 2.95 | 3.55 |
| Top-100-sales firms | — | 2.95 | 3.32 |
| Top-100-attention firms | — | 1.19 | 1.35 |
| **Lucas Island** | | | |
| Total volatility | 4.07 | 4.51 | 4.71 |
| Top-100 firms/total | 2.21% | 41.91% | 48.41% |
| Top-100 firms | 0.09 | 1.89 | 2.28 |
| Top-100-sales firms | — | 1.89 | 2.13 |
| Top-100-attention firms | — | 0.76 | 0.87 |

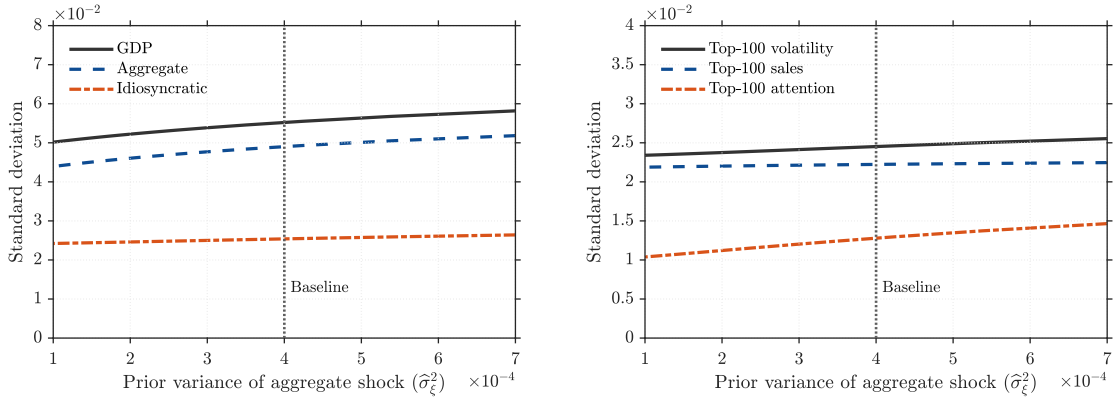
*Notes*: Volatilities are multiplied by 100. "—" indicates not applicable.

## C.3 Adding Public Signals

Figure C.1 shows how GDP volatility, the volatilities of the aggregate and idiosyncratic components, and the volatilities of the 100 most volatile firms, the top-100 sales firms, and the top-100 attention firms vary with

the prior variance of aggregate shock, as described in Section 6.3. The granular effects remain quantitatively important across these alternative values of the prior variance.

Figure C.1: Robustness on Prior Variance




# D. Proofs of Results in Section 3 and Section 4

## D.1 Asymptotic and Probability Notations

Throughout the main text and the appendix, we follow the standard definitions of asymptotic notation used in textbooks on random graph theory, probability theory, and asymptotic analysis (e.g., Frieze and Karoński (2026), Billingsley (1995)). Specifically, for two nonnegative functions $f(N)$ and $g(N)$ defined on an economy with size $N \in \mathbb{N}$,

- $f(N) = O\left(g(N)\right)$ **[asymptotically bounded above]**: $f(N) \leq cg(N)$ for some constant $c > 0$ and for all $N$ sufficiently large.
- $f(N) = \Omega\left(g(N)\right)$ **[asymptotically bounded below]**: $f(N) \geq cg(N)$ for some constant $c > 0$ and for all $N$ sufficiently large.
- $f(N) = \Theta\left(g(N)\right)$ **[asymptotically bounded above and below]**: $c_1 g(N) \leq f(n) \leq c_2 g(N)$ for some constants $c_1 > 0$, $c_2 > 0$ and for all $N$ sufficiently large.
- $f(N) = o\left(g(N)\right)$ **[asymptotically negligible]**: $\lim_{N\to\infty} \frac{f(N)}{g(N)} = 0$.
- $f(N) \approx g(N)$ **[asymptotic equivalence in ratios]**: $\lim_{N\to\infty} \frac{f(N)}{g(N)} = 1$.

In addition, let $\{X_N\}_{N\in\mathbb{N}}$ be a sequence of random variables and let $\{\mathcal{A}_N\}_{N\in\mathbb{N}}$ be a sequence of events in economies indexed by $N$. We define

- w.h.p. **[occurrence with high probability]**: A sequence of events $\mathcal{A}_N$, where $N = 1, 2, \ldots$, is said to occur with high probability (w.h.p.) if $\lim_{N\to\infty} \mathbb{P}(\mathcal{A}_N) = 1$.
- $X_N = \Theta_p(1)$ **[bounded in probability]**: A sequence of random variables $\{X_N\}$ is said to be $\Theta_p(1)$ if it is bounded in probability, i.e.,

$$X_N = \Theta_p(1) \quad \Longleftrightarrow \quad \forall \varepsilon > 0,\ \exists \delta > 0,\ \exists M > 0,\ \exists N_0 \in \mathbb{N},\ \forall N \geq N_0 :$$

$$P(|X_N| < \delta) < \varepsilon \quad \text{and} \quad P(|X_N| > M) < \varepsilon.$$

## D.2 Proof of Lemma 3.1

*Proof.* Under quasi-linear preferences, $U'(C_t) = 1$. The household's labor-supply condition in island $i$ is

$$\ell_{it}^{1/\nu} = W_{it}\mathbb{E}_{it}\left[\frac{1}{P_t}\right].$$

Using goods-market clearing, $C_t = Q_t$, and substituting the labor-supply schedule into firm $i$'s problem gives

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

subject to

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

Inverse demand implies

$$\frac{P_{it}}{P_t} = \left(\frac{NQ_{it}}{Q_t}\right)^{-1/\rho}.$$

Hence firm $i$'s real revenue is

$$\frac{P_{it}Q_{it}}{P_t} = N^{-1/\rho}Q_t^{1/\rho}Q_{it}^{1-1/\rho}.$$

Substituting $Q_{it} = \exp(z_{it})\ell_{it}$, the first-order condition is

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

Taking logs and dropping constants and higher-order terms yields

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

Since $q_{it} = z_{it} + \ell_{it}$, we obtain

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

where

$$\delta = \frac{\rho(1+\nu)}{\rho+\nu}, \qquad \alpha = \frac{\nu}{\rho+\nu}.$$

Finally, log-linearizing aggregate output around the symmetric steady state gives

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

This completes the proof. □

## D.3 Proof of Proposition 4.1

*Proof.* Suppose we are given an attention graph $\mathbf{A}$, and conjecture the following solution for an individual firm's decision,

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

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

$$q_t = \mathbf{M} \boldsymbol{z}_t. \tag{D.2}$$

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

$$M_{ij} = 0; \text{ if } A_{ij} = 0. \tag{D.3}$$

Thus, the equilibrium conjecture is given by

$$q_{it} = \sum_{j \in \mathcal{D}_i}^{N} M_{ij} z_{jt}. \tag{D.4}$$

To derive the solution, it is necessary to use the general representation (D.1). Using the in-out degree construction in the main text, we obtain a compact representation for the expected aggregate output,

$$\frac{1}{N} \mathbb{E}_{it} \left[ \sum_{j=1}^{N} q_{jt} \right] = \frac{1}{N} \left[ \sum_{j \in \mathcal{D}_i} \left( \sum_{k=1}^{N} A_{kj} M_{kj} \right) z_{jt} + \left\{ \sum_{h \notin \mathcal{D}_i} \sum_{k=1}^{N} A_{kh} M_{kh} \right\} \mathbb{E}_{it} \left[ \xi_t | S_{\mathcal{D}_i} \right] \right]. \tag{D.5}$$

Next, we apply the inference formula (3.10) and obtain

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

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

$$q_{it} = \delta z_{it} + \alpha \frac{1}{N} \sum_{j \in \mathcal{D}_i} \left[ \sum_{k=1}^{N} A_{kj} M_{kj} + \lambda_i \sum_{h \notin \mathcal{D}_i} \left( \sum_{k=1}^{N} A_{kh} M_{kh} \right) \right] z_{jt} \tag{D.6}$$

given that $i \in \mathcal{D}_i, \forall i \in \mathcal{I}$. Comparing (D.6) with the refined conjecture (D.3) and (D.4), 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

$$M_{ij} = \begin{cases} 0; & j \notin \mathcal{D}_i \\ \alpha \frac{1}{N} \left[ \sum_{k \in \mathcal{M}_j} A_{kj} M_{kj} + \lambda_i \sum_{h \notin \mathcal{D}_i} \left( \sum_{k \in \mathcal{M}_h} A_{kh} M_{kh} \right) \right]; & j \in \mathcal{D}_i,\ j \neq i \\ \delta + \alpha \frac{1}{N} \left[ \sum_{k \in \mathcal{M}_j} A_{kj} M_{kj} + \lambda_i \sum_{h \notin \mathcal{D}_i} \left( \sum_{k \in \mathcal{M}_h} A_{kh} M_{kh} \right) \right]; & j = i \in \mathcal{D}_i. \end{cases} \tag{D.7}$$

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 (D.7) for matrix $\mathbf{M}$ is a linear system of equations; however, characterizing individual-level solutions in closed-form is difficult, if not impossible. It would require keeping track of the exact attention distribution of $\mathbf{A}$, but the system can be solved numerically without difficulty.

We now derive a closed-form solution for the aggregate output. Recall from (D.3), 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 an $N \times 1$ 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 (D.7) over $k \in \mathcal{M}_j$, we get

$$g_j = \frac{1}{N} \sum_{k \in \mathcal{M}_j} M_{kj} = \frac{1}{N} \left( \delta + \alpha \frac{1}{N} m_j \sum_{k \in \mathcal{M}_j} M_{kj} + \alpha \frac{1}{N} \sum_{k \in \mathcal{M}_j} \left[ \lambda_k \sum_{h \notin \mathcal{D}_k} \left( \sum_{l \in \mathcal{M}_h} M_{lh} \right) \right] \right). \tag{D.8}$$

We apply algebraic tricks to simplify the last term,

$$\frac{1}{N} \sum_{h \notin \mathcal{D}_k} \left( \sum_{l \in \mathcal{M}_h} M_{lh} \right) = \frac{1}{N} \sum_{h=1}^{N} (1 - A_{kh}) \left( \sum_{l \in \mathcal{M}_h} M_{lh} \right) = \left( \mathbf{1}_{1 \times N} - \mathbf{A}_{(k,:)} \right) \cdot \mathbf{g},$$

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

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

where $\operatorname{diag}(\boldsymbol{\lambda})$ is the diagonal matrix of signal loading vector $\{\lambda_1, \lambda_2, \cdots, \lambda_N\}$, which depends on the attention network's out-degree $\{d_1, d_2, \cdots, d_N\}$. $\mathbf{e}_j$ is $j$th standard basis row vector. $\mathbf{1}$ is $N \times N$ matrix of 1s. Now equation (D.8) becomes

$$g_j = \frac{1}{N} \delta + \alpha \frac{1}{N} m_j g_j + \alpha \frac{1}{N} \mathbf{e}_j \mathbf{A}^T \operatorname{diag}(\boldsymbol{\lambda}) (\mathbf{1} - \mathbf{A}) \mathbf{g}, \tag{D.9}$$

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 $j = 1, 2, \ldots, N$, we obtain the unique solution for $\mathbf{g}$,

$$\mathbf{g} = \frac{\delta}{N}\left(\mathbf{I} - \frac{\alpha}{N}\operatorname{diag}(\boldsymbol{m}) - \alpha\frac{1}{N}\mathbf{A}^T\operatorname{diag}(\boldsymbol{\lambda})(\mathbf{1}-\mathbf{A})\right)^{-1}\mathbf{1}_{N\times 1}, \tag{D.10}$$

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

## D.4 Proof of Proposition 4.2

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 $\lambda$ 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 4** (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 nonnegative matrix (every element of $B$ is nonnegative) whose spectral radius (the largest absolute value or complex modulus of its eigenvalues) is strictly less than $k$.*

The following are properties of M-matrices.

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

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

$$\mathbf{A}'\operatorname{diag}(\boldsymbol{\lambda})(\mathbf{1}-\mathbf{A}) = \sum_{i=1}^{N}\lambda_i\mathbf{A}_i(\mathbf{1}-\mathbf{A}_i'). \tag{D.11}$$

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

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

$$u_{kj} = \begin{cases} 1, & \text{if } A_{ik} = 1 \text{ and } A_{ij} = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{D.12}$$

*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 D.1 we derive the following result.

**Lemma D.2.** *Let* $\mathbf{V} = (\mathbf{v}_{ij})_{N\times N} = \mathbf{T}_1 + \mathbf{T}_2$.

*(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* $\frac{\alpha}{N} \cdot \sum_{i\in\mathcal{M}_k} \frac{1}{d_i}$*, and 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* $\alpha^j$*. 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) Consider the $j$-th column of $\mathbf{V}$. We first show that its $j$-th column sum is at most $\alpha < 1$. First, by Lemma D.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 sum of non-diagonal elements in $j$-th column of $\mathbf{V}$, by Lemma D.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 sum of non-diagonal elements in $j$-th column of $\mathbf{V}$ is equal to

$$\frac{\alpha}{N} \cdot \sum_{i\notin\mathcal{M}_j} \sum_{k\in\mathcal{D}_i} \lambda_i. \tag{D.13}$$

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

$$\frac{\alpha}{N} \cdot m_j + \frac{\alpha}{N} \cdot \sum_{i\notin\mathcal{M}_j} \sum_{k\in\mathcal{D}_i} \lambda_i = \frac{\alpha}{N} \cdot m_j + \frac{\alpha}{N} \cdot \sum_{i\notin\mathcal{M}_j} d_i\lambda_i \leq \frac{\alpha}{N} \cdot m_j + \frac{\alpha}{N} \cdot \sum_{i\notin\mathcal{M}_j} 1 = \frac{\alpha}{N} \cdot m_j + \frac{\alpha}{N} \cdot (N - m_j) = \alpha.$$

By Theorem 1, the spectral radius of $\mathbf{V}$ is at most $\alpha$, thus by Definition 4 we obtain that $\mathbf{I} - \mathbf{V}$ is an $M$-matrix.

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

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

For the non-diagonal element $k$-th row sum of $\mathbf{V}$, by Lemma D.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

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

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

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

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) Similar 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

$$\mathbf{v}_{kj} = \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k \text{ and } i \notin \mathcal{M}_j} \lambda_i \leq \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} \lambda_i \leq \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} \frac{1}{d_i} \leq \frac{\alpha}{N} \cdot \sum_{i \in \mathcal{M}_k} 1 \leq \frac{\alpha}{N} \cdot m_k. \tag{D.14}$$

(4) We prove it by induction. The case $j = 1$ is proved by (1). Now assume that it is true for $j$. Let $\mathbf{V}^j = (\mathbf{V}_1^j \ \ \mathbf{V}_2^j \ \ \ldots \ \ \mathbf{V}_N^j)$ 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 $\alpha^j$ 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, note 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

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

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 follow directly from (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 $\mathbf{I}+\mathbf{V}$. □

We are now 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 D.2 we obtain that $g_i$ has the following lower and upper bounds for $1 \le i \le N$, i.e.,

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

□

## D.5 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 D.3.** *Let $\mathbf{A} = (a_{ij})_{N\times N}$ be the attention graph where $a_{ii} = 1$. Let $\mu_i = \sum_{j\in\mathcal{M}_i}\frac{1}{d_j}$ 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 D.2, we obtain

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

An important observation is that

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

where the second equality follows from the definition of the information network. Next, we apply the Cauchy-Schwarz inequality,

$$\left(\sum_{i=1}^N 1\mu_i\right)^2 \le \left(\sum_{i=1}^N 1^2\right)\left(\sum_{i=1}^N\mu_i^2\right) = N\sum_{i=1}^N\mu_i^2,$$

which immediately implies that

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

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 that $\sum_{i=1}^{N} \mu_i^2$ is maximized when $\frac{N}{\mu}$ number of the $\mu_i$ equal $\mu$ and the remaining $N - \frac{N}{\mu}$ number of $\mu_i$ 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 D.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 $\mu_i$ 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, 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 $\max_i m_i = \Theta(N)$.

□

# Online Supplementary Appendix

## Micro Shocks and Macro Fluctuations in the Information Network

Zhen Huo  Minghao Li  Jieran Wu  Huan Xiong

June 2026

Contents

S.1 Proof of Lemma 5.1 . . . . . S-2
S.2 Proof of Proposition 5.1 . . . . . S-3
S.3 Proof of Proposition 5.2 . . . . . S-6
S.4 Proof of Lemma 5.2 . . . . . S-8
S.5 Proof of Lemma 5.3 . . . . . S-9
S.6 Technical Results of Random Graph Models . . . . . S-13
S.7 Additional Lemmas for Proofs of Propositions 5.3 and 5.4 . . . . . S-50
S.8 Proof of Proposition 5.3 . . . . . S-55
S.9 Proof of Proposition 5.4 . . . . . S-65

**References** **77**

# Proof of results in Section 5

## S.1 Proof of Lemma 5.1

*Proof.* Under quasi-linear preferences with heterogeneous taste parameters, $U'(C_t) = 1$. The household's labor-supply condition in island $i$ remains identical to the baseline model,

$$\ell_{it}^{1/\nu} = W_{it}\mathbb{E}_{it}\left[\frac{1}{P_t}\right].$$

Using goods-market clearing, $C_t = Q_t$, and substituting the labor-supply schedule into firm $i$'s problem gives

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

subject to

$$Q_{it} = \left(\frac{P_{it}}{P_t}\right)^{-\rho}\zeta_i Q_t, \qquad Q_{it} = \exp(\bar{z}_i + z_{it})\ell_{it}.$$

We normalize the price index to unity: $P_t = \left[\sum_{i=1}^N \zeta_i P_{it}^{1-\rho}\right]^{\frac{1}{1-\rho}} \equiv 1$, and the first-order condition is given by

$$\left(1+\frac{1}{\nu}\right)\ell_{it}^{\frac{1}{\nu}} = \frac{\rho-1}{\rho}(\zeta_i)^{\frac{1}{\rho}}\,\mathbb{E}_{it}\left[Q_t^{\frac{1}{\rho}}\exp(\bar{z}_i + z_{it})^{1-\frac{1}{\rho}}\ell_{it}^{-\frac{1}{\rho}}\right].$$

Log-linearizing the above equation and using the relation $q_{it} = z_{it} + \ell_{it}$, we obtain the beauty-contest representation,

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

where $\delta$ and $\alpha$ are defined as before. To derive the firm sizes and log-linearized aggregate output, we compute the modified steady state. In steady state, the individual firm's output is given by

$$Q_i = \iota_i Q^\alpha; \qquad \iota_i = \left(\frac{\nu}{1+\nu}\frac{\rho-1}{\rho}(\zeta_i)^{\frac{1}{\rho}}\right)^{\frac{\rho\nu}{\rho+\nu}} = \underbrace{\left(\frac{\nu}{1+\nu}\frac{\rho-1}{\rho}\right)^{\frac{\rho\nu}{\rho+\nu}}}_{\iota}\zeta_i^{\frac{\nu}{\rho+\nu}}\left(\exp(\bar{z}_i)\right)^{\frac{\rho(1+\nu)}{\rho+\nu}}.$$

We substitute the above equation into the definition of aggregate output,

$$Q = \left[\sum_{i=1}^N (\zeta_i)^{\frac{1}{\rho}} Q_i^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}} = \iota Q^\alpha\left[\sum_{i=1}^N (\zeta_i)^{\frac{1+\nu}{\rho+\nu}}\left(\exp(\bar{z}_i)\right)^{\frac{(\rho-1)(1+\nu)}{\rho+\nu}}\right]^{\frac{\rho}{\rho-1}}.$$

Therefore, the steady-state aggregate output is given by

$$Q = \left( \iota \left[ \sum_{i=1}^{N} (\zeta_i)^{\varkappa} \left( \exp(\bar{z}_i) \right)^{(\rho-1)\varkappa} \right]^{\frac{\rho}{\rho-1}} \right)^{\frac{1}{(1-\alpha)}}, \qquad \varkappa = \frac{1+\nu}{\rho+\nu}.$$

The individual output is given by

$$Q_i = \iota^{\frac{1}{1-\alpha}} (\zeta_i)^{\frac{\varkappa\nu}{1+\nu}} \left( \exp(\bar{z}_i) \right)^{\rho\varkappa} \left[ \sum_{i=1}^{N} (\zeta_i)^{\varkappa} \left( \exp(\bar{z}_i) \right)^{(\rho-1)\varkappa} \right]^{\frac{\alpha\rho}{(\rho-1)(1-\alpha)}}.$$

Now, standard log-linearization implies that

$$q_t = \sum_{i=1}^{N} \frac{\zeta_i^{\frac{1}{\rho}} Q_i^{\frac{\rho-1}{\rho}}}{\sum_{j=1}^{N} \zeta_j^{\frac{1}{\rho}} Q_j^{\frac{\rho-1}{\rho}}} q_{it}.$$

It follows that

$$\zeta_i^{\frac{1}{\rho}} Q_i^{\frac{\rho-1}{\rho}} = \iota^{\frac{\rho-1}{(1-\alpha)\rho}} (\zeta_i)^{\frac{1}{\rho} + \frac{\varkappa\nu(\rho-1)}{(1+\nu)\rho}} \left( \exp(\bar{z}_i) \right)^{(\rho-1)\varkappa} \left[ \sum_{j=1}^{N} \left( \zeta_j \right)^{\varkappa} \left( \exp\left( \bar{z}_j \right) \right)^{(\rho-1)\varkappa} \right]^{\frac{\alpha}{(1-\alpha)}}.$$

Notice that the first component and the last component are common across all islands. In addition, after simple algebra,

$$\frac{1}{\rho} + \frac{\varkappa\nu(\rho-1)}{(1+\nu)\rho} = \varkappa.$$

Therefore, the equilibrium firm share is given by

$$\omega_i = \frac{\zeta_i^{\frac{1}{\rho}} Q_i^{\frac{\rho-1}{\rho}}}{\sum_{j=1}^{N} \zeta_j^{\frac{1}{\rho}} Q_j^{\frac{\rho-1}{\rho}}} = \frac{(\zeta_i)^{\varkappa} \left( \exp(\bar{z}_i) \right)^{(\rho-1)\varkappa}}{\sum_{j=1}^{N} \left( \zeta_j \right)^{\varkappa} \left( \exp\left( \bar{z}_j \right) \right)^{(\rho-1)\varkappa}}.$$

By construction, $\omega_i \in [0, 1]$, $\sum_{i=1}^{N} \omega_i = 1$. This completes the proof. □

## S.2 Proof of Proposition 5.1

*Proof.* In this section, we present the solution to the quantitative model that allows two dimensions of heterogeneity: firm size and the volatility of their idiosyncratic technology shocks. The results in Proposition 5.1 serve as a special case that considers only the firm-size heterogeneity. Suppose the variances of firm-specific economic fundamentals differ,

$$z_{it} = \xi_t + \varepsilon_{it}, \qquad \varepsilon_{it} \sim N(0, \tau_i^{-1}).$$

Given an information network $\mathbf{A}$, the inference about the common fundamental is given by

$$\mathbb{E}_{it}[\xi_t] = \frac{1}{\tau_\xi + \sum_{k\in\mathcal{D}_i}\tau_k}\sum_{j\in\mathcal{D}_i}\tau_j z_{jt}. \tag{S.1}$$

Without loss of generality, we define

$$v_i = \frac{1}{\tau_\xi + \sum_{k\in\mathcal{D}_i}\tau_k}. \tag{S.2}$$

Similar to our homogeneous model, we conjecture the equilibrium strategy as

$$y_{it} = \sum_{j\in\mathcal{D}_i} M_{ij} z_{jt}.$$

It follows from the beauty-contest game that

$$\begin{aligned}
q_{it} &= \delta z_{it} + \alpha\mathbb{E}_{it}\left[\sum_{j=1}^N \omega_j \sum_{k=1}^N M_{jk} z_{kt}\right] \\
&= \delta z_{it} + \alpha\sum_{j=1}^N \omega_j \sum_{k\in\mathcal{D}_i} M_{jk} z_{kt} + \alpha\sum_{j=1}^N \omega_j \sum_{k\notin\mathcal{D}_i} M_{jk}\mathbb{E}_{it}[\xi_t] \\
&= \delta z_{it} + \alpha\sum_{j=1}^N \omega_j \sum_{k\in\mathcal{D}_i} M_{jk} z_{kt} + \alpha\frac{1}{\tau_\xi + \sum_{k\in\mathcal{D}_i}\tau_k}\sum_{j=1}^N \omega_j \sum_{k\notin\mathcal{D}_i} M_{jk}\sum_{q\in\mathcal{D}_i}\tau_q z_{qt} \\
&= \delta z_{it} + \sum_{j\in\mathcal{D}_i}\alpha\left(\sum_{k=1}^N \omega_k A_{kj} M_{kj} + \tau_j v_i \sum_{h\notin\mathcal{D}_i}\left[\sum_{k=1}^N \omega_k A_{kh} M_{kh}\right]\right) z_{jt}.
\end{aligned}$$

GDP in this economy is given by

$$q_t = \sum_{j=1}^N \omega_j q_{jt} = \sum_{j=1}^N \omega_j \sum_{k=1}^N A_{jk} M_{jk} z_{kt} = \sum_{k=1}^N \sum_{j\in\mathcal{M}_k} \omega_j M_{jk} z_{kt} = \sum_{k=1}^N g_k z_{kt},$$

where

$$g_k = \sum_{j\in\mathcal{M}_k} \omega_j M_{jk}$$

denotes the GDP elasticity with respect to micro shocks. Using the beauty-contest equilibrium properties of $M_{ij}$, we have

$$g_k = \sum_{j\in\mathcal{M}_k} \omega_j M_{jk} = \omega_k\delta + \alpha\sum_{j\in\mathcal{M}_k} \omega_j g_k + \alpha\sum_{j\in\mathcal{M}_k}\left[\omega_j v_j \tau_k \sum_{h\notin\mathcal{D}_j}\left(\sum_{\ell\in\mathcal{M}_h}\omega_\ell M_{\ell h}\right)\right]. \tag{S.3}$$

We note that

$$\sum_{h\notin\mathcal{D}_j}\left(\sum_{\ell\in\mathcal{M}_h}\omega_\ell M_{\ell h}\right)=\sum_{h=1}^{N}(1-A_{jh})\left(\sum_{\ell\in\mathcal{M}_h}\omega_\ell M_{\ell h}\right)=\begin{bmatrix}1-A_{j1} & \dots & 1-A_{jN}\end{bmatrix}\begin{bmatrix}\sum_{\ell\in\mathcal{M}_1}\omega_\ell M_{\ell 1}\\ \vdots \\ \sum_{\ell\in\mathcal{M}_N}\omega_\ell M_{\ell N}\end{bmatrix}=\left(\mathbf{1}_{1\times N}-\mathbf{A}_{(j,:)}\right)\cdot\mathbf{g}.$$

Therefore, we simplify the last term in (S.3) as

$$\begin{aligned}
&\sum_{j\in\mathcal{M}_k}\left[\omega_j v_j\tau_k\sum_{h\notin\mathcal{D}_j}\left(\sum_{\ell\in\mathcal{M}_h}\omega_\ell M_{\ell h}\right)\right]\\
=&\sum_{j\in\mathcal{M}_k}\omega_j v_j\tau_k\left[\left(\mathbf{1}_{1\times N}-\mathbf{A}_{(j,:)}\right)\cdot\mathbf{g}\right]\\
=&\sum_{j=1}^{N}\omega_j v_j\tau_k A_{jk}\left[\left(\mathbf{1}_{1\times N}-\mathbf{A}_{(j,:)}\right)\cdot\mathbf{g}\right]\\
=&\mathbf{e}_k\,\mathrm{diag}\,(\boldsymbol{\tau})\,\mathbf{A}'\,\mathrm{diag}\,(\boldsymbol{\omega})\,\mathrm{diag}\,(\boldsymbol{v})\,(\mathbf{1}-\mathbf{A})\,\boldsymbol{g}.
\end{aligned}$$

It follows that the elasticity is given by

$$\begin{aligned}
g_k=&\omega_k\delta+\alpha\sum_{j\in\mathcal{M}_k}\omega_j g_k+\alpha\mathbf{e}_k\,\mathrm{diag}\,(\boldsymbol{\tau})\,\mathbf{A}'\,\mathrm{diag}\,(\boldsymbol{\omega})\,\mathrm{diag}\,(\boldsymbol{v})\,(\mathbf{1}-\mathbf{A})\,\boldsymbol{g}\\
=&\omega_k\delta+\alpha\sum_{j=1}^{N}A_{jk}\omega_j g_k+\alpha\mathbf{e}_k\,\mathrm{diag}\,(\boldsymbol{\tau})\,\mathbf{A}'\,\mathrm{diag}\,(\boldsymbol{\omega})\,\mathrm{diag}\,(\boldsymbol{v})\,(\mathbf{1}-\mathbf{A})\,\boldsymbol{g}.
\end{aligned}$$

Stacking over $k=1,2,\dots,N$, we obtain the unique solution for $\boldsymbol{g}$,

$$\boldsymbol{g}=\delta\left(\mathbf{I}-\alpha\,\mathrm{diag}(\mathbf{A}'\boldsymbol{\omega})-\alpha\,\mathrm{diag}\,(\boldsymbol{\tau})\,\mathbf{A}'\,\mathrm{diag}\,(\boldsymbol{\omega})\,\mathrm{diag}\,(\boldsymbol{v})\,(\mathbf{1}-\mathbf{A})\right)^{-1}\boldsymbol{\omega}. \tag{S.4}$$

The solution in (S.4) is then used in our quantitative analysis. Finally, when firms' idiosyncratic volatility is identical, $\boldsymbol{\tau}$ in (S.1) and $\boldsymbol{v}$ in (S.2) degenerate to $\boldsymbol{\lambda}$. The inference formula implies that our model solution reduces to

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

The proof is now complete. □

## S.3 Proof of Proposition 5.2

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

$$\delta\left(\omega_k + \alpha\omega_k \cdot \sum_{i\in\mathcal{M}_k} \omega_i + \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1} \sum_{i\in\mathcal{M}_k} \frac{\omega_i}{d_i} \sum_{j\notin\mathcal{D}_i} \omega_j\right) \le g_k \le \delta\left(\omega_k + \frac{\alpha}{(1-\alpha)^2} \cdot \left(\omega_k \cdot \sum_{i\in\mathcal{M}_k} \omega_i + \sum_{i\in\mathcal{M}_k} \frac{\omega_i}{d_i}\right)\right). \tag{S.5}$$

(1) $\Leftarrow$: 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 $1 \le k \le N$; or (ii) $w = o(1)$ and $\sum_{j\in\mathcal{M}_k} \frac{\omega_j}{d_j} = \Theta(1)$ for some $1 \le k \le N$. Thus, we prove the claim for the two cases.

(i) If $\omega_k = \Theta(1)$, then by equation (S.5) we obtain

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

Thus

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

(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 (S.5) we obtain

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

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) $\Rightarrow$: Assume $\sum_{i=1}^{N} g_i^2 = \Theta(1)$.
by (S.5) we obtain

$$\sum_{i=1}^{N} g_i^2 = O\left(\sum_{i=1}^{N}\left(\omega_i + \sum_{j\in\mathcal{M}_i}\frac{\omega_j}{d_j}\right)^2\right) = O\left(\sum_{i=1}^{N}(\mu_i^\omega)^2\right). \tag{S.6}$$

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 that $\sum_{i=1}^{N}(\mu_i^\omega)^2$ is maximized when $\frac{2}{\mu^\omega}$ number of the $\mu_i^\omega$ equals $\mu^\omega$ and remaining $N - \frac{2}{\mu^\omega}$ number of $\mu_i^\omega$ equal $0$. Therefore,

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

Therefore, by Eq. (S.6) we have

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

This means that

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

The proof is now complete. □

**Remark.** When $\omega_i = \frac{1}{N}$ for $1 \le i \le N$, 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,

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

## S.4 Proof of Lemma 5.2

*Proof.* We focus on the case with $\beta_m \ge 1$ so that all probabilities are well-defined. We then extend our analysis to the case with $\beta_m < 1$. First, for $1 \le j \le N$, by (5.5), the expected in-degree is given by

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

Similarly, for $1 \le i \le N$, the expected out-degree is given by

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

where we define

$$\zeta\left(\beta_m^{-1}, N\right) \equiv \sum_{j=1}^{N} j^{-\beta_m^{-1}} \tag{S.7}$$

as the finite-order ($N$) partial sum of the Riemann zeta function $\zeta$ with power coefficient $\beta_m^{-1}$. When

$N$ is sufficiently large, the Riemann zeta function can be approximated asymptotically as[22]

$$\zeta\left(\beta_m^{-1}, N\right) \approx \begin{cases} \frac{N^{1-\beta_m^{-1}}}{1-\frac{1}{\beta_m}}, & \beta_m > 1; \\ \ln N, & \beta_m = 1; \\ \zeta(\beta_m^{-1}), & \beta_m \in [0, 1). \end{cases} \tag{S.8}$$

Here $\zeta(\beta_m^{-1})$ represents the convergent Rieman zeta constant when $\beta_m < 1$. (S.8) is a well-known result and can be found on Page 77 of Titchmarsh and Heath-Brown (1986) and Page 443 of Graham et al. (1994).

**Remark:** When $\beta_m < 1$, the formula (5.5) may lead to negative probability. In this case, we reformulate the probability as

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

This change does not affect any of our results. Specifically, the expected degrees are increased by a constant 1:

$$\mathbb{E}[m_j] = \frac{N}{j^{1/\beta_m}} + 1; \qquad \mathbb{E}[d_i] = \frac{N}{N-1}\left[\zeta(1/\beta_m, N) - \frac{1}{i^{1/\beta_m}}\right] + 1.$$

In this case, the expected in-degrees excluding self-loops (the diagonal elements) still satisfy exact power laws. □

## S.5 Proof of Lemma 5.3

*Proof.* We begin with the first statement of this lemma, which characterizes the asymptotic behavior of order statistics from the Pareto distribution.

Fix an economy of size $N$. Let $\{X_i\}_{i=1}^N$ be a sequence of i.i.d. random samples drawn from the Pareto distribution. It follows that for any $X \in \{X_i\}_{i=1}^N$,

$$\mathbb{P}(X > x) = \frac{c^\kappa}{N} \cdot x^{-\kappa}, \qquad x \geq x_{\min} := \frac{c}{N^{1/\kappa}}.$$

We define a canonical transformation:

$$Y := \frac{c^\kappa}{N} X^{-\kappa} \in (0, 1].$$

[22] As we defined before, $f(N) \approx g(N)$ if $\lim_{N\to+\infty} \frac{f(N)}{g(N)} = 1$.

The probability distribution of $Y$ is given by

$$\begin{aligned}\mathbb{P}(Y < y) &= \mathbb{P}\left(\frac{c^\kappa}{N} X^{-\kappa} < y\right) = \mathbb{P}\left(X^{-\kappa} < \frac{Ny}{c^\kappa}\right) \\ &= \mathbb{P}\left(X > \left(\frac{Ny}{c^\kappa}\right)^{-1/\kappa}\right) = \frac{c^\kappa}{N}\left(\left(\frac{Ny}{c^\kappa}\right)^{-1/\kappa}\right)^{-\kappa} \\ &= \frac{c^\kappa}{N} \cdot \frac{Ny}{c^\kappa} = y, \qquad y \in (0,1].\end{aligned}$$

Therefore, for each $N \in \mathbb{N}$,

$$Y \sim \text{uniform}(0,1)$$

follows the uniform distribution. Let $x_{(1)} > x_{(2)} > \cdots > x_{(N)}$ be the order statistics from $X \in \{X_i\}_{i=1}^N$ in decreasing order. We define the corresponding order statistics in $Y$ with reverse order, and we denote them as $Y_{1:N} \leq \cdots \leq Y_{N:N}$, in increasing order, following the convention in Arnold et al. (2008).

$$x_{(k)} = \frac{c}{N^{1/\kappa}} \left(Y_{k:N}\right)^{-1/\kappa}, \qquad k \in [N]. \tag{S.9}$$

The order statistics of uniformly distributed random variables follow a Beta distribution with the moment formula

$$Y_{k:N} \sim \text{Beta}\left(k, N+1-k\right); \qquad \mathbb{E}\left[\left(Y_{k:N}\right)^{-1/\kappa}\right] = \frac{\Gamma\left(k - \frac{1}{\kappa}\right)\Gamma(N+1)}{\Gamma(k)\Gamma\left(N+1-\frac{1}{\kappa}\right)}; \qquad k\kappa > 1,$$

where the Gamma function is defined in integral form for $z \in \mathbb{C}$,

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt, \quad \text{Re}(z) > 0.$$

Notably, the condition $k \cdot \kappa > 1$ ensures that the moment $\mathbb{E}\left[\left(Y_{k:N}\right)^{-1/\kappa}\right]$ is finite and well-defined.[23]

Using (S.9), we obtain

$$\mathbb{E}\left[x_{(k)}\right] = \frac{c}{N^{1/\kappa}} \frac{\Gamma\left(k - \frac{1}{\kappa}\right)\Gamma(N+1)}{\Gamma(k)\Gamma\left(N+1-\frac{1}{\kappa}\right)}. \tag{S.10}$$

Next, we apply the Stirling approximation to derive the asymptotic expansion of the Gamma function (Abramowitz and Stegun (1964), eq(6.1.47)),

$$\frac{\Gamma(z+a)}{\Gamma(z)} = z^a \left(1 + \frac{a(a-1)}{2z} + O\left(z^{-2}\right)\right), \quad z \to \infty.$$

[23] Even if we relax this assumption, it does not affect our asymptotic results. For $k$ sufficiently large, the moments of order statistics are finite.

Let $N \to \infty$, $z = N + 1 - \frac{1}{\kappa}$, and $a = \frac{1}{\kappa}$. We obtain

$$\frac{\Gamma(N+1)}{\Gamma\left(N+1-\frac{1}{\kappa}\right)} = \frac{\Gamma\left(z+\frac{1}{\kappa}\right)}{\Gamma(z)} \approx N^{1/\kappa},$$

where " $\approx$ " denotes asymptotic equivalence in ratios as before. By construction, we have

$$\lim_{N\to\infty} \mathbb{E}\left[x_{(k)}\right] = c \cdot \frac{\Gamma\left(k-\frac{1}{\kappa}\right)}{\Gamma(k)}; \qquad \forall k > \frac{1}{\kappa}. \tag{S.11}$$

The above equation is an exact identity for each $k > \frac{1}{\kappa}$, when $N \to \infty$. To establish the asymptotic behavior of order statistics for large $k$, we apply the approximation of Gamma function again by letting $z = k$ and $a = -\frac{1}{\kappa}$,

$$\frac{\Gamma\left(k-\frac{1}{\kappa}\right)}{\Gamma(k)} = k^{-1/\kappa}\left(1 + \frac{\frac{1}{\kappa}\left(\frac{1}{\kappa}+1\right)}{2k} + O\left(\frac{1}{k^2}\right)\right).$$

Therefore,

$$\lim_{N\to\infty} \mathbb{E}\left[x_{(k)}\right] = ck^{-1/\kappa}\left(1 + O\left(\frac{1}{k}\right)\right), \quad k \to \infty \Longleftrightarrow \lim_{N\to\infty} \mathbb{E}\left[x_{(k)}\right] = ck^{-1/\kappa} + o\left(k^{-1/\kappa}\right), \quad k \to \infty. \tag{S.12}$$

To prove the almost-sure convergence, it is necessary to construct a single probability space that defines order statistics for all $N \in \mathbb{N}$. Consider a sequence of i.i.d, exponentially distributed random variables and their partial sum,

$$E_1, E_2, E_3, \cdots \sim \exp(1); \qquad S_m = \sum_{i=1}^{m} E_i.$$

The classical Reńyi representation (Arnold et al. (2008), ch.2) states that

$$(Y_{1:N}, \ldots, Y_{N:N}) \overset{d}{=} \left(\frac{S_1}{S_{N+1}}, \ldots, \frac{S_N}{S_{N+1}}\right),$$

where $\overset{d}{=}$ denotes the equality in distribution. Using the transformation (S.9),

$$\frac{x_{(k)}^{(N)}}{ck^{-1/\kappa}} = \left(\frac{S_k/k}{S_{N+1}/N}\right)^{-1/\kappa}, \tag{S.13}$$

where $(N)$ denotes order statistics in the economy of size $N$. Next, we apply the Kolmogorov strong law of large numbers (SLLN) to deduce that[24]

$$\frac{S_N}{N} \to 1 \text{ a.s.}$$

[24] See, e.g., Billingsley (1995), Theorem 22.1.

Let $k = k(N)$ be any sequence of order statistics such that $k(N) \to \infty$.[25], by the SLLN, we obtain that

$$\frac{S_k/k}{S_{N+1}/N} \to 1 \quad \text{a.s.} \implies \frac{x_{(k)}^{(N)}}{ck^{-1/\kappa}} \to 1 \quad \text{a.s.}$$

Combing the above result with (S.12), we arrive at (5.10),

$$x_{(k)} \approx \mathbb{E}\left[x_{(k)}\right] = ck^{-1/\kappa} + o\left(k^{-1/\kappa}\right), \quad \text{a.s.} \quad k \to \infty,$$

as desired. In Appendix S.9, we formally present the proof of the transitivity of almost-sure convergence via asymptotic equivalence in ratios. We refer the reader to that section for technical details.

To establish the universal bound on order statistics in our second statement, we need the Wendel/Gautschi inequality (Gautschi (1959)) for Gamma functions,

$$\left(\frac{x}{x+s}\right)^{1-s} \leq \frac{\Gamma(x+s)}{x^s\Gamma(x)} \leq 1, \qquad x > 0; \quad s \in (0,1).$$

By our assumption $\kappa > 1$, we let $x = k - \frac{1}{\kappa}$ and $s = \frac{1}{\kappa}$. In our context, the above inequality implies

$$\left(\frac{k-\frac{1}{\kappa}}{k}\right)^{1-\frac{1}{\kappa}} \leq \frac{\Gamma(k)}{\left(k-\frac{1}{\kappa}\right)^{1/\kappa}\Gamma\left(k-\frac{1}{\kappa}\right)} \leq 1.$$

We rearrange terms to derive

$$\left(k-\frac{1}{\kappa}\right)^{-1/\kappa} \leq \frac{\Gamma\left(k-\frac{1}{\kappa}\right)}{\Gamma(k)} \leq \left(k-\frac{1}{\kappa}\right)^{-1/\kappa}\left(\frac{k-\frac{1}{\kappa}}{k}\right)^{\frac{1}{\kappa}-1}.$$

Simple algebra reveals that $\left(k-\frac{1}{\kappa}\right)^{-1/\kappa} = k^{-1/\kappa}\left(\frac{k}{k-\frac{1}{\kappa}}\right)^{1/\kappa} = k^{-1/\kappa}\left(\frac{k\kappa}{k\kappa-1}\right)^{1/\kappa}$. As such, the above inequality reduces to

$$k^{-\frac{1}{\kappa}}\left(\frac{k\kappa}{k\kappa-1}\right)^{\frac{1}{\kappa}} \leq \frac{\Gamma\left(k-\frac{1}{\kappa}\right)}{\Gamma(k)} \leq k^{-\frac{1}{\kappa}}\left(\frac{k\kappa}{k\kappa-1}\right). \tag{S.14}$$

The upper and lower bounds of (S.14) are $k$-dependent and shrink as $k$ grows larger. Finally, we note that

$$\frac{k\kappa}{k\kappa-1} \leq \frac{\kappa}{\kappa-1}; \qquad \left(\frac{k\kappa}{k\kappa-1}\right)^{\frac{1}{\kappa}} \geq 1; \qquad \forall k \geq 1, \ \ \forall \kappa > 1.$$

Therefore, by (S.11) we obtain

$$k^{-\frac{1}{\kappa}} \leq \frac{\Gamma\left(k-\frac{1}{\kappa}\right)}{\Gamma(k)} \leq \frac{\kappa}{\kappa-1}k^{-\frac{1}{\kappa}}, \qquad \forall k \geq 1, \ \ \forall \kappa > 1. \tag{S.15}$$

[25]For example, $k = N$, $\sqrt{N}$, $\log N$.

(S.15) is a weaker but $k-$ independent bound than (S.14) that holds universally for all order statistics. To conclude, when $N$ tends to infinity, by (S.11), we derive

$$ck^{-\frac{1}{\kappa}} \leq \lim_{N\to\infty} \mathbb{E}\left[x_{(k)}\right] \leq \frac{\kappa}{\kappa-1} ck^{-\frac{1}{\kappa}}, \qquad \forall k \geq 1, \ \ \forall \kappa > 1.$$

Equivalently, $\mathbb{E}\left[x_{(k)}\right] = \Theta\left(ck^{-\frac{1}{\kappa}}\right)$. Our proof is complete. □

## S.6 Technical Results of Random Graph Models

In this section, we present a set of technical results that establish the rigorous foundation of our random graph model (5.5) with firm-size distribution (5.8). We use these results extensively in the proofs of Propositions 5.3 and 5.4.

### S.6.1 Probabilistic Formulation of Random Graph

We begin with the construction of the probability measure space for inhomogeneous Erdos–Renyi random graphs. Consider a sequence of economies indexed by the number of firms $N = 1, 2, \cdots$. Let $[N] = \{1, 2, \cdots, N\}$. We define

$$S^N = \{0,1\}^{[N]\times[N]}, \qquad \mathcal{F}_s^N = \bigotimes_{(i,j)\in[N]^2} 2^{\{0,1\}}$$

where $\mathcal{F}_s^N$ is the product $\sigma$-algebra generated by all cylinder sets. An element $s \in S^N$ is simply the stochastic adjacency matrix with coordinate mapping: $s : [N] \times [N] \to \{0, 1\}$. It follows that $(S^N, \mathcal{F}_s^N, \mathbb{P}_N)$ forms a probability measure space where

$$\mathbb{P}_N(s) = \prod_{i=1}^{N}\prod_{j=1}^{N} p_{ij}^{s_{ij}} \left(1 - p_{ij}\right)^{1-s_{ij}},$$

and independent Bernoulli probabilities $p_{ij}$ are given by (5.5). The family of probability measures $\{\mathbb{P}_N\}_{N\geq 1}$ is consistent under coordinate restrictions. Hence, by Kolmogorov's extension theorem, when $N \to \infty$, there exists a unique probability measure $\mathbb{P}$ on

$$S := \{0,1\}^{\mathbb{N}\times\mathbb{N}}, \qquad \mathcal{F}_s := \bigotimes_{(i,j)\in\mathbb{N}\times\mathbb{N}} 2^{\{0,1\}},$$

which is consistent with finite-dimensional restrictions. We define $(S, \mathcal{F}_s, \mathbb{P})$ as the canonical probability space in our proof. Given this construction, key statistics of the random graphs are $\mathcal{F}_s-$ measurable.

For example, the degrees of firms for events $s \in S^N$:

$$d_i(s) = \sum_{j=1}^{N} s(i,j); \qquad m_k(s) = \sum_{j=1}^{N} s(j,k)$$

are $\mathcal{F}_s-$ measurable as the countable sum of measurable coordinate maps $s \mapsto s(i,j)$.

### S.6.2 Three Regularity Conditions of the Power-Law Random Graphs

Next, we establish three regularity conditions of the power-law random graphs. These conditions are central pieces in our estimate of GDP volatility and size-attention-ordering results. In this subsection, we start with the baseline case. In the rest of Appendix S.6, we extend the baseline results to scenarios with fewer restrictions on the power-law exponent $(\beta_m, \beta_\omega)$ and to scenarios that focus on particular subsets of the economy.

**Proposition S.1.** *Given the random graph structure* (5.5) *and the distribution of firm sizes* $\boldsymbol{\omega}$ *following* (5.8)*, for any arbitrary permutation* $\pi$ *on* $\boldsymbol{\omega}$*, if* $\beta = \min\{\beta_m, \beta_\omega\} > 1$*, the following technical conditions*

1. $d_i \approx d = o(N), \quad \forall i = 1, 2, \ldots, N.$
2. $\sum_{j \notin \mathcal{D}_k} \omega_{\pi(j)} \approx 1, \quad \forall k = 1, 2, \ldots, N.$

*holds almost surely as* $N \to \infty$*, for all firms simultaneously. If, in addition,*

$$\frac{1}{\beta_m} + \frac{1}{\beta_\omega} < 1, \tag{S.16}$$

*the following additional technical condition*

3. $\sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \approx \frac{m_k}{N}, \quad \forall k = 1, 2, \ldots, N.$

*holds almost surely as* $N \to \infty$*, for all firms simultaneously.*

*Proof.* Given a numerical distribution of firm sizes $\boldsymbol{\omega}$ following (5.8), we fix an arbitrary permutation $\pi$ on $\boldsymbol{\omega}$. For each indexed economy $N = 1, 2, \ldots$, we define three classes of (random) events for each firm (node) $k = 1, 2, \ldots, N \in [N]$,

$$\begin{aligned}
S^A_{N,k} &:= \left\{ s \in S^N, \ \left| \frac{\sum_{i \in \mathcal{M}^{(N)}_k(s)} \omega_{\pi(i)}}{m^{(N)}_k(s)/N} - 1 \right| < \varepsilon_N \right\}, \qquad S^A_N = \bigcap_{k=1}^{N} S^A_{N,k}; \\
S^B_{N,k} &= \left\{ s \in S^N, \ \left| \frac{d^{(N)}_k(s)}{d_N} - 1 \right| < \varepsilon_N, \ \ d_N = o(N) \text{ is deterministic} \right\}, \qquad S^B_N = \bigcap_{k=1}^{N} S^B_{N,k}; \\
S^C_{N,k} &= \left\{ s \in S^N, \ \left| \sum_{j \notin \mathcal{D}^{(N)}_k(s)} \omega_{\pi(j)} - 1 \right| < \varepsilon_N \right\}; \qquad S^C_N = \bigcap_{k=1}^{N} S^C_{N,k};
\end{aligned} \tag{S.17}$$

where $\varepsilon_N \to 0$ as $N \to \infty$.

It is immediate that each subset in the sequences $\{S_N^A\}_{N=1}^{\infty}$, $\{S_N^B\}_{N=1}^{\infty}$, and $\{S_N^C\}_{N=1}^{\infty}$ are $\mathcal{F}_s^N$– measurable and hence $\mathcal{F}_s$– measurable. The following definition formally establishes the notion of almost-sure occurrence.

**Definition S.1.** *Given* (S.17) *with* $\varepsilon_N \to 0$, *the three regularity conditions imposed in Proposition S.1 hold almost surely when* $N \to \infty$ *if the event set*

$$S_N^* \equiv S_N^A \cap S_N^B \cap S_N^C \tag{S.18}$$

*holds* $\mathbb{P}$*-a.s. as* $N \to \infty$. *That is,*

$$\mathbb{P}\left(\lim_{N\to\infty} \inf S_N^*\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} \bigcap_{N\geq n} S_N^*\right) = 1. \tag{S.19}$$

$\mathbb{P}$-a.s. occurrence requires that for all sufficiently large $N$, the three regularity conditions hold in almost all sample realizations of the information networks (eventually always), and the failure only occurs a finite number of times along the sequence $N \to \infty$.

In addition to Definition S.1, we will also use an equivalent notion of almost-sure convergence in the upcoming proofs. In particular, fix $\varepsilon > 0$, let

$$G_N(\varepsilon) := \bigcap_{k=1}^{N} \left\{ \left| \frac{\sum_{i\in\mathcal{M}_k^{(N)}} \omega_{\pi(i)}}{m_k^{(N)}/N} - 1 \right| < \varepsilon \right\} \bigcap \left\{ \left| \frac{d_k^{(N)}}{d_N} - 1 \right| < \varepsilon \right\} \bigcap \left\{ \left| \sum_{j\notin\mathcal{D}_k^{(N)}} \omega_{\pi(j)} - 1 \right| < \varepsilon \right\}. \tag{S.20}$$

Then the three regularity conditions occur almost surely if

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\liminf_{N\to\infty} G_N(\varepsilon)\right) = 1. \tag{S.21}$$

To simplify notation, we suppress the economy index $(N)$ and the sample path $(s)$ whenever convenient.

**Step 1: Convergence with High Probability (w.h.p.)**

We start with the sequence of events $\{S_N^A\}_{N=1}^{\infty}$ that corresponds to the most subtle condition (3). We show that it occurs with high probability (w.h.p.), which provides technical tools for our final results of a.s. convergence.[26] That is,

$$\lim_{N\mapsto\infty} \mathbb{P}\left(S_N^A\right) = 1;$$

with $\varepsilon_N \to 0$.

[26]See the definition of w.h.p. in Frieze and Karoński (2026), pp.xiii. w.h.p. is a much weaker condition than a.s. convergence that concerns only the limit of probability.

We utilize the following concentration inequalities for the summation of independent Bernoulli random variables in probability theory.

**Lemma S.1** (Chung et al. (2003)). *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 $v = \sum_i a_i^2 p_i$. Then we have*

$$\mathbb{P}(X < \mathbb{E}[X] - t) \le \exp\left(-\frac{t^2}{2v}\right),$$

$$\mathbb{P}(X > \mathbb{E}[X] + t) \le \exp\left(-\frac{t^2}{2(\mathrm{var}(X) + at/3)}\right),$$

*where $a = \max_k a_k$.*

Under the random graph formulation given by (5.5) and the assumption $\beta_m > 1$, we obtain $\mathbb{E}[m_j(s)] = \frac{N}{j^{1/\beta_m}}$ for an arbitrary firm $j \in [N]$. Thus by Lemma S.1 we have

$$\mathbb{P}\left(m_j(s) < \frac{N}{j^{1/\beta_m}} - t\right) \le \exp\left(-\frac{t^2}{2N/j^{1/\beta_m}}\right), \tag{S.22}$$

since here $v = \sum_{i=1}^N p_{ij} = \mathbb{E}[m_j(s)]$ with $a_i = 1, \forall i$. Meanwhile,

$$\mathbb{P}\left(m_j(s) > \frac{N}{j^{1/\beta_m}} + t\right) \le \exp\left(-\frac{t^2}{2\,\mathrm{var}(m_j(s)) + 2t/3}\right) \le \exp\left(-\frac{t^2}{2N/j^{1/\beta_m} + 2t/3}\right), \tag{S.23}$$

where the second inequality follows from the fact that for the sum of independent Bernoulli random variables,

$$\mathrm{var}(m_j(s)) = \sum_{i=1}^N p_{ij}(1 - p_{ij}) = \mathbb{E}[m_j(s)] - \sum_{i=1}^N p_{ij}^2 \le \mathbb{E}[m_j(s)] = \frac{N}{j^{1/\beta_m}}.$$

We will use this trick repeatedly. Combining (S.22) and (S.23),

$$\mathbb{P}\left(\frac{N}{j^{1/\beta_m}} - t \le m_j(s) \le \frac{N}{j^{1/\beta_m}} + t\right) \ge 1 - \exp\left(-\frac{t^2}{2N/j^{1/\beta_m}}\right) - \exp\left(-\frac{t^2}{2N/j^{1/\beta_m} + 2t/3}\right). \tag{S.24}$$

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

$$\mathbb{P}\left(\frac{N}{j^{1/\beta_m}} - \left(\frac{N}{j^{1/\beta_m}}\right)^{2/3} \le m_j(s) \le \frac{N}{j^{1/\beta_m}} + \left(\frac{N}{j^{1/\beta_m}}\right)^{2/3}\right)$$

$$\geq 1-\exp\left(-\frac{1}{2}\left(\frac{N}{j^{1/\beta_m}}\right)^{1/3}\right)-\exp\left(-\frac{\left(N/j^{1/\beta_m}\right)^{4/3}}{2N/j^{1/\beta_m}+2\left(N/j^{1/\beta_m}\right)^{2/3}/3}\right). \tag{S.25}$$

By inspection,

$$\lim_{N\to\infty}\mathbb{P}\left(\frac{N}{j^{1/\beta_m}}-\left(\frac{N}{j^{1/\beta_m}}\right)^{2/3}\leq m_j(s)\leq\frac{N}{j^{1/\beta_m}}+\left(\frac{N}{j^{1/\beta_m}}\right)^{2/3}\right)=1.$$

Since $\left(\frac{N}{j^{1/\beta_m}}\right)^{2/3}\to\infty$ for all $j=1,2,\ldots,N$ when $N$ tends to infinity, we have

$$m_j(s)=\frac{N}{j^{1/\beta_m}}+o\left(\frac{N}{j^{1/\beta_m}}\right),\quad \text{w.h.p.}$$

when $N\to\infty$, which implies

$$m_j(s)\approx\frac{N}{j^{1/\beta_m}},\quad \text{w.h.p.}\rightleftharpoons\frac{m_j(s)}{N}\approx\frac{1}{j^{1/\beta_m}},\quad \text{w.h.p.} \tag{S.26}$$

for every firm (node) $j=1,2,\ldots,N$. Note the condition (S.26) does not require assumption (S.16).

On the other hand, for a given permutation $\pi$ and a firm $k=1,2,\ldots,N$, we have

$$\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]=N\mathbb{E}\left[\sum_{i=1}^{N}\omega_{\pi(i)}A_{ik}\right]=N\left[\left(1-\omega_{\pi(k)}\right)\frac{Nk^{-1/\beta_m}-1}{N-1}+\omega_{\pi(k)}\right] \tag{S.27}$$

using (5.5). Note that the second component of (S.27) is derived from the diagonal element (self-loop) that occurs with certainty. We exclude this deterministic component and use Lemma S.1's definition to denote

$$X=N\sum_{i\in\mathcal{M}_k(s),i\neq k}\omega_{\pi(i)}=\sum_{i\neq k}N\omega_{\pi(i)}A_{ik};\qquad a_i=N\omega_{\pi(i)};\qquad X_i=A_{ik},\ \ i=1,2,\ldots,N,\ \ i\neq k$$

for each firm $k\in[N]$. It follows that

$$v_k=N^2\sum_{i\neq k}\omega_{\pi(i)}^2p_{ik}=N^2\left(\sum_{i=1}^{N}\omega_{\pi(i)}^2-\omega_{\pi(k)}^2\right)\frac{Nk^{-1/\beta_m}-1}{N-1}\leq N^2\left(\sum_{i=1}^{N}\omega_{\pi(i)}^2\right)\frac{Nk^{-1/\beta_m}-1}{N-1}, \tag{S.28}$$

which holds for any permutation $\pi$ and firm $k=1,2,\ldots,N$. Next, we employ the definition of firm size distribution (5.8) and the definition of $N$-th partial sum of the Riemann zeta function

$$\sum_{i=1}^{N}\omega_{\pi(i)}^2=\sum_{i=1}^{N}\frac{\left(\pi(i)^{-\frac{1}{\beta\omega}}\right)^2}{\left(\sum_{j=1}^{N}j^{-\frac{1}{\beta\omega}}\right)^2}=\frac{1}{\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]^2}\sum_{j=1}^{N}\left(j^{-\frac{1}{\beta\omega}}\right)^2=\frac{\zeta\left(\frac{2}{\beta_\omega},N\right)}{\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]^2};\qquad \omega_{\max}=\frac{1}{\zeta\left(\frac{1}{\beta_\omega},N\right)}, \tag{S.29}$$

since any permutation $\pi$ is a bijective map and $\beta_\omega > 1$.

Using (S.28) and (S.29), we derive the lower concentration inequality for firm $k = 1, 2, \ldots, N$,

$$\mathbb{P}\left(N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)} < \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)}\right] - t\right) \leq \exp\left(-\frac{t^2}{2N^2 \frac{\zeta\left(\frac{2}{\beta_\omega}, N\right)}{\left[\zeta\left(\frac{1}{\beta_\omega}, N\right)\right]^2} p_k}\right), \tag{S.30}$$

where the $k$-th probability is given by

$$p_k = \frac{Nk^{-1/\beta_m} - 1}{N - 1}, \qquad k \in [N] \tag{S.31}$$

which is k-dependent. Here we use the fact that

$$\begin{aligned}
&\mathbb{P}\left(N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)} < \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)}\right] - t\right) \\
&= \mathbb{P}\left(N \sum_{i\in\mathcal{M}_k(s), i\neq k} \omega_{\pi(i)} + N\omega_{\pi(k)} < \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)}\right] + N\omega_{\pi(k)} - t\right) \\
&= \mathbb{P}\left(N \sum_{i\in\mathcal{M}_k(s), i\neq k} \omega_{\pi(i)} < \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)}\right] - t\right), \qquad \forall k \in [N],\ \forall t > 0,
\end{aligned}$$

since for any possible permutation $\pi$, $N\omega_{\pi(k)}$ is deterministic.

The variance of $X$ in the current context follows

$$\mathrm{var}_k(X) = N^2 \left[\sum_{i\neq k} \omega_{\pi(i)}^2\right] \left(\frac{Nk^{-1/\beta_m} - 1}{N - 1} - \left(\frac{Nk^{-1/\beta_m} - 1}{N - 1}\right)^2\right) \leq N^2 \left[\sum_{i=1}^{N} \omega_{\pi(i)}^2\right] \frac{Nk^{-1/\beta_m} - 1}{N - 1},$$

since $\sum_{i\neq k} \omega_{\pi(i)}^2 \in [0, 1]$. The second concentration inequality then implies

$$\mathbb{P}\left(N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)} > \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)}\right] + t\right) \leq \exp\left(-\frac{t^2}{2N^2 \frac{\zeta\left(\frac{2}{\beta_\omega}, N\right)}{\left[\zeta\left(\frac{1}{\beta_\omega}, N\right)\right]^2} p_k + \frac{2}{3} \frac{Nt}{\zeta\left(\frac{1}{\beta_\omega}, N\right)}}\right), \tag{S.32}$$

where $a = \max_i a_i = N\omega_{\max}$. There are two observations about (S.30) and (S.32). First, the upper

probability bound is strictly larger than the lower probability bound,

$$\exp\left(-\frac{t^2}{2N^2\frac{\zeta\left(\frac{2}{\beta_\omega},N\right)}{\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]^2}p_k}\right)\leq\exp\left(-\frac{t^2}{2N^2\frac{\zeta\left(\frac{2}{\beta_\omega},N\right)}{\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]^2}p_k+\frac{2}{3}\frac{Nt}{\zeta\left(\frac{1}{\beta_\omega},N\right)}}\right);\qquad\forall k,N,t>0.$$

Therefore, we only need to prove the uniform convergence of the upper probability bound. Second, both probability bounds are $k$-dependent, and $p_k$ defined in (S.31) is strictly decreasing with $\lim_{k\to\infty}p_k=0$. Consequently, to prove convergence in probability that is uniform across all firms simultaneously, we need to consider the worst-case scenario when $k=N$, i.e., when the deviation probability is the largest.

Letting $N\to\infty$, (S.8) implies the following asymptotic equivalence when $\beta=\min\{\beta_m,\beta_\omega\}>1$,

$$\frac{\zeta\left(\frac{2}{\beta_\omega},N\right)}{\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]^2}\approx\begin{cases}\frac{\left(1-\beta_\omega^{-1}\right)^2}{\left(1-\frac{2}{\beta_\omega}\right)}\frac{1}{N}; & \beta_\omega>2;\\ \frac{\ln N}{4N}; & \beta_\omega=2;\\ \frac{\left(1-\beta_\omega^{-1}\right)^2\cdot\zeta(\frac{2}{\beta_\omega})}{N^{2-2\beta_\omega^{-1}}}; & 1<\beta_\omega<2,\end{cases}\qquad\frac{N}{\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]}\approx\left(1-\beta_\omega^{-1}\right)N^{1/\beta_\omega}.\tag{S.33}$$

We pick $t_k=\left(\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]\right)^\eta$ with $0<\eta<1$. The limit of expectation is given by

$$\lim_{N\to\infty}\mathbb{E}\left[\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]=\lim_{N\to\infty}\left[\left(1-\omega_{\pi(k)}\right)\frac{Nk^{-1/\beta_m}-1}{N-1}+\omega_{\pi(k)}\right]=k^{-1/\beta_m},\tag{S.34}$$

since

$$\lim_{N\to\infty}\omega_{\pi(k)}=\lim_{N\to\infty}\frac{\pi(k)^{-\frac{1}{\beta_\omega}}}{\zeta\left(\frac{1}{\beta_\omega},N\right)}=0;\qquad\forall k=1,2,\ldots,N\tag{S.35}$$

with diverging Riemann zeta function under $\beta=\min\{\beta_\omega,\beta_m\}>1$ and $\pi(k)\in[N]\geq1$ for every firm $k\in[N]$.

For large k, the expected sum $\mathbb{E}\left[\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]$ converges to 0. In the worst-case scenario with $k=N$,

$$p_N=\Theta\left(N^{-\frac{1}{\beta_m}}\right);\qquad t_N=\Theta\left(N^{\eta\left(1-\frac{1}{\beta_m}\right)}\right),$$

and $p_k\geq\Theta\left(N^{-\frac{1}{\beta_m}}\right)$, $t_k\geq\Theta\left(N^{\eta\left(1-\frac{1}{\beta_m}\right)}\right)$, $\forall k\in[N]$. Therefore, the numerators of both upper-tail and lower-tail probability satisfy

$$t_k^{2\eta}\geq\Theta\left(N^{\left(2\eta(1-1/\beta_m)\right)}\right),\qquad\forall k\in[N].\tag{S.36}$$

We now discuss the denominator of the upper bound (S.32) in the worst-case scenario with $k = N$. Since $p_k$ is strictly decreasing, the minimal magnitude of the first component in the denominator is given by,

$$2N^2 \frac{\zeta\left(\frac{2}{\beta_\omega}, N\right)}{\left[\zeta\left(\frac{1}{\beta_\omega}, N\right)\right]^2} p_k \geq \begin{cases} \Theta\left(N^{(1-1/\beta_m)}\right); & \beta_\omega > 2; \\ \Theta\left(\ln N N^{(1-1/\beta_m)}\right); & \beta_\omega = 2; \\ \Theta\left(N^{(2/\beta_\omega - 1/\beta_m)}\right); & 1 < \beta_\omega < 2; \end{cases} \qquad \forall k \in [N]. \tag{S.37}$$

If the first component dominates the second component (i.e., the upper-tail and the lower-tail bounds share the same magnitude), we only need to consider the first component. In this case, it is easy to see from (S.36) and the above equation that when $\beta_\omega \geq 2$, there is a probability bound that is converging to 0 when $N \to \infty$,

$$\begin{aligned} &\lim_{N\to\infty} \mathbb{P}\left(\left|N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)} - \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)}\right]\right| < o\left(\mathbb{E}\left[N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)}\right]\right)\right) \\ &\leq \exp\left(-\frac{\Theta\left(N^{2\eta(1-1/\beta_m)}\right)}{2\Theta\left(\ln N N^{(1-1/\beta_m)}\right)}\right) \to 0, \end{aligned} \tag{S.38}$$

if $\eta \in (1/2, 1)$, and this converging probability bound holds uniformly for all firms $k = 1, 2, \ldots, N$. This result derives from basic asymptotic theory: any polynomial of $N$ dominates $\ln N$.

If $\beta_\omega \in (1, 2)$, we consider the probability bound in the worst-case scenario,

$$\lim_{N\to\infty} \mathbb{P}\left(\left|N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)} - \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)}\right]\right| < o\left(\mathbb{E}\left[N \sum_{i\in\mathcal{M}_k(s)} \omega_{\pi(i)}\right]\right)\right) \leq \exp\left(-\frac{\Theta\left(N^{2\eta(1-1/\beta_m)}\right)}{2\Theta\left(N^{(2/\beta_\omega - 1/\beta_m)}\right)}\right). \tag{S.39}$$

The probability bound in the limit converges to 0 if we can pick an $\eta \in (0, 1)$ sufficiently close to 1 such that

$$\frac{1}{2} \frac{2\frac{1}{\beta_\omega} - \frac{1}{\beta_m}}{1 - \frac{1}{\beta_m}} < \eta < 1.$$

The parameter restriction then implies that in order to have a feasible set for $\eta$, we must have

$$\frac{1}{\beta_\omega} + \frac{1}{2}\frac{1}{\beta_m} < 1,$$

which is satisfied given our assumption (S.16).[27]

We next consider the second component of the denominator in the upper-tail probability. If the

[27] We only consider the worst-case scenario here. For firms with in-degree $k \leq N^\gamma$, $\gamma \in (0, 1)$, feasible $\eta$ requires only weaker conditions: $\frac{1}{\beta_\omega} + \frac{1}{2}\frac{\gamma}{\beta_m} < 1$.

second component dominates the first component, we can rewrite the upper-tail probability as

$$\exp\left(-\frac{t_k}{\frac{2}{3}\frac{N}{\zeta\left(\frac{1}{\beta_\omega},N\right)}}\right)=\exp\left(-\frac{t_k}{\frac{2}{3}\left(1-\beta_\omega^{-1}\right)N^{1/\beta_\omega}}\right)\leq\exp\left(-\Theta\left(N^{\eta\left(1-\frac{1}{\beta_m}\right)-\frac{1}{\beta_\omega}}\right)\right);\qquad\forall k\in[N].$$

This uniform upper bound is derived based on (S.33) and (S.36). Consequently, there exists an $\eta\in(0,1)$ sufficiently close to 1 such that

$$\lim_{N\to\infty}\mathbb{P}\left(\left|N\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}-\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]\right|<o\left(\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]\right)\right)\to 0,\qquad\forall k\in[N]\tag{S.40}$$

uniformly if (S.16) holds[28], i.e.,

$$\frac{1}{\beta_m}+\frac{1}{\beta_\omega}<1.$$

In this case, we can pick $\eta\in(0,1)$ such that

$$\frac{\frac{1}{\beta_\omega}}{1-\frac{1}{\beta_m}}<\eta<1.$$

Therefore, following the same techniques as in (S.25) with the fact that $\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]\to\infty$ as $N$ tends to infinity,

$$N\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}=\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]+o\left(\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]\right),\quad\text{w.h.p.},\quad\forall k\in[N].$$

Therefore,

$$\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\approx\mathbb{E}\left[\sum_{i\in\mathcal{M}_k(s)}\omega_{\pi(i)}\right]=\left[\left(1-\omega_{\pi(k)}\right)\frac{Nk^{-1/\beta_m}-1}{N-1}+\omega_{\pi(k)}\right];\quad\text{w.h.p.},\quad\forall k\in[N].\tag{S.41}$$

Consolidating (S.26), (S.41), and (S.34), we observe that the right-hand sides (RHSs) of these equations coincide and are deterministic in the limit $N\to\infty$. Next, we present a technical lemma.

**Lemma S.2 (Transitivity of Asymptotic Equivalence w.h.p.).** *Let $(x_n)_{n\geq1}$ and $(y_n)_{n\geq1}$ be real-valued random variables, and let $(c_n)_{n\geq1}$ and $(d_n)_{n\geq1}$ be deterministic real sequences such that when $n\to\infty$*

$$c_n\to L,\qquad d_n\to L,\qquad\text{for some } L\in\mathbb{R}\setminus\{0\}.$$

[28](S.16) is a structural restriction for convergence in probability for firms with sufficiently small attention ($p_k\to0$ and $k\to N$). While the Bernstein-type inequality in Lemma S.1 can be refined to a stronger version of the Bennett-Hoeffding inequality and the Chernoff bound, they cannot resolve this structural issue.

*Meanwhile,*

$$x_n \approx c_n, \ \ w.h.p. \qquad y_n \approx d_n, \ \ w.h.p.$$

*Then*

$$x_n \approx y_n, \ \ w.h.p.$$

*Proof.* Since $c_n \to L \neq 0$ and $d_n \to L \neq 0$, we have $\frac{c_n}{d_n} \to 1$ deterministically. Fix $\varepsilon > 0$. Choose $\delta \in (0,1)$ sufficiently small such that

$$\frac{(1+\delta)^2}{1-\delta} < 1+\varepsilon \qquad \text{and} \qquad \frac{(1-\delta)^2}{1+\delta} > 1-\varepsilon. \tag{S.42}$$

This is possible because the left-hand sides converge to 1 as $\delta \to 0$. Define the event set

$$E_n := \left\{\left|\frac{x_n}{c_n} - 1\right| < \delta\right\} \cap \left\{\left|\frac{y_n}{d_n} - 1\right| < \delta\right\} \cap \left\{\left|\frac{c_n}{d_n} - 1\right| < \delta\right\}.$$

On $E_n$ we have the bounds $1-\delta < \frac{x_n}{c_n} < 1+\delta$, $1-\delta < \frac{y_n}{d_n} < 1+\delta$, $1-\delta < \frac{c_n}{d_n} < 1+\delta$. In particular, on $E_n$ the second inequality implies $\frac{d_n}{y_n} \in \left(\frac{1}{1+\delta}, \frac{1}{1-\delta}\right)$.

Using the factorization

$$\frac{x_n}{y_n} = \frac{x_n}{c_n} \cdot \frac{c_n}{d_n} \cdot \frac{d_n}{y_n},$$

we obtain on $E_n$:

$$\frac{(1-\delta)^2}{1+\delta} \ \leq \ \frac{x_n}{y_n} \ \leq \ \frac{(1+\delta)^2}{1-\delta}.$$

By (S.42), this implies on $E_n$ that

$$1-\varepsilon < \frac{x_n}{y_n} < 1+\varepsilon,$$

which implies the event sets satisfy $E_n \subseteq \left\{1-\varepsilon < \frac{x_n}{y_n} < 1+\varepsilon\right\}$. By reverse inclusion of complement set,

$$\left\{\left|\frac{x_n}{y_n} - 1\right| \geq \varepsilon\right\} \subseteq E_n^c.$$

Therefore, by the union bound and De Morgan's laws,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{x_n}{y_n} - 1\right| \geq \varepsilon\right) &\leq \mathbb{P}(E_n^c) \\ &\leq \mathbb{P}\left(\left|\frac{x_n}{c_n} - 1\right| \geq \delta\right) + \mathbb{P}\left(\left|\frac{y_n}{d_n} - 1\right| \geq \delta\right) + \mathbf{1}\left\{\left|\frac{c_n}{d_n} - 1\right| \geq \delta\right\},\end{aligned}$$

where $\mathbf{1}$ denotes the indicator function for the deterministic set. By assumption, the first two probabilities tend to 0 as $n \to \infty$. Since $\frac{c_n}{d_n} \to 1$ deterministically, the indicator term is 0 for all sufficiently large $n$. Hence

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{x_n}{y_n} - 1\right| \geq \varepsilon\right) = 0.$$

Because $\varepsilon > 0$ was arbitrary, this proves $x_n \approx y_n$, w.h.p. □

Consolidating (S.26), (S.41), and (S.34) and invoking Lemma S.2,

$$\lim_{N\to\infty} \mathbb{P}\left(\left|\frac{\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}}{m_k/N} - 1\right| \geq \varepsilon\right) = 0, \quad \forall \varepsilon > 0, \qquad \forall k \in [N]. \tag{S.43}$$

Note that (S.43) is a uniform convergence condition in probability. To show that this condition holds simultaneously for all firms, we note that by (S.25) and (S.38)-(S.40),

$$\sum_{k=1}^{N} \mathbb{P}\left(\left\{\left|\frac{\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}}{m_k/N} - 1\right| \geq \varepsilon\right\}\right) \leq 4N \exp^{\left(-\Theta(P_{\max}(N))\right)} \to 0, \tag{S.44}$$

where $P_{\max}(N)$ denotes the largest tail probability in terms of positive polynomial functions among (S.25) and (S.38)-(S.40) when $k = N$. Here we use the fact that an exponentially decaying function dominates any polynomial. We present this convergence result formally in Lemma S.4.

Therefore, using De Morgan's law again for countable intersections,

$$\begin{aligned}\mathbb{P}\left(\bigcap_{k=1}^{N}\left\{\left|\frac{\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}}{m_k/N} - 1\right| < \varepsilon\right\}\right) &= 1 - \mathbb{P}\left(\bigcup_{k=1}^{N}\left\{\left|\frac{\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}}{m_k/N} - 1\right| \geq \varepsilon\right\}\right) \\ &\geq 1 - \sum_{k=1}^{N} \mathbb{P}\left(\left\{\left|\frac{\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}}{m_k/N} - 1\right| \geq \varepsilon\right\}\right).\end{aligned}$$

Taking limits on both sides of the above equation and using (S.43) and (S.44),

$$\lim_{N\to\infty} \mathbb{P}\left(\bigcap_{k=1}^{N}\left\{\left|\frac{\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}}{m_k/N} - 1\right| < \varepsilon\right\}\right) = \lim_{N\mapsto\infty} \mathbb{P}\left(S_N^A\right) = 1, \quad \forall \varepsilon > 0^{29}. \tag{S.45}$$

Therefore, by definition,

$$\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} \approx \frac{m_k}{N}, \quad \text{w.h.p.}$$

for all firms $k = 1, 2, \ldots, N$ simultaneously.

The proofs of w.h.p. convergence for the second and third conditions follow the same procedure and are much simpler. Using the result in the proof of Lemma 5.2, the expected out-degree for a given firm $i \in [N]$ is given by

$$v_i(N) \equiv \mathbb{E}[d_i] = \frac{N}{N-1}\left(\zeta(\frac{1}{\beta_m}, N) - \frac{1}{i^{\frac{1}{\beta m}}}\right). \tag{S.46}$$

Without loss of generality, we define the deterministic Riemann zeta partial-sum function for the

[29] We will use this technique of simultaneous convergence several times in the subsequent proof, and thus will omit details hereafter.

economy indexed by $N$,

$$d_N \equiv \zeta(\frac{1}{\beta_m}, N) = o(N), \qquad \beta_m > 1.$$

Invoking the concentration inequality in Lemma S.1 again,

$$\mathbb{P}\left(v_i(N) - t \leq d_i \leq v_i(N) + t\right) \geq 1 - \exp\left(-\frac{t^2}{2v_i(N)}\right) - \exp\left(-\frac{t^2}{2v_i(N) + 2t/3}\right),$$

where use the trick $\text{var}\left(d_i\right) \leq \mathbb{E}\left[d_i\right]$ again. Let $t = \left(v_i(N)\right)^{2/3}$ in the above inequality. We obtain

$$\begin{aligned}
&\mathbb{P}\left(v_i(N) - \left[v_i(N)\right]^{2/3} \leq d_i \leq v_i(N) + \left[v_i(N)\right]^{2/3}\right) \\
&\geq 1 - \exp\left(-\frac{1}{2}\left(v_i(N)\right)^{1/3}\right) - \exp\left(-\frac{\left(v_i(N)\right)^{4/3}}{2v_i(N) + 2\left(v_i(N)\right)^{2/3}/3}\right).
\end{aligned} \tag{S.47}$$

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

$$\lim_{N\to\infty} \mathbb{P}\left(v_i(N) - \left[v_i(N)\right]^{2/3} \leq d_i \leq v_i(N) + \left[v_i(N)\right]^{2/3}\right) = 1.$$

Since $\inf_{i\in[N]} v_i(N) \to \infty$ uniformly when $N$ tends to infinity,

$$d_i = \zeta(\frac{1}{\beta_m}, N) + o\left(\zeta(\frac{1}{\beta_m}, N)\right), \quad \text{w.h.p.} \tag{S.48}$$

or equivalently

$$d_i \approx d = \zeta(\frac{1}{\beta_m}, N), \quad \text{w.h.p.}$$

for each firm $i = 1, 2, \ldots, N$ uniformly, where $\zeta(\frac{1}{\beta_m}, N)$ denotes the divergent Riemann zeta series as $N \to \infty$. Using De Morgan's law again for countable intersections,

$$\lim_{N\to\infty} \mathbb{P}\left(\bigcap_{i=1}^{N}\left\{\left|\frac{d_i}{d_N} - 1\right| < \varepsilon\right\}\right) = \lim_{N\mapsto\infty} \mathbb{P}\left(S_N^B\right) = 1, \quad \forall \varepsilon > 0$$

for all firms $i \in [N]$ simultaneously. The simultaneous convergence in probability follows the same logic as in (S.44).

Next, we transform the last condition as

$$\sum_{j\notin\mathcal{D}_k^{(N)}(s)} \omega_{\pi(j)} \approx 1 = \sum_{j=1}^{N} \omega_{\pi(j)} \Longleftrightarrow \lim_{N\to\infty} \sum_{j=1}^{N} \omega_{\pi(j)} A_{kj} = 0; \qquad \forall k = 1, 2, \ldots, N.$$

Therefore, we reformulate the last event in (S.17) as

$$S_{N,k}^{C} = \left\{ s \in S^N, \quad \left| \sum_{j=1}^{N} \omega_{\pi(j)} A_{kj}(s) \right| < \varepsilon_N \right\}; \qquad S_N^C = \bigcap_{k=1}^{N} S_{N,k}^C$$

with $\varepsilon_N \to 0$. It is convenient to define the size-weighted row sum as

$$\varrho_k(N) \equiv \sum_{j=1}^{N} \omega_{\pi(j)} A_{kj}, \qquad \forall k = 1, 2, \ldots, N. \tag{S.49}$$

Using (5.5) and (5.8), the expectation of $\varrho_k(N)$ is given by

$$\mathbb{E}\left[\varrho_k(N)\right] = \sum_{j=1}^{N} \omega_{\pi(j)} p_{kj} = \frac{\sum_{j=1}^{N} \pi(j)^{-\frac{1}{\beta_\omega}} \left( \frac{N}{j^{1/\beta_m}} - 1 \right)}{\zeta\left(\frac{1}{\beta_\omega}, N\right)(N-1)} + \omega_{\pi(k)} \left( 1 - \frac{N k^{-\frac{1}{\beta_m}} - 1}{N-1} \right). \tag{S.50}$$

After simple algebra, (S.50) reduces to

$$\mathbb{E}\left[\varrho_k(N)\right] = \frac{N \sum_{j=1}^{N} \left[ \pi(j)^{-\frac{1}{\beta_\omega}} j^{-\frac{1}{\beta_m}} \right]}{(N-1)\,\zeta\left(\frac{1}{\beta_\omega}, N\right)} - \frac{\sum_{j=1}^{N} \left[ \pi(j)^{-\frac{1}{\beta_\omega}} \right]}{(N-1)\,\zeta\left(\frac{1}{\beta_\omega}, N\right)} + \omega_{\pi(k)} \frac{N}{N-1} \left( 1 - k^{-\frac{1}{\beta_m}} \right).$$

Since $\pi$ is an arbitrary bijective permutation function, $\sum_{j=1}^{N} \left[ \pi(j)^{-\frac{1}{\beta_\omega}} \right] = \sum_{j=1}^{N} j^{-\frac{1}{\beta_\omega}} = \zeta(\frac{1}{\beta_\omega}, N)$ by definition. Consequently,

$$\mathbb{E}\left[\varrho_k(N)\right] = \frac{N \sum_{j=1}^{N} \left[ \pi(j)^{-\frac{1}{\beta_\omega}} j^{-\frac{1}{\beta_m}} \right]}{(N-1)\,\zeta\left(\frac{1}{\beta_\omega}, N\right)} - \frac{1}{(N-1)} + \omega_{\pi(k)} \frac{N}{N-1} \left( 1 - k^{-\frac{1}{\beta_m}} \right); \qquad \forall k = 1, 2, \ldots, N. \tag{S.51}$$

By construction, $\varrho_k(N)$ is a nonnegative random variable. Hence by (S.50), $\mathbb{E}\left[\varrho_k(N)\right] \geq 0, \quad \forall k \in [N]$. In addition, for any possible permutation $\pi$, the rearrangement inequality akin to (S.99) applies to (S.51),

$$\sum_{j=1}^{N} \left[ \pi(j)^{-\frac{1}{\beta_\omega}} j^{-\frac{1}{\beta_m}} \right] \leq \sum_{j=1}^{N} \left[ \pi^*(j)^{-\frac{1}{\beta_\omega}} j^{-\frac{1}{\beta_m}} \right] = \sum_{j=1}^{N} j^{-\left(\frac{1}{\beta_m} + \frac{1}{\beta_\omega}\right)} = \zeta\left( \frac{1}{\beta_m} + \frac{1}{\beta_\omega}, N \right), \qquad \forall \pi,$$

where $\pi^*(j) = j$ is the identity permutation mapping. The rearrangement inequality leads to the

following asymptotic behavior,

$$0 \leq \lim_{N\to\infty} \mathbb{E}\left[\varrho_k(N)\right] \leq \lim_{N\to\infty} \left( \frac{N\zeta\left(\frac{1}{\beta_m}+\frac{1}{\beta_\omega},N\right)}{(N-1)\,\zeta\left(\frac{1}{\beta_\omega},N\right)} - \frac{1}{(N-1)} + \omega_{\pi(k)}\frac{N}{N-1}\left(1-k^{-\frac{1}{\beta_m}}\right)\right) = 0,\ \forall \pi,\ \forall k \in [N], \tag{S.52}$$

where we use the fact that $\beta = \min\{\beta_\omega, \beta_m\} > 1$ and $\pi(k) \in [N] \geq 1$ so that $\lim_{N\to\infty} \omega_{\pi(k)} = \lim_{N\to\infty} \frac{\pi(k)^{-\frac{1}{\beta_\omega}}}{\zeta\left(\frac{1}{\beta_\omega},N\right)} = 0$ as illustrated before.

Now we are ready to invoke the concentration inequality. In the current context, by (S.50),

$$\operatorname{var}\left(\varrho_k(N)\right) = \sum_{j=1}^{N} \omega_{\pi(j)}^2 p_{kj}\left(1-p_{kj}\right) \leq \nu_k = \sum_{j=1}^{N} \omega_{\pi(j)}^2 p_{kj} \leq \frac{1+\zeta\left(\frac{2}{\beta_\omega},N\right)}{\zeta\left[\left(\frac{1}{\beta_\omega},N\right)\right]^2}.$$

It follows that

$$\begin{aligned}
&\mathbb{P}\left(\mathbb{E}\left[\varrho_k(N)\right] - t \leq \varrho_k(N) \leq \mathbb{E}\left[\varrho_k(N)\right] + t\right) \\
&\geq 1 - \exp\left(-2\frac{t^2\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]^2}{1+\zeta\left(\frac{2}{\beta_\omega},N\right)}\right) - \exp\left(-\frac{2t^2\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]^2}{2\left(1+\zeta\left(\frac{2}{\beta_\omega},N\right)\right)+\frac{2}{3}\zeta\left(\frac{1}{\beta_\omega},N\right)t}\right).
\end{aligned}$$

Note that the above inequality holds uniformly for all firms and the tail probability bounds are independent of $k \in [N]$. Using (S.8) and (S.33), when $\beta_\omega > 2$,

$$\frac{\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]^2}{\zeta\left(\frac{2}{\beta_\omega},N\right)} = \Theta(N); \qquad \zeta\left(\frac{1}{\beta_\omega},N\right) = \Theta\left(N^{1-\frac{1}{\beta_\omega}}\right).$$

Without loss of generality, we define

$$t \equiv \varepsilon_N = N^{-\varsigma}; \qquad 0 < \varsigma < \min\left\{\frac{1}{2}, 1-\frac{1}{\beta_\omega}\right\}.$$

Then, the concentration inequality transforms into

$$\mathbb{P}\left(\left|\varrho_k(N) - \mathbb{E}\left[\varrho_k(N)\right]\right| \leq \varepsilon_N\right) \geq 1 - \exp\left(-2\Theta\left(N^{\left(1-2\varsigma\right)}\right)\right) - \exp\left(-\Theta\left(N^{\left(1-\frac{1}{\beta_\omega}-\varsigma\right)}\right)\right). \tag{S.53}$$

By construction,

$$\lim_{N\to\infty}\mathbb{P}\left(\left|\varrho_k(N)-\mathbb{E}\left[\varrho_k(N)\right]\right|\leq\varepsilon_N\right)=1;\qquad\lim_{N\to\infty}\varepsilon_N=0;\qquad\forall k\in[N].$$

On the other hand, by (S.52) $\mathbb{E}\left[\varrho_k(N)\right]\to 0$ as $N$ tends to infinity. To conclude this step, fix $\varepsilon>0$, there exists a deterministic $N_0$ such that for all $N\geq N_0$,

$$\left|\mathbb{E}\left[\rho_k(N)\right]\right|+\varepsilon_N<\varepsilon.$$

Define random events $F_{N,k}=\left\{s\in S^N,\left|\varrho_k(N)(s)-\mathbb{E}\left[\varrho_k(N)\right]\right|\leq\varepsilon_N\right\},\forall k\in[N],\ \forall N\geq N_0$. By the triangle inequality,

$$\begin{aligned}\left|\rho_k(N)(s)\right|&\leq\left|\mathbb{E}\left[\rho_k(N)\right]\right|+\left|\rho_k(N)(s)-\mathbb{E}\left[\rho_k(N)\right]\right|\\&\leq\left|\mathbb{E}\left[\rho_k(N)\right]\right|+\varepsilon_N<\varepsilon,\qquad\forall k\in[N],\ \forall s\in S^N;\ \forall N\geq N_0.\end{aligned}$$

Therefore, by reverse inclusion of the complement set, $\forall N\geq N_0$,

$$\left\{s\in S^N,\left|\varrho_k(N)(s)\right|\geq\varepsilon\right\}\subseteq F_{N,k}^c:=\left\{s\in S^N,\left|\varrho_k(N)(s)-\mathbb{E}\left[\varrho_k(N)\right]\right|>\varepsilon_N\right\}.$$

Taking probability limit,

$$\lim_{N\to\infty}\mathbb{P}\left(\left|\rho_k(N)\right|\geq\varepsilon\right)\leq\lim_{N\to\infty}\mathbb{P}\left(\left|\rho_k(N)-\mathbb{E}\left[\rho_k(N)\right]\right|>\varepsilon_N\right)=0.\qquad\forall k\in[N],\ \forall\varepsilon>0.$$

Equivalently,

$$\lim_{N\to\infty}\mathbb{P}\left(\left|\rho_k(N)\right|<\varepsilon\right)=1;\qquad\forall k\in[N],\ \forall\varepsilon>0.$$

Therefore, for each firm $k\in[N]$, $\rho_k(N)\to 0$ with high probability (w.h.p.) uniformly. Finally, we establish simultaneous convergence in probability by taking the union bound of countable intersections,

$$\lim_{N\to\infty}\mathbb{P}\left(\bigcap_{k=1}^{N}\left\{\left|\rho_k(N)=\sum_{j=1}^{N}\omega_{\pi(j)}A_{kj}\right|<\varepsilon\right\}\right)=\lim_{N\mapsto\infty}\mathbb{P}\left(S_N^{\mathrm{C}}\right)=1,\ \ \forall\varepsilon>0.$$

Again, we apply the summability of exponentially decaying tail probabilities as in (S.44). If $1<\beta_\omega\leq 2$, $\zeta\left(\frac{2}{\beta_\omega},N\right)$ is either $\ln N$ or finite (convergent). In this case, $\frac{\left[\zeta\left(\frac{1}{\beta_\omega},N\right)\right]^2}{\zeta\left(\frac{2}{\beta_\omega},N\right)}=\Omega\left(\frac{N^{\left(2-\frac{2}{\beta_\omega}\right)}}{\ln N}\right)$, the proof of simultaneous convergence in probability is identical. We omit these details.

**Step 2: Almost-Sure Convergence**

In step 1, we have developed the necessary tools for proving the almost-sure occurrence of the three

technical conditions as defined in (S.17), (S.18), and (S.20). Additionally, we need the Borel-Cantelli lemma in probability theory.

**Lemma S.3 (The First Borel-Cantelli Lemma, Billingsley (1995) Theorem 4.3).** *Let $\{\mathcal{A}_N\}_{N=1}^{\infty}$ be a sequence of measurable sets, suppose*

$$\sum_{N=1}^{\infty} \mathbb{P}(\mathcal{A}_N) < \infty,$$

*then the probability that infinitely many of them occur is 0, that is*

$$\mathbb{P}\left(\limsup_{N\to\infty} \mathcal{A}_N\right) = \mathbb{P}\left(\lim_{N\to\infty} \bigcap_{N=1}^{\infty} \bigcup_{k\geq N} \mathcal{A}_k\right) = \mathbb{P}(\{\mathcal{A}_N \ i.o\}) = 0.$$

We begin with the first event $\left\{S_N^A\right\}_{N=1}^{\infty}$. Our goal is to show that the cumulative probability of failures is summable when $N \to \infty$. We follow the same procedure as in the w.h.p. convergence. By (S.25),

$$\begin{aligned}
&\mathbb{P}\left(\left\{\left|\frac{m_j(s)}{\frac{N}{j^{1/\beta_m}}} - 1\right| \leq \varepsilon_{N,j}\right\}\right) \\
&\geq 1 - \exp\left(-\frac{1}{2}\left(\frac{N}{j^{1/\beta_m}}\right)^{1/3}\right) - \exp\left(-\frac{\left(N/j^{1/\beta_m}\right)^{4/3}}{2N/j^{1/\beta_m} + 2\left(N/j^{1/\beta_m}\right)^{2/3}/3}\right) \\
&\geq 1 - \exp\left(-\frac{1}{2}N^{\frac{1}{3}\left(1-\frac{1}{\beta_m}\right)}\right) - \exp\left(-\Theta\left(N^{\frac{2}{3}\left(1-\frac{1}{\beta_m}\right)}\right)\right); \qquad \forall j = 1, 2, \ldots, N.
\end{aligned}$$

The above equation provides a uniform probability bound for each firm $j \in [N]$, and we define

$$\varepsilon_{N,j} = \left(\frac{N}{j^{1/\beta_m}}\right)^{-1/3} \leq \varepsilon_N = N^{-\frac{1}{3}\left(1-\frac{1}{\beta_m}\right)} \to 0, \qquad \forall j \in [N].$$

For a fixed economy with index $N$, a failure event occurs with probability

$$\mathbb{P}\left(\left\{\bigcap_{j=1}^{N}\left\{\left|\frac{m_j(s)}{\frac{N}{j^{1/\beta_m}}} - 1\right| \leq \varepsilon_N\right\}\right\}^c\right) = \mathbb{P}\left(\bigcup_{j=1}^{N}\left\{\left|\frac{m_j(s)}{\frac{N}{j^{1/\beta_m}}} - 1\right| > \varepsilon_N\right\}\right) \leq \sum_{j=1}^{N}\mathbb{P}\left(\left\{\left|\frac{m_j(s)}{\frac{N}{j^{1/\beta_m}}} - 1\right| > \varepsilon_N\right\}\right).$$

By the uniform probability bound,

$$\begin{aligned}
\sum_{j=1}^{N}\mathbb{P}\left(\left\{\left|\frac{m_j(s)}{\frac{N}{j^{1/\beta_m}}} - 1\right| > \varepsilon_N\right\}\right) &< \sum_{j=1}^{N}\left[\exp\left(-\frac{1}{2}N^{\frac{1}{3}\left(1-\frac{1}{\beta_m}\right)}\right) + \exp\left(-\Theta\left(N^{\frac{2}{3}\left(1-\frac{1}{\beta_m}\right)}\right)\right)\right] \\
&= N\left[\exp\left(-\frac{1}{2}N^{\frac{1}{3}\left(1-\frac{1}{\beta_m}\right)}\right) + \exp\left(-\Theta\left(N^{\frac{2}{3}\left(1-\frac{1}{\beta_m}\right)}\right)\right)\right].
\end{aligned}$$

Following the same logic, a failure event in the entire sequence of economies $N = 1, 2, \ldots$ has probability

$$\mathbb{P}\left(\left\{\bigcap_{N=1}^{\infty} F_N\right\}^c\right) \leq \sum_{N=1}^{\infty} \mathbb{P}\left(F_N^c\right) \leq \sum_{N=1}^{\infty}\sum_{j=1}^{N} \mathbb{P}\left(\left\{\left|\frac{m_j(s)}{\frac{N}{j^{1/\beta_m}}} - 1\right| > \varepsilon_N\right\}\right); \qquad F_N \equiv \bigcap_{j=1}^{N}\left\{\left|\frac{m_j(s)}{\frac{N}{j^{1/\beta_m}}} - 1\right| \leq \varepsilon_N\right\}. \tag{S.54}$$

An important observation is that with exponentially decaying tail probability, the double-index probability sum converges,

$$\sum_{N=1}^{\infty}\sum_{j=1}^{N} \mathbb{P}\left(\left\{\left|\frac{m_j(s)}{\frac{N}{j^{1/\beta_m}}} - 1\right| > \varepsilon_N\right\}\right) < \sum_{N=1}^{\infty} N\left[\exp\left(-\frac{1}{2}N^{\frac{1}{3}\left(1-\frac{1}{\beta_m}\right)}\right) + \exp\left(-\Theta\left(N^{\frac{2}{3}\left(1-\frac{1}{\beta_m}\right)}\right)\right)\right] < \infty. \tag{S.55}$$

(S.55) follows from the next Lemma, which is celebrated in asymptotic analysis.

**Lemma S.4.** *Let $P$ be a real polynomial in $N$ and let $c > 0$, $\vartheta > 0$. Then*

$$P(N)e^{-cN^{\vartheta}} \to 0 \qquad \text{as } N \to \infty.$$

*In addition*

$$\sum_{N=1}^{\infty} |P(N)|e^{-cN^{\vartheta}} < \infty.$$

Using (S.54) and (S.55), we deduce from the Borel-Cantelli Lemma S.3 that

$$\sum_{N=1}^{\infty} \mathbb{P}\left(F_N^c\right) < \infty \Rightarrow \mathbb{P}\left(\limsup_{N\to\infty} F_N^c\right) = 0.$$

Next, we apply a basic set-theory identity to obtain

$$\limsup_{N\to\infty} F_N^c = \left(\liminf_{N\to\infty} F_N\right)^c \implies \mathbb{P}\left(\liminf_{N\to\infty} F_N\right) = 1 - \mathbb{P}\left(\limsup_{N\to\infty} F_N^c\right) = 1. \tag{S.56}$$

Therefore, we proved that when $N \to \infty$,

$$m_j \approx \frac{N}{j^{1/\beta_m}}, \quad \text{a.s.} \quad \forall j \in [N] \tag{S.57}$$

for all firms simultaneously. Following the same procedure as above, we use Lemma S.3, Lemma S.4, (S.34), and the uniform probability bound (S.38)-(S.40) to deduce that as $N$ tends to infinity,

$$\sum_{i\in\mathcal{M}_j(s)} \omega_{\pi(i)} \approx \mathbb{E}\left[\sum_{i\in\mathcal{M}_j(s)} \omega_{\pi(i)}\right] = \left[\left(1 - \omega_{\pi(j)}\right)\frac{Nj^{-1/\beta_m} - 1}{N-1} + \omega_{\pi(j)}\right] \longrightarrow j^{-1/\beta_m}; \quad \text{a.s.} \qquad \forall j \in [N],$$

for all firms simultaneously. Given the deterministic limit (S.34), it is easy to show that the transitivity of asymptotic equivalence akin to Lemma S.2 also holds with respect to almost-sure convergence.[30] As such,

$$\mathbb{P}\left(\liminf_{N\to\infty}\left\{\bigcap_{k=1}^{N}\left\{\left|\frac{\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}}{m_k/N}-1\right|<\varepsilon\right\}\right\}\right)=\mathbb{P}\left(\liminf_{N\to\infty}S_N^A\right)=1,\quad\forall\varepsilon>0, \tag{S.58}$$

or equivalently, $\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\approx m_k/N$ a.s. for all firms simultaneously when $N\to\infty$. Intuitively, in the entire sequence, failures of the first technical condition only occur a finite number of times, so that when the economy $N$ is sufficiently large, it always holds with probability 1.

Consolidating (S.46), (S.47), (S.52), and (S.53), the remaining two events share an identical mathematical structure with the first event. Specifically, fix an economy of size $N$, for each firm $k=1,2,\ldots,N\in[N]$ in this economy, there exist (1) concentrations around means with $\varepsilon_N$ distance shrinking at a polynomial rate in $N$; (2) a uniform probability bound with exponentially decaying tail probability in the form of $e^{-cN^{\vartheta}}$. Therefore, we apply the above procedure with Lemma S.3 and Lemma S.4 to deduce the almost-sure occurrence,

$$\mathbb{P}\left(\liminf_{N\to\infty}\left\{\bigcap_{k=1}^{N}\left\{\left|\frac{d_k}{d_N}-1\right|<\varepsilon\right\}\right\}\right)=\mathbb{P}\left(\liminf_{N\to\infty}S_N^B\right)=1,\quad\forall\varepsilon>0; \tag{S.59}$$

$$\mathbb{P}\left(\liminf_{N\to\infty}\left\{\bigcap_{k=1}^{N}\left\{\left|\sum_{j=1}^{N}\omega_{\pi(j)}A_{kj}\right|<\varepsilon\right\}\right\}\right)=\mathbb{P}\left(\liminf_{N\to\infty}\left\{\bigcap_{k=1}^{N}\left\{\left|\sum_{j\notin\mathcal{D}_k^{(N)}}\omega_{\pi(j)}-1\right|<\varepsilon\right\}\right\}\right)=\mathbb{P}\left(\liminf_{N\to\infty}S_N^C\right)=1,\quad\forall\varepsilon>0. \tag{S.60}$$

For brevity, we omit the details of the derivation. Finally, by (S.58)-(S.60),

$$\mathbb{P}\left(\lim_{N\to\infty}\inf S_N^*\right)=\mathbb{P}\left(\bigcup_{n=1}^{\infty}\bigcap_{N\geq n}\left\{S_N^A\cap S_N^B\cap S_N^C\right\}\right).$$

By the distributive rule of intersections,

$$\bigcap_{N\geq n}\left(S_N^A\cap S_N^B\cap S_N^C\right)=\left(\bigcap_{N\geq n}S_N^A\right)\cap\left(\bigcap_{N\geq n}S_N^B\right)\cap\left(\bigcap_{N\geq n}S_N^C\right).$$

It follows that

$$\begin{aligned}\bigcup_{n=1}^{\infty}\bigcap_{N\geq n}\left\{S_N^A\cap S_N^B\cap S_N^C\right\}&=\left(\bigcup_{n=1}^{\infty}\bigcap_{N\geq n}S_N^A\right)\bigcap\left(\bigcup_{n=1}^{\infty}\bigcap_{N\geq n}S_N^B\right)\bigcap\left(\bigcup_{n=1}^{\infty}\bigcap_{N\geq n}S_N^C\right)\\&=\left(\liminf_{N\to\infty}S_N^A\right)\bigcap\left(\liminf_{N\to\infty}S_N^B\right)\bigcap\left(\liminf_{N\to\infty}S_N^C\right),\end{aligned}$$

where the first equality follows from the fact that $\left\{\bigcap_{N\geq n}S_N^i\right\}_{n=1}^{\infty}$, $i=A,B,C$ are nested increasing

[30] Standard $\varepsilon-\delta$ techniques with the triangle inequality will lead to this result. We omit these details to avoid redundancy.

sequences of sets. Therefore, by (S.56) and De Morgan's law,

$$\mathbb{P}\left(\lim_{N\to\infty}\inf S_N^*\right)=\mathbb{P}\left(\left(\liminf_{N\to\infty}S_N^A\right)\bigcap\left(\liminf_{N\to\infty}S_N^B\right)\bigcap\left(\liminf_{N\to\infty}S_N^C\right)\right)=1.$$

In other words, we have proved the almost-sure occurrence of the three technical conditions as defined in (S.17), (S.18), and (S.20) when $N\to\infty$. The proof of this proposition is now complete. □

### S.6.3 Regularity Conditions under Alternative Assumptions

In this subsection, we present a set of additional results regarding the three regularity conditions proposed in the previous section. These results will be used in our subsequent proofs of Propositions 5.3 and 5.4.

**Corollary S.1.** *If $\beta_m>1,\beta_\omega>1$ and $\frac{1}{\beta_m}+\frac{1}{\beta_\omega}=1$, for any arbitrary permutation $\pi: i\longmapsto[j]$ on $\boldsymbol{\omega}$, the following technical condition*

$$\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\approx\Theta\left(\frac{m_k}{N}\right),\qquad\forall k=1,2,\ldots,N$$

*holds almost surely as $N\to\infty$, for all firms simultaneously.*

*Proof.* In the proof of Proposition S.1, we prove that

$$m_k\approx\frac{N}{k^{1/\beta_m}},\ \text{ a.s. }\ \forall k\in[N].$$

This result only requires $\beta_m>1,\beta_\omega>1$. Meanwhile, by (S.34), we only need to prove the following result:

$$N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\approx\Theta\left(\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right),\ \text{ a.s.}\qquad\forall k\in[N],$$

when $\frac{1}{\beta_m}+\frac{1}{\beta_\omega}=1$. In this boundary case, we invoke a stronger concentration inequality, namely, the Bennett-Hoeffding inequality (Bennett (1962) and Hoeffding (1963)). As before, we isolate the deterministic self-loop

$$N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}=N\omega_{\pi(k)}+X_k,\qquad X_k:=\sum_{i\neq k}N\omega_{\pi(i)}A_{ik}.$$

Then

$$N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}-\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]=X_k-\mathbb{E}[X_k].$$

**Step 1: Uniform lower bound on the mean.** Since

$$\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]=N\Big[(1-\omega_{\pi(k)})p_k+\omega_{\pi(k)}\Big],\qquad p_k:=\frac{Nk^{-1/\beta_m}-1}{N-1},$$

and $p_k$ is decreasing in $k$, we have

$$\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\geq Np_k\geq Np_N=N\frac{N^{\frac{1}{\beta_\omega}}-1}{N-1}. \tag{S.61}$$

Hence there exists $c_0>0$ such that for all sufficiently large $N$,

$$\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\geq c_0N^{\frac{1}{\beta_\omega}},\qquad \forall k=1,\dots,N,$$

with $\frac{1}{\beta_\omega}\in(0,1)$.

**Step 2: Lower-tail estimate.** By the lower-tail concentration bound,

$$\mathbb{P}\left(N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}<\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]-t\right)\leq\exp\left(-\frac{t^2}{2v_k}\right),$$

where

$$v_k=N^2p_k\sum_{i\neq k}\omega_{\pi(i)}^2\leq N^2p_k\sum_{i=1}^{N}\omega_{\pi(i)}^2\leq N^2p_k\frac{\zeta(\frac{2}{\beta_\omega},N)}{\left[\zeta(\frac{1}{\beta_\omega},N)\right]^2}.$$

We pick

$$t=\frac{1}{2}\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right].$$

It follows from (S.61) that

$$\mathbb{P}\left(N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}<\frac{1}{2}\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right)\leq\exp\left(-\frac{\left(\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right)^2}{8v_k}\right)\leq\exp\left(-\frac{p_k}{8\frac{\zeta(\frac{2}{\beta_\omega},N)}{\left[\zeta(\frac{1}{\beta_\omega},N)\right]^2}}\right)\leq\exp\left(-\frac{p_N}{8\frac{\zeta(\frac{2}{\beta_\omega},N)}{\left[\zeta(\frac{1}{\beta_\omega},N)\right]^2}}\right).$$

The above inequality holds uniformly for all firms $k\in[N]$. Since $\frac{1}{\beta_m}+\frac{1}{\beta_\omega}=1$, we estimate $p_N=$

$\Theta\left(N^{-\frac{1}{\beta_m}}\right) = \Theta\left(N^{\frac{1}{\beta_\omega}-1}\right)$. By definition, $\frac{1}{\beta_\omega} \in (0,1)$, therefore, by (S.33),

$$\frac{p_N}{\frac{\zeta(\frac{2}{\beta_\omega},N)}{\left[\zeta(\frac{1}{\beta_\omega},N)\right]^2}} = \begin{cases} \Theta\left(N^{\frac{1}{\beta_\omega}}\right), & \frac{1}{\beta_\omega} < \frac{1}{2}, \\ \Theta\left(N^{1/2}/\ln N\right), & \frac{1}{\beta_\omega} = \frac{1}{2}, \\ \Theta\left(N^{1-\frac{1}{\beta_\omega}}\right), & \frac{1}{\beta_\omega} > \frac{1}{2}, \end{cases}$$

which diverges to $+\infty$ in all cases. Hence there exists $c_1 > 0$ such that

$$\mathbb{P}\left(N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} < \frac{1}{2}\mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right]\right) \le \exp\left(-c_1 \frac{N^{\min\{\frac{1}{\beta_\omega}, 1-\frac{1}{\beta_\omega}\}}}{\ln N}\right), \qquad \forall k \in [N],$$

for all large $N$, uniformly in $k$. By the union bound,

$$\begin{aligned}\mathbb{P}\left(\left\{\exists k \in [N] : \ N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} < \frac{1}{2}\mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right]\right\}\right) &= \mathbb{P}\left(\bigcup_{k=1}^{N}\left\{ N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} < \frac{1}{2}\mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right]\right\}\right) \\ &\le N \exp\left(-c_1 \frac{N^{\min\{\frac{1}{\beta_\omega}, 1-\frac{1}{\beta_\omega}\}}}{\ln N}\right) \to 0.\end{aligned}$$

Using De Morgan's law for countable intersections, with high probability (w.h.p.), we have

$$\lim_{N\to\infty} \mathbb{P}\left(N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} \ge \frac{1}{2}\mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right]\right) = 1, \qquad \forall k \in [N] \tag{S.62}$$

for all firms simultaneously.

**Step 3: Upper-tail estimate via Bennett-Hoeffding's inequality.** As before, we consider the stochastic component of the weighted sum of firm sizes. Define

$$a_N := \max_{i\neq k} N\omega_{\pi(i)} \le N\omega_{\max} = \frac{N}{\zeta(\frac{1}{\beta_\omega}, N)} = \Theta\left(N^{\frac{1}{\beta_\omega}}\right),$$

and

$$v_k := \mathrm{Var}(X_k) = N^2 p_k(1-p_k) \sum_{i\neq k} \omega_{\pi(i)}^2 \le N^2 p_k \frac{\zeta(\frac{2}{\beta_\omega}, N)}{\left[\zeta(\frac{1}{\beta_\omega}, N)\right]^2}.$$

We apply Bennett-Hoeffding's inequality:

$$\mathbb{P}(X_k - \mathbb{E}[X_k] \ge t) \le \exp\left(-\frac{v_k}{a_N^2}\, h\left(\frac{a_N t}{v_k}\right)\right), \qquad h(u) := (1+u)\log(1+u) - u.$$

We choose

$$t = (C-1)\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right],$$

where $C > 1$ will be chosen later. Then

$$\mathbb{P}\left(N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)} \geq C\,\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right) = \mathbb{P}\left(X_k - \mathbb{E}[X_k] \geq (C-1)\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right) \leq \exp\left(-B_{k,N}\right), \tag{S.63}$$

where we define

$$B_{k,N} := \frac{v_k}{a_N^2}\, h\left(\frac{a_N(C-1)\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]}{v_k}\right).$$

We set

$$u_{k,N} := \frac{a_N(C-1)\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]}{v_k}.$$

Using $\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right] \geq Np_k$ in (S.61) and $v_k \leq N^2 p_k \frac{\zeta(\frac{2}{\beta_\omega},N)}{\left[\zeta(\frac{1}{\beta_\omega},N)\right]^2}$, we obtain

$$u_{k,N} \geq (C-1)\frac{a_N}{N\frac{\zeta(\frac{2}{\beta_\omega},N)}{\left[\zeta(\frac{1}{\beta_\omega},N)\right]^2}} = \begin{cases}\Theta\left(N^{\frac{1}{\beta_\omega}}\right), & \frac{1}{\beta_\omega} < \frac{1}{2},\\ \Theta\left(N^{1/2}/\ln N\right), & \frac{1}{\beta_\omega} = \frac{1}{2},\\ \Theta\left(N^{1-\frac{1}{\beta_\omega}}\right), & \frac{1}{\beta_\omega} > \frac{1}{2}.\end{cases} \tag{S.64}$$

Since $a_N = \Theta\left(N^{\frac{1}{\beta_\omega}}\right)$. Hence $u_{k,N} \to \infty$ in all possible cases. In particular, for all large $N$, $u_{k,N} \geq e^2$ uniformly in $k \in [N]$. For $u \geq e^2$, it is easy to derive that

$$h(u) \geq \frac{1}{2}u\log u.$$

Therefore, for all large $N$,

$$B_{k,N} \geq \frac{v_k}{a_N^2}\cdot\frac{1}{2}u_{k,N}\log u_{k,N} = \frac{(C-1)\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]}{2a_N}\log u_{k,N}.$$

We have shown that $\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right] \geq c_0 N^{\frac{1}{\beta_\omega}}$ and $a_N \leq C_0 N^{\frac{1}{\beta_\omega}}$ for some $C_0 > 0$. Hence there exists $c_2 > 0$ such that

$$\frac{\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]}{a_N} \geq c_2 \qquad \forall k \in [N],$$

for all large $N$. Hence, by (S.64),

$$B_{k,N} \geq \frac{c_2(C-1)}{2} \log u_{k,N} \geq \frac{c_2(C-1)}{2} \log \left( (C-1) \frac{a_N}{N \frac{\zeta(\frac{2}{\beta_\omega}, N)}{\left[\zeta(\frac{1}{\beta_\omega}, N)\right]^2}} \right).$$

Since any positive polynomials in $N$ dominates $\log N$ , there exists $c_3 > 0$ such that

$$\log \left( (C-1) \frac{a_N}{N \frac{\zeta(\frac{2}{\beta_\omega}, N)}{\left[\zeta(\frac{1}{\beta_\omega}, N)\right]^2}} \right) \geq c_3 \log N$$

for all large $N$. Thus

$$B_{k,N} \geq c_4 (C-1) \log N$$

for some constant $c_4 > 0$, uniformly in $k \in [N]$. Using (S.63),

$$\mathbb{P}\left( N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \geq C \, \mathbb{E}\left[ N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \right] \right) \leq N^{-c_4(C-1)}$$

uniformly in $k \in [N]$, for all large $N$. Next, we pick $C > 1$ large enough so that $c_4(C-1) > 2$, it follows by the union bound

$$\begin{aligned}
\mathbb{P}\left( \left\{ \exists k : \; N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \geq C \, \mathbb{E}\left[ N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \right] \right\} \right) &= \mathbb{P}\left( \bigcup_{k=1}^{N} \left\{ N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \geq C \, \mathbb{E}\left[ N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \right] \right\} \right) \\
&\leq \sum_{k=1}^{N} P\left( \left\{ N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \geq C \, \mathbb{E}\left[ N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \right] \right\} \right) \\
&\leq N^{1-(c_4(C-1))} \to 0.
\end{aligned}$$

By countable intersections,

$$\lim_{N \to \infty} \mathbb{P}\left( \bigcap_{k=1}^{N} \left\{ N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} < C \, \mathbb{E}\left[ N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \right] \right\} \right) = 1 \tag{S.65}$$

simultaneously for all firms $k \in [N]$.

Consolidating (S.62) and (S.65), we obtain

$$\lim_{N\to\infty}\mathbb{P}\left(\bigcap_{k=1}^{N}\left\{\frac{1}{2}\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\leq N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\leq C\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right\}\right)=1.$$

This proves

$$N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}=\Theta\left(\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right)\qquad\text{w.h.p.}$$

for all firms $k\in[N]$ simultaneously. Dividing by $N$ gives

$$\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}=\Theta\left(\mathbb{E}\left[\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right)=\Theta\left(\frac{m_k}{N}\right)\qquad\text{w.h.p.}$$

for all firms $k\in[N]$ simultaneously.

**Step 4: Almost-sure convergence.** Since $C$ can be chosen large enough so that $c_4(C-1)>2$, we have the summability condition

$$\sum_{N=1}^{\infty}\sum_{k=1}^{N}P\left(\left\{N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\geq C\,\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right\}\right)\leq\sum_{N=1}^{\infty}N^{-[(c_4(C-1))-1]}=\zeta\Big((c_4(C-1))-1\Big)<\infty.$$

by the convergence of the Riemann zeta function. The lower-tail error probabilities are even smaller, so they are summable as well by Lemma S.4. Therefore, using the first Borel–Cantelli lemma S.3, with probability one, for all sufficiently large $N$,

$$\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}=\Theta\left(\frac{m_k}{N}\right)\qquad\text{a.s.}$$

holds almost surely eventually, and for all firms $k\in[N]$ simultaneously. □

**Corollary S.2.** *Suppose $\beta_\omega>1$ and $\beta_m\geq1$, for any arbitrary permutation $\pi:i\longmapsto[j]$ on $\boldsymbol{\omega}$, there exists a subset of firms with cardinality $N^\Lambda$ with*

$$0<\Lambda<\min\left\{\beta_m\left(1-\frac{1}{\beta_\omega}\right),1\right\}.$$

*In this subset of the economy,*

$$\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\approx\frac{m_k}{N}\qquad a.s.$$

*holds almost surely eventually, and for all firms $k$ in this subset simultaneously.*

*Proof.* If $\beta_m\left(1-\frac{1}{\beta_\omega}\right)>1$, Proposition S.1 shows that this technical condition holds a.s. for all firms

simultaneously. So we focus on the case

$$0 < \Lambda < \beta_m\Big(1 - \frac{1}{\beta_\omega}\Big) \le 1.$$

under the random graph structure

$$\mathbb{P}(A_{ik} = 1) = p_k = \frac{Nk^{-1/\beta_m} - 1}{N - 1} \quad (i \neq k), \qquad A_{kk} = 1.$$

We pick a subset $K_N$ of firms with highest attentions, i.e., the first $N^\Lambda$ columns in the stochastic adjacency matrix,

$$K_N := \{1, 2, \ldots, \lfloor N^\Lambda \rfloor\}.$$

It follows that

$$1 - \frac{\Lambda}{\beta_m} > \frac{1}{\beta_\omega}.$$

Hence we can choose $\eta \in (1/2, 1)$ such that

$$\eta\Big(1 - \frac{\Lambda}{\beta_m}\Big) > \frac{1}{\beta_\omega}.$$

Note that when $\beta_m \ge 1$, for all $k \in K_N$, (S.25) and (S.55) hold, so

$$m_j \approx \frac{N}{k^{1/\beta_m}}, \quad \text{a.s.} \quad \forall k \in K_N.$$

Define

$$t_k := \left(\mathbb{E}\left[N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)}\right]\right)^\eta, \qquad k \in K_N.$$

For $k \in K_N$, we have $k \le N^\Lambda$, and therefore, $k^{-1/\beta_m} \ge N^{-\Lambda/\beta_m}$. Hence

$$p_k = \frac{Nk^{-1/\beta_m} - 1}{N - 1} \ge \frac{N^{1-\Lambda/\beta_m} - 1}{N - 1}.$$

Since $\Lambda < \beta_m$, the exponent $1 - \Lambda/\beta_m$ is positive, and so for all sufficiently large $N$,

$$p_k \ge c\, N^{-\Lambda/\beta_m}$$

for some constant $c > 0$, uniformly in $k \in K_N$. It follows that

$$\mathbb{E}\left[N \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)}\right] = N\Big[(1 - \omega_{\pi(k)})p_k + \omega_{\pi(k)}\Big] \ge Np_k \ge c\, N^{1-\Lambda/\beta_m}, \qquad \forall k \in K_N$$

for all large $N$. In particular,

$$\inf_{k\in K_N} \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] \to \infty.$$

Since $\eta < 1$, it follows that

$$t_k = o\left(\mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right]\right)$$

uniformly in $k \in K_N$. The rest of our proof strategy parallels Proposition S.1, so we omit details to avoid redundancy. In particular, for the lower-tail probability, there exists a deterministic sequence $\Delta_N \to \infty$ such that

$$\sup_{k\in K_N} \mathbb{P}\left(N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} < \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] - t_k\right) \le e^{-\Delta_N}, \qquad \frac{\Delta_N}{\ln N} \to \infty.$$

By the union bound,

$$\begin{aligned}
\mathbb{P}\left(\left\{\exists k \in K_N : N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} < \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] - t_k\right\}\right) &= \mathbb{P}\left(\bigcup_{k=1}^{N^\Lambda}\left\{N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} < \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] - t_k\right\}\right) \\
&\le \sum_{k=1}^{N^\Lambda} P\left(\left\{N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} < \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] - t_k\right\}\right) \\
&\le |K_N| e^{-\Delta_N} = N^\Lambda e^{-\Delta_N} \to 0
\end{aligned}$$

for all firms $k \in K_N$ simultaneously. Similarly, for the upper-tail probability, there exists a deterministic sequence $\widetilde{\Delta}_N \to \infty$ with[31]

$$\sup_{k\in K_N} \mathbb{P}\left(N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} > \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] + t_k\right) \le e^{-\widetilde{\Delta}_N}, \qquad \frac{\widetilde{\Delta}_N}{\ln N} \to \infty.$$

Hence, by the union bound,

$$\begin{aligned}
\mathbb{P}\left(\left\{\exists k \in K_N : N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} > \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] + t_k\right\}\right) &\le \sum_{k=1}^{N^\Lambda} P\left(\left\{N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} > \mathbb{E}\left[N \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] + t_k\right\}\right) \\
&\le N^\Lambda e^{-\widetilde{\Delta}_N} \to 0.
\end{aligned}$$

[31] The derivations of these two inequalities are available upon request.

Combining these two bounds yields

$$\mathbb{P}\left(\bigcup_{k=1}^{N^{\Lambda}}\left\{\left|N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}-\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right|<o\left(\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right)\right\}\right)$$
$$\leq N^{\Lambda}e^{-\Delta_N}+N^{\Lambda}e^{-\widetilde{\Delta}_N},\qquad \frac{\widetilde{\Delta}_N}{\ln N}\to\infty;\ \ \frac{\Delta_N}{\ln N}\to\infty.$$

Note that we can pick $\varepsilon>0$ such that

$$\Delta_N>(\Lambda+1+\varepsilon)\ln N;\qquad \widetilde{\Delta}_N>(\Lambda+1+\varepsilon)\ln N.$$

Therefore,

$$\sum_{N=1}^{\infty}\mathbb{P}\left(\bigcup_{k=1}^{N^{\Lambda}}\left\{\left|N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}-\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right|<o\left(\mathbb{E}\left[N\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\right)\right\}\right)\leq 2\sum_{N=1}^{\infty}N^{-(1+\varepsilon)}<\infty.$$

Then by the first Borel-Cantelli lemma S.3,

$$\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\approx\frac{m_k}{N}\qquad\text{a.s.}$$

holds almost surely eventually, and for all firms $k\in K_N$ simultaneously. The proof is now complete.

□

**Corollary S.3.** *Suppose $\beta_m=1\leq\beta_\omega$. Given the random graph structure (5.5) and the numerical distribution of firm sizes $\boldsymbol{\omega}$ following (5.8), for any arbitrary permutation $\pi: i\longmapsto[j]$ on $\boldsymbol{\omega}$,*

$$d_i=\Theta\left(\zeta\left(\frac{1}{\beta_m},N\right)\right)=\Theta\left(\ln N\right);\qquad \textit{w.h.p.}$$

*as $N\to\infty$, for any subset of the economy with cardinality of $\left|\widetilde{K}_N\right|=N^{1/24}$ firms [32] simultaneously.*

*Proof.* Using the result in the proof of Lemma 5.2, the expected out-degree for a given firm $i\in[N]$ is given by

$$v_i(N)\equiv\mathbb{E}\left[d_i\right]=\frac{N}{N-1}\left(\zeta(\frac{1}{\beta_m},N)-\frac{1}{i^{\frac{1}{\beta_m}}}\right)\approx\ln N \tag{S.66}$$

when $N\to\infty$, since $\beta_m=1$. Note that

$$\inf_{i\in[N]}v_i(N)=\frac{N}{N-1}\left(\zeta(\frac{1}{\beta_m},N)-1\right)\approx\ln N.$$

[32] Here the power $1/24$ can be improved to any constant smaller than $1/2$ by more accurate estimations, we just adopt $1/24$ to simplify the calculation.

Therefore, for any $h < 1$, there exists $N_0 \in \mathbb{N}$ such that

$$v_i(N) \geq h \ln N, \quad \forall i \in [N], \quad \forall N \geq N_0.$$

Invoking the concentration inequality in Lemma S.1 again,

$$\mathbb{P}\left(v_i(N) - t \leq d_i \leq v_i(N) + t\right) \geq 1 - \exp\left(-\frac{t^2}{2v_i(N)}\right) - \exp\left(-\frac{t^2}{2v_i(N) + 2t/3}\right),$$

where we use the trick $\mathrm{var}\left(d_i\right) \leq \mathbb{E}\left[d_i\right]$ again. Let $t = \frac{1}{2}v_i(N)$ in the above inequality, we obtain

$$\begin{aligned}&\mathbb{P}\left(v_i(N) - \frac{1}{2}v_i(N) \leq d_i \leq v_i(N) + \frac{1}{2}v_i(N)\right)\\ &\geq 1 - \exp\left(-\frac{1}{8}v_i(N)\right) - \exp\left(-\frac{3}{28}v_i(N)\right),\end{aligned}$$

when $N$ is large enough. Next, we note that

$$\exp\left(-\frac{1}{8}v_i(N)\right) + \exp\left(-\frac{3}{28}v_i(N)\right) \leq 2\exp\left(-\frac{3}{28}v_i(N)\right) \leq 2\exp\left(-\frac{3}{28}h\ln N\right), \qquad \forall i \in [N]$$

uniformly. We pick $h > \frac{7}{9}$ so that

$$\frac{3}{28}h > \frac{1}{12}.$$

It follows that

$$\exp\left(-\frac{1}{8}v_i(N)\right) + \exp\left(-\frac{3}{28}v_i(N)\right) \leq 2\exp\left(-\frac{3}{28}h\ln N\right) \leq \exp\left(-\frac{1}{12}\ln N\right), \qquad \forall i \in [N]$$

uniformly, provided that

$$N \geq 2^{\left(\frac{1}{\frac{3}{28}h - \frac{1}{12}}\right)}.$$

Now, pick any subset $\widetilde{K}_N$ of firms with cardinality $\left|\widetilde{K}_N\right| = N^{1/24}$ firms, the probability that $\frac{1}{2}v_i(N) \leq d_i \leq \frac{3}{2}v_i(N)$ hold simultaneously is at least

$$1 - \left|\widetilde{K}_N\right| N^{-\frac{1}{12}} = 1 - N^{\frac{1}{24}}N^{-\frac{1}{12}} = 1 - N^{-\frac{1}{24}} \to 1$$

when $N$ tends to infinity. Therefore,

$$d_i = \Theta\left(\ln N\right), \qquad \text{w.h.p.} \qquad \forall i \in \widetilde{K}_N,$$

simultaneously. The proof is now complete. □

**Corollary S.4.** *Suppose $\beta_\omega = 1$ and $\beta_m > 0$. Given the random graph structure (5.5) and the numerical*

*distribution of firm sizes $\boldsymbol{\omega}$ following (5.8), for any arbitrary permutation $\pi$ on $\boldsymbol{\omega}$,*

$$\sum_{j \notin \mathcal{D}_k} \omega_{\pi(j)} = \Theta(1), \qquad a.s.$$

*for all firms $k \in [N]$ simultaneously.*

*Proof.* We follow closely the proof strategy of Proposition S.1. By definition,

$$\sum_{j \notin \mathcal{D}_k} \omega_{\pi(j)} \leq 1$$

is a sure event. Therefore, we only need to show that the above sum has a uniform lower bound for all firms simultaneously. That is, there exists $t \in (0,1)$ such that

$$\sum_{j \notin \mathcal{D}_k} \omega_{\pi(j)} \geq t, \qquad \text{a.s. } \forall k \in [N]$$

uniformly when $N \to \infty$. This is equivalent to the condition that

$$1 - \lim_{N\to\infty} \sum_{j=1}^{N} \omega_{\pi(j)} A_{kj} \geq t \Leftrightarrow \varrho_k(N) \leq 1 - t \qquad \text{a.s. } \forall k \in [N].$$

By (S.52),

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}[\varrho_k(N)] &\leq \lim_{N\to\infty} \left( \frac{N\zeta\left(\frac{1}{\beta_m} + \frac{1}{\beta_\omega}, N\right)}{(N-1)\,\zeta\left(\frac{1}{\beta_\omega}, N\right)} - \frac{1}{(N-1)} + \omega_{\pi(k)} \frac{N}{N-1} \left(1 - k^{-\frac{1}{\beta_m}}\right) \right), \\
&= \lim_{N\to\infty} \left( \frac{N\zeta\left(\frac{1}{\beta_m} + 1, N\right)}{(N-1)\ln N} - \frac{1}{(N-1)} + \omega_{\pi(k)} \frac{N}{N-1} \left(1 - k^{-\frac{1}{\beta_m}}\right) \right) = 0, \qquad \forall \pi, \ \ \forall k \in [N].
\end{aligned}$$

Note that the above limit is uniform (consider the largest limit with $k = 1$). In Proposition S.1, we derive the probability upper bound:

$$\begin{aligned}
\mathbb{P}(\varrho_k(N) \leq \mathbb{E}[\varrho_k(N)] + 1 - t) &\geq 1 - \exp\left( - \frac{2(1-t)^2 \left[\zeta\left(\frac{1}{\beta_\omega}, N\right)\right]^2}{2\left(1 + \zeta\left(\frac{2}{\beta_\omega}, N\right)\right) + \frac{2}{3}\zeta\left(\frac{1}{\beta_\omega}, N\right)(1-t)} \right) \\
&= 1 - \exp\left( - \frac{2(1-t)^2 \left[\ln N\right]^2}{2\left(1 + \frac{\pi^2}{6}\right) + \frac{2}{3}\ln N\,(1-t)} \right)
\end{aligned}$$

$$= 1 - \exp\left(-\Theta\left(3(1-t)\ln N\right)\right), \qquad \forall k \in [N]$$

for all firms uniformly. Next, we choose a small $\varepsilon > 0$ and $t > 0$ such that

$$0 < t < \frac{(1-\varepsilon)}{3}.$$

Then by the union bound in the limit,

$$\lim_{N\to\infty} \mathbb{P}\left(\bigcap_{k=1}^{N} \left\{ \varrho_k(N) = \sum_{j=1}^{N} \omega_{\pi(j)} A_{kj} < 1 - t \right\}\right) \geq 1 - \Theta\left(N^{-(1+\varepsilon)}\right) = 1.$$

Since $\varepsilon > 0$, the summability condition is satisfied so that we also have almost-sure convergence by Borel-Cantelli lemma. The proof is now complete. □

### S.6.4 Volatility Estimates on Random Graphs

In this section, we present a set of results on volatility estimates in random graphs. These results will be used in the proof of Proposition 5.3.

**Some notation and conventions**

Let $0 < \frac{1}{\beta_m} < 1$ and $0 < \frac{1}{\beta_\omega} < 1$. For $1 \leq i, j \leq N$, let $\{A_{ij}\}_{i,j\in[N]}$ be a sequence of independent Bernoulli random variables with

$$\mathbb{P}(A_{ij} = 1) = p_j = j^{-1/\beta_m}.$$

We define size-related sums of Bernoulli random variables as

$$a_i = i^{-1/\beta_\omega}, \qquad \mathcal{Y}_j = \sum_{i=1}^{N} a_i A_{ij}, \qquad \mathcal{S}_N = \sum_{j=1}^{N} \mathcal{Y}_j^2.$$

We also define the power-sums of sizes as

$$\mathcal{A}_N = \sum_{i=1}^{N} a_i, \qquad \mathcal{B}_N = \sum_{i=1}^{N} a_i^2, \qquad \mathcal{C}_N = \sum_{i=1}^{N} a_i^3, \qquad \mathcal{D}_N = \sum_{i=1}^{N} a_i^4,$$

and the power-sums of probabilities as

$$\mathcal{P}_1 = \sum_{j=1}^{N} p_j, \qquad \mathcal{P}_2 = \sum_{j=1}^{N} p_j^2, \qquad \mathcal{P}_3 = \sum_{j=1}^{N} p_j^3.$$

These variables will be used extensively in our computation of higher-order moments, leading to the following lemma.

**Lemma S.5** (Central Moment Formulas of $\mathcal{Y}_j$)**.** *Fix a $j \in [N]$ and for simplicity we write $p = p_j$. Then the*

*central moments of $\mathcal{Y}_j$ are given by*[33]

$$\boldsymbol{\mu}_j = \mathbb{E}[\mathcal{Y}_j], \qquad \boldsymbol{\sigma}_j^2 = \mathrm{Var}(\mathcal{Y}_j), \qquad \boldsymbol{\tau}_j = \mathbb{E}[(\mathcal{Y}_j - \boldsymbol{\mu}_j)^3], \qquad \boldsymbol{\kappa}_j = \mathbb{E}[(\mathcal{Y}_j - \boldsymbol{\mu}_j)^4].$$

*In particular,*

$$\boldsymbol{\mu}_j = p\mathcal{A}_N; \qquad \boldsymbol{\sigma}_j^2 = p(1-p)\mathcal{B}_N; \qquad \boldsymbol{\tau}_j = p(1-p)(1-2p)\mathcal{C}_N \tag{S.67}$$

*and*

$$\boldsymbol{\kappa}_j = p(1-p)\Big(1 - 3p(1-p)\Big)\mathcal{D}_N + 3p^2(1-p)^2(\mathcal{B}_N^2 - \mathcal{D}_N). \tag{S.68}$$

*Proof.* Note that for any Bernoulli random variable, by definition, $A_{ij}^k = A_{ij}$ for any positive integer $k$. Define

$$\mathcal{U}_{i,j} = a_i(A_{i,j} - p).$$

Then

$$\mathcal{Y}_j - \boldsymbol{\mu}_j = \sum_{i=1}^N \mathcal{U}_{i,j}.$$

Since $A_{i,j} \sim \mathrm{Bernoulli}(p)$, we have

$$\mathbb{E}[A_{i,j} - p] = 0, \qquad \mathbb{E}[(A_{i,j} - p)^2] = p(1-p),$$

and

$$\mathbb{E}[(A_{i,j} - p)^3] = p(1-p)(1-2p), \qquad \mathbb{E}[(A_{i,j} - p)^4] = p(1-p)\big(1 - 3p(1-p)\big).$$

Therefore

$$\mathbb{E}[\mathcal{U}_{i,j}] = 0, \qquad \mathbb{E}[\mathcal{U}_{i,j}^2] = a_i^2 p(1-p),$$

and

$$\mathbb{E}[\mathcal{U}_{i,j}^3] = a_i^3 p(1-p)(1-2p), \qquad \mathbb{E}[\mathcal{U}_{i,j}^4] = a_i^4 p(1-p)\big(1 - 3p(1-p)\big).$$

For the mean, it follows that

$$\boldsymbol{\mu}_j = \mathbb{E}[\mathcal{Y}_j] = \sum_{i=1}^N a_i \mathbb{E}[A_{i,j}] = p\sum_{i=1}^N a_i = p\mathcal{A}_N.$$

For the variance, since the variables $\mathcal{U}_{i,j}$ are independent and centered,

$$\boldsymbol{\sigma}_j^2 = \mathrm{Var}\Big(\sum_{i=1}^N \mathcal{U}_{i,j}\Big) = \sum_{i=1}^N \mathbb{E}[\mathcal{U}_{i,j}^2] = p(1-p)\sum_{i=1}^N a_i^2 = p(1-p)\mathcal{B}_N.$$

For the third centered moment, we use the binomial expansion to expand $\Big(\sum_{i=1}^N \mathcal{U}_{i,j}\Big)^3$. Since $\mathbb{E}\Big[\mathcal{U}_{i,j}^2 \mathcal{U}_{k,j}\Big] =$

[33] We use bold symbols to avoid notation inconsistency.

0 for $i \neq k$, only the terms $\mathcal{U}_{i,j}^3$ survive. That is, every mixed term has expectation zero because $\mathcal{U}_{i,j}$ are independent and centered. Hence

$$\tau_j = \sum_{i=1}^{N} \mathbb{E}[\mathcal{U}_{i,j}^3] = p(1-p)(1-2p) \sum_{i=1}^{N} a_i^3 = p(1-p)(1-2p)\mathcal{C}_N.$$

For the fourth centered moment, for independent centered variables we observe that if we expand $\left(\sum \mathcal{U}_{i,j}\right)^4$, the only nonzero terms are $\mathcal{U}_{i,j}^4$ (coefficient 1) and $\mathcal{U}_{i,j}^2 \mathcal{U}_{k,j}^2$ with $i \neq k$ (coefficient $\binom{4}{2} = 6$ ). Hence

$$\mathbb{E}\Big[\Big(\sum_{i=1}^{N} \mathcal{U}_{i,j}\Big)^4\Big] = \sum_{i=1}^{N} \mathbb{E}[\mathcal{U}_{i,j}^4] + 6 \sum_{1 \leq i < k \leq N} \mathbb{E}[\mathcal{U}_{i,j}^2]\mathbb{E}[\mathcal{U}_{k,j}^2].$$

Thus

$$\kappa_j = \sum_{i=1}^{N} a_i^4 p(1-p)\big(1 - 3p(1-p)\big) + 6 \sum_{1 \leq i < k \leq N} a_i^2 a_k^2 p^2 (1-p)^2.$$

Since

$$6 \sum_{1 \leq i < k \leq N} a_i^2 a_k^2 = 3\left(\left(\sum_{i=1}^{N} a_i^2\right)^2 - \sum_{i=1}^{N} a_i^4\right) = 3(\mathcal{B}_N^2 - \mathcal{D}_N),$$

we conclude that

$$\kappa_j = p(1-p)\big(1 - 3p(1-p)\big)\mathcal{D}_N + 3p^2(1-p)^2(\mathcal{B}_N^2 - \mathcal{D}_N).$$

This proves the lemma. □

**Proposition S.2** (Convergence of $\mathcal{S}_N$ in probability)**.** *With the notation defined above, suppose* $0 < \frac{1}{\beta_m} < 1$ *and* $0 < \frac{1}{\beta_\omega} < 1$*, a weak law of large numbers (WLLN) applies to* $\mathcal{S}_N$*:*

$$\frac{\mathcal{S}_N}{\mathbb{E}[\mathcal{S}_N]} \xrightarrow{\mathbb{P}} 1 \qquad (N \to \infty).$$

*Moreover,*

$$\mathbb{E}[\mathcal{S}_N] = \mathcal{A}_N^2 \mathcal{P}_2 + \mathcal{B}_N(\mathcal{P}_1 - \mathcal{P}_2) = \Theta\Big(\mathcal{A}_N^2 \mathcal{P}_2 + \mathcal{B}_N \mathcal{P}_1\Big),$$

*and*

$$\mathrm{Var}(\mathcal{S}_N) = O\Big(\mathcal{A}_N^2 \mathcal{B}_N \mathcal{P}_3 + \mathcal{A}_N \mathcal{C}_N \mathcal{P}_2 + \mathcal{D}_N \mathcal{P}_1 + \mathcal{B}_N^2 \mathcal{P}_2\Big).$$

*Proof.* We first compute the expectation exactly. Since $\mathcal{S}_N = \sum_{j=1}^{N} \mathcal{Y}_j^2$, we have

$$\mathbb{E}[\mathcal{S}_N] = \sum_{j=1}^{N} \mathbb{E}[\mathcal{Y}_j^2] = \sum_{j=1}^{N} (\mu_j^2 + \sigma_j^2).$$

Using Lemma S.5, we obtain the following.

$$\mathbb{E}[\mathcal{S}_N] = \sum_{j=1}^{N} \left( \mathcal{A}_N^2 p_j^2 + \mathcal{B}_N p_j (1 - p_j) \right) = \mathcal{A}_N^2 \mathcal{P}_2 + \mathcal{B}_N (\mathcal{P}_1 - \mathcal{P}_2).$$

Next, we show

$$\mathcal{P}_1 - \mathcal{P}_2 = \Theta(\mathcal{P}_1).$$

Indeed, $p_1 = 1$, while for every $j \geq 2$,

$$0 < p_j = j^{-1/\beta_m} \leq 2^{-1/\beta_m} < 1.$$

Hence, there exists a constant $\delta_0 \in (0, 1)$ such that

$$1 - p_j \in [\delta_0, 1] \qquad \text{for all } j \geq 2.$$

Therefore,

$$\delta_0 p_j \leq p_j (1 - p_j) \leq p_j \qquad \text{for all } j \geq 2.$$

Summing over $j = 2, \ldots, N$ gives

$$\delta_0 \sum_{j=2}^{N} p_j \leq \sum_{j=2}^{N} p_j (1 - p_j) \leq \sum_{j=2}^{N} p_j.$$

Since

$$\mathcal{P}_1 - \mathcal{P}_2 = \sum_{j=1}^{N} p_j (1 - p_j) = \sum_{j=2}^{N} p_j (1 - p_j), \qquad \sum_{j=2}^{N} p_j = \mathcal{P}_1 - 1 = \Theta(\mathcal{P}_1),$$

and $\mathcal{P}_1 \to \infty$ when $0 < 1/\beta_m < 1$, it follows that

$$\mathcal{P}_1 - \mathcal{P}_2 = \Theta\big(\mathcal{P}_1\big).$$

Consequently,

$$\mathbb{E}[\mathcal{S}_N] = \Theta\left( \mathcal{A}_N^2 \mathcal{P}_2 + \mathcal{B}_N \mathcal{P}_1 \right).$$

We now estimate the variance. For any random variable $Z$, $\mathrm{Var}(Z^2) = \mathbb{E}[Z^4] - \big(\mathbb{E}[Z^2]\big)^2$. Applying this identity with $Z = \mathcal{Y}_j$, and using the identity

$$\mathcal{Y}_j = \Big( \mathcal{Y}_j - \boldsymbol{\mu}_j \Big) + \boldsymbol{\mu}_j,$$

we obtain

$$\mathbb{E}[\mathcal{Y}_j^4] = \boldsymbol{\kappa}_j + 4\boldsymbol{\mu}_j \boldsymbol{\tau}_j + 6\boldsymbol{\mu}_j^2 \boldsymbol{\sigma}_j^2 + \boldsymbol{\mu}_j^4, \qquad \mathbb{E}[\mathcal{Y}_j^2] = \boldsymbol{\sigma}_j^2 + \boldsymbol{\mu}_j^2.$$

Hence,

$$\mathrm{Var}(\mathcal{Y}_j^2) = \boldsymbol{\kappa}_j - \boldsymbol{\sigma}_j^4 + 4\boldsymbol{\mu}_j\boldsymbol{\tau}_j + 4\boldsymbol{\mu}_j^2\boldsymbol{\sigma}_j^2.$$

Substituting the formulas from Lemma S.5 gives

$$4\boldsymbol{\mu}_j^2\boldsymbol{\sigma}_j^2 = 4\mathcal{A}_N^2\mathcal{B}_N p_j^3(1-p_j), \qquad 4\boldsymbol{\mu}_j\boldsymbol{\tau}_j = 4\mathcal{A}_N\mathcal{C}_N p_j^2(1-p_j)(1-2p_j),$$

and

$$\boldsymbol{\kappa}_j - \boldsymbol{\sigma}_j^4 = p_j(1-p_j)\big(1-3p_j(1-p_j)\big)\mathcal{D}_N + (2\mathcal{B}_N^2 - 3\mathcal{D}_N)p_j^2(1-p_j)^2.$$

Combining these two equations leads to

$$\begin{aligned}\mathrm{Var}(\mathcal{Y}_j^2) &= 4\mathcal{A}_N^2\mathcal{B}_N p_j^3(1-p_j) + 4\mathcal{A}_N\mathcal{C}_N p_j^2(1-p_j)(1-2p_j) \\ &\quad + p_j(1-p_j)\Big(1-3p_j(1-p_j)\Big)\mathcal{D}_N + \Big(2\mathcal{B}_N^2 - 3\mathcal{D}_N\Big)p_j^2(1-p_j)^2.\end{aligned}$$

Now we note that the collections $\{A_{i,j}\}_{i=1}^N$ and $\{A_{i,k}\}_{i=1}^N$ are independent whenever $j \neq k$. Hence $\mathcal{Y}_j$ and $\mathcal{Y}_k$ are independent whenever $j \neq k$, and therefore $\mathcal{Y}_j^2$ and $\mathcal{Y}_k^2$ are independent. It follows that

$$\mathrm{Var}(\mathcal{S}_N) = \sum_{j=1}^N \mathrm{Var}(\mathcal{Y}_j^2).$$

Next, we observe that

$$0 \le p_j \le 1, \qquad 0 \le 1-p_j \le 1, \qquad |1-2p_j| \le 1, \qquad 0 \le 1-3p_j(1-p_j) \le 1.$$

It follows that

$$\mathrm{Var}(\mathcal{Y}_j^2) \le 4\mathcal{A}_N^2\mathcal{B}_N p_j^3 + 4\mathcal{A}_N\mathcal{C}_N p_j^2 + \mathcal{D}_N p_j + 2\mathcal{B}_N^2 p_j^2.$$

Summing over $j$ gives

$$\mathrm{Var}(\mathcal{S}_N) \le 4\mathcal{A}_N^2\mathcal{B}_N\mathcal{P}_3 + 4\mathcal{A}_N\mathcal{C}_N\mathcal{P}_2 + \mathcal{D}_N\mathcal{P}_1 + 2\mathcal{B}_N^2\mathcal{P}_2.$$

Therefore,

$$\mathrm{Var}(\mathcal{S}_N) = O\Big(\mathcal{A}_N^2\mathcal{B}_N\mathcal{P}_3 + \mathcal{A}_N\mathcal{C}_N\mathcal{P}_2 + \mathcal{D}_N\mathcal{P}_1 + \mathcal{B}_N^2\mathcal{P}_2\Big). \tag{S.69}$$

We now show that

$$\mathrm{Var}(\mathcal{S}_N) = o\Big(\big(\mathbb{E}[\mathcal{S}_N]\big)^2\Big).$$

Since $\mathbb{E}[\mathcal{S}_N] = \Theta\left(\mathcal{A}_N^2\mathcal{P}_2 + \mathcal{B}_N\mathcal{P}_1\right)$, it suffices to show

$$\mathcal{A}_N^2\mathcal{B}_N\mathcal{P}_3 + \mathcal{A}_N\mathcal{C}_N\mathcal{P}_2 + \mathcal{D}_N\mathcal{P}_1 + \mathcal{B}_N^2\mathcal{P}_2 = o\left(\left(\mathcal{A}_N^2\mathcal{P}_2 + \mathcal{B}_N\mathcal{P}_1\right)^2\right).$$

To simplify the exposition, we define

$$\mathcal{L}_N := \mathcal{A}_N^2\mathcal{P}_2 + \mathcal{B}_N\mathcal{P}_1. \tag{S.70}$$

By definition,

$$\mathcal{L}_N^2 \geq \mathcal{A}_N^4\mathcal{P}_2^2, \qquad \mathcal{L}_N^2 \geq \mathcal{B}_N^2\mathcal{P}_1^2.$$

By Riemann zeta asymptotics,

$$\mathcal{A}_N \approx \frac{N^{1-\frac{1}{\beta_\omega}}}{1-\frac{1}{\beta_\omega}} = \Theta(N^{1-\frac{1}{\beta_\omega}}).$$

Similarly, we derive that

$$\mathcal{B}_N = \zeta\left(\frac{2}{\beta_\omega}, N\right) = \begin{cases} \Theta\left(N^{1-\frac{2}{\beta_\omega}}\right), & \beta_\omega > 2, \\ \Theta\left(\ln N\right), & \beta_\omega = 2, \\ \Theta(1), & 1 < \beta_\omega < 2, \end{cases} \qquad \mathcal{C}_N = \zeta\left(\frac{3}{\beta_\omega}, N\right) = \begin{cases} \Theta(N^{1-\frac{3}{\beta_\omega}}), & \beta_\omega > 3, \\ \Theta(\ln N), & \beta_\omega = 3, \\ \Theta(1), & 1 < \beta_\omega < 3, \end{cases}$$

and

$$\mathcal{D}_N = \zeta\left(\frac{4}{\beta_\omega}, N\right) = \begin{cases} \Theta(N^{1-\frac{4}{\beta_\omega}}), & \beta_\omega > 4, \\ \Theta(\ln N), & \beta_\omega = 4, \\ \Theta(1), & 1 < \beta_\omega < 4. \end{cases}$$

In the same spirit, we obtain

$$\mathcal{P}_1 \approx \frac{N^{1-\frac{1}{\beta_m}}}{1-\frac{1}{\beta_m}} = \Theta\left(N^{\left(1-\frac{1}{\beta_m}\right)}\right),$$

and

$$\mathcal{P}_2 = \zeta\left(\frac{2}{\beta_m}, N\right) = \begin{cases} \Theta\left(N^{1-\frac{2}{\beta_m}}\right), & \beta_m > 2, \\ \Theta\left(\ln N\right), & \beta_m = 2, \\ \Theta(1), & 1 < \beta_m < 2, \end{cases} \qquad \mathcal{P}_3 = \zeta\left(\frac{3}{\beta_m}, N\right) = \begin{cases} \Theta(N^{1-\frac{3}{\beta_m}}), & \beta_m > 3, \\ \Theta(\ln N), & \beta_m = 3, \\ \Theta(1), & 1 < \beta_m < 3. \end{cases}$$

Given these approximations, we estimate the four variance contributions one by one.

*First term.* Using $\mathcal{L}_N^2 \geq \mathcal{A}_N^4 \mathcal{P}_2^2$,

$$\frac{\mathcal{A}_N^2 \mathcal{B}_N \mathcal{P}_3}{\mathcal{L}_N^2} \leq \frac{\mathcal{B}_N \mathcal{P}_3}{\mathcal{A}_N^2 \mathcal{P}_2^2}.$$

We need to show that

$$\frac{\mathcal{B}_N}{\mathcal{A}_N^2} \to 0, \qquad N \to \infty.$$

This is indeed the case since

$$\frac{\mathcal{B}_N}{\mathcal{A}_N^2} = \begin{cases} \Theta(N^{-1}), & \beta_\omega > 2, \\ \Theta\left(\frac{\ln N}{N}\right); & \beta_\omega = 2 \\ \Theta\left(N^{-2\left(1-\frac{1}{\beta_\omega}\right)}\right); & 1 < \beta_\omega < 2. \end{cases} \tag{S.71}$$

Meanwhile, we want to show that

$$\frac{\mathcal{P}_3}{\mathcal{P}_2^2} = O(1).$$

We discuss different possible cases:

If $\beta_m > 3$, then

$$\frac{\mathcal{P}_3}{\mathcal{P}_2^2} = \Theta\left(\left(N^{-\left(1-\frac{1}{\beta_m}\right)}\right) \to 0.$$

If $2 < \beta_m \leq 3$, then $\mathcal{P}_3 = \Theta(1)$ or $\Theta(\ln N)$, whereas $\mathcal{P}_2^2 \to \infty$.

If $\beta_m = 2$,

$$\mathcal{P}_3 = \Theta(1) \qquad \text{and} \qquad \mathcal{P}_2^2 = \Theta((\ln N)^2).$$

If $1 < \beta_m < 2$, then both $\mathcal{P}_2$ and $\mathcal{P}_3$ are $\Theta(1)$.

These discussions lead to the conclusion that $\frac{\mathcal{P}_3}{\mathcal{P}_2^2} = O(1)$. Consequently,

$$\frac{\mathcal{A}_N^2 \mathcal{B}_N \mathcal{P}_3}{\mathcal{L}_N^2} \to 0. \tag{S.72}$$

*Second term.* Using again $\mathcal{L}_N^2 \geq \mathcal{A}_N^4 \mathcal{P}_2^2$,

$$\frac{\mathcal{A}_N \mathcal{C}_N \mathcal{P}_2}{\mathcal{L}_N^2} \leq \frac{\mathcal{C}_N}{\mathcal{A}_N^3 \mathcal{P}_2}.$$

Since $\mathcal{P}_2 \geq 1$ by definition, it suffices to show

$$\frac{\mathcal{C}_N}{\mathcal{A}_N^3} \to 0, \qquad N \to \infty.$$

This is indeed the case since

$$\frac{\mathcal{C}_N}{\mathcal{A}_N^3} = \begin{cases} \Theta(N^{-2}); & \beta_\omega > 3 \\ \Theta\left(\frac{\ln N}{N^2}\right); & \beta_\omega = 3 \\ \Theta(N^{-3\left(1-\frac{1}{\beta_\omega}\right)}); & 1 < \beta_\omega < 3. \end{cases}$$

Therefore, we conclude that

$$\frac{\mathcal{A}_N \mathcal{C}_N \mathcal{P}_2}{\mathcal{L}_N^2} \to 0. \tag{S.73}$$

*Third term.* Using $\mathcal{L}_N^2 \geq \mathcal{B}_N^2 \mathcal{P}_1^2$, we derive that

$$\frac{\mathcal{D}_N \mathcal{P}_1}{\mathcal{L}_N^2} \leq \frac{\mathcal{D}_N}{\mathcal{B}_N^2 \mathcal{P}_1}.$$

Since $\mathcal{P}_1 \to \infty$, it is enough to show

$$\frac{\mathcal{D}_N}{\mathcal{B}_N^2} = O(1).$$

This holds in all cases:
If $\beta_\omega > 4$, then $\frac{\mathcal{D}_N}{\mathcal{B}_N^2} = \Theta(N^{-1})$.
If $\beta_\omega = 4$, $\frac{\mathcal{D}_N}{\mathcal{B}_N^2} = \Theta\left(\frac{\ln N}{N}\right)$.
If $2 < \beta_\omega < 4$, then $\mathcal{D}_N = \Theta(1)$ while $\mathcal{B}_N^2 \to \infty$.
If $\beta_\omega = 2$, $\frac{\mathcal{D}_N}{\mathcal{B}_N^2} = \Theta\left(\frac{1}{(\ln N)^2}\right)$.
If $1 < \beta_\omega < 2$, both $\mathcal{D}_N$ and $\mathcal{B}_N$ are $\Theta(1)$. Consequently, we conclude that $\frac{\mathcal{D}_N}{\mathcal{B}_N^2} = O(1)$ and

$$\frac{\mathcal{D}_N \mathcal{P}_1}{\mathcal{L}_N^2} \to 0. \tag{S.74}$$

*Fourth term.* Using $\mathcal{L}_N^2 \geq \mathcal{B}_N^2 \mathcal{P}_1^2$, we derive that

$$\frac{\mathcal{B}_N^2 \mathcal{P}_2}{\mathcal{L}_N^2} \leq \frac{\mathcal{P}_2}{\mathcal{P}_1^2}.$$

So it suffices to show that

$$\frac{\mathcal{P}_2}{\mathcal{P}_1^2} \to 0, \qquad N \to \infty.$$

This is indeed the case since

$$\frac{\mathcal{P}_2}{\mathcal{P}_1^2} = \begin{cases} \Theta(N^{-1}); & \beta_m > 2 \\ \Theta\left(\frac{\ln N}{N}\right); & \beta_m = 2 \\ \Theta\left(N^{-2\left(1-\frac{1}{\beta_m}\right)}\right); & 1 < \beta_m < 2. \end{cases} \tag{S.75}$$

In all cases, $\frac{\mathcal{P}_2}{\mathcal{P}_1^2} \to 0$. It follows that

$$\frac{\mathcal{B}_N^2 \mathcal{P}_2}{\mathcal{L}_N^2} \to 0. \tag{S.76}$$

Combining the four estimates (S.72)-(S.76) with (S.69) and (S.70) yields

$$\frac{\mathrm{Var}(\mathcal{S}_N)}{\mathcal{L}_N^2} \to 0.$$

Since we have shown that $\mathbb{E}[\mathcal{S}_N] = \Theta(\mathcal{L}_N)$, we conclude that

$$\frac{\mathrm{Var}(\mathcal{S}_N)}{(\mathbb{E}[\mathcal{S}_N])^2} \to 0, \qquad N \to \infty.$$

Finally, by the celebrated Chebyshev's inequality, for every $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{\mathcal{S}_N}{\mathbb{E}[\mathcal{S}_N]} - 1\right| > \varepsilon\right) \leq \frac{\mathrm{Var}(\mathcal{S}_N)}{\varepsilon^2 (\mathbb{E}[\mathcal{S}_N])^2} \to 0.$$

This immediately gives the convergence in probability for the volatility estimate $\mathcal{S}_N$

$$\frac{\mathcal{S}_N}{\mathbb{E}[\mathcal{S}_N]} \xrightarrow{\mathbb{P}} 1, \qquad N \to \infty.$$

This completes the proof. □

## S.7 Additional Lemmas for Proofs of Propositions 5.3 and 5.4

In this section, we provide deterministic characterizations of the model in Section 5 that accounts for both information salience and firm size. Similarly to Lemma D.1, we derive the following result.

**Lemma S.6.** *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*

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

*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 weighted 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,*

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

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

*(4) Let* $j\ge1$*. Then every column sum of* $\mathbf{V}^j$ *(the j-th power of* $\mathbf{V}$*) is at most* $\alpha^j$*. 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* $j\ge1$ *and* $k\ge1$*. Then the k-th weighted row sum* $v_k^{(j)}=\sum_{i=1}^N\omega_i\mathbf{V}_{ki}^j$ *of* $\mathbf{V}^j$ *is upper bounded by*

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

*(6) Let* $k\ge1$*. 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) Consider the $j$-th column of $\mathbf{V}$ and show that its $j$-th column sum is at most $\alpha<1$. First, by Lemma D.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 sum of non-diagonal elements in $j$-th column of $\mathbf{V}$, by Lemma D.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 sum of non-diagonal elements in $j$-th column of $\mathbf{V}$ is equal to

$$\alpha\cdot\sum_{i\notin\mathcal{M}_j}\sum_{k\in\mathcal{D}_i}\omega_i\lambda_i. \tag{S.77}$$

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

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

By Theorem 1, the spectral radius of $\mathbf{V}$ is at most $\alpha$, thus by Definition 4 we obtain that $\mathbf{I}-\mathbf{V}$ is an $M$-matrix.

(2) First, by Lemma D.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 D.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

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

Adding 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

$$v_k = \sum_{i=1}^{N} \omega_i \mathbf{v}_{ki} = \alpha \omega_k \cdot \sum_{i \in \mathcal{M}_k} \omega_i + \alpha \cdot \sum_{i \in \mathcal{M}_k} \sum_{j \notin \mathcal{D}_i} \omega_i \lambda_i \omega_j. \tag{S.78}$$

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 $\mathbf{v}_{kj} (1 \leq j \leq N,\ \ j \neq k)$ in the $k$-th row of $\mathbf{V}$ satisfies

$$\mathbf{v}_{kj} = \alpha \cdot \sum_{i \in \mathcal{M}_k \text{ and } i \notin \mathcal{M}_j} \omega_i \lambda_i \leq \alpha \cdot \sum_{i \in \mathcal{M}_k} \omega_i \lambda_i \leq \alpha \cdot \sum_{i \in \mathcal{M}_k} \frac{\omega_i}{d_i} = \frac{\alpha}{N} \cdot \mu_k' \leq \alpha \mu_k''. \tag{S.79}$$

(4) We prove it by induction. The case $j = 1$ is proved by (1). Now assume that it is true for $j$. Let $\mathbf{V}^j = (\mathbf{V}_1^j \ \ \mathbf{V}_2^j \ \ \ldots \ \ \mathbf{V}_N^j)$ 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 $\alpha^j$ by the 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. When $j = 1$, by (2) we have

$$v_k^{(1)} \leq \alpha \mu_k''.$$

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

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

which completes the proof.

(6) The upper bound is obtained directly from (5). The lower bound follows directly from (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 $\mathbf{I} + \mathbf{V}$. □

In addition, the following lemma is easy to prove.

**Lemma S.7.** *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.

## S.8 Proof of Proposition 5.3

*Proof.* In this proof, we will cite results in Appendix S.6 and Appendix S.7 extensively, in which we present the rigorous mathematical foundation of our random graph model.

Under the random graph formulation

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

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[34]

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

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 $\zeta(1/\beta_m)$ is the convergent Riemann zeta constant when $\beta_m<1$.

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

$$\omega_{\pi(i)}=\frac{\pi(i)^{-\frac{1}{\beta_\omega}}}{\sum_{k=1}^{N}k^{-\frac{1}{\beta_\omega}}}=\frac{1}{\zeta\left(\frac{1}{\beta_\omega},N\right)}\pi(i)^{-\frac{1}{\beta_\omega}}\in[0,1], \qquad \text{such that} \qquad \sum_{i=1}^{N}\omega_{\pi(i)}=1, \tag{S.80}$$

where firm $i$ in the information network is ranked $\pi(i)$ in terms of firm size. Notably, after permutation, the permuted sequences of firm sizes are monotonically decreasing $\omega_1>\omega_2>\cdots>\omega_N$. The largest firm size is linked to some firm $i$ such that $i=\pi^{-1}(1)$. We define the Riemann zeta partial sum as

$$\zeta\left(\frac{1}{\beta_\omega},N\right)\equiv\sum_{k=1}^{N}k^{-\frac{1}{\beta_\omega}}.$$

[34] As we noticed in the main text, to avoid negative probability, when $\beta_m<1$, we define the probability as $\mathbb{P}\left(A_{ij}=1\right)=p_{ij}=\begin{cases}\frac{\frac{N}{j^{1/\beta_m}}}{N-1}\in[0,1]; & \forall j\neq i\\ 1; & j=i\end{cases}$. This change does not affect any of our results and the expected degrees are increased by a constant 1: $\mathbb{E}[m_j]=\frac{N}{j^{1/\beta_m}}+1$; $\mathbb{E}[d_i]=\frac{N}{N-1}\left[\zeta(1/\beta_m,N)-\frac{1}{i^{1/\beta_m}}\right]+1$.

Next, we divide the proof into the following scenarios.

**Scenario 1**: $\beta = \min\{\beta_m, \beta_\omega\} < 1$.

**Case (i):** $\beta_\omega < 1$**.**

In this case,

$$\omega_1 = \max_i \omega_i = \frac{1}{\zeta\left(\frac{1}{\beta_\omega}, N\right)} = \Theta(1).$$

Using (S.5) and (S.6), we obtain

$$\sum_{i=1}^{N} g_i^2 = O\left(\sum_{i=1}^{N}\left(\omega_{\pi(i)} + \sum_{j\in\mathcal{M}_i}\frac{\omega_{\pi(j)}}{d_j}\right)^2\right) = O\left(\sum_{i=1}^{N}(\mu_i^\omega)^2\right).$$

In Appendix S.3, we show that

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

This total sum is independent of the numerical distribution and the combinatorial ordering of $\boldsymbol{\omega}$. By definition,

$$\sum_{i=1}^{N}(\mu_i^\omega)^2 \le \left(\sum_{i=1}^{N}\mu_i^\omega\right)^2 = 4. \tag{S.81}$$

It then follows that

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

On the other hand,

$$\sum_{k=1}^{N} g_k^2 \ge g_{k_0}^2 \ge \Theta\left(\omega_1^2\right) = \Theta(1),$$

where $\pi(k_0) = 1 = \operatorname{argmax}_i \omega_i$. Therefore, in this sub-case,

$$\sigma_{\text{GDP}} = \Theta(1)$$

is a sure event.

**Case (ii):** $\beta_\omega \ge 1$**,** $\beta_m < 1$**.**

Using Lemma S.6, we consider the bound for GDP elasticity with a given size permutation $\pi$,

$$\delta\left(\omega_{\pi(k)} + \alpha\omega_{\pi(k)}\cdot\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)} + \frac{\alpha}{\tau_\xi/\tau_\varepsilon + 1}\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\sum_{j\notin\mathcal{D}_i}\omega_{\pi(j)}\right)$$

$$\le g_k$$

$$\leq \delta\left(\omega_{\pi(k)} + \frac{\alpha}{(1-\alpha)^2}\cdot\left(\omega_{\pi(k)}\cdot\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)} + \sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)\right). \tag{S.82}$$

We have shown that the volatility of GDP $\sum_{k=1}^{N} g_k^2$ has a constant upper bound that is independent of the numerical distribution and the combinatorial ordering of $\boldsymbol{\omega}$. Therefore, it suffices to estimate a probabilistic lower bound for the volatility. If $\beta_\omega > 1$, in the proof of Proposition S.1, we show that

$$\sum_{j\notin\mathcal{D}_i}\omega_{\pi(j)} \approx 1, \qquad \text{a.s.,} \qquad \forall i \in [N],$$

for all firms simultaneously (see the proof around (S.52) and (S.53)). If $\beta_\omega = 1$, we use the Corollary S.4 so that

$$\sum_{j\notin\mathcal{D}_i}\omega_{\pi(j)} = \Theta\left(1\right); \quad \text{a.s.,} \quad \forall i \in [N],$$

for all firms simultaneously.

Meanwhile, in this case, the expected out-degree is given by

$$\mathbb{E}[d_i] = \frac{N}{N-1}\left[\zeta(1/\beta_m, N) - \frac{1}{i^{1/\beta_m}}\right] + 1 = \Theta(1).$$

Next, we use the concentration inequality in Lemma S.1 to derive

$$\mathbb{P}\left(d_i > \mathbb{E}[d_i] + t\right) \leq \exp\left(-\frac{t^2}{2\,\mathrm{Var}(d_i) + 2t/3}\right) \leq \exp\left(-\frac{t^2}{2\mathbb{E}[d_i] + 2t/3}\right). \tag{S.83}$$

When $t$ tends to infinity, $\exp\left(-\frac{t^2}{2\mathbb{E}[d_i]+2t/3}\right)$ tends to $0$. On the other hand, $d_i \geq 1$ always holds, thus we obtain

$$d_i = \Theta_p(1), \text{ bounded in probability,} \qquad \forall i \in [N].$$

Without loss of generality, we consider the first column of the random graph. It follows that

$$g_1^2 \geq \Theta_p\left(\left(\sum_{i\in\mathcal{M}_1}\frac{\omega_{\pi(i)}}{d_i}\right)^2\right).$$

By the random-graph structure, the first column of the adjacency matrix contains all 1's. Therefore,

$$\sum_{i\in\mathcal{M}_1}\frac{\omega_{\pi(i)}}{d_i} = \Theta_p\left(\sum_{i\in\mathcal{M}_1}\omega_{\pi(i)}\right) = \Theta_p\left(\sum_{i=1}^{N}A_{i1}\omega_{\pi(i)}\right) = \Theta_p\left(\sum_{i=1}^{N}\omega_{\pi(i)}\right) = \Theta_p(1), \text{ bounded in probability.}$$

To summarize, we arrive at the desired lower bound for GDP volatility,

$$\sum_{k=1}^{N} g_k^2 \geq g_1^2 \geq \Theta_p\left(\left(\sum_{i\in\mathcal{M}_1}\frac{\omega_{\pi(i)}}{d_i}\right)^2\right) = \Theta_p(1), \quad \text{bounded in probability.}$$

**Scenario 2:** $\beta = \min\{\beta_m, \beta_\omega\} > 2$.
In this scenario, we note that (S.16) holds, that is,

$$\frac{1}{\beta_m} + \frac{1}{\beta_\omega} < 1.$$

Therefore, we directly apply Proposition S.1. In particular, all of the following results hold a.s.:

1. $\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)} \approx \frac{1}{k^{1/\beta_m}}$;
2. $d_k \approx d = \zeta(\frac{1}{\beta_m}, N) = o(N)$;
3. $\sum_{j\notin\mathcal{D}_k}\omega_{\pi(j)} \approx 1$;

$\forall k = 1, 2, \ldots, N$ simultaneously.

Therefore, by Lemma S.6 we obtain

$$g_k = \Theta\left(\omega_{\pi(k)} + \sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\cdot\left(\omega_{\pi(k)} + \frac{1}{d_i}\right)\right) = \Theta\left(\omega_{\pi(k)} + \sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right), \quad \text{a.s.}$$

$\forall k = 1, 2, \ldots, N$ simultaneously.

Furthermore, by Lemma S.7 we derive that for each $1 \leq k \leq N$,

$$\omega_{\pi(k)}^2 + \left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2 \leq \left(\omega_{\pi(k)} + \sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2 \leq 2\left(\omega_{\pi(k)}^2 + \left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2\right), \quad \text{a.s.}$$

Therefore, taking the sum over $k = 1$ to $N$ in the above inequality, we obtain

$$\sum_{k=1}^{N} g_k^2 = \Theta\left(\sum_{k=1}^{N}\omega_{\pi(k)}^2 + \sum_{k=1}^{N}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2\right) = \Theta\left(\sum_{k=1}^{N}\omega_k^2 + \sum_{k=1}^{N}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2\right), \quad \text{a.s.} \tag{S.84}$$

The second equality uses the fact that for any bijective permutation $\pi$, $\sum_{k=1}^{N}\omega_{\pi(k)}^2 = \sum_{k=1}^{N}\omega_k^2$.

For the first term $\sum_{k=1}^{N}\omega_k^2$ in the last equality, by (S.33) we have a.s.,

$$\sum_{k=1}^{N}\omega_k^2 = \frac{\sum_{k=1}^{N}k^{-\frac{2}{\beta_\omega}}}{\left(\zeta\left(\frac{1}{\beta_\omega}, N\right)\right)^2} = \frac{\zeta\left(\frac{2}{\beta_\omega}, N\right)}{\left(\zeta\left(\frac{1}{\beta_\omega}, N\right)\right)^2} \approx \begin{cases} \Theta\left(\frac{1}{N}\right); & \beta_\omega > 2; \\ \Theta\left(\frac{\ln N}{N}\right); & \beta_\omega = 2; \\ \Theta\left(\frac{1}{N^{2-2/\beta_\omega}}\right). & 1 < \beta_\omega < 2. \end{cases} \tag{S.85}$$

For the second term $\sum_{k=1}^{N}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2$ in the last equality, we have a.s.,

$$\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2 \approx \frac{1}{d^2}\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right)^2 = \Theta\left(\left(\frac{1}{\zeta(\frac{1}{\beta_m},N)\cdot k^{1/\beta_m}}\right)^2\right). \tag{S.86}$$

Therefore, we have a.s.,

$$\sum_{k=1}^{N}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2 = \Theta\left(\sum_{k=1}^{N}\left(\frac{1}{\zeta(\frac{1}{\beta_m},N)k^{1/\beta_m}}\right)^2\right) = \Theta\left(\frac{\zeta\left(\frac{2}{\beta_m},N\right)}{\left(\zeta\left(\frac{1}{\beta_m},N\right)\right)^2}\right) = \begin{cases} \Theta\left(\frac{1}{N}\right); & \beta_m > 2; \\ \Theta\left(\frac{\ln N}{N}\right); & \beta_m = 2; \\ \Theta\left(\frac{1}{N^{2-2/\beta_m}}\right). & 1 < \beta_m < 2. \end{cases} \tag{S.87}$$

Naturally, in this scenario, we have

$$\sigma_{\text{GDP}} = \Theta\left(\frac{1}{\sqrt{N}}\right), \qquad \text{a.s.}$$

**Scenario 3:** $\beta = \min\{\beta_m, \beta_\omega\} = 2$

**Case (i)**: $\beta_m > 2, \beta_\omega = 2$ or $\beta_m = 2, \beta_\omega > 2$.

In this case, (S.16) still holds. Hence we apply Proposition S.1 again with (S.84), (S.85), (S.86), and (S.87) to conclude that

$$\sigma_{\text{GDP}} = \Theta\left(\sqrt{\frac{\ln N}{N}}\right), \qquad \text{a.s.}$$

**Case (ii)**: $\beta_m = \beta_\omega = 2$.

This is a knife-edge boundary case. We note that by Proposition S.1, the conditions $\sum_{j\notin\mathcal{D}_k}\omega_{\pi(j)} \approx 1$ and $d_k \approx d = \zeta(\frac{1}{\beta_m}, N) = o(N)$ still hold. As such the volatility estimate of GDP is still given by (S.84). Moreover, we observe that

$$\frac{1}{\beta_m} + \frac{1}{\beta_\omega} = 1.$$

Therefore, we apply Corollary S.1 so that

$$\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)} \approx \Theta\left(\frac{m_k}{N}\right), \qquad \forall k = 1, 2, \ldots, N.$$

holds almost surely as $N \to \infty$, for all firms simultaneously. It is then straightforward to show that (S.86) is still valid:

$$\sum_{k=1}^{N}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2 = \Theta\left(\sum_{k=1}^{N}\left(\frac{1}{\zeta(\frac{1}{\beta_m},N)k^{1/\beta_m}}\right)^2\right).$$

It is immediate that we arrive at the same volatility estimate,

$$\sigma_{\text{GDP}} = \Theta\left(\sqrt{\frac{\ln N}{N}}\right), \qquad \text{a.s.}$$

**Scenario 4:** $\beta = \min\{\beta_m, \beta_\omega\} \in (1, 2)$.
In this scenario, we face a key difficulty that the technical condition $\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} \approx \Theta\left(\frac{m_k}{N}\right)$ does not hold, even individually, for $k$ sufficiently large. To overcome this problem, we use Proposition S.2 to provide a volatility estimate with high probability (w.h.p).

To maintain consistency, we use notation in that proposition. In particular, we observe the following facts. First, we define size-related statistics as

$$\omega_i = \frac{a_i}{\mathcal{A}_N}, \qquad a_i = i^{-1/\beta_\omega}, \qquad \mathcal{A}_N = \sum_{i=1}^{N} a_i.$$

Second, we define probability-related statistics as

$$p_j = j^{-1/\beta_m}, \qquad \mathcal{P}_1 = \sum_{j=1}^{N} p_j = \sum_{j=1}^{N} j^{-1/\beta_m}.$$

Third, when $1 < \beta < 2$, by Proposition S.1 the conditions $\sum_{j\notin\mathcal{D}_k} \omega_{\pi(j)} \approx 1$ and $d_k \approx d = \zeta(\frac{1}{\beta_m}, N) = o(N)$ still hold almost surely. Therefore, there exist constants $0 < c_1 \le c_2 < \infty$ such that, w.h.p., for all sufficiently large $N$ and all $i$,

$$c_1\mathcal{P}_1 \le d_i \le c_2\mathcal{P}_1.$$

Equivalently, w.h.p.,

$$d_i = \Theta(\mathcal{P}_1) \qquad \text{uniformly in } i.$$

Meanwhile, the volatility bound (S.84) still holds.

Fourth, in our random graph model, for each $k \in [N]$ and any permutation $\pi$, we have

$$\sum_{i\in M_k} \omega_{\pi(i)} = \sum_{i=1}^{N} \omega_{\pi(i)} A_{i,k}.$$

Since

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

when $i \neq j$, it is easy to see that

$$\mathbb{P}(A_{ij} = 1) = p_{ij} = \frac{Nj^{-1/\beta_m} - 1}{N-1} \approx j^{-1/\beta_m} = p_j.$$

When $N$ is sufficiently large. Consequently, for any permutation $\pi$,

$$\begin{aligned}
\sum_{k=1}^{N} \left( \sum_{i \in M_k} \omega_{\pi(i)} \right)^2 &= \sum_{k=1}^{N} \left( \sum_{i=1}^{N} \omega_{\pi(i)} A_{i,k} \right)^2 \\
&= \sum_{k=1}^{N} \left( \omega_{\pi(k)} + \sum_{i \neq k} \omega_{\pi(i)} A_{i,k} \right)^2 \\
&= \Theta\left( \sum_{k=1}^{N} \left( \omega_{\pi(k)}^2 + \left( \sum_{i \neq k} \omega_{\pi(i)} A_{i,k} \right)^2 \right) \right) \\
&= \Theta\left( \sum_{k=1}^{N} \left( \sum_{i=1}^{N} \omega_i A_{i,k} \right)^2 \right) + O\left( \sum_{k=1}^{N} \omega_k^2 \right).
\end{aligned} \tag{S.88}$$

Note that we utilize Lemma S.7 in the derivation.

Given the above facts, we now show that the GDP volatility estimate is given by

$$\sum_{k=1}^{N} g_k^2 = \Theta\left( \frac{\mathcal{B}_N}{\mathcal{A}_N^2} + \frac{\mathcal{P}_2}{\mathcal{P}_1^2} \right) \qquad \text{w.h.p.} \tag{S.89}$$

We refer the reader to Lemma S.5 and Proposition S.2 in Appendix S.6.4 for the notation. To prove the claim, we note that w.h.p.,

$$\frac{1}{d_i} = \Theta\left( \frac{1}{\mathcal{P}_1} \right) \qquad \text{uniformly in } i \in [N].$$

Hence w.h.p., for each $k \in [N]$,

$$\sum_{i \in M_k} \frac{\omega_{\pi(i)}}{d_i} = \Theta\left( \frac{1}{\mathcal{P}_1} \sum_{i \in M_k} \omega_{\pi(i)} \right).$$

Squaring both sides and summing over $k$ yields

$$\sum_{k=1}^{N} \left( \sum_{i \in M_k} \frac{\omega_{\pi(i)}}{d_i} \right)^2 = \Theta\left( \frac{1}{\mathcal{P}_1^2} \sum_{k=1}^{N} \left( \sum_{i \in M_k} \omega_{\pi(i)} \right)^2 \right).$$

Using (S.88),

$$\sum_{k=1}^{N}\left(\sum_{i\in M_k}\omega_{\pi(i)}\right)^2=\Theta\left(\sum_{k=1}^{N}\left(\sum_{i=1}^{N}\omega_i A_{i,k}\right)^2\right)+O\left(\sum_{k=1}^{N}\omega_k^2\right).$$

Since we defined $\omega_i = a_i/\mathcal{A}_N$,

$$\sum_{i=1}^{N}\omega_i A_{i,k}=\frac{1}{\mathcal{A}_N}\sum_{i=1}^{N}a_i A_{i,k}.$$

Therefore,

$$\sum_{k=1}^{N}\left(\sum_{i=1}^{N}\omega_i A_{i,k}\right)^2=\frac{1}{\mathcal{A}_N^2}\sum_{k=1}^{N}\left(\sum_{i=1}^{N}a_i A_{i,k}\right)^2=\frac{\mathcal{S}_N}{\mathcal{A}_N^2}.$$

where the definition of $\mathcal{S}_N$ is given in Appendix S.6.4. It follows that

$$\sum_{k=1}^{N}\left(\sum_{i\in M_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2=\Theta\left(\frac{\mathcal{S}_N}{\mathcal{P}_1^2\mathcal{A}_N^2}\right)+O\left(\frac{1}{\mathcal{P}_1^2}\sum_{k=1}^{N}\omega_k^2\right),\qquad\text{and}\qquad\sum_{k=1}^{N}\omega_k^2=\frac{1}{\mathcal{A}_N^2}\sum_{k=1}^{N}a_k^2=\frac{\mathcal{B}_N}{\mathcal{A}_N^2}.$$

Hence, by (S.84),

$$\sum_{k=1}^{N}g_k^2=\Theta\left(\frac{\mathcal{B}_N}{\mathcal{A}_N^2}+\frac{\mathcal{S}_N}{\mathcal{P}_1^2\mathcal{A}_N^2}\right)\qquad\text{w.h.p.}\tag{S.90}$$

In Proposition S.2, we proved a weak law of large numbers (WLLN) for $\mathcal{S}_N$:

$$\mathcal{S}_N=(1+o_P(1))\mathbb{E}[\mathcal{S}_N].$$

Moreover,

$$\mathbb{E}[\mathcal{S}_N]=\mathcal{A}_N^2\mathcal{P}_2+\mathcal{B}_N(\mathcal{P}_1-\mathcal{P}_2)=\Theta(\mathcal{A}_N^2\mathcal{P}_2+\mathcal{B}_N\mathcal{P}_1).$$

It follows that

$$\frac{\mathcal{S}_N}{\mathcal{P}_1^2\mathcal{A}_N^2}=\Theta\left(\frac{\mathcal{A}_N^2\mathcal{P}_2+\mathcal{B}_N\mathcal{P}_1}{\mathcal{P}_1^2\mathcal{A}_N^2}\right)\qquad\text{w.h.p.,}$$

or equivalently,

$$\frac{\mathcal{S}_N}{\mathcal{P}_1^2\mathcal{A}_N^2}=\Theta\left(\frac{\mathcal{P}_2}{\mathcal{P}_1^2}+\frac{\mathcal{B}_N}{\mathcal{A}_N^2\mathcal{P}_1}\right)\qquad\text{w.h.p.}$$

Since $\mathcal{P}_1\to\infty$ whenever $\beta_m>1$,

$$\frac{\mathcal{B}_N}{\mathcal{A}_N^2\mathcal{P}_1}=o\left(\frac{\mathcal{B}_N}{\mathcal{A}_N^2}\right).$$

Therefore, by (S.90), we obtain

$$\sum_{k=1}^{N}g_k^2=\Theta\left(\frac{\mathcal{B}_N}{\mathcal{A}_N^2}+\frac{\mathcal{P}_2}{\mathcal{P}_1^2}\right)\qquad\text{w.h.p.,}$$

which proves our claim in (S.89).

It remains to compute the two deterministic orders. These quantities were calculated in (S.71) and (S.75). Namely,

$$\frac{\mathcal{B}_N}{\mathcal{A}_N^2} = \begin{cases} \Theta(N^{-1}), & \beta_\omega > 2, \\ \Theta\left(\frac{\ln N}{N}\right); & \beta_\omega = 2 \\ \Theta\left(N^{-2\left(1-\frac{1}{\beta_\omega}\right)}\right); & 1 < \beta_\omega < 2 \end{cases}; \qquad \frac{\mathcal{P}_2}{\mathcal{P}_1^2} = \begin{cases} \Theta(N^{-1}); & \beta_m > 2 \\ \Theta\left(\frac{\ln N}{N}\right); & \beta_m = 2 \\ \Theta\left(N^{-2\left(1-\frac{1}{\beta_m}\right)}\right); & 1 < \beta_m < 2. \end{cases}$$

Since $\beta = \min\{\beta_m, \beta_\omega\}$, the slower-decaying term is determined by $\beta$, and therefore

$$\sum_{k=1}^{N} g_k^2 = \Theta\left(N^{-(2-2/\beta)}\right), \qquad 1 < \beta < 2, \qquad \text{w.h.p.}$$

Taking square roots yields the following results.

$$\sigma_{GDP} = \Theta\left(N^{-(1-1/\beta)}\right), \qquad \text{when } 1 < \beta < 2, \qquad \text{w.h.p.}$$

Note that we can use this method to derive similar w.h.p. estimates for the case $\beta \geq 2$, but they are weaker than the almost sure estimates. We omit them for brevity.

**Scenario 5:** $\beta = \min\{\beta_m, \beta_\omega\} = 1$.

We first notice that when $\beta_\omega \geq 1$ and $\beta_m > 0$, Proposition S.1 and Corollary S.4 imply that for any arbitrary permutation $\pi$ on $\boldsymbol{\omega}$,

$$\sum_{j \notin \mathcal{D}_k} \omega_{\pi(j)} = \Theta(1); \qquad \text{a.s.}$$

for all firms $k \in [N]$ simultaneously. Therefore, (S.84) still holds a.s. in this scenario:

$$\sum_{k=1}^{N} g_k^2 = \Theta\left(\sum_{k=1}^{N} \omega_k^2 + \sum_{k=1}^{N}\left(\sum_{i \in \mathcal{M}_k} \frac{\omega_{\pi(i)}}{d_i}\right)^2\right), \quad \text{a.s.} \tag{S.91}$$

We then discuss different possible cases.

**Case (i)**: $\beta_\omega = 1, \beta_m \geq 1$.

By (S.91),

$$\sum_{k=1}^{N} g_k^2 \geq \Theta\left(\sum_{k=1}^{N} \omega_k^2\right), \qquad \text{a.s.}$$

The deterministic sum is given by

$$\sum_{k=1}^{N} \omega_k^2 = \frac{\sum_{k=1}^{N} k^{-\frac{2}{\beta_\omega}}}{\left(\zeta\left(\frac{1}{\beta_\omega}, N\right)\right)^2} = \frac{\zeta\left(\frac{2}{\beta_\omega}, N\right)}{\left(\zeta\left(\frac{1}{\beta_\omega}, N\right)\right)^2} = \frac{\frac{\pi^2}{6}}{(\ln N)^2}, \qquad \beta_\omega = 1.$$

Therefore, in this case,

$$\sigma_{\text{GDP}} \geq \Theta\left(\frac{1}{\ln N}\right), \qquad \text{a.s.}$$

**Case (ii)**: $\beta_\omega > 1, \beta_m = 1$.
In this case, we analyze the second component of the volatility estimate in (S.91) as

$$\sum_{k=1}^{N} g_k^2 \geq \Theta\left(\sum_{k=1}^{N}\left(\sum_{i\in\mathcal{M}_k} \frac{\omega_{\pi(i)}}{d_i}\right)^2\right), \qquad \text{a.s.}$$

To estimate the RHS of this equation, we employ Corollaries S.3 and S.2. Specifically, by Corollary S.2, when $\beta_m = 1, \beta_\omega > 1$, we pick a subset of economy $K_N$ of firms with the highest attention, i.e., the first $N^\Lambda$ columns in the stochastic adjacency matrix,

$$K_N := \{1, 2, \ldots, \lfloor N^\Lambda \rfloor\},$$

where the cardinality $\Lambda$ satisfies

$$0 < \Lambda < 1 - \frac{1}{\beta_\omega}.$$

In this subset of economy, we know that

$$\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} \approx \frac{m_k}{N} \approx \frac{1}{k^{1/\beta_m}} \qquad \text{a.s.}$$

holds almost surely eventually and for all $k \in [K_N]$ simultaneously. In addition, by Corollary S.3, we pick another subset of the economy $\widetilde{K}_N$ that corresponds to the first $N^{\frac{1}{24}}$ columns in the stochastic adjacency matrix,

$$\widetilde{K}_N := \{1, 2, \ldots, \lfloor N^{\frac{1}{24}} \rfloor\}.$$

In this subset, we have

$$d_i = \Theta\left(\zeta\left(\frac{1}{\beta_m}, N\right)\right) = \Theta\left(\ln N\right); \qquad \text{w.h.p.}$$

for all $N$ sufficiently large and $i \in \widetilde{K}_N$ simultaneously. Now take the intersection of these two subsets,

$$K_N^* = K_N \bigcap \widetilde{K}_N =: = \{1, 2, \ldots, \lfloor N^{\Lambda^*} \rfloor\}; \qquad \Lambda^* = \min\{\Lambda, \frac{1}{24}\}.$$

It follows that for every firm $k \in K_N^*$, when $\beta_m = 1$,

$$\sum_{k\in K_N^*}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2 = \sum_{k=1}^{N^{\Lambda^*}}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2 = \Theta\left(\frac{1}{\zeta(1,N)^2}\sum_{k=1}^{N^{\Lambda^*}}k^{-2}\right) = \Theta\left(\frac{1}{(\ln N)^2}\right); \qquad \text{w.h.p.}$$

In the derivation of the above equation, we use the fact that when $N \to \infty$, $N^{\Lambda^*} \to \infty$ so that $\sum_{k=1}^{N^{\Lambda^*}} k^{-2} \to \frac{\pi^2}{6}$. By definition,

$$\sum_{k=1}^{N}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2 \geq \sum_{k\in K_N^*}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2.$$

Therefore, in this case we have

$$\sum_{k=1}^{N} g_k^2 \geq \Theta\left(\sum_{k=1}^{N}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2\right) \geq \Theta\left(\sum_{k\in K_N^*}\left(\sum_{i\in\mathcal{M}_k}\frac{\omega_{\pi(i)}}{d_i}\right)^2\right) = \Theta\left(\frac{1}{(\ln N)^2}\right); \qquad \text{w.h.p.}$$

Taking square roots gives

$$\sigma_{\text{GDP}} \geq \Theta\left(\frac{1}{\ln N}\right), \qquad \text{w.h.p.}$$

Since almost-sure convergence implies convergence in probability (w.h.p.), combining the two cases completes our proof of Scenario 5. In particular, when $\beta = 1$, with high probability, GDP volatility decays at a rate no faster than $\frac{1}{\ln N}$.

To summarize, we note that all of our proofs are based on an arbitrary choice of the permutation $\pi : i \longmapsto [j]$. Hence, our proof of Proposition 5.3 is now complete. □

## S.9 Proof of Proposition 5.4

*Proof.* We break down the proof of this proposition into several steps. First, given a (deterministic) information network $\mathbf{A}$ and a numerical distribution of $\boldsymbol{\omega}$, we show that the set of joint regularity conditions for $\mathbf{A}$ and $\boldsymbol{\omega}$ proposed in Proposition S.1 lead to the three combinatorial properties of the volatility of GDP.

**Step 1: Deterministic Results on Combinatorial Ordering**

Given a fixed information network $\mathbf{A}$ and a numerical distribution of $\boldsymbol{\omega}$, without loss of generality, we arrange the firm index such that $m_1 > m_2 > \cdots > m_N$ is monotonically decreasing. Such relabeling would not affect the computation of GDP volatility, and the firm indexes correspond to in-degree sorting. We then arrange the firm size sequence in monotonically decreasing order: $\omega_{[1]} > \omega_{[2]} > \cdots > \omega_{[N]}$, and $\sum_{j=1}^{N}\omega_{[j]} = 1$.

A permutation $\pi$ is defined as a bijective map that assigns a size $\omega_{[j]}$ for each firm $i$ with corre-

sponding in-degree $m_i$,

$$\pi : i \in [1, 2, \ldots, N] \longmapsto [j] \in [1, 2, \ldots, N]. \tag{S.92}$$

We start with the following lemma.

**Lemma S.8.** *Consider an information network* $\mathbf{A}$ *and a firm-size distribution* $\boldsymbol{\omega}$*. Suppose that for any possible permutation of* $\pi$ *under* $(\mathbf{A}, \boldsymbol{\omega})$*, as* $N \to \infty$*,*

1. $\sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \approx \frac{m_k}{N}, \quad \forall k = 1, 2, \ldots, N.$
2. $d_i \approx d = o(N), \quad \forall i = 1, 2, \ldots, N.$
3. $\sum_{i \notin \mathcal{D}_k} \omega_{\pi(i)} \approx 1, \quad \forall k = 1, 2, \ldots, N.$

*for all firms simultaneously. Then up to a leading-order approximation, the three combinatorial properties of GDP volatility in Proposition 5.4 hold.*[35]

*Proof.* Given $(\mathbf{A}, \boldsymbol{\omega})$, it suffices to consider the following set of permutations:

$$\Pi = \left\{ \pi : i \longmapsto [j], \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \approx \frac{m_k}{N}, \sum_{i \notin \mathcal{D}_k} \omega_{\pi(i)} \approx 1, \forall k = 1, 2, \ldots, N \right\}. \tag{S.93}$$

That is, we consider the collection of permutations that are consistent with our two restrictions imposed in the lemma. By assumption, $\Pi$ contains all possible permutations with $|\Pi| = N!$.

Given a permutation $\pi \in \Pi$, GDP volatility conditional on micro shocks is given by

$$\sigma_{\text{GDP}}(\pi) = \sum_{i=1}^{N} g_i^2(\pi) \tau_\varepsilon^{-1}. \tag{S.94}$$

Now fix an arbitrary $\pi \in \Pi$. By (S.78), the $k$-th weighted row sum of $\mathbf{V}$ is given by

$$v_k = \sum_{i=1}^{N} \omega_{\pi(i)} \mathbf{v}_{ki} = \alpha \omega_{\pi(k)} \cdot \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} + \alpha \cdot \sum_{i \in \mathcal{M}_k} \sum_{j \notin \mathcal{D}_i} \omega_{\pi(i)} \lambda_i \omega_{\pi(j)}.$$

When N is sufficiently large, $d_i \approx d = o(N)$ for each $1 \leq i \leq N$, it follows that

$$\lambda_i = \frac{1}{d_i + \tau_\xi / \tau_\varepsilon} = \frac{1}{\widehat{d}_i}; \qquad \widehat{d}_i = d\,(1 + o_i(1)) + \tau_\xi / \tau_\varepsilon.$$

[35] In our random graph model (5.5) and (5.8), Proposition S.1 shows that under appropriate assumptions of $(\beta_m, \beta_\omega)$, the entire set of possible size permutations satisfies these restrictions in almost all sample realizations of graph $\mathbf{A}$. The restrictions on $(\mathbf{A}, \boldsymbol{\omega})$ in this lemma can also be relaxed so that only a subset of permutations $\pi$ satisfy these properties. Within this subset, the ordering properties of GDP volatility remain valid.

Note that the remainder errors $o_i(1)$ are in general heterogeneous across firms but uniformly bounded. Therefore,

$$v_k = \alpha\omega_{\pi(k)} \cdot \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} + \alpha \sum_{i\in\mathcal{M}_k} \frac{\omega_{\pi(i)}}{\widehat{d}_i} \sum_{j\notin\mathcal{D}_i} \omega_{\pi(j)}. \tag{S.95}$$

Similarly, when $N$ is sufficiently large, $\sum_{j\notin\mathcal{D}_i} \omega_{\pi(j)} \approx 1$ and $\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} \approx \frac{m_k}{N}$ for all $\pi \in \Pi$. Hence,

$$\begin{aligned} v_k &= \alpha\omega_{\pi(k)} \cdot \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} + \alpha \sum_{i\in\mathcal{M}_k} \frac{\omega_{\pi(i)}}{\widehat{d} + o_i(d)} \left(1 + o_{i,\pi}(1)\right) \\ &= \alpha\omega_{\pi(k)} \frac{m_k}{N} \left(1 + o_{k,\pi}(1)\right) + \alpha \sum_{i\in\mathcal{M}_k} \frac{\omega_{\pi(i)}}{\widehat{d} + o_i(d)} \left(1 + o_{i,\pi}(1)\right) \end{aligned}$$

where $\widehat{d} = d + \tau_\xi/\tau_\varepsilon$ and $o_{i,\pi}(1)$ and $o_{k,\pi}(1)$ are uniformly bounded. After the leading-order approximation, we drop the $o$ terms. It follows that[36]

$$v_k = \alpha\omega_{\pi(k)} \frac{m_k}{N} + \frac{\alpha}{\widehat{d}} \frac{m_k}{N}.$$

By Proposition 4.2 and Lemma S.6, $(\mathbf{I} - \mathbf{V})^{-1} = \mathbf{I} + \mathbf{V} + O\left(\mathbf{V}^2\right)$. Up to the leading (first)-order approximation, we drop higher-order terms. Then,

$$\begin{aligned} g_k &= \delta(\omega_{\pi(k)} + v_k) \\ &= \delta\omega_{\pi(k)} + \alpha\delta\omega_{\pi(k)} \cdot \frac{m_k}{N} + \frac{\alpha\delta}{\widehat{d}} \cdot \frac{m_k}{N} \\ &= \left(\delta + \frac{\alpha\delta}{N} m_k\right) \omega_{\pi(k)} + \left(\frac{\alpha\delta}{N\widehat{d}}\right) m_k, \qquad \forall k \in [N]. \end{aligned} \tag{S.96}$$

We define two positive, non-increasing sequences $\{\phi_k\}_{k=1}^N$ and $\{\varphi_k\}_{k=1}^N$ as

$$\phi_k = \left(\delta + \frac{\alpha\delta}{N} m_k\right)^2 > 0; \qquad \varphi_k = 2\left(\frac{\alpha\delta}{N\widehat{d}}\right) m_k \left(\delta + \frac{\alpha\delta}{N} m_k\right) > 0; \quad \forall k = 1, 2, \ldots, N, \tag{S.97}$$

[36] Keeping the remainder term may flip our ordering results under certain permutations, and in this case we can construct an admissible set of permutations $\Pi^{\text{adm}} \subseteq \Pi$ that is closed under inversion-reducing adjacency swaps; in particular, for any $\pi \in \Pi^{\text{adm}}$, we write $v_k(\pi) = l_k(\pi) + r_k(\pi)$. $l_k(\pi) = \alpha\omega_{\pi(k)} \frac{m_k}{N} + \frac{\alpha}{\widehat{d}} \frac{m_k}{N}$ denotes the leading-order term and $r_k(\pi)$ denotes the remainder error. Such an admissible class of permutations preserves our combinatorial properties if

$$\sup_{k\in[N],\pi\in\Pi^{\text{adm}}} |r_k(\pi)| = o\left(\inf_{k\in[N],\pi\in\Pi^{\text{adm}}} l_k(\pi)\right).$$

Because the economic messages are the same, we leave out the higher-order terms for the sake of simplicity.

which preserve the order of $\boldsymbol{m}$. Therefore, the volatility measure is given by

$$\Psi(\pi) = \sum_{k=1}^{N} g_k^2(\pi) = \sum_{k=1}^{N} \phi_k \left(\omega_{\pi(k)}\right)^2 + \sum_{k=1}^{N} \varphi_k \omega_{\pi(k)} + \underbrace{\sum_{k=1}^{N} \left(\frac{\alpha\delta}{N\widehat{d}}\right)^2 m_k^2}_{\pi-\text{invariant}}. \tag{S.98}$$

The last term is permutation-invariant. Therefore, we focus on the first two terms. Note that $\{\phi_k\}_{k=1}^N$ and $\{\varphi_k\}_{k=1}^N$ are co-monotone,

$$\phi_1 > \phi_2 > \cdots > \phi_N; \qquad \varphi_1 > \varphi_2 > \cdots > \varphi_N.$$

We then apply the classical rearrangement inequality ( Hardy et al. (1952), Theorem 368 and Marshall et al. (2011), pp. 203-239) to deduce that $\forall \pi \in \Pi$,

$$\sum_{k=1}^{N} \phi_k \left(\omega_{[N+1-k]}\right)^2 \leq \sum_{k=1}^{N} \phi_k \left(\omega_{\pi(k)}\right)^2 \leq \sum_{k=1}^{N} \phi_k \left(\omega_{[k]}\right)^2; \qquad \sum_{k=1}^{N} \varphi_k \omega_{[N+1-k]} \leq \sum_{k=1}^{N} \varphi_k \omega_{\pi(k)} \leq \sum_{k=1}^{N} \varphi_k \omega_{[k]}; \tag{S.99}$$

with equality on the right (left) when $\phi$, $\varphi$ and $\omega$ are co-monotone (counter-monotone). That is, the leading-order approximation of volatility is maximized when $\boldsymbol{m}$ and $\boldsymbol{\omega}$ are organized in the same order, and is minimized when $\boldsymbol{m}$ and $\boldsymbol{\omega}$ are organized in the exact opposite order.[37] The classical rearrangement inequality applies when $N$ is finite. When $N \to \infty$, by basic real analysis, we have the inequality

$$\lim_{N\to\infty} \inf \Psi_{\min} \leq \lim_{N\to\infty} \inf \Psi(\pi) \leq \lim_{N\to\infty} \sup \Psi(\pi) \leq \lim_{N\to\infty} \sup \Psi_{\max}; \qquad \forall \pi \in \Pi \tag{S.100}$$

where $\Psi_{\min}$ and $\Psi_{\max}$ denote counter-monotone minimum and co-monotone maximum. Therefore, rearranged sums of any permutation lie between sums of sorted and reversed-sorted matchings as $N \to \infty$.

We prove the third statement regarding the inversion monotonicity of the GDP volatility using the standard inversion-swap argument in rearrangement inequalities. Consider a permutation $\pi$, and suppose that $m_k > m_j$ (so that $k < j$) and $\omega_{\pi(k)} > \omega_{\pi(j)}$. If we swap the order of $\omega_{\pi(k)}$ and $\omega_{\pi(j)}$ while keeping all other parameters unchanged, this leads to a new permutation $\pi'$, in particular,

$$g_k' = \left(\delta + \frac{\alpha\delta}{N} m_k\right) \omega_{\pi(j)} + \left(\frac{\alpha\delta}{N\widehat{d}}\right) m_k;$$
$$g_j' = \left(\delta + \frac{\alpha\delta}{N} m_j\right) \omega_{\pi(k)} + \left(\frac{\alpha\delta}{N\widehat{d}}\right) m_j.$$

[37] Alternatively, the mappings $\boldsymbol{\omega} \mapsto \sum_{k=1}^{N} \phi_k \left(\omega_{\pi(k)}\right)^2$ and $\boldsymbol{\omega} \mapsto \sum_{k=1}^{N} \varphi_k \omega_{\pi(k)}$ are both Schur-convex functionals with $\boldsymbol{\omega} \geq 0$ (Marshall et al. (2011), pp. 79–154). Consequently, $\Psi(\pi)$ is maximized at the majorization-top arrangement (co-monotone) and minimized at the majorization-bottom arrangement (counter-monotone).

Therefore, the volatility difference is given by

$$\Delta\Psi = \Psi(\pi) - \Psi(\pi') = (\omega_{\pi(k)}^2 - \omega_{\pi(j)}^2) \cdot (\phi_k - \phi_j) + (\omega_{\pi(k)} - \omega_{\pi(j)}) \cdot (\varphi_k - \varphi_j) > 0.$$

That is, after we interchange the order of $\omega_{\pi(k)}$ and $\omega_{\pi(j)}$, the summation $g_k^2 + g_j^2$ will decrease, while all other $g_i^2$ $(i \neq k, j)$ remain the same. Also, note that after the above interchange, the $(k, j)$ pair becomes a new inversion pair (i.e., $(\omega_{\pi(k)} - \omega_{\pi(j)}) \cdot (m_j - m_k) < 0$). On the other hand, for all $s \neq j, k$, 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, the number of inversion pairs strictly increases by at least one. Starting from any permutation of $\boldsymbol{\omega}$, if there exists some non-inversion pair $(k, j)$, then we can repeat the above process by interchanging $\omega_{\pi(k)}$ and $\omega_{\pi(j)}$, and we will always reach the case in which $\omega_1 < \omega_2 < \ldots < \omega_N$, and $\sigma_{\text{GDP}}$ is minimized when the ordering of $\{m_i\}$ is the exact opposite of $\{\omega_i\}$. Reversing this process iteratively leads to the co-monotone maximum with perfectly-sorted matching. This completes the proof of the lemma.

□

**Step 2: Ordering Results for Expected Volatility**

We now prove the main statements in Proposition 5.4 in terms of expected GDP volatility, $\mathbb{E}[\sigma_{\text{GDP}}]$. When $\beta = \min\{\beta_m, \beta_\omega\} > 1$, by Proposition S.1, the second and third conditions in Lemma S.8 hold almost surely eventually for all firms simultaneously. Therefore, they also hold in expectation when $N \to \infty$. The only remaining obstruction is the first condition in Lemma S.8, which does not hold probabilistically for $k \approx N$ when we merely have $\beta > 1$.

To bypass this obstruction, our proof strategy is to provide an expectation-based equivalent to S.96. Consider the canonical probability space $(S, \mathcal{F}_s, \mathbb{P})$ we established in Appendix S.6.1 for random graph $\mathbf{A}$ defined by (5.5). Using (S.95),

$$\mathbb{E}[v_k] = \mathbb{E}\left[\alpha\omega_{\pi(k)} \cdot \sum_{i\in\mathcal{M}_k} \omega_{\pi(i)} + \alpha \sum_{i\in\mathcal{M}_k} \frac{\omega_{\pi(i)}}{\widehat{d}_i} \sum_{j\notin\mathcal{D}_i} \omega_{\pi(j)}\right] = \alpha\omega_{\pi(k)}\mathbb{E}\left[\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] + \alpha\mathbb{E}\left[\sum_{i\in\mathcal{M}_k} \frac{\omega_{\pi(i)}}{\widehat{d}_i} \sum_{j\notin\mathcal{D}_i} \omega_{\pi(j)}\right].$$

Next, applying (S.34) as $N \to \infty$ gives

$$\mathbb{E}\left[\sum_{i\in\mathcal{M}_k} \omega_{\pi(i)}\right] = \left[\left(1 - \omega_{\pi(k)}\right) \frac{Nk^{-1/\beta_m} - 1}{N - 1} + \omega_{\pi(k)}\right] \approx k^{-1/\beta_m}; \qquad \forall k \in [N],$$

where we utilize the vanishing property of firm size $\boldsymbol{\omega}$ (S.35) under any permutation. On the other hand, when $\beta > 1$, by (S.57),

$$m_k \approx \frac{N}{k^{1/\beta_m}}, \quad \text{a.s. } \forall k \in [N].$$

Combining these two equations gives[38]

$$\mathbb{E}\left[\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right]\approx\mathbb{E}\left[\frac{m_k}{N}\right]=\mathbb{E}\left[\frac{m_k}{N}\right](1+o(1)),\qquad\forall k\in[N]. \tag{S.101}$$

By construction, $\mathbb{E}\left[\frac{m_k}{N}\right]$ is monotonically decreasing.

Next, we analyze the expectation of the second term. We will apply the results in Proposition S.1. In particular, we define an event set

$$\mathcal{G}_N\subset\mathcal{F}_s:=\underbrace{\left\{s\in S,d-d^{2/3}\le d_k\le d+d^{2/3}\ \forall k\in[N]\right\}}_{\equiv\mathcal{G}_N^1}\bigcap\underbrace{\left\{s\in S,\left|\varrho_k(N)-\mathbb{E}[\varrho_k(N)]\right|\le\varepsilon_N,\ \forall k\in[N]\right\}}_{\equiv\mathcal{G}_N^2},\qquad\varepsilon_N:=N^{-\varsigma},$$

with

$$0<\varsigma<\min\left\{\frac{1}{2},1-\frac{1}{\beta_\omega}\right\}.$$

$\varrho_k(N)$ is defined in (S.49) and $d=\zeta(\frac{1}{\beta_m},N)$. Using (S.46)

$$\mathbb{E}[d_k]=\frac{N}{N-1}\left(\zeta(\frac{1}{\beta_m},N)-\frac{1}{k^{\frac{1}{\beta_m}}}\right)\approx d=\zeta(\frac{1}{\beta_m},N),$$

uniformly for all $k\in[N]$ as $N\to\infty$. Then by the concentration inequality (S.47), the tail probabilities for all $k\in[N]$ are bounded above by

$$\mathbb{P}\left(\left(\mathcal{G}_N^1\right)^c\right)\le N\exp\left(-\frac{1}{2}d^{1/3}\right)+N\exp\left(-\frac{d^{4/3}}{2d+2d^{2/3}/3}\right)=2N\exp\left(-\Theta\left(N^{\frac{1}{3}\left(1-\frac{1}{\beta_m}\right)}\right)\right).$$

Similarly, by (S.53) we have

$$\mathbb{P}\left(\left(\mathcal{G}_N^2\right)^c\right)\le\begin{cases}N\exp\left(-2\Theta\left(N^{1-2\varsigma}\right)\right)+N\exp\left(-\Theta\left(N^{1-\frac{1}{\beta_\omega}-\varsigma}\right)\right); & \beta_\omega>2\\ N\exp\left(-2\Theta\left(\frac{N^{2\left(1-\frac{1}{\beta_\omega}-\varsigma\right)}}{\ln N}\right)\right)+N\exp\left(-\Theta\left(N^{1-\frac{1}{\beta_\omega}-\varsigma}\right)\right); & \beta_\omega=2\\ N\exp\left(-2\Theta\left(N^{2\left(1-\frac{1}{\beta_\omega}-\varsigma\right)}\right)\right)+N\exp\left(-\Theta\left(N^{1-\frac{1}{\beta_\omega}-\varsigma}\right)\right); & 1<\beta_\omega<2.\end{cases}$$

Therefore, by De Morgan's law,

$$\mathbb{P}(\mathcal{G}_N)=\mathbb{P}\left(\mathcal{G}_N^1\bigcap\mathcal{G}_N^2\right)=1-\mathbb{P}\left(\left(\mathcal{G}_N^1\right)^c\bigcup\left(\mathcal{G}_N^2\right)^c\right)$$

[38] For simplicity, we drop the subscript $k$ and $\pi$ for the approximation error $o$, since we focus on the leading order.

and by the union bound,

$$\mathbb{P}\left((\mathcal{G}_N)^c\right) = \mathbb{P}\left((\mathcal{G}_N^1)^c \bigcup (\mathcal{G}_N^2)^c\right) \le \mathbb{P}\left((\mathcal{G}_N^2)^c\right) + \mathbb{P}\left((\mathcal{G}_N^2)^c\right) = o\left(\exp\left(-\Theta\left(N^{\varepsilon_0}\right)\right)\right)$$

for any

$$0 < \varepsilon_0 < \min\left\{\frac{1}{3}\left(1 - \frac{1}{\beta_m}\right),\, 1 - 2\varsigma,\, 1 - \frac{1}{\beta_\omega} - \varsigma\right\},$$

as exponential decay dominates any polynomial by Lemma S.4. It follows that

$$\mathbb{P}\left(\mathcal{G}_N\right) = 1 - o\left(\exp\left(-\Theta\left(N^{\varepsilon_0}\right)\right)\right).$$

Given the probability estimates, we define

$$H_N^{(k)} := \sum_{i \in \mathcal{M}_k} \frac{\omega_{\pi(i)}}{d_i + \tau_\xi/\tau_\varepsilon} \sum_{j \notin \mathcal{D}_i} \omega_{\pi(j)}. \tag{S.102}$$

By definition,

$$0 \le \sum_{j \notin \mathcal{D}_i} \omega_{\pi(j)} \le 1, \qquad d_i + \tau_\xi/\tau_\varepsilon \ge 1, \qquad \forall i \in [N]$$

we have

$$0 \le H_N^{(k)} \le \sum_{i \in \mathcal{M}_k} \omega_{\pi(i)} \le 1, \qquad \forall k \in [N].$$

Hence $H_N^{(k)}$ is uniformly bounded, which is a sure event. Next, we split the expectation according to the event $\mathcal{G}_N$:

$$\mathbb{E}\left[H_N^{(k)}\right] = \mathbb{E}\left[H_N^{(k)} \mathbf{1}_{\mathcal{G}_N}\right] + \mathbb{E}\left[H_N^{(k)} \mathbf{1}_{\mathcal{G}_N^c}\right].$$

The second term is bounded by

$$\left|\mathbb{E}\left[H_N^{(k)} \mathbf{1}_{\mathcal{G}_N^c}\right]\right| \le \mathbb{P}(\mathcal{G}_N^c) = o\left(\exp(-\Theta(N^{\varepsilon_0}))\right). \tag{S.103}$$

On the event $\mathcal{G}_N$, we have uniformly

$$\widehat{d}_i = \widehat{d} + o(\widehat{d}), \qquad \sum_{j \notin \mathcal{D}_i} \omega_{\pi(j)} = 1 + o(1), \qquad \forall i \in [N],$$

where the second equation is derived from the uniform limit (S.52). Therefore,[39]

$$H_N^{(k)}\mathbf{1}_{\mathcal{G}_N} = \frac{1+o(1)}{\widehat{d}}\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right)\mathbf{1}_{\mathcal{G}_N}.$$

Taking expectations gives

$$\mathbb{E}\left[H_N^{(k)}\mathbf{1}_{\mathcal{G}_N}\right] = \frac{1+o(1)}{\widehat{d}}\mathbb{E}\left[\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right)\mathbf{1}_{\mathcal{G}_N}\right].$$

Since $0 \le \sum_{i\in\mathcal{M}_k}\omega_{\pi(i)} \le 1$ with certainty, we have

$$\left|\mathbb{E}\left[\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right)\mathbf{1}_{\mathcal{G}_N^c}\right]\right| \le \mathbb{P}(\mathcal{G}_N^c) = o\left(\exp(-\Theta(N^{\varepsilon_0}))\right).$$

Hence,

$$\mathbb{E}\left[\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right)\mathbf{1}_{\mathcal{G}_N}\right] = \mathbb{E}\left[\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right] - o\left(\exp(-\Theta(N^{\varepsilon_0}))\right).$$

Therefore,

$$\mathbb{E}\left[H_N^{(k)}\mathbf{1}_{\mathcal{G}_N}\right] = \frac{1+o(1)}{\widehat{d}}\mathbb{E}\left[\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right], \tag{S.104}$$

where the exponentially-decaying errors are absorbed in the $o(1)$ term. Combining the good-event and bad-event estimates (S.103) and (S.104) yields

$$\mathbb{E}\left[H_N^{(k)}\right] = \frac{1+o(1)}{\widehat{d}}\mathbb{E}\left[\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right] = \mathbb{E}\left[\frac{\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}}{\widehat{d}}\right](1+o(1)), \qquad \forall k\in[N]. \tag{S.105}$$

Again, the exponentially decaying errors are absorbed in the $o(1)$ terms.

Consolidating equations (S.101), (S.102), and (S.105), we have

$$\begin{aligned}\mathbb{E}[v_k] &= \alpha\omega_{\pi(k)}\mathbb{E}\left[\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right] + \alpha\mathbb{E}\left[H_N^{(k)}\right]\\ &= \alpha\omega_{\pi(k)}\mathbb{E}\left[\frac{m_k}{N}\right](1+o(1)) + \frac{\alpha}{\widehat{d}}\mathbb{E}\left[\frac{m_k}{N}\right](1+o(1)).\end{aligned}$$

[39] The $o(\widehat{d})$ terms drop out of the expression since we have $d\to\infty$ and

$$\frac{1}{\hat{d}_i} = \frac{1}{\hat{d}\{1+o(1)\}} = \frac{1+o(1)}{\hat{d}}, \qquad (1+o(1))(1+o(1)) = 1+o(1).$$

Next, we apply the leading-order approximation identical to the proof of Lemma S.8. It follows that we have an expectation-based analogue of (S.96),

$$\mathbb{E}\left[g_k\right] = \delta\left(\omega_{\pi(k)} + \mathbb{E}\left[v_k\right]\right) = \left(\delta + \frac{\alpha\delta}{N}\mathbb{E}\left[m_k\right]\right)\omega_{\pi(k)} + \left(\frac{\alpha\delta}{N\widehat{d}}\right)\mathbb{E}\left[m_k\right], \qquad \forall k \in [N]. \tag{S.106}$$

We now complete the argument for the aggregate second moment. Under the leading-order approximation of $(\mathbf{I} - \mathbf{V})^{-1}$, we have

$$g_k = \delta\left[\omega_{\pi(k)} + \alpha\omega_{\pi(k)}\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)} + \alpha H_N^{(k)}\right].$$

The preceding argument implies that, on the high-probability event $\mathcal{G}_N$,

$$H_N^{(k)} = \frac{1 + o(1)}{\widehat{d}}\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}$$

uniformly over the relevant indices, while the contribution from $\mathcal{G}_N^c$ is exponentially small. Therefore, after summing over $k$, the aggregate second moment satisfies

$$\sum_{k=1}^N \mathbb{E}[g_k^2] = \delta^2\sum_{k=1}^N \mathbb{E}\left[\left\{\omega_{\pi(k)} + \alpha\left(\omega_{\pi(k)} + \frac{1}{\widehat{d}}\right)\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right\}^2\right] + \delta^2 o\left(\sum_{k=1}^N \omega_{\pi(k)}^2\right).$$

Expanding the square gives

$$\sum_{k=1}^N \mathbb{E}[g_k^2] = \delta^2\sum_{k=1}^N\left[\omega_{\pi(k)} + \alpha\left(\omega_{\pi(k)} + \frac{1}{\widehat{d}}\right)\mathbb{E}\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right)\right]^2 + \alpha^2\delta^2\sum_{k=1}^N\left(\omega_{\pi(k)} + \frac{1}{\widehat{d}}\right)^2\mathrm{Var}\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right) + \delta^2 o\left(\sum_{k=1}^N \omega_{\pi(k)}^2\right).$$

We next show that the aggregate variance correction is negligible. Since the edge indicators are independent across ordered pairs and $A_{kk} = 1$ is deterministic,

$$\mathrm{Var}\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right) = \sum_{i\neq k} p_{ik}(1 - p_{ik})\omega_{\pi(i)}^2.$$

Using the probability structure in (5.5), the off-diagonal probabilities in column $k$ are of order $\mathbb{E}[m_k/N]$. Therefore,

$$\mathrm{Var}\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right) \leq C\,\mathbb{E}\left[\frac{m_k}{N}\right]\sum_{i=1}^N \omega_{\pi(i)}^2.$$

for some constant $C > 0$. It follows that

$$\sum_{k=1}^{N}\left(\omega_{\pi(k)}+\frac{1}{\widehat{d}}\right)^2 \mathrm{Var}\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right) \le C\left(\sum_{i=1}^{N}\omega_{\pi(i)}^2\right)\sum_{k=1}^{N}\mathbb{E}\left[\frac{m_k}{N}\right]\left(\omega_{\pi(k)}+\frac{1}{\widehat{d}}\right)^2$$
$$\le C\left(\sum_{i=1}^{N}\omega_{\pi(i)}^2\right)\left[\sum_{k=1}^{N}\omega_{\pi(k)}^2+\frac{2}{\widehat{d}}\sum_{k=1}^{N}\omega_{\pi(k)}+\frac{1}{\widehat{d}^2}\sum_{k=1}^{N}\mathbb{E}\left(\frac{m_k}{N}\right)\right].$$

Note that $\sum_{k=1}^{N}\mathbb{E}\left(\frac{m_k}{N}\right) \approx \sum_{k=1}^{N} k^{-\frac{1}{\beta_m}} = \zeta(\frac{1}{\beta_m}, N) \approx \widehat{d}$. Meanwhile, $\sum_{k=1}^{N}\omega_{\pi(k)} = 1$, $\sum_{k=1}^{N}\omega_{\pi(k)}^2 \to 0$ when $\beta_\omega > 1$, and $\widehat{d} \to \infty$ when $\beta_m > 1$, we have

$$\sum_{k=1}^{N}\left(\omega_{\pi(k)}+\frac{1}{\widehat{d}}\right)^2 \mathrm{Var}\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right) = o\left(\sum_{k=1}^{N}\omega_{\pi(k)}^2\right).$$

Therefore,

$$\sum_{k=1}^{N}\mathbb{E}[g_k^2] = \delta^2\sum_{k=1}^{N}\left[\omega_{\pi(k)}+\alpha\left(\omega_{\pi(k)}+\frac{1}{\widehat{d}}\right)\mathbb{E}\left(\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right)\right]^2 + \delta^2 o\left(\sum_{k=1}^{N}\omega_{\pi(k)}^2\right).$$

By the expectation-based equivalent (S.101), $\mathbb{E}\left[\sum_{i\in\mathcal{M}_k}\omega_{\pi(i)}\right] \approx \mathbb{E}\left[\frac{m_k}{N}\right]$. Consequently, the leading-order expected volatility is

$$\sum_{k=1}^{N}\mathbb{E}[g_k^2] = \delta^2\sum_{k=1}^{N}\left[\omega_{\pi(k)}+\alpha\left(\omega_{\pi(k)}+\frac{1}{\widehat{d}}\right)\mathbb{E}\left(\frac{m_k}{N}\right)\right]^2 + \delta^2 o\left(\sum_{k=1}^{N}\omega_{\pi(k)}^2\right). \tag{S.107}$$

By definition,

$$\left(\omega_{\pi(k)}+\alpha\left(\omega_{\pi(k)}+\frac{1}{\widehat{d}}\right)\mathbb{E}\left[\frac{m_k}{N}\right]\right)^2 \ge \omega_{\pi(k)}^2, \qquad \forall k\in[N],$$

as the terms on LHS are all nonnegative. Therefore,

$$\sum_{k=1}^{N}\left(\omega_{\pi(k)}+\alpha\left(\omega_{\pi(k)}+\frac{1}{\widehat{d}}\right)\mathbb{E}\left[\frac{m_k}{N}\right]\right)^2 \ge \sum_{k=1}^{N}\omega_{\pi(k)}^2.$$

We express (S.107) as

$$\sum_{k=1}^{N}\mathbb{E}[g_k^2] = \delta^2\sum_{k=1}^{N}\left[\omega_{\pi(k)}+\alpha\left(\omega_{\pi(k)}+\frac{1}{\widehat{d}}\right)\mathbb{E}\left(\frac{m_k}{N}\right)\right]^2(1+o(1)). \tag{S.108}$$

Under the leading-order approximation, we ignore the $o(1)$ term. It follows that

$$\sum_{k=1}^{N}\mathbb{E}[g_k^2]=\sum_{k=1}^{N}\left[\delta+\alpha\delta\,\mathbb{E}\left(\frac{m_k}{N}\right)\right]^2\omega_{\pi(k)}^2+\sum_{k=1}^{N}2\left[\frac{\alpha\delta}{\widehat{d}}\mathbb{E}\left(\frac{m_k}{N}\right)\right]\left[\delta+\alpha\delta\,\mathbb{E}\left(\frac{m_k}{N}\right)\right]\omega_{\pi(k)}+\underbrace{\sum_{k=1}^{N}\left[\frac{\alpha\delta}{\widehat{d}}\mathbb{E}\left(\frac{m_k}{N}\right)\right]^2}_{\pi\text{-invariant}},\tag{S.109}$$

which is the expectation analogue of (S.98). The last term is permutation-invariant. Moreover, $\mathbb{E}[m_k/N]$ is monotonically decreasing in $k$ by construction. Therefore both coefficient sequences

$$\phi_k^E=\left[\delta+\alpha\delta\,\mathbb{E}\left(\frac{m_k}{N}\right)\right]^2,\qquad \varphi_k^E=2\left[\frac{\alpha\delta}{\widehat{d}}\mathbb{E}\left(\frac{m_k}{N}\right)\right]\left[\delta+\alpha\delta\,\mathbb{E}\left(\frac{m_k}{N}\right)\right].$$

are monotonically decreasing in $k$. The same rearrangement inequality and adjacent-swap argument from Step 1 therefore apply. Consequently, the leading-order expected GDP volatility is maximized under the co-monotone matching between expected in-degrees and firm sizes, minimized under the counter-monotone matching, and decreases along adjacent swaps that increase the inversion number. The proof of the Proposition 5.4 is now complete. □

**Extensions**

The statements in Proposition 5.4 are expressed in terms of the expected volatility of GDP. Now we discuss two possible extensions that strengthen the ordering results to almost -sure occurrence.

**Corollary S.5.** *Given the random graph structure (5.5) and the numerical distribution of firm sizes $\boldsymbol{\omega}$ following (5.8), for any arbitrary permutation $\pi: i \longmapsto [j]$ on $\boldsymbol{\omega}$, if $\beta=\min\{\beta_m,\beta_\omega\}>1$, and $\frac{1}{\beta_m}+\frac{1}{\beta_\omega}<1$, then up to a leading-order approximation the GDP volatility conditional on micro shocks satisfies the following properties almost surely when $N\to\infty$:*

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

*Proof.* In Proposition S.1, we proved that under the current restrictions on $\beta_m$ and $\beta_\omega$, the three technical conditions proposed in Lemma S.8 hold almost surely as $N\to\infty$, for all firms simultaneously. We use Definition S.1 and denote $S_N^*$ as the event of the simultaneous occurrence of three technical conditions. Now define the event

$$\mathcal{E}_N\equiv\left\{s\in S^N,\text{The three combinatorial properties in Proposition 5.4 hold}\right\}.$$

By Lemma S.8, the event sets satisfy: $S_N^* \subseteq \mathcal{E}_N$. Therefore, Proposition S.1 implies

$$\liminf_{N\to\infty} S_N^* \subseteq \liminf_{N\to\infty} \mathcal{E}_N \Longrightarrow \mathbb{P}\left(\liminf_{N\to\infty} \mathcal{E}_N\right) = 1,$$

which proves the almost sure statements. □

The second extension does not impose additional restrictions on $(\beta_m, \beta_\omega)$. Instead, we focus on volatility contributions driven by an infinite subset of firms receiving the largest attention. In the following lemma, the random graph structure and firm-size distribution remain the same.

**Corollary S.6.** *For any permutation $\pi : i \longmapsto [j]$ on $\boldsymbol{\omega}$, if $\beta = \min\{\beta_m, \beta_\omega\} > 1$, there exists a subset of most salient firms with cardinality $N^\Lambda$, $\Lambda \in (0,1)$ such that up to a leading-order approximation, the GDP volatility driven by micro shocks in these high-salience firms satisfies the three combinatorial properties almost surely when $N \to \infty$.*

*Proof.* We apply Proposition S.1 and Corollary S.2. In particular, we pick a subset $K_N$ of firms with highest attentions, i.e., the first $N^\Lambda$ columns in the stochastic adjacency matrix,

$$K_N := \{1, 2, \ldots, \lfloor N^\Lambda \rfloor\}.$$

Then by Proposition S.1 and Corollary S.2, the three technical conditions proposed in Lemma S.8 hold almost surely as $N \to \infty$, for all firms $k \in K_N$ simultaneously. Then by Lemma S.8 and the last Corollary, it is straightforward to see that the ordering results in Proposition 5.4 are preserved. However, we need to revise the statement as permutations are defined in the entire set $[N]$ but we only focus on the subset $K_N$. In particular, we focus on GDP volatility driven by firms in $K_N$ and the localized images of permutations from $K_N$. Fixing a permutation $\pi$, we consider the localized image set,

$$R_\Lambda := \left\{\pi(1), \pi(2), \ldots, \pi(N^\Lambda)\right\}.$$

We sort the elements in $R_\Lambda$ in decreasing order,

$$r_{(1)} > r_{(2)} > \cdots > r_{(N^\Lambda)}.$$

Then the rearrangement inequalities akin to (S.99) are expressed as

$$\sum_{k=1}^{N^\Lambda} \phi_k \left(\omega_{[r_{(N+1-k)}]}\right)^2 \leq \sum_{k=1}^{N^\Lambda} \phi_k \left(\omega_{\pi(k)}\right)^2 \leq \sum_{k=1}^{N^\Lambda} \phi_k \left(\omega_{[r_{(k)}]}\right)^2; \ \sum_{k=1}^{N^\Lambda} \varphi_k \omega_{[r_{(N+1-k)}]} \leq \sum_{k=1}^{N^\Lambda} \varphi_k \omega_{\pi(k)} \leq \sum_{k=1}^{N^\Lambda} \varphi_k \omega_{[r_{(k)}]}. \tag{S.110}$$

Equation (S.110) holds almost surely for firms in $K_N$. The rest of the proof is identical to that of Lemma S.8, so we omit the details. □

# References


Abramowitz, M. and I. A. Stegun (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Ninth Dover printing, tenth GPO printing ed.). New York: Dover.

Arnold, B. C., N. Balakrishnan, and H. N. Nagaraja (2008). *A First Course in Order Statistics*, Volume 54 of *Classics in Applied Mathematics*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM).

Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *Journal of the American Statistical Association 57*(297), 33–45.

Billingsley, P. (1995). *Probability and Measure* (3 ed.). Wiley Series in Probability and Mathematical Statistics. New York: John Wiley & Sons.

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

Frieze, A. and M. Karoński (2026). *Introduction to Random Graphs*. Cambridge University Press.

Gautschi, W. (1959). Some elementary inequalities relating to the gamma and incomplete gamma function. *Journal of Mathematics and Physics 38*(1-4), 77–81.

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

Hardy, G. H., J. E. Littlewood, and G. Pólya (1952). *Inequalities* (2 ed.). Cambridge: Cambridge University Press.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association 58*(301), 13–30.

Marshall, A. W., I. Olkin, and B. C. Arnold (2011). *Inequalities: Theory of Majorization and Its Applications* (2 ed.). Springer Series in Statistics. New York: Springer.

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